Lecture Notes in Electrical Engineering 355

R. Jagadeesh Kannan
S. Geetha
Sravanthi Sashikumar
Carl Diver *Editors*

International Virtual Conference on Industry 4.0

Select Proceedings of IVCI4.0 2020

Springer

Lecture Notes in Electrical Engineering

Volume 355

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at http://www.springer.com/series/7818

R. Jagadeesh Kannan · S. Geetha ·
Sravanthi Sashikumar · Carl Diver
Editors

International Virtual Conference on Industry 4.0

Select Proceedings of IVCI4.0 2020

Springer

Editors
R. Jagadeesh Kannan
School of Computer Science
and Engineering
Vellore Institute of Technology
Chennai, Tamil Nadu, India

S. Geetha
School of Computer Science
and Engineering
Vellore Institute of Technology
Chennai, Tamil Nadu, India

Sravanthi Sashikumar
Department of Engineering
Manchester Metropolitan University
Manchester, UK

Carl Diver
Manchester Metropolitan University
Manchester, UK

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-16-1243-5 ISBN 978-981-16-1244-2 (eBook)
https://doi.org/10.1007/978-981-16-1244-2

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

Preface

Industry 4.0 marks the new phase in the Industrial Revolution which concentrates profoundly on automation, interconnectivity, machine learning, and real-time data. Industry 4.0, also referred to as IIoT or smart manufacturing, bridges physical production and operations with smart digital technology, big data, and machine learning to craft a better connected and more holistic ecosystem for industries that can focus on manufacturing and supply chain management. Industry 4.0 symbolizes a compelling transformation in the way the products are produced by digitization in manufacturing.

Industry 4.0 is perceived as the computerized version of Industry 3.0. It was disruptive when computers were introduced into the Industry 3.0. Presently, and into the future as Industry 4.0 unleashes, computers stay connected and effectively communicate with each other to eventually make decisions without human participation. An amalgamation of cyber-physical systems, the Internet of Things, and the Internet of Systems lead to the realization of Industry 4.0 and the smart factory a possibility. On account of the sustenance of smart machines which grow smarter as they get access to more data, the factories will become more efficient and productive and less wasteful. Eventually, it is the network of these machines that stay digitally connected with each other to generate and share information which results in the true potential of Industry 4.0.

This book is the collection of remarkable works related to Industry 4.0 and cyber-physical systems that were submitted and presented in the First International Virtual Conference on Industry 4.0—IVCI4.0. This prestigious event was hosted by the School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus, India, and Manchester Metropolitan University, UK. The conference hosted the main theme—"Next Generation Computing Systems" with three sub-themes, viz., (1) Industry 4.0; (2) cyber-physical systems; and (3) allied topics, use cases, and applications. Over the two days of July 6 and 7, 2020, IVCI4.0 brought together a large number of stakeholders, including engineers, faculty members, researchers, students, industry, governmental organizations, private sectors, etc., to learn, share, and build fruitful and long-term collaborations. Ninety-one papers were submitted from different countries including India, UK, Dubai, Indonesia, and Australia. Of these papers, only 39 papers were accepted for presentation and publications—keeping the acceptance ratio to be 42%.

The conference would not have been successful without the support of the program committee members, organizing committee members, keynote/plenary speakers, authors, tutorial presenters, special events organizers, reviewers, and the conference attendees. We greatly appreciate the meticulous reviews provided by the program committee members and additional reviewers and thank the authors for having revised their papers to address the comments and suggestions by the referees. Our sincere thanks to the steering committee members for their timely help and supervision.

We want to thank all the keynote speakers and plenary speakers of the conference. The conference program was enriched by the outstanding talks by Mr. Robin Phillips (President of Greater Manchester Chamber of Commerce, Head of Siemens Advanta Consulting UK, and Ex-CFO of Siemens Industry UK), Mr. Paul Gilbert (Quanser), Dr. Farook Hussain (University of Technology Sydney), Prof. Andrea M. Tonello (University of Klagenfurt), Dr. Madhevan Balasundaram (Austria), Dr. Rupak Kharel (Cyber Foundry Lead and Reader in Cyber Security), Dr. Sabu M. Thampi (IIITM-K), Dr. Eric Lou (Reader in Cyber Security, MMU), Dr. Iain Reid (Reader in Operations Management, MMU), Dr. Iniya Nehru (Senior Technical Director, National Informatics Centre (NIC)), Dr. Soufiene Djahel and Dr. Keeley Crockett (Reader in Computational Intelligence), Dr. Mohammad Hammoudeh (Reader in Future Networks and Security), Prof. Yonghong Peng (MMU), Dr. Paul Battersby (MIDAS, Manchester), Dr. Scott Pepper (GAMBICA), Dr. Martin Vares (Fractory), Mr. Suresh Perinjery (PTC), Mr. Sai Krishna Rao (Schneider), Dr. Zhipeng Wang (CEO, TIPoT Technologies Inc., Ottawa, ON, Canada), Prof. Muhammad Ali Imran (University of Glasgow, Scotland), Dr. Umashankar Subramaniam (Prince Sultan University, Saudi Arabia), Dr. C. P. Ravikumar (Technical Director, Texas Instruments India), Dr. B. K. Murthy (Scientist and Group Coordinator, Meity, India), Dr. Khalil Drira (CNRS Research Director, LAAS-CNRS, University of Toulouse, France), Dr. Rajeev R. Raje (Professor and Associate Dean for Planning, Finance and Faculty Affairs, Department of Computer and Information Science, IUPUI, Indianapolis), Prof. Liangxiu Han (Professor of Computer Science, Co-director of CIACS, and Deputy Director of MMU Bigdata Centre), Dr. Alhussein Albarbar (IIOT and Smart Sensors), Dr. Narasimhan Sundararajan (Professor (Retd.), EEE, NTU, Singapore), Prof. David Bamford (Professor of Operations Management), Mr. Muraleedharan Manningal (VP and Head of Technology and Innovation Center, SE2 Digital Service LLP, Pune, India), and Dr. Sakkaravarthi Ramanathan (Cegep Gaspesie, Canada).

We would like to express our sincere thanks to the Chancellor, Vice Presidents, Assistant Vice President, Executive Director, Vice-Chancellor, Pro-Vice-Chancellor, Registrar, Additional Registrar and other members of VIT University family for extending all facilities and support for the smooth conduct of the conference.

We wish to express our thanks to the Springer Publisher for producing and publishing the IVCI4.0 proceedings and especially to Ms. Kamiya Khatter and Mr. Maniarasan Gandhi for their excellent support. We would also like to thank all those who contributed to the success of IVCI4.0 but whose names may not be listed.

We hope you will find this book informative and resourceful.

Chennai, India R. Jagadeesh Kannan
Chennai, India S. Geetha
Manchester, UK Sravanthi Sashikumar
Manchester, UK Carl Diver
February 2021

Contents

About the Editors

R. Jagadeesh Kannan is Professor and Dean of the School of Computer Science and Engineering at Vellore Institute of Technology, India. He completed his Ph.D. degree in Handwritten Character Recognition using Hybrid Techniques from Anna University, Chennai, India. He got his M.E. degree in Computer Science and Engineering from National Engineering College, Tamil Nadu, and B.E. in Instrumentation and Control Engineering from Madurai Kamaraj University, Tamil Nadu, India. Prof. Kannan has over 18 years of teaching and industrial experience in reputed organizations. Prof. Kannan has got several publications in conference proceedings and journals of National and International repute. His research interests are neural networks, fuzzy logic, neuro-fuzzy systems, soft computing tools, pattern recognition, natural language processing, image processing, networking, printed, handwritten and cursive character recognition, and artificial intelligence. Dr. Kannan is an active member of several Indian and International societies such as IEEE, ISTE, IACSIT, SDIWC, IFRSA, and IAENG.

S. Geetha is Professor and Associate Dean of the School of Computer Science and Engineering at Vellore Institute of Technology, Chennai Campus, India. Dr. Geetha received her Master's (Computer Science and Engineering) and doctoral (Information and Communication Engineering) degrees from Anna University in 2004 and 2011, respectively. Her areas of interest are computer vision, information and network security, and machine learning. Dr. Geetha has co-authored 06 books, 06 book chapters, and over 100 journal and conference papers.

Sravanthi Sashikumar is Deputy Head of Department of Engineering at Manchester Metropolitan University, UK. She received a Bachelor's degree in mechanical engineering from the University of Madras, India, in 2002 and a Master's degree in Engineering Control Systems and Instrumentation from the University of Huddersfield, in 2005. She completed her Ph.D. from the University of Bolton in 2012, and her research interest areas are structural analysis vehicle crashworthiness, passive safety of road vehicles, vehicle occupant safety, and impact biomechanics. Dr. Sravanthi has got several publications to her credit.

Carl Diver is the academic lead on Industry 4.0 and a reader in Industrial Digitalization at Manchester Metropolitan University. Carl joined Manchester Metropolitan University in 2018 to lead the University's Industry 4.0 activities. Before academia, Carl worked in the manufacturing sector for over 20 years, firstly for a large multinational and then establishing his consultancy. At Delphi Automotive Systems, Carl's focus was on fuel injection equipment from both an R&D and manufacturing angle. More recently, Carl led Industry 4.0 activity, research, and teaching for the manufacturing group in the School of Mechanical Aerospace and Civil Engineering at the University of Manchester. He helped establish the academic conference at the Industry 4.0 summit in Manchester, and his Ph.D. students won international competitions such as the Siemens Open Space challenge at Hannover Messe in 2018. Carl has been speaking about Industry 4.0 at conferences and events in the UK, Europe, the Middle East, and Asia. He is an associate co-editor of the Virtual and Physical Prototyping Journal and continues helping businesses with more advanced, efficient, and controlled processes.

Pre-diagnosis, Prediction and Report Generation of a Disease

O. Pandithurai, K. Dhinakaran, D. Jayashree, M. Nivetha, M. Nithya Sree, and S. Priyadharshini

Abstract In recent days, people face numerous diseases due to many factors that includes an injury, unhealthy lifestyle, infection and environmental conditions. For this reason, earlier prediction and timely diagnosis of a human disease become a crucial as well as a demanding task in the field of health care. The paper has proposed a solution in the form of a cloud application that predicts the disease of a patient based on symptoms, Electronic health record (EHR) and a set of vital signs such as temperature, heart rate, blood pressure (BP) and electrocardiogram (ECG) measured with the help of specific sensor devices and finally generates a report based on the disease. Data mining, also termed as knowledge discovery in data (KDD), plays a vital role in correct prediction of a disease. With the aid of human diseases and conditions data sets, KDD detects an unknown pattern information in the enormous volume of clinical data. Health care is among the most significant fields, and artificial intelligence is intending to reconstruct. Here, we use AI identified medical records, one hundred thousand data points each person. So, beyond any medical practitioner could analyze, we can essentially predetermine the disease. The main aim of the system is to predict a disease of a human being in a more efficient and easier way.

O. Pandithurai (✉) · K. Dhinakaran · D. Jayashree · M. Nivetha · M. Nithya Sree · S. Priyadharshini
Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai, India
e-mail: pandics@ritchennai.edu.in

D. Jayashree
e-mail: jayashree.d@ritchennai.edu.in

M. Nivetha
e-mail: nivetha.m.2017.cse@ritchennai.edu.in

M. Nithya Sree
e-mail: nithyasree.m.2017.cse@ritchennai.edu.in

S. Priyadharshini
e-mail: priyadharshini.s.2017.cse@ritchennai.edu.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_1

Keywords Disease prediction · Symptoms · Data mining · Natural language processing

1 Introduction

In today's world, constructing a disease prediction model is of larger necessity and a greatly challenging task. Generally, disease prediction and report generation involve too many manual processes and cause a delay. Prediction based on collective intelligence, utilizing general human movement models and other such related models does not predict diseases in a shorter span of time and fails to assure accurate results. To surmount these issues, we have addressed a solution in the form of a cloud application that predicts the disease of a patient based on the symptoms [1, 2], EHR and vital signs, also, it suggests a set of treatment measures for the disease.

2 Background Study

2.1 K-Nearest Neighbors Algorithm

Machine learning algorithms engage a range of statistical, probabilistic and optimization techniques to learn from previous action and identify effective patterns from large, unstructured and complicated data sets. The ML algorithms for disease prediction include Naive Bayesian (NB), K-nearest neighbor (KNN), decision tree (DT) algorithm, support vector machine (SVM) and random forest (RF) algorithm [3, 4]. KNN is one of the best data mining algorithms which is considerably used in disease prediction system. KNN is a simple algorithm that gathers all available patient data and classifies diseases based on a similarity measure. Classification is acquired by determining the nearest neighbor to decide on the class of an unknown data set. KNN is preferred over other classification algorithms because of its high convergence speed and simplicity.

2.2 Data Mining

Data mining is a process of collecting, cleaning, processing and gaining useful insights from data. Health care is one of many industries benefitting from data mining. In hospitals, data mining techniques save more time as it reduces the number of tests to be performed. Data mining has high potential for analyzing the unknown patterns in the data sets of medical domains. With the ease of data mining, the efficiency and quality of disease prediction models are optimized [5].

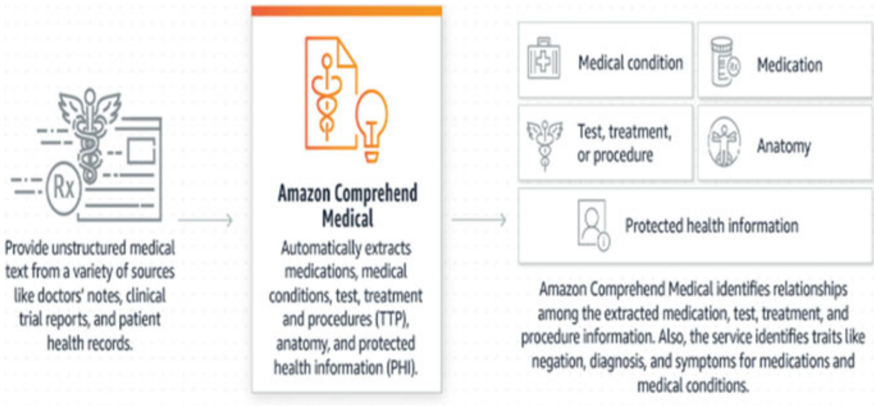


Fig. 1 Architecture—Amazon comprehend medical

2.3 Natural Language Processing

Natural language processing (NLP) employs computational techniques for the purpose of learning, understanding and producing human language content. Development of natural language processing methods is essential to automatically transform clinical text into structured clinical data that can be directly processed using machine learning algorithms. Natural language processing in the healthcare sector can help increase the accuracy and completeness of EHRs by transforming the free text into standardized data. There is many NLP software available in the market. Amazon comprehend medical is one of the best NLP tools which process the written documents to find the needed text. The tool processes the unstructured clinical data, extracts the necessary key terms and identifies the medical information and their relationship to one another, for example the medical dose, symptoms and medical state as shown in the Fig. 1.

2.4 WEKA Tool

Waikato Environment for Knowledge Analysis (Weka) is a widely known suite of machine learning software written in Java, developed at University of Waikato, New Zealand. It is an unpaid software authorized under GNU General Public License. Weka tool is an accumulation of machine learning algorithms for data mining actions. The advantages of WEKA include free availability under GNU General Public License, portability, ease of use, platform independent, feature selection and integrated data mining algorithms.

3 Proposed Solution

Accurate and timely predictions of a disease are increasingly important for several reasons. Building a real-time disease prediction application can increase the quality of life, reduce stigma associated with the disease, lead to robust treatment options, stimulate longer lives and more. On a complete study on the existing disease prediction methods, we address the following solution that accurately predicts a disease based on set of prerequisites, including symptoms, EHR/scanned reports and a few parameters such as temperature, heart rate, blood pressure and electrocardiogram measured with the help of suitable sensors. The proposed solution is in the form of a cloud application which is capable of predicting a disease with the help of machine learning algorithms and data mining methods. New technologies such as AI, ML, advanced data science-based decision support systems could assist and support a registered medical practitioner (RMP) on patient evaluation, diagnosis or management, but the final prescription or counseling is directly delivered by the RMP. Healthcare information holds a collection of textual and numeric medical data including symptoms, diagnosis, treatments, medications, outcomes and other such related data. Data mining is employed to derive valuable knowledge from textual data. Text mining performs two important tasks, namely entity recognition and relation extraction using some of the best and well-suited computational algorithms such as K-means clustering, Naive Bayes classifier, K-nearest neighbor (KNN), support vector machines (SVM) and decision tree [6]. In addition to the prediction of a disease, the cloud application generates a report comprising of suggestions on how to treat the disease which helps a doctor to give an improved remedy. Figure 2 illustrates the overall working of the proposed system.

Additional features:

Authentication, retrieval and track of health data:

- The Aadhar number serves as a unique ID to retrieve a patient's health record [7].
- Once a stranger tries to access the Electronic Health Record (EHR) of a patient using the Aadhar id, an OTP gets generated and sent to the registered patient's

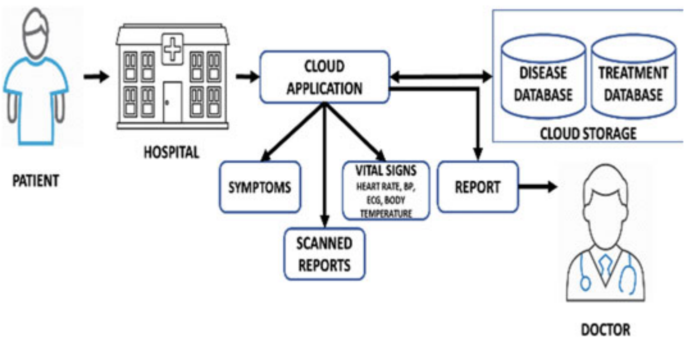


Fig. 2 Overview of the system

mobile number. This feature ensures that only certain authorized or authenticated software or devices or personnel access the data.

- The cloud application facilitates the patient to view the list of persons who had already viewed the patient's profile.

4 Working Process

Firstly, the patient who visits the hospital needs to sign up or get registered with the application by using email ID, phone number, name and password. The next step involves the user to request for an OTP, and the user gets authenticated with the generated OTP. From the next visit onward, the user can directly login with the username and password.

A login is a collection of digital credentials employed to validate a user. It is also used to gain access and control of computers, networks as well as online accounts and other services. All the user credentials get stored in the database folder of the application. Once the initial application registration is successfully completed, the user gets accessed to the system. With the assistance of a hospital staff, the user feeds the symptoms and past medical reports (if any) into the system. Also, the basic parameters such as body temperature, heart rate, BP and ECG are measured using a suitable device, and these medical data are recorded in the system. After the data entry process, the symptoms, the patient past medical reports and the measured parameters are stored and analyzed in the database.

Data mining and machine learning algorithm extract the usable data, automatically predict meaningful pattern based on analysis and output the disease accurately. Data mining involves a five-step process to extract useful information:

1. Source of information is determined.
2. Appropriate data points are selected for analysis process.
3. Relevant information is derived from the data.
4. Key values are identified from the derived data set.
5. Results are interpreted, and output is generated.

5 Database Work

The cloud application provides an instantaneous database, distinctive application programming interfaces, several authentication patterns and hosting channel. The Firebase real-time database, a cloud-managed database is used to access the data [8]. Data are collected as JavaScript Object Notation and integrated in real time to each associated client. As an alternative to conventional hypertext transfer protocol (HTTP) requests, the Firebase platform employs synchronization of data, i.e., at any time data modify every single associated system encounter that update in less than

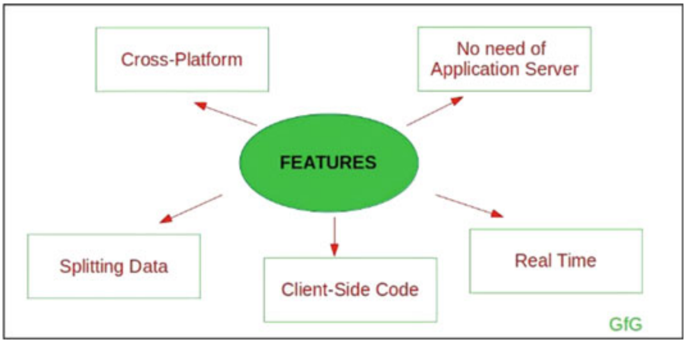


Fig. 3 Features of Firebase real-time database

seconds. The Firebase platform comprises a wide range of features as illustrated in Fig. 3.

It delivers collective and immersive exposures with no consideration about networking code. The Firebase console is helpful to create substantial applications by enabling safe approach to the database straightly from the client-side code, data is maintained handily, and even when in a disconnected state, real-time occurrences proceed to function, providing with the end user an approachable experience [9]. When the system retrieves connection, the real-time database integrates the data modifications with the remote updates that resulted while the client was offline, affixing any problems accordingly.

Diseases prediction using traditional risk model involves machine learning and supervised learning which store data in the database to give the accurate results for each of the symptoms given. The patient's information, results and history of diseases are recorded in EHR which makes information available instantly and securely to the authorized user.

5.1 Data Retrieval

Retrieving data in the admin software development kit is established in a distinct method across various programming languages.

1. **Asynchronous listeners:** Data saved in a Firebase console are re-acquired by coupling an asynchronous listener to a database reference. Initially, the listener is generated once, then it is actuated each time, when the data get altered. Many kinds of events are received by the event listener. Especially in Python, Java and Node.js Admin SDKs, asynchronous listeners are supported.
2. **Blocking reads:** Data saved in a Firebase console are re-acquired by requesting a blocking technique on a database reference, which sends back the data stored

at the reference. Each method request is a onetime activity. This technique of retrieving the data is accepted in Python and Go Admin SDKs.

5.2 Security in Firebase

The process of developing and going over a cloud application with **Firebase** is way straightforward, so is its protected design and data retrieval **security** rules [10].

Secure socket layer shortly known as SSL is a typical security technology in which Firebase is hosted. It is used to establish a secured and encrypted link between a client and the server. To secure a Firebase database, there are two options:

- To incorporate a conditional access pattern as a rule
- To verify the data before it is being introduced to the database.

In addition to security, privacy is also an important consideration, particularly in the domain of health care. The system collects and stores intricate details of patient's vital parameters. For privacy considerations, a right encryption and digital signature mechanism is employed, so that only the people with the right to know get access to the healthcare data which is coming from a healthcare system.

6 Implementation of WEKA Tool

In general, the real-world data are incomplete, noisy and inconsistent. Due to this fact, the data need to be preprocessed. Major tasks involved in preprocessing include data cleaning, data integration, data transformation, data reduction and data discretization [12]. Initially, the data set is loaded in the 'weka explorer'. Immediately when the data set is inputted, WEKA will identify the attributes and at the time of analysis of the data will evaluate some primary statistics on every attribute. The selecting or filtering process is done using the attribute filters in weka. Feature selection involves two methods, namely attribute evaluator and search method. Feature selection—ChiSquaredAttributeEval—the most common correlation measure for categorical data is the chi-squared test. Chi-square is a statistical test applied to the groups of categorical features to evaluate the likelihood of correlation or association between them using their frequency distribution. Search method—best first search—is used to navigate attribute subsets as illustrated in (Fig. 4).

A few approaches, namely association rule mining, might only be executed on categorical data. This demands executing discretization on continuous or numeric attributes. The next step is the concept of classification which is basically to distribute data among the various classes defined on a data set. Classification algorithms acquire a knowledge of this type of distribution from a provided set of training and then try to classify it correctly when it comes to test data for which the class is not specified. The values that specify these classes on the data set are given a label name and are

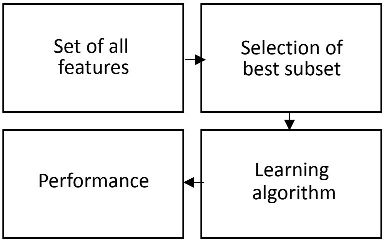


Fig. 4 Feature selection approach

used to determine the class of data to be given during the test. In the classify section, the lazy, K-nearest neighbors algorithm is selected. Configuration is optional, but highly recommended. Configuration is established to get the best results from an algorithm.

7 Algorithm

K-Nearest Neighbors Algorithm

K-nearest Neighbor algorithm is most popularly used in data mining and predictive analysis. Classification is achieved by detecting the nearest neighbor to decide on the type of an unknown model. This algorithm stores the information about the previous data and generates the report based on the new cases being generated. KNN algorithm is based on memory, and it is the basic form of simple machine learning method [11].

KNN splits the data from the data set into two categories, namely medical parameters and non-medical parameters.

Medical parameters: blood pressure, electrocardiogram, body temperature, heart rate, symptoms
Non-medical parameters: age, gender, height, weight.

Medical parameters gain higher priority than non-medical parameters. As a result, only the essential data are extracted from the input given by the patient.

The KNN algorithm involves three basic steps: distance calculation; finding closest neighbors and voting for labels. Euclidean distance method is used to calculate the distance.

Euclidean $$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$$

In the next step, an optimal ‘k’ value is selected using cross-validation [13].

Steps to compute KNN algorithm are as follows:

1. Decide on parameter $K =$ number of nearest neighbors.
2. Evaluate the distance between the query instance and all the training patterns.
3. Rank the evaluated distance and identify nearest neighbors derived from the K-th minimum distance.
4. Collect the class of the nearest neighbors.
5. Employ a simple majority of the category of nearest neighbors as the prediction value of the query.

Symptom Data Set

A training data set is a collection of data used to train a machine for achieving a set of actions. An enriched data set will lead to the enhanced performance of the machine. The structured data set as shown in Table 1 is used by the application to produce a desired output. The ML algorithm enables the application to predict diseases based on the given symptoms.

Table 1 Sample symptom data set

Disease	Symptom 1	Symptom 2	Symptom 3	Symptom 4	Symptom 5
Stroke	Sudden numbness	Blurred vision	Speech loss	Dizziness	Severe head ache
Tuberculosis	Coughing up blood	Chest pain	Unintentional weight loss	Night sweats	Fever
Dengue	High fever	Head ache	Skin rash	Severe joint and muscle pain	Nausea, vomiting
Malaria	Fever	Chills	Abdominal pain	Nausea	Headache
Cholera	Diarrhea	Dehydration	Low blood pressure	Rapid heart rate	Vomiting
Cardiac arrest	Loss of consciousness	Fatigue	Chest pain	Fainting	Heart palpitations
Typhoid	High fever	Head ache	Stomach pain	Loose stools	Vomiting
Influenza	Fever	Chills	Muscle aches	Runny nose	Congestion
Ischemic heart disease	Chest pain	Extreme fatigue	Heart palpitations	Shortness of breath	Swelling in legs, feet or abdomen
Corona virus	Fever	Cough	Shortness of breath	Sore throat	Tiredness

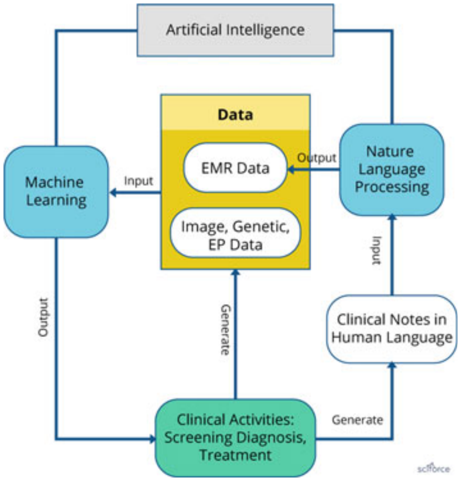


Fig. 5 Data processing and output generation

8 Data Mining Technique

Electronic health record (EHR) is mainly classified into three types—structured data, semi-structured data and unstructured data. Structured data comprise basic information such as birth data and vital signs. Semi-structured data include name, value and timestamp. Unstructured data are a kind of complex data consisting of clinical notes, surgical records, pathology reports, etc. Data are extracted from EHRs using data mining techniques. It is difficult to carry out data mining and analysis on EHR as it holds the characteristics of redundancy, diversity and incompleteness. Therefore, the data require various preprocessing techniques. Preprocessing of data includes data cleansing, data integration, data reduction, data transformation and privacy protection [14]. Data mining is mainly for the semi-structured and unstructured data. The fundamental approach is to convert semi-structured and unstructured texts into computer-readable format by means of information extraction and natural language processing (NLP) mechanisms [15]. In this process, the important mechanisms involved include named entity recognition (NER) and relation extraction (RE) as shown in (Fig. 5).

9 Illustration

The data requirements of the cloud application, classification of the input data and report generation are illustrated in the below section.

9.1 Electronic Health Record

Figures 6 and 7 depict the copy of a patient's health record, which holds several clinical data such as patient's age, sex, complete blood count, differential count, erythrocyte sedimentation rate (ESR), hemoglobin, blood sugar and CRP (c-reactive protein).

Patient : Mrs. VANI
SID .No : 22003381 Age / Sex : 40 Y / Female
Referrer : M.C.SRIDHAR

Patient ID : 2200008250
Rec Date/Time : 16/03/2020 08:38
Prn Date/Time : 16/03/2020 18:14
Page : 1 / 3

Final Test Report

		Biological Reference Interval
HAEMATOLOGY		
COMPLETE BLOOD COUNT		
TOTAL WBC COUNT (Method : Automated Cellcounter) (Specimen: EDTA BLOOD)	: 9560 cells/cumm	at Birth:10000 - 26000 cells/cumm up to 1 Year : 6000 - 16000 cells/cumm 2 - 6 yrs : 5000 - 15000 cells/cumm 6 - 12 yrs : 5000 - 13000 cells/cumm Adult : 4000 - 10000 cells/cumm
DIFFERENTIAL COUNT		
Neutrophils (Specimen: EDTA BLOOD)	: 66 %	40 - 80 %
Lymphocytes (Specimen: EDTA BLOOD)	: 27 %	20 - 40 %
Eosinophils (Specimen: EDTA BLOOD)	: 2 %	1 - 6 %
Monocytes (Specimen: EDTA BLOOD)	: 4 %	2 - 10 %
Basophils (Specimen: EDTA BLOOD)	: 1 %	0 - 1%
ESR 1 Hour (Specimen: CITRATED BLOOD)	: 19 mm	Male (< 70 Years) : 5 - 15 mm. Male (> 70 Years) : 5 - 30 mm. Female (< 70 Years) : 5 - 20 mm. Female (> 70 Years) : 5 - 35 mm. In pregnancy (First half) : Upto 48 mm. In pregnancy (Second half) : Upto 70 mm.
HAEMOGLOBIN (Method : Automated Cellcounter) (Specimen: EDTA BLOOD)	: 14.1 gms/dl	Male :13 - 17.0 g/dl Female :12.0 - 15 g/dl

Fig. 6 Sample health record—page 1

Patient : Mrs. VANI
SID .No : 22003381 Age / Sex : 40 Y / Female
Referrer : M.C.SRIDHAR

Patient ID : 2200008250
Rec Date/Time : 16/03/2020 08:38
Prn Date/Time : 16/03/2020 18:14
Page : 3 / 3

Final Test Report

		Biological Reference Interval
BLOOD - BIOCHEMISTRY		
PLASMA GLUCOSE (FASTING)		
BLOOD SUGAR (F) (Method :Hexokinase) (Specimen: FLUORIDE PLASMA)	: 128 mgs/dl	70 - 110 mgs/dl
CRP (Method : IMMUNO TURBIDIMETRY) (Specimen: SERUM)	: 23.50 Mg/l	Less than 5 mg/L

P. Balapriya.

Dr.P.Bala Priya.,M.D

End of Report

Fig. 7 Sample health record—page 2

9.2 *Patient's Symptoms*

A detailed view of the symptoms undergone by the patient is collected and listed as presented in Table 2.

Table 2 Patient symptoms

Symptoms	
Symptom 1	Suffered high fever 2 weeks ago
Symptom 2	Fever subsided after the intake of antibiotics and paracetamol (suggested by medical practitioner)
Symptom 3	Allergy
Symptom 4	Pain at all joints, especially at foot

9.3 Measured Parameters

Some of the vital clinical parameters, including height, weight, temperature, BP and heart rate, are measured using sensor devices, and the values are recorded as displayed in Table 3.

Table 3 Measured parameters of the patient

Parameter name	Value
Height	155
Weight	60
Temperature	98°F
Blood pressure	124/83 mm Hg
Heart rate	70 bpm

Table 4 Medical parameter

Medical parameter
Total WBC count
Differential count
ESR
Hemoglobin
Plasma glucose
CRP
Temperature
Blood pressure
Heart rate

Table 5 Non-medical parameter

Non-medical parameter
Patient
SID. No.
Age/Sex
Height
Weight
Referrer
Patient ID
Rec Date/Time
Prn Date/Time

9.4 Parameter Classification

Using KNN algorithm, the given clinical parameters are classified into two categories, namely medical parameters and non-medical parameters as shown in Tables 4 and 5. This classification enables the cloud application to work only with the relevant data.

9.5 Report Generation

Based on the inputted data comprising of the patient's health record, symptoms and measured clinical parameters, a detailed report gets generated as illustrated in Table 6. The report mainly consists of the predicted disease, suggestions for the disease, cause and the effect of the disease.

Table 6 Patient report generated by the cloud application

Patient Id	2200008446
Patient Name	Mrs. Vani
Age/Gender	40 Y/Female
Predicted Disease	Post viral arthritis
Suggestion	Consult a rheumatologist
Suggested medication (Dosage as per doctor's advice)	HCQS (Hydroxychloroquine Sulfate)
	Steroid tablets
	Paracetamol (if required)
Cause	Viral infection
Effect	Joint pain may last for a month

10 Conclusion

The proposed system serves as a crucial utilization in the healthcare field. The implementation of the system is simple and flexible. It is becoming very important to take care of one self by performing an accurate prediction of a disease and obtaining the necessary treatment at the right time, which can be easily done with the help of this cloud-based application.

References

1. Shofi M, Wardhani LK, Anisa G (2016) Android application for diagnosing general symptoms of disease using forward chaining method. In: 2016 4th International conference on cyber and IT service management. Bandung, pp 1–7
2. Luo X, Chang Y, Yang J (2015) An automatic disease diagnosis method based on big medical data. In: 2015 2nd International conference on information science and security (ICISS). Seoul, pp 1–3
3. Sharmila L, Dharuman C, Venkatesan P (2017) Disease classification using machine learning algorithms-a comparative study. Int J Pure Appl Math 114:1–10
4. Maliha SK, Ema RR, Ghosh SK, Ahmed H, Mollick MRJ, Islam T (2019) Cancer disease prediction using Naive Bayes, K-nearest neighbor and J48 algorithm. In: 2019 10th International conference on computing, communication and networking technologies (ICCCNT). Kanpur, India, pp 1–7
5. Gomathi K, Shanmuga Priya D (2016) Multi disease prediction using data mining techniques. Int J Syst Softw Eng
6. Nithya B, Ilango V (2017) Predictive analytics in health care using machine learning tools and techniques. In: 2017 International conference on intelligent computing and control systems (ICICCS). Madurai, pp 492–499
7. Sreesaila B, Abinaya K, Swarnalatha M, Sugumar R (2018) Aadhaar card based health records monitoring system. Int J Innov Res Sci Eng Technol 7(2). (A High Impact Factor, Monthly Peer Reviewed Journal). ISSN (Online): 2319-8753, ISSN (Print): 2347-6710. Visit: www.ijirset.com
8. Khedkar S, Thube S (2017) Real time databases for applications. Int Res J Eng Technol (IRJET) 04(06). p-ISSN: 2395-0072, e-ISSN: 2395-0056. www.irjet.net
9. Khawas C, Shah P (2018) Application of firebase in android app development-a study. Int J Comput Appl 179:49–53. https://doi.org/10.5120/ijca2018917200
10. Shin KI, Park JS, Lee JY, Park JH (2017) Design and implementation of improved authentication system for android smartphone users. 2012 26th International conference on advanced information networking and applications workshops. Fukuoka, pp 704–707
11. Xing W, Bei Y (2020) Medical health big data classification based on KNN classification algorithm. IEEE Access 8:28808–28819
12. Mohamed T (2020). Heart diseases prediction using WEKA. https://doi.org/10.13140/RG.2.2.15160.08960
13. Nagesh Singh Chauhan Om Classifying heart disease using K-nearest neighbors
14. Sun W, Li Y, Liu F, Fang S, Wang G (2018) Data processing and text mining technologies on electronic medical records: a review. J Healthcare Eng 2018:9. Article ID 4302425. https://doi.org/10.1155/2018/4302425
15. Velupillai S, Suominen H, Liakata M, Roberts A, Shah A, Morley K, Osborn D, Hayes J, Stewart R, Downs J, Chapman W, Dutta R (2018) Using clinical natural language processing for health outcomes research: overview and actionable suggestions for future advances. J Biomed Inform. https://doi.org/10.1016/j.jbi.2018.10.005

An Audio-Aided Face and Text Recognition System for Visually Impaired

Manjusha Sreedharan, Shalini Mohanraj, and Lakshmi Sutha Kumar

Abstract One of the biggest difficulties for a visually impaired person is to identify people and read text. This limits the visually impaired to interact socially and puts them at risk. In recent years, several deep learning techniques have been used for face and text recognition. This paper proposes a portable system for face and text recognition using the Raspberry Pi 3B+, Raspberry Pi camera module, customized dataset, few push button switches, and earphones. Multi-task cascaded convolutional neural network and support vector machine are used for face detection and face recognition, respectively. The proposed idea involves efficient and accurate scene text detector for text detection and Tesseract optical character recognition (OCR) for text recognition. Finally, converting the text or face labels to speech by e-Speak tool, a process which allows visually impaired people hear the name of the person/text-written in front of them.

Keywords Face recognition · Text recognition · MTCNN · LSTM · Tesseract · SVM

1 Introduction

According to WHO, approximately 285 million people are visually impaired worldwide [1]. Visually impaired people experience many difficulties throughout their life. These include difficulty in communicating with people, perceiving the atmosphere, and confining motion as they are unaware of the dangers in front of them [2]. If the visually impaired person can "see" the world with the support of present technological advancements, they will attain improved independence and freedom. Simple prototype can be designed to support visually impaired using camera to detect and recognize texts and faces.

Face recognition is considered as one of the most important issues because face plays an important role in recognizing people in social life. It should be customized

M. Sreedharan (✉) · S. Mohanraj · L. S. Kumar
Department of Electronics and Communication Engineering, National Institute of Technology Puducherry, Karaikal 609609, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_2

so that the visually impaired can alter the face dataset according to his/her needs. Face detection is done using deep learning and recognition using support vector machine, and text recognition is done using Tesseract-OCR engine in [3].

Different methods of face detection have been discussed in [4]. It is concluded in [5] that for real-time applications, deep learning methods like multi-task convolutional neural network (MTCNN) are better than conventional methods designed by Viola and Jones. A new method in [6] cascades two processes of face detection and face alignment using multi-task CNN. Face recognition is done using a multi-class SVM classifier in [7]. Python scikit learn libraries are used as mentioned in Scikit documentation (2011) [8]. Text detection is done using an efficient and accurate scene text detector (EAST). Text recognition is done by Tesseract OCR which includes a new neural net called long short-term memory (LSTM)-based OCR engine, which concentrates on line recognition and recognizes character pattern. The recognized face or text is converted to audio output by a text-to-speech (TTS) system that converts text into speech which is explained in [10].

Krishna et al. [11] implementation includes a smartphone that obtains images using the phone's camera and wirelessly transceives to a remote server for recognition. PCA and Microsoft speech engine were used. Chillaron [12] implemented a face detection and recognition application developed using Raspberry Pi hardware and an interface for Android smartphones. The object detection and face recognition are based on the boosted cascade and eigenfaces, respectively. Rani [13] describes a smart glass that can automatically detect, examine, and recognize text and provides voice hints for visually impaired. The prototype was based on the Intel Edison Platform, and recognition was done using Tesseract software.

Aravind and Roshna [14] have constructed a model for text recognition that continuously acquires images, finds textual characters, converts it into speech, and reads out loud. The shortcoming of this model is that the computing device is not mobile which is disadvantageous to the visually impaired. Balduzzist [15] implemented model using a compact computational system that acquires a video and inspects it to detect faces in the frame. The face recognition module uses local binary patterns (LBP) to determine the detected faces. Zhou et al. [9] implemented a text detector called efficient and accurate scene text detector. The model gives the arbitrary location of the textual data and is highlighted by a boundary box in the image.

This paper aims to build a customized camera-based assistant for face and text recognition to help the visually impaired. The idea of the design is in such a way that it initiates with a choice between face or text recognition by the visually impaired. He/she can toggle between both as per his/her necessities. Even though earlier works support visually affected people, most of the prevailing systems use MATLAB as [16]. Handful of the designs use laptops which are not portable. Algorithms used in the previous systems are not accurate and efficient.

The rest of the paper is organized as follows. Section 2 deals with the design of prototype and explanation of the model; a detailed description of the algorithms used in the project is described in Sect. 3; Sect. 4 includes the hardware components used in the design and discusses the experimental implementation of the proposed system. Based on the design and analyses, conclusions and future scope are provided in Sect. 5.

2 Proposed System

The block diagram in Fig. 1 shows the overall working of the prototype. Face dataset for a single person is developed by taking around 20 photos of him/her in different angles. 80 photos of 4 known persons (friends/relatives of visually impaired) are used as the customized dataset in this paper. During the testing phase, a new photo of the person (any one of 4 persons available in the customized dataset) is taken and recognized. It is classified as unknown if it does not belong to existing dataset. Python programming language is used to program the Raspberry pi. Face recognition and text recognition programs are coded separately and later integrated. The GPIO pins of the Raspberry pi are programmed for mode selection using push buttons.

B1, B2, and B3, three push buttons, used to initiate face recognition, text recognition, and to stop Raspberry Pi, respectively. The user can choose between either face recognition or text recognition as per the requirement. The buttons are connected using the jumper wires from a breadboard to RPi. The camera is switched ON when the button 1 is pressed. During the face recognition session, face detection uses multi-task CNN to detect faces in the scene, and the detected face is passed on for recognition. Face recognition is done by the SVM classifier which recognizes the person (He/She must be one of the persons in the customized dataset for recognition).

Similarly, when the push button selects the text recognition, the camera is switched on for text recognition. When text is detected by EAST detector, it is captured and pre-processed. The pre-processing and recognition are done by Tesseract. Then, the

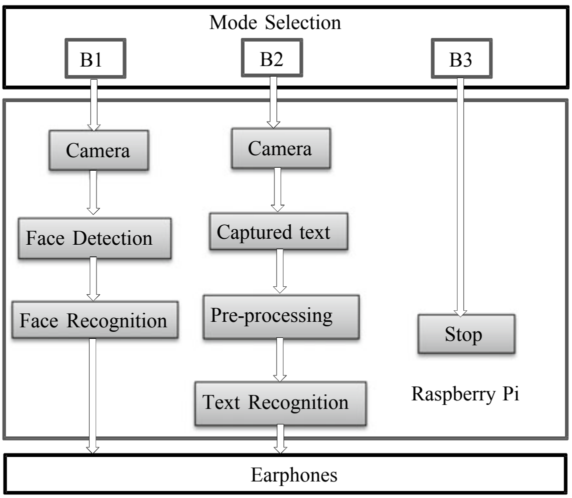


Fig. 1 Block diagram for the face and text recognition system

recognized text is converted to voice. The third button, B3 stops the process and when pressed once more shuts down the Raspberry Pi.

3 Algorithms

3.1 MTCNN

Multi-task CNN, proposed in [19], is used from the integration of face detection and alignment tasks. An image pyramid is formed in which the images are rescaled into many sizes in order to detect faces of all sizes. This is then passed through the cascade of networks- proposal network or P-Net, refine network or R-Net, and output network or O-Net. This then outputs the probability of a face being in the box, the coordinates of the bounding box, and the coordinates of the facial landmarks (locations of the eyes, nose, and mouth). An additional process called non-maximum suppression is performed between these stages to reduce the number of bounding boxes. After the isolation of the image from the background and pre-processing using Dlib, the face needs to be represented in numerical embedding. This is done using pretrained deep neural network OpenFace which is explained next.

3.2 Model Used—nn4.Small2.V1

Two largest publicly available datasets are used for training the current face recognition model as of August 2015. These include FaceScrub and CASIA-WebFace. It uses triplet loss function which is explained in the next section.

3.2.1 Triplet Loss

The embedding is represented by $f(x) \in$ Rd. It embeds an image into a d-dimensional Euclidean space which is constrained into a d-dimensional hypersphere, i.e., $\|f(x)\|2 = 1$. Model is trained such that it learns an image embedding which is done by calculating the L2 distance between image of the same identity (anchor image and positive image) and between different (anchor image and negative image) identities which will be small and large, respectively. Here, a stands for anchor image, *p* stands for positive image, *n* for negative image for any identity *i,* and α stands for least margin.

$$L = \sum_{i=1}^{m} [\left|| f(x_i^a) - f(x_i^p) |\right|_2^2 - \left\| f(x_i^a) - f(x_i^p) \right\|_2^2 + \alpha]_+ \quad (1)$$

Means max (z, 0) and m is the number of triplets in the training set.

The receiver operating characteristics (ROC) curve in Fig. 1 from [20] shows comparison of ROC curves among recognition algorithms. The ROC curve of the nn4.small2.v1model covers much area than of OpenBR v1.1.0, eigenfaces, OpenFace nn4.small2.v1 folds. This is a smaller announcement of the latest model, nn4.small2.v1, which improves the LFW accuracy from 91.5 to 93.6%. It uses a smaller model and improves the runtime performance. To identify new inputs, the classifier trained with the database entries (embedding) is required. In this paper, a linear support vector machine is used for this purpose.

3.3 Support Vector Machine

The purpose of SVM algorithm is to identify a hyperplane in an N-dimensional space that classifies the data points precisely. Linear SVM is the newest exceptional machine learning algorithm used to solve multi-class classification for extremely large database. For both binary classification and multi-class classification ($k = 2$ and $k \geq 3$, respectively), support vector classification (SVC) is very promising.

3.4 EAST

The Efficient and Accurate Scene Text Detection (EAST) architecture is a two-step process to detect the region of interest. The EAST detector consists of a fully convolutional network (FCN) model that produces word or text-line predictions. The detected text regions can be either rotated rectangles or quadrangles that are subjected to non-maximum suppression, after thresholding to return final results. The architecture was created by taking different sizes of word regions into account. Small word regions that use low-level features are detected in initial stages, while large word regions are detected from the later stage of the network.

3.5 Tesseract OCR

Tesseract software is labeled as one of the most accurate open-source OCR engines. Optical character recognition systems transform a two-dimensional image of text, from its image representation into machine-readable text. It makes use of long short-term memory (LSTM) which is a part of recurrent neural networks. Graves introduced bidirectional LSTM architectures for examining text in both forward and backward directions. These layers are later connected to a single output layer, and thereby, it ensures the most accurate prediction [17].

3.6 Pyttsx

There are several APIs available to convert text-to-speech in Python. One of such APIs is Python text-to-speech (pyttsx). The advantage of pyttsx is that it works offline, unlike other text-to-speech libraries. Rather than saving the text as an audio file, the pyttsx speaks it there. This makes it more reliable. The volume, pitch, and gender of the audio output can be changed as per the needs of the user.

4 Hardware Implementation and Working

The user collects the photos of his friends and relatives. As explained in Sect. 2, this model uses 20 images of four people. These collected photos of each person are saved in separate folders. A Python script which converts these photos into matrices using the Numpy library is written. The nn4.smallv2 model is then loaded. Using this pretrained network, embedding vectors are calculated. From [19], it can be seen that the validate rate of 128 dimensions is the highest, i.e., 87.9% $\pm$ 1.9. Figure 2 shows the squared L2 distance calculated using (1) between anchor-negative and anchor-positive pairs and L2 distance of anchor-negative pair is higher than the L2 distance of anchor-positive pair. After the embedding is calculated, the classifier is trained by them.

The Scikit library in Python with the function of linear SVM classifier (SVM) is used to train the dataset and classification. Therefore, the whole script is divided into four parts. First MTCNN detects the face as explained in Sect. 3.1. The second part is for pre-processing using OpenCV and to find the embedding using Numpy libraries. The third part deals with recognition, and it is done as SVM classifies it. Finally, the last part aims to convert the name of the person (label of the classifier) to voice. The text recognition script consists of three sections. The first section is used for text detection using EAST detector, to produce bounding boxes of these textual content. The second part recognizes the text using Tesseract. The last part converts the recognized text to voice.

An SD card with Raspbian OS and along with the models used for face recognition is loaded in the micro-SD card slot of the raspberry pi. The designed prototype along

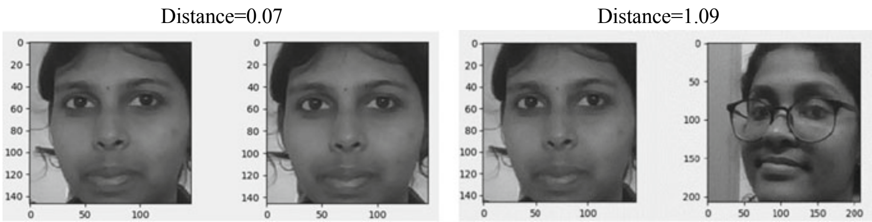


Fig. 2 Distance between similar and dissimilar images

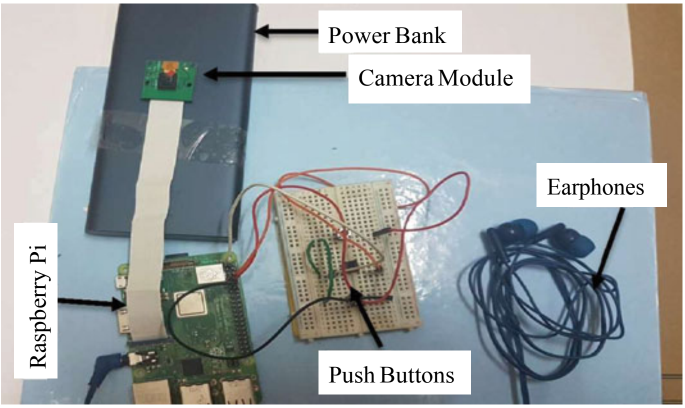


Fig. 3 Hardware setup

with its connections is shown in Fig. 3 where the camera module is connected to the raspberry pi using a ribbon cable. The push buttons are connected to the GPIO pins using jumper wires. The power bank is connected to the raspberry pi power input to start the process. The user can choose the face or text recognition as per needs.

If the first button is pressed, face recognition is started. This is initiated with the switching ON of the camera. The camera then searches the frame for a face. The face detection is done by MTCNN as explained. Each face in the frame is extracted using bounding box regression and face landmark localization. The coordinates of the face are produced as the output. The 128-d vector of the detected face is calculated. In embedding space, Euclidean distances are calculated as explained in Sect. 3.2. The SVC then classifies the face. The label of the classified output and the probability is calculated by Platt's scaling. The next step is to convert the output (label) into audio signals. Thus, the name of the person is given to the Python text-to-speech function. The voice output can be heard using earphones. The stop button is pressed to end the process. It can be started again whenever the user wishes.

When the second button is pressed, the same procedure follows. The camera is turned ON. This can be identified as the red LED is ON. Again, the frame searches for text. The EAST detector produces predictions as explained in Sect. 3.4. These bounding boxes are sent to Tesseract OCR for recognition. The working of Tesseract has been explained in the previous section. Tesseract uses LSTM to predict the next word. Thus, with time, the model gets improved. The user gets the audio using the earphones and can stop the text reading by pressing the stop button.

Four identities have been trained for face recognition. When still images are given to the program, 100% accuracy has been verified as this is a small dataset. In real-time processing, accuracy is represented by the time taken to detect, process, and classify all faces in a frame. Durr [18] suggests that for a video stream, the time for real-time recognition is the time between consecutive frames. We have about 3FPS which is better than [10, 21] which are about 1–2 FPS. The probability of real-time face recognition averages to about 58% as shown in Fig. 4e whereas [21] and [22]

Fig. 4 Screenshots of outputs of **a** face detection, **b** face recognition, **c** text detection, **d** text recognition, **e** face recognition labels and respective probabilities

got around 33%. This paper shows 75% improvement in the probability of real-time face recognition as compared to [21] and [22]. For text recognition, despite EAST and Tesseract being the most recommended software, there are always persistent difficulties in analyzing moving scenes; the model used in this paper given improved accuracy in recognition.

5 Conclusions

The goal of this paper is to build a video face and text recognition entity operating in real-time using a low power embedded system. This has been accomplished using Raspberry Pi, Raspberry Pi camera module, few push button switches, and earphones. The camera is used to get the real-time video, push button to select between face and text recognitions and earphones to hear the output. MTCNN and linear SVM are used for face detection and face classification, respectively. Similarly EAST, one of the best text detector models, is used for text detection, and Tesseract OCR is used for text recognition and word prediction. Python text-to-speech is used for conversion of recognized outputs to audio. English is used in the designed prototype out of the 116 available languages. The face recognition accuracy of the prototype reaches 100% for still images. The probability of real-time face recognition averages to about 58% which is much better than the existing systems. For text recognition, the model designed in this paper has given a fairly good accuracy in recognition of signboards.

The major limitation of this model is the recognizing text of smaller font size. Further work can be done to overcome this and to increase real-time accuracy of face recognition. This paper consolidates the most effective methods in the detection and recognition of face and text. These methods have been used to design and implement a novel hardware prototype. This portable device enables visually impaired user to move around without constraints.

References

1. Pascolini D, Mariotti SP (2011) Global estimates of visual impairment: 2010. Br J Ophthalmol
2. Chaudhry S, Chandra R (2015) Design of a mobile face recognition system for visually impaired persons. ArXiv, abs/1502.00756
3. Smith R (2007) An overview of the tesseract OCR engine. In: Ninth International conference on document analysis and recognition (ICDAR). IEEE Computer Society, pp 629–633
4. Aza, IV, Areni IS (2019) Face recognition using local binary pattern histogram for visually impaired people. In: 2019 International seminar on application for technology of information and communication (iSemantic). Semarang, Indonesia, pp 241–245
5. Ma M, Wang J (2018) Multi-view face detection and landmark localization based on MTCNN. In: 2018 Chinese Automation Congress (CAC), Xi'an, China, pp. 4200–4205
6. Zhang K et al (2016) Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process Lett 23(10):1499–1503
7. Zhao L, Song Y, Zhu Y, Zhang C, Zheng Y (2009) Face recognition based on multi-class SVM. In: 2009 Chinese control and decision conference. Guilin, pp 5871–5873
8. Pedregosa et al. (2011) Scikit-learn: machine learning in python. J Mach Learn Res JMLR 12:2825–2830
9. Zhou X et al (2017) EAST: An efficient and accurate scene text detector. In: Proceedings of conference on computer vision and pattern recognition CVPR, pp 2642–2651
10. Rajesh M et al (2017) Text recognition and face detection aid for visually impaired person using Raspberry PI. In: International conference on circuit, power and computing technologies (ICCPCT). Kollam, pp 1–5
11. Krishna S, Little G, Black J, Panchanathan S (2005) A wearable face recognition system for individuals with visual impairments. In: 7th international ACM SIGAC-CESS conference on computers and accessibility. Baltimore, MD, USA
12. Dunai L, Chillarón Pérez M, Peris-Fajarnés G, Lenua I (2015) Face detection and recognition application for android. https://doi.org/10.1109/IE-CON.2015.7392581
13. Rani K (2017) An audio aided smart vision system for visually impaired. In: International conference on nextgen electronic technologies: silicon to software, pp 22–25. https://doi.org/10.1109/ICNETS2.2017.8067889
14. Sasikumar A (2013) A text reading system for the visually disabled. Int J Res Comput Appl Manage
15. Balduzzi L, Fusco G, Odone F, Dini S, Mesiti M, Destrero A, Lovato A (2010) Low-cost face biometry for visually impaired users. In: Biometric measurements and systems for security and medical applications (BIOMS), IEEE workshop, pp 45–52
16. Mareeswari V, Ramaraj V, Uma Maheswari G, Sujatha R, Preethi E (2019) Accessibility using human face, object and text recognition for visually impaired people. Int J Innov Technol Explor Eng (IJITEE) 8(6). ISSN: 2278-3075
17. Graves A, Liwicki M, Fernandez S, Bertolami HB, Schmidhuber J (2008) A novel connectionist system for unconstrained handwriting recognition. IEEE Trans Pattern Anal Mach Intell 31(5):855–868

18. Dürr O, Pauchard Y, Browarnik D, Axthelm R, Loeser M (2015) Deep learning on a Raspberry Pi for real time face recognition. https://doi.org/10.2312/egp.20151036
19. Schroff F, Kalenichenko D, Philbin J (2015) FaceNet: a unified embedding for face recognition and clustering. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR). Crossref. Web
20. Amos B, Ludwiczuk B, Satyanarayanan M (2016) Openface: A general-purpose face recognition library with mobile applications. Technical Reports No. CMU-CS-16-118. CMU School of Computer Science
21. Rosebrock A (2018) Raspberry Pi face recognition pyimagesearch https://www.pyimagesearch.com/2018/06/25/raspberry-pi-face-recognition. Last Accessed 01 Oct 2019
22. Rovai (2018) Real-time face recognition: an end-to-end project. https://www.hackster.io/mjrobot/real-time-face-recognition-an-end-to-end-project-a10826#toc-introduction-0. Last accessed 02 Jan 2020

Improving Prediction Accuracy Using Machine Learning Classification Techniques for Alzheimer's Disease in Healthcare Services

L. Shakkeera and K. Sowmiya

Abstract The populace is quickly maturing, and the people who suffer from dementia is to spread around 35 million by 2030, to 65 million today. Alzheimer's disease (AD) is where dementia begins. Person with AD will have memory impairment and finally the people lose his/her ability to carry out the day to day works. The existing model used CNN for classifying Alzheimer patient. Major drawback of CNN is overfitting and it is expensive because it has to take a large dataset for training. To alleviate this, we developed a model for accurate prediction of AD. The proposed system implemented with ROI-based feature extraction, NCA feature selection and tenfold satisfied cross-validation techniques to predict the AD. The advantage of this system is better accuracy which is obtained using SVM classifier. In present days, no treatment can anticipate and fix the Alzheimer's disease in early times. Therefore, this has become the research direction of our proposed work. The accuracy, sensitivity, specificity of the proposed model ranges from 98.15%, 98%, 97%, respectively where the existing model value ranges from 92.50%, 93%, 85%, respectively.

Keywords Predictive modeling · Alzheimer's disease · Region of interest · Correlation-based feature selection · Cross-validation · Support vector machine

1 Introduction

Alzheimer's disease is a chronic condition causing to degenerate and death of the brain cells. Alzheimer's disease is the most common cause of dementia, and it is

L. Shakkeera (✉)
School of Computing Science and Engineering, VIT Bhopal University, Sehore, Madhya Pradesh, India
e-mail: shakkeera.l@vitbhopal.ac.in

K. Sowmiya
Department of Information Technology, School of Computer, Information and Mathematical Sciences, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_3

a gradual deterioration in the ability of an individual to work independently. There is no therapy that prevents the Alzheimer's disease or improves the progression of the disorder in the brain. Complications from the severe brain functions loss can cause dehydration, malfunction or infection and result in death in the late stages of the disease. There are believed to be about 44 million people have Alzheimer's disease or a similar form of dementia worldwide. Therefore, this problem statement becomes our research direction and for many researchers. Identifying Alzheimer's disease (AD) at its initial stages is crucial in the therapeutic trails. Treatment options for Alzheimer's disease can temporarily control the symptoms or slow down the rate of decline. Often, these medications can help those people with Alzheimer's disease to improve their function and preserve independence for a while. Prediction plays a major role in this problem, which can be done with the help of methods of machine learning.

In recent days, in fast growing technological world, machine learning (ML) becomes popular among many of researchers, industries and government sectors. It is the contrivance of automating and cultivating the learning process of systems/computers based on their capabilities/understanding capabilities without being actually programmed by the human beings assistance. The ML process starts with feeding the raw data as an input to train the machines by creating/building machine learning models or learning tasks. Using different ML algorithms uses the input data and targeted output builds ML models or ML tasks. The ML algorithms are based on type of data and kind of task which are trying to automate the machine learning models. The major issues in ML such as understanding which processes need mechanization, nonexistence of excellence data inadequate infrastructure and implementation, and deficiency of skilled possessions or assets (both man and machines). In ML, the process to identify which set of categories are belong to one group is called as classification method, based on the relevant observations. The classification algorithm classifies the data into each of the classes concerned. The plotted is interpreted as a function with the band result for that class is spectral signature or spectral response curve Classification techniques for ML, such as logic-based techniques, Perceptron-based techniques, statistical learning techniques, SVM, and instance-based techniques.

In machine learning algorithms, the approach that uses data mining methods and probability techniques to obtain the expected results is described as predictive modeling. Each model development is prepared with number of predictors, which are prospective attributes or variables for controlling potential outcomes. Nowadays, predictive modeling methods have many applications development in business and medical sectors.

The main contribution of this proposed work to predict the Alzheimer's disease (AD) in earlier stages by increasing the prediction accuracy, sensitivity, specificity values based on machine learning classification techniques.

2 Related Works

Ding et al. [1] "A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG-PET of the Brain". This paper has the prospective of 18F-FDG-PET brain scan images taken from Alzheimer's neuroimaging initiative (ADNI) and unbiased retrospective test set. Final clinical diagnosis was recorded at follow-up. The proposed model CNN is trained with 90% of the ADNI sample group and validated on the rest 20% with standalone test set, performance compared to radiological readers. This model has been analyzed with receiver operating characteristic (ROC), sensitivity, specificity, saliency map and t-distributed stochastic neighbor embedding. By using brain F-FDG, a deep learning algorithm developed to predict Alzheimer's disease early reached 82% specificity at 100% sensitivity, an average of 75.8 months prior to the actual final diagnosis. F-FDG itself is not a specific biomarker for either AD or MCI imaging. Several early diagnosis tools have been developed over the last decade, including increasingly specific disease biomarkers.

Kwon et al. [2] "Prediction and classification of Alzheimer's disease based on combined features from Apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers". The proposed framework in this paper consists of three processing stages: A feature extraction and an early fusion technique, combining multiple features into a single form, the optimal selection of subset feature using a truncated SVD and method of dimensional reduction and classification. Participants were arbitrarily divided into two groups in a 75%: 25% ratio as training and testing sets, respectively, before passing them to the multiclass [3] SVM classifier based on the kernel. In addition, a gray matter atrophy (from the SMR image) and average intensity of each region (FDG-PET image) was automatically extracted from the Nifty Reg toolbox, as well as a set of CSF scores (from the biochemical level). They have obtained unbiased performance estimates for each classification problem. The number of subjects in each group was not identical. Therefore, a comparison of the performance of the different classification experiments is not possible only by calculating accuracy.

Hassan and Khan [4] "A Machine Learning model to predict the onset of Alzheimer disease using potential cerebrospinal fluid (CSF) biomarkers". In this paper, the classification model based on J48 can be used effectively to discriminate mentally disabled AD patient from ordinary well personalities with an accuracy of 98.82%, curve area (AUC) value of 0.992 and sensitivity, specificity of 99.19% and 97.87%, respectively. Based on adequate training attributes chosen using feature selection method, this model predicts patients with early stages of AD.

The efficiency of the model constructed using NB, J48 and SMO was evaluated and compared using various statistical performance evaluators. Neuroimaging data is widely used to classify subject with early stage of AD, and the novelty in the approach is to apply low cost CSF biomarkers adequately to detect AD in its initial stages. The major drawback in this paper is, the present research offers a novel CSF biomarker-based classification method to accurately distinguish a subject from healthy subjects with higher precision and sensitivity with an early stage of cognitive impairment.

Wang et al. [5] "Predictive Modeling of the Progression of Alzheimer's Disease with Recurrent Neural Networks." To predict AD, RNN model of 2 hidden layers with 100 hidden units of each layer was implemented in this paper. In the training process, the learning rate decay and the moving average mechanism of decay were applied. Regularization of L2 was added to the loss function, and Adam optimizer was used to optimize the loss function. Tenfold cross-validation has been used in the experiment. The advantage of the model is an adequate and supportive therapy involves ways to promote lifestyle improvements and brain stimulation for AD patients with the right and successful rates. Hence, understanding and predicting how AD develops over time on an individual patient basis is the key to success in enabling early AD intervention and effectively delivering personalized healthcare [6] services accordingly. The major drawback of this paper is, LSTM requires 4 linear layers (MLP layers) per cell to run at and for each time-step sequence. Linear layers allow large quantities of memory bandwidth to be measured as the device does not have sufficient memory bandwidth to feed the computing units.

Moore et al. [7] "Random forest prediction of Alzheimer's disease using pairwise selection from time series data." They have used random forest to learn about the relationship between pairs of data pairs of data points at different time partings. The input vector is a description of the past of the time series, which includes variable ranging from demographic to non-time such as genetic data. Using demographic, physical and cognitive input data to test the method they use. Data from the TADPOLE grand challenge, an initiative which aims to predict the evolution of the subjects at risk of Alzheimer's disease. For diagnosis, a mAUC of 0.82 and a classification accuracy of 0.73 were obtained following results as compared to a benchmark SVM predictor that gives mAUC = 0.62 and BCA = 0.5. The results show how effective and comparable the methods are with other methods. The PET imaging is expensive and that is major drawback of the system.

3 Existing Predictive Model Using CNN

Figure 1 shows predictive model using CNN classification technique. The existing system develops classification and predictive model that can account for accurate classification grouping and prediction of Alzheimer's disease. This approach will focus on prediction of the disease in the early stage. This system contemplates mechanized diagnosis of Alzheimer's disease in three-dimensional operational MRI brain scans. We use a deep convolution neural network-based classifier to examine the brain MRI images from scans. The system extracts the needed information from MRI scan three-dimensional images and learns and trains the significant information related AD. We perform Gaussian filtering and gray scale conversion in the raw MRI scans. Then, the segmentation process is used as fuzzy C-means techniques to segment the image into several pixels. Then, the pixels or image objects have been sent into feature extraction and the module applied with DWT and GLCM techniques previously transient them to the CNN classifier. We use the datasets of ADNI for classification of

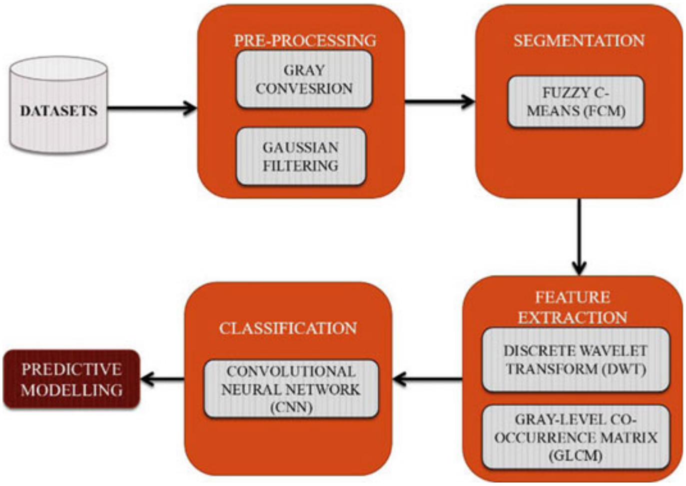


Fig. 1 Predictive model using CNN classification technique

the AD between normal or healthy persona to evaluate the proposed model. The existing system implemented with combination of two techniques in feature extraction module, discrete wavelet transform (DWT) and gray level co-occurrence matrix (GLCM). The major drawback of DWT is that it gets computational intensive for analysis and it is less effective and normal. Where GLCM concerned its high dimensionality of the matrix and the weak correlation of the Haralick texture features are one of the weaknesses of the GLCM method.

4 Proposed Predictive Model Using SVM

A combination of various machine learning methods will be used to implement and develop the model. In the proposed system, we have used a combination of extract superpixel and region of interest techniques for feature selection. Brain MRI images contain enormous amount of information because it has high resolution. Region of interest (ROI) is defined as the regions so at first sight it attracts the attention of the viewer and the focal point of the image. ROI selection allows us to explore the area for unique objects in the region. Traditional approaches for selecting ROI are complex in computational terms and are inaccurate. In this work, a hybrid approach is proposed to select the ROI in brain images which combines the spatially weighted intensity contrasting based on the superpixel. SVM is used in prediction module. Region of interest (ROI) and feature selection neighborhood component analysis are the star of the feature extraction and feature selection modules, respectively. To develop the model, ROI are samples selected for a specific reason within a dataset and it is a way of labeling specific parts of an image. Tenfold satisfied cross-validation technique

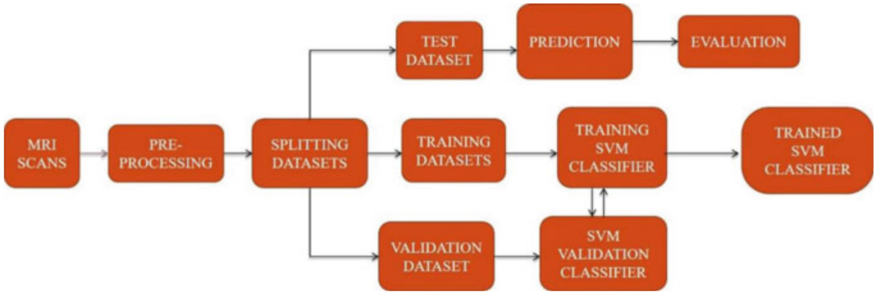


Fig. 2 Operational workflow

is used to estimate predictive models by partitioning the unique samples to train the model and test dataset to gauge it. The predictive modeling in the proposed system uses support vector machine (SVM) and is a formally constructed discriminative classifier with a separate hyperplane. Provided the labeled trainingdata, the SVM algorithm generates an ideal hyperplane in which the new examples are categorized. SVM perform fairly well when there is a large margin of class separation.

4.1 Operational Workflow

Figure 2 shows operational workflow of the proposed system. The input of the proposed system is MRI scans, then it is given into pre-processing function, then it is pre-processed to remove the unwanted noise and distortions, and then the datasets are splitted. 75% of the data are considered as training data and 25% of data are taken as testing data and all the data will be cross-validated. The SVM classifier trained and then it will predict the output based on the trained SVM classifier.

4.2 Predictive Model Using SVM

Figure 3 depicts the predictive model using SVM classification technique.

4.3 Pre-processing

The pre-processing is a mutual term for image processing at bottom abstract level; both input and output are concentration images. The objective of pre-processing the image is to improve the data that removes the unwanted distortions or improves some important features for further proceedings. The proposed

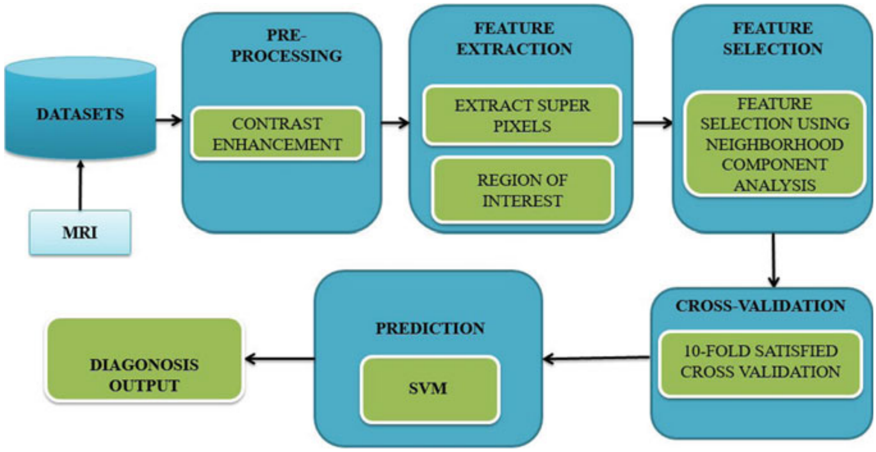


Fig. 3 Predictive model using SVM classification technique

system overwhelms unwanted misrepresentations and enhances the image features for further processing using pre- processing techniques. The pre-processing module use contrast enhancement to pre-process the data. Figure 4 shows pre-processing in predictive modeling.

Contrast Enhancement

Contrast enhancement [8] is a process that enhances the image features by allowing maximum use of colors. Manipulations of contrasts involve changing the range of values within an image to increase contrast. To improve dynamic range of gray level images into quality of high contrast images, contrast enhancement is used. Nonetheless, improving the image contrast without disrupting other image parameters is one of the difficult roles in image processing. The purpose of image enhancement is to improve the interpretability or interpretation of information for human viewers or to provide the best feedback for other image processing techniques.

pout_imadjust = imadjust(I);
pout_histeq = histeq(I);
pout_adapthisteq = adapthisteq(I);

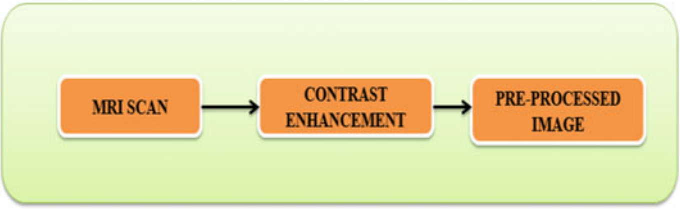


Fig. 4 Pre-processing

Imadjust improves the image contrast by converting the input intensity image values to new values, such that 1 percent of the data becomes saturated at low and high input data intensities by default. Histeq does the equalization of histograms. It improves image contrast by translating values into an intensity map, such that the output image histogram approximately matches a specified histogram (default uniform distribution).

Adapthisteq performs adaptive histogram equalization with contrast limit. Unlike histeq, it works on a small region of data rather than the whole image. Each tile's contrast is improved, so each output region's histogram matches roughly the defined histogram (default uniform distribution). The contrast may be limited to avoid amplification of the noise that may be present in the image.

4.4 Feature Extraction

Figure 5 describes feature extraction module. Extraction of the function defines the related type details in a sequence, so that a standard procedure helps to make the task of classifying pattern easy. In image processing, feature extraction process is a special form of reducing the dimensionality. In this module, we have used two techniques for the process of feature extraction.

- Extracting superpixels.
- Region of interest.

The function of superpixels uses the simple linear iterative clustering algorithm (SLIC) [9]. This algorithm groups pixels into regions with similar values. Using this, regions may reduce the complexity in segmentation. A region of interest (ROI) is a segment of an image that needed to be filtered or performed on. The ROI is defined by creating a binary mask, which is known as binary images of the same size of the images to process with the pixels. To represent the ROI, set to 1 and for others pixel set to 0.

Fig. 5 Feature extraction

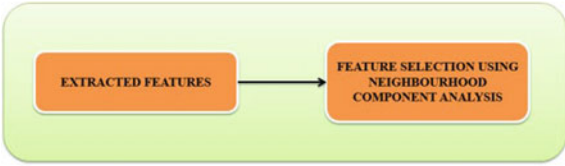


Fig. 6 Feature selection

4.5 *Feature Selection*

Selection of features is to weed out obsolete or redundant features from the dataset. The key difference between selection and extraction of features is that the selection of features maintains a subset of the original features while extraction of features creates brand new ones. Feature selection is the process in which we select those features automatically or manually that mostly contribute to our predictive variable or output we are interested in. With irrelevant features in our data, model's accuracy can be reduced and our model will learn based on the irrelevant features. Figure 6 shows the feature selection process.

Feature selection using neighborhood component analysis: We have used a dataset which consist of 216 images, and based on the segmentation and ROI of image, we have extracted 12 features which are cA, cH, cV, cD, mean, standard deviation, entrophy, variance, smoothness, kurtosis, skewness and IDM. A good feature will always be strongly associated with the class and not redundant with any other features appropriate. Selection of correlation-based features consists of the following important phases, namely.

- Selecting relevant class features.
- Identify redundant features.
- Removes the same from the original dataset.

In this module, we have used feature selection using neighborhood component analysis. It learns the feature weights by regularizing it and by using a diagonal adaptation of neighborhood component analysis (NCA) with regularization. The selected features will be plotted. Zero weight is assigned for irrelevant features. Feature selection using neighborhood component analysis correctly detects the relevant features. Neighborhood component analysis (NCA) is a non-parametric method of feature selection with the goal of optimizing regression and classification [10] algorithm with prediction accuracy.

4.6 *Cross-validation*

The predictive model needs to be trained the machine learning method and to be tested the method. But a terrible approach would be using all of the data to train the algorithm because we would not have any data left to test the method. Reusing the

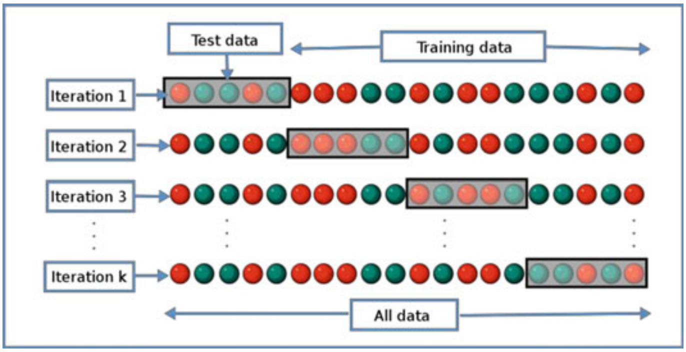


Fig. 7 Cross-validation

same data for both training and testing is a bad idea because we need to know how the method will work on data it was not trained. A better idea is to use 75% of data for training and last 25% of data for testing we could then compare methods by seeing how well each one categorized the test data. But how we will come to a conclusion of using first 75% of the data for training and the last 25% of data for testing is a best way to divide up the data. So to find which block would be best for testing, here, we use cross-validation [11] uses them all, one at a time and summarize the result at the end. Figure 7 shows cross-validation process.

Tenfold Satisfied Cross-Validation: To cross-validate, the data cvpartition is used. It creates a cross-validation partition for data. Cvpartition class object defines a random partition on a set of data of a given size. This partition is used to describe test and training datasets for statistical model validation using cross-validation.

cvp = cvpartition(ytrain,'kfold',10);

It would start by using the first block for testing and the remaining blocks for training and keep the track of how well the method performs. Repeat this process with all the blocks till the k iterations. In the end, every block of the data is used for testing and training by comparing methods by seeing how well they performed. Here, we used tenfold cross-validation it is used where tuning parameters are extracted, i.e., a parameter that is not estimated but just sort of guessed.

4.7 Prediction

Predictive analytics is the use of data, mathematical algorithms and techniques of machine learning to determine the probability of the future outcomes based on historical data. The objective is to go beyond the understanding what has happened and have to provide a predictive capability of what will happen in future. For predictive modeling, we use support vector machine (SVM) algorithm. The classes are identified from the specific training sites on the images are known as prediction process. Mean and variances values of the each training sites are used to classify all the pixels. In this process, initially, build the image representation from all the features mentioned, then train the model with representations. Finally, obtain the class predictions [12] for the input images.

Support Vector Machine (SVM): SVM [13, 14] is a relatively new supervised classification technique for the mapping of land cover communities. When your data is in exactly two classes, we can use a support vector machine (SVM). An SVM classifies data by finding the best hyperplane which separates all data points from those of the other one-class. The best hyperplane for an SVM means the one with the biggest margin between the two classes. Margin means the maximum width of the slab parallel to the hyperplane which has no internal data points. As for any supervised learning model, we first train a vector support system, and then cross-validate the classifier. Using the qualified machine, identify (predict) new data. We can also use different functions of the SVM kernel to achieve adequate predictive accuracy, and we need to tune the function parameters of the kernel.

Training an SVM Classifier

Train, and optionally cross-validate, an SVM classifier using **fitcsvm**. The most common syntax is:

SVMModel = fitcsvm (X,Y,'KernelFunction','rbf',...
Standardize',true,'ClassNames',{'negClass','posClass'});
The inputs are:

X—Predictor data matrix, in which each row is one observation and each column is a predictor.
Y—Class label array with each row corresponding to the value of the respective row in X. Y can be a categorical, character or string array, a logical or numerical vector or a range of vector character cells.
Kernel Function—For two-class learning, the default value is "linear," separating the data by a hyperplane. Choosing a proper kernel function is an important step to successfully train an SVM classifier.
Standardize—Flag stating whether the software should standardize the predictors before the classifier is trained. Class Names—Distinguishes between the classes negative and positive or specifies which classes to include in the data. The first element (or row of a character array) is the negative class, and the second element (or row of a character array) is the positive class.

Algorithm—SVM

Step 1: By using fitcsvm we can train or cross validate support vector machine (SVM) model where obs is table which has the predictor variables and the class labels in the label. svm = fitcsvm(obs,label) returns an SVM classifier trained using the predictor variables in the table obs and the class labels in label.
Step 2: Result = predict(svm,featext_fcm) predicts for the image data in SVM using the trained network featext_svm.

Step 3: If result is zero, it predicts it as Alzheimer or else it will predict it as normal.

5 Results Discussion and Analysis

5.1 Experimental Setup

The proposed system framework has been implemented by utilizing software like, MATLAB 2018a. The data has been collected from Alzheimer's disease neuroimaging initiative (ADNI). The proposed work mainly focuses on developing and providing a system which can diagnose the Alzheimer's disease (AD) in early stages with more accuracy. The complete setup has been implemented and deployed in Windows 10 pro, 64-bit operating system, $\times$ 64-based processor with high-end hardware configurations in the computer system. Thus, the proposed system is implemented with pre-processing, feature extraction, feature selection, cross-validation and prediction.

5.2 Performance Metrics

The relevant performance metrics processes and examination the process involved in creating, testing and providing an efficient model for Alzheimer's disease prediction. The metrics are compared with the existing techniques of classification and prediction models of AD.

Prediction Accuracy: It is the ratio of number of accurate predictions to the total number of sample of input image.
True Positive Rate (Sensitivity): True positive rate mentions the percentage of positive input data points which are properly marked as positive, with respect to all positive input data point.
False Positive Rate (Specificity): False positive rate mentions the percentage of negative input data points which are falsely considered as positive, with respect to all negative data points.

5.3 Performance Analysis

Pre-processing: Figure 8 shows pre-processed images on the basis of three concept adjustment, equalization and histogram equalization, but for further proceeding, we need to select any one of the concept. I have chosen the adjustment image for further process.

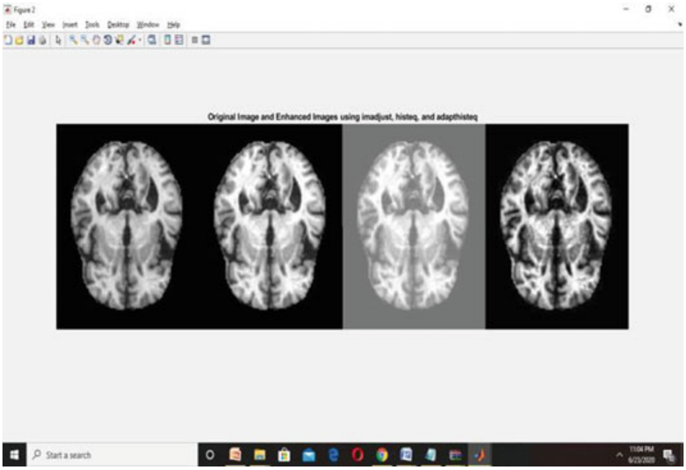


Fig. 8 Pre-processed images

Feature Extraction: Extraction of features is the name of the methods which pick or merge variables into features, essentially reducing the amount of data to be processed while the original dataset is still correctly and thoroughly represented. Here, we have used two concepts for feature extraction, superpixels and region of interest (ROI). Figures 9 and 10 show segmented images and result of ROI images, respectively.

Feature Selection: The selection of features in machine learning and statistics, also known as variable selection, selection of attribute or variable selection of subsets, is the method of selecting a subset of specific feature for using in model construction. Here, we have used feature selection in neighborhood using component analysis (FSNCA) for selecting features and we have selected around 12 features. Figure 11 shows the results of feature selection.

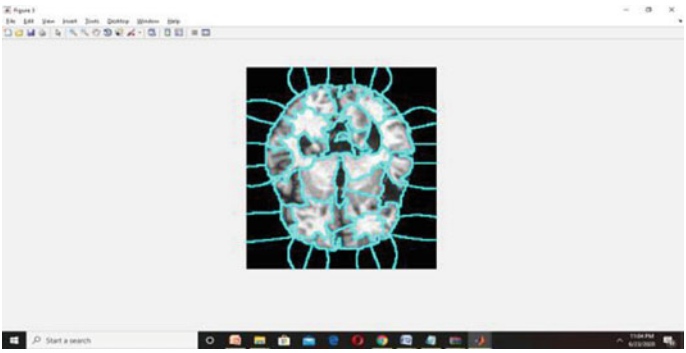


Fig. 9 Segmented image

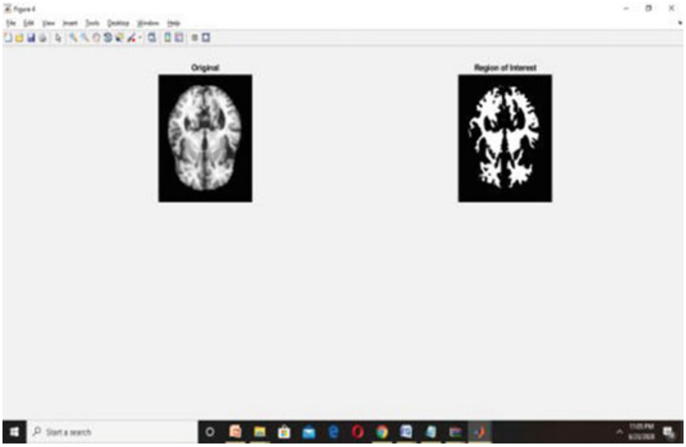


Fig. 10 ROI image

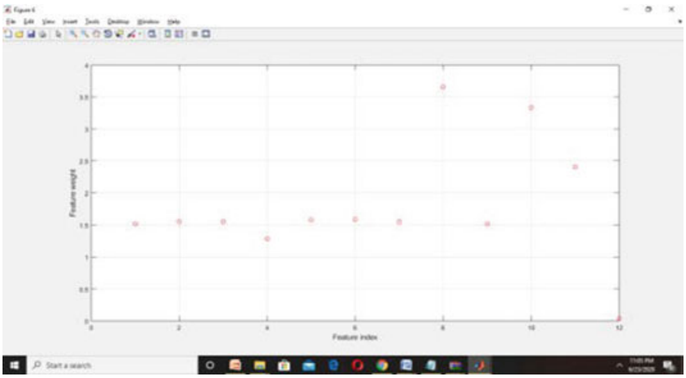


Fig. 11 Feature selection

Based on feature weights and relative threshold, the features are selected. Classifying the observation using selected features. From the training data, extract the features with feature weights greater than zero. Applying a SVM classifier technique using the selected features to the reduced training set. Estimate the classification accuracy of the trained classifier on the test data which has not been used for selecting features.

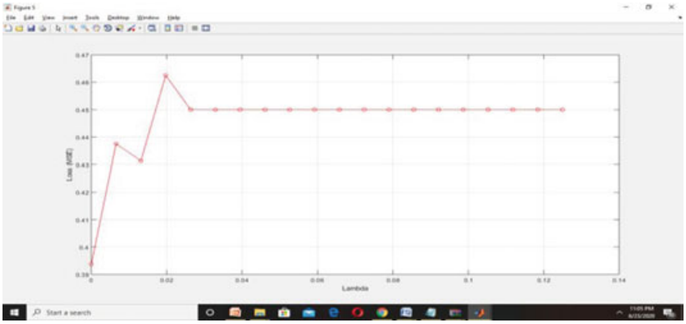


Fig. 12 Cross-validated images

5.4 Cross-validation

Cross-validation is a technique used to assess how statistical analytical results generalize to an independent set of data. Cross-validation is used largely in settings where the target is predicted and the accuracy of the predictive models performance must be estimated. Here, we have used k-fold cross-validation and where $k = 10$. Finding the best lambda value that corresponds to minimum average loss. Fitting the NCA model on all data using best λ and plot the feature weights. Figure 12 describes cross-validated images.

5.5 Prediction

The predicted output is based on the selected features and cross-validated data and we have used SVM as the prediction algorithm. Figure 13 mentions the diagnosis output of the disease.

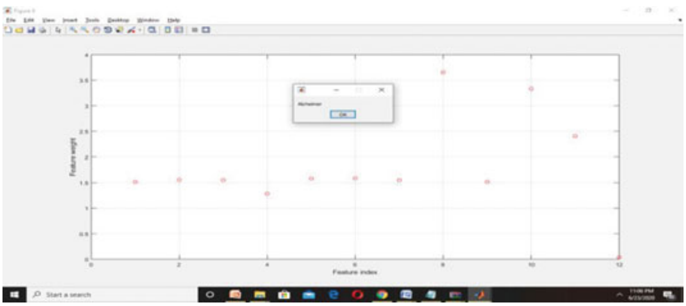


Fig. 13 Diagnosis output

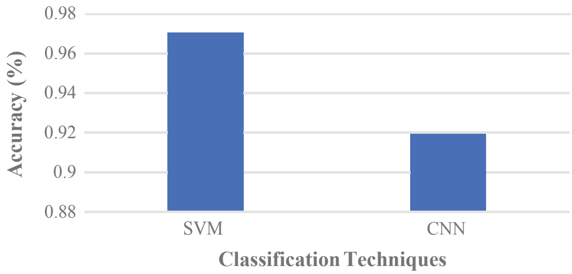


Fig. 14 Prediction accuracy

5.6 Prediction Accuracy

This graph depicts the difference between accuracy between the existing model and proposed model. The existing model has the accuracy of 92.50% where the proposed model has 98.15%, and we have used CNN that has the prediction algorithm in the existing model which has the following drawbacks.

- Overfitting
- It is computationally expensive because it has to take a large dataset for training. Figure 14 shows the prediction accuracy for classification techniques such as SVM and CNN.

5.7 True Positive Rate (Sensitivity)

The sensitivity has been increased in the proposed system the higher the value the system will identify the right person who has AD. This graph depicts the difference between sensitivity in the existing model and the proposed model. The existing model has the sensitivity of 93% where the proposed model has 98%. Figure 15 shows sensitivity percentage values for both classification techniques.

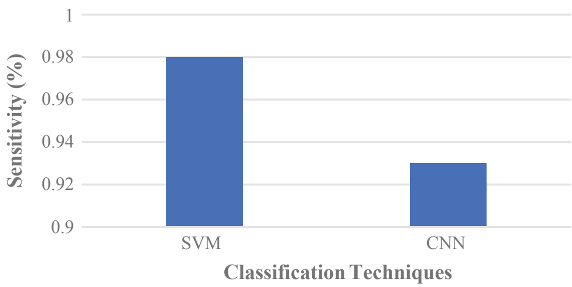


Fig. 15 Sensitivity

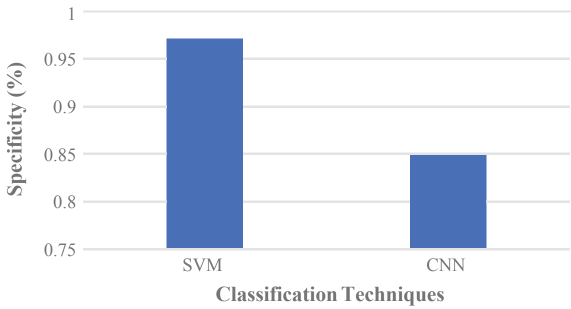


Fig. 16 Specificity

5.8 False Positive Rate (Specificity)

The specificity has been increased in the proposed system the higher the value the system will easily identify the normal patient. This graph depicts the difference between specificity in the existing model and the proposed model. The existing model has the specificity of 97% where the proposed model has 85%. Figure 16 shows specificity percentage values for both classification techniques.

5.9 Performance Analysis of CNN and SVM

Figure 17 shows the performance analysis of CNN and SVM in terms of comparison of accuracy, specificity and sensitivity of existing system and proposed system. SVM is better than CNN because SVM is a margin classifier that supports various kernels to perform this classification. We can use CNN for sequence data but with a large amount of picture it faces difficulties and considers nonlinear correlations.

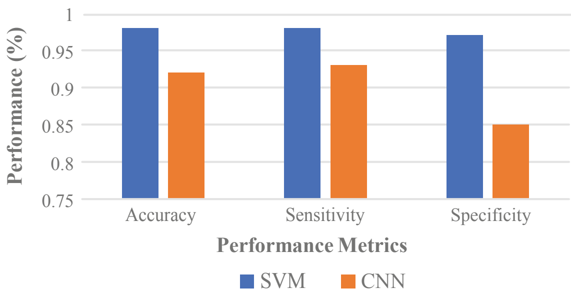


Fig. 17 Performance analysis of CNN and SVM

6 Conclusion and Future Work Directions

6.1 Conclusion

The work proposed in the system focuses on the important challenge of classifying the Alzheimer's disease and develops intelligent classifiers, which can successfully classify that whether the patient has Alzheimer's disease or not. The proposed work implements the concepts of superpixels and region of interest for feature extraction and SVM classification technique and results obtained have shown to be promising. We have shown how a machine learning method can predict AD at early stages. The advantage of this system is better accuracy of 98.15%, sensitivity of 98% and specificity of 97% which is obtained using SVM classifier. The results are better than the existing system and they validate the method as effective.

6.2 Future Work Directions

The future works as follows:

- To predict the disease in future based on the brain images and history of the patient.
- To predict the disease six years prior based on multiple datasets and to obtain maximum results.

References

1. Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, Lituiev D, Copeland TP, Aboian MS, Mari Aparici C, Behr SC, Flavell RR, Huang SY, Zalocusky KA, Nardo L, Seo Y, Hawkins RA, Hernandez Pampaloni M, Hadley D, Franc BL (2019) A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the Brain. Radiology 290(2):456–464
2. Kwon G-R, Gupta Y, Lama RK (2019) Prediction and classification of Alzheimer's disease based on combined features from Apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers. Front Comput Neurosci 13(72)
3. Ramzan F, Khan MUG, Rehmat A, Iqbal S, Saba T, Rehman A, Mehmood Z (2019) A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. J Med Syst 44(2)
4. Hassan SA, Khan T (2018) A machine learning model to predict the onset of Alzheimer Disease using potential cerebrospinal fluid (CSF) biomarkers. RSNA Radiology 290(2)
5. Wang T, Qiu RG, Yu M (2018) Predictive modeling of the progression of Alzheimer's disease with recurrent neural networks. Sci Rep 8:9161
6. Zerka F, Barakat S, Walsh S, Bogowicz M, Leijenaar RTH, Jochems A, Miraglio B, Townend D, Lambin P (2020) Systematic review of privacy-preserving distributed machine learning from federated databases in health care. JCO Clin Cancer Inform 184–200

7. Moore PJ, Lyons TJ, Gallacher J (2018) Random forest prediction of Alzheimer's disease using pairwise selection from time series data. PLoS ONE 14(2)
8. Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, Busa E, Pacheco J, Albert M, Killiany R, Maguire P, Rosas D, Makris N, Dale A, Dickerson B, Fischl B (2006) Reliability of MRI derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. Neuroimaging 32(1):180–194
9. Pan D, Zeng A, Jia L, Huang Y, Frizzell T, Song X (2020) Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning. Front Neurosci 14(259)
10. Mehmood A, Maqsood M, Bashir M, Shuyuan Y (2020) A deep Siamese convolution neural network for multi-class classification of Alzheimer disease. Brain Sci 10(2)
11. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, Folio LR, Summers RM, Rubin DL, Lungren MP (2020) Preparing medical imaging data for machine learning. Radiology 295(1):4–15
12. Islam J, Zhang Y (2019) Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. Brain Inf 5(2)
13. Zhen SH, Cheng M, Tao YB, Wang YF, Juengpanich S, Jiang ZY, Jiang YK, Yan YY, Lu W, Lue JM, Qian JH, Wu ZY, Sun JH, Lin H, Cai XJ (2020) Deep learning for accurate diagnosis of liver tumor based on magnetic resonance imaging and clinical data. Front Oncol
14. Zhu W, Xie L, Han J, Guo X (2020) The application of deep learning in cancer prognosis prediction. Cancers (Basel), 12(3)

Evolutionary Computing-Based Feature Selection for Cardiovascular Disease: A Review

J. Jasmine Gabriel and L. Jani Anbarasi

Abstract Cardiovascular diseases (CVDs) are one of the most lives threatening primary reasons in reducing the mortality of human beings, where one-third of global death is due to CVDs. People who fall under the age group of 75 were mostly affected by CVD. It causes quite 30% of the deaths throughout the world between the ages of 30 and 70. It is high time to identify the disease at the early stage to enhance the expectancy rate through the accuracy of disease prediction. To develop a better prediction, the input to the system needs to refine first, i.e., the selection of appropriate, relevant features. Feature selection is important in the detection of CVDs for better classification resulting in better prediction. However, to meet the challenges in the feature selection process, the evolutionary computing method obtained more attention with improved results comparatively. This paper describes a survey on evolutionary computing-based feature selection techniques, its merits, demerits, and contribution in the classification of CVDs. This detailed, comprehensive work might help the researcher a better understanding of evolutionary computing in heart disease prediction.

Keywords Heart disease dataset · Feature selection · Evolutionary computing · CVDs · Prediction

1 Introduction

Living in this epoch and alter of lifestyle play an important role in affecting the human's physical state and mental condition. Any disorder or malfunctioning of the body or mind that destroys healthiness is known as a disease. Diseases are caused due to various reasons. Every disease has certain traits to spot the categories of diseases.

J. Jasmine Gabriel (✉) · L. Jani Anbarasi
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
e-mail: jasminegabriel.j2019@vitstudent.ac.in

L. Jani Anbarasi
e-mail: janianbarasi.l@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_4

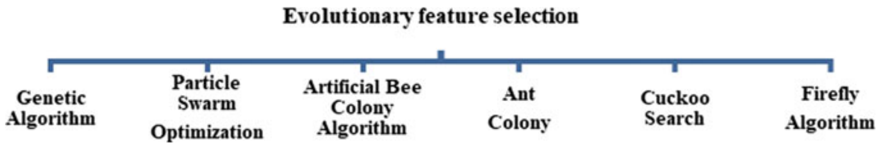


Fig. 1 Evolutionary computing feature selection techniques

World Health Organization defined some virulent diseases in humans are cardiovascular disease a heart disease, cancer, lower respiratory infections, HIV, trachea, bronchus, tuberculosis, and diabetes mellitus. CVDs are a collection of malfunctions in the heart, and the blood vessels connected to it, i.e., the heart does not function properly due to fatty deposits on the stream of blood vessels. Fat prevents the flow of blood to the heart or brain [1]. The traditional methods of diagnosis of CVD are electrocardiogram (ECG) (uneven heartbeat), Holter monitoring (handy ECG device), echo cardiogram (ultrasound of your chest) [2, 3], stress test (raising your heart rate with exercise), cardiac catheterization (heart abnormalities is checked by catheter into an artery), multidetector computerized tomography (CT) scans (different angled X-ray picture of heart and chest), and cardiac magnetic resonance imaging (MRI magnetic field pictures of the heart).

This traditional way of diagnosis may lead to false results due to human error and delayed results due to the number of different tests we take based on the availability of the doctor and the radiologist. The limitation also includes side effects and high cost. It is high time the CVD has to be diagnosed and treated at the earliest. The severity of the disease depends on how early it was detected to increase the mortality of humans. Early detection can be automated by using the heart disease dataset and use them for prediction mechanism. Many researchers and academicians have already proposed various prediction methods. The demanding need for a better prediction model with high accuracy of disease prediction with less cost and less time is required for early diagnosis.

In machine learning, input data is preprocessed by choosing relevant feature selection, and missing value imputation [4] plays an essential role in some real-life applications. Of which, identifying relevant features helps in different fields like false detection, text mining, disease diagnosis, face detection, image processing, bioinformatics, and socio-industrial application. In those applications, the quantity and the quality of features are more critical in the classification and prediction process. The higher numbers of features are challenging to visualize. Irrelevant features create noisy data that affect the accuracy of the prediction model [5]. The traditional feature selection techniques are filter, wrapper, and embedded method. FS can also be a hybrid method (the combination of filter and wrapper methods) and the ensemble method (combining many classifiers into one desired model by aggregating the results) where those methods of feature selection are popular among the researchers in selecting the better feature. Another fast-growing feature selection technique is evolutionary computing, which is inspired by the natural behavior of animals and birds.

Figure 1 shows the 6 meta-heuristic evolutionary computing like genetic algorithm (GA), particle swarm optimization (PSO), artificial bee colony algorithm (ABCA), ant colony optimization (ACO), cuckoo search (CS), and firefly algorithm (FA) specifically carried on in this survey.

2 Literature Survey

Researchers analyzed numerous techniques in developing an automated method for finding a global optimal solution for heart disease prediction by improving its accuracy. This section gives a brief literature survey of different research works using the evolutionary computing-based feature selection specifically for heart disease prediction.

Ismail et al. [6] proposed a novel FS method using BPSO in treating coronary heart disease with EST dataset (ECG data) of 23 attributes of 480 instances. Two feature selection methods like GA and binary particle swarm optimization (BPSO) methods are used for finding the relevant features. The selected features were trained with three different classifiers namely BPSO with support vector machine (SVM) kernel, GA with SVM kernel, and simple SVM method. The BPSO with SVM classifier showed better results in terms of selecting 11 relevant features and improved classification accuracy of 81.46%.

Jabbar et al. [7] proposed a GA-based heart disease prediction for Andhra Pradesh dataset with 12 features using GA-based association rule. Each feature is represents an individual items with the frequent itemset that are derived with minimum support of 6. Fitness function using a G-mean measure generates frequent itemset with two different sets of 7 frequent itemsets are found to be relevant. Frequent itemset creates a lesser number of rules for disease prediction. The author prefers GA for a higher level of prediction and better feature interaction than the greedy algorithm.

Subanya et al. [8, 9] proposed a novel feature selection method with ABC and SVM classifiers. The classifier sets the processing parameters manually for the number of iteration, limit, number of dimensions, and number of employer bee and onlooker bee values. The proposed method showed good classification accuracy compared with forward ranking and reverse ranking. The ABC-SVM showed improved accuracy with seven features. The author identified features like age and fasting blood sugar as notable features in improving the accuracy of 86.76%. The same author enhanced the model with binary artificial bee colony algorithm-K-nearest neighbor (BABC-KNN), for heart disease classification. This model uses the BABC method for feature selection which resulted in selecting 6 prominent features. The selected features are trained using the KNN model with eight existing models for comparison and achieved improved accuracy of 92.4%

Long et al. [10] proposed a new diagnosis method for feature selection using fuzzy classification. The feature selection was carried out by two different methods chaos firefly algorithm rough set-attribute reduction (CFARS-AR) and BPSORS-AR. Comparatively, CFARS-AR is efficient with only four relevant features tested with Cleveland dataset and SPECTF. Fuzzy c-means clustering is used on selected features to obtain the required number of fuzzy and structure of the fuzzy rule. Those values are input to the newly defined approach, interval type-2 fuzzy logic system (IT2FLS) classifier. The proposed model showed improved accuracy of 88.3% over other classifiers like Naïve Bayes, SVM, and artificial neural networks (ANN).

Verma et al. [11] developed a hybrid model with correlation-based FS (CFS) and PSO as the feature selection method for the real-time dataset from Indira Gandhi medical college, Shimla and Cleveland dataset. The k-means clustering tests the selected feature for eliminating incorrectly classified clusters. The resultant features are classified using multilayer perceptron (MLP), fuzzy unordered rule induction algorithm (FURIA), C4.5, and multinomial logistic regression (MLR) classifier. MLP classifier with 7 features achieved an accuracy of 90.28%.

Thippa Reddy et al. [12] proposed a diagnosis model using cuckoo search and classification with fuzzy logic. Cuckoo search (CS) and rough set (RS) helped to pick the most exceptional feature through fitness function without losing its precision value. CS with RS shows reduced 6 features when compared with FA with RS, BAT with RS, and locality preserving projection (LPP). The selected features are fed into the fuzzy classifier to obtain its fuzzy score. The author used three different datasets for comparison. The proposed method CS with RS feature selection with RS classifier of Cleveland heart disease dataset resulted in improved accuracy with 91.5% when compared with other fuzzy models and other datasets used.

Iftikhar et al. [13] proposed a new healthcare model for identifying heart disease risk factors using Cleveland dataset. The proposed model includes 4 test scenarios. In the first and second methods, the model tests with the least square linear method (SVM-LS) and sequential minimum optimization (SVM-SMO-RBF) without feature selection. The third and fourth method tests GA-SVM-SMO-RBF and PSO-SVM-SMO-RBF with feature selection. Of all the four methods, the GA-SVM-SMO-RBF method showed improved accuracy with only seven selected features.

Dulhare [14] developed a prediction system with Cleveland dataset using PSO and GA for feature selection. PSO eliminated features with low PSO values resulting in the selection of 7 relevant features which is better than GA. The selected relevant features are fed into NB classifier for 100 iterations to obtain better accuracy.

Gokulnath et al. [15] defined a wrapper-based feature selection with GA-SVM. The feature selection of GA is compared with existing algorithms like relief, correlation-based, filtered subset, information gain, consistency subset, chi-squared, one attribute-based, filtered attribute, and gain ratio. GA algorithm efficiently chooses 7 relevant features out of 13. The selected features use Z score optimization for preprocessing. The system is tested with classifiers like SVM, MLP, J48, and KNN. SVM classifier showed improved accuracy of 88.34% when compared with the other classifier.

The literature survey focuses on existing EC-based feature selection for diagnosing the heart disease dataset, a lesser number of relevant features, classifier used, accuracy achieved, and the different types of heart disease dataset used. Table 1 explains the overall comparison of this survey.

Table 1 Summarized of evolutionary computing-based feature selection

Reference	FS	Classifier used	SF/TF	List of selected features	Accuracy (%)	Advantage	Disadvantage	Dataset used
Ismail et al. [6]	BPSO	SVM	11/23	NA	81.46	• Minimal feature • Less training and testing time	200 iterations for BPSO	EST dataset
Jabbar et al. [7]	GA	Association rule based	7/12	Sex, blood pressure, resting ECG, smoking, alcohol, family history of CAD, rural/urban	–	• Higher level prediction • Better feature interaction	No performance metric was evaluated	Andhra Pradesh dataset
Subanya et al. [8]	ABC	SVM	7/13	Age, Chol, Fbs, Restecg, Thalach, slope, ca	86.76	Good classification	Parameters are manually chosen	Cleveland
Subanya et al. [9]	BABC	KNN	6/13	Cp, Trestbps, Chol, Thalch, Slope and Thal	92.4	Improved accuracy	–	Cleveland
Long [10]	CFARS- (AR)	IT2FLS	4/13	NA	88.3	• High accuracy • Reduce computation expenses with high dimension data	• Cost is more when used on high dimensional dataset • Training time is slow	Cleveland

(continued)

Table 1 (continued)

Reference	FS	Classifier used	SF/TF	List of selected features	Accuracy (%)	Advantage	Disadvantage	Dataset used
Verma [11]	CFS + PSO	Clustering + MLP	7/13	cp, thalch, exang, old peak, slope, ca, and thal	90.28	• K-means clustering is used	–	Cleveland, Switzerland, and Hungarian
Gadekallu [12]	CS + RS	Fuzzy	6/13	NA	91.5	Less computational cost	• Space complexity can be considered	Cleveland, Switzerland, and Hungarian
Iftikhar et al. [13]	GA	SVM-SMO-RBF	7/13	Cp, Restecg, Thalach, Exang, Slope, Ca, Thal,	88.10	• Better accuracy • Reduced search space	–	Cleveland
Dulhare [14]	PSO	NB	7/13	cp, Restecg, Thalach, Exang, Old peak, Ca, Thal	87.91	• Minimal FS • Better classification	Features were selected at the 70^{th} iteration	Statlog
Gokulnath [15]	Wrapper-based GA	SVM	7/13	Cp, Restecg, Thalach, Exang,Oldpeak, Ca, Thal	88.34	• ROC curve show good performance with SVM	Accuracy is not more than 90%	Cleveland

3 Results and Observation on EC-Based FS

The following are the datasets used by various researchers in this survey for preprocessing the heart disease are EST dataset, Andhra Pradesh dataset, Cleveland, Switzerland, Hungarian, and Statlog heart disease dataset. Table 1 gives the summary of literature survey performed on 10 different papers on heart disease dataset, where all the authors aimed at selecting the lesser number of feature through EC-based FS method for improving the classification accuracy with a suitable classifier.

Figure 2 shows the comparison of the achieved accuracy by processing the heart disease dataset by various researchers. It shows that the BABC method of EC-based feature selection has the highest accuracy of 92.4% with a minimum number of 6 relevant features. Moreover, the accuracy rate ranges from 87 to 92%. Still, a better way of using the EC method is needed to improve the accuracy of heart disease prediction.

Based on the observation, the most of the datasets have these 13 common features in the heart disease dataset. 10 out of 7 authors use the Cleveland heart disease dataset for their proposed work. The attribute of the dataset are age, gender, cp (chest pain), trestps (level of blood pressure at resting), chol (serum cholesterol), fbs (fasting blood sugar), restecg (resting value of ECG), thalach (maximum heart rate), exang (exercise -induced angina), oldpeak (St depression induced by exercise vs. rest), slope (ST segment in terms of a slope), ca (number of major vessels colored by fluroscopy), and thal (defect type). Figure 3 shows the most selected feature using evolutionary computing methods as thalach, cp, restecg, ca, and thal. Dataset referred in this paper is available in the UCI repository [16, 17].

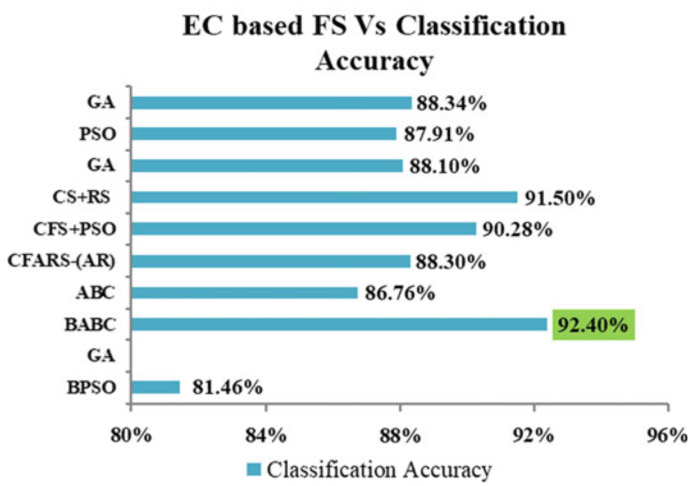


Fig. 2 EC-based FS versus classification accuracy

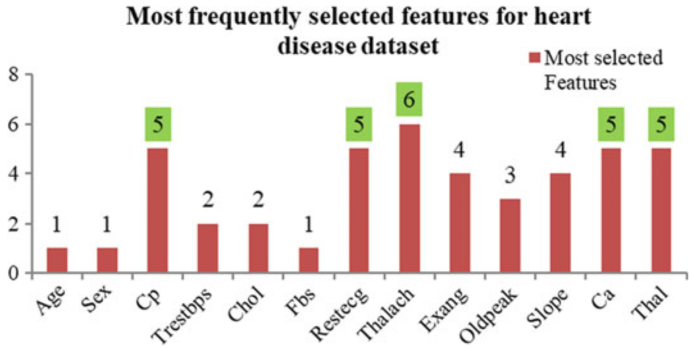


Fig. 3 Most frequently selected features based on the survey

Table 2 gives the insight knowledge about the methods with their strength and weakness of EC-based FS techniques. This brief view can help the researchers in choosing the best method for their problem. Particularly, in dealing with a complex problem and high dimensional dataset, the EC method approach would make a better choice for feature selection.

4 Conclusion and Future Work

The comprehension survey states, more number of features might result in misclassification and overfitting. Most of the researchers focused on two important issues. A suitable feature selection method to obtain less number of relevant features and a better classifier for improving accuracy. The evolutionary computing-based feature selection method met the expectation of a researcher in solving the mentioned issues with efficient identification of global optimal solution, better interaction between features, higher level of prediction, flexibility, and better visual output. The evolutionary computing-based feature selection built a better classifier, a better classifier results in better prediction. Hence, the growing need for handling high dimensional dataset with better prediction automated systems can be constructed for early detection and diagnosis of CVDs. In future work, a fully automated EC-based feature selection can be carried out with a real-time dataset for better heart disease prediction.

Table 2 Strength and weakness of evolutionary computing methods

EC methods		Key points
GA	Strength	• Identify the best attribute by a fitness function • Higher level prediction and find search space easily • Evaluate each subset by executing the model • Better interaction among features
	Weakness	• Sometime, result is in invalid states • Computationally expensive • Poor performance with a high dimensional dataset • Does not guarantee minimal features
PSO	Strength	• Deal with real numbers optimization problems • Implementation is simple and easy continuous optimization • Lesser time complexity
	Weakness	• Does not guarantee minimal features • In high dimensional dataset can fall into the local optimum solution • Slow during iteration
ABCA	Strength	• Fewer control parameters • Flexible and fast during iteration
	Weakness	• Premature convergence where cannot produce offspring's • Low accuracy
ACO	Strength	• Can use for dynamic approach, inherit parallelism • Positive feedback information and discrete optimization
	Weakness	• Challenging to work with high dimension data • Too many parameters were theoretically difficult • Distribution change over iteration
CS	Strength	• Aim to find the global optimal solution • Use fewer parameters • Use of levy flight process finds the optimal solution
	Weakness	• Continuous optimization • Difficult in solving a complex problem
FA	Strength	• Simple mathematical operator • Find global optima faster with a higher success rate • Find the globally optimal solution • Does not worry about initial iteration and velocity as PSO
	Weakness	• Computational expensive both by memory and runtime

References

1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC530409
2. Sharon JJ, Anbarasi LJ (2018) Diagnosis of DCM and HCM heart diseases using neural network function. Int J Appl Eng Res 13(10):8664–8668
3. Sharon JJ, Anbarasi LJ, Edwin Raj B (2018) DPSO-FCM based segmentation and classification of DCM and HCM heart diseases. In: 2018 Fifth HCT information technology trends (ITT). IEEE
4. Jasmine Gabriel J, Valarmathie P (2012) Unified clustering technique for microarray gene expression data. In: International conference on computing and control engineering. ISBN: 978-1-4675-2248-9

5. Hira ZM, Gillies DF (2015) A review of feature selection and feature extraction methods applied on microarray data. Adv Bioinform 1–13
6. Babaoglu I, Findik O, Ulker E (2010) A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. Expert Syst Appl 37(4):3177–3183
7. Jabbar Akil M, Deekshatulu BL, Chandra P (2012) Heart disease prediction system using associative classification and genetic algorithm. Elsevier 1, 183–192
8. Subanya B, Rajalaxmi RR (2014) A novel feature selection algorithm for heart disease classification. Int J Comput Intell Inf 4(2):117–124. ISSN: 2349-6363
9. Subanya B, Rajalaxmi R (2014) Feature selection using artificial bee colony for cardiovascular disease classification. In: 2014 International conference on electronics and communication systems (ICECS)
10. Long NC, Meesad P, Unger H (2015) A highly accurate firefly based algorithm for heart disease prediction. Expert Syst Appl 1–9. Verma L, Srivastava S, Negi PC (2016) A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data. J Med Syst 40(7):1–7
11. Gadekallu TR, Khare N (2017) Cuckoo search optimized reduction and fuzzy logic classifier for heart disease and diabetes prediction. Int J Fuzzy Syst Appl 6(2):25–42
12. Iftikhar S, Fatima K, Rehman A, Almazyad AS, Saba T (2017) An evolution based hybrid approach for heart diseases classification and associated risk factors identification. Biomed Res Int J Med Sci 28(8)
13. Dulhare UN (2018) Prediction system for heart disease using Naive Bayes and particle swarm optimization. Biomed Res 29(12)
14. Gokulnath CB, Shantharajah SP (2018) An optimized feature selection based on genetic approach and support vector machine for heart disease. Cluster Comput 1–11
15. Espejo PG, Ventura S, Herrera F (2010) A survey on the application of genetic programming to classification. IEEE Trans Syst Man Cybernet Part C App Rev 40(2):121–144
16. The UCI machine learning repository [online]. Available at: https://archive.ics.uci.edu/ml/datasets/Heart+Disease
17. http://archive.ics.uci.edu/ml/datasets/Statlog+%28Heart%29

Readiness and Maturity Assessment Model to Measure the Industry 4.0 Ecosystem

Alagiri Govindasamy and Arivarasi Arularasan

Abstract Industry 4.0 is a pragmatic phenomenon happening across the world and clearly defining the silver line between winners and losers from today's market place. There is heightened awareness and seriousness found among the organizations to embrace the Industry 4.0 ecosystem to take first-mover advantage from the industry, but this process is not a smooth one due to inherent challenges from this long journey. Many times, organizations are lacking the necessary competency and skillsets to implement Industry 4.0 framework, and often they are looking for external partners help to chart out tactical and strategic digital road maps. One of the gaps found from the literature is the non availability of the comprehensive Industry 4.0 readiness and maturity assessment mechanism to support organizations. In this article, we are proposing an extensive Industry 4.0 readiness and maturity assessment framework with necessary building blocks to provide an accurate and robust decision support mechanism to organizations. The proposed industry 4.0 mechanism has unique propositions to assess nine different vital organizational constructs-like digital leadership and digital operations and classify companies into five maturity quadrants to make an informed decision on their quest for industry 4.0 journey. This theoretical model will provide an unbiased reflection about the organization's current technical capability and socio-cultural approval for Industry 4.0 ecosystem adoption. Limitations and potential future research areas are discussed in this article.

Keywords Industry 4.0 · Cyber-physical systems (CPS) · Assessment and maturity model

A. Govindasamy (✉)
PMC Global Systems, Chennai, Tamilnadu 600901, India

A. Arularasan
Vellore Institute of Technology, Chennai 632014, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_5

1 Introduction

Organizations across the world are sitting in a crossroad where there never witnessed the high impact from emerging technologies as they are facing on today. Incidentally, Industry 4.0 is a promising framework to address various stakeholders' expectations from present questionable global economy and to excel in the market place [1]. Berman and Bell [2] argued that the success of organizations will get determined by the degree of Industry 4.0 framework adoption and this stand explains the irreversible digital transformation journey option available to all organizations. Various studies highlighted the constructive and larger role from Industry 4.0 on the broader society and labor market as a whole, apart from obvious benefits to organizations like increased productivity improvement and reduced new product development time.

Many times, organizations are lacking the necessary competency and skillsets to measure their current state of readiness to implement Industry 4.0 ecosystem, and often they are looking for external expert's help to chart out tactical and strategic digital road maps. We intend to address the below research questions from this study. Research question 1 (RQ 1): How to define a comprehensive Industry 4.0 assessment framework to measure the current state of organizational readiness? Research question 2 (RQ 2): How to determine Industry 4.0 maturity level for an organization based on assessment outcome? The proposed industry 4.0 assessment mechanism has unique propositions to assess various vital organizational elements to make an informed decision to adopt Industry 4.0 journey. This article explains the innovative industry 4.0 maturity classification model based on assessment outcome.

2 Literature Review

2.1 The Role of Industry 4.0

Industry 4.0 is the strategic initiative from the German government as part of the 'High tech strategy 2020 action plan' in 2011 [3]. The fundamental concept for Industry 4.0 is interconnectedness, to establish a real connection among various stakeholders from Cyber-Physical System [4]. Six important design principles from Industry 4.0 are interoperability, virtualization, decentralization, real-time capability, service orientation, and modularity [5]. One main differentiator of Industry 4.0 from earlier three industrial revolutions is that the fourth industrial revolution is affecting entire organizational functions, not limiting with productivity improvement objectives noticed from the previous three industrial revolutions [6].

The motto of Industry 4.0 is to shrink the gap between physical and digital paradigms by the autonomous exchange of data among critical elements like a machine, man, sensors, actuator, and products and establish intelligent coordination among them without manual intervention. Mrugalska and Wyrwicka [5] explain the important role played by three important components namely the Smart Product, the Smart Machine, and the Augmented Operator and this author explained the critical role played by each of these components. In the conventional business context, human beings are wielding the upper hand in controlling other resources, but in case of Industry 4.0, the cyber-physical system (CPS) will results in the seamless communication between all stakeholders like machines, man, sensors, and products, thus diminishes the dominating role by employees [7]. Important elements like data acquisitions, data processing, machine-to-machine communication, human–machine interaction, and decentralized decision-making environment with intelligent mechanisms are hallmarks from an able cyber-physical system as part of Industry 4.0 [8].

2.2 The Current State of Industry 4.0 Assessment and Maturity Mechanism

'The Industry 4.0 readiness online self-check for business' developed by RWTH Aachen University, Germany is a promising mechanism to assess various parameters from an organization's readiness [9, 10]. The 'Industry 4.0 maturity index' by Acatech given a framework to measures the digital footprint for an organization by comparing the current state and future state [9]. This assessment tool is designed for manufacturing companies and no scope to factor existing business philosophies in a place like a Lean system. 'The System Integration Maturity Model for Industry 4.0' proposed by Leyh et al. [11] offers a mechanism to measure an organization's existing IT landscape with 5 stage model in place. This model is not ideal to assess large organizations and non-manufacturing environments. Provisions to measure existing operational philosophies and inner aspects of the digital system are missing from this model. 'The Industry 4.0 Quick Check-up assessment model' offers a window to measure five dimensions. Few notable shortcomings are non-available metrics to assess existing operational philosophies and no provisions to assess the readiness of all organizational entities.

'The Guideline industry 4.0' developed by VDMA offers a five-stage assessment model to measure the digital maturity model for an organization [12]. One notable limitation from this model is that there is no provision to capture critical organizational parameters like culture, digital strategy, and employees. 'The Industry 4.0 Maturity Model' proposed by the Vienna University of Technology offers a model to measure organizational digital maturity [10]. This model contains nine factors with 62 internal dimensions with a provision to assess each dimension with a corresponding weighting factor. One limitation is that this model does not have any provision to measure existing operational philosophies used and no maturity model suggestions found from this model. 'The Production Assessment 4.0' Developed by Fraunhofer IAO [13] contains five focus areas with a two-stage process. One advantage of this model is that there is a provision to factor the expectations from existing operational systems like total quality management.

3 Gaps from the Literature

Most of the existing Industry 4.0 assessment models are focusing on measuring existing IT maturity levels and the importance of managerial inclination. Cultural and operational parameters are not considered from these existing assessments. Another gap found from the literature is that existing industry 4.0 assessment tools do not have provisions to measure the option to integrate existing operational philosophies into Industry 4.0 implementation. Most existing assessment models are either focused on manufacturing organizations or addressing major organizations alone. There is a gap found to have a universal Industry 4.0 assessment tool to fulfill requirements from all industry segments in all sizes. Finally, many existing Industry 4.0 assessment tools do not proclaim the strategies to reach the desired digital maturity state after the initial assessment is done.

4 Proposed I 4.0 Assessment and Maturity Mechanism

The proposed Industry 4.0 assessment mechanism is poised to fill the gaps identified from the previous research studies and provide informed guidance to the management team to identify driving forces, challenges, and potential areas for future investments. In reality, Industry 4.0 implementation is a challenging one irrespective of organization size and complexity, so the proposed mechanism provides comprehensive coverage of all vital organizational parameters to take an informed decision. Refer to Fig. 1 to see the details. The appendix shows the complete questionnaire from the proposed assessment tool.

Here, we address the 'Research question 1 (RQ 1): How to define a comprehensive Industry 4.0 assessment framework to measure the current state of organizational readiness?' with a potential industry 4.0 assessment mechanism.

The digital leadership dimensions will measure the incumbent executive team's awareness about Industry 4.0, tolerance level, and readiness to embrace industry 4.0 ecosystem and their strategic direction to allocate sufficient budget to address competitor's actions. ***The strategy*** dimension will address the expectations from current operations philosophies, tactical and strategic focus areas, any priority areas identified, emerging technologies utilization, and potential pilot study areas to adopt Industry 4.0 ecosystem. ***Business model*** dimensions will capture the potential revenue generation from the digital way of doing business, explore options to introduce digital features to existing product portfolio, and to leverage other supply chain stakeholders to achieve higher digital maturity. ***Product portfolio*** factor explains the product flexibility to leverage digital technologies, availability of necessary infrastructure in place to leverage data, agile approach in new product introduction, and companies' agility and flexibility to adopt digital mode of business. ***Digital operations*** perspective explores the organization's existing infrastructure readiness to use emerging technologies, good data generation mechanisms through a sufficient

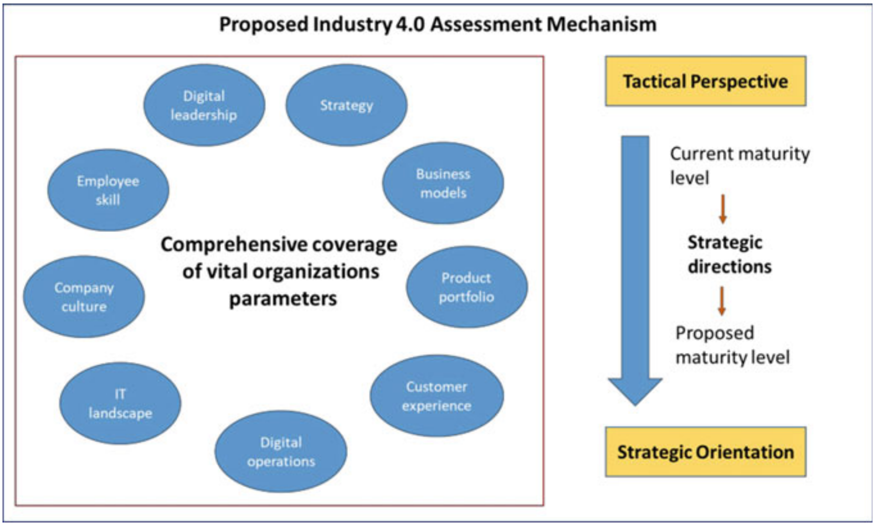


Fig. 1 Proposed industry 4.0 assessment mechanism

array of sensors, access critical parameters from anywhere, the degree of digital maturity from current operations and the extent of width and breadth of organizational functions covered by industry 4.0 system. ***IT Landscape*** is a crucial function to assess whether existing IT infrastructure is good enough to address Industry 4.0 requirements, sufficient IT flexibility to connect with all stakeholders from the supply chain, and enough security features to address ever-increasing cyber threats. ***Company culture*** is perhaps the deciding factor to determine the outcome from any major initiative like Industry 4.0 as the proposed system cannot last long without full approval from employees across the hierarchies. This assessment block will find out any dedicated training and mentoring program to prepare the employees for Industry 4.0 projects. ***Employee skill*** dimension will explore whether existing employees need additional training programs for Industry 4.0 challenge or organization can go outside to choose capable third party agencies to bring industry 4.0 skilled resources. The ***customer experience*** dimension will examine the existing digital mechanism's capability to address ever-increasing customers' desire to do business through digital means and potential digital solutions identified to ensure seamless communication with customers.

Organizations can be classified into five maturity stages as per the outcome from the proposed Industry 4.0 assessment mechanism (Refer Fig. 2). Here, we address the '*Research question 2 (RQ 2): How to determine Industry 4.0 maturity level for an organization based on assessment outcome?'* with the proposed Industry 4.0 maturity model.

Stage 1: Digitization: This is the minimum required phase to kick-start the digital transformation in any organization. The idea behind this stage is to put an IT system to replace manual activities into the digital mode. ***Stage 2: Digitalization:*** This is

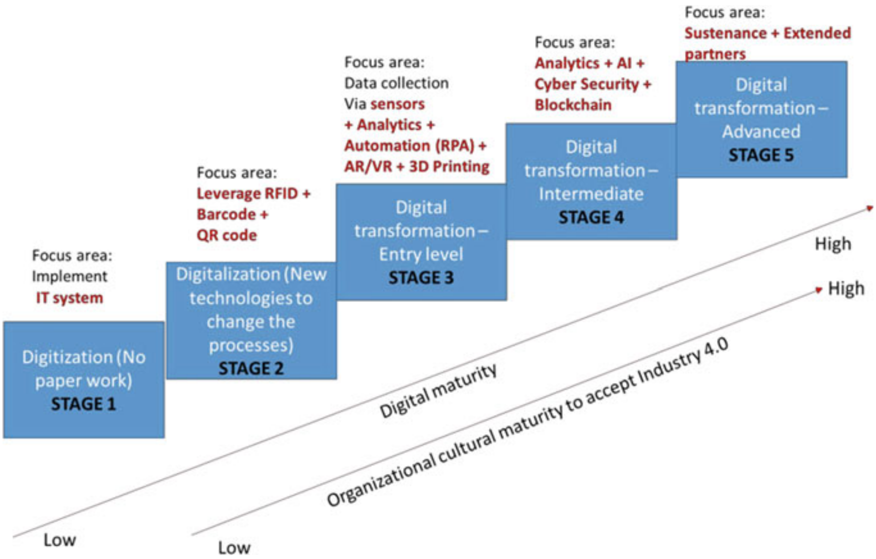


Fig. 2 Proposed industry 4.0 maturity model

the phase to implement higher-order technologies like RFID, bar code, and QR code to bring the next level of automation to enhance existing business processes and introduce new revenue generation platforms.

Stage 3: Digital Transformation—Entry Level: Data is the fundamental element for critical Industry 4.0 elements like big data platform, data analytics including machine learning and deep learning and artificial intelligence. The viable option is to generate such a massive amount of data is through installing sensors as per given requirements. Robotic Process Automation (RPA) technology and the usage of virtual reality and augmented reality technologies can be leveraged across various business functions. ***Stage 4: Digital Transformation—Intermediate Level:*** Industry 4.0 elements like data analytics and artificial intelligence can be used in this phase to leverage the extensive availability of data from various sources collected from the previous phase. The usage of advanced functions like Blockchain can be leveraged here to bring a high degree of transparency and singe truth across the organizational boundaries. The necessity of strong cybersecurity to protect organization IT and the proposed digital system is an important criterion here. ***Stage 5: Digital Transformation—Advanced Level:*** This is the phase where organizations have to start sustaining their Industry 4.0 investments and benefits. The role of solid cybersecurity is crucial here as the entire organization starts to work in a near smart manufacturing concept with an autonomous decision support system in place. Another driving factor is to integrate the organization's industry 4.0 ecosystem with other stakeholders like suppliers and customers to get true benefits.

5 Limitations and Future Research Areas

One of the limitations is this proposed industry 4.0 assessment model has to be implemented across many organizations to validate the expected outcome and to ensure the effectiveness. The chance to test this industry 4.0 assessment mechanism across countries will enhance the general acceptance of this model by the industry. Another limitation is this assessment model has to be filled by many people from an organization to avoid individual bias. This assessment model is giving the same treatment to all existing operational strategies like Lean and TQM, there may be some necessity to design a dedicated assessment mechanism for each dedicated operational systems in place. Potential future areas for research activities may include provisions for measuring nonoperational parameters like financial aspects to gain a comprehensive assessment for an organization.

6 Conclusions

Days are not far away to witness Industry 4.0 will alter the competitive positioning from the industry and a deciding factor to determine industry leaders. The non-availably of necessary data leads to the management team's biased approach toward Industry 4.0 and may hamper their decision-making abilities to frame necessary polices to proceed with Industry 4.0 ecosystem implementation. The proposed industry 4.0 assessment mechanism augments the management team's decision-making process backed by solid data points. The management team can chart out strategic plans with clear identification of challenges from the Industry 4.0 journey, which helps them to adopt reasonable expectations and clarity on the benefits from Industry 4.0 investment decisions. This proposed solution address the executive team's desire to acquire ground realities, provide necessary data points, and give confidence to them to embark Industry 4.0 journey.

Appendix: Digital Transformation (Industry 4.0) Assessments Instrument

Industry:
Nature of business:
MNC/Domestic Company:
Country:
City (Operations):
Turnover (In USD):
Number of employees:
Number of years in operation:

Type of company:

Sl. No.	Parameter	Questions
1	Digital leadership	Does the leadership team believe that the industry 4.0 paradigm will make a profound impact on the traditional way of doing business from your organization?
2		Are you ready to implement all I 4.0 digital components to become a leader in your industry?
3		The leadership team has taken some attempts to implement I 4.0 digital solutions across the organization?
4		Are you determined to start investing I 4.0 platform in this year?
5		There is a budget provision for I 4.0 other than typical IT budget for your organization
6		Are you witnessing competition from your peers for Industry 4.0 adoption? What is your immediate expectation from I 4.0? (Like Increase revenue, reduce cost, increase the quality, increase customer connect, create transparent in business operation)
7	Strategy	Do you have a dedicated designation like Chief Digital Officer in your organization?
8		Are you sure that you want to formulate various management strategies based on real-time data from the Cyber-Physical System environment from the organization?
9		Are you clear on how to kick start Industry 4.0 journey from your organization?
10		Which I 4.0 technology you are going to year adopt in this and next year
11		Do you need an experienced Industry 4.0 expert outside from your organization to implement I 4.0
12		What are the business areas prioritized for 4.0 investment?
13		You want to take a cautious approach to identify potential issues to choose the relevant I 4.0 digital component accordingly
14		Do you want to go for a pilot study to understand the potential benefits from I 4.0 before adopting company-wide implementation
15		Your focus areas from last 1 year for I 4.0 deployment
16		Your focus areas from next 1 year for I 4.0 deployment
17	Business models	Is your entire product life cycle embracing digital solutions (Starting from design to warranty?)

(continued)

(continued)

Sl. No.	Parameter	Questions
18		Are you determined to generate significant revenue from digital-based products and services in the coming years?
19		Do you foresee higher demands from your customers for digital features from your products and services?
20		Do you feel strong digital collaboration with your partners like suppliers and customers as a critical factor to decide your success in the future?
21	Product portfolio	We have the agility and flexibility capabilities in place to fulfill customer expectations by introducing innovative features in the product portfolio
22		We made provision to generate real-time data from our product portfolio, hence we can perform real-time product-related analysis
23		We leveraged digital solutions to introduce prototypes in a fast manner as per changing customer requirements
24		We are ready to explore new business models that run purely on digital solutions
25	Digital operations	What's your organization's maturity in terms of a digital-driven and intelligent automation system?
26		We have a unified digital platform to visualize companywide operations without any manual interventions
27		There is a mechanism available to generate the necessary data from all machines
28		Right now, we can able to access and control data from our regional/corporate offices far away from our factories
29		All critical operational parameters like KPI's are automatically updated through digital channels
30		Our internal functions are relying upon digital channels for coordination and collaboration
31		How do you rate the digital penetration for your vertical value chain?
32	IT landscape	How do you rate the digital penetration for your horizontal chain value chain?
33		Is your existing IT landscape is agile enough to accommodate dynamic needs from critical business parameters like reduce cost and enhance quality?
34		Your IT landscape is digitally connected with your partners like suppliers and customers?
35		Will you form a dedicated I 4.0 team or enhance the skill of existing IT team to execute I 4.0 projects

(continued)

(continued)

Sl. No.	Parameter	Questions
36		Are you sure that the existing IT team is matured enough to address I 4.0 requirements?
37		Are you ready to face cybersecurity risks if you go for a higher maturity level in I 4.0 operations in your organization?
38	Company culture	As a company, do you respect the hidden value of the raw data?
39		How do you rate your I 4.0 maturity in your organization?
40		Do you agree that I 4.0 journey is a significant change management program for your organization?
41		Are you sure that your employees are culturally matured enough to accept I 4.0 adoption?
42		Do you have any mechanism to encourage I 4.0 suggestions from your employees
43	Employee skills	Are you confident that your internal team can execute I 4.0 projects?
44		We already using digital solutions for employee competency building activities like e-learning and inline training classes
45		We already using the digital repository for knowledge management
46	Customer experience	Our organization adopted Omni digital strategy to reach out the customer
47		We determined to leverage emerging technology to engage with the customer
48		We have a digital mechanism to capture customer feedback from all digital channels, ex: social media
49		We give high importance to leverage digital solutions to address customer feedback to enhance our products and services
50	Others	Tell us your challenges to embrace I 4.0 in your organization

Note The quantitative questions are measured in 1–5 Likert scale

References

1. Stock T, Seliger G (2016) Opportunities of sustainable manufacturing in industry 4.0. Procedia CIRP 40:536–554
2. Berman SJ, Bell R (2011) Digital transformation: creating new business models where digital meets physical. IBM Inst Bus Value 1–17

3. Kagermann H, Wahlster W, Helbig J (2013) Recommendations for implementing the strategic initiative Industrie 4.0. In: Final report of the Industrie 4.0 Working Group. Acatech, München, pp 19–26
4. Aichholzer G, Gudowsky N, Saurwein F, Weber M (2015) Industry 4.0. Background paper on the pilot project "Industry 4.0. foresight & technology assessment on the social dimension of the next industrial revolution"
5. Mrugalska B, Wyrwicka MK (2017) Towards lean production in industry 4.0. Procedia Eng 182:466–473
6. Russwurm S (2013) Software: Die Zukunft der Industrie. In: Industrie 4.0 – Beherrschung der industriellen Komplexität mit SysLM, pp 21–36
7. Einsiedler Ingrid (2013) Embedded Systeme für Industrie 4.0. Prod Manage 18:26–28
8. Siepmann D, Graef N (2016) Industrie 4.0–Grundlagen und Gesamtzusammenhang. In: Einführung und Umsetzung von Industrie 4.0. Springer Gabler, Berlin, Heidelberg, pp 17–82
9. Kolla SSVK, Minoufekr M, Plapper P (2017) Deriving essential components of lean and industry 4.0 assessment model for manufacturing SMEs. In: 52nd CIRP conference on manufacturing systems (CMS), vol 81, pp 753–758
10. Schumacher A, Erol S, Sihn W (2016) A maturity model for assessing industry 4.0 readiness and maturity of manufacturing enterprises. Procedia CIRP 52(1):161–166
11. Leyh C, Bley K, Schäffer T, Forstenhäusler S (2016) September. SIMMI 4.0-a maturity model for classifying the enterprise-wide it and software landscape focusing on Industry 4.0. In: 2016 federated conference on computer science and information systems (fedcsis), pp 1297–1302
12. Anderl R, Picard A, Wang Y, Fleischer J, Dosch S, Klee B (2015) Guideline Industrie 4.0-guiding principles for the implementation of Industrie 4.0 in small and medium-sized businesses 4:1–31
13. Pokorni B, Schlund S, Findeisen S, Tomm A, Euper D, Mehl D, Brehm N, Ahmad D, Ohlhausen P, Palm D (2017) Produktionsassessment 4.0. ZWF Zeitschrift wirtschaftlichen Fabrikbetrieb 112(1–2):20–24

An Insight on Context-Aware Mobile Application Execution in Mobile Cloud IoT (MCIoT)

L. Shakkeera, Y. Sharmasth Vali, and Santosh Karthikeyan Viswanathan

Abstract The present-day significance of mobile cloud and Internet of things (IoT) in the evolving digital world, MCIoT framework affords the flexibility service for IoT-based mobile applications and also bids highly scalable service to the storage and computing of large streaming data. Owing to the several mobile devices, high mobility, change in network overlay and bandwidth, conserving the seamless connection is the challenging tasks in MCIoT environment, resulting in higher handover delay. To overcome the above constraints, this paper discusses the perceptive with context-aware mobile application execution that effectually provides the service guarantee, improving the user's experience, enhancing processing capability by reducing interaction delay. In future, MCIoT possibly can enable the mobile users to openly access the applications in the mobile cloud through SaaS provisioning rather than installing the applications on the mobile devices. Finally, MCIoT dominates the mobile industry due to its significant factor and context-aware seamless mobile application execution.

Keywords Context-aware execution · Mobile cloud computing · Mobile networking · Service accessibility · MCIoT

1 Introduction

The client–server models were used as traditional method which is very successful for using the mobile devices and for also holding the resources in the cloud. Onloading and offloading work during server runtime is more beneficial. Offloading may also bring security threats. Many Web or cloud apps can slower the execution process in mobile devices. So online cloud services are introduced to process and

L. Shakkeera (✉) · Y. Sharmasth Vali
School of Computing Science and Engineering, VIT Bhopal University, Sehore, Madhya Pradesh, India

S. K. Viswanathan
AstraZeneca, Chennai, Tamil Nadu, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_6

access seamless application execution. For cloud providers and network operators, it provides huge career opportunities. It also has a rich computational technology that supports unique versatile resource of different clouds and network technology. It also serves a host for devices anytime, anywhere by using the medium called Internet in diverse environment and all platforms. State-of-the-art application for mobile phones and cloud are major services in today's market, and also we suspect that possible services are yet to be explored.

MCC frameworks [1, 2] are mainly modeled for energy-efficiency mobile applications and mobile-based IoT application execution in smartphones. The framework was developed based on task scheduling, resource allocation, optimization, and adaptive workflow management approaches with the limitations of context-aware mobile application execution based on mobile networking concepts. To overcome the above-mentioned constraints, this paper gives insights on context-aware mobile application execution in MCIoT environment by reducing congestion rate, and eliminating handover delay during service accessibility from remote cloud server. For IoT-based applications, the application structuring depends on dynamic analysis of request characteristics, diverse nature of workflow, resource availability, device characteristics, and task dependency at a runtime environment.

1.1 Mobile Computing

Mobile computing which permits for transmission of data, voice, and video. Mobile computing mainly comprises of three strategies: mobile hardware; mobile software; and mobile communication. The mobile communication includes protocols, data formats, network properties, and communication technologies. Mobile hardware embraces mobile devices (smartphones) or device components. Mobile software compacts requirements and characteristics of mobile applications. The main principles of mobile computing are: portability; connectivity; interactivity; and individuality. The furthermost communal forms of mobile computing devices are: smartphones; wearable computer; tablets, etc. The important limitations such as transmission interferences, power consumption; energy efficiency; human-device interface; communication range and bandwidth; expandability; and security standards. Mobile computing uses mobile/wireless data connections, it uses technologies like GSM, GPRS, 3G or 4G networks.

1.2 Mobile Cloud Computing

Mobile cloud computing (MCC) [3] is a set of cloud computing and mobile computing technology which is combined for rich computational resources for a mobile user as well as network operators and to cloud service providers. The major goal of a mobile cloud is enabling an execution process of applications in mobile

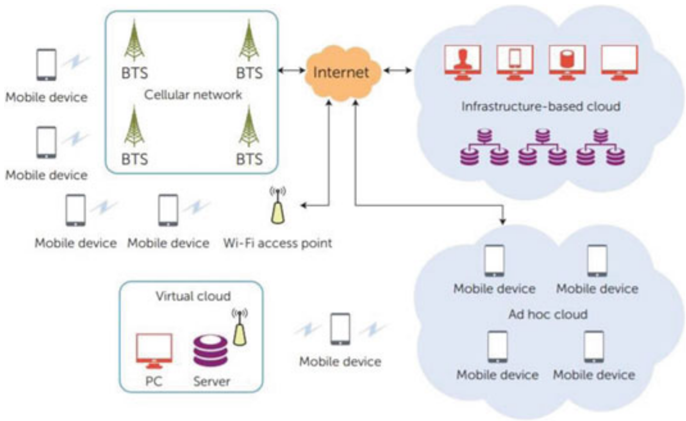


Fig. 1 Mobile cloud architecture with infrastructure-based cloud, virtual cloud, and ad hoc cloud [5]

devices and remote cloud server, in which user can also have a rich experience. The vogue of mobile device and mobile networks has significantly changed the way of people to access network services and computers. The main merits of MCC such as battery lifetime can be extended [4], improvement in processing power and data storage capacity, dynamic provisioning, improving reliability and scalability, multi-tenancy, and ease of integration. The advantages of MCC are directly linked with the present mobile computing and cloud computing and accessing mobile applications. The applications are categorized as image processing, multimedia such as audio and video streaming, sensor-based data application in IoT entities, crowd computing, online gaming, and text processing. Figure 1 [5] shows mobile cloud architecture with infrastructure-based cloud, virtual cloud, and ad hoc cloud-based on which mobile devices computation offloading.

Technological Challenges of MCC. MCC has several advantages for mobile users and cloud service providers, but by combining various fields, which comprises of cloud computing, mobile computing, mobile networks, and mobile communication, MCC necessities to address several technological challenges. The challenges include resource scarcity in mobile devices (smartphones), network bandwidth, latency, network availability, computation offloading, heterogeneity, data access efficiency, ownership on data, elastic application models, and finally security and privacy for mobile applications. These issues focused in many of the MCC research studies that directly connected to cloud computing and mobile communication.

1.3 MCIoT

The mobile cloud IoT is rising in many Internet applications [6]. IoT equipment from vehicles to smartphones which can be used as day to day life. The generation of used amount of data selection and distribution of IoT devices is necessary to achieve accurate results and efficient way of working the application system is therefore important. Data collections are located from outside the devices using cloud computing with mobile cloud IoT which reduces the pressure on the mobile devices in terms of energy efficiency. Wireless cloud connection is also needed for mobile application. The combination of mobile computing includes telephones, devices, and how they wirelessly connect with other devices. The MCIoT needs to support the services while processing several dynamic mobile applications such as video streaming, financial transaction, health care, and online gaming rely on the instant response to the lively variation.

1.4 Contributions

Context-aware remote execution is the progression of examining the mobile application based on the device constraints to achieve dynamic and seamless execution of an application on a mobile device and a remote server even in the increase of congestion rate in mobile networks. The main overview and contributions of this paper discusses context-aware mobile application execution in MCIoT, the process of mobile application execution in cloud server, service accessibility using mobile devices in MCC communication, the process of context-aware healthcare service system, and IoT-based remote application execution.

2 Related Research Works

The related research works in this section describe the concepts and methodologies of context-aware mobile application execution, context-aware energy-efficient communication, context-aware strategy for IoT-cloud services, energy-efficient approach for application execution in mobile cloud IoT environment, and context-aware mobile application recommendation in remote cloud server.

Abusair et al. [7] “Context-aware adaptation of mobile applications driven by software quality and user satisfaction,” this paper discusses context-aware approach. Which is mainly for the mobile applications that totally work on the quality of software as well as satisfaction of mobile users. Particularly, they focus on services that should be available at any time and to acquire a term that selects the optimal application adaptation that is to be implemented in a given context by incorporating a devised methodology which has the concept of product availability and user feedback.

Masango et al. [8], "Context-aware mobile application for mobile devices," this paper discusses android smart devices that offers smart lock capability which utilize various methods for authentication to unlock the device by user's location. Though android smart lock does not allow individual applications to lock. So they have come up with an application uses a context-aware approach as solution. This paper proposes a framework that offers multiple user authentication mechanisms which are set by the user to auto-detect areas identified as safe zones. This application focuses on improving the security level of a given device content by safeguarding the individual apps.

Aksanli et al. [8], "Context-aware and user-centric residential energy management," this paper describes hierarchical model for the concept of energy management in residents which is IoT-based, which involves a general functional framework to guide processing of data and reasoning. They extended this hierarchy to describe the power supply network to all individual households. This system collects additional data generated as user context from multiple devices and uses to determine flexibility to handle the energy by the user. Further, the paper demonstrates that consumers experience simulation provides 14% increase in forecast accuracy for electricity efficiency and 12% decrease in average grid energy costs.

Sen et al. [9], "Invited: Context-aware energy-efficient communication for IoT sensor nodes," the paper describes that wireless communication is a major energy consumption in this data-based IoT environment, so paying particular attention to differences of local area and remote area connectivity is crucial to total usage of energy. In changing contexts such as channel conditions and data-rate requirements, the communication connectors which will control this huge of IoT workload needed to be an energy efficient. In addition, IoT devices include several multiple parallel communication materials like proximity, wired, mm-wave, 5G, etc. The author discusses how self-learning facilitates context-aware activity in these communication systems that allow minimal energy data for every communication. Need for context-aware activity inside and within multiple physical layers would be stressed in upcoming IoT research challenges. Hence, these energy-efficient networking along with the low-power computing will harness the promise of the IoT system revolution and generate major impacts.

Sneps-Sneppe et al. [10], "On context-aware proxy in mobile cloud computing for emergency services," in this paper for emergency services, they define a new paradigm with 5G networks for the context-aware computation. They introduced a different mobile services model that effectively enhances 5G network power. The ultimate strength of 5G networks is always to allow applications that previously were not possible to use. They build a duplicate copy for mobile data sensing in their model and enable server-side processing to make efficient use of context information during processing. They are proposing a context-aware proxy centered on the proximity-related cloud-targeted applications and services.

Liang et al. [11], "A broad learning approach for context-aware mobile application recommendation," this paper discusses with tensor analysis. They suggest large learning approach for context-aware application recommendations. In particular, they use a tensor-based framework to incorporate new app category information and a multi-view features for users and apps to facilitate the performance of prediction ratings. The multilayered framework is used to capture the relationships between app categories and multi-view features that are hidden. They also create an effective factorization method that uses Tucker decomposition to realize the entire-order relationships between device's groups and functions.

Kumari et al. [12], "Energy efficient approach for application execution in mobile cloud IoT environment," states and explains that Internet of things (IoT) provides dynamic energy resource control and monitoring facilitating stakeholders to save more energy use at granular levels. A new technology called mobile cloud IoT (MCIoT) is being developed for the advancement of IoT technologies and the mobile cloud, which is the integration of both. The main objective of MCIoT is to provide appropriate resources for the application to allow it to execute the mobile application without any deficiencies. The SMDs battery life significantly increases when a device is partitioned complex component is on the cloud and the remainder is on the SMD. In this paper, they provided an empirical model in the MCIoT context to show the correct partitioning rate that affects the time and energy usage of SMD and cloud execution. Their experiments are conducted based on the different partition rate, processing power and also data size of an application. Furthermore, the influence of these parameters is evaluated on the residual energy of SMD.

3 Mobile Application Execution in Cloud Server

Online cloud services become increasingly popular through these years, and the user has to believe the cloud service provider with their personal data. The user's private information is kept as highly confidential. In android application, streaming of data is good and it is very much beneficial for the user to store the information. It is impossible for the user to monitor the data usage in users mobile devices. In recent years, Samsung Galaxy has been partner with Drop-box, while Microsoft similarly offers Microsoft One-drive. i-Cloud is cloud storage portal which is pre-loaded and designed for Apple iOS device. Most of the companies offer cloud storage with highly secured Web sites for access files and any device can also browse to the Internet. By using these cloud services, backing up process can be done easily in case of emergency. Figure 2 shows mobile application execution in cloud server environment. Based on user preferences and application requirements, the mobile applications are executed in mobile devices. Resource-intensive applications are offloaded into remote cloud server and cloud execution manager allocates computational and storage resources in the form virtual and physical resources. The mobile devices and cloud server process the data to and from between mobile and servers.

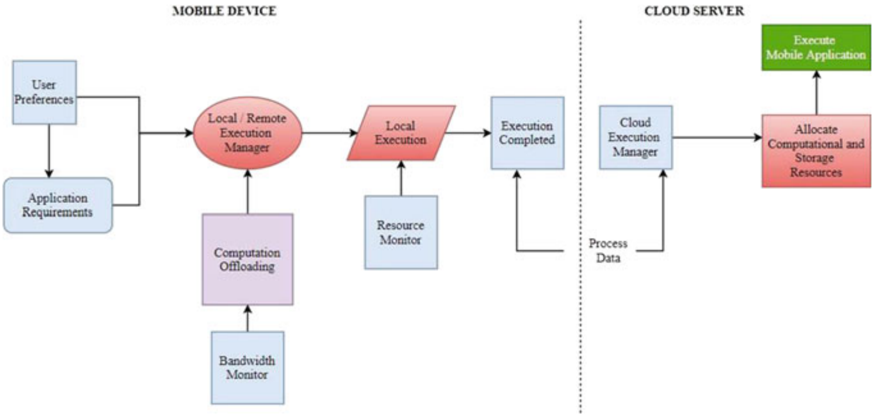


Fig. 2 Mobile application execution in cloud server

4 Mobile Networking in MCC

4.1 Mobile Networking in MCC Communication

For MCC and telecommunications services, wireless and mobile networks quickly become the most important type of network access. There have now been two different paths for the progress of the wireless network and the future generation of mobile networks in MCC. All of them are following the direction of mobile cellular networks with their worldwide long-standing infrastructure. Mobile communication is also one of the most sophisticated and challenging technologies for transmission. Wireless technologies are on the verge of being developed in their fourth generation. In the past, wireless access systems have taken different evolutionary paths for the same goals such as high mobile performance and capacity.

In a mobile technology generation, Universal Mobile Phone Service (UMPS) is a third system that offers code division multiple access wideband application. Universal Mobile Telecommunications System (UMTS) has greater flexibility, enhanced storage capacity, and much broader range of services with storage rates of up to 2 Mbps. Code division multiple access (CDMA) infrastructures is capable of delivering images, video messaging to location services, and also offers superior performance and distribution capability.

Mobile networks in the fourth generation are the advanced edition of Internet services in the third generation. The last generation of broadband services has high power, fast transmission data that enables uses to connect interactive digital services in high quality which includes telecommunication, 3D animation games, audio content, and facilities. Furthermore, 4G networks should offer portability of services and lower-cost scalability. It is based on the multiplexing of orthogonal frequency division with thousands of simultaneous channels available. The data transfer rates are scheduled in mobile mode in the range of 20–100 Mbps. This framework enables the fast and efficient convergence of various mobile networks that allows telephone, radio, and satellite communications to switch from indoor networks such as Wi-Fi and Bluetooth. Wireless networks always have changed topologies, mainly made up of wireless links with restricted bandwidth. It has wireless sender and receivers, with a network like a router.

4.2 Service Accessibility Using Mobile Devices

New levels of utilizability, performance, and computation power have been attained by mobile devices such as smartphones and personal digital assistant. More and more wireless communication and location technology are being added to these devices in a wireless phone like Subscriber Identity Module card. This module is also available for smartphones. After the mobile card is entered, network begins checking the validity of SIM card by verification process connecting to the center for authentication. Mobile phones and their size are portable the other performance metrics of mobile phones and new technologies make these phones smaller by weight and battery life. GSM systems are comparable to analog systems, and interference is weak singles that are related under unfavorable circumstances. GSM systems produce good quality under conditions that range from moderate to good speech quality in GSM systems.

There are several other technologies that can be grouped into two sections of radio transmission, one width frequency challenges and other with technology of transmission. For sending the information, radio devices do not work with radio equipment and cables. Intensity is the limitless resource that can be utilized efficiently. Three key issues need to be addressed in system design is technical effect on routers and sensors which includes propagation features such as ability to reflect and decrease the system. Any device connected to and running network of mobile devices, including pages and cellphone if commonly classified as a mobile network. However, there is a large number of valuable electronic information for mobile devices similar elements will be maintained by a network service provider using report and billing records, and the researches can capture mobile network traffic. It is also effective for people to analyze the fundamental concepts of network, types of data it can collected, and analyzed mobile network. Figure 3 depicts Service accessibility using mobile devices through mobile networking.

4.3 Issues in Mobile Networking Communication

In mobile networking, many of the techniques are identified in order to manage and obtain the data from mobile cloud environment. For example, obtaining geo-location data from mobile networks. It includes the components and evaluation of details and utility of textual and visual messages. Cellular services are also built

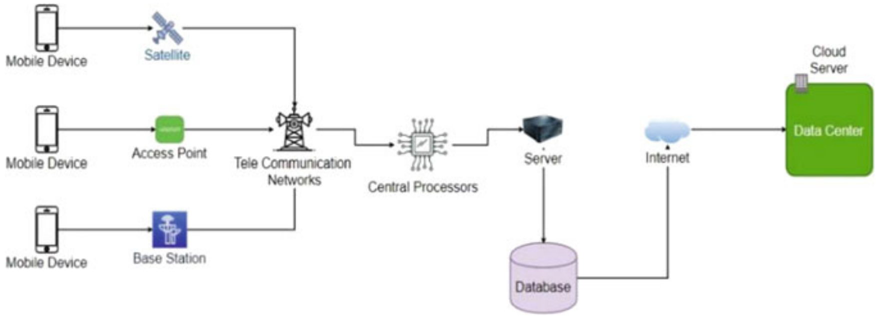


Fig. 3 Service accessibility using mobile devices through mobile networking

in network node environment and frequently implemented in areas with unsecured nodes, which provide less protection against attack. Many other strategies to safe guard client data privacy and their data confidentiality on cellular networks have been suggested. The dynamic movement of the network nodes is an another difficult issue in mobile networks. Different classification in membership location must be expected to deal with data integrity when client enters sub-net and enters the new sub-net.

Network congestion in mobile data networking [13] is an important issue that will reduce quality of service when network nodes carry more data that it can manage to handle. This effects network communication and leads to packet loss and canceling connections by using same mobile networks. Consequence of network congestion leads to increase the network load and decreases the throughput. Several network protocols are suggested to compensate packet loss and use destructive retransmission of packets. Federated architecture of management therefore suggested to encourage communication between mobile network entities.

5 Application Context-Aware Remote Execution

5.1 Context-Aware Application Execution

Context-aware derived from concept of pervasive computing which uses heterogeneous source to adapt and provide a service based on user needs. It offers new possibility in decreasing the necessity of user interaction instead of requiring the operator to choose the needed functionalities. The context-aware mobile application execution [14] provides the desired service automatically. Context-aware execution based on different types of parameter like location, identity, activity, and time. Based on this, there are three models are there for context- aware application execution: activity determination, process definition, and context elicitation.

Activity Determination: Context-aware execution is to identify the activities. Therefore, it may help to decide whether the contextual data could assist the users. They are represented using modeling language like UML. *Process Definition:* Based on the result of activity determination, the actions are categorized into two parts such as active execution and passive execution. This execution triggers the user a new process flow which changes the functionality. It also requires to define the changes in working procedures. The process definition ends in validating of preliminary results.

Context Elicitation: After defining context, it validates to ensure the following properties: uniqueness (each activity must be unique) and measurability (all the contextual data must be measurable). If any of this condition is violated, the execution will not operate in the intendant way. So it required new elements such as sensor, additional conceptual data ensuring the uniqueness and measurability of all contextual data.

Figure 4 depicts context-aware application execution with three difference model process. Following the activity determination, response and corresponding contextual are consider concurrently. If necessary, process adjustment is performed to change the requirements. After the adjustment, it is necessary for each newly created process flow to repeat a response and context definitions steps. Once the process definition is performed, then validate the preliminary results. During context elicitation, define and validate contexts. If system architecture adjustment is necessary based on contexts, then incorporate system architecture requirements otherwise end the activity which leads to execution model.

Context-aware computing system [15] provides few key issues like high dimensionality different format, missing data, repetitive data, and high dispersion due to large volume of data. This kind of issues leads to data inconsistency (detect to avoid error context analysis) and difficult to validate. Context-activity matrix is enhanced by integrating the different abstraction level and easing the validation of properties like uniqueness and measurability while focusing on identification of critical situations. In context-activity matrix, labeling the rows with type of activity, column is labeling by sensed contextual data, specific data are noted in remaining field, and finally, top left field is used to identify the state that allow differentiation between different matrices. This construction is efficient for two reasons, first reasons is to identify the activities with limited set of context type. It helps us to allow the usage of same header for different matrices. Secondly, the uniqueness of contextual data valid by comparing with other rows and measurability is secured by ensuring functionalities and sensors to measure the contextual data. Context types are used for separation. Introducing different level of abstraction for extensive of number of activity and context data help to improve the clarity. Encourage to follow use case breakdown while there is no restriction on splitting matrix so it provides clear line for separation and their subdivision is aligned with other developed approaches. For example, various locations are mapped, each matrix contains the activity of one location. It also helpful for specified applications using separate matrix with unique set context like monitoring the characteristics of machine with various internal sensors result in set of specialized in contextual data. Therefore, clear separation is also possible for another matrix.

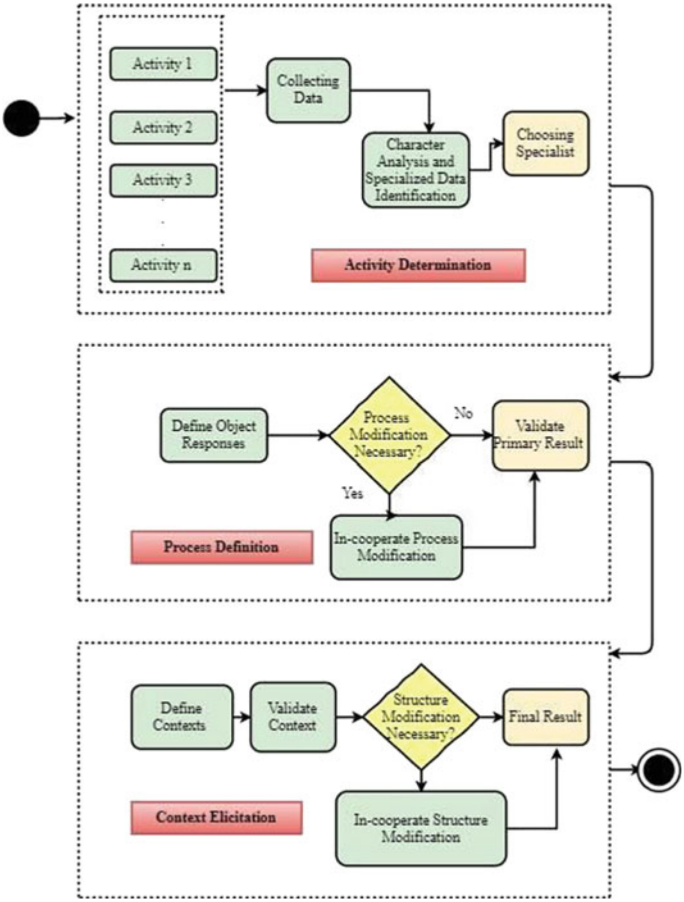


Fig. 4 Context-aware application execution

5.2 Efficient Remote Execution-Based Approaches

In current scenario, the usage of smartphones by the mobile end users become essential with emerging mobile cloud computing concepts with offloading the code from resource constrained mobile devices to resource-rich cloud servers. The remote execution of mobile cloud applications aims at enhancing the seamless application execution [16], application performance, and achieves energy efficiency in mobile devices. There were several approaches developed for efficient remote execution (computation offloading) of applications from mobile devices to cloud. The part of the application code needs to be offloaded from smartphones to remote cloud in the form of partial offloading or full offloading. In full offloading, the full application code along with all the necessary data need to be offloaded to the cloud where the entire computation process should be taken place, and finally, the computation results are sent back to the mobile devices. In the partial offloading, only part of the application code that ingests more energy and leads to computation complexity should be offloaded to the cloud. Some of the tasks are executed locally which eliminates runtime complexity between mobile devices and cloud. The efficient remote execution leads to offloading benefits such as improvement in lifespan of battery of mobile devices, reduced computation complexity, and the approaches includes static and dynamic-based code partitioning, dynamic resource allocations, execution optimization techniques, workflow-aware execution [17], energy-efficient execution, workflow-based computing techniques for IoT in mobile cloud includes application structuring, application partitioning, and computation offloading. Mobile cloud IoT systems implements remote execution for influencing the data storage and computation requirements of the IoT applications such as mobile gaming, smart health care, cloud-assistive vehicular transportation, smart network, and smart city [18].

5.3 Context-Aware Application Development Models

Mobile cloud computing technologies are evolved, however, this requires specific context-aware application development models that can expedite the development of cloud and IoT-based applications adept of taking context-aware computation offloading decisions. Context-aware applications are trending applications that are able to sense the entire environment and manage the situational context to facilitate better end users experience. To enable effective computation offloading in MCC, context-aware application development models are designed to enhance performance of the system, energy-efficiency, and local and remote execution support on mobile devices and cloud server. Model-driven context-aware application models are developed for Web applications on a Web service context management architecture [19]. In context-aware application models, they are two aspects are there such as context-aware computation offloading and context-aware application partitioning. The context-aware computation offloading is developed for system performance enhancement, energy-efficiency, and effective way of application execution. The main entities involved context-aware computation offloading includes smartphones, cloud resources (both physical and virtual), communication technologies such as Wi-Fi or cellular, data size and location, network bandwidth, and application model for computation offloading technique. The context-aware application partitioning to enable selective offloading between smartphones and remote cloud. Because entire application execution requires smartphones hardware support, high amount of communication bandwidth and it increases handover delay and energy consumption in mobile devices. The partitioning the applications can be static or dynamic based on development and runtime environment conditions.

The current mobile cloud application development models are: CloneCloud, μ Cloud, eXCloud, MAUI, and ThinkAir. Paper [20] discusses an application development model (Mobibyte) for mobile cloud computing. These application models offload resource-intensive applications from mobile device to cloud server based on different procedures. The model uses computation offloading, method-level offloading, on-demand data migration, context-aware offloading decisions, energy-level modeling decisions, application partitioning, and application structuring depending on the available mobile and cloud resources and user preferences.

5.4 Context-Aware Mobile Application Execution

The context-aware (contextual) information relevant between user and an application. These information parameters will change during dynamic execution. Based on set of environment states and settings, context- awareness defined and determines application's behavior. It is an ability of a system component to collect the information from its surroundings at any given time and embrace user's behaviors accordingly. The context-aware computing utilizes both hardware and software to routinely collect and analyze the data and to monitor the responses. The context in mobile devices includes various fields (mobile computing, wearable computing, and ubiquitous computing). Context-aware computation offloading in MCC [21] faces many challenges such as application partitioning, computational time, computational energy, offloading time, offloading energy, and application support. These challenges obstruct mobile cloud application models from effective context-aware offloading decisions.

Context-aware mobile application execution [22] of mobile devices defines location awareness where location determines how the firm the processes contribute the device to operate. The term context-awareness derived from the form ubiquitous computing or also called as pervasive computing. Context-aware mobile application execution linked with contextual application design and business process management. For example, Fig. 5 depicts context-aware mobile application execution for healthcare service system. Using mobile health monitoring apps access the context-aware healthcare service system [23] through communication interface. The system is connected with database server to access the Web service cache, user requirement database, geographic information database and connection and control module.

6 Mobile Cloud Computing Over Dynamic IoT Environments

In recent times, both academic and industry have shown serious attention to integration between Internet of things with IoT sensor cloud. The coordination is driven by the benefits of effective cloud computing processing and storage for data sensing. With this integration, cloud sensing allows the cloud application to provide sensing of data. Moreover, restricted nodes in the sensor pass the energy saving computer activities to the cloud. The system for enabling the IoT-cloud in the situation of different tasks to reduce the impact of queries shared to sensing devices to maximize

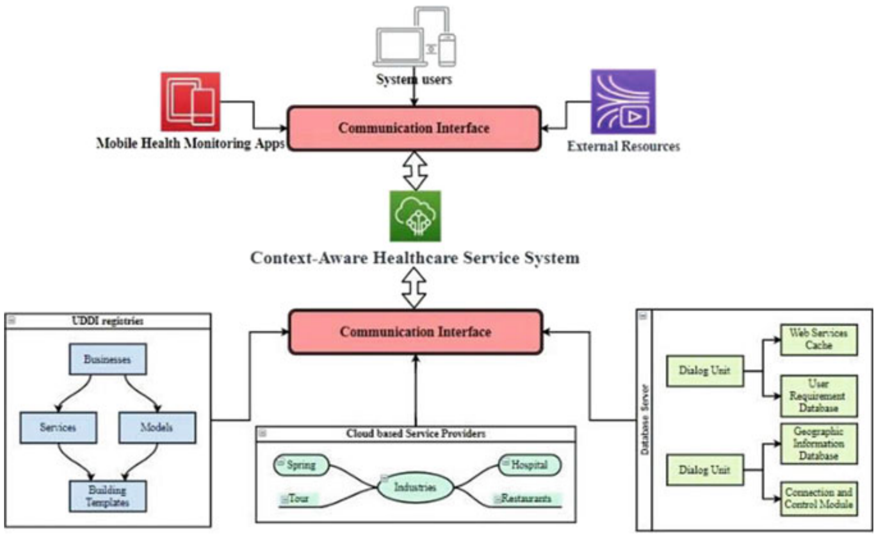


Fig. 5 Context-aware healthcare service system

the energy consumption of sensor. Their devices sensing also fulfills application requirements. For reducing the transmission of data, processes such as data catching and request aggregation will be helpful and data is transmitted regularly by the use of sensors. For periodically reporting their sensing data, their devices are needed. The report of predictive analysis which does not understand the sensing demands, results in extreme loss of energy due to duplicate transfer.

The position of mobile users is generally recognizable by cloud in much mobile cloud application such as vehicle safety and computer hacker systems. The cloud will know where the users require sensing services. The cloud therefore has knowledge about the location needed and the location where sensing services are required. The location-based (context-aware) collaborative model between IoT and cloud [24, 25] can be introduced for a successful area of application and therefore it will be modeled. The model allowing on-demand interaction of the cloud with the IoT enabled cloud to save power by controlling and scheduling the sensors as a coordinator for resource-restricted sensors. The cloud-controlled sensor services offer three main advantages. The data are collected on-demand based upon request interest, thus, reducing data redundancy and the quality of service of the sensing system can be set on the basis of application users requirements. IoT sensor cloud has recently been put forward as a hopeful and research-related approach. The IoT-cloud virtualizes the physical sensors into digital cloud sensors. Sensing and transmission of their sensing data to the cloud are responsible for physical sensor network.

6.1 IoT-Based Remote Application Execution

In recent years, variety of applications have been programmed in connection with IoT devices which plays a huge role in market today and those applications can be installed in any type of android or iOS devices according their norms and conditions. These IoT devices created a huge business platform. Some type of IoT devices are programmed to be accessed from anywhere and anytime and it also connected to cloud storage for quick backing up process. The users of IoT devices are also demanding such type of facilities from the developers. IoT devices also include health care, fitness, home appliances, etc.

For example, nowadays, IoT started contributing to healthcare industry. It allows medical devices to stay connected with cloud and to its applications. It plays a pivotal role in devices which collects patient records, data, staff records, inventory, facilities, smart beds, etc. It directly benefits to patients and healthcare providers for quick accessing and detecting failures before they become serious and can provide personalized therapy soon. Managing hospital equipment: It is also easy for physicians or healthcare professionals to check via mobile devices as it is connected to cloud storage. Stock management: apart from all the records, their supplies stock can also be maintained in real time with use of RFID tags and mobile scanners. By making these devices connected, the manufacturing cost can be easily reduced. Machine learning techniques are also combined the data for identifying new patterns which can lead to new product or services.

Currently, IoT application is created and its execution is done in cloud platform. For example, Amazon has created an IoT device with an application for both android users can download it in "Google Play Store" and for iOS "Apple Play Store" which is connected to cloud database for storage which is now major moving product on online market. The specialty of that product is it can connected through Bluetooth and also Wi-Fi technology can track all type of actions and movements, using that product, we can access find your phone, calls, messages, emails, and other specified applications. And all type of fitness applications which is connected to IoT can be tracked and backup for good results.

6.2 IoT-Based Applications in MCC

Mobile cloud computing platform is the important step to implement precise planting to collect information on spatially time variance that effects plant crop growth and development in a timely manner. The accurate environmental monitoring and early warning analysis is used to intelligently surveillance the situation of large- scale planting. Firstly, continuous collection of data and information such as carbon-dioxide, concentration, soil humidity, and NPK soil content. These data are indexed so that mobile query application can easily access the information in the second location, and we begin to store and upload and transfer all data sets to the cloud computing center. The mobile cloud IoT database manages the data relating to quality characteristics which is used to analyze and calculate production plant. The development of mobile cloud IoT can also be recognized to use pattern recognition technology and carry out complex plant-growing monitoring using other measuring instruments, and then, mobile front end service provides online investigation to users. Agricultural workers understand the specific real-time planting state using mobile devices. This mobile network provides analysts with the clear view of plant production and also reduces the problems and which causes natural disaster, viruses, yield of crops, and impact of crop production.

7 Conclusion

Context-aware mobile application execution incorporates context information, application requirements, and user preferences in cloud server database information systems. The adaptive, responsive, reactive, and environmental- directed applications have special emphasis on the location context. Context in mobile devices based on user contexts that comprises mainly personal attributes of the mobile users, point-of-interest, and environmental context. So, based on the incorporation of context information, the mobile applications are provided to the mobile users may expressively enhance the quality of service accessibility in terms of finding more related results. In present days, the context-awareness [26] in mobile cloud computing take advantages in various mobile contexts, namely, microblogging (e.g., Instagram, Twitter), social news feed, and mainly recommendation system services. This paper discusses a perception on context-aware mobile application execution in MCIoT environment by considering the issues of mobile networking in MCC communication, process of service accessibility using mobile devices, efficient remote execution-based approaches, different context-aware application development models, and MCC over dynamic IoT environments. Finally, by achieving context-aware mobile application execution in MCIoT, it reduces network delay, conserving energy-efficiency in mobile devices, and cloud server with better user experience while accessing seamless service accessibility.

References

1. Shuja J, Gani A, ur Rehman MH, Ahmed E, Madani SA, Khan MK, Ko K (2016) Towards native code offloading based MCC frameworks for multimedia applications. A survey. J Netw Comput Appl 75:335–354
2. Panneerselvam J, Hardy J, Liu L, Yuan B, Antonopoulos N (2016) Mobilouds: an energy efficient MCC collaborative framework with extended mobile participation for next generation networks. IEEE Access 4:9129–9144
3. Dinh HT, Lee C, Niyato D, Wang P (2013) A survey of mobile cloud computing: architecture, applications, and approaches. Wirel Commun Mobile Comput 13:1587–1611

4. Shakkeera L, Tamilselvan L (2016) QoS and load balancing aware task scheduling framework for mobile cloud computing environment. Int J Wirel Mobile Comput 10(4):309–316
5. Khan AUR, Othman M, Xia F, Khan AN (2015) Context-aware mobile cloud computing and its challenges. IEEE Cloud Comput 2:42–49
6. Kim S (2015) Nested game-based computation offloading scheme for mobile cloud IoT systems. EURASIP J Wirel Commun Netw 229:2–11
7. Abusair M, Di Marco A, Inverardi P (2017) Context-aware adaptation of mobile applications driven by software quality and user satisfaction. In: IEEE International conference on software quality, reliability and security companion (QRS-C) proceedings. Prague, pp 31–38
8. Masango M, Mouton F, Nottingham A, Mtsweni J (2016) Context aware mobile application for mobile devices. Information Security for South Africa (ISSA), Johannesburg, pp 85–90
9. Aksanli B, Venkatesh J, Chan C, Akyurek AS, Rosing TS (2013) Context-aware and user-centric residential energy management. In: IEEE International conference on pervasive computing and communications workshops (PerCom Workshops). Hona, HI, pp 455–460
10. Sen S (2016) Invited: context-aware energy-efficient communication for IoT sensor nodes. In: 53rd ACM/EDAC/IEEE design automation conference (DAC). Austin, TX, pp 1–6
11. Sneps-Sneppe M, Namiot D (2017) On context-aware proxy in mobile cloud computing for emergency services. In: 24th International conference on telecommunications (ICT). Limassol, pp 1–5
12. Liang T, He L, Lu C, Chen L, Yu PS, Wu J (2017) A broad learning approach for context-aware mobile application recommendation. In: IEEE International conference on data mining (ICDM). New Orleans, LA, pp 955–960
13. Kumari R, Kaushal S (2017) Energy efficient approach for application execution in mobile cloud IoT environment. In: 2nd international conference on internet of things and cloud computing (ICC 2017). Cambridge, UK
14. Mangiante S, Schapira M, Navon A, Dias Silva M, Godfrey B, Wang W, Smith K, Pechtalt I (2018) Congestion control for future mobile networks. In: 13th Workshop on challenged networks, CHANTS'18. New Delhi, pp 55–61
15. Gómez-Torres ER, Challiol C, Gordillo SE (2018) Context-aware mobile applications: taxonomy of factors for building approaches. In: IEEE International conference on electronics, electrical engineering and computing (INTERCON) proceedings, pp 1–4
16. Gandodhar PS, Chaware SM (2018) Context aware computing systems: a survey. In: 2nd International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) proceedings, pp 605–608
17. Shakkeera L, Tamilselvan L (2017) Satisfying SLA objectives of seamless execution of mobile applications in cloud with net profit. J Theor Appl Inf Technol 95(11):2577–2588
18. Shakkeera L, Tamilselvan L (2017) Towards maximum resource utilization and optimal task execution for gaming IoT workflow in mobile cloud. Int J Intell Eng Syst 10(1)
19. Abu-Issa A, Nawawreh H, Shreteh L, Salman Y, Hassouneh Y, Tumar I, Hussein M (2017) A smart city mobile application for multitype, proactive, and context-aware recommender system. In: International conference on engineering and technology (ICET). Antalya, pp 1–5
20. Kapitsaki GM, Venieris LS (2009) Model-driven development of context-aware web applications based on a web service context management architecture. In: International conference on model driven engineering languages and systems, models in software engineering. Lecture Notes in Computer Science Book Series (LNCS), vol 5421, pp 343–355
21. ur Rehman Khan A, Othman M, Khan AN, Abid SA, Madani SA (2015) MobiByte: An application development model for mobile cloud computing. J Grid Comput 13:605–628
22. Orsini G, Bade D, Lamersdorf W (2015) Context-aware computation offloading for mobile cloud computing; requirements analysis, survey and design guideline. In:12th International conference on mobile systems and pervasive computing (MobiSPC) proceedings. Procedia Computer Science, vol 56. Elsevier, pp 10–17
23. Berrocal J, Garcia-Alonso J, Murillo JM, Mendes D, Fonseca C, Lopes M (2018) Context-aware mobile app for the multidimensional assessment of the elderly. In: 13th Iberian conference on information systems and technologies (CISTI). Caceres, pp. 1–6

24. Wang X, Jin Z (2019) An overview of mobile cloud computing for pervasive healthcare. IEEE Access 7:66774–76679
25. Sezer OB, Dogdu E, Ozbayoglu AM (2018) Context-aware computing, learning, and big data in internet of things: a survey. IEEE Internet Things J 5(1):1–27
26. Chatterjee B, Cao N, Raychowdhury A, Sen S (2019) Context-aware intelligence in resource-constrained IoT nodes: opportunities and challenges. IEEE Design Test 36(2):7–40

Breast Cancer Detection in Histology Images Using Convolutional Neural Network

Soumya Sara Koshy and L. Jani Anbarasi

Abstract Breast cancer has been one of the leading causes of death among women in the world. The death rates due to breast cancer can be reduced by early detection. The normal symptoms of breast cancer are masses or lumps that feel different from the other tissues, and there are two types of masses- benign and malignant. Benign masses are abnormal growths which are not cancerous, whereas malignant masses are cancerous. Various methods have been proposed to detect breast cancer. A detailed survey has been done in this paper. Deep learning techniques are now widely used for cancer detection, and in deep learning, features are extracted from input data using multiple layers. In this paper, transfer learning techniques have been used to detect breast cancer, and pretrained models like InceptionResnetV2, VGG-16 and VGG-19 are used. In transfer learning, pretrained models are used to solve a new problem where training can be done using a small amount of data. A comparison of those techniques have also been done. Publicly available dataset named BreaKHis dataset has been used for the study.

Keywords Breast cancer detection · Deep learning · Convolutional neural network · Histology images

1 Introduction

Breast cancer has been one of the deadliest disease affecting women. It affects the breast cells [1]. Breast cancer affects both men and women, but the breast cancer rate is found to be higher among women. Main symptoms of breast cancer are masses or lumps that are different from the surrounding tissue. Masses can be either benign or malignant. Benign masses are unusual growth that are not cancerous and are not life

S. Sara Koshy (✉) · L. Jani Anbarasi
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

L. Jani Anbarasi
e-mail: janianbarasi.l@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_7

threatening. Benign masses do not grow outside the breast. Malignant masses are cancerous and are life threatening. Malignant masses can damage the surrounding tissues.

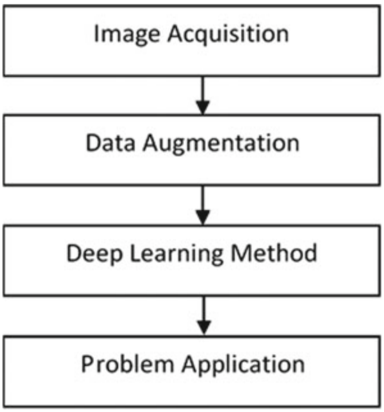
Detection at an earlier stage can reduce mortality rate due to breast cancer to a great extent. Mammograms, breast ultrasound, biopsy, and MRI are the techniques commonly used to diagnose breast cancer. X-ray images of the breast are taken in mammogram. Ultrasound uses sound waves to distinguish between tumour cells and benign cyst. Different images of the breast are combined in MRI to help the doctors to detect tumours and are usually performed as a follow up of mammogram and ultrasound [2]. Pictures of the interior of the breast are created using radio waves and magnet in MRI.

Traditional methods of detecting breast cancer involve analysing the breast images manually by a pathologist which is a time consuming process and often the results obtained may not be that accurate. With the technological advancements, various computer-aided design methods have been proposed for breast cancer detection with increased accuracy, of which the prominent ones are the machine learning and deep learning techniques. In machine learning, machines are provided with the ability to learn themselves without being programmed explicitly and can be supervised learning or unsupervised learning.

Deep learning is a type of machine learning in which features are extracted using multiple layers. Deep learning mimics the working of neurons in brain where different layers are used to learn from the data. Deep learning model can have any number of layers, and the number of layers determines the depth of the model. Data is given to the input layer, and output is obtained from the output layer. Between the input layer and output layer, any number of hidden layers can be present. Deep learning requires a lot of data, and the data is fed through neural networks [3]. The steps involved in deep learning are shown in Fig. 1.

The paper includes the following sections: Various deep learning methods for breast cancer detection have been reviewed in Sect. 2. Dataset used is discussed in

Fig. 1 Process flow of deep learning [4]



Sect. 3. Data augmentation used is discussed in Sect. 4. Section 5 covers the methods used, and Sect. 6 is discussions and conclusion.

2 Related Works

Khan et al. [5] used GoogLeNet, VGGNet, and ResNet to extract different low-level features separately, and later, the extracted features are combined. An accuracy of 97.525% is achieved, without training from scratch, thus, improving the classification efficiency. The deep learning framework used by Khan et al. [5] is shown in Fig. 2.

Wang et al. [6] used a mass detection method for extraction of region of interest, and the features were extracted. With the extracted features and labels, classifiers are trained. The method combined objective features and subjective features, that is the doctor's experience and the mammogram features. Extreme learning machine classifier is used.

Li et al. [7] used a fully convolutional autoencoder to learn the prominent patterns among normal image patches. Then, the patches that are different from the normal patches are detected and analysed.

Perre et al. [8] used transfer learning approach. Three pretrained models- CNN-F, CNN-M, and Caffe have been used, and the model is pretrained using ImageNet dataset. Handcrafted features including intensity features, textures features, and shape features were used, resulting in improved classification efficiency.

Ragab et al. [9] used two segmentation approaches, where initially manual ROI is determined, and later, region-based and threshold segmentation were performed. For feature extraction, deep convolutional neural networks were used in which

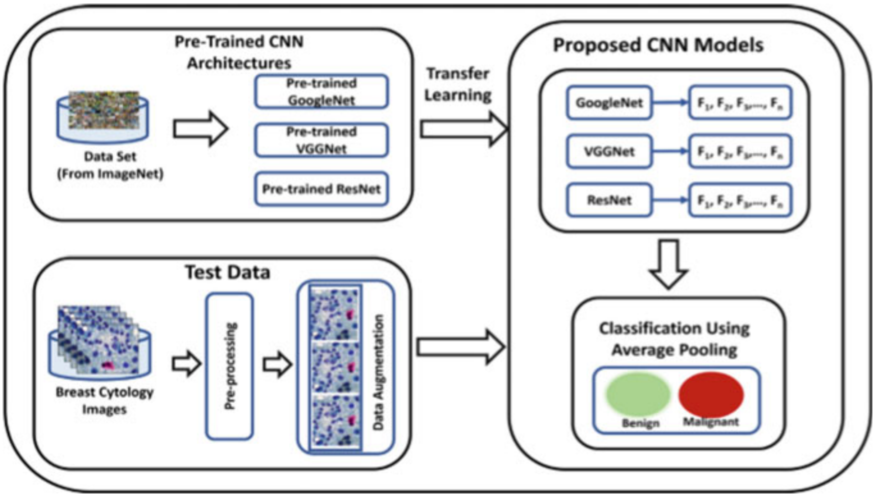


Fig. 2 Deep learning framework used by Khan et al. [5]

the last layer is replaced with SVM. An accuracy of 79% is achieved for manual segmentation, and for the automated ROI process, an efficiency of 94% is obtained.

A method for detecting invasive ductual carcinoma of breast cancer is proposed by Brancati et al. [10]. The area of the region of interest is identified and not the exact boundaries of the region of interest. The whole slide images are divided into patches which are marked with either invasive ductual carcinoma label or non-invasive ductual carcinoma label. Initially, training was performed using an unsupervised manner for extracting features and to reconstruct the input images. A stochastic gradient descent algorithm was used to implement back propagation algorithm. Sparsity constraint is included in hidden units to prevent overfitting. Later, supervised classification was performed using a convolutional autoencoder named supervised encoder FusionNet (SEF), where training is done only on the encoding part. This method achieved an accuracy of 97.67% and had lower standard deviation of accuracy.

Gecer et al. [11] used four fully connected convolutional neural networks that can handle images at different magnifications to remove the irrelevant details and localize the region of interest. Whole slide images are classified into five classes using another convolutional neural network. Later, labelling of the whole slide is performed pixel wise, and an overall slide level classification accuracy of 55% was obtained.

Rakhlin et al. [12] used a deep convolutional feature representation method where unsupervised feature representation extraction is performed, and deep convolutional neural networks trained on ImageNet were used. Sparse descriptors of low dimensionality are obtained followed by supervised classification using LightGBM for implementing gradient boosted trees. Two-class classification and four-class classification were performed and achieved an accuracy of $93.8 \pm 2.3\%$ for two-class classification, and for four-class classification, the accuracy obtained was $87.2 \pm 2.6\%$.

Vesal et al. [13] used transfer learning technique in which two pretrained convolutional neural networks, namely Inception-V3 and Resnet50 are used. Yap et al. [14] investigated three different methods—U-Net, a patch-based LeNET, and a transfer learning approach with AlexNet. Two datasets are used, and the LeNET achieved an F-score of 0.91 on both datasets, and U-Net achieved an F-score of 0.89 and 0.78 and AlexNet achieved an accuracy of 0.92 and 0.88 on first and second datasets, respectively.

Spanhol et al. [15] extracted features from images and used them as the input to classifier. Output of a previously trained convolutional neural network is fed into these classifiers that are trained on problem-specific data. The pretrained BVLC CaffeNet model is used and was trained on the ImageNet dataset. Spanhol et al. [16] extracted non-overlapping grid patches either randomly or using a sliding window. The results of all the patches of image are combined using fusion rules which are sum, product, and max.

Sun et al. [17] used both labelled and unlabelled data. The proposed method is shown in Fig. 3 and is helpful in cases where it is difficult to obtain labelled data, and an accuracy of 82.43% was obtained.

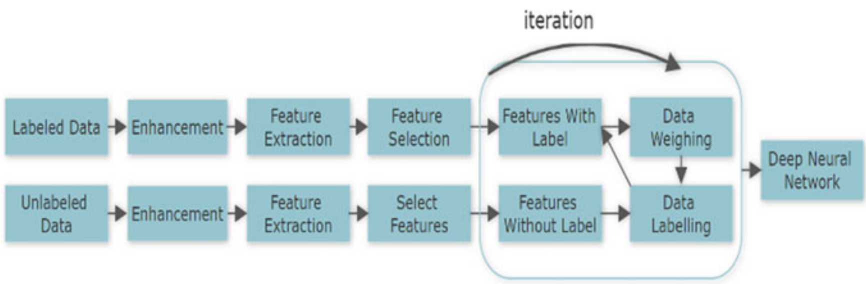


Fig. 3 Method used by Sun et al. [17]

Abdel-Zaher [18] proposed a deep belief path followed by a back propagation path for breast cancer detection. Deep belief path is an unsupervised path, and back propagation path is a supervised path, and back propagation neural network is constructed using Liebenberg Marquardt learning algorithm. An accuracy of 99.68% was obtained.

Ciresan et al. [19] detected breast cancer by detecting the presence of mitosis. The central coordinate of single mitosis is found out, and training is performed using that information. DNN-based pixel classifier is used for detection and obtained an F-score of 0.782. Methods used by different authors are summarized in Table 1.

3 Dataset

The publicly available dataset named BreakHis [20] is used for the experiments. BreakHis dataset is composed of 9109 microscopic images of which 2480 are benign and 5429 are malignant. Images are collected from 82 patients using varying magnifying factors. Images are of PNG format with size 700 × 460 pixels. Images consist of three channels red, green, and blue with each channel of 8 bit depth. Benign images from the dataset are shown in Fig. 4a, and malignant images from the dataset are shown in Fig. 4b.

4 Data Augmentation

Images with magnification factor 400X are used for the study. The size of the dataset is increased by applying various geometric transformations. The augmentation techniques used are rotation, width shifting, height shifting, shearing, horizontal, and vertical flipping.

Table 1 Methods used by various researchers

References	Method
[5, 8, 13, 14]	Transfer learning
[6]	ROI using mass detection Feature extraction Labelled classification
[7, 10]	Detection of invasive carcinoma breast cancer Patch formation Supervised and unsupervised classification Supervised encoder FusionNet
[9]	Threshold and region-based segmentation Feature extraction and classification using deep convolutional neural network
[11]	Multiclass classification Five classes
[12]	Deep convolutional feature extraction Supervised and unsupervised classification
[15]	Pretrained convolutional neural network
[16]	Supervised training using patches Classification by combining the patches
[17]	Semi-supervised learning Convolutional neural network
[18]	Deep belief path followed by back propagation path
[19]	Mitosis detection

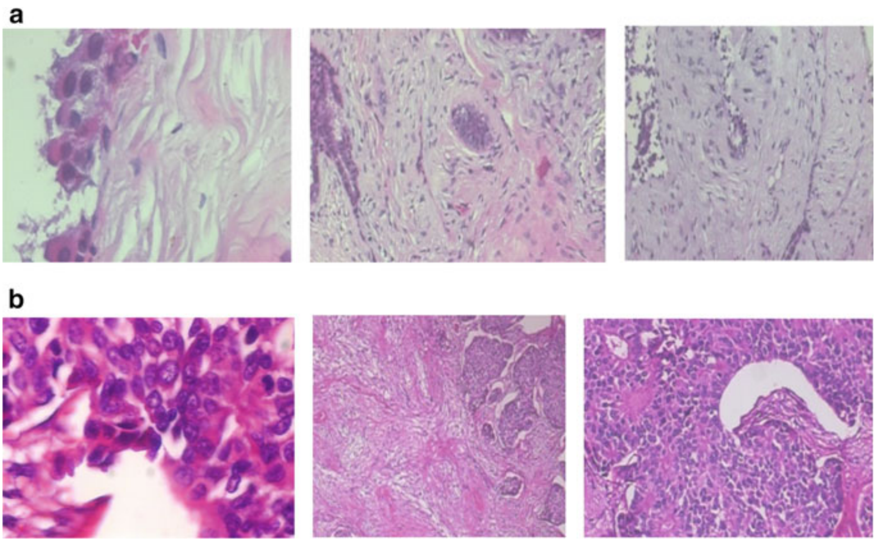


Fig. 4 **a** Benign images, **b** malignant images

5 Method

In this paper, the technique of transfer learning has been used to detect the cancer in the histology images. Three pretrained convolutional neural networks, namely VGG-16, VGG-19, and inceptionresnetv2 have been used. VGG-16 has been used in many classification problems and is easy to implement. VGG-19 is a variant of VGG-16 that includes 19 layers. The concept of batch normalization is introduced in inception resnetv2, and higher learning rate can also be used. Transfer learning is a method in which pretrained networks have been used to solve new problems. The knowledge gained by solving one problem can be used to solve another problem.

5.1 *Performance Evaluation*

Accuracy [6] is calculated as,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

Sensitivity [6] is calculated as,

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \tag{2}$$

Specificity [6] is calculated as,

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \tag{3}$$

where TP is the true positive, TN is the true negative, FP is the false positive, and FN is the false negative.

5.2 *Convolutional Neural Network*

Convolutional neural network is a type of deep neural networks. It primarily includes four layers—convolution, pooling, flattening, and fully connected layer which is shown in Fig. 5.

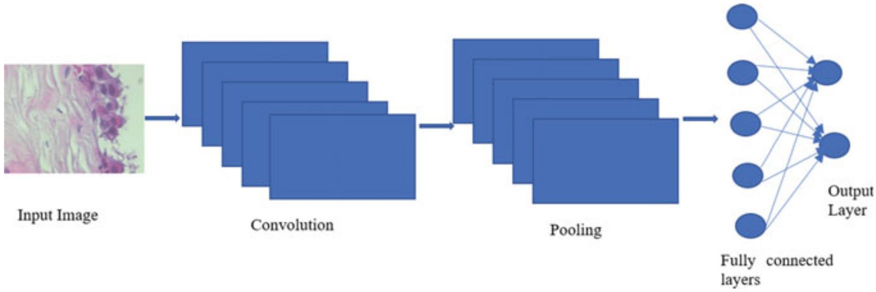


Fig. 5 Convolutional neural network

```
Model: "sequential_4"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg16 (Model)                (None, 7, 7, 512)         14714688
_________________________________________________________________
flatten_4 (Flatten)          (None, 25088)             0
_________________________________________________________________
dense_6 (Dense)              (None, 1024)              25691136
_________________________________________________________________
dropout_4 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_7 (Dense)              (None, 1)                 1025
=================================================================
Total params: 40,406,849
Trainable params: 25,692,161
Non-trainable params: 14,714,688
_________________________________________________________________
```

Fig. 6 Model parameters—VGG 16

5.3 VGG-16

VGG-16 is a 16 layer deep neural network trained on Imagenet dataset. The training is done for 30 epochs, and the optimizer used is RMSprop. The loss function used is binary cross entropy, and sigmoid activation function is used in the fully convolutional layer. Figure 6 shows the model parameters. The accuracy for the model and the loss is shown in Fig. 7.

5.4 VGG-19

VGG-19 is a 19 layers deep convolutional neural network trained on Imagenet dataset. The training is done for 30 epochs, and the optimizer used is RMSprop. The loss function used is binary cross entropy, and sigmoid activation function is used in the

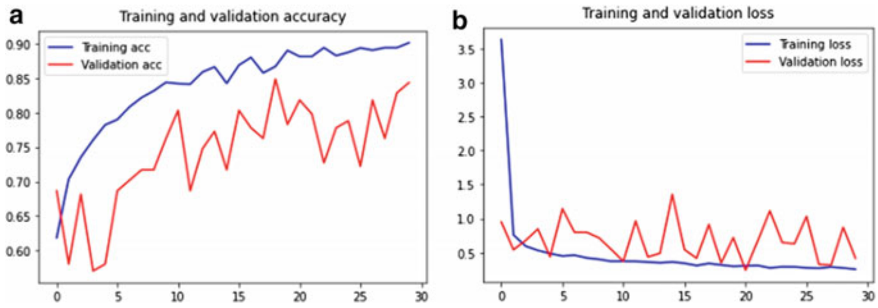


Fig. 7 **a** Training accuracy and validation accuracy-VGG 16, **b** training loss and validation loss—VGG 16

```
Model: "sequential_6"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg19 (Model)                (None, 7, 7, 512)         20024384
_________________________________________________________________
flatten_6 (Flatten)          (None, 25088)             0
_________________________________________________________________
dense_13 (Dense)             (None, 1024)              25691136
_________________________________________________________________
dropout_6 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_14 (Dense)             (None, 1)                 1025
=================================================================
Total params: 45,716,545
Trainable params: 25,692,161
Non-trainable params: 20,024,384
_________________________________________________________________
```

Fig. 8 Model parameters—VGG 19

fully convolutional layer. Figure 8 shows the model parameters. The accuracy for the model and the loss is shown in Fig. 9.

5.5 *Inception-Resnet-V2*

Inception-Resnet-V2 is a 164 layers deep convolutional neural network trained on Imagenet dataset which contains millions of images. The model was trained for 15 epochs, and the optimizer used is Adam. Figure 10 shows the model parameters. The accuracy for the model and the loss is shown in Fig. 11.

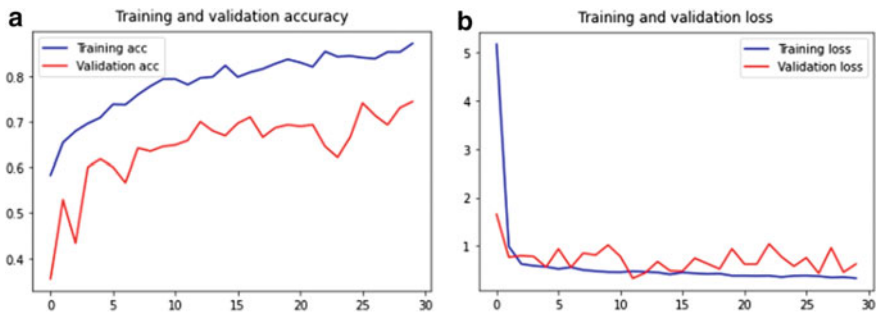


Fig. 9 **a** Training accuracy and validation accuracy—VGG 19, **b** training loss and validation loss—VGG 19

```
block8_9_ac (Activation)        (None, 8, 8, 2080)  0          block8_9[0][0]
____________________________________________________________________________________________
conv2d_200 (Conv2D)             (None, 8, 8, 192)   399360     block8_9_ac[0][0]
____________________________________________________________________________________________
batch_normalization_200 (BatchN (None, 8, 8, 192)   576        conv2d_200[0][0]
____________________________________________________________________________________________
activation_200 (Activation)     (None, 8, 8, 192)   0          batch_normalization_200[0][0]
____________________________________________________________________________________________
conv2d_201 (Conv2D)             (None, 8, 8, 224)   129024     activation_200[0][0]
____________________________________________________________________________________________
batch_normalization_201 (BatchN (None, 8, 8, 224)   672        conv2d_201[0][0]
____________________________________________________________________________________________
activation_201 (Activation)     (None, 8, 8, 224)   0          batch_normalization_201[0][0]
____________________________________________________________________________________________
conv2d_199 (Conv2D)             (None, 8, 8, 192)   399360     block8_9_ac[0][0]
____________________________________________________________________________________________
conv2d_202 (Conv2D)             (None, 8, 8, 256)   172032     activation_201[0][0]
____________________________________________________________________________________________
batch_normalization_199 (BatchN (None, 8, 8, 192)   576        conv2d_199[0][0]
____________________________________________________________________________________________
batch_normalization_202 (BatchN (None, 8, 8, 256)   768        conv2d_202[0][0]
____________________________________________________________________________________________
activation_199 (Activation)     (None, 8, 8, 192)   0          batch_normalization_199[0][0]
____________________________________________________________________________________________
activation_202 (Activation)     (None, 8, 8, 256)   0          batch_normalization_202[0][0]
____________________________________________________________________________________________
block8_10_mixed (Concatenate)   (None, 8, 8, 448)   0          activation_199[0][0]
                                                               activation_202[0][0]
____________________________________________________________________________________________
block8_10_conv (Conv2D)         (None, 8, 8, 2080)  933920     block8_10_mixed[0][0]
____________________________________________________________________________________________
block8_10 (Lambda)              (None, 8, 8, 2080)  0          block8_9_ac[0][0]
                                                               block8_10_conv[0][0]
____________________________________________________________________________________________
conv_7b (Conv2D)                (None, 8, 8, 1536)  3194880    block8_10[0][0]
____________________________________________________________________________________________
conv_7b_bn (BatchNormalization) (None, 8, 8, 1536)  4608       conv_7b[0][0]
____________________________________________________________________________________________
conv_7b_ac (Activation)         (None, 8, 8, 1536)  0          conv_7b_bn[0][0]
____________________________________________________________________________________________
global_average_pooling2d (Globa (None, 1536)        0          conv_7b_ac[0][0]
____________________________________________________________________________________________
dense (Dense)                   (None, 512)         786944     global_average_pooling2d[0][0]
____________________________________________________________________________________________
dense_1 (Dense)                 (None, 256)         131328     dense[0][0]
____________________________________________________________________________________________
dense_2 (Dense)                 (None, 2)           514        dense_1[0][0]
============================================================================================
Total params: 55,255,522
Trainable params: 918,786
Non-trainable params: 54,336,736
____________________________________________________________________________________________
```

Fig. 10 Model parameters—InceptionResnetV2

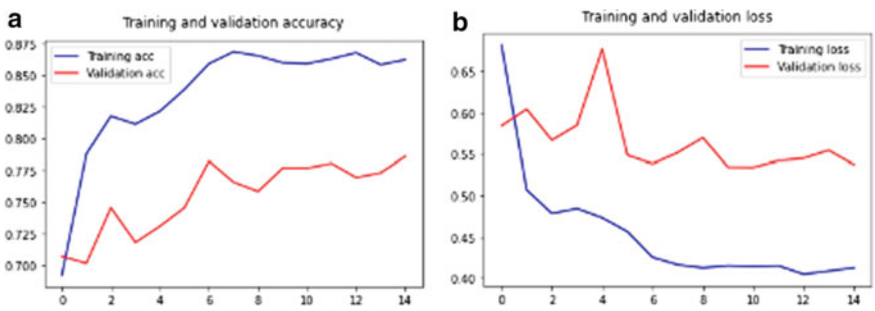


Fig. 11 **a** Training accuracy and validation accuracy-InceptionResnetV2, **b** training loss and validation loss-InceptionResnetV2

6 Discussions and Conclusion

This paper analysed various methods for detecting breast cancer in histology images using deep learning techniques. Transfer learning technique is used in the proposed paper where the network is trained using the pretrained models. Three convolutional neural networks are used to analyse the histology images. The highest accuracy is obtained for VGG-16 which is 82.83%, and for VGG-19 and ResNet-50, the accuracies obtained are 73.04 and 78.57%, which can be successfully used to classify benign and malignant images. The work is going on, to further improve the accuracy by changing the network architecture and fine-tuning the hyper parameters.

References

1. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, Jemal A, Siegel RL (2019) Breast cancer statistics, 2019. CA A Cancer J Clin 69:438–451. https://doi.org/10.3322/caac.21583
2. Sharon JJ, Anbarasi LJ (2018) Diagnosis of DCM and HCM heart diseases using neural network function. Int J Appl Eng Res 13(10):8664–8668
3. Anbarasi LJ, Anandha Mala GS, Narendra M (2015) DNA based multi-secret image sharing. Procedia Comput Sci 46:1794–1801
4. Jimenez-del-Toro O, Otálora S, Andersson M, Eurén K, Hedlund M, Rousson M, Müller H, Atzori M (2017) Chapter 10 Analysis of histopathology images from traditional machine learning to deep learning, biomedical texture analysis. Elsevier
5. Khan S, Islam N, Jan Z, Ud Din I, Rodrigues JJPC (2019) A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recogn Lett 125:1–6. ISSN: 0167-8655
6. Wang Z et al (2019) Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features. IEEE Access 7:105146–105158
7. Li X et al (2019) Discriminative pattern mining for breast cancer histopathology image classification via fully convolutional autoencoder. IEEE Access 7:36433–36445
8. Perre AC, Alexandre LA, Freire LC (2018) Lesion classification in mammograms using convolutional neural networks and transfer learning. Comput Methods Biomech Biomed Eng: Imaging Vis. https://doi.org/10.1080/21681163.2018.1498392

9. Ragab DA, Sharkas M, Marshall S, Ren J (2019) Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ 7:e6201. https://doi.org/10.7717/peerj.6201
10. Brancati N, De Pietro G, Frucci M, Riccio D (2019) A deep learning approach for breast invasive ductal carcinoma detection and lymphoma multi-classification in histological images. IEEE Access 7:44709–44720. https://doi.org/10.1109/ACCESS.2019.2908724
11. Gecer B, Aksoy S, Mercan E, Shapiro LG, Weaver DL, Elmore JG (2018) Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. Pattern Recogn 84:345–356. ISSN: 0031-3203
12. Rakhlin A et al (2018) Deep convolutional neural networks for breast cancer histology image analysis. Image Anal Recogn 737–744
13. Vesal S et al (2018) Classification of breast cancer histology images using transfer learning. In: International conference image analysis and recognition. Springer, Cham
14. Yap MH et al (2018) Automated breast ultrasound lesions detection using convolutional neural networks. IEEE J Biomed Health Inform 22(4):1218–1226. https://doi.org/10.1109/JBHI.2017.2731873
15. Spanhol FA, Oliveira LS, Cavalin PR, Petitjean C, Heutte L (2017) Deep features for breast cancer histopathological image classification. In: 2017 IEEE International conference on systems, man, and cybernetics (SMC). Banff, AB, pp 1868–1873. https://doi.org/10.1109/SMC.2017.8122889
16. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) Breast cancer histopathological image classification using convolutional neural networks. In: 2016 International joint conference on neural networks (IJCNN). Vancouver, BC, pp 2560–2567
17. Sun W, (Bill) Tseng T-L, Zhang J, Qian W (2017) Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Comput Med Imaging Graph 57:4–9. ISSN: 0895-6111
18. Abdel-Zaher AM, Eldeib AM (2016) Breast cancer classification using deep belief networks. Expert Syst Appl 46:139–144. ISSN: 0957-4174
19. Cireşan D, Giusti A, Gambardella L, Schmidhuber J (2013) Mitosis detection in breast cancer histology images with deep neural networks. Medical image computing and computer-assisted intervention: MICCAI. In: International conference on medical image computing and computer-assisted intervention, vol 16, pp 411–418. https://doi.org/10.1007/978-3-642-40763-5_51
20. Spanhol F, Oliveira LS, Petitjean C, Heutte L (2016) A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng (TBME) 63(7):1455–1462

A Novel Approach on Auto-Scaling for Resource Scheduling Using AWS

I. George Fernandez and J. Arokia Renjith

Abstract The mostly wide adoptions in cloud computing for businesses have the several reason; among that the elasticity in cloud computing is the virtual infrastructures which will be the dominating leader. The most important area is elasticity that allows to auto-scale the resources which are on-demand. On the other hand, Web applications typically comprise dynamic workload and are hard to predict. The cloud service providers and researchers are jointly working together to condense the cost at the same time as maintaining quality of service (QoS). AWS command-line interface (CLI): the command-line tools, which is Python written, introduce the efficient use cases for managing the AWS services through the sets precisely modest in command. Forecasting on historical data by way of inputting in order to make the up-to-date assessment which are analytical in modeling a path for forthcoming trends. Prophet which acts as a tool will be built to report such concerns, and it delivers real-world methodology in forecasting that is at scalable. From the experimental result, it is found that when predicted value is greater than 51% of CPU usage, then a new EC2 instance is created.

Keywords Amazon web services · Amazon EC2 · Auto-scaling · Prophet · Time series · Elastic load balance

1 Introduction

Scalability in cloud applications is being main reasons behind the wide adoption for cloud computing. Most of the IaaS cloud services providers (CSP) offer auto-scaling services to familiarize the VMs to the shifting demand. A virtual entity of new type called container, which allows user to proposal the loosely coupled applications

I. George Fernandez (✉)
Department of Information Technology, Jerusalem College of Engineering, Chennai, Tamil Nadu, India

J. Arokia Renjith
Department of Computer Science and Engineering, Jeppiaar Engineering College, Chennai, Tamil Nadu, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_8

containing of several small building blocks. These building block such as microservices implements a minor fixed of functions and interconnect with other microservices. In auto-scaling, the decision-making technique provisioned to allocate the number of resources to divergent process is classified as reactive and proactive. The reactive technique regularly monitors events, namely CPU utilization, workloads, queues, etc., and it performs the elastics operation on resources based on threshold. In proactive techniques, forecasting methods are used to predict traffic from past workload.

Bitbrains, a service provider, which is specialized in business computation and managed hosting for enterprises. The requirements would come in form of request such as data transfers among the customer, and the data center through protected networks, the calculate nodes are rented as virtual machines (VM) cutting-edge data center which delivers a expectable routine and provide in elevation obtainability for consecutively of business-critical simulation.

Prophet is a technique aimed at foretelling, time series statistics which are constructed on an additive model such that nonlinear trends is appropriate with the annual or weekly and on everyday seasonality, in addition to leave effects. This work finest through the time series that invent healthy periodic effects and numerous spells of the past data.

Time series forecasting finds significant problem with various industrial applications such as the trade request estimating, economic forecasts, prediction of stream of traffic, or the climate patterns. In broad, these shows a crucial part in the industrializing commercial developments. These methods are included in the well-known auto-regression (AR), ARIMA, the classical Box-Jenkins methodology, exponential smoothing, and more precisely linear state-space models. Conversely, these approaches will not certainly accessible to enormous datasets through the lots of time series, owed the essential for the specific training. Furthermore, they will not help from the mutual temporal configurations in whole dataset although the prediction and training.

Amazon Web services (AWS): Amazon provides auto-scaling services in particular IaaS public cloud. Amazon Web services auto-scaling is precisely auto-scaling group (ASG). The auto-scaling group is a fixed of various Amazon elastic compute cloud (EC2) instance, the virtual machines (VM's) which are sharing the similar. Consequently, each VM pioneering the collection consumes the similar Amazon machine image in addition the identical hardware features. Load distribution among the virtual machines (VMs) are automated by elastic load balancer (ELB). Amazon cloud watch is performance monitoring tool, be responsible for the performance data cast-off in scaling rules.

1.1 Reactive Auto-Scaling

Reactive auto-scaling (automatic scaling) technology that enabled for automatic provisioning as well as termination of the virtual entities to familiarize the resource capacity for changing demand. The automation is accomplished via a monitoring service to retrieve related resource utilization metrics by which the alarms and also triggers can be defined. Reactive auto-scaling is one such type of auto-scaling which will either deploy or will terminate based on predefined amount of the virtual entities as a response to change of some metrics. Therefore, reactive auto-scaling proceeds only on the basis of given parameters, which can be either provided or measured by the monitoring solution or by the cloud administrator. The number of changes will be the number of virtual entities which are allocated or terminated is encoded based set of rules.

1.2 Predictive Scaling

Predictive scaling or proactive scaling influences historical data just about the virtual infrastructure and the application which is collected via the monitored solutions. Collected historical data are in different forms, namely: logs, application traces, time series, etc. These data will be required in the derivation of the prototypes that are used to generalize, future values of the particular metrics. For example, the poised requests per second time series are used to develop a model and to forecast such as predict or extrapolate, the requests per second values for a particular service at some instant in near future.

2 Literature Review

Papadopoulos et al. [1] proposed an auto-scaling policy on performance evaluation approach constructed on a chance constrained optimization problem explained by means of scenario theory. The approach was implemented in performance evaluation framework for auto-scaling (PEAS) and tested on the several existing auto-scaling policies using 796 real workload traces. The work introduced a numerous distinct metrics to estimate the autoscaling performance with the core metrics of the average number of under provisioned resources and over provisioned resources. Ilyushkin et al. [2] proposed a set of performance metrics to estimate the auto-scaling policy. The set which includes, namely wrong-provisioning timeshare, under- and over-provisioning accuracy, instability as well as the other user-oriented metrics such as elastic slow-down, response time, wait time, average task throughput, or average number of resources. Bauer et al. [3] investigated predictive auto-scaling solution: (a) collects the monitoring data for forecasted parameters, (b) derives forecast models, (c) derives the virtual infrastructure and the application performance models, (d) derives the scaling policy to guarantee provision of virtual entities that are be able to assist forecast workload, and (e) finally executes the scaling action for forecast [4]. Anshul et al. [5] proposed experimental evaluations of the multilayered auto-scaling, performance with the mixture on virtual setup auto-scaling of Google compute engine, AWS, and Microsoft Azure by the pods for flat auto-scaling of the Kubernete via demonstrating ScaleX based on four typical patterns of load. Benjamin et al. [6] demonstrated performance analysis [7] and for sake of comparisons and evaluated the forecasted procedures and also introduced a tool which enable the analyst to make use of their expertise reliable and forecasting business time series more practically. Fang et al. [8] analyzed that prophet and LSTM which are used in prediction for the trends of the time series data, and author studied that prediction trends can be united by the contrary neural network exemplary aimed at the prediction. Shen et al. [9] analyzed business-critical workloads hosting in distributed data center with 1,750 virtual machines workload traces of long term and large scale. Traces are also analyzed on actual resource usage as well as requested resources, in terms of CPU, memory, disk I/O, and network I/O. Walid et al. [10] proposed the optimization framework which can adaptively solve the joint VM-to-PM packing problem and VM auto-scaling. Prathanra et al. [11] investigated that the performances of the machine knowledge prototypes can forecast recital of the Jupyter notebook on the JupyterHub popular relationships of their response time.

3 System Architecture

In the proposed model, complete architecture is depicted (Fig. 1 shows proposed system architecture), which contains five components such as AWS auto-scaling group, elastic load balancer (ELB), prophet, prediction manager, and cloud watch. The various works are allocated according to the various requests by cloud users.

3.1 AWS Auto-Scaling Group

The AWS launch configuration is used to form the auto-scaling group (ASG) in Amazon Web service cloud using the user metrics. The auto-scaling group is a usual of dissimilar Amazon elastic compute cloud (EC2) instances.

3.2 Elastic Load Balancer (ELB)

Load distribution among the virtual machines (VMs) is automated by elastic load balancer (ELB). To direct the loads to the auto-scaling group, the elastic load balancer (ELB) remains added. It functions as a distinct endpoint aimed at the load generations workload, which in turn, the load is disseminated between the ASG instances.

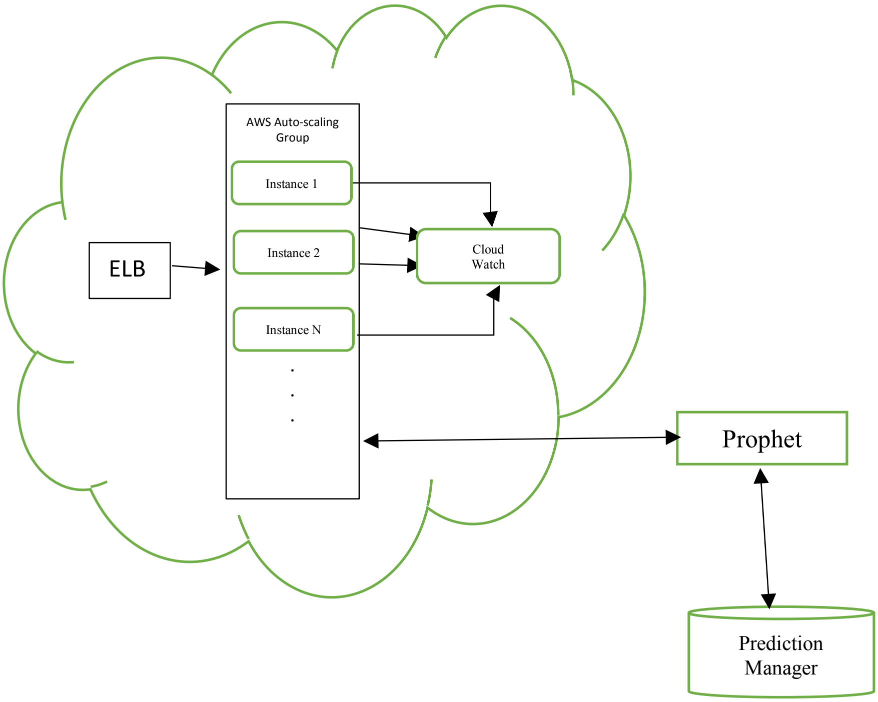


Fig. 1 Proposed system architecture

3.3 Prophet

Prophet is sklearn exemplary API, which can generate an instances of prophet class, call the fit besides prediction methods expectially. The input parameter to the prophet exists data frame has two columns, namely y and ds. The y column necessity must be numeric. The date stamp (ds) column must be of a setup likely by Pandas. It represents the magnitude that anticipated to estimate.

3.4 Prediction Manager

The prediction manager manages the prediction rules that are available to predict according to the prophet predict method to forecast.

3.5 AWS Cloud Watch

Amazon cloud watch is performance monitoring tool, be responsible for the performance data castoff in scaling rules.

4 Proposed Work

In our proposed work, we investigate the active traces explanatory intended at business-critical capabilities. Business-critical capabilities regularly comprise requests on creditworthiness areas, and these are repeatedly Monte Carlo imitation balanced marketable exposing requests. The Auxiliary requests that illustrate business-critical capabilities are database, CRM, email, and management plus collaborative amenities.

4.1 Traces Collected

In proposed work, we obtain traces which cover both actually used and requested resources, basically four types of resource such as memory, CPU, network, and disk. The traces collected from August to September 2013, the bit fastStorage which comprises of 1250 virtual machines (VM's), and fastStorage trace contains a complex section of application servers and the compute nodes.

4.2 Dataset

The records are systematized into three subdirectories in terms of month that each metrics remain documented. The setup specification of individually file is racket wise; each such racket signifies an opinion for concert metrics such as timestamp in milliseconds, CPU usage in percentage, the CPU's are in cores, the CPU's capability in MHZ, memory's capability in kilobyte (KB), the memory's usage in KB, the disk read throughput in kilobytes per second, the disk write throughput in kilobytes per second, network transmitted throughput in kilobyte per sec network received throughput in kilobytes per second.

4.3 AWS Auto-Scaling

Boto3 which is the AWS SDK used for Python that provides a low-level direct access and object-based API's to AWS services such as elastic compute cloud. AWS command-line Iinterface (CLI): The command-line tool which is Python programmed that familiarizes a well-organized use cases for managing the Amazon Web services through established of precise modest instructions.

4.4 Experimental Setup

We install the AWS CLI and the Python Boto3 libraries, in order to create our own consumer authorizations for the Amazon Web service solace, since AWS amenities will remain accessible programmable. After creating the user and obtaining the credentials (Access ID and Secret key). Python scripting environment is configured with credentials for managing EC2 instances. We now build a key pair for EC2 instances, which can be accessed in later stages, so that virtual machines (VMs) are launched programmatically by using the Python.

```
Ec2_instances = elasticcomputecloudec2.Create_AW instances (
node_Id = 'ami-0323c3dd2da7fb37d',
Count_MIN = 1,
Count_MAX = 3,
Type_Instances = 't1.micro',
Name_Key='auto-key pair')
```

From the above program code, Node_ID which stipulates, Amazon machine image (AMI) ID of the AWS EC2 illustration. Count_MIN and Count_MAX that are used to express, the quantity of EC2 occurrences which are to be launched. For example, Count_MIN = 1 and Count_MAX = 3, number of instances launched are said to be 3. The Typt_Instances specify instance size such as: t1.micro, t1.small, or M3.large. The Name_key is name distinct for key pair which will permit to[12] admittance EC2 instances; for example, in our proposed work, we use the name "auto-key pair" which is created in the AWS proposed scenario. The virtual machine configuration in proposed on clouds is given in Table 1. The AWS AMI image for operating system we used for the VM configuration is centos 7.

From Table 2, we encounter minimum instances and maximum instances that remain cast-off to explain, the quantity of EC2 instances which are near be launched. For example, minimum instances are 1, and maximum instances are 3; number of maximum instances launched are said to be 3. The scaling metrics are CPU utilization, and the threshold to launch a new instance is said to be 50%.

Table 1 Virtual machine configuration of proposed

Storage memory	Type_Instance	Virtual CPUs
1 GB	t1.micro	1 CPU

Table 2 Configuration of Amazon web service auto-scaling

Metrics for scaling	Threshold (%)	Minimum instance	Maximum instances
CPU usage	50	1	3

5 Results and Discussion

In proposed work, we obtain traces which are the used and requested resources, here four types of resource such as memory, CPU, network, and disk. The traces collected from August to September 2013, the trace fastStorage which consists of 1250 virtual machines (VMs).

5.1 Dataset Importing

To prophet involvement stays continually a data frame through two stakes, namely y and ds. The ds stands for date stamp stake, always in format predictable in Pandas. For the analysis in our proposed, we are using an excel file that contains a total of "CPU Usage in Percentages," traces collected from August to September 2013, the trace fastStorage which consists of 1250 virtual machines (VMs).

5.2 Converting the Dataset to Prophet Compliant

We can convert the historical dataset to be prophet compliant. Now convert the given data into the formats that are desired by prophet from Table 3. It is renamed as the date: ds and the CPU Usage (%): y.

Table 3 Converted dataset to prophet compliant

	Date	CPU usage (%)
0	2013-08-12 13:40:46	40.866667
1	2013-08-12 13:45:46	42.100000
2	2013-08-12 13:50:46	40.733333
3	2013-08-12 13:55:46	44.000000
4	2013-08-12 14:00:46	42.300000

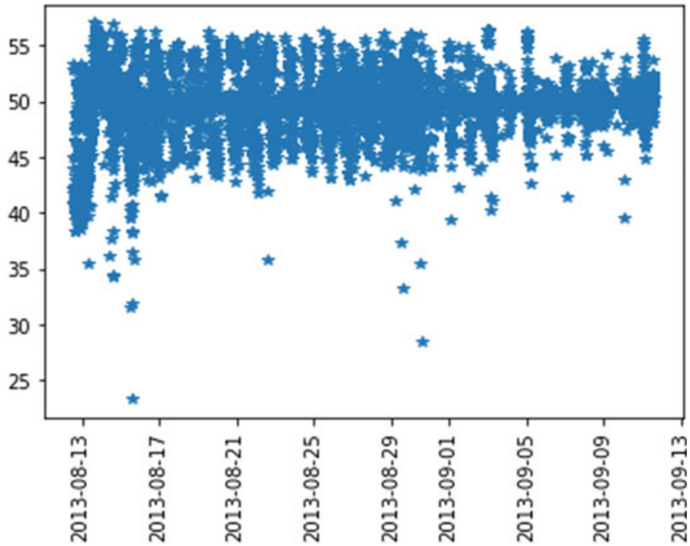


Fig. 2 Forecast plotting of CPU usage versus timestamp

5.3 *The Forecasting in Prophet*

To forecast, in prophet for prediction in near future. A data frame is created for the purpose of future predictions by making use of make_future_dataframe. Predict the CPU Usage (%) for upcoming 300 s. Forecast plotting using prophet: Forecast can be plotted by calling method as Prophet.plot, and it is passed to forecast data frame.

From (Fig. 2 CPU Usage versus timestamp), we can forecast the CPU Usage in percentage with timestamp, logs from August to September 2013.

5.4 *The Proposed Future Prediction*

When predicted value is greater than 51% of CPU usage for 300 s in next 50 min, then create new EC2 instance. The forecast technique shall allocate for, respectively, rows in the forthcoming of prediction assessment which it is named as yhat.

Algorithm for Future Prediction

```
Step 1: if len(out[out['yhat']>51])>300
Step 2: r=requests.get(url=\"http://localhost:5000/api/start\")
Step 3:data = r.json()
Step 4:print(data)
Step 5: else
Step 6: x=requests.get(url=\"http://localhost:5000/api/stop\")
Step 7:data2 = x.json()
```

```
Step 8:print(data2)
Step 9: end
```

From the (Fig. 3 future prediction), it represents the quantity we desire to forecast, and we can predict that when the CPU Usage increased above 51 percentages; then, a new instance of AWS EC2 is created.

From the (Fig. 4 CPU Utilization), the cloud watch monitoring tool of AWS experimental result found that only about 9% of CPU utilization is done.

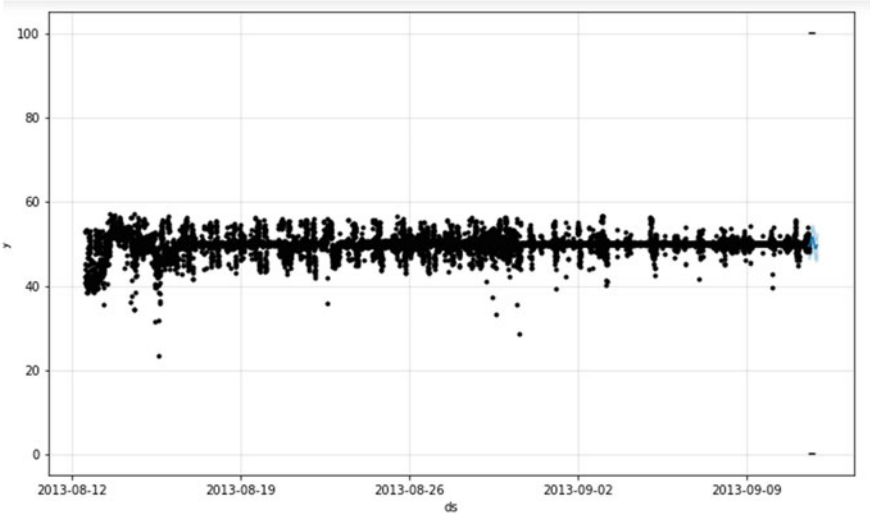


Fig. 3 Future prediction when CPU Usage above 51%

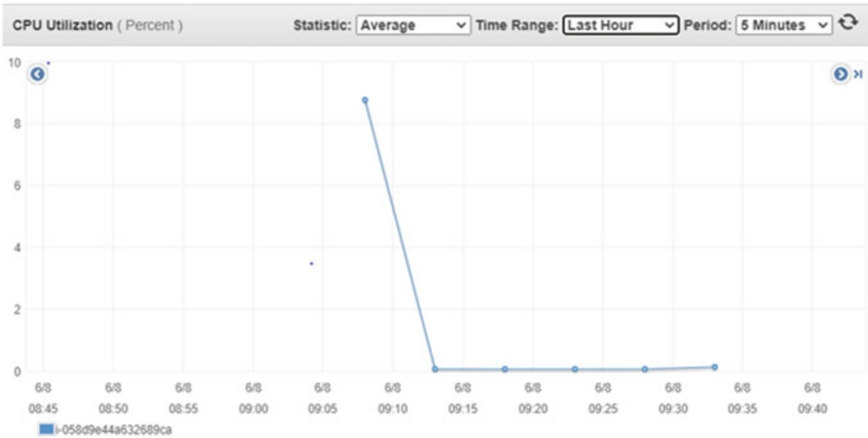


Fig. 4 CPU utilization in terms of percentage vs time interval

6 Conclusion

In this paper, the traces collected from August to September 2013, the trace fastStorage which consists of 1250 virtual machines (VMs). We use the Python, namely: Pandas, Numpy, and Matpoltlib are the modules which are used for analysis and the transformation. In our proposed work, we have converted the historical dataset to be prophet compliant. Implemented AWS auto-scaling for predicting the future workload by applying algorithm for future prediction. Prophet acts as tool, real-world methodology in forecasting that are at scalable.

The main finding of the proposed work is:

(i) When predicted value is greater than 51% of CPU usage for 300 s in next 50 min, then create new EC2 instance.
(ii) We found that only about 9% of CPU utilization is done.

References

1. Papadopoulos AV, Ali-Eldin A, Arzen KE, Tordsson J, Elmroth E (2016) PEAS: A performance evaluation framework for auto-scaling strategies in cloud applications. In: ACM transactions on modeling and performance evaluation of computing systems
2. Ilyushkin A, Ali-Eldin A, Herbst N, Papadopoulos AV, Ghit B, Epema D, Iosup A (2017) An experimental performance evaluation of autoscaling policies for complex workflows. In: Proceedings of the 8th ACM/SPEC on International conference on performance engineering, ICPE'17. L'Aquila, Italy, pp 75–86
3. Bauer A, Herbst N, Kounev S (2017) Design and evaluation of a proactive, application-aware auto-scaler: tutorial paper. In: Proceedings of the 8th ACM/SPEC on International conference on performance engineering, ICPE'17. L'Aquila, Italy, pp 425–428
4. Podolskiy V, Jindal A, Gerndt M (2019) Multilayered autoscaling performance evaluation: can virtual machines and containers co–scale. Int J Appl Math Comput Sci
5. Jindal A, Podolskiy V, Gerndt M (2017) Multilayered cloud applications autoscaling performance estimation. In: IEEE 7th International symposium on cloud and service computing
6. Sean J. Taylor, Benjamin Letham (2017) Forecasting at Scale https://doi.org/10.7287/peerj.preprints.3190v2
7. Taylor SJ, Letham B (2017) Forecasting at scale 10.7287
8. Fang W-X, Lan P-C, Lin W-R (2019) Combine Facebook prophet and LSTM with BPNN forecasting financial markets: Morgan Taiwan Index. In: International symposium on intelligent signal processing and communication systems (ISPACS)
9. Shen S, Van Beek V, Iosup A (2015) Statistical characterization of business-critical workloads hosted in cloud datacenters
10. Guo Y, Stolyar A, Walid A (2018) Online VM auto-scaling algorithms for application hosting in a cloud. IEEE Trans Cloud Comput 1–1
11. Prathanrat P, Polprasert C (2018) Performance prediction of Jupyter notebook in JupyterHub using machine learning. ICIIBMS
12. https://aws.amazon.com/sdk-for-python

A Blockchain-Based COVID-19 Protection Framework

Basundhara Chakrabarty and Harish Krishnamoorthy

Abstract Diseases like COVID-19, SARS-CoV2 and Ebola have led to loss of lives, economic breakdown and mass hysteria. Governments and healthcare workers are on an incessant strive to isolate and curb the community spread of the virus. Several countries have employed mobile-based applications for tracing contacts, distributing COVID protection guidelines, helping people self-diagnose and forming infection patterns and control groups. These existing strategies, however, do not provide end-to-end privacy in the collection and handling of critical patient data. Moreover, they are vulnerable to linkage attacks. In this paper, we introduce a privacy-preserving COVID-19 protection strategy resting on a decentralized, blockchain-based framework, that aids in contact tracing and social distancing. Our simulation validates that our proposed system maintains the confidentiality of the user.

Keywords COVID-19 · Pandemic control · Contact tracing · Surveillance · Blockchain technology · Ethereum

1 Introduction

On December 31st 2019, The Wuhan Municipal Health Commission, China, reported a cluster of pneumonia cases in Hubei Province [1]. The causative agent was soon identified as novel coronavirus (COVID-19). It belonged to a large family of virus that gave rise to symptoms varying from mild flu to severe respiratory distress [2].

The number of COVID cases burgeoned through 2020, and as of today, COVID-19 has affected more than 11 million people in the world and has caused more than 4 lakh deaths [3]. With no promise of a vaccine in the near few months, countries struggle to diagnose the afflicted and identify people who have been exposed to

B. Chakrabarty (✉) · H. Krishnamoorthy
Cisco Systems Inc, Bangalore, India
e-mail: baschakr@cisco.com

H. Krishnamoorthy
e-mail: hkrishn5@cisco.com

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_9

the virus to prevent community spread. A multitude of existing pandemic handling schemes resort to mobile-based contact tracing and mass surveillance applications, like the ones used in Australia, China, Israel, South Korea and Singapore. However, these applications have suffered criticism due to serious privacy threats. Parallelly, nationwide lockdowns have halted economies and given rise to unemployment and recession [4]. The world needs a strategy that improves COVID protection systems, while allowing businesses and corporates to carry out their activity in a sustained manner.

Blockchain is among the most promising technologies of the twenty-first century and has caused a major paradigm shift in the healthcare sector [5]. The paper applies blockchain to COVID protection and delineates three major features of our approach: the contact tracing subsystem, blockchain network and COVID status checking scheme. The contact tracing scheme allows stakeholders to identify people who have come in proximity with individuals diagnosed with COVID-19, to test and isolate them, and treat them when applicable, preserving user privacy while doing so. The COVID status checking scheme allows citizens to stay protected while they carry out their daily activities in a controlled manner. The blockchain network forms the skeleton of the proposed system, stores patient records on a distributed ledger and coordinates access management. Experimental simulation shows the framework to be highly efficient in preserving privacy.

2 Related Work

Several countries have come up with ingenious contact tracing and surveillance methods in the wake of COVID-19. The World Health Organization adopted contact tracing to contain the Ebola outbreak in Africa. WHO defined it as the ‘identification and follow-up of people who may have come into contact with a person infected with Ebola virus’ [6]. It enforced rules for conducting systematic contact tracing for frontline epidemiologists, surveillance officials, healthcare specialists and other volunteers. No mobile applications were used during the Ebola outbreak.

Danquah introduced an application-based contact tracing system to track Ebola cases in Sierra Leone [7] and demonstrated how it improved on data storage and accuracy over paper-based systems.

Several organizations have, thence, engaged in developing contact tracing applications for pandemic control. With the rise of COVID-19, countries have adopted mobile technologies to conduct business in the safest manner possible. The Indian Government conceived the Aarogya Setu app for the same purpose [8]. Aarogya Setu utilizes the smartphone’s Bluetooth and GPS features to determine the user’s infection risk and perform contact tracing. It generates a randomized device identifier (DiD), maps it with the mobile number of a registering user and stores an encrypted version on it. This method gives the government authority access to the user’s epidemiological and GPS information, which poses serious security concerns.

Similarly, China, Singapore South Korea and Australia have urged citizens to install apps for contact tracing and surveillance. However, the concerns of privacy have not been addressed in most of these endeavors. The South Korean surveillance app utilizes GPS data. Furthermore, it requires users to provide their real names and government-issued identity numbers, which infringes on user privacy. Singapore's TraceTogether app operates by exchanging random tokens between nearby phones via Bluetooth, which are then sent to a central server [9]. When an individual is diagnosed with COVID-19, the tokens recorded by the patient's apps are released. Since the government authority maintains a mapping of tokens and associated phone numbers, it can use the released list of tokens to trace the list of exposed users. TraceTogether maintains privacy from other user by using anonymized tokens; however, it provides little privacy for infected citizens. Additionally, the system banks on a central authority (the Singaporean government) and is, hence, less scalable and secure.

Owing to the disadvantages of centralized systems, decentralized COVID protection projects have emerged. Some notable endeavors include COVID watch, PACT and that of Google/Apple [10]. COVID watch is a group of volunteers spread over various continents, comprising security and public health experts [11], and uses Bluetooth Beacon technology for contact tracing. This framework can potentially collect more data than necessary for surveillance and is also susceptible to man-in-the-middle attacks.

Hekmati's work introduces CONTAIN, a privacy-centric mobile contact tracing application that has no dependency on GPS or any other form of location-sensing and reduces the quantity of personally identifiable data logged on a server [12]. The simulation study outlined in the paper emphasizes the efficiency of the system. However, CONTAIN users can choose to reveal their COVID status in an opt-in fashion to the concerned authorities. This stands on the unrealistic assumption that all users shall operate in good faith and shall cooperate and furnish information wherever necessary.

Torky describes a four-tiered blockchain-based COVID containment system [13]. The system adopts the concepts of regex to digitally represent and verify infection patterns under the 'Infection Verifier Subsystem.' A 'blockchain subsystem' acts as a backbone and stores data about confirmed COVID-19 cases in real time, and a peer-to-peer app is used by users to stay abreast of COVID developments. The paper also introduces a mass surveillance subsystem which works in tandem with the infection verifier subsystem, but it provides scarce technical details about the system design and the protocols involved.

3 Blockchain Technology

The groundwork for blockchain technology was laid out by Nakamoto when his whitepaper on bitcoin shook the world of cryptocurrency in 2008 [14].

A blockchain essentially stores a list of transaction records, with each block pointing to the previous block via a hash reference. The block header contains, among other notable fields, the block version, the hash of the parent block, the nonce and a merkle tree root hash [15] of all transactions in the block. The block body contains the transaction list. New blocks are added to the ledger by the block miners after a consensus mechanism, the most common being proof-of-work, which is based on a cryptographic block-racing game [16]. Being a decentralized ledger, blockchains prevent a single point of failure. Furthermore, since each block stores the hash of the previous block, it is computationally infeasible to tamper with blockchains.

4 The Proposed System

In this paper, we propose a blockchain-based framework for protection against COVID that has the following salient features:

- A contact tracing subsystem backed by the Bluetooth beacon technology. The proposed system differs from existing systems in using a decentralized consortium blockchain for storing patient data. Additionally, it uses a highly secure Diffie-Hellman (DH) key exchange algorithm instead of the symmetric algorithms used by existing frameworks.
- A dynamic COVID status record of participating users that need a mandatory daily update, and which can be leveraged by offices, grocery stores, restaurants, etc., to maintain social distancing norms.

The various stakeholders in the system are as follows:

- Medical practitioners, hospitals and testing centers reserve the right to update the blockchain with the COVID status of patients based on their test report. (positive, negative, etc.)
- Research and development centers and government nodes can request to read the epidemiological data in the blockchain for research purposes.
- Shops, restaurants, malls, offices and government personnel maintain read-only access to the blockchain.

4.1 Dapp Registration

Every user, during registration to the Dapp, is assigned a public-private key pair. On registration, the following additional information is collected:

- A collective dataset of the user's epidemiological information ('Info') comprising the age, gender, blood type, pre-existing diseases and the location, encrypted with the user's private key. This ensures that only agencies with the user's public key (e.g., treating medical agencies) can decrypt this information.

Table 1 COVID states and their meaning

COVID state	Meaning
Positive	User shows active strain of COVID-19 in his blood
Negative	User has never been diagnosed with COVID1-9/in contact with a diagnosed patient
Has_to_be_quarantined	User has been in contact with a COVID-19 positive patient and been advised to self-quarantine
Quarantined	User has begun self-quarantine
Off_quarantine	User has served the quarantine period and has not been diagnosed with COVID-19
Recovered	User has been diagnosed with COVID-19 and has recovered

- The COVID-19 status of the patient (status) can take either of the values described in Table 1.

4.2 *The Contact Tracing Subsystem*

Let us assume that Alice and Bob have registered themselves to the Dapp and have received the public keys (A and B) and private keys (a and b), respectively. When they come in proximity, Alice sends her public key to Bob and requests his public key. When Bob accepts and sends over the same, the two users use the Diffie-Helman (DH) algorithm to create a shared secret key. This shared secret is known only to Alice and Bob. Being an asymmetric key exchange algorithm, DH is more secure than the symmetric key exchange used by Bluetooth beacon-based COVID systems (Fig. 1).

Both Alice and Bob then periodically send beacons using Bluetooth. These beacons are pseudorandom numbers encrypted by the DH-shared secret key generated in the previous step. The beacons are encrypted by the user's private key and are stored locally by the receiver.

When Bob is diagnosed with COVID, the certifying medical practitioner interacts with the consortium blockchain to update his COVID status as 'positive,' and at the same time, it downloads the encrypted beacons from his Dapp and updates the list on the consortium blockchain.

All users are required to perform a 'status check' on their Dapps every 24 h. This 'status check' basically interacts with the consortium blockchain, obtains the list of encrypted beacons (from the patients marked COVID positive in that city) and attempts to decrypt the same using the repository of public keys which it has collected. When Alice performs the same, she shall be able to decrypt Bob's beacons (as she has Bob's private key). She is, thus, advised to take the necessary steps to quarantine herself for 21 days. Her COVID status is automatically changed from positive to 'has_to_be_quarantined' and the 'status check' is marked complete. Since

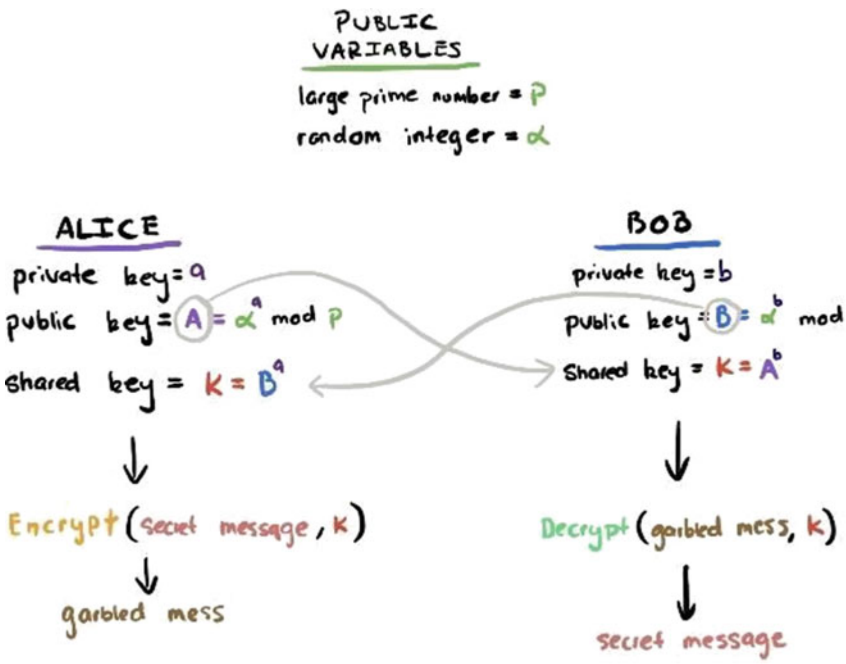


Fig. 1 Diffie-Helman key exchange

the beacons used are anonymized, the identity of the COVID infected patients is not revealed.

4.3 The Blockchain Network

The consortium blockchain stores transactions only on users that have either of the following COVID states: 'positive,' 'has_to_be_quarantined,' 'quarantined,' 'off_quarantine' and 'recovered.' Only medical agencies, testing centers and government agencies certified with a miner ID (ID_p) can add transactions to the blockchain. Table 2 illustrates the fields contained in each transaction in the proposed blockchain.

The epidemiological information of each afflicted/quarantined user is stored in an encrypted format. Government/Research authorities that aim to perform epidemiological surveys are required to request the public key of the transaction creator with a justification, in order to decrypt the data and obtain the content. This ensures that agencies access epidemiological data only when they absolutely need to do so.

Each set of transactions is bundled into a block. The mined blocks are verified by other nodes, spanning healthcare agencies and government bodies, and a consensus is reached via proof-of-work. Since the block verification system is decentralized and involves various stakeholders, linkage attacks are infeasible.

Table 2 Block contents

Field	Value/description
ID	Unique identifier of the patient
H(P)	Cryptographic hash of the user's public key
info	Info = E_p(age, gender, blood type, disease history, location information) where E denotes encryption with key 'p,' the private key of the user
state	The individual's COVID-19 infection state
org_public_key	Public key of the organization who made the last update to the patient's record
updated	A yes/no value indicating whether the beacon decryption check has been performed
allow	If 'yes,' then the user is COVID negative or recovered, and if 'no,' then the user is contagious and is not allowed in shops, offices, etc
public_keys	List of public keys that the user has collected over past 14 days

4.4 *Role of COVID State in Maintaining Social Distancing*

The COVID state of an individual can help people practice social distancing norms. Offices, shopping stores, malls, restaurants, grocery stores, etc., may reserve the right to deny any individual whose COVID state is 'Positive,' 'Quarantined' or 'Has_to_be_quarantined' from entering their premises. Moreover, they may reserve the right to deny access to any individual who has not performed his COVID

'status check' for the day, or whose status check 'updated' parameter reflects as 'no.'.

An example would be the case of Alice, who wants to go to her office. The office can verify Alice's COVID state following these simple steps:

- The office generates a nonce and transmits it to Alice with a request for her COVID state
- Alice calculates the following:

$$Y = E_{\mathrm{a}}\,(\text{nonce}) \tag{1}$$

where a is Alice's private key, and E is an encryption function.

- Alice transmits 'Y' and her public key 'A' to her office authorities
- Her office decrypts 'Y' with the public key and verifies whether it obtains the 'nonce' value
- The office then queries the consortium COVID-19 blockchain for an entry against H(A)
- If the office observes that Alice's entry shows 'updated = Yes' and 'state = negative/off_quarantine/recovered,' it allows her in. In other words, if Alice's entry shows 'allow = yes,' she is allowed in. In all other cases, Alice is denied access.

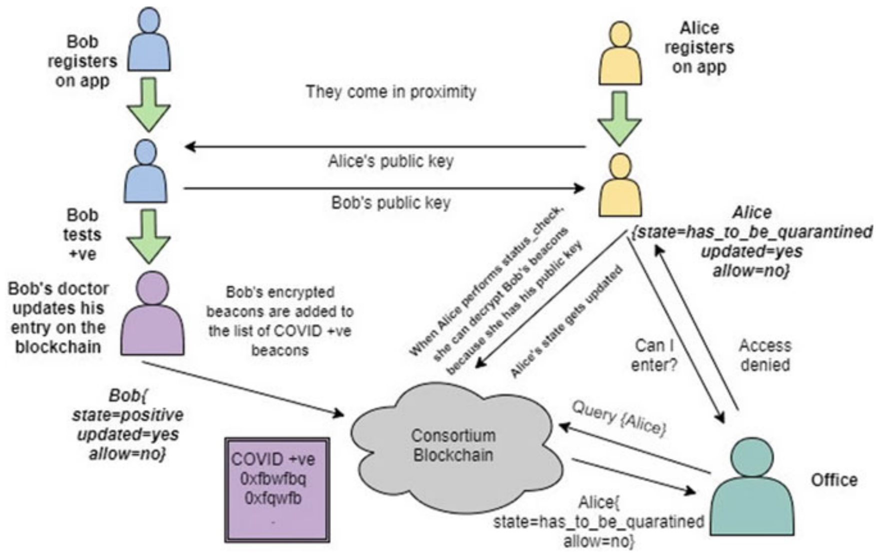


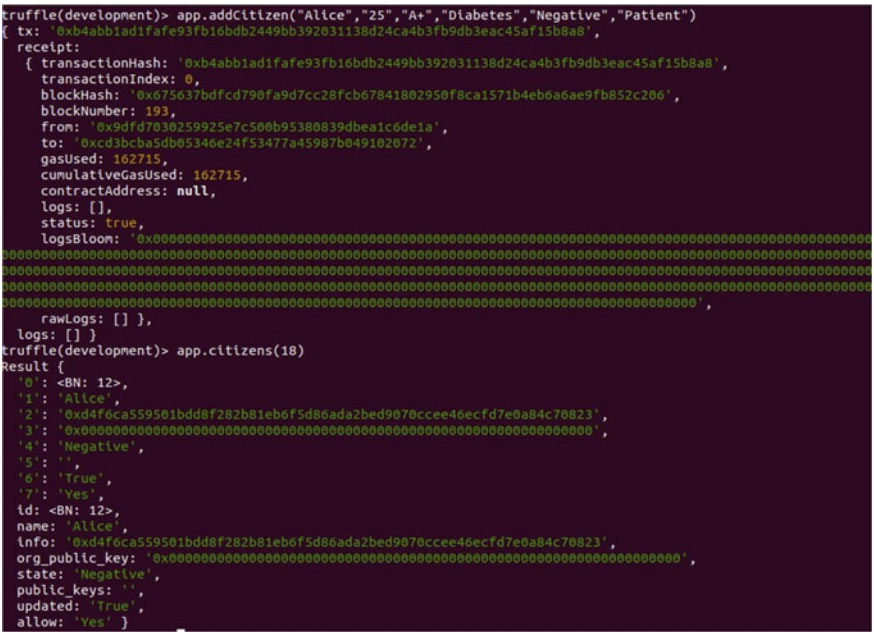
Fig. 2 Schematic representation of the contact tracing and status checking scheme

This implementation allows people to engage in their economic or business activities in a sustained manner, while offering an extra layer of protection against the pandemic in their day-to-day lives. The implementation also facilitates user privacy because an individual can obtain only the COVID state of another individual from the blockchain by using the ID/public key, and for any further information, one must reach out to the miner authority with a proper justification (Fig. 2).

The system provides sufficient incentive to businesses, offices, shopping centers, delivery agencies for using and updating the Dapp because in using the same, they are allowed to carry out their normal operations in a safe manner.

5 Simulation

Ethereum solidity 0.5.16 has been used with Ganache on Ubuntu 18.04 LTS to simulate the smart contract and observe the block structure as maintained and modified by the proposed framework. Truffle framework 5.1.33 is used to interact with the Ganache blockchain testbed. A modular approach has been followed in solidity, with separate functions coded for user registration, status check and for updating the COVID states. Figure 3 shows the usage of the addCitizen() module to implement user registration. The function addCitizen() contains access control checks to verify the role of the invoking user and throws an error if anybody but a 'Patient/Citizen' invokes it. In the smart contract, a total of 30 identities are created, including a doctor, a testing agency and a government official. The entire epidemiological information



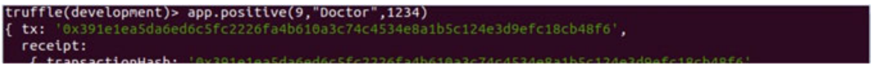
```
truffle(development)> app.addCitizen("Alice","25","A+","Diabetes","Negative","Patient")
{ tx: '0xb4abb1ad1fafe93fb16bdb2449bb392031138d24ca4b3fb9db3eac45af15b8a8',
  receipt:
   { transactionHash: '0xb4abb1ad1fafe93fb16bdb2449bb392031138d24ca4b3fb9db3eac45af15b8a8',
     transactionIndex: 0,
     blockHash: '0x675637bdfcd790fa9d7cc28fcb67841802950f8ca1571b4eb6a6ae9fb852c206',
     blockNumber: 193,
     from: '0x9dfd7030259925e7c500b95380839dbea1c6de1a',
     to: '0xcd3bcba5db05346e24f53477a45987b049102072',
     gasUsed: 162715,
     cumulativeGasUsed: 162715,
     contractAddress: null,
     logs: [],
     status: true,
     logsBloom: '0x000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000',
     rawLogs: [] },
  logs: [] }
truffle(development)> app.citizens(18)
Result {
  '0': <BN: 12>,
  '1': 'Alice',
  '2': '0xd4f6ca559501bdd8f282b81eb6f5d86ada2bed9070ccee46ecfd7e0a84c70823',
  '3': '0x0000000000000000000000000000000000000000000000000000000000000000',
  '4': 'Negative',
  '5': '',
  '6': 'True',
  '7': 'Yes',
  id: <BN: 12>,
  name: 'Alice',
  info: '0xd4f6ca559501bdd8f282b81eb6f5d86ada2bed9070ccee46ecfd7e0a84c70823',
  org_public_key: '0x0000000000000000000000000000000000000000000000000000000000000000',
  state: 'Negative',
  public_keys: '',
  updated: 'True',
  allow: 'Yes' }
```

Fig. 3 User registration

of the user stored in an encrypted format. Please note that a separate variable is created to store the plaintext 'name' only for understanding purposes on this paper. In the actual framework, the 'name' shall also be encrypted within 'info.'

Similarly, Bob registers himself to the blockchain. When he tests positive, his doctor invokes the positive() function to update his COVID state to 'positive.' This triggers the status_check() function, which sets the 'allow' value to 'no.' Figure 4 shows the status update, and Fig. 5 shows the block state after the update. Note that the doctor's public key is now stored in 'org_public_key.'

Alice has maintained the list of public keys obtained from users in the last 14 days, including Bob's key. When she does her daily status update, the citizen_update() function is triggered (Fig. 6), and her status is updated on the blockchain, to 'state=has_to_be_quarantined' and 'allow=no.' Hereafter, both Alice and Bob are not allowed in offices, shops, etc., (Fig. 7).



```
truffle(development)> app.positive(9,"Doctor",1234)
{ tx: '0x391e1ea5da6ed6c5fc2226fa4b610a3c74c4534e8a1b5c124e3d9efc18cb48f6',
  receipt:
```

Fig. 4 Doctor updates Bob as COVID 'positive'

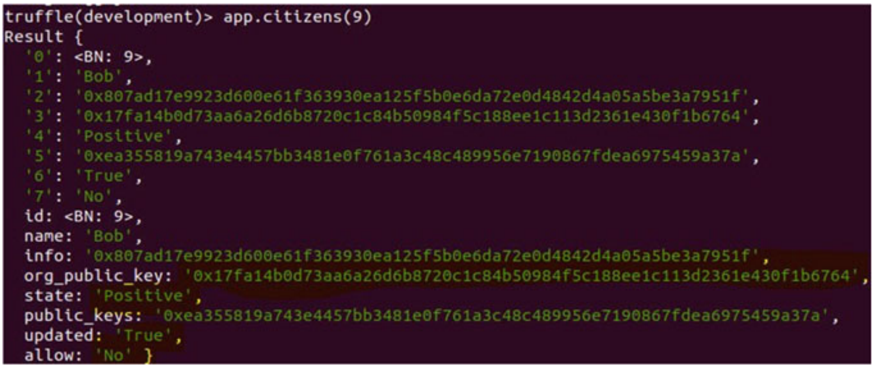


Fig. 5 Block structure (Bob) after testing positive

Fig. 6 Alice performs her daily status update

Fig. 7 Alice's block structure after being found to have been in contact with Bob

6 Conclusion and Future Work

The staggering increase of COVID-19 cases has caused a major healthcare crisis, and a blockchain-based approach goes a long way in solving it. The proposed framework and accompanying simulations show that user's personal or epidemiological information is never dispelled to any stakeholder without justified reason. Additionally, the blockchain backbone maintains the reliability and immutability of user data. Our work can be extrapolated by designing a more layered access control mechanism roping in bed/ambulance availability checks to bridge the gap between the user, hospital and the government. Moreover, a JavaScript interface can be designed to ease the user interaction, and payment mechanisms via cryptocurrencies can be incorporated.

References

1. WHO COVID Timeline. https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19
2. Zheng Y, Ma Y, Zhang J et al (2020) COVID-19 and the cardiovascular system. Nat Rev Cardiol 17:259–260
3. WHO Situation Report-158. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200626-covid-19-sitrep-158.pdf?sfvrsn=1d1aae8a_2
4. Fernandes N (2020) Economic effects of coronavirus outbreak (COVID-19) on the World Economy (March 22, 2020). Available at SSRN: https://ssrn.com/abstract=3557504
5. Gordon C (2018) Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. Comput Struct Biotechnol J 16:224–230
6. Olu O, Kargbo B, Kamara S et al (2015) Epidemiology of Ebola virus disease transmission among health care workers in Sierra Leone, May to December 2014: a retrospective descriptive study. BMC Infect Dis 15:416. https://doi.org/10.1186/s12879-015-1166-7
7. Danquah et al (2019) Use of a mobile application for Ebola contact tracing and monitoring in northern Sierra Leone: a proof-of-concept study. BMC Infect Dis 19:810. https://doi.org/10.1186/s12879-019-4354-z
8. Aarogya Setu. https://www.mygov.in/aarogya-setu-app/
9. Cho I, Yu (2020) Contact tracing mobile apps for COVID-19: privacy considerations and related trade-offs, cryptography and security
10. Newsroom. https://www.apple.com/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/
11. COVID Watch. https://covid-watch.org/
12. Hekmati, Ramachandran, Krishnamachari (2020) CONTAIN: privacy-oriented contact tracing protocols for epidemics, cryptography and security
13. Torky H (2020) COVID-19 Blockchain framework: innovative approach, cryptography and security
14. Nakamoto Bitcoin Whitepaper. https://bitcoin.org/bitcoin.pdf
15. Szydlo M. (2004) Merkle tree traversal in log space and time. In: Cachin C, Camenisch JL (eds.) Advances in cryptology—Eurocrypt 2004. Lecture notes in computer science, vol 3027. Springer, Berlin, Heidelberg
16. Wang W, Hoang DT, Xiong Z, Niyato D, Wang P, Hu P, Wen Y (2018) A survey on consensus mechanisms and mining management in blockchain networks. arXiv preprint arXiv:1805.02707, 1–33

A State of Art of Machine Learning Algorithms Applied Over Language Identification and Speech Recognition Models

A. Razia Sulthana and Aakansha Mathur

Abstract Over the years, we have observed several applications of machine learning. One of the applications is in speech recognition. One of the sub-topics in speech recognition is language identification. A typical language identification system involves a feature extraction and classification stage. Machine learning classifiers have been extensively used in the classification stage of the LID system. The current paper examines the chronological involvement of machine learning in the LID system. The paper also details the chronological development and refinement of the LID system.

Keywords Machine learning · Speech recognition · Support vector machines · Classification · Language

1 Introduction

The objective of a language identification (LID) system is to determine the correct language in the audio speech signal. A LID system generally consists of two stages, feature extraction stage and classification stage. The feature extraction stage also consists of some pre-processing steps such as pre-emphasis, framing and windowing. The feature extraction stage involves different types of speech features. The speech features are broadly classified in high-level features and low-level features. Acoustic, phonotactic and prosodic features are low-level features. Lexical and syntactic features are high-level features. The classification stage involves using different machine learning classifiers like support vector machine (SVM), Gaussian mixture models (GMM) and K-nearest neighbors.

Most researchers have proposed LID systems with low-level feature extraction techniques. An additional stage, normalization, has been applied to refine the LID

A. Razia Sulthana · A. Mathur (✉)
BITS Pilani Dubai Campus, Dubai, United Arab Emirates

A. Razia Sulthana
e-mail: razia@dubai.bits-pilani.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_10

system results. Other refinements to the LID system include splitting the classification stage into the learning and recognition stage. The current paper looks at chronological development of research in language identification in recent years. The next six sections explain a few of the LID models and finally the conclusion.

2 Literature Review

2.1 Language Identification System—I

The research proposal in [1] develops an automatic language identification system that classifies languages into modern standard Arabic (MSA) and Kabyle, two commonly spoken languages in Algeria. The research uses two different databases. One of the databases is in MSA, and the other database is in Kabyle dialect. The databases do not have an equal number of samples. All the speakers were Algerian, so they were fluent in MSA and Kabyle. Moreover, the speakers spoke the sentences with dialect. The research creates a bilingual database by taking 10 sentences from MSA. The MSA 10 sentences were spoken by 13 different speakers and were repeated 10 times. This makes the total number of utterances to be 1300. The sampling frequency of each MSA sentence is 44,100 Hz, and each sentence is coded in 16 bits. The sentences selected were phonetically balanced. This was done so that the greatest number of phonemes can be obtained without varying the frequency of each sentence. For the Kabyle database, the research took dialogs spoken by four male and two female speakers. And, these dialogs were repeated five times. This makes the total number of Kabyle utterances to be 720. Moreover, the Kabyle utterances were sampled at a frequency of 16,000 Hz. The Kabyle utterances, like MSA utterances, were coded in 16 bits. The system proposed by the research involves two phases: the learning phase and recognition phase. Both the phases have the same steps. The only difference is that in the learning phase, the language of the speech sentence sample is known while in the recognition phase, the language of the sentence is unknown. The steps for both the phases are as follows: pre-processing, feature extraction, Arabic and Kabyle Language Model, decision and identify language. The research has extracted two types of low-level features: prosodic features and acoustic features. Melody and stress are two types of prosodic features extracted by the research. The research extracted MFCC acoustic features. A melody is made up of high- and low-pitch notes. Melody is a prosodic feature which represents the speech signal frequency in function of time. Melody is a speaker-dependent feature and is related to the speaker's emotions. The research used PRAAT software to extract the melody feature. Stress is also a prosodic feature which represents the speech signal's intensity. The research used support vector machines as classifiers for its language identification system. For the learning phase, the research took 16 Kabyle sentences from 24 Kabyle sentences and 6 MSA sentences from 10 MSA sentences. The research first used prosodic features to classify unknown audio signals into one of the two languages.

2.2 *Language Identification System—II*

A learning model is proposed in [2] for a language identification system with optimized machine learning algorithms. The input signals are converted into frames of length 25 s. These frames had an overlap of 20 s. Following which vocal tract length normalization (VTLN) is applied on the input speech signal to obtain seven mel-frequency cepstral coefficients (MFCC). The research proceeded by performing cepstral mean and variance normalization along with representations relative spectra RASTA filtering. Next, the shifted delta cepstral (SDC) features are calculated.

The extreme learning machine (ELM) approach involves calculating the output weights of a neural network using least-squares solutions. The technique involves generating biases randomly for hidden layer weights. A disadvantage of ELM, as mentioned by the research, is that it does not have a specific approach for determining the input-hidden layer weights. Therefore, the ELM becomes subject to local minima. So, optimal weights must be identified. The identification of optimal weights can be done by using an optimization approach. This will ensure that ELM gives its best performance. Further, the ameliorated teaching–learning-based optimization (ATLBO) algorithm is applied. For the language identification system, the research worked with eight different languages. The research obtained audio files for the eight languages from broadcasting channels. The dataset consisted of 120 records for the 8 languages. And, the total number of features extracted was 600. There were 15 utterances for each language, and each utterance was of 30 s length. The dataset was split into training and testing dataset. The training set was 67% of the entire dataset. The testing set was 33% of the entire dataset. The research selected enhanced ameliorated teaching–learning-based optimization (EATLBO) and enhanced self-adjusting extreme learning machine (ESA-ELM) as optimization approaches for the language identification system. The research evaluated EATLBO by comparing it with ATLBO. Each experiment involved applying a different mathematical function for optimization.

2.3 *Language Identification System—III*

The objective of this proposal [3] was to build a learning model for language identification systems. The study used an extreme learning machine (ELM) as a learning model for the language identification system. The first step was to extract the features. A drawback of ELM is that it gets subjected to local minima due to the absence of a specific approach to determine input-hidden layer weights. The drawback can be overcome by using optimization approaches which will give optimal weights for the ELM. The research used genetic algorithms for optimization tasks. The purpose of the genetic algorithms is to discover the optimal hidden layer weights and biases. A genetic algorithm involves: initial population, evaluation, selection and genetic operators. The genetic algorithm will enhance the performance of the ELM. Generally, only one criteria are used by genetic algorithms during the selection process for optimization. Due to the shortcomings of genetic algorithms, the research optimized the genetic algorithm and proposed a new optimized genetic algorithm (OGA). The optimized genetic algorithm (OGA) has three selection criteria: roulette wheel, K-tournament and random. These criteria are used for parent selection. The parents are selected for generation of two new offspring for a new population in the crossover operation. The OGA is now used to enhance the performance of the ELM. The OGA-ELM has the three different selection criteria: selection, crossover and mutation process which optimizes the input weight values and hidden nodes bias. The parameter settings for ELM are as follows: The number of hidden nodes is 700–900, the output neurons are class values assigned to each language, the activation function applied is sigmoid activation function, the input weights range between –1 and 1, and the bias values range 0–1. The parameter settings for OGA are as follows: The number of iterations is 100, the population size is 50, the crossover operation is arithmetic operation, the mutation operation is uniform mutation, the crossover population is 70% of the population, the mutation population is 30% of the population, the gamma value is 0.4, and the tournament size is 3. The research worked with eight different languages. The class values for each language are as follows: 1 is for Arabic, 2 is for English, 3 is for Malay, 4 is for French, 5 is for Spanish, 6 is for German, 7 is for Persian, and 8 is for Urdu. The research architecture performs the following steps. The first step is random initialization of input weights and bias, determining the population size, objective function and maximum iteration. The third step is tabulating the fitness value. The fourth step is initialization of the crossover and mutation populations. The fifth step involves two parts. The first part is if the crossover population is less than or equal to 70%, then parents are selected based on the selected criteria. After which, crossover is performed, and the two offspring generated are saved in the new crossover population. The new crossover population is sent to the first part of the fifth step. Suppose the crossover population is greater than 70%, then the algorithm goes to the second part of the fifth step. The second part itself has if–then condition. If the population is less than or equal to 30%, then select parents randomly and perform mutation. The child created is added to the mutation population. After which merge the mutation and crossover population. However, if the mutation population is greater than 30%, then simply merge the crossover and mutation population. After merging, the algorithm checks if the termination criteria are satisfied. If it is not satisfied, then the algorithm goes back to calculation of the fitness value step (third step). However, if the termination criteria are satisfied, then the global optimum parameters are discovered. This marks the end of the OGA-ELM algorithm.

2.4 Language Identification System—IV

The research proposed [4] a system that performs three tasks: speaker identification, gender identification and language identification. The system classifies speakers into eight categories, languages into three categories and gender into two categories. The authors created their own database. The database primarily consists of speech recordings in Arabic, English and Polish. The speech recordings were taken from online and TV broadcasts. There were eight speakers for each language. So, in total, the dataset consisted of 24 speech recordings. The speakers were male and female. The database consisted of talks of 2 h 20 min and 27 s. The speech recordings had a frequency of 44 kHz. Firstly, the speech recordings are pre-processed. The pre-processing step involved eliminating the silence in the recordings. Moreover, low-energy fragments are eliminated from the input file. After pre-processing, the input file time length was reduced to 2 h 7 min and 22 s. The research used radial neural network (RNN), probabilistic neural network (PNN) and the used long short-term memory recursive neural network LSTM neural network (NN) architecture for classification. The LSTM NN architecture consists of six layers. The input layer consists of input feature vectors. The length of the input features depends upon the features extracted. The research extracted Burg's estimation, TM-eigenvalues and MFCC features. The feature vector length of Burg's estimation is 129. The feature vector length of TM-eigenvalues is 125. The feature vector length of MFCC is 23. After the input layer, the LSTM NN consists of 24 hidden units. After the hidden layer, there is a dropout layer with dropout possibility of 50%. Then, there is a fully connected layer (fc) consisting of 24 neurons. Following the fc layer, there is a SoftMax layer. The purpose of the SoftMax layer is to apply SoftMax function to the input. Finally, there is an output layer which calculates the cross-entropy loss function. The number of neurons in LSTM and output layer has been reduced for language and gender identification tasks because of the small number of final categories. The dataset was distributed in the following way. Firstly, the dataset was split into Arabic, English, Polish, Male and Female subset. Each subset was further split into a training set and testing set. Furthermore, 30 s of each speaker was taken. So, the Arabic training set consisted of 30 speakers. This totaled to 4 min for Arabic training set. For the Arabic testing set, the remaining speech signal is taken. This process is repeated for English, Polish, Male and Female subsets. Hence, the training data was 9% of the complete database. While the testing data was 91% of the complete database, the training–testing ratio was 9:91. These features were inputted into probabilistic NN and radial NN for speaker, language and gender identification.

2.5 Language Identification System—V

A two-stage language identification system (TS-LID) [5] is proposed for Indian languages. The research used the two databases for experimentation. The first database consists of telephone speech 11 languages. The telephone speech is at a frequency of 8 kHz. The speech includes 90 utterances from 90 different speakers per language. Out of the 11 languages, only 2 are tonal languages. The remaining languages are non-tonal languages. The second database consists of studio quality 12 Indian languages. There were five tonal languages and seven non-tonal languages. The research has analyzed the performance of MHEC and MFCC features individually and in combination. The first step in the classification stage involves a tonal/non-tonal pre-classification. Here, i-vector-based SVM will be used as classifiers. The second step involves language identification. The architecture proposed by the research is subjected to three conditions. The first condition is a conventional language identification system where each language has a model, and the model is trained. So, during the identification stage, it is determined which language model is the most likely for a given test sample. The second condition involves the languages to be first pre-classified into tonal and non-tonal. The pre-classification step is the first step. Depending upon the outcome of the first step, irrespective of its correctness, the speech signal is routed to either the tonal or non-tonal module of the second step. The third condition also involves a pre-classification step but only those samples which are correctly classified into tonal or non-tonal are forwarded to the next step, which is the language identification step.

2.6 Language Identification System—VI

The research [6] proposed a language identification system that recognizes four languages: Tamil, Malayalam, Hindi and English. In fact, a number of recommendation systems [7–9] are supportive to be modified into speech recognition systems; in future, 50 utterances of each language are present in the dataset. Moreover, the dataset consists of equal numbers of male and female speakers. The total number of male speakers for each language is five. The total number of female speakers for each speaker is five. Studio quality acoustic systems are used to record the speeches by these speakers. The research used Sound Forge software for isolation of words and file organization. The audio recordings are saved in Mono Wav File Format. The sampling frequency of the speech recordings is 16 kHz. Once the words are isolated from each speech, a total of 200 input speech signals is obtained. This is primarily the pre-processing step. The speech signals now undergo pre-emphasis. The next step is the feature extraction step. The research extracts several features and experiments upon the combination of the features as well. The features extracted by research are MFCC, perceptual linear prediction features (PLP) relative perceptual linear prediction features (RASTA-PLP) and shifted delta cepstrum (SDC). 40 filters filter bank is used in MFCC. Of these 40 filters, 13 are linear filters and 27 are logarithmic filters. For each isolated word, a thirteen-coefficient matrix is obtained. The next step is the classification step. The research uses feed-forward back-propagation neural networks (FFBPNN) as a classifier. The features extracted from the earlier step are stored as feature matrices. The learning algorithms used are trainlm and trainscg. The objective

Table 1 Comparative analysis of speech features and machine learning method used by different papers

References	Speech features	Machine learning
[1]	MFCC, melody and stress	SVM
[2]	MFCC, cepstral mean, variance normalization along with RASTA filtering, SDC	Extreme learning machine
[3]	MFCC, SDC, cepstral mean and variance normalization with RASTA filtering	Extreme learning machine
[4]	MFCC	Radial neural network (RNN), probabilistic neural network (PNN) and LSTM
[5]	MFCC, MHEC	SVM
[6]	MFCC, PLP, SDC	Feed-forward backpropagation neural networks

of using two different learning algorithms is to identify which learning algorithm is the best to train the neural network. Moreover, nonlinear sigmoid activation function was used and softmax activation function. To evaluate the performance of FFBPNN, the research obtains error and accuracy rate for training, testing and validation sets. The research uses 100 hidden neurons in FFBPNN when the "trainlm" algorithm is used. The research uses 30 hidden neurons in FFBPNN when the "trainscg" algorithm is used.

3 Comparative Analysis

From Table 1, we can observe that MFCC features are extracted by all the papers. This is because MFCC features give better performance for the various LID systems built by different researchers.

4 Results Analysis

This section primarily discusses the results obtained by different researchers. Moreover, the limitation of the research is also addressed.

The melody features when classified using SVM by [1], the accuracy for MSA and Kabyle was 90.41 and 82.08, respectively. The average accuracy for melody features was 86.25. When stress features were used for language identification, the accuracy for MSA and Kabyle was 98.75 and 91.67, respectively. The average accuracy for melody features was 95.2. The research then combined the melody and stress features. When the combined features were used for language identification, the accuracy for MSA and Kabyle was 99.17 and 91.67, respectively. The average accuracy for combined features was 95.42. 238 MSA samples were correctly classified into MSA language when combined features were used. 220 Kabyle samples were correctly classified into Kabyle language when combined features were used. 2 MSA samples were incorrectly classified into Kabyle language when combined features were used. 20 Kabyle samples were incorrectly classified into MSA language when combined features were used. The research then proceeded by using only MFCC features for language identification. The research obtained an accuracy of 97.91 for MSA language and 93.75 for Kabyle language. The average accuracy was 95.83. 235 MSA samples were correctly classified into MSA language when MFCC features were used. 225 Kabyle samples were correctly classified into Kabyle language when MFCC features were used. 5 MSA samples were incorrectly classified into Kabyle language when MFCC features were used. 15 Kabyle samples were incorrectly classified into MSA language when MFCC features were used. The research then combined melody, stress and MFCC features. The research obtained an accuracy of 98.75 for MSA language and 96.25 for Kabyle language. The average accuracy was 97.5. The research reinstated the fact that hybridization of acoustic and prosodic features increased the accuracy of language identification. The research used only two types of prosodic characteristics, melody and stress. Moreover, the research did not attempt to optimize the learning phase.

Albadr et al. [2] observed that EATLBO performs better than ALTBO. The research also evaluated ESA-ELM on the basis of multiple parameters of the learning model. The highest accuracy for ESA-ELM was 96.25%. The research inferred that ESA-ELM is better suited for language identification systems. The research also used the elitist genetic algorithm to enhance extreme learning machines. This approach was called the elitist genetic algorithm extreme learning machine (EGA-ELM). The highest accuracy for EGA-ELM was obtained when 750 hidden neurons were used. The lowest accuracy for EGA-ELM was obtained when 900 hidden neurons were used. However, the accuracy of ESA-ELM was higher than the accuracy of EGA-ELM for all iterations. Thus, the research reinstated the fact that ESA-ELM is better suited for language identification systems. The research did not experiment on time and cost optimization of front-end feature extraction. The research did not explore the use of metaheuristic algorithms in ESA-ELM for optimization of weights.

The OGA-ELM achieved 100% accuracy for K-tournament, 99.50% accuracy for roulette wheel and 99.38% for random. The research does not consider real-time aspects such as noise that could affect the performance of the model proposed by [3].

The accuracy or recognition rate for the nine cases (3 features $\times$ 3 identifications) was tabulated by [4]. The research deduced that probabilistic NN works better with longer speech signals. The LSTM NN performed well for all three identification cases. The research took a simple dataset, and the performance of LSTM NN was not observed on complex datasets.

Bhanja et al. [5] noted that the pre-classification step gave better performance for OGI-multilingual database than NIT Silchar language database (NITS-LD). That is, the language identification accuracy was higher for OGI-MLTS than that of NITS-LD when the pre-classification step was applied. The research did not study the use of modified neural networks to help reduce processing time.

Deshwal et al. [6] deduced that highest FFBPNN accuracy and lowest testing error are achieved (0.10). The research also varied the number of epochs to observe the performance of the neural network. The epochs range from 30 to 60. It was observed that as the number of epochs increased, the classification accuracy also increased. It was also inferred that the "trainlm" algorithm gives better performance than "trainscg" algorithm. A limitation and future expansion of the research are using the proposed language identification system for continuous speech signals and using more training functions.

5 Conclusion

The paper examined the contribution of machine learning in LID systems from 2016 to 2020. It can be observed that the contribution of machine learning has increased in LID systems over the recent years. In this paper, we have seen how different machine learning classifiers such as SVM, neural networks and decision trees have been used in the classification stage of the LID system. Moreover, we have also seen how using a machine learning classifier in a pre-classification stage can improve the accuracy of the model. Some researchers have optimized the training parameters and observed the performance of the LID system. Further research can involve using continuous speech signals in the LID system.

References

1. Lounnas K, Demri L, Falek L, Teffahi H (2018, October). Automatic language identification for Berber and Arabic languages using prosodic features. In: 2018 International conference on electrical sciences and technologies in Maghreb (CISTEM). IEEE, pp 1–4
2. Albadr MAA, Tiun S, AL-Dhief FT, Sammour MA (2018) Spoken language identification based on the enhanced self-adjusting extreme learning machine approach. PLoS ONE 13(4)
3. Albadr MAA, Tiun S, Ayob M, AL-Dhief FT (2019) Spoken language identification based on optimised genetic algorithm–extreme learning machine approach. Int J Speech Technol 22(3):711–727
4. Nammous MK, Saeed K (2019) Natural language processing: speaker, language, and gender identification with LSTM. In: Advanced computing and systems for security. Springer, Singapore, pp 143–156
5. Bhanja CC, Laskar MA, Laskar RH, Bandyopadhyay S (2019) Deep neural network based two-stage Indian language identification system using glottal closure instants as anchor points. J King Saud Univ-Comput Inf Sci
6. Deshwal D, Sangwan P, Kumar D (2020) A language identification system using hybrid features and back-propagation neural network. Appl Acoust 164:107289

7. Sulthana AR, Gupta M, Subramanian S, Mirza S (2020) Improvising the performance of image-based recommendation system using convolution neural networks and deep learning. Soft Comput 1–14
8. Sulthana AR, Ramasamy S (2019) Ontology and context based recommendation system using neuro-fuzzy classification. Comput Electr Eng 74:498–510
9. Sulthana R, Ramasamy S (2017) Context based classification of Reviews using association rule mining, fuzzy logics and ontology. Bull Electr Eng Inf 6(3):250–255

Feedback-Based Real—Time Surveillance for MidDay Meal Scheme

Kartik Joshi and Shyam Vadalia

Abstract There is a growing demand of a proper supervision of the midday meal scheme that is enforced in India. Last decade statistics and news articles tend to suggest that there are many loopholes in the current system that needs immediate attention to help effective enforcement of the scheme at the grassroots level. Our proposed system provides a holistic approach towards supervision of the scheme. This system, when implemented in compliance with the Government regulations, can help the concerned authorities in monitoring the execution of the midday meal scheme eliminating the scope of error. Our proposed system is capable of performing real-time food detection, counting number of students through state-of-the-art surveillance cameras with machine learning capabilities. The application is trained on Indian food that can be deployed in Indian School. In case of discrepancy, the user can report discrepancy with a thorough follow-up mechanism that enables authorities to take quick action against the reported complaint. The proposed system enhances the current monitoring process. The system is robust to the changes in the menu. This helps in monitoring process supervised by the government authority. This standalone application can help the governing authorities to have an all-round inclusive approach and magnify the supervision process under a single platform.

Keywords Midday meal scheme · Image classification · Convolutional neural network · Video surveillance · Chatbot support

1 Introduction

Midday meal is a visionary scheme that initiates the fight to combat malnutrition among school going students. Students get free meals in the school premises. This

K. Joshi (✉) · S. Vadalia
School of Electronics Engineering, Vellore Institute of Technology, Chennai, India
e-mail: nckd.joshi@gmail.com

S. Vadalia
e-mail: skvadalia2014@gmail.com

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_11

has helped India keep a check on the malnutrition of children. Looking back at the journey from its inception, the scheme had its fair share of recognition and criticism. While there are lots of organizations such as Akshaya Patra, Anna Amrita Foundation are trying to provide exemplary service to schools, and there are some schools in parts of India which still find it difficult to effectively implement the scheme as per the government regulations. According to a study of nutritional health status of primary school children (7–9 years of age) in Bareilly district [10], it is found that the children belonging to the rural area are suffering more as compared to that of in the urban area. Also, a recent study on the impact of midday meal scheme on the nutritional status of rural middle school children (9–5 years of age) [12] shows that nutritional intake of children was lower than recommended dietary allowance of Indian Council of Medical Research (ICMR). According to that study, the midday meal food was low on calcium, iron and vitamin A (Fig. 1).

Unhygienic food and inadequate supply of food against the estimated demand are major issues faced. The unhygienic food simply compromises with the goal and principle of the scheme. And inadequate food prevents the scheme to scale up to appreciable level to eradicate malnutrition. These issues need to be addressed properly in order to make the scheme a success.

If we look at the scope of error causing these pitfalls, it is clear that there is a lot of potential for improvement. Unhygienic food could also result in food losing its nutritional value. Our proposed method uses third party Application program interface for detecting the food and obtaining the calorie count of the food from the image of the food. The food detection algorithm is trained on Indian food. This can

Fig. 1 School students having food under midday meal scheme

be matched with the nutritional chart provided by the government in compliance with the Food safety and standards authority of India. This can help in comparing the food prescribes against the meals that are actually delivered. The lapse can be evident from the camera infrastructure made available.

The lapse in valid information is yet another aspect that is contributing to the failure of the scheme. The regular count of the number of students is very important to maintain the availability of meals for all the students. The camera capabilities have improved over the course of time, and the deep learning boom has helped in tackling this issue. Installation of cameras can help keep a track of the students and help the authorities to make a sound decision on the stock of the food. The figures reported now are a result of mere manual observation which is prone to error. This loophole makes the scheme vulnerable to fail. We present a system to count the number of students availing the meal everyday using the camera and machine learning algorithms. This data could be then forwarded to the authorities, so that the gap between the demand and supply is reduced.

The surveillance system is also enabled with a feedback mechanism, backed by Chatbot support. Any discrepancy in the delivered meal, either qualitative or quantitative, can be reported to the authority through Chatbot. This approach makes the system fast and responsive to the needs at the grassroots level. This attempt makes the system flawless and accounts for the responsibility of the officials. This platform could also be used to address everything related to the scheme and communicate with the people.

2 Related Work

2.1 People Counting Applications

Counting number of people in the surrounding area has been interesting problem for study in the field of computer vision. Dittrich introduced the idea of array of cameras in people counting in crowded and outdoor scenes using a hybrid multi-camera approach [5]. This could be very useful in crowded places like schools. Another approach proposed by Deepak Babu Sam [4] was using modular paradigm, detecting head in the densely crowded surrounding.

But the hybrid multi-camera approach is not possible in the current scenario due to the poor camera infrastructure of the Indian government school. Their approach is unable to distinguish students availing the scheme from others. Drawing inspiration from their work, we intend to deploy the crowd counting system customized for school. Considering the characteristic feature of people availing midday meals, we extract the visual leads such as uniform, identity card to leverage the security aspect. Such system also helps in keeping a track of people on a regular basis. This is a crucial module in effective functioning of the surveillance system enabling the maintenance of the meal count for availability of sufficient stock of meals. Shortage of meals is one major contributor of failure of the scheme and resulting in malnutrition.

2.2 Nutritional Value Supervision

In recent years, there has been increasing concern over obesity, and as a consequence, we see a surge in development of calorie counting systems. With more and more datasets related to the food images being generated, there has been advancement in the food detection and processing algorithms. Initial works have suggested for having nutritional fact tables populated corresponding to the food item. Pouladzadeh et al. [11] are based on this idea. There have been many third party applications offering the service of daily calorie count based on image upload by the user. ClarifAI [7], Calorie Mama are few of the products that we have for reference.

In [8], Chokr presents a three-phase pipeline at approaching to this problem. This is an improvement to the previous work as it involves a supervised method for arriving at the nutritional count of the food. Our implementation combines these techniques and banks on the application programming interface to get the calorific value of the food being offered at the school as a part of the midday meal scheme. The challenge to this approach is to customize it for the Indian food. This is tackled by training on several images of each of the meals and mapping it to the nutritional fact table of the same. This is very crucial for the system to perform well. This will help in verifying the nutritional value of the meals being served.

2.3 Feedback Mechanism Using Chatbot Support

In this day and age, availability is quite essential. All time availability of staff in government organizations could be a costly affair. So, Chatbot is finding its significance in these scenarios. The idea is to collect information. There have been several platforms and frameworks explored by Kar and Haldar in [13] that also reflected the commercial success of such platforms.

We explore a couple of such platforms Snatchbot [14] and Collect.chat [3] and compare their performances in case of deployment. Chatbot provides a robust solution to make the system all time available and introduces a feedback mechanism. We implement the design suggested by Adam et al. [9] in building a human centric conversational artificial intelligence agent that helps the user report discrepancy in the midday meal scheme, schedule an appointment with the concerned authorities.

3 Simulation

We present a single digital platform to solve the midday meal scheme management issues. The proposed system (1) identifies people in a densely populated hall, (2) counts the number of people, (3) computes the nutritional information of the meals being offered and (4) provides a feedback mechanism using Chatbot platform to make the system flawless.

Multiple cameras embedded in the school hall can help in recreating the exact scenario giving more accurate details. As proposed by [2], the system has multiple camera feeds. The cameras are responsible for tracking the number of people present in the hall and counting the number of calories present in the meals being offered. The website is maintained with the weekly report being generated. These can be monitored by the government officials, eliminating the scope of error. In case of any discrepancy related to the food or the services, the concerned authorities could be informed using Chatbot platform.

The overview of our proposed system is reflected in Fig. 2

3.1 People Tracking and Facial Recognition

We use the one shot detector for facial recognition which uses triplet loss function to compare positive, negative and anchor image. Results from [1] are encouraging to bank upon triplet loss as the optimizer. Facial features of individual students are unique and are mapped with the identity card. Following is the equation of the triplet loss function derived from FaceNet: [6]

$$\left\| f\left(x_i^a\right) - f\left(x_i^{p^p}\right) \right\|_2^2 + \alpha < \left\| f\left(x_i^a\right) - f\left(x_i^n\right) \right\|_2^2 \tag{1}$$

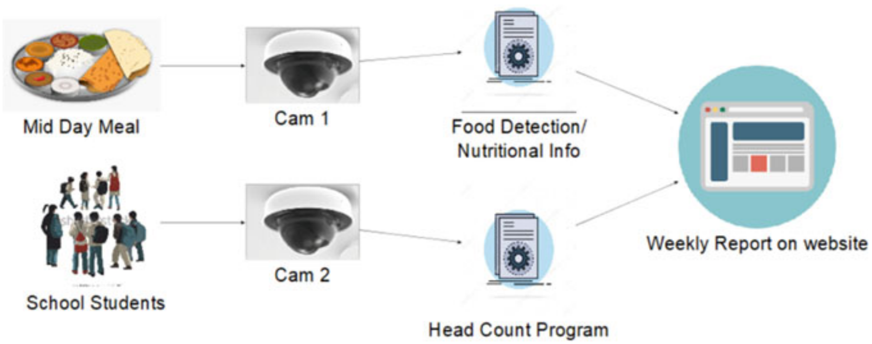


Fig. 2 Block diagram of our proposed system

Fig. 3 Photo of Indian bread

Another approach that could be helpful is the DocFace implementation presented in [15] and [16] which involves mapping of student's identity with their face. A combination of transfer learning and using a domain-specific dataset of ID-Selfie pairs helps in building a robust facial recognition module. This is very critical for sufficient supply of meals to the students registered for the scheme.

3.2 Nutritional Information using Food Images and Fact Tables

Recently, there has been a lot of work done on mapping the food images with the nutritional information. As a result, there have been several application programming interface for the same. With a combination of application programming interface and dataset of Indian food images, we perform the task of food detection and nutrition information extraction (Figs. 3 and 4).

3.3 Discrepancy Report Using Chatbot Support

Our system digitizes the process of midday meal supervision and drastically reduces the scope of error. We also consider the possibility of prospective error and provide means of rectifying it using all time available Chatbot. With the conversational agents getting more smart using the sheer scale of the corpus available and leveraging natural language processing techniques, producing near human results for conversations is not a difficult task and falls beyond the scope of our work. We, however, explore

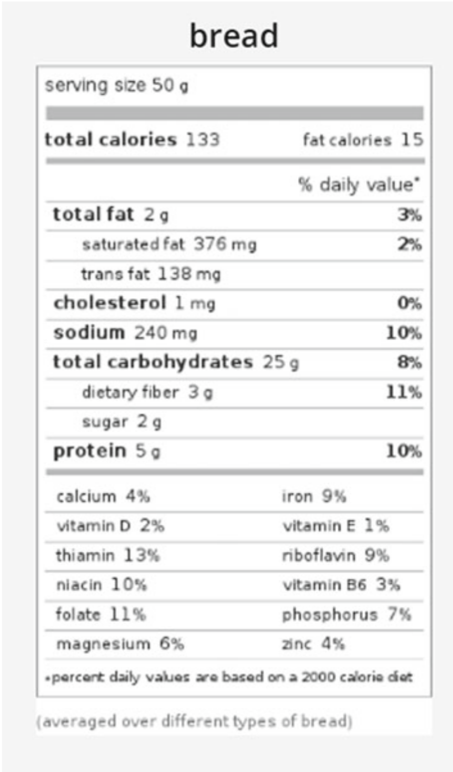


Fig. 4 Nutritional info of bread (roti) using food detection model

the digital platforms available for making a customized Chatbot, with capabilities to personalize the human interaction. This includes omnichannel support, scheduling appointments et al (Fig. 5).

4 Result

In this section, the results of simulation of all the features will be presented. Based on this, we draw our conclusion and provide insights towards building a robust system.

4.1 People Counting Using Facial Recognition

Using the vanilla FaceNet model for the facial recognition helped in establishing a baseline for the task at hand. Other approaches including [4] and [16] could be considered to enhance the model for accommodating Indian faces as the FaceNet

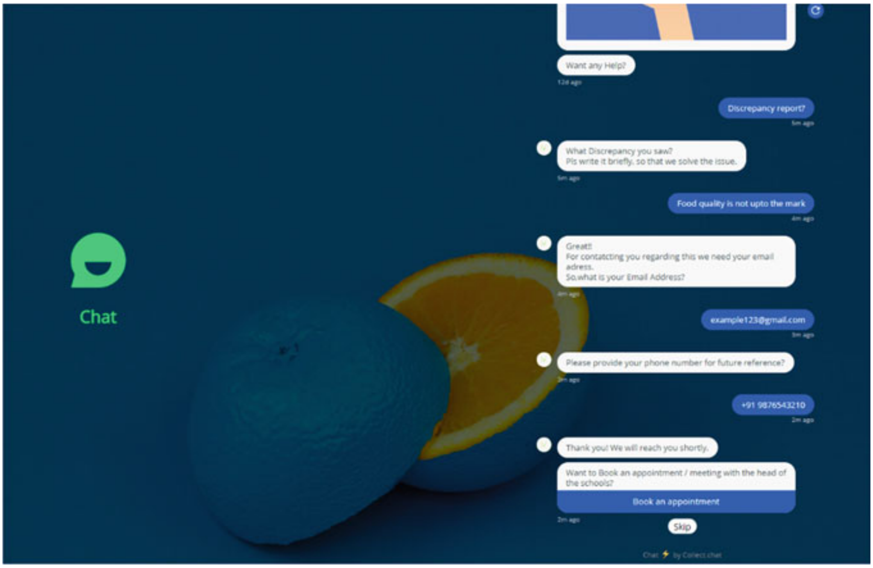


Fig. 5 Reporting a discrepancy using collect.chat

model has a potential bias towards Siamese faces. This also gives us insight about the future work towards Indian face recognition. With this idea in mind, we can integrate the database consisting of student's identity card and their photograph to come up with a solution similar to [16].

4.2 Food Detection/Nutritional Information

The goal of this feature was to identify the food items being served under the MDM scheme. We use off-the-shelf applications for the purpose of food detection and nutritional information extraction. Services offered by ClarifAI, Wolfram and Calorie Mama have shown exemplary results with their artificial intelligence techniques. We have used that model to train on custom Indian food images. The results obtained are shown in figure below are the result from using the clarifai's food model for extracting nutritional information from the image of rice provided as per the midday meal scheme menu (Fig. 6).

As an extension to the food detection and nutritional information, the velocity of the data could be accounted by regularly updating the menu as shown in Fig. 7.

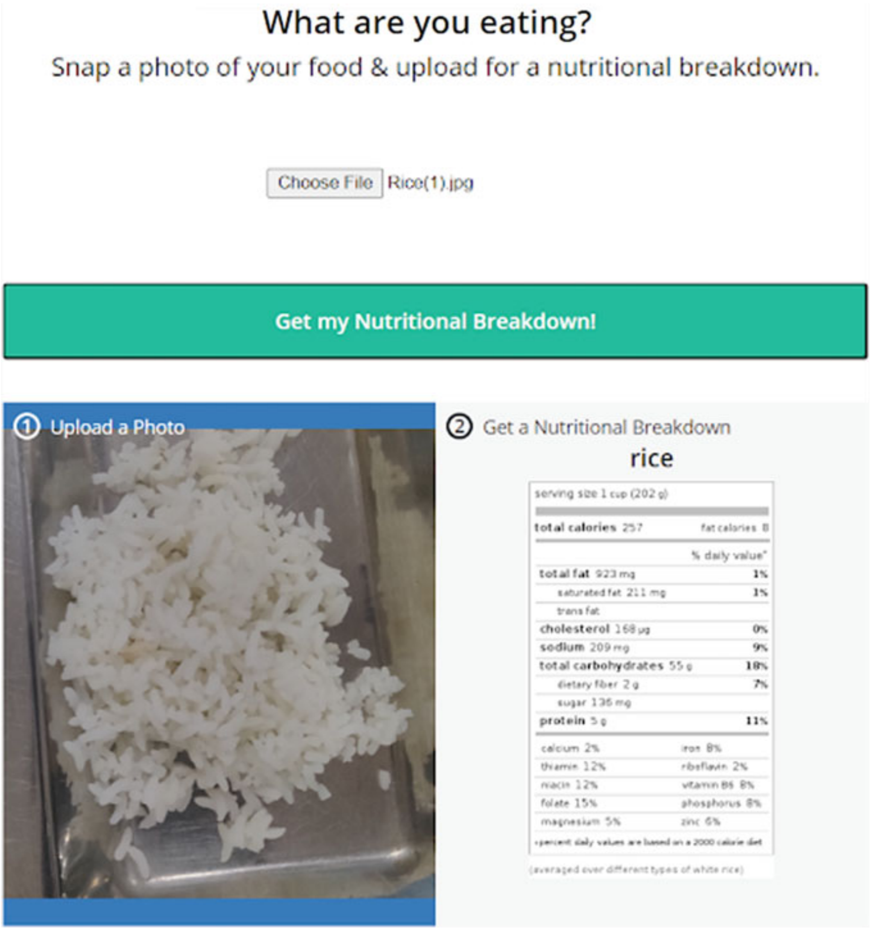


Fig. 6 Nutritional info of rice using food detection model

4.3 Discrepancy Report Using Chatbot

All time available Chatbot is achieved using the existing platforms that enable user to schedule a meeting with the concerned authority.

According to user reviews as collected by peer-to-peer review digital platform—g2.com, businesses have found collect.chat a better platform. Even though Snatchbot platform outperforms in terms of meeting requirements, but Collect.chat is found to be better in terms of ease of setup and ease of admin.

While Snatchbot is useful for large-scale enterprise, Collect.chat retains its popularity at the grassroots level with 80 percentage of small businesses against 26.7 percentage in case of Snatchbot. This clearly favours Collect.chat for the use case of midday meal scheme schools having 50 or fewer employees.

Fig. 7 Dynamic updation of weekly menu on website

Table 1 As reported by peer reviewers at g2.com, business software and services review

Collect. Chat	Snatchbot
Marketing and advising (30%)	Information technology and services (50%)
Internet (30%)	Retail (6.7%)
Printing (10%)	Supermarkets (3.3%)

Table 1 shows the reviewer industry distribution (Top 3) for the two platforms:

5 Discussion and Conclusion

With aim of improving the existing monitoring infrastructure, we provide our insights towards building a flawless system for midday meal surveillance and analytics. The existing monitoring procedure is manual and has a scope of error. We introduce techniques to automate the process in order to reduce the error. Our paper explains that by automating the monitoring task, there will be a significant reduction in error.

Despite the fund allocated for the scheme, it still fails at the execution phase. When it comes to ground level, there is a need of a feedback mechanism which is responsive to the demands and complaints of the authorities. The paper thus presents ways to curb the loopholes by implementing a transparent feedback approach to make sure that the concerned authorities are reported in a timely manner with mechanism to take a thorough follow up with them. In future, we plan to enhance the involved modules to cater to a larger audience with customized solutions.

References

1. Hermans A, Beyer L, Leibe B (2017) In defense of the triplet loss for person re-identification
2. Danilkina A, Allard G, Baccelli E, Bartl G, Gendry F, Hahm O, Hege G, Kriegel U, Palkow M, Petersen H et al Multi-camera crowd monitoring: the safest approach. 2015.16/11/21
3. Collect.chat: Chatbot tool for conversational interfaces for websites
4. Sam DB, Peri SV, Sundararaman MN, Kamath A, Babu RV (2019). Locate, size and count: accurately resolving people in dense crowds via detection
5. Dittrich F, de Oliveira LES, Britto Jr AS, Koerich AL (2017) People counting in crowded and outdoor scenes using a hybrid multi-camera approach
6. Schroff F, Kalenichenko D, Philbin J. Facenet (Jun 2015) A unified embedding for face recognition and clustering. 2015 IEEE conference on computer vision and pattern recognition (CVPR)
7. Sood G Clarifai: R client for the clarifai api, 2015. R package version 0.2
8. Chokr M, Elbassuoni S (2017) Calories prediction from food images. In *proceedings of the thirty-first AAAI conference on artificial intelligence*, AAAI'17, AAAI Press, p 4664
9. Adam M, Wessel M, Benlian A (Mar 2020) AI-based chatbots in customer service and their eects on user compliance. Electronic Markets
10. Veenu N (2011) Nutritional health status of primary school children. Indian Educ Rev 48(1):18
11. Pouladzadeh P, Shirmohammadi S, Al-Maghrabi R (2014) Measuring calorie and nutrition from food image. IEEE Trans Instrum Meas 63(8):1947–1956
12. Upadhyaya P, Bisla G (2019) Impact of mid day meal on the nutritional status of rural middle school children (age 9–15 years) of Auraiya district, Uttar Pradesh. Indian Journal of Health & Wellbeing, 10
13. Kar R, Haldar R (2016) Applying chatbots to the internet of things: opportunities and architectural elements. Int J Adv Comput Sci Appl 7(11)
14. Snatchbot: A bot builder platform designed to add multi-channel messaging to any system
15. Shi Y, Jain AK (2018) Docface: matching id document photos to selfies
16. Shi Y, Jain AK (2018) Docface+: Id document to selfie matching

Compact Planar Retrodirective Antenna with a Rat-Race Coupler for Autonomous Car Applications

P. Mahalakshmi and R. Vallikannu

Abstract In radiofrequency, spectrum 30–300 GHz is allocated for millimeter-wave communications. For overcoming the traffic access issue in cloud services over 4G networks, mm-Waves have been proposed. It has been used for many applications including industrial and academic purposes. To achieve high data rate communication and low-latency 5G has been designed to replace the bulk processing unit into planar structures mainly suited for low-power environments in automation. The features such as obstacle detection and automatic braking operation are important factors for self-driving cars. The proposed work is to design retrodirective antenna modes to work for vehicular communication (self-driving cars), which comes under mission-critical communication in 5G. The retrodirective mode is obtained using hybrid (rat-race) couplers simulated in ANSYS, which operates at 2.5 GHz.

Keywords Rat-race couplers · Retrodirective antenna · Mm-wave communications · Return loss · ANSYS · FR4 substrate

1 Introduction

In recent years, interest in autonomous automation is increasing due to the development of vehicle communication technology. Since the 1920s, the first radio-controlled electric was designed [1]. This self-autonomous vehicle is powered by embedded circuits on the roads. By the 1960s, electronic guide systems were integrated inside the autonomous cars. From 1980s, vision-guided autonomous vehicles came into the market, which was a milestone in automation technology [2]. For safer transportation in an autonomous car, the feature such as obstacle detection and automatic braking is an important factor [3]. The researches on obstacle detection and tracking mainly focus on two methods, namely computer vision methods and LiDAR, where

P. Mahalakshmi (✉) · R. Vallikannu
Hindustan Institute of Technology and Science, Chennai, India

R. Vallikannu
e-mail: vallikannu@hindustanuniv.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_12

computer vision method used deep learning algorithms to detect obstacles, and it causes poor detection when light is weak [4]. The next method LiDAR uses pulses of light to detect obstacles; due to bad weather conditions, it causes poor detection because of refraction.

Retrodirective array operates by responding toward a source upon incident interrogation with a directivity base on array theory [5]. Because of high speed, high directivity, and self-tracking process, retrodirective antennas can also be used for millimeter communications, military communications, radars, and civilian applications. The scope of using this method compare to traditional retrodirective methods such as Van atta array and heterodyne method is its size and low power consumption [6]. These couplers have a wide bandwidth and can be used for high-frequency applications. To realize in planar structures, passive retrodirective array is used. Because of this planar structure, propagation loss is reduced, and it requires no power to scan multiple antennas angles.

2 Literature Survey

Retrodirective operates by responding toward a source upon incident interrogation with a directivity based on array theory. The two types of retrodirective topologies are Van atta array and Pon arrays. Figure 1 represents the architecture of Van atta array. This array consists of a pair of antennas connected by equal lengths of transmission lines creating the symmetric array [7]. The received signal creates a phase gradient across the elements in the array based on the source direction. It can operate over a wide bandwidth but require that the incoming signal and the array be planar. Besides, for large arrays, the design of the interweaving feed lines can become intricate. Van

Fig. 1 Van atta array

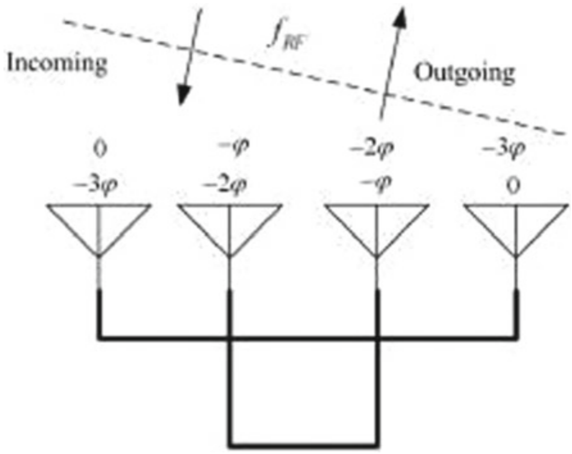
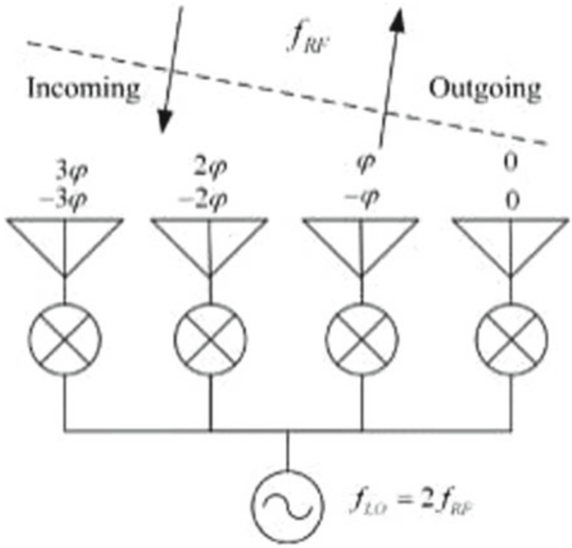


Fig. 2 Pon array



atta arrays are restricted to even mode of antennas, which limits the low-profile environments.

The next method to achieve retrodirectivity by phase conjugation is shown in Fig. 2. This implementation is sometimes known as the heterodyne approach achieves the phase reversal effect in an array at each element using phase conjugating mixers. Since they were using active devices, power consumption will be more, and it cannot be suited for low-power environments [8].

The Pon array method is an alternative way to achieve retrodirectivity using 90° hybrid couplers. To achieve the retrodirective mode, antennas to be connected to the input and isolated ports and the coupled and direct ports are terminated with reflection coefficient [9]. It has been designed for both even and odd modes. Due to the presence of sharp edges in the signal degrade its performance, it is not suited for low-power environments. To overcome these limitations, rat-race couplers or 180—degree hybrid couplers were used to overcome the limitation of signal degradation.

3 Proposed Model

The block diagram of the proposed model is shown in Fig. 3. To overcome the limitations of the existing methods computer vision-based method (causing poor detection when light is weak) and LiDAR technique (bad weather conditions), an antenna unit is integrated into on-board systems. During nighttime and bad weather conditions, signals detected from the obstacle sensor are sent to rat race with antenna unit. If a signal is detected from the obstacle sensor, then port 4 is terminated with 50 Ω; the retrodirective mode with out of phase signal is obtained where braking

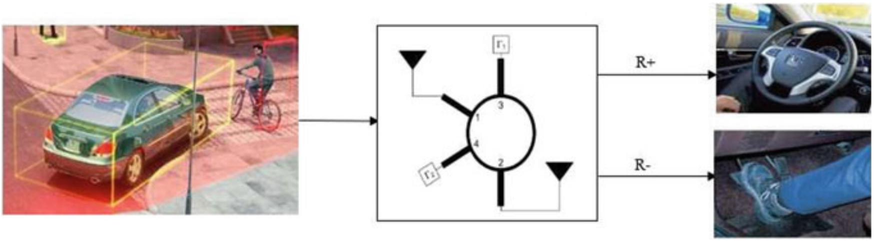


Fig. 3 Block diagram for proposed work

operation in an autonomous car. When port 3 is excited and port 4 is terminated, then non-retrodirective mode occurs, which implies that no signal is detected from an obstacle sensor, then safe driving operations occur.

3.1 *Rat-Race as Retrodirective*

The proposed model of the rat-race couplers is shown in Fig. 4. The port 1 and port 2 are connected to the microstrip patch, while other two ports are terminated with the same reflection coefficient. The antennas can act as transmitters and receivers. To derive the retrodirectivity conditions, the hybrid rat-race couplers [10] method can be used. It is obtained from the scattering matrix, as

$$[\mathrm{s}] = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & \ell^{\frac{-j\tau}{2}} & \ell^{\frac{-j}{2}} \\ 0 & 0 & \ell^{\frac{-j3\pi}{2}} & \ell^{\frac{-j\pi}{2}} \\ \ell^{\frac{-j\pi}{2}} & \ell^{\frac{-j3\pi}{2}} & 0 & 0 \\ \ell^{\frac{-j}{2}} & \ell^{\frac{-i}{2}} & 0 & 0 \end{bmatrix}$$

The output waves (b_3 and b_4) at port 3 and port 4 are the input waves (a_1 and a_2) at port 1 and port 2 are

$$b3 = \frac{1}{\sqrt{2}}\left(a_1 e^{-\frac{j\pi}{2}} + a_2 e^{\frac{j3\pi}{2}}\right) \tag{1}$$

$$b4 = \frac{1}{\sqrt{2}}\left(a_1 e^{-\frac{j\pi}{2}} + a_2 e^{-\frac{j\pi}{2}}\right) \tag{2}$$

$$a1 = \Gamma 1\, b3 \tag{3}$$

$$a2 = \Gamma 2\, b4 \tag{4}$$

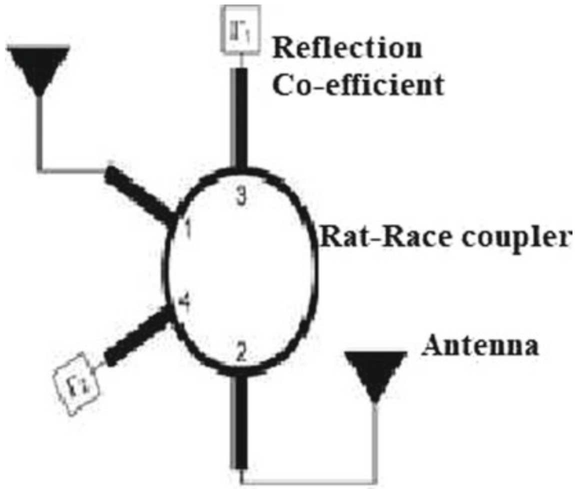


Fig. 4 Rat-race coupler as retrodirective

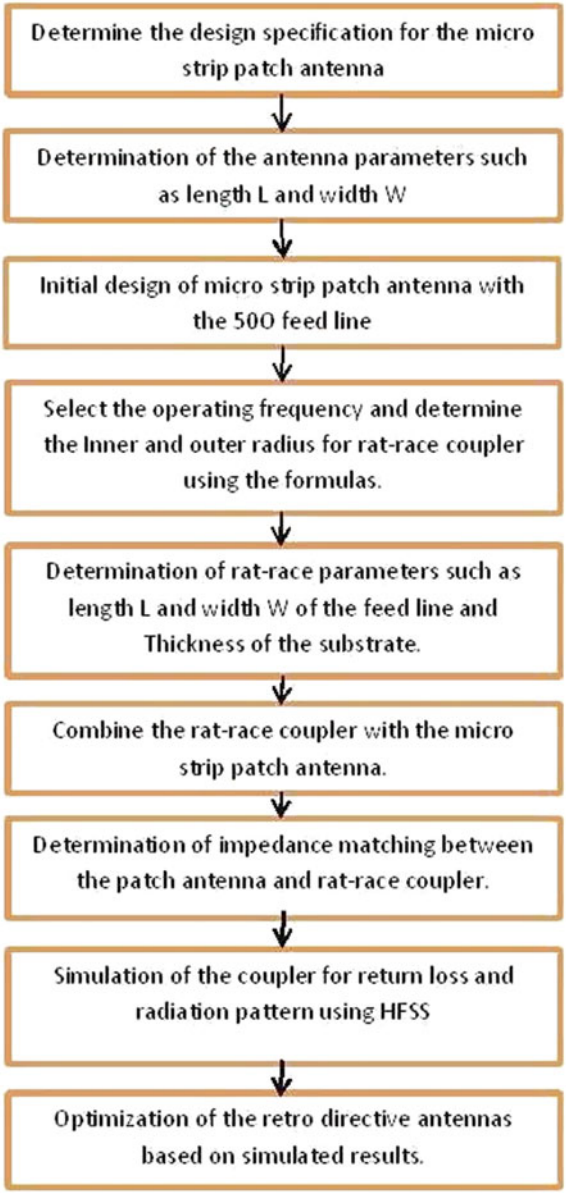
By setting the condition $\Gamma 2 = -\Gamma 1$, retrodirectivity condition is obtained [11]. By comparing the phase difference of the input waves to that of the output waves, the phase of the output is negative concerning input. Thus, the rat-race couplers reverse the incoming phase and act as a retrodirective. When port 4 is excited with 50 Ω impedance and port 3 acts as 0 Ω impedance, then coupler acts as a non-retrodirective mode of operation. Port 3 is excited with 50 Ω impedance and port 4 acts as 0 Ω impedance. Since the port 4 is terminated with 0 Ω impedance, this port will be isolated, and then, only three ports are there, and it will act as circulator operation, traveling in a clockwise direction. From the operation of circulators, the maximum power will be reached to port 2, it acts as a transmitter, and the remaining power will be reached to port 1 and act as a receiver. The radiation pattern obtained will be an in-phase signal. When port 3 acts as 0 Ω impedance and port 4 is excited with 50 Ω impedance, then port 3 is isolated and circulator operations occurred between remaining ports, the maximum power will reach port 1, act as a transmitter and port 2 as a receiver, and then, phasor output obtained will be 180-degree out of phase, where retrodirecitivity condition is obtained.

3.2 Design and Evaluation

To realize in planar structure, microstrip antennas are widely used. The patch antennas are designed with relative permittivity of $\varepsilon_r = 4.4$, and it is separated from the ground plane with FR4 epoxy substrate.

Figure 5 represents the design flow of the proposed system. It is carried out in three steps. First microstrip patch antenna is designed at 2.5 GHz. The simulation parameters used to design patch antenna are length and width are 40 mm and 60 mm;

Fig. 5 Design flow



the thickness of the substrate is 1.58 mm [12]. The second step is to design a hybrid coupler with inner and outer radius which are 17.56 mm and 16.46 mm; length and width of the feed line are 3 mm and 9.36 mm. Now combine the designed microstrip patch antenna with rat-race couplers. Determination of the impedance matching between the antenna and rat-race couplers is obtained, and simulation is performed using ANSYS-EM simulation software, return loss, and the radiation pattern is obtained.

3.3 Coupler Design

This proposed model presents a comprehensive description to model rat-race couplers [13]. The dielectric material is glass epoxy substrate Fr4, and the thickness is h = 1.58 mm at 2.5 GHz.

- The radius (inner and outer) of the rat-race coupler is calculated using

$$R = \frac{3\lambda g}{4\pi} \tag{5}$$

$$\text{Inner radius of the ring } r_1 = \left(r - \frac{w}{2}\right) \tag{6}$$

$$\text{outer radius of the ring } r_2 = \left(r + \frac{w}{2}\right) \tag{7}$$

- The wavelength of the coupler is calculated with the velocity of light C = 3 × *10*^{8} *m/s* at the operating frequency f = 2.5 GHz.

$$\lambda 0 = \frac{c}{f} \tag{8}$$

$$\lambda g = \frac{\lambda 0}{\sqrt{\varepsilon \text{eff}\, 2}} \tag{9}$$

- Effective dielectric constant $\varepsilon_{\text{eff}1}$ and $\varepsilon_{\text{eff}2}$ is calculated using

$$\varepsilon \text{eff} = \frac{\varepsilon + 1}{2} + \frac{\varepsilon \text{r} - 1}{2}\left(\frac{1}{\sqrt{1 + 12\frac{h}{w}}}\right) \tag{10}$$

- The width for the port input and width for the branches are calculated using the formula

$$\frac{w}{h} = \frac{8e^{A}}{e^{2A} - 2} \text{for} \frac{w}{d} < 2 \tag{11}$$

$$\text{where } A =; \frac{z}{60}\sqrt{\frac{\varepsilon r + 1}{2}} + \frac{\varepsilon r - 1}{\varepsilon r + 1}\left(0.23 + \frac{0.11}{\varepsilon r}\right) \tag{12}$$

Table 1 Design parameters for hybrid couplers

Parameters	Value (mm)
Inner radius (r1)	17.56 mm
Outer radius (r2)	16.46 mm
Length of the feed line	3 mm
Width of the feed line	9.36 mm

where λg represents the guided wavelength, r and w represent radius and width, ε_r represents the dielectric permittivity, and h denotes the thickness of the substrate. Table 1 represents the values obtained from the above Eqs. (6), (7), (10), and (11).

4 Results and Discussion

4.1 Retrodirective Mode

The optimization is done for this coupler dimension to evaluate the results of scattering parameters. The return loss is defined as the ratio of power reflected from the port to the power communicated into port. The couplers operate at a center resonant frequency of 2.5 GHz, and very low return losses are observed as shown in Fig. 6, nearly as—22 dB in retrodirective mode of operation.

Antennas are connected with port 1 and port 2, while other ports are terminated with reflection coefficient. For the coupler to act as retrodirective mode port 4 is excited with 50 Ω, it gets equally split into two ports, but traveling in a clockwise direction since it acts as a circulator. Port 3 is terminated with zero impedance. From the results shown in Fig. 7, 2.96 dB gain is obtained in retrodirective mode, where both E-plane and H-plane co-polarizations are considered.

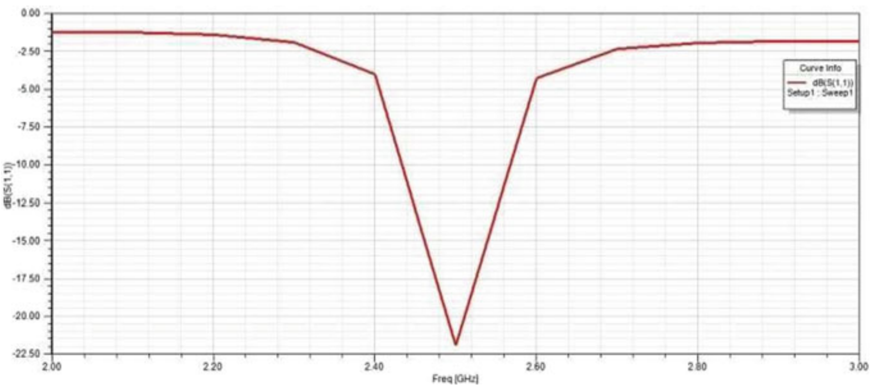


Fig. 6 Frequency versus scattering parameters

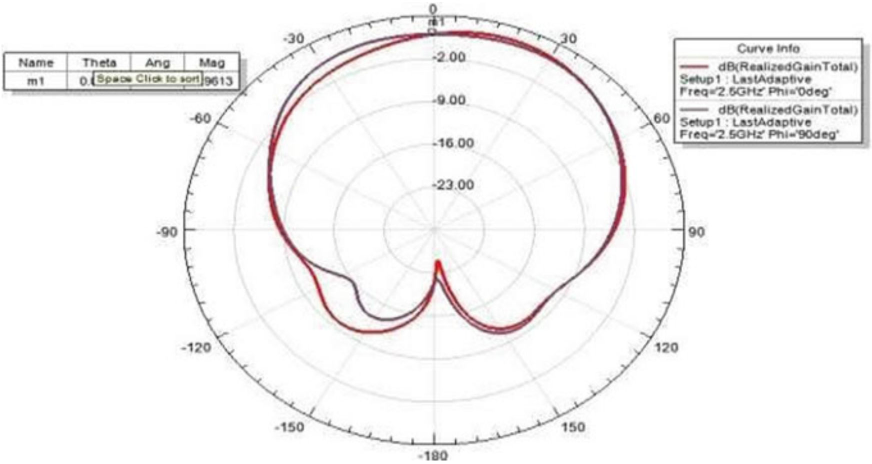


Fig. 7 Radiation pattern when port 4 is excited and port 3 is terminated

4.2 Non-Retrodirective Mode

The couplers operate at a center resonant frequency of 2.5 GHz, and low return losses are observed as shown in Fig. 8 nearly as −19 dB in non-retrodirective mode of operation.

Antennas are connected to port 1 and port 2, while other ports are terminated with reflection coefficient. For the coupler to act as non-retrodirective mode port 3 is excited with 50 Ω, it gets equally split into two ports, but traveling in a clockwise direction since it acts as a circulator. Port 4 is terminated with zero impedance. From Fig. 9, simulated radiation pattern for non-retrodirective mode 2.12 dB gain is obtained.

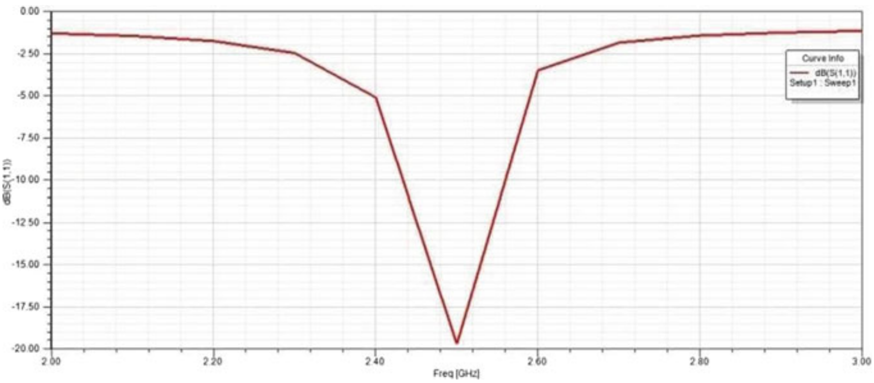


Fig. 8 Frequency versus scattering parameters

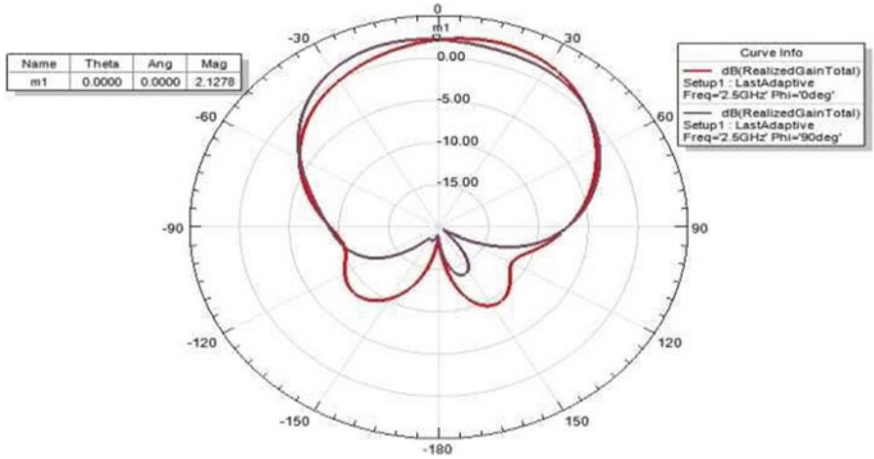


Fig. 9 Radiation pattern when port 3 is excited and port 4 is terminated

From the above radiation pattern of both modes, the phase difference between the input waves and output waves is negative or 180° out of phase of incoming waves. From this, retrodirective condition is obtained using couplers. Whenever out of phase signals are produced, automatic braking operation is performed.

5 Conclusion

Thus, the retrodirecitivity can be realized in planar structures using rat-race couplers and evaluated using ANSYS simulation software. The two modes of operation retrodirective mode and non-retrodirective mode are obtained from the simulated results return loss and radiation pattern. We can replace the processor-controlled sensor unit as couplers with an antenna because the coverage of the antenna will be high, and it is also very light when compared to the bulky processor unit. Multiple usages of this planar structure can be connected in an array format and will act as a feeding network to work as the function of the processor unit. To improve the gain of the antenna, series fed array is interconnected along with each antenna in an array for millimeter-wave communications.

References

1. Manoharan S (2019) Quot.: An improved safety algorithm for artificial intelligence enabled processors in self driving cars. &quot. J Artific Intell 1(02)
2. Williams DA (1992) Millimeter wave radars for automotive applications. IEEE MIT-S Int.Microwave Symp. Dig., Albuquerque, NM

3. Meinel HH (1995) Commerical applications of millimeter waves: history, present status and future trends, IEEE Trans Microwave Theory Tech 43
4. Dobias F, Grabow W (1994) Adaptive arrays antennas for 5.8 GHz vehicle to roadside communications. IEEE 44th Veh Tech Conf. Stockholm, Sweden
5. Andre SN, Leonard DJ (Feb 1964) An active retro directive array for satellite communications. IEEE Trans Antennas Propag 12(2)
6. Yen SC, Chu TH (Jan 2004) A retrodirective antenna array with phase conjugation circuit using sub- harmonically injection-locked self-oscillating mixers. IEEE Trans Antennas Propag 42(1)
7. Nielsen ED (Jan 1970) Square van atta reflector with conducting mounting plane. IEEE Trans Antennas Propagation AP-18
8. Pon CY (Mar 1964) Retrodirective array using the heterodyne technique. IEEE Trans Antennas Propagation, vol 12
9. Pon CY (1961) Hybrid-ring directional couplers for arbitarary power divisions. IRE Trans Microw Theory Techn 9(6):529–535
10. Lin Z, Jan-Dong., Wu B, Li JM (2015) Hybrid type rat-race coupler designs, APMEC
11. Thomas S, Reynolds MS (Apr 2010) QAM backscatter for passive UHF RFID tags. IEEE Int Conf RFID
12. Al-Hiti AS (Jan 2019) Design of rectangular microstrip patch antenna for WLAN and WIMAX applications. ARPN J Eng Appl Sci 14
13. Chiou YC, Kuo JT, Chan CH (2009) New miniature dual band rat-race coupler with microwave c-sections. IEEE MIT-S Int Microwave Sympo Dig, pp 701–704

Digital Watermarking Technique in Transform Domain Using Meta-Heuristic Algorithm

Ayesha Shaik and M. Nivedita

Abstract The digital data transmission or sharing through the inter-networked systems has been increased heavily in recent decades as Internet access is reachable. Due to that, illegal digital data distribution becomes the main concern as the copyrights of the owner can be breached through the tools available on the Internet. Digital multimedia (in the form of audio, text, image, and video) can be easily modified and shared for profit. To address this issue, digital watermarking (WMG) has been introduced as one of the solutions. In this article, an optimized blind dual WMG technique for the digital images in the transform domain using the Slantlet transform and discrete cosine transform (DCT) is proposed. The consideration factors of WMG such as robustness and imperceptibility are both contradicting. So, to provide a trade-off between these factors, a meta-heuristic algorithm is used to balance these factors. The visual quality of the watermarked image (WMI) is verified with the reference image quality (RIQ) measures and no-reference image quality (NRIQ) measures and shown that the proposed method produces high-quality WMIs.

Keywords Digital watermarking · Image · Meta-heuristic optimizer · Imperceptibility · No-reference image quality

1 Introduction

In today's digital era, data communication through the inter-networked multimedia systems became very popular. In this regard, the digital data that has been transmitted or communicated over the Internet is susceptible to third-party attacks. Malicious parties may generate the illegal copies of the data as the duplication of digital data has become easier with the available data editing tools and re-distribute them. Then, the copyrights of the content owner will be at trouble, and digital WMG is one such solution to protect the copyrights of the owner and verify the authentication of the data.

A. Shaik (✉) · M. Nivedita
School of Computing Science and Engineering, Vellore Institute of Technology, Chennai 600127, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_13

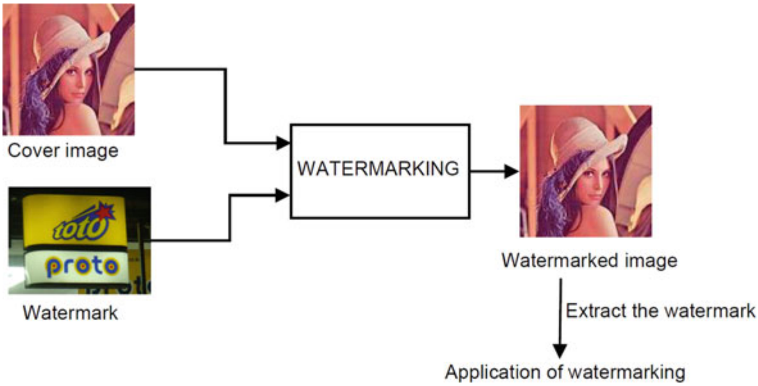


Fig. 1 Basic watermarking system

Digital WMG technique is a branch of information hiding domain (other branches are steganography and cryptography), where copyright as watermark (digital data such as text, audio, image, and video) is hidden inside the digital data that needs to be protected and extracted for verification when the copyright of the owner is in trouble [1]. The basic block diagram for WMG is as shown in Fig. 1. Digital WMG can be classified based on the following categories: (i) embedding domain: spatial and transform domain, (ii) robustness towards attacks: Robust, Fragile, Semi-fragile, (iii) Embedding method: additive and multiplicative, (iv) requirement of the original image during the extraction of the watermark: non-blind and blind, and (v) significance: adaptive and non-adaptive WMG schemes [2].

In this article, a digital watermarking scheme for digital images in the transform domain using SLT and DCT is proposed such that the combined characteristics of both the transforms can be captured for embedding the watermark in a hybrid manner. The proposed method is not using a constant watermark embedding strength and uses optimal embedding strengths selected with moth flame optimizer to balance the contradicting factors of WMG, which are robustness and imperceptibility (These two factors have conflicting properties such as improving one factor will degrade another). The watermark is decomposed into m-shares and embedded as a share instead of its original format such that the visual quality of the WMI can be improved by reducing the degradation done to the original image. Further, the article is organized as follows: Sect. 2 details the state-of-art WMG techniques available in the literature. Section 3 provides the proposed methodology and the preliminaries in detail, Sect. 4 gives experimental results and implementation analysis, and Sect. 5 provides the conclusion.

2 Literature Review

Tampering of digital images is an extremely significant concern in this digital era, as the modification of the digital data is easy. It is urgent concern for the researchers working in this domain to protect the copyrights and tampering detection. The state-of-art WMG techniques have been developed to address the above issues, and a few of those watermarking schemes are detailed in this section. The primary factors of WMG technique are robustness and imperceptibility, but it is very difficult to develop a WMG system which will be robust and imperceptible.

It is important to provide a trade-off between these factors when a WMG system is being developed. A discussion on standard WMG system frameworks and the requirements required to design the WMG techniques, and their limitations are discussed in [3].

The capability of WMG techniques can be discussed in terms of security, authenticity, recovery. The basics of digital WMG techniques along with their strengths and weaknesses for the tampered image detection and the method to recover them have been presented in [4]. A WMG technique for digital images using two-level discrete wavelet transform (DWT), discrete cosine transform (DCT), and QR decomposition has been presented in [5]. The watermark is inserted in the selected blocks of the original image depending on the entropy values. The authors of the technique in [5] discussed the imperceptibility and robustness of their method with the help of experimental results.

For protecting copyrights of the owner, the distribution of the authenticated documents should be available for the authorized user such that they are not available for the unauthorized user, as there is a need to protect the digital documents from the tampering and re-distribution in an illegitimate manner. In [6], the author has implemented the DWT and SVD with the least significant bit (LSB)-based WMG technique to protect the digital rights of the content owner. This technique is to provide access to the authenticated documents to legitimate users and preventing access to illegitimate users. In this scheme, the cover image is decomposed into sub-bands by using two-level DWT, and the sub-bands are transformed using SVD. This method gives higher visual quality and provides robustness towards the attacks such as JPEG compression, Gaussian noise, to compromise the security of the intellectual property.

A digital WMG technique that uses non-subsampled contourlet transform (NSCT), DCT, and multi-resolution singular value decomposition (MSVD) has been discussed in [7]. Because of this combination of NSCT, DCT, and MSVD in a hybrid manner, this technique is robust against signal processing and geometrical attacks as it has the advantage of the characteristics of all the transforms. A DWT and PCA-based digital image WMG technique using Bhattacharya distance and Kurtosis has been discussed in [8]. Here, PCA is used to compress the watermark, and the compressed watermark is embedded in the transformed original image to improve the visual quality of the WMI. This technique uses Bhattacharyya distance and Kurtosis to select the optimal scaling factors to make the WMG adaptive when compared to the constant scaling factor. The main aim to select the optimal scaling factors as the images and their characteristics are not the same, and selecting a static embedding factor is not serving the purpose of providing a trade-off between robustness and imperceptibility.

A digital WMG technique that aims at copyright protection, data security, and the illegal accessing of data from unauthorized users has been presented in [9]. Another digital WMG technique using bat optimizer (BA) and SURF features is discussed in [10]. The high-frequency coefficients of the stationary wavelet transform (SWT) of the original image are inserted with the watermark scaled with optimal scale factors using the BA framework to tolerate attacks. The SURF detector is processed on the WMI to extract the feature points that can be used for corrections to geometric distortions. To extract the watermark, the probable geometrical distortions need to be corrected, using the SURF rotation and scaling invariance property [10].

SVD transformation was used very often in the WMG techniques, but they suffer from the false-positive problem for protection of copyrights. So, a solution to counter this issue has been presented in [11] with the hashing techniques with 128-bit values, and the hash values are verified to claim the ownership. A reversible WMG scheme for CT scan images with high capacity has been presented in [12] based on the region of interest and region of non-interest.

3 Proposed Methodology

In this section, the required preliminaries and the proposed WMG technique are discussed in detail. The proposed techniques are divided into two sections: (i) watermark embedding and (ii) watermark extraction.

3.1 Preliminaries

1. *Slantlet Transform*: Slantlet transformation (SLT) is an equivalent representation to DWT [13, 14], where the filters are shorter for SLT, and it provides better time localization and smoothness. The equivalence of DWT and SLT is shown in Fig. 2a and Fig. 2b, respectively.
2. *Moth Flame Optimizer (MFO):* The MFO algorithm [15, 16] is a three-tuple function that finds global optimum, and it is defined as follows:

MFO = (*A, B, C*), where *A* is a function that generates moth population and their fitnesses randomly, *B* is a function for the moths that move around the search space, and *C* is a function that returns true and false if termination criterion is satisfied and not satisfied, respectively. *A* is defined as function of moths *M* and moths fitness OM; where m_{ij} corresponds to *j*th index of *i*th moth; $1 \leq i \leq m;\ 1 \leq j \leq n.$

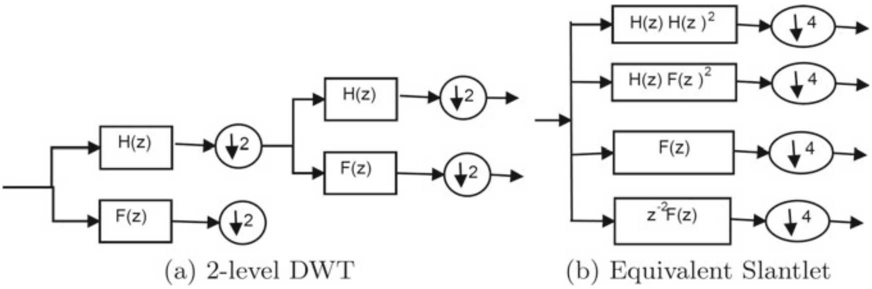


Fig. 2 Slantlet transform filters and its equivalent DWT filters

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & \cdots & m_{1n} \\ m_{21} & m_{22} & m_{23} & \cdots & m_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{m1} & m_{m2} & m_{m3} & \cdots & m_{mm} \end{bmatrix} \quad (1)$$

The moth fitness's are stored as shown below.

$$OM = \begin{bmatrix} f(m_1) \\ f(m_2) \\ \vdots \\ f(m_m) \end{bmatrix} \quad (2)$$

where m_i denotes ith moth and $m_i = (m_{i1}, m_{i2}, m_{i3}, \ldots\ldots\ldots.., m_{in})$, and n is number of moths.

3.2 *Watermark Embedding*

The block diagram of the proposed WMG technique is as shown in Fig. 3. The original image is decomposed into sub-bands using the Slantlet transform, and the entropies of the sub-bands are measured. The sub-band with highest entropy is selected, and DCT is applied on that sub-band, by transforming the sub-band into frequency coefficients. The highest K frequency coefficients are inserted with the watermark. The watermark is decomposed into two shares (S1, S2), and one of the share Si, where $i = 1$ or 2, is selected randomly for inserting into the selected cosine transform coefficients using the optimal scale factors generated with the moth flame optimizer. The optimal scale factors are preferred in WMG to balance the robustness and imperceptibility factors of the WMG scheme.

Use of constant scale factor for all the transform coefficients of the image is not preferable; considering the characteristics of the image is not static. So, in the

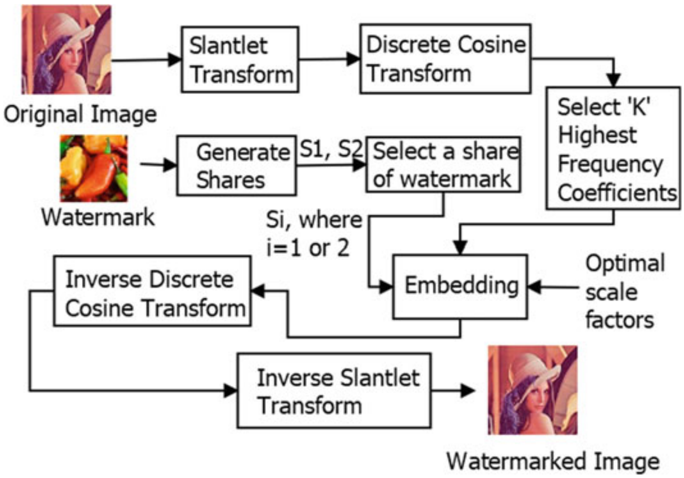


Fig. 3 Proposed watermark embedding algorithm

proposed technique, moth flame optimizer is used to select the optimal embedding strengths for the watermark. The watermark embedding algorithm is detailed in Algorithm 1.

Algorithm 1 Watermark embedding

Input: Original Image I, Watermark W, K, and optimal scaling factors a_i
Output: Watermarked Image, I_W

1: Apply 3-level Slantlet Transform on I decompose into sub bands (SBs) S_i, where $i = 1, 2..., 8$
2: Apply DCT on the approximate sub-band S_{approx}, and the result is F_{approx}
3: Select K highest coefficients from F^{High}_{approx}, where $High$=$1, 2, ..., K$
4: W is decomposed into shares, say $W_{share1}, W_{share2}, ..W_{shareM}$ and select a share, say W_{select} randomly to embed
5: Embed W_{select} into F^{High}_{approx}

$$F^{High}_{approx}{}^{WM} = F^{High}_{approx} + a_i \times W_{select} \tag{3}$$

6: Perform inverse DCT on $F^{High}_{approx}{}^{WM}$, result is say S^{WM}_{approx}
7: Perform inverse slantlet transform on S^{WM}_{approx}, result is say I_W

3.3 Watermark Extracting

The proposed watermark extracting block diagram is shown in Fig. 4. The WMI is decomposed into sub-bands using the Slantlet transform, and the entropies of the

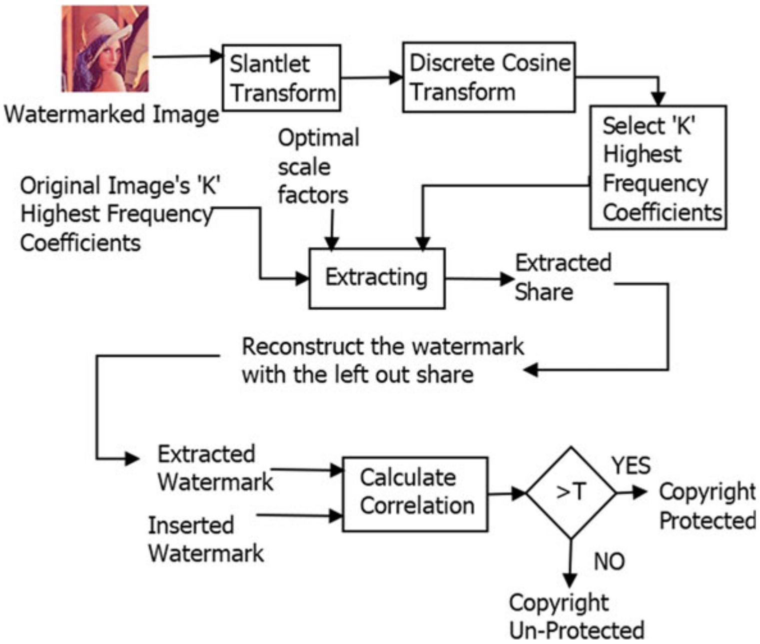


Fig. 4 Proposed watermark extracting algorithm

sub-bands are measured. The sub-band with highest entropy is selected, and DCT is applied on that sub-band, by transforming the sub-band into frequency coefficients. The highest K frequency coefficients are selected to extract the watermark share and reconstruct the whole watermark. The correlation is calculated between the inserted and extracted watermark, and if the correlation is higher than the predefined threshold, then the copyrights are protected,; otherwise, the copyrights are not protected. The watermark extracting algorithm is detailed in Algorithm 2.

Algorithm 2 Watermark extraction

Input: Possibly modified WMI, I_W^*,a_i,K, W_{select},F_{approx}^{High}
Output: ownership proetcted or not
1: Apply 3-level Slantlet Transform on I_W^* decompose into sub bands (SBs) S_i^*, where $i = 1, 2..., 8$
2: Apply DCT on the approximate sub-band S_{approx}^*, and the result is F_{approx}^*
3: Select K highest coefficients from $F_{approx}^{High^*}$, where $High$=1, 2, ..., K
4: Extract W_{select}^* into $F_{approx}^{High^*}$

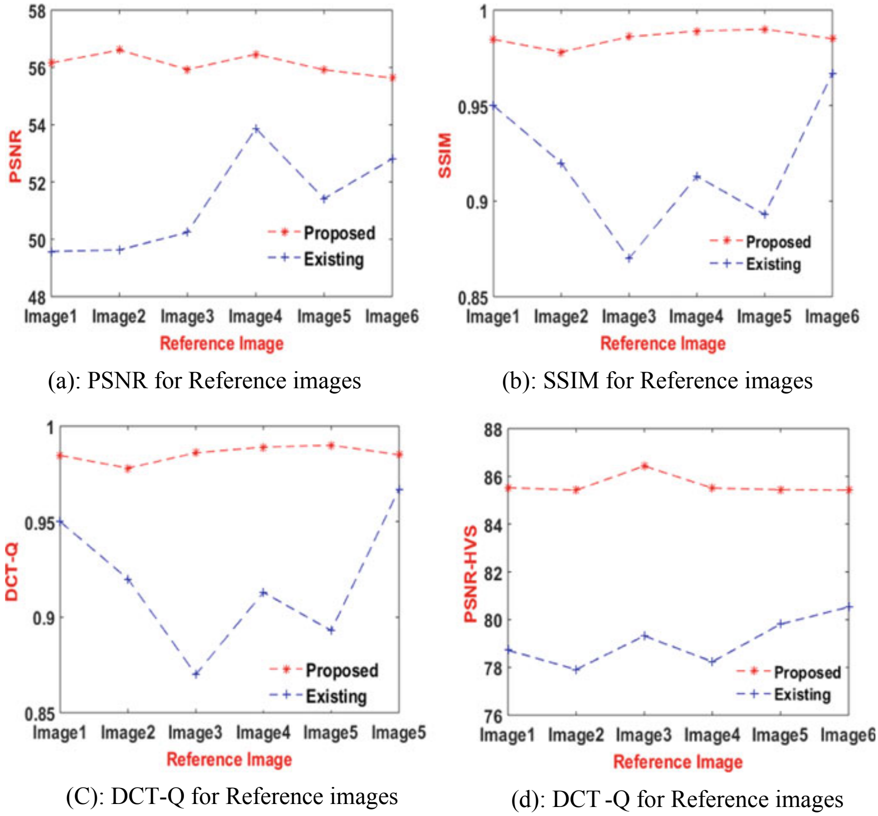
$$W_{select}^* = \frac{F_{approx}^{High^*} - F_{approx}^{High}}{a_i} \tag{4}$$

5: Find normalized correlation (NC) between inserted and extracted watermarks, say W_{select} and W_{select}^* respectively
6: If the NC is greater than the predefined threshold then the ownership is protected. Otherwise it is not protected.

4 Results and Analysis

The efficiency of the proposed method is tested on the images from BOWS [17] dataset. The peak signal-to-noise ratio (PSNR) values achieved with the proposed method and the method in [18] are compared and plotted in Fig. 5a. The PSNR for the proposed method is in the range of 55–57 dB, and the PSNR for the existing method is around 49–53 dB. From Fig. 5a, one can say that the proposed method gives better PSNR (If PSNR is high, WMI quality is good, and if it is less, then the WMI is more distorted). The structural similarity index measure (SSIM) values achieved with the proposed method, and the method in [18] is compared and plotted in Fig. 5b. The SSIM for the proposed method is in the range of 0.98–1, and the existing method is having SSIM in the range of 0.86–0.96. From Fig. 5b, one can say that the proposed method gives better SSIM (high SSIM denotes that the technique is better). The NIRQ measures such as DCT quality (DCT-Q) measure is also calculated for the proposed method and compared with the existing method [18].

The DCT-Q values obtained for the proposed method and the method in [18] are compared and plotted in Fig. 5c. The proposed method efficacy is also justified with one more image quality measure PSNR-HVS which are compared with existing method and plotted in Fig. 5d. The PSNR-HVS for the proposed method is in the range of 85–87 dB, and the PSNR for the existing method is around 77–83 dB. From Fig. 5d, one can say that the proposed method gives better PSNR-HVS. The average values computed for the performance measures are listed in Table 1 for the proposed method and the existing method [18]. The existing method is the method given in [18], and the PSNR, SSIM, and PSNR-HVS values will not be the same for other existing methods. In Table 1, only the method in [18] has been implemented and compared with the proposed technique. The PSNR, SSIM, and PSNR-HVS values are



(a): PSNR for Reference images

(b): SSIM for Reference images

(C): DCT-Q for Reference images

(d): DCT -Q for Reference images

Fig. 5 Implementation results for proposed method and existing method [18]

Table 1 Average values obtained for the proposed method and the existing method

Average values for quality measure for the images in the dataset	Proposed method	Existing method
PSNR	56.515	48.63
SSIM	1	0.98
DCT-Q	0.985	0.923
PSNR-HVS	86.515	79.627
Computational time (in sec)	1.21	1.6699

higher for the proposed method compared to the existing method as in the proposed method, only a share of the watermark is embedded and not the whole watermark itself. But in the existing method,, the whole watermark is inserted which degrades the visual quality of the WMI. The computational time of the proposed method is also less compared to the existing method because of following reasons: (i) In the existing method, DCT, DWT, and SVD transforms are used for watermarking, and in the proposed method, SLT and DCT transforms are used, (ii) the computational complexity of the proposed method is $O(N^2)$ (maximum of DCT-$O(N^2)$, SLT-$O(N^2)$), and the computational complexity of the existing method is $O(N^3)$ (maximum of DCT-$O(N^2)$, DWT-$O(N^2)$, SVD-$O(N^3)$).

The image quality measures that are used for calculating the watermarked image visual quality are PSNR, SSIM, DCT-Q, and PSNR-HVS.

The PSNR is computed as follows:

$$PSNR\left(O, \hat{O}\right) = 10.log_{10} \frac{\left(MAX\left(\hat{O}\right)^2\right)}{MSE} \tag{5}$$

$$MSE\left(O, \hat{O}\right) = \frac{1}{MN} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left[O(i, j) - \hat{O}(i, j)\right]^2 \tag{6}$$

where O and $\hat{O}$ are original image and watermarked images, and (i, j) are the pixel locations, and MXN is the size of the image.

The SSIM can be calculated as follows:

$$SSIM(P, \mathrm{Q}) = \frac{\left(2\mu_P \mu_{\mathrm{Q}} + C1\right)\left(2\mu_{PQ} + C2\right)}{\left(\mu_P^2 + \mu_Q^2 + C1\right)\left(\sigma_P^2 + \sigma_Q^2 + C1\right)} \tag{7}$$

where *P, Q* are original and watermarked images of size $M \times N$, respectively, μ_P and μ_{Q} are average of P and Q, respectively, σ_P^2 and σ_Q^2 are variance of P and Q, respectively, and *C1* and *C2* are used for stabilization.

The DCT quality measure [19] (DCT-Q) techniques use SSIM and DCT, where the frequency components are provided weightage based on their luminance, contrast and brightness comparisons. The PSNR-HVS quality measure [20] uses PSNR and human visual characteristics (HVS) into consideration to calculate the image quality.

5 Conclusion

In this article, a digital watermarking scheme for digital images in order to protect the copyrights of the owner from the malicious parties during digital data distribution over the Internet is proposed. The proposed method uses the Slantlet transform and discrete cosine transform for embedding the watermark. The watermark is decomposed into shares, and the randomly chosen share of the watermark is inserted into the hybrid transform coefficients. The embedding strengths of the watermark are selected using moth flame optimizer to ensure trade-off between the contradicting factors of watermarking scheme such as imperceptibility and robustness. The experimental results demonstrate that the proposed scheme performs better compared to the listed schemes in the article. The robustness of the proposed scheme is verified, and it is shown robust towards a list of attacks. Overall, the proposed approach provides better watermarking quality and robustness. The proposed method is best suited for the applications where ownership protection is required with the balance in robustness and imperceptibility. The visual quality of the watermarked image obtained through the proposed scheme is better, and it is justified with the reference image quality measures such as PSNR, SSIM, and PSNR-HVS and no-reference image quality measure such as DCT-Q.

References

1. Ahmadi SBB, Zhang G, Wei S (2020) Robust and hybrid svd-based image water-marking schemes. Multimedia Tools Appl 79(1):1075–1117
2. Kumar S, Singh BK, Yadav M (2020) A recent survey on multimedia and database watermarking. Multimedia Tools Appl 1–49
3. Begum M, Uddin MS (2020) Digital image watermarking techniques: a review. Information 11(2):110
4. Swain M, Swain D, Paikaray BK (2020) A competitive analysis on digital image tamper detection and its secure recovery techniques using watermarking. In: Machine learning and information processing. Springer, pp 477–484
5. Zairi M, Boujiha T, Abdelhaq O (2020) An algorithm for digital image watermarking using 2-level dwt, dct and qr decomposition based on optimal blocks selection. In: Proceedings of the 3rd International conference on networking, information systems & security, pp 1–4
6. Kumar A (2020) A review on implementation of digital image watermarking techniques using lsb and dwt. In: Information and communication technology for sustainable development. Springer, pp 595–602
7. Singh R, Nigam S, Singh AK, Elhoseny M (2020) Integration of wavelet transforms for single and multiple image watermarking. In: Intelligent wavelet based techniques for advanced multimedia applications. Springer, pp 51–63
8. Narvekar A, Swain A, Satpurkar A, Ambadekar S (2020) Adaptive digital image watermarking using dwt-pca. Available at SSRN 3568133
9. Shalom RG, Kavitha P, Murugan E (2020) A study on digital image watermarking characteristics, types and its applications. Stud Indian Place Names 40(20):737–744
10. Pourhadi A, Mahdavi-Nasab H (2020) A robust digital image watermarking scheme based on bat algorithm optimization and surf detector in swt domain. Multimedia Tools Appl 1–25
11. Singh N, Joshi S, Birla S (2020) Color image watermarking with watermark hashing. In: Intelligent computing techniques for smart energy systems. Springer, pp 673–681
12. Memon NA, Alzahrani A (2020) Prediction-based reversible watermarking of CT scan images for content authentication and copyright protection. IEEE Access 8:75448–75462
13. Rayachoti E, Tirumalasetty S, Prathipati SC (2020) Slt based watermarking system for secure telemedicine. Cluster Comput 1–10
14. Selesnick IW (1999) The slantlet transform. IEEE Trans Signal Process 47(5):1304–1313
15. Faris H, Aljarah I, Mirjalili S (2017) Evolving radial basis function networks using moth–flame optimizer. In: Handbook of neural computation. Elsevier, pp 537–550
16. Xu Y, Chen H, Luo J, Zhang Q, Jiao S, Zhang X (2019) Enhanced moth-flame optimizer with mutation strategy for global optimization. Inf Sci 492:181–203
17. http://bows2.ec-lille.fr. Accessed June 2020
18. Zhang L, Wei D (2019) Dual dct-dwt-svd digital watermarking algorithm based on particle swarm optimization. Multimedia Tools Appl 78(19):28003–28023

19. Lv D, Bi D-Y, Wang Y (2010) Image quality assessment based on DCT and structural similarity. In: 2010 6th International conference on wireless communications networking and mobile computing (WiCOM). IEEE, pp 1–4
20. Egiazarian K, Astola J, Ponomarenko N, Lukin V, Battisti F, Carli M (2006) New full-reference quality metrics based on HVS. In: Proceedings of the second international workshop on video processing and quality metrics, vol 4

Centralized Smart Air Purifier System for Industrial Applications

Aditya Asabe, Deepak Tiwari, Manu Dubey, Vedant Joshi, Shardul Shrikhande, and R. Mohan

Abstract Small-to-medium scale enterprises (SMEs) in India involve many manufacturing operations which involves emission of air pollutants, gases with foul odor, etc. inside the workplace which results in an unhealthy environment. This in turn affects the workers' health on the long term. There are certain industry grade air purifiers out there in the market. But, they are not economical and not smart enough to meet day-to-day operational requirements. So this article aims to present the concept, design and analysis of the prototype of a smart centralized air purifier system. The simulations are carried out in Ansys Fluent software to substantiate our model and empirical data. The proposed centralized air purifier is able to eliminate air pollutants of 10 microns and above from the polluted air in the industrial environments and it operates on the concepts of connected devices. Thus an economical, smart and healthy air purifier is developed to suit industrial applications.

Keywords Connected devices · Industrial air purifier · IoT · Smart devices · Air pollution

1 Introduction

World Health Organization estimated worldwide about 8 billion premature deaths every year due to cardiovascular and respiratory diseases as a result of exposure to particulate matter which are 2.5 μm or less in contaminated ambient air [1]. The indoor air pollutants are more dangerous and cause diseases in affected places [2]. This adequate filtration is essential to maintain good indoor air quality to avoid eye irritation and headache [3]. Air purification based on principles of filtration and catalytic oxidation types are being used to prevent the harmful contaminants [4, 5]. The filtration type air purifiers remove air pollutants efficiently, whereas fouled filters increases energy consumption by resisting the air passing through it, thereby increase in operating cost [6]. So, manual intervention is inevitable to restore the

A. Asabe · D. Tiwari · M. Dubey · V. Joshi · S. Shrikhande · R. Mohan (✉)
Vellore Institute of Technology, Chennai, India
e-mail: mohan.r@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_14

filtering efficiency. In general, operating costs contributes 81% of filter's total life cycle cost [7]. So, there is a need for an economical and efficient indoor air purifier with minimal human interference for instance, to set them up, gives them instructions or accesses the data.

In this research work, an attempt was made to design efficient filtration-based air purifier with IoT based automation for better serving its purpose in an industrial environment as well as Covid-19 quarantine centers during pandemic, where no physical contact will be made between user and air purifier control unit. Air purifier prototype design iterations were carried out using DS SolidWorks. A sample control volume of 1000 mm × 700 mm × 400 mm was generated corresponding to the dimensions of the prototype. CFD analysis and simulations were carried out in ANSYS Fluent to analyze the nature of flow of air within the duct and a HEPA filter. To illustrate that the system holds its place in the industry, an industrial schematic simulated operating conditions using automation studio is presented.

2 Design and Analysis of Air Purifier Prototype

2.1 Design of Air Purifier

In order to design the purifier, the whole system was divided into three parts, first intake of air, purification system and lastly the outlet of the purified air. For the air inlet, an outer casing with open top was designed where the fan can be assembled. The fan is an induced draft type which can suck in air. This way air gets inside the purifier. For the purification system of air just after the fan, filters are placed such that the air particles flow through the filters. There are two beds of filters designed; first being the normal filter which will filter out the coarser particles from the air and the second filter bed is HEPA filter which then filters the finer particles up to 10 μm in size and several harmful gases [8]. After air passes through the filters, a duct is provided to channel air further which is optimized based on the results of CFD analysis of the system. The air from duct passes through water which increases the moisture content of the air. Humidity of air can be manipulated by passing it through silica filter installed above the water level. For killing the germs and microbes, two UV emitters are also provided [9]. Lastly for the outlet of purified air, some openings were provided in the form of cuts in outer casing. Table 1 shows the dimensions of all components arrived based on air flow, and Fig. 1 represents the model of air purifier.

The blower fan designed in DS Solidworks 2018 was imported to Ansys Design Modeler and then an enclosure was inserted around the blower fan to set a flow restriction which resembled and replicated the optimized air duct. The cushion radius, positive and negative cushion direction were taken to be 1, 1.1 and 0.1 mm according to the prototype design. A conical duct toward the outlet was designed in order to accelerate the particles through the outlet, using continuity equation. Two planes were created 10 mm away from each other to divide the cylindrical part of the duct in

Table 1 Specifications of air purifier

Items	Specifications
Fan	170 mm $\times$ 170 mm $\times$ 50 mm (along with the fan casing)
Outer casing	Height: 420 mm; inner diameter: 195 mm; wall thickness: 3 mm
Duct	Height: 200 mm; diameter: 195 mm; draft angle: 71.35°; thick: 2 mm
Filters	Normal filter: thickness: 10 mm & HEPA filter: thickness: 15 mm

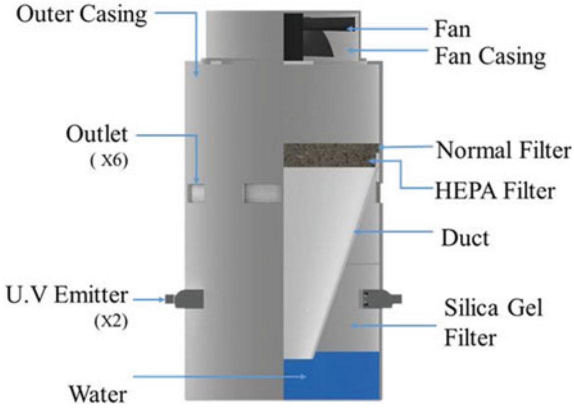


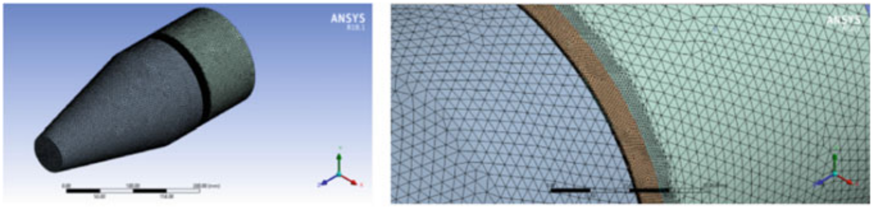
Fig. 1 Model of air purifier

to a smaller cylindrical region which was considered as a porous region resembling a HEPA filter. The slice command was used to split the duct in to separate entities giving a separate body for the porous region.

2.2 *CFD Analysis of Air Purifier System*

Computational fluid dynamics (CFD) has undoubtedly been of utmost importance, while dealing with fluid flow problems. After the model was designed, it was split into small finite elements to analyze each and every portion of the model. Different size mesh was applied to different regions based on its importance and criticality.

The wall of the frustum was given a face mesh with a patch conforming method to optimize the mesh for most accurate and precise details. A tetrahedral mesh shows more diffusive character in contrast to a hexahedral mesh and hence it was taken into account. The porous region is main focus and the flow of air particles through the outlet invariably depended on the flow characteristics of the same air particles through a HEPA filter. Hence, the size of the tetrahedral mesh on the curved surface of the porous region was drastically reduced to allow increased precision. Moreover, 3 times refinement was given on the 2 edges of the region dividing the porous region and the wall which is illustrated in Fig. 2. This was carried out to accurately observe



(a) (b)

Fig. 2 **a** Meshing of the model and **b** mesh refinement at the region of interest

the transition of the characteristics of air particles from the filter zone to the normal zone.

To analyze the air flow through a HEPA filter and duct, a viscous model was taken into account. The energy model was not taken into consideration as the problem had not dealt with any temperature changes from the inlet and outlet. Further, a k-epsilon model and realizable were chosen as it was deemed to be the most putative pertaining to the required analysis. The realizable model satisfies general mathematical constraints on the Reynolds stresses, consistent with the physics of turbulent flows, which cannot be carried out using the standard or RNG model. Moreover, Enhanced Wall Treatment was deployed as the simulation involved a detailed study of air flow and accumulation along the walls of the duct. Alternate models like multiphase, discrete phase, radiation, heat exchanger were impertinent and hence not taken into consideration. Air was the primary flowing fluid in the following simulation. Hence, the values of density and viscosity of air were inserted from the fluent database. As the whole model comprised of 4 parts, i.e., the inlet and outlet side of the casing, the porous region, and the blower fan, distinct cell zone conditions were applied to each region. The inlet and outlet part of the casing were standard fluid regions with zero frame motions. The blower fan was taken as a solid zone with a certain rotational frame motion. The fan which was chosen had a rotational velocity of 2600 rpm which is equivalent to 272 rad/s. The porous media model is used to define a cell zone and the pressure loss in the flow is determined. The porous jump model is applied to a face zone, not to a cell zone for robust and yields better convergence. To insert accurate coefficients for the porous region, a pressure drop versus face velocity graph of a HEPA filter was used. For a porous cylinder, values 0, 0, 1 were entered as the axis vector, and 0, 0, 0 were entered for the cone axis since the frustum has its axis parallel to the z axis. The cone angle of the cylinder was taken as 0° as it was a perfect cylinder. Furthermore, directions 1, 2 and 3 were entered accordingly. For the Conical option Direction 1 is tangential, i.e., along its length, direction 2 is normal to the cone surface and direction 3 is circumferential. A pressure inlet and a pressure outlet were taken into account for the simulation. The fan was the only solid zone in the CFD model, whereas all other sections were considered to be fluid zones along with the porous zone of the HEPA filter. The duct wall was construed as adiabatic involving no heat or mass transfer between the control volume and the surrounding.

Solution methods, residuals, initialization and running calculations: Simple scheme method was chosen for the analysis. Least-squares cell-based gradient was incorporated as it proved to be the most efficient, while dealing with similar fluid flow problems. Second-order upwind was taken into consideration for the turbulent kinetic energy and turbulent dissipation rate as well as the pressure and momentum. Residual with a precision of 0.0001 was considered to provide highly accurate and converged results. Hybrid Initialization was performed for a total of 10 iterations, and the comprehensive simulations was done for a total of 150 iterations to provide maximum accuracy and precision. The velocity vector and stream line of the HEPA filter is shown in Figs. 3 and 4, respectively. The velocity streamline of model is compared with three different speed of exhaust fan. At 200 rad/s the fan is not effective in developing uniform flow across the purifier as the speed is low for this model, we increased this speed to 320 rad/s and there is turbulence created across purifier this resulted in increasing kinetic energy of the air molecules resulting in the loss.

Now, from the theoretical calculations we got 272 rad/s as boundary condition when plotting streamlines there is no overlapping of airflow as a result, we uniformity is maintained throughout the purifier and no turbulence is created. This shows that the losses are less and no rise in temperature inside the purifier.

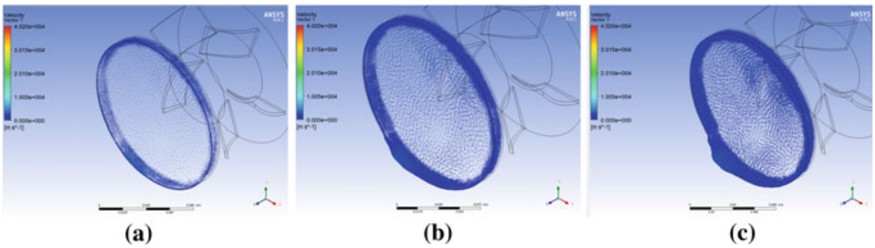


Fig. 3 HEPA filter velocity vector **a** 200 rad/s, **b** 272 rad/s, **c** 320 rad/s

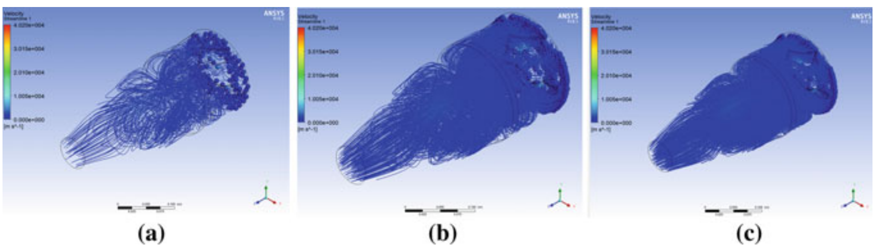


Fig. 4 HEPA filter velocity streamlines **a** 200 rad/s, **b** 320 rad/s, **c** 272 rad/s

3 IoT Based Air Purifier Prototype

IoT is an interrelated system consists of network-enabled smart devices that use embedded systems, such as processors, sensors and mechanical systems, to receive, transfer and act on data they acquire from the environments. IoT enabled devices share the sensor data they collect by connecting to a network where data is either sent to the cloud to be analyzed or analyzed locally [10]. IoT based automation of air purifier prototype requires various components including Arduino UNO, Inlet Fan, UVC Light, DHT11, Humidity Sensor, Bluetooth Module, 9v Battery which are represented in Fig. 5. After the circuit is assembled, Arduino coding for giving instructions to microcontroller to control the connected components in a specific manner is done. Provisions are made in the design to perform both purifier and humidifier functions simultaneously or individually as per need. A Bluetooth module in the system is designed to receive specific input from user's smartphone or any other smart devices to operate the prototype. Interface shown in Fig. 6 for the operation of device is designed using MIT App Inventor. It has various options to operate manually as well as automatically as per user requirements.

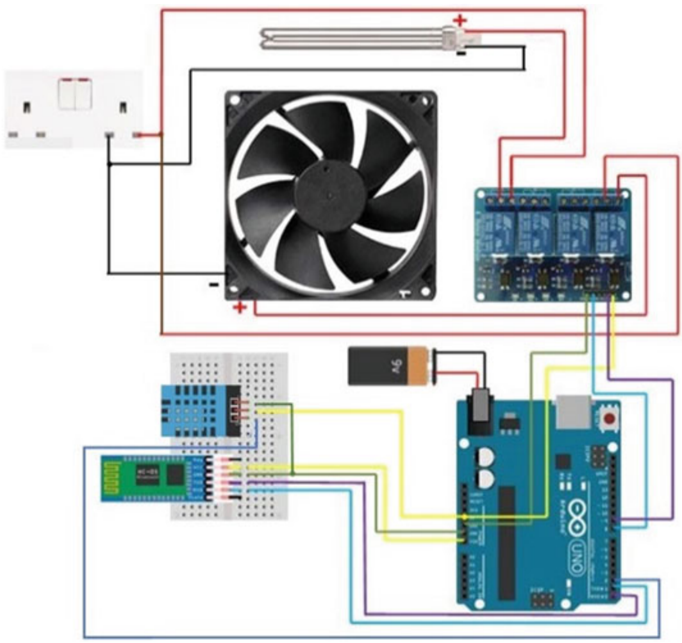


Fig. 5 Components of IoT based automatic system

Fig. 6 Operational interface

4 PLC Based Simulation of Industrial Air Purifier

The concept for industrial air purifier can serve as a centralized system in many industries like manufacturing, medical, office, shopping complex and theaters. The schematic representation of industrial air purifier is shown in Fig. 7. It operates as purifier with self-cleaning capabilities and controlled by any internet enabled smart devices. In purifying mode, fan act sucks air due to negative pressure created inside the intake manifold. Then, the air passing through three layers of filters (including Activated Carbon filter to remove both NO_2 and HONO gases) before it is accelerated

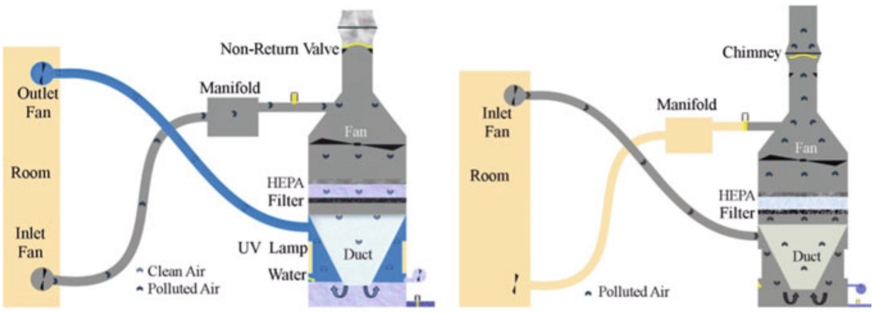


Fig. 7 Schematic representation of **a** purifying mode and **b** cleaning mode

in a duct to pass through the water bank and humidified [11]. Finally, filtered, odorless air suitable for the occupants is delivered to workplace. The cleaning mode is provided as a maintenance measure to clean the filter using reverse flow of the air. At the beginning of this process, an electronically actuated normally open valve is closed to prevent the polluted air from entering the workplace. Then, the fan sucks the air from the outlet part which is below the Filter bed. So, the sucked in Air passes through the filter and brings the particles along with its path. A gate hinged at the exhaust port opens up due to the air pressure. This is how the air will leave the purifier taking the filtered particles and hence cleaning it.

A programmable logic controller (PLC) is a ruggedized computer used for industrial automation. These controllers are capable to automate a specific process, machine functions or even entire production lines [12]. The automation studio software provides an excellent platform to design circuits and is used to do Ladder Logic Programming for PLC in this research work. Toggle switches are used as Human Machine Interface (HMI), while designing the circuits for PLCs.

Design of electrical circuits with components and the associated ladder logic program are shown in Figs. 8 and 9, respectively. The same is verified in the software itself using simulation function to ensure that all functional requirements are fulfilled. Table 2 highlights the labels and its functions used in PLC programs for simulation.

Simulation of the PLC automation for purification and cleaning process: When the purification toggle switch is activated, switch s_1 gets closed and current flows through OUT0, OUT1 and OUT4 which are corresponding to inlet fan located at inside the room, UVC lights as well as outlet fans inside the room and the main inlet fan inside the purifying unit. Thus, the purification process continues till the s_1 switch is opened by toggle switch. During operation, level switch s_2 keep monitoring water level in the humidifying chamber; get closed if below the specified level by passing current to OUT2. OUT2 is connected to a water pump to fill the water inside the

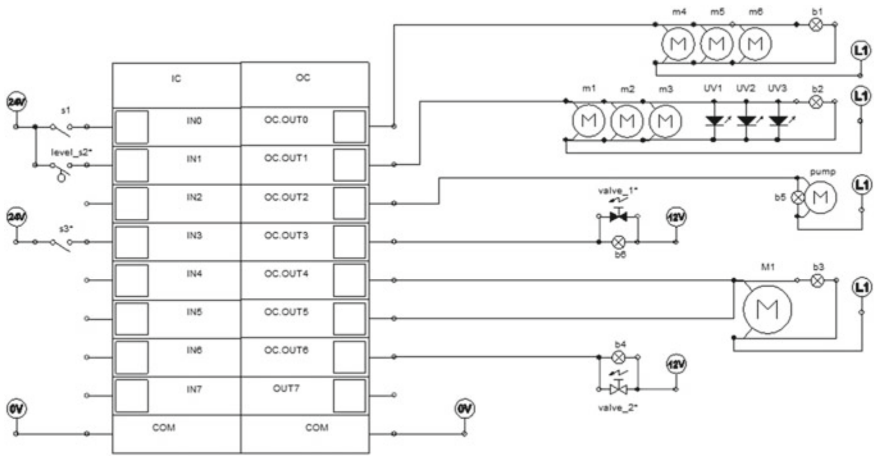


Fig. 8 PLC circuit diagram

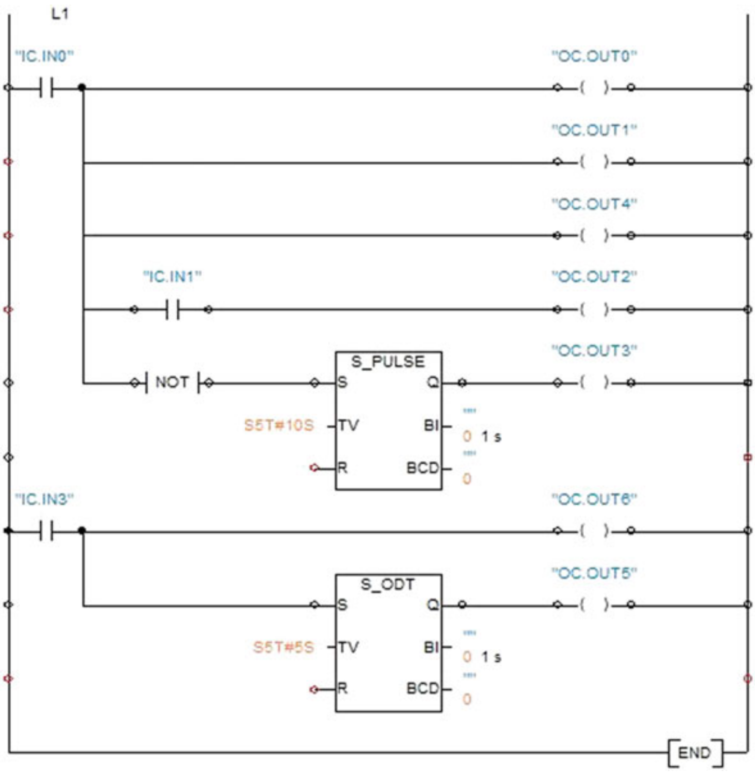


Fig. 9 Ladder diagram of PLC coding

Table 2 Defining the labels in PLC circuit

Labels	Components
m_1, m_2, m_3	Intake fan inside the room
m_4, m_5, m_6	Outlet fan inside room
UV1, UV2, UV3	UVC lights inside purifying unit
M1	Main inlet fan inside purifying unit
Pump	Pump to fill water inside purifying unit as humidifier
valve_1	Normally closed electronic valve
Valve_2	Normally open electronic valve
s_1	Initiate purifying
s_2	Level switch
s_3	Initiate cleaning

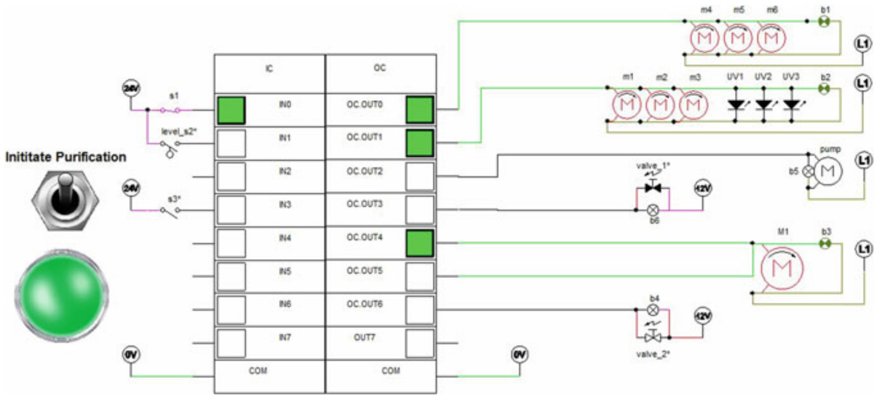


Fig. 10 Simulation of initiate purification command

humidifying chamber to allowable level so that level switch s_2 can again get open. When the users again toggle OFF the switch s_1, S_PULSE counter will get activated which will provide current to OUT3 for 10 s before completely opening the circuit. OUT3 is connected to electrically actuate normally closed valve which gets open to drain the existing water inside the chamber. The cleaning process is incorporated to facilitate cleaning of filters periodically per maintenance practice of the industry. To initiate this process, inlet valve located at inside the room is closed and main inlet fan inside the purifying unit starts rotate in opposite direction to force the pollutants from the filters to escape into atmosphere through chimney. When users activate toggle switch s_3, provides current to OUT6 and OUT5. OUT6 is connected to electrically actuated normally open valve will get closed which block the only inlet path to the room. OUT5 is connected to main inlet fan inside the purification unit, before switching on the fan in opposite direction S_ODT counter give 5 s to properly close the inlet valve. After, 5 s the fan will rotate in opposite direction to successfully force all pollutant away from filter. Figures 10 and 11 show the simulation results carried out using automation studio.

5 Conclusions

In this article, the concept, design and analysis of the prototype of a smart centralized air purifier system are presented. The simulation carried out in Ansys Fluent software ensures optimal design of air purifier for its intended use. IoT based operation of the purifier provide safe and efficient operation environment, Simulated operating conditions using Automation Studio facilitates industrial applications of centralized air purifier that eliminates air pollutants of 10 μm and above from the polluted air in the industrial environments. Thus an economical, smart and healthy air purifier is developed to suit industrial applications. The model is scalable and the effective

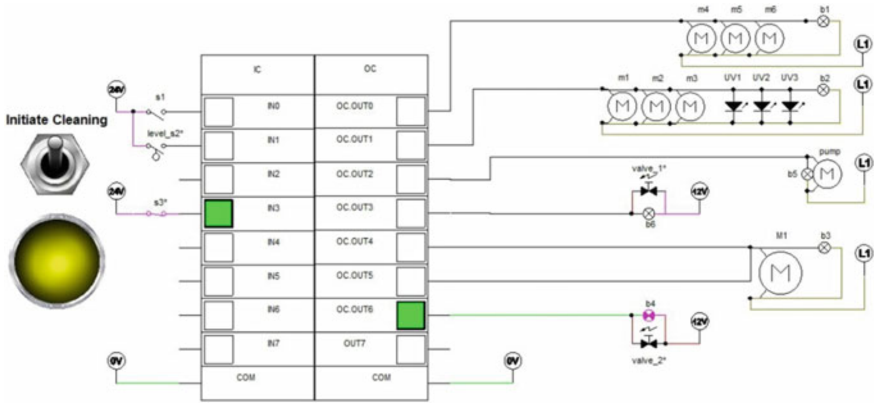


Fig. 11 Simulation of initiate cleaning command

area of the purifier can be altered with respect to specific industrial or large area air purifying requirements.

References

1. Anderson JO, Thundiyil JG, Stolbach A (2011) Clearing the air: a review of the effects of particulate matter air pollution on human health. J Med Toxicol 8:166–175
2. Soreanu G, Dixon M, Darlington A (2013) Botanical biofiltration of indoor gaseous pollutants—a mini-review. Chem Eng J 229:585–594
3. Gunnar DN, Lea FH, Peder W (1997) Chemical and biological evaluation of building material emissions, II. Approaches for setting indoor air standards or guidelines for chemicals. Indoor Air 7:17–32
4. Fang J, Zhu Z (2019) Material analysis and design of air purifier. In: IOP conference series: materials science and engineering 585, p 012107
5. Diette GB, McCormack MC, Hansel MN, Breysse PN, Matsui EC (2008) Environmental issues in managing asthma. Respir Care 53(5):602–617
6. Liu G, Xiao M, Zhang X, Gal C, Chen X, Liu L, Pan S, Wu J, Tang L, Clements-Croome D (2017) A review of air filtration technologies for sustainable and healthy building ventilation. Sustain Cities Soc 32:375–396
7. Matela D (2006) Air filtration: green and clean - how to improve indoor air quality. Filtr Sep 43:24–27
8. Sublett JL (2011) Effectiveness of air filters and air cleaners in allergic respiratory diseases: a review of the recent literature. Curr Allergy Asthma Rep 11:395–402
9. Mahalalakshmi G (2017) IOT based home automation using Arduino. Int J Eng Adv Res Technol 3(7)
10. Medical Advisory Secretariat (2005) Air cleaning technologies: an evidence-based analysis. Ontario Health Technol Assessment Ser 5:17
11. Yoo JY, Park CJ, Kim KY, Son Y-S, Kang C-M, Wolfson JM, Jung I-H, Lee S-J, Koutrakis P (2015) Development of an activated carbon filter to remove NO_2 and HONO in indoor air. J Hazard Mater 289:184–189
12. Hudedmani M, Umayal R, Kabberalli S, Hittalamani R (2017) Programmable logic controller (PLC) in automation. Adv J Grad Res 2(1):37–45

Rapid Forest Cover Detection Using Ensemble Learning

M. Mohammed Al Sameer, T. Prasanth, and R. Anuradha

Abstract In this competing environment, many different intruding factors interfere in the forest cover region. Statistics show that forests cover more than 30% of the earth's land surface, and these forested areas can provide food, medicine, and fuel for more than a billion people. Worldwide, forests provide 13.4 million people with jobs within the forest sector, and another 41 million people have jobs associated with forests. Detecting the forest cover type using conventional field measurements are time-consuming and resource intensive. The paper's aim is to use cartographic data to predict the forest cover type provided by the UCI machine learning repository and ease the whole situation rapidly through artificial intelligence. The model trained by examining through exploratory data analysis and implementing feature engineering techniques and ensemble learning methods improved the overall accuracy by 93% of the model compared to the existing model and the proposed methodology has given the most excellent way to find the forest cover type.

Keywords Feature engineering · Exploratory data analysis · Forest cover · Ensemble learning

1 Introduction

Many initiatives to expand the urban near the forest cover region or to expand and upgrade the rural near forest cover region face the same issue of forest cover getting deforested in an unethical manner. Growth is essential, but not at the cost of destruction. In some scenarios, deforestation is required, but it should be in a compensable

M. Mohammed Al Sameer (✉) · T. Prasanth · R. Anuradha
Sri Ramakrishna Engineering College, Coimbatore, India
e-mail: sameer.1601122@srec.ac.in

T. Prasanth
e-mail: prasanth.1601145@srec.ac.in

R. Anuradha
e-mail: anuradha.r@srec.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_15

or easily recoverable state for the forest. The approach to this scenario needs to be assertive, in which it creates a win–win situation that causes no harm on both sides. Using the conventional method requires more workforce and is time intensive. For an optimal solution, we approach this issue by using feature engineering and ensemble learning method from artificial intelligence.

Feature engineering [1] is the method which help our model by creating new input features from the existing dataset. Ensemble learning is a powerful machine learning method which has exhibited several advantages in various applications. It is the process by which multiple models combine several classifiers to produce better predictive performance than utilizing a single classifier. The main principal behind the ensemble model [2, 3] is that a bunch of weak learners come together to make a powerful learner.

1.1 Features and Data

The dataset used in this paper was taken from the machine learning repository [4] of UCI, and the test data are a stratified sample. The data set consisted of 30 m $\times$ 30 m patches of 15,120 samples of forest located in northern Colorado's Roosevelt National Forest. There are 54 attributes in each data sample which includes: Slope, elevation (in meters), slope, aspect (compass direction of slope face), the vertical distance and the horizontal distance from water, the intensity of sunlight at 9 am, 12 and 3 pm, 40 soil type and four wilderness area. Each sample was classified into one among seven forest cover types: Spruce/fir, lodgepole/shore pine, ponderosa, Cottonwood/Willow, Aspen, Douglas Fir, or Krummholz. In each of the seven cover types, there are 2160 samples classified. The training set (15,120 observations) contains both the cover type and the features. The test set contains only the features in which it has a total of 565,892 observations.

2 Existing Methodology

There are several approaches to deal with the forest cover type. The existing methodology gave us an idea of which method was successful and which is not provides us with a sense of what to choose and what not to choose, while handling the forest cover type dataset [4]. In this paper [5], they have examined the ability of an ANN [Artificial Neutral Network] model and discriminant analysis [DA] to predict forest cover type, which uses the same forest cover type dataset.

In comparison, two alternative techniques for predicting forest cover types from cartographic variables. The results of the comparison indicated that a feed-forward artificial neural network model more accurately predicted forest cover type than a traditional statistical model based on Gaussian discriminant analysis. During this process, ANN model has its own drawback it had over-fitting issue which decreased the performance of the model, but performed well compared with Gaussian discriminant analysis. Classification accuracies produced by each model, as calculated from the test data set, were also examined. The Gaussian model's predictions for the forest cover type produced an 58.38% and ANN model's predictions for the forest cover type produced an overall classification accuracy of 70.58%. In this methodology, they have examined the ability of principal component analysis, [PCA] multi-class support vector machine and K-means clustering to predict the forest cover type, which also uses the forest cover type dataset.

In comparison, these techniques for predicting forest cover types from cartographic variables. The results of the comparison indicated that a principal component analysis performed dimensionality reduction, which improved the performance of the model multi-class SVM [Support Vector Machine] more accurately predicted forest cover type than the K-means Clustering. While using PCA reducing the dimensions resulted in loss of variance. Classification accuracies produced by each model, as calculated from the test data set, were also examined. The SVM model's predictions for the forest cover type produced an overall accuracy of 78.64%. Ensemble learning [6–8] is a method by which multiple models combine several classifiers to produce better predictive performance than utilizing a single classifier.

3 Proposed Methodology

The processing of the proposed methodology is given Fig. 1. The dataset used in the proposed methodology is preprocessed using exploratory data analysis (EDA) and feature engineering for improvement. We classify the data and train the features to validate and fit extra tree classifier which helps us to remove correlation features. Principal component analysis (PCA)is performed for reducing the dimensions and feature scaling helps to normalize the distribution of the data and speeds up the calculation of the algorithm. This will help us to recognize nature of the algorithm and identify the strong and weak feature.

During this estimation, we can predict the way for approaching these classifiers for a better outcome by contacting them through ensemble methods to combine the weak and strong learners and by using sequential (Adaboost) and parallel (Random Forest) ensemble method the weak and the strong learners are combined. Finally, using extreme gradient boost algorithm increases the speed of the execution and performance of the model by providing an accuracy of 93%.

3.1 Proposed Module Description

In the rapid forest cover detection, detecting the forest cover type is split into different modules which include the following:

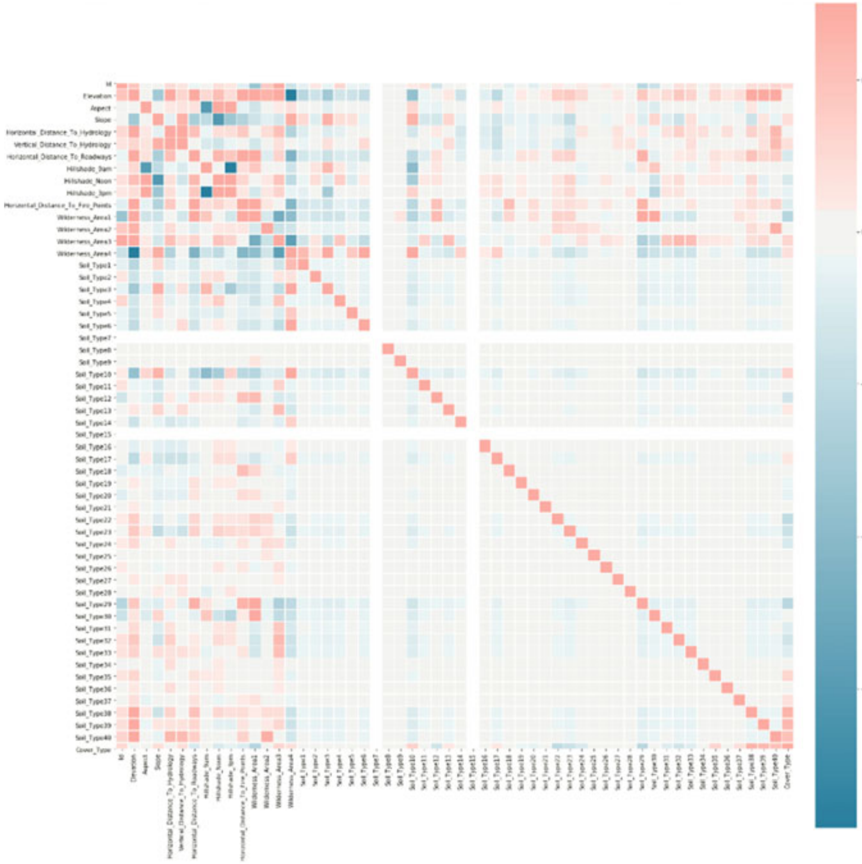


Fig. 1 Heatmap for data visualization

- Import and analyze the dataset.
- Perform feature engineering.
- Train the dataset and apply classification algorithm.
- Perform PCA and Feature scaling.
- Perform Ensemble learning and XGBoost algorithm.

Module 1—Import and Analyze the Dataset

The dataset is collected to classify them into different objectives. The initial step is to collect forest cover type data and analyse their characteristics. After analysing the authenticity of the dataset, then we used it as the trained data. Exploratory data analysis (EDA) is an approach to analysing data sets to summarize their main characteristics, often with visual methods like statistical visualization for analysing the data. Exploratory data visualization is implemented through seaborn and matplotlib using sns and plot function for the statistical data visualization for the Heatmap of

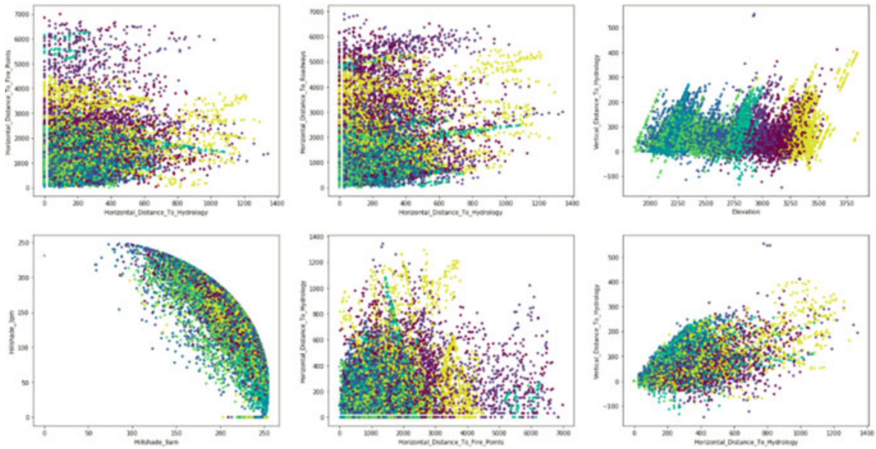


Fig. 2 Plot visualization of data fields for correlation between features

the data to understand (Fig. 1) shows which data is missing and which data is strong or weak. Validating the dataset gives us an idea of what is missing and what needs to be focused.

Matplotlib uses plot function for the statistical data visualization of data getting plotted. The Matplot (Fig. 2) has a visualization of plots for the features with correlation from the datasets.

The vertical distance contains some negative numbers and it gave better performance when taken the absolute for the column.

Module 2—Perform Feature Engineering

In many cases, machine learning algorithms do not perform well without feature engineering, which is the process of filling NaNs (Not a Number), missing values and creating new features. Feature engineering is used here because of its method of using the domain knowledge to extract the features from raw data via data mining techniques. In turn, it helps us to explore forest cover type data and be able to know which features to use and which new features need to be incorporated.

Module 3—Train the Dataset and Apply Classification Algorithm

Training the forest cover type dataset makes it possible to predict the probability of approximate occurrence. While training the dataset, the data points in each class are almost balanced, so using accuracy as a metric to measure yields a better prediction with the help of a classification algorithm. Applying the classification algorithm for the forest cover type dataset [4] with different initial conditions will result in different models to an extent for analysing the other combining strategies to avoid the generalization error and examine the variance. So, by using an extra-trees classifier which implements a meta estimator that fits several extra-trees on various subsamples of the forest cover type dataset and uses averaging to improve predictive accuracy and control over-fitting. Parameter tuning is done using GridsearchCV (Cross Validation) instead of manual tuning.

Module 4—Perform PCA and Feature Scaling
PCA (Principal Component Analysis) is used to perform Linear dimensionality reduction on the input which is fed as Cartographic variables or Integers to predict the particular forest cover type. Using singular value decomposition of the forest cover type data to project it to a lower-dimensional space and remove the linear correlation across features because PCA helps us to overcome the unnecessary features in the dataset.

Scaled for each feature but the input data is centered, but implementing feature scaling on the forest cover type dataset that standardizes the approach of extracting the features that are considerate of their independency in a fixed range of features. Standard scaler assumes that the forest cover type data is normally distributed within each feature and will scale them such that the distribution is now centered around 0, with a standard deviation of 1.

Module 5—Perform Ensemble Learning and XGBoost Algorithm
The classifier of AdaBoost is an meta-estimator that initiate by fitting a classifier on the forest cover type dataset then fits a further copy of the classifier on an equivalent forest cover type dataset but where the weights of the falsely classified instances are balanced in such a way that the subsequent classifiers focus more on severe cases.

Random forest algorithm is performed on this model for providing higher accuracy. Random forest classifier is the meta estimator which fits several decision tree classifiers on the forest cover type dataset. It uses averaging to control over-fitting and also improves the predictive accuracy.

Voting classifier is a machine learning algorithm which is implemented on our model since it does the ensemble of the several models and predicts the output based on their highest probability of chosen class as the output. It only aggregates the findings of each classifier passed into Voting Classifier and predicts the output class based on the highest majority of voting.

Gradient boosting algorithm is implemented for boosting the model by adding predictors sequentially to an ensemble, each one correcting its predecessor. Instead of altering the weights for all falsely classified observation at every iteration like AdaBoost. The gradient boosting method fixes the residual errors made by the previous predictor by fitting the new predictor.

Gradient boosting algorithm is used to combine the weak learning models to make it a strong predictive model on the forest cover type dataset.

XGBoost algorithm which is the implementation of the gradient boosted decision trees created for performance and faster execution. Gradient boosting algorithms are slow in implementation because its training model is sequential. So, it has scalability issues. XGBoost is implemented in our model because it is determined on performance of the model and its computational speed.

4 Results

The results were taken using hardware configuration of x86-64bit processor and packages used where NumPy, Pandas, Scikit learn, matplotlib, Seaborn and XGBoost. Rapid forest cover detection follows a sequential process of training and evaluating by focusing on which part needs more priority and which needs less priority.

Tree-based models only fits vertical and horizontal lines so to train some features like slope, distance, elevation, Hill shade are important. When using extra-trees classifier to make the model converge faster, feature scaling was performed. Parameter tuning is done using GridsearchCV and Performed Ensemble Learning. Figure 3 shows the comparison between the proposed methodology and existing methodology.

A number of points in the classes is balanced. If the number of points in the classes is not balanced cannot use accuracy as a performance metric. The number of occurrences in the data points in each class is almost balanced so using accuracy as shown in Fig. 4 as a metric to measure the performance will be a better option.

Though accuracy may be the best metric to measure, but using confusion matrix in our rapid forest cover detection helps us to explain the performance of its classification model on a group of test data that the true positive and negative values are known.

The color intensity indicates the probability of each element in a row which is stated in Fig. 5. Confusion matrix for seven forest cover types.

In Table 1, it states the precision which gives the ratio of positively predicted values from the entire predicted positive values. The recall gives the ratio of precisely predicted positive values to all the values in the actual class true, f1-score ranges from 86 to 97 percentage for each forest cover type and it is the weighted average of both

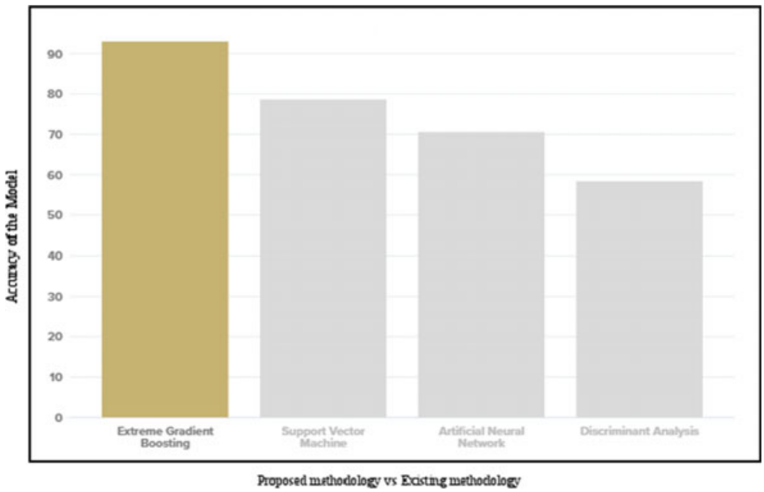


Fig. 3 Comparison of proposed methodology versus existing methodology

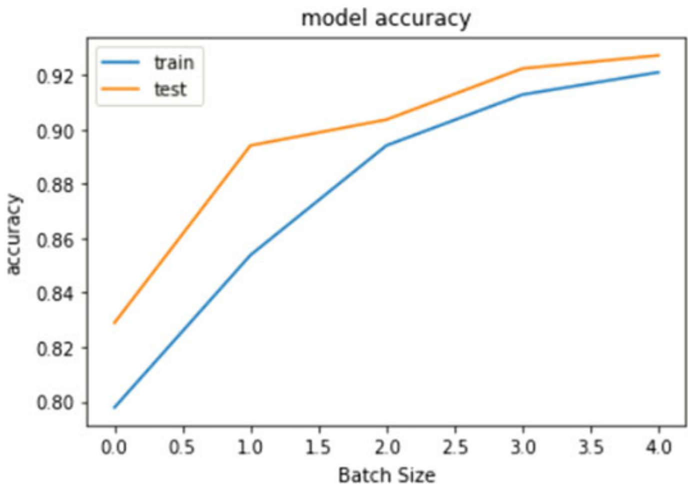
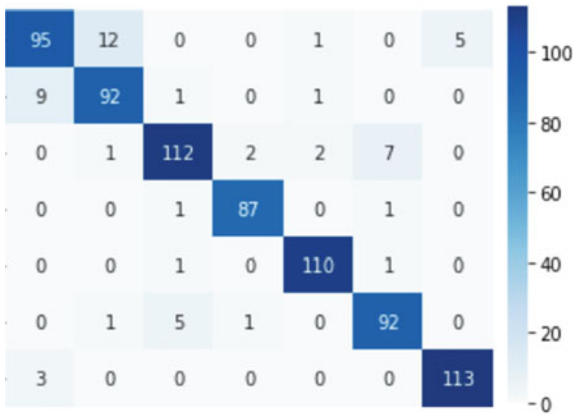


Fig. 4 Model accuracy of rapid forest cover detection

Fig. 5 Confusion matrix of rapid forest cover detection



precision and recall which gives us an average of 93% and the support achieved by each forest cover type.

5 Conclusions and Future Enhancements

Conclusion

The algorithms and classifiers are performed well on both training and testing data. The head start gave by feature engineering improved the dataset and provided us with strong foundation. The Ensemble model surpassed other models used in studies [5,

Table 1 Experimental result of rapid forest cover detection

Cover type	Precision	Recall	f1-score	Support
1	0.89	0.84	0.86	113
2	0.87	0.89	0.88	103
3	0.93	0.90	0.92	124
4	0.97	0.98	0.97	89
5	0.96	0.93	0.92	112
6	0.91	0.97	0.97	99
7	0.96			116
Accuracy			0.93	756
Macro average	0.93	0.93	0.93	756
Weighted average	0.93	0.93	0.93	756

9] having the same data set [4]. These results show that an optimized, and ensemble model is a accurate model for the classification of forest type. Similarly, our extra tree classifier and random forest tree classifier performed equally well in both training and test data. This helps us to conclude that the Rapid forest cover detection was not struggling with over-fitting and performance which yielded 93% accuracy. Thus, the rapid forest cover detection provides an optimal solution for detecting forest cover type using the ensemble learning model.

Future Enhancements

Rapid forest cover detection can be extended for monitoring changes in the forest cover such as decision-making, forest planning and management, climate change studies and wildlife habitat. The more dataset in the forest cover type and other related datasets will help yielding more accuracy than produced and can cover more domains which have similar data fields and helps to ease the situation.

Acknowledgements We would like to thank Sri Ramakrishna Engineering College for providing us resources for deploying the project.

References

1. Sharma S, Jain A (2020) Hybrid ensemble learning with feature selection for sentiment classification in social media 10:40–58. https://doi.org/10.4018/IJIRR.2020040103
2. Mamudur K, Kattamuri MR (2020) Application of boosting-based ensemble learning method for the prediction of compression index. J Inst Eng (India): Ser A https://doi.org/10.1007/s40030-020-00443-7
3. Yu H, Ni J (2014) An improved ensemble learning method for classifying high-dimensional and imbalanced biomedicine data. IEEE/ACM Trans Comput Biol Bioinform 11(4)
4. UCI Machine Learning Repository—forest cover type dataset acquired from http://archive.ics.uci.edu/ml//datasets/Covertype.

5. Blackard JA, Dean DJ (1999) Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. Comput Electron Agric
6. Pandey M, Taruna S (2014) A comparative study of ensemble methods for students' performance modeling. Int J Comput Appl 103(8):0975–8887
7. Graczyk M, Lasota T, Trawiński B, Trawiński K (2010) Comparison of bagging, boosting and stacking ensembles applied to real estate appraisal. In: Nguyen NT, Le MT, Świątek J (eds) ACIIDS 2010 , Part II, LNAI 5991. Springer, Berlin Heidelberg, pp 340–350
8. Sun S (2007) Ensemble learning methods for classifying EEG signals. In: Haindl M, Kittler J, Roli F (eds) MCS 2007, LNCS 4472. Springer, Berlin Heidelberg, pp 113–120
9. Crain K, Davis G (2014) Stanford University—CS 229 classifying forest cover type using cartographic features machine learning

Analysis of Connectivity Model to Study the Neurophysiological Process for Autism Detection

Naimish P. Mehta, R. Menaka, R. Karthik, and M. Thanga Aarthy

Abstract Studying network communication within the neurons is the next step toward exploring the complexities of the human brain. In this research, we used electroencephalography (EEG) to study the neurophysiological processes because of their high temporal and spatial resolution. EEG stands out to be a vital modality in assessing patients with brain abnormalities like Autism, Epilepsy, Dementia, and Parkinson's. Nowadays, a large number of children worldwide is affected by Autism Spectrum Disorder, which impairs the ability to communicate and behave. In this research, we generate the connectivity models using the EEG signal dataset of Autism and normal children. Here, connectivity models are presented into a graphical form using different measures like phase synchronization, classical measures, granger causality, and information theory. These parameters were used to analyze the variation between Autistic and typically developed children.

Keywords Autism · EEG · Connectivity model

1 Introduction

In the recent years, researchers have showed their interest in neurodevelopmental disability due to its complexity in identifying a promising biomarker. An exceptional objective in the field is to recognize typically developing subjects and children with neuro disabilities. Relationship of human being with others and the external environment is influenced through the understanding of expressions [1]. Autism spectrum

N. P. Mehta · M. Thanga Aarthy
School of Electronics Engineering, Vellore Institute of Technology, Chennai, India
e-mail: thanga.aarthy@vit.ac.in

R. Menaka (✉) · R. Karthik
Centre for Cyber Physical Systems, Vellore Institute of Technology, Chennai, India
e-mail: menaka.r@vit.ac.in

R. Karthik
e-mail: r.karthik@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_16

disorder (ASD) refers to a wider range of conditions including social skill problems, repetitive behaviors, speech, and nonverbal communication. EEG is a promising modality for contemplating neuro physiological procedures. The global prevalence rate of ASD is estimated as 1–2% in children [2]. There is a need for low-cost tool for screening children with ASD in low and middle-income countries. Diagnosis of ASD at an early age is challenging with the screening techniques available at present. This demands a distinguishing feature for children with ASD that could lead to a clarity in diagnosis and also helpful for early intervention for ASD children. An association between brain's functional connectivity and the cognitive performance leads a way to the in-depth understanding of the neurodevelopmental disorder [3]. The cognitive behavior of the brain can be well identified by EEG. Quicker advancements in clinical instruments and structures have prompted the exact and non-intrusive estimation of the mind's electrical action. EEG is considered as a potential tool for numerous neurodevelopmental studies due to its high-temporal resolution and accurate brain's response to various stimuli. The neuronal interactions in the brain are depicted as wave patterns in EEG.

The connectivity in the brain network is categorized into auxiliary, functional (FC), and effective connectivity (EC) [4–10]. Anatomical network (AN): Neighboring neurons are connected through the synaptic contacts. Functional connectivity (FC): Which is characterized as the worldly reliance of neuronal actuation examples of anatomically isolated mind areas that depends on such correlation, covariance, or phase locking [7–9]. Effective connectivity (EC): The influence of one neuron framework will depend upon another is called EC. Granger causality is indicator of effective connectivity [6].

The connection of brain regions and association among them can be represented by the connectivity model [1–3, 11]. Through this connectivity model, the investigation to identify and describe the dynamic connection and interaction among the brain region can be explored [12]. Parameters like Granger Causality, Phase synchronization, Information Theory, and Classic Measure will give various measures of dynamic behavior of the brain in different EEG channels [4]. Analyzing the brain pattern and parameters will help to develop a mathematical model to describe the difference between brain activities of autistic children and normal children. Here in this work, we have used HERMES toolbox to explore this possibilities [13]. The information flow in the human brain is analyzed through the connectivity model. Various experiments have been made to arrive different pattern and is found that proposed method is a good fit to distinguish between autism one and typical children.

This paper has been organized as: Sect. 2 gives a concise prolog to our proposed work. In Sect. 3, the experimentation technique and results are discussed. In Sect. 4, the concluding remarks and scope for future work are presented.

Fig. 1 Block diagram of proposed methodology

2 Proposed Model

This research work proposes an approach to identify the patterns present in the EEG signals and potentially detect the presence of ASD. The recognition of these patterns is conditioned on the connectivity diagram generated by HERMES toolkit. Figure 1 entitles the high-level schematic diagram of the proposed research work. Our design utilizes the information from five different parameters to fuse the time series EEG signals into the connectivity model and identify the level with corresponding the presence of Autism in a subject.

3 Materials and Methods

3.1 Electroencephalogram (EEG)

An EEG can record and track the brain wave patterns. Small electrodes are appended to the scalp with wires [1]. These electrodes investigate the electrical driving forces in the mind and impart signs to a computer that records the outcomes. The analysis and understanding of the relation between time series has gained significant traction in the recent years for EEG signals. Our research utilizes an open source HERMES toolkit for analyzing the time series data to effectively analyse the brain connectivity modeling for neurological data (e.g., MEG, EEG, iEEG) [10].

Six parameters are analyzed in this work namely, Granger causality, PS records ("Phase Synchronization Indexes"), Classical measures, Information theoretic files ("Data Theoretic Measures"), and GS indexes ("General Synchronization Indexes").

3.2 Data Acquisition

EEG data from 3 typically developing children and 3 children with ASD were acquired through Nihon Kohden MEB9000 with sensitivity of 7 μv. With the sampling frequency of 500 Hz, 21 channel signal are recorded and 13,000 samples of EEG signals are acquired and filtered with low pass channel & high pass channel at a range of 0.53–70 Hz frequencies. These electrodes placements are Fp1, F7, T3,

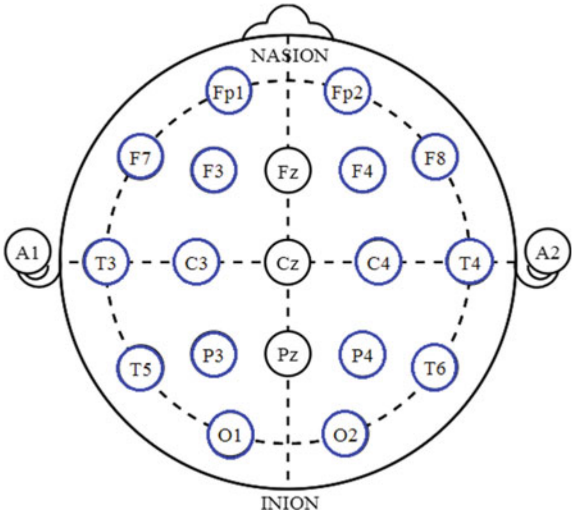


Fig. 2 Electrode positions applied for this research [10–20 system]

T5, F3, C3, P3, O1, Fp2, F8, T4, T6, F4, C4, P4, O2, Fz, Cz, Pz, A1, A2 (ch21) as shown in Fig. 2.

The children are made to sit and watch a cartoon video for 10 min. The cartoon is played with audio and without audio. The children are exposed to rhymes alone for 10 min. The response of the children is recorded in the EEG. This signal is preprocessed for eye blink removal, power line artifact removal.

Six different data sets are available for analysis as shown below.

1. EEG recording with only rhymes for Autism children and Normal children.
2. EEG recording while viewing cartoon video for Autism children and Normal children.
3. EEG recording with combined audio and video for both Autism children and Normal children.

Using HERMES tool, connectivity model is generated with the above mentioned signals. The parameters as directionality phase index, Grangner causality, phase locking value, correlation coefficient, and Mutual information are calculated from the connectivity model. The different parameters are explained below.

3.3 Directionality Phase Index (DPI)

Directionality measure among two signs $X(t)$ and $Y(t)$ can be acquired by the addition of the periodic elements of the two phases of the signals. This is measured by DPI parameter [10]. Three methods available to calculate directionality phase index are EMA (Evolution map approach), instantaneous period approach (IPA) and information theoretic approach (ITA). The DPI measure by ITA approach is expressed by

Eq. (1).

$$d^{xy} = \frac{C_x - C_y}{C_x + C_y} \tag{1}$$

Where, C_x = cross reliance of the elements of $X(t)$ on the period of signal $Y(t)$.
C_y = cross reliance of the elements of $Y(t)$ on the period of signal $X(t)$.
This cross coupling records C are then calculated for a group of substitutes and associate with the original information.

3.4 Granger Causality (GC)

Granger causality is a measure of causality that depends on expectation. Granger causality, predicts the value of a signal X_2 from the past values of a signal X_1, if X_1 contain some data that to predicts X_2 well beyond the data contained in past estimations of X_2 alone [3]. Granger causality depends on linear auto-regression modeling. Granger causality (GC) predicting x from y is given in below Eq. 2:

$$GC_{y \to x} = \ln\left(\frac{V_{x|\overline{x}}}{V_{x|\overline{x,y}}}\right) \tag{2}$$

Where, $V_{x|\overline{x}} = \mathrm{var}(U_x)$ and $V_{x|\overline{x,y}} = \mathrm{var}(U_{xy})$, where var() is the difference after some time and where $x|x,y$ is the gage of signal $x(t)$ by the past example of estimations of signals $x(t)$ and $y(t)$. Range of GC is from 0 to ∞ and if GC is equal to 0 then the $Y(t)$ does not improve the prediction from $X(t)$ and also, for more than 0 the past of $Y(t)$ is improves the forecast of signal $X(t)$.

3.5 Phase Locking Value (PLV)

The PLV evaluates how the relative stage is set over the unit circle. When the relative stage includes a little piece of the circle and the PLV is almost 1 then we say that there is a strong PS among X and Y. [10]. PLV is an estimation that can used to look at task-actuated changes in long-extend synchronization of neural development from the EEG data. Equation 3 is used to find the PLV value.

$$\mathrm{PLV} = \left|\mathrm{e}^{i\,\Delta\emptyset_{\mathrm{rel}}(t_n)}\right| = \left|\frac{1}{N}\sum_{n=1}^{N}\mathrm{e}^{i\,\Delta\emptyset_{\mathrm{rel}}(t_n)}\right| \tag{3}$$

RANGE: $0 \leq \mathrm{PLV} \leq 1$ here 0 for unsynchronized frameworks. For example, this spread has two high values which are differentiated by π and 1 if the condition of demanding phase locking is aggregated then phase distinction is consistent, and subsequently called total phase synchronization.

3.6 Pearson's Correlation

Pearson's correlation is a measurement that estimates direct association between two variables $X(t)$ and $Y(t)$ and that can be calculated from Eq. 4. RANGE: $-1 \leq$ Rxy ≤ 1_-1 for reverse connection among $X(t)$ and $Y(t)$, 0 for no straight relationship and 1 for complete straight connection between the $X(t)$ and $Y(t)$.

$$R_{xy} = \frac{1}{N} \sum_{k=1}^{N} x(k)y(k) \tag{4}$$

3.7 Transfer Entropy (TE)

Transfer entropy is estimating the measure of coordinated move of information between two irregular Processes.

$$T_{X \to Y} = \sum_{y_{t+1}, y_t^n, x_t^m} p(y_{t+1}|y_t^n, x_t^m) \log\left(\frac{p(y_{t+1}|y_t^n, x_t^m)}{p(y_{t+1}|y_t^n)}\right) \tag{5}$$

Equation 5 shows the measure of data stream from $X(t)$ to $Y(t)$. This parameter is theoretically the same as the possibility of Granger Causality. The range of TE is from 0 to ∞, here on the off chance that 0, at that point there is no causality among $X(t)$ and $Y(t)$ and if greater than 0 then $X(t)$ is 'causing' $Y(t)$.

3.8 Mutual Information (MI)

When two different variables is proportion by the common dependence between the two factors that is called mutual information [10]. Mutual information is used to identify high correlation between two channels and this value is calculated by Equation 6. The range of MI is from 0 to ∞ and if equal to 0 then $X(t)$ and $Y(t)$ both are totally autonomous and for greater than 0 $X(t)$ and $Y(t)$ signals are dependent on each other.

$$\mathrm{MI}_{xy} = \sum_{i} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \tag{6}$$

4 Result and Discussion

The objective of this examination was to show the potential of nonlinear characteristics of the EEG signal in discriminating the normal and autistic response for various stimuli. This will serve as an early identifier and for designing customized training for the children with autism. Figure 3 shows the 3 columns and 2 rows where three column represent three different stimuli like only video, only audio, and audio + video and rows are for normal children and autistic children.

DPI and PLV are phase synchronization parameter and these are used to get local and long rage connectivity. PLV is not strong against the nearness of normal sources. Correlation gives the linear relation between two signals it is consider under the classical measures parameter and it will be used to get low-level relationship between channels. TE and MI are the information theory measures where TE is the conditional MI and it gives the directed information flow from one signal to another, where MI is used to get mutual dependency as discussed earlier and main use of MI is gives the higher order relationship between two signal therefore it does not depends

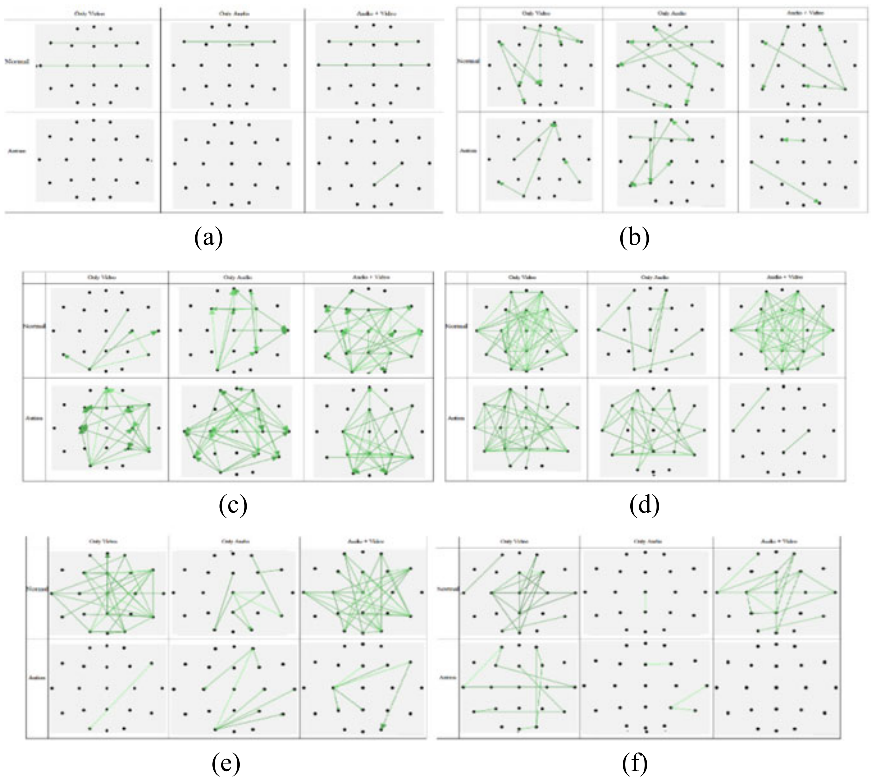


Fig. 3 These figures show the connectivity model of each measure **a** PLV, **b** GC, **c** PLV, **d** COR, **e** TE, **f** MI

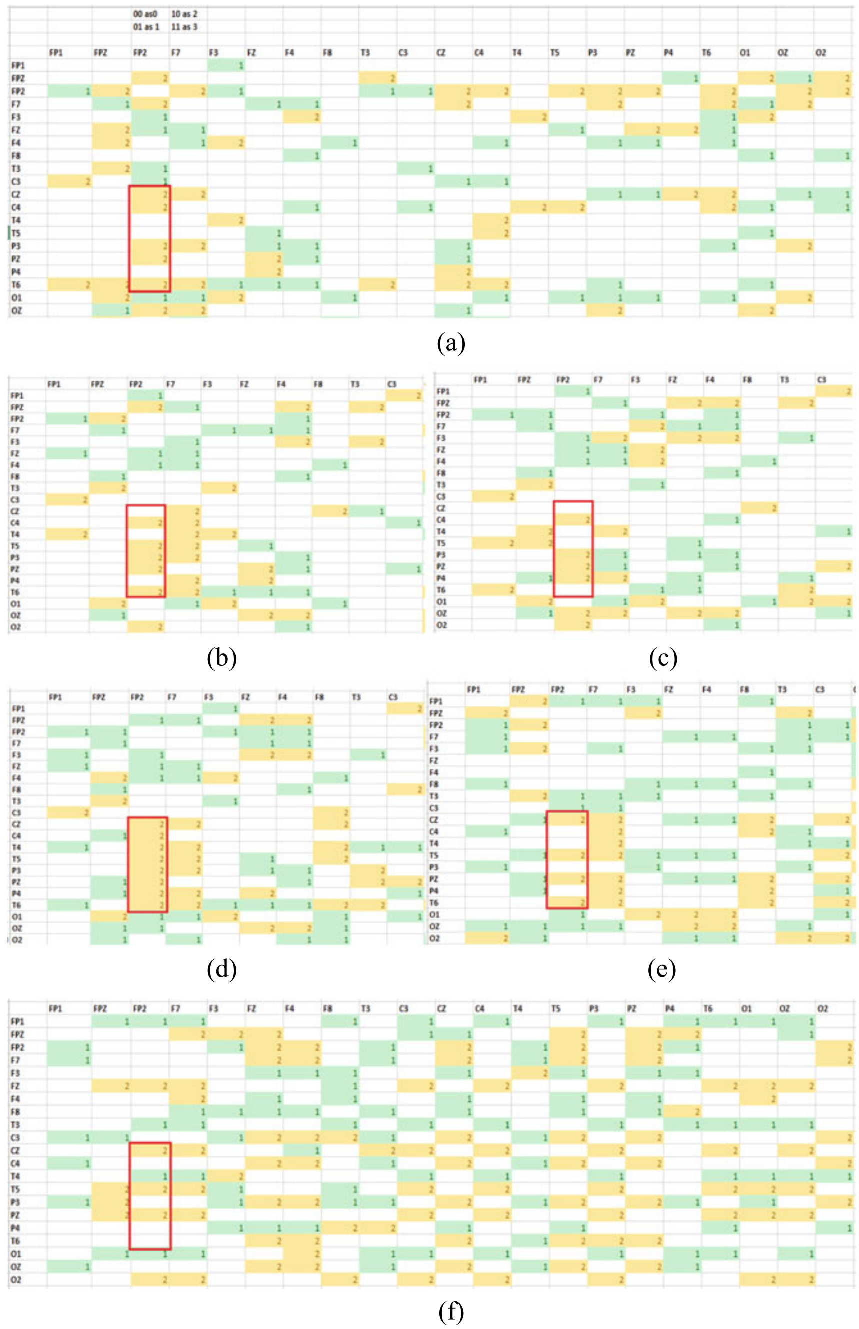


Fig. 4 **a–g** Shows the normalized averaged values of COR for each 21 × 21 channels

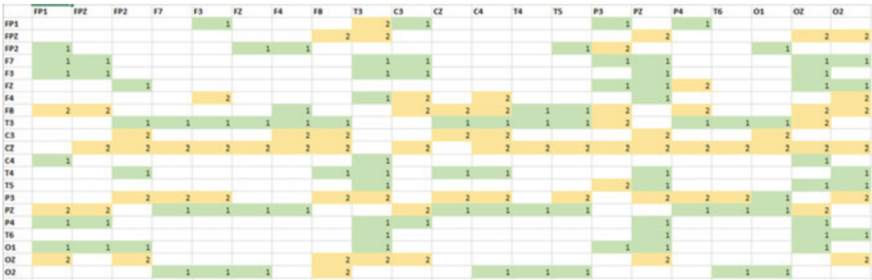


Fig. 5 Shows the normalized averaged value of MI for each 21 × 21 channels

on any specific model. The above connectivity diagrams are converted into weighted graphs. Dissimilar patterns between normal and autistic children are observed.

For pattern detection, we observed only audio + video stimuli with the different datasets to detect better patterns. Here, we have used 7 datasets for both autistic and normal children with the help of showing audio + video stimuli to 3 autism and 3 normal children. Then by applying Correlation connectivity measures to that dataset and we got the value of correlation for each specific 21 channel in the form of a 21 × 21 matrix. Average normalization is applied by getting the average value for each node then whichever has less average value is replaced with 0 and whichever has greater value is replaced by 1 and at the end, the last step is to combine both autism and normal values to distinguish their differences. Here, Fig. 4 shows the combined data which has value for 00 as 0, 01 as 1, 10 as 2, and 11 as 3, and here value 3 and 0 is removed because both autism and normal has the same relationship where 1 and 2 have differences.

The yellow-colored blocks having the value '2' means '10' and green-colored blocks having the value '1' means '01'. In this binary pattern, 1 represents the greater value than average and 0 represents the lesser than average. In Fig. 4, we have inferred that FP2 node has all the connectivity with the parietal cortex so it represents that dense network is present around the prefrontal cortex. These correlation gives the linear relationship and it is also depends on the position and distance between two nodes of scalp. Figure 5 shows the value of mutual information for 21 × 21 channels and in MI here no any such common pattern like correlation is observed so here no possibility to focus any particular node.

5 Conclusion

There are many researches trending toward ASD, which can be useful and provides a broad knowledge about the pathological condition. At present, there is no known cure for ASD, but the level of the pathological condition can be reduced when it is detected early. Early detection of this disorder is complicated and challenging till date. In this work, the connectivity model is transformed into a graphical representation and different objective measures like Phase synchronization, Classic Measure, Granger Causality, and Information Theory-based parameters were extracted from it. These parameters were examined to study the variation between autistic and typically developing controls. Higher correlation value in autistic children represents that similar activities are present in brain and less value represents the independent activities. The behavioral, planning and communication skill will more affected by the frontal part of brain and that is clarify in this research. Also, we conclude that parietal part of brain is more useful to analyze connectivity because of low noise into the central brain region and it gives more accurate connectivity than other regions. Analysis with more parameters will lead to better results in the early identification of autism.

References

1. Habeeb E, Ghazal N, Majzoub S (2019) Behavior analysis tool for autistic children using EEG signals. In: 2019 advances in science and engineering technology international conferences (ASET). IEEE, pp 1–5
2. Lang EW, Tomé AM, Keck IR, Górriz-Sáez JM, Puntonet CG (2012) Brain connectivity analysis: A short survey. Comput Intell Neurosci
3. Zhang J, Xu R, Belkacem AN, Shin D, Wang K, Wang Z, Chen C (2019) Brain network analysis of hand motor execution and imagination based on Granger causality. In: 2019 IEEE MTT-S international microwave biomedical conference (IMBioC), vol 1. IEEE, pp 1–4
4. Brunner C, Billinger M, Seeber M, Mullen TR, Makeig S (2016) Volume conduction influences scalp-based connectivity estimates. Front Comput Neurosci 10:121
5. Van de Steen F, Faes L, Karahan E, Songsiri J, Valdes-Sosa PA, Marinazzo D (2019) Critical comments on EEG sensor space dynamical connectivity analysis. Brain Topogr 32(4):643–654
6. Tyng CM, Amin HU, Saad NM, Abdul Rahman M, Malik AS, Tang TB (2019) Exploring EEG effective connectivity network in estimating influence of color on emotion and memory. Front Neuroinform 13:66
7. Bastos AM, Schoffelen JM (2016) A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. Front Syst Neurosci 9:175
8. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10(3):186–198
9. Friston KJ (2011) Functional and effective connectivity: a review. Brain Connectivity 1(1):13–36
10. Niso G, Bruña R, Pereda E, Gutiérrez R, Bajo R, Maestú F, del-Pozo F (2013) HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity. Neuroinformatics 11(4):405–434
11. Ahmadi SMM, Mohajeri N, Soltanian-Zadeh H (2014) Connectivity abnormalities in autism spectrum disorder patients: a resting state fMRI study. In: 2014 22nd Iranian conference on electrical engineering (ICEE). IEEE, pp 1878–1882
12. Olejarczyk E, Marzetti L, Pizzella V, Zappasodi F (2017) Comparison of connectivity analyses for resting state EEG data. J Neural Eng 14(3):036017
13. Olejarczyk E, Jernajczyk W (2017) Graph-based analysis of brain connectivity in schizophrenia. PloS one 12(11)

An Image Processing Based American Sign Language Fingerspelling Interpreter

Gopireddy Sirisha, Parmar Harshdeepsinh, Hemanth Konduri, and Sathiya Narayanan

Abstract Owing to limitations such as high cost, discomfort, and connectivity issues, existing fingerspelling interpreters fail to present a feasible and comfortable medium to efficiently communicate with communities that rely on sign language that includes hand gestures, facial expressions and orientations, for communication. Biological measurements, based on aspects such as hand or arm movements and acoustic variations at wrist, are unique to an individual and require customized systems to be designed for many groups of people. However, developing a general system that would not only incorporate sign language translation, but also additional features would bring a revolutionary change in differently abled community. In this article, we propose an image processing based system with machine learning algorithms, implemented as a WebApp on Raspberry Pi, which is an efficient, cost-effective, and user-friendly bridge between communities that understand sign language and the communities that do not.

Keywords American sign language · Machine learning · Support vector machines · Feed forward neural networks · Genetic algorithm

1 Introduction

To develop a sustainable society, we need innovations. This proposed system aims to contribute to the high quality of life of future societies. We believe that social inclusion, cohesion, environmental sustainability, and economic prosperity are the attributes of a sustainable society. These attributes are closely related and need an integrated way to develop a coherent society. We are tackling an issue in the field of computer vision. This article focuses on the development of a sign language interpreter that is highly efficient and effective. The highlighting objectives of this model are creating a portable system for the differently abled people that would interpret american sign language (ASL) and performing user defined functions that

G. Sirisha · P. Harshdeepsinh · H. Konduri · S. Narayanan (✉)
School of Electronics Engineering, Vellore Institute of Technology, Chennai 600127, India
e-mail: sathiyanarayanan.s@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_17

would facilitate the usability of additional features. Sign languages use the visual-manual technique to convey the meaning. There are static and dynamic gestures that are determined by a particular configuration of fingers, thumb, and palm. Using advanced techniques to develop a human–computer interface will prove to be a vital milestone in the domain of interpreting a sign language. We are particularly focusing on developing an interpreter for american sign language, because it is predominant in differently abled community and often regarded as the most widely learned second language. We have analyzed various methods by going through the last ten years of research work in this field and have examined the models that translate the ASL signs. In this project, we intend to elaborate the system developed that is solely designed for the communication of differently abled community with the world.

2 Literature Review

The categorized methods are.

2.1 Data Glove

Arslan Arif et al. have designed an electronic hand glove that recognizes static gestures and dynamic motion of hand to give visual and audible output [1]. Authors in [2] propose the GesTALK system, a system that uses data glove to translate gestures into voice. It recognizes 24 signs with 90% accuracy. These ideas are expensive and demand glove with integrated hardware to be worn on hand, which feels uncomfortable.

2.2 Surface Electromyography (sEMG)

Surface electromyogram signals from the user forearm are processed to detect the signs. Nine features are considered with two channel sEMG signals which gave accuracy of 97.7% and only 9 ASL signs are detected [3].

2.3 Acoustic Measurements

Sensitive microphones placed on wrist are utilized to capture feeble sound signals that are produced due to the geometric changes that result from movement of fingers. This helps in identifying 36 signs with 80% accuracy [4].

Dependence on the biological aspects, which are not accurate and do not cover the entire range of alphabets, is expensive and demands customization. There is more to sign language, which includes facial expressions and orientations, than just hand gestures and the above proposed methodologies, even if enhanced further, cannot present a feasible and comfortable medium to efficiently communicate with differently abled community. Image processing looks promising in capturing all the required features in less clumsy, more cost efficient and comfortably feasible way.

2.4 Image-Processing

Using Smartphone and Android Application. Android application is used to capture the image with camera of smartphone and to send it to server, where it is classified in Python. From comparison, SVM is faster and uses less memory than MLP (Multi-layer perceptron) in categorizing 24 signs with 91% accuracy [5]. In [6], a new algorithm for features extraction is proposed that combines both K-Curvature and Convex hull algorithms to detect fingertip with high accuracy and can distinguish 37 signs with an accuracy of 94.32%. Certain android-based methods require uninterrupted connectivity as they rely on servers. Instead of sending the data to a server, local computations on recent powerful smart phones would eliminate the connectivity issue, and for complex computations that cannot be performed locally, especially the model-training task, can be done elsewhere and just the resultant classifier can be incorporated into the application so that the system functions independent of the server, except for the training of new gestures.

Using Customized Hardware. In [7], photo mixing device (PMD) ToF Camera is used to segment hand from background of an image using depth information. Random Projection and KPCA are used for dimension reduction and perform the classification in lower-dimension space. The time complexity is $O(n_2)$ during the training of dataset, so they have implemented random sampling and computed only for sampled points. A fivefold cross validation was done on the dataset and accuracy was 99.8%. (PMD ToF camera sensor board is very expensive).

In [8], a new approach was discussed for training large scale ASL applications called Parallel Hidden Markov Models in which training can be done independently and does not require different combinations of phonemes during training. A new algorithm was discussed called token passing algorithm for sign recognition and it can carry additional path information.

2.5 Radar System

Authors in [9] used Soli, a solid-state short-range radar, and proposed a novel algorithm to recognize 11 gestures, based on the energy reflected by hands, with 87%

accuracy. For now, image processing differentiates more signs with better accuracy than this approach.

After carefully examining almost all those approaches that translate the american sign language, we found that Image processing is the best. Image processing with smartphone and android application requires uninterrupted connectivity (Including the training of model). Customized hardware seems to be a suitable method for our goal.

3 Proposed Model

The objective of the presented model is to interpret the american sign language sign and present its respective English language alphabet. Using complex sensors such as Kinect and PMD ToF camera, covers the whole range of features the sign language has and improves accuracy, but is expensive. Though there are many advanced machine learning models, support vector machine (SVM) works better with low-processing power that we can afford. Raspberry Pi seems to be a perfect choice with its decent processing power and wide range of compatibility features at an affordable range. As shown in Fig. 1, the system block diagram shows how the hand gestures are captured and processed using Raspberry Pi 3 B+ and Pi Camera.

3.1 Design Approach

For demonstration, Pi Camera module is connected to the Raspberry Pi 3 Camera Serial Interface Port and Camera Interfacing in Raspberry Pi configuration settings is enabled. VNC Viewer is used on the laptop to mirror the Raspberry Pi GUI to view the recognized gestures output. Training data set consists of 400 Images of each alphabet

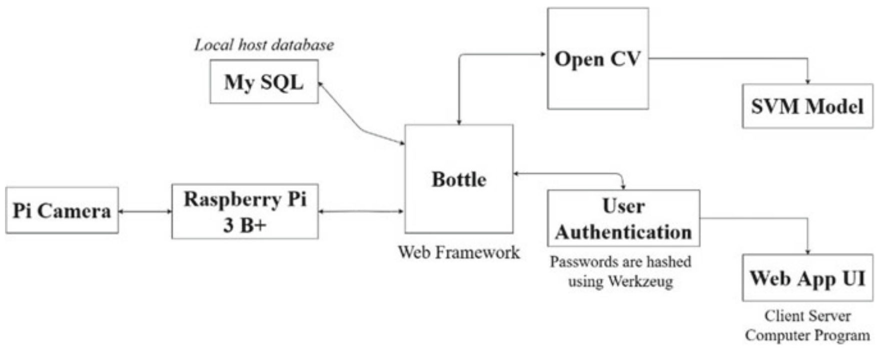


Fig. 1 Block diagram of ASL interpreter system

of American Sign Language. The server-side approach is implemented in the system using Bottle python library and the login credentials are stored and retrieved form MySQL Database and is stored locally in the Raspberry Pi. Werkzeug is a Web Server Gateway Interface that helps in adding salt and generating hashed passwords and it is available as library in python.

3.2 User Authentication

A user should be registered to use this app. A session is hosted for the user to login, if already registered or else to register, if the user is not registered before. These user details get stored so that the experience is user specific. A log file specific to user is maintained to follow the requests from one to the next until the user is logged out. Cookies are used to host the session, and they are signed cryptographically.

3.3 Password Hashing

First and foremost, while handling user accounts that are secured by a unique password, we must make sure that the passwords should be hashed before storing them on the database. Or else those passwords get highly vulnerable. Even when they are hashed, the basic hashing won't be able to secure them from attacks such as Rainbow table attacks, where a list of common words and passwords are hashed with MD5 or SHA functions and then are formed into a list then matched with the stored passwords to crack the password by finding a perfect match.

To further secure passwords from dictionary and rainbow effects, hashed passwords are added with a salt, which is a random data that is generated to be added to the hashing function. In bottle, Werkzeug helps in performing the above described task. Werkzeug is a WSGI (Web Server Gateway Interface) library for Python that takes the password and adds salt to the hashing function so that the one-way function returns a non-reversible password. Where salt_length sets the specified length of the salt in letters and check_password_hash() makes sure that the hashed password passed is equal to the one in the database created in MySQL. MySQL is a database service, which is completely managed locally in Raspberry Pi. MySQL has a connecter for python to access the database and it is very efficient and has high performance when compared to other local database services on the Raspberry Pi. Usernames and encrypted passwords get stored in this database.

3.4 Webpage UI Design

The Raspberry pi acts as the server and a web app has been designed to interact with the system. It contains a login page, index page, and a registration page through which a user can either login using existing credentials or register for a new account.

The login page consists of username and password fields and a hyperlink for signup page if the user is new, as shown in Fig. 2. In the registration page, user must create a new username and password. If the username is available then the registration will be successful, else the page says, "username already exists" and user has to enter a different username.

Once the registration is successful, it navigates directly to the login page and user can enter their credentials. After logging in, user can see ASL Interpreter option and the name of the user logged in.

Then clicking on the ASL interpreter, the training data set is loaded, and the camera window opens as shown in Fig. 3 and can recognize the gestures. If the

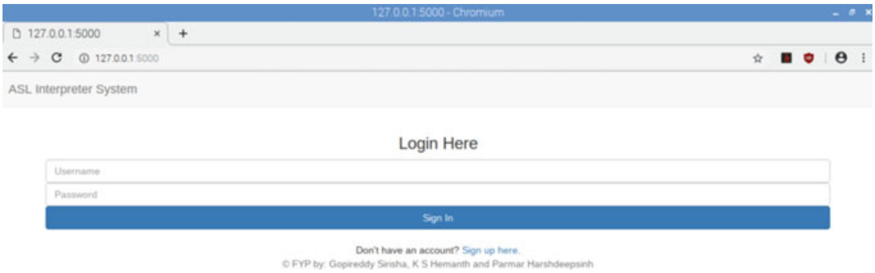


Fig. 2 Web login page of ASL interpreter system

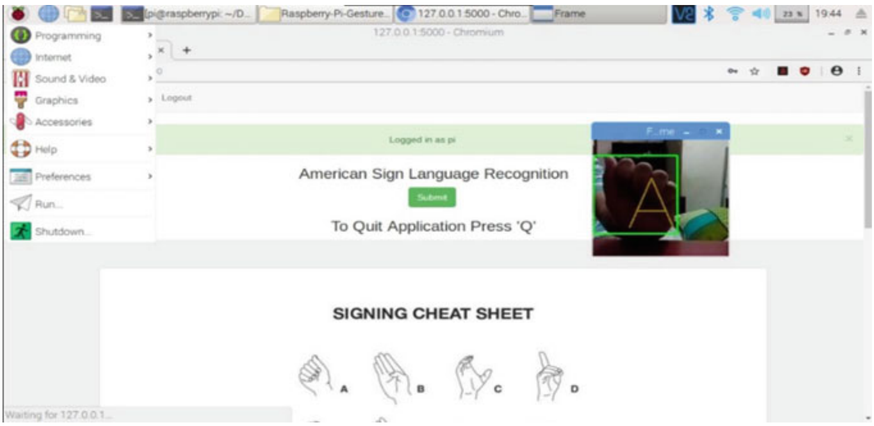


Fig. 3 ASL interpreter system showing live camera stream

user needs to quit, '*q*' can be pressed on keyboard and the programs gets killed and clicking on logout button user is successfully logged out.

3.5 Supervised Machine Learning Algorithm

It is a classifier, in which a discriminative hyperplane defines the segregation. This classifier takes a properly labeled training data set and then performs defined calculations on that data to give a function of hyperplane that helps in categorizing future data. This model calculates the error between the predicted output and the true output from the test data to optimize the values of variables of the function to minimize the error. Though it is a basic approach, the SVM model is highly effective in identifying various objects. Depending on the user, the results may be different.

3.6 Image Processing Algorithms

Initially, we have implemented the following algorithm (From here onwards, it will be referred as algorithm 1):

- **Step 1: Image Scaling**—Bell Bicubic interpolation function. It is applied on the Grayscale image obtained from the captured RGB image.
- **Step 2: Smoothing—Gaussian Low Pass Filter**. It removes the noise and blur, and further helps in enhanced edge detection.
- **Step 3: Edge Detection (Canny Edge Detection)** [10]. This set of algorithms performs edge thinning and reduces the number of false edges by calculating two threshold values—higher and lower, through Otsu's method and classifying edges as true and false edges.
- **Step 4: Improved Sobel Edge Detection** [11]. It is a mixture of Sobel Operator and functions of Dilation and Erosion (Open and Close functions).
- **Step 5: Wavelet Image Fusion** [12]. The Canny Edge and the Improved Sobel Edge are fused together through Wavelet Transform (Haar wavelet 2—Mother wave). The finally obtained edge after fusion is used for feature extraction through Zone based approach [13].

Each image is divided into 100 zones and the mean and standard deviation are calculated for the extracted features in each zone. Thus, 100 values for each image are calculated, and the data base is created for training the model. Features are extracted from the captured gestures in the same way and are classified by SVM. Later, we have tried a wavelet-based feature vector method utilized in 2D Image recognition [14] (From here onward, it will be referred as Algorithm 2). Genetic algorithm that uses a fitness function, which optimizes the variables that enhance the image, unique to each problem, is used for image preprocessing and wavelet optimization. Acquired image after being preprocessed is then decomposed (through wavelet decomposition) to its lowest level possible to form the feature vector that represents the respective sign. Classifier used here is feed forward neural network (FFNN), with 18 nodes in the input layer, 5 nodes in the hidden layer and 24 nodes in the output layer. This method produces a set of feature vectors with a high variance, thus decreasing the probability of error in recognizing a sign. We propose to use this method for better accuracy. But we cannot train the data on Raspberry Pi through this method, as it is computationally intensive. Adding new gestures will be a problem. We have to train the model elsewhere and include the classifier obtained, into our system or replace Pi with a better device, this would shoot up the expenses. We have trained on our laptop and implemented on Pi.

4 Results and Discussions

To classify the ASL, we have selected SVM as the classifier that has very high precision for algorithm 1 and FFNN for algorithm 2. As our main objective is to develop a portable system to promote the mobile usage of the system, we have also developed a user authentication interface that lets the user login with safe credentials and use the features. The user interface uses password hashing, MySQL Database and real time image capturing to securely get the hand sign (as shown in Fig. 4). The support vector machine uses optimal hyperplane computation to get the promising results faster and uses less memory. As our focus is just on simple hand sign recognition with limited dataset and resources, SVM and FFNN models are deployed. Greater accuracies on complex extractions can be obtained through deep-convolutional neural networks, but they require high-processing power and large memory. In addition, they cannot be deployed on Raspberry Pi 3B + for now.

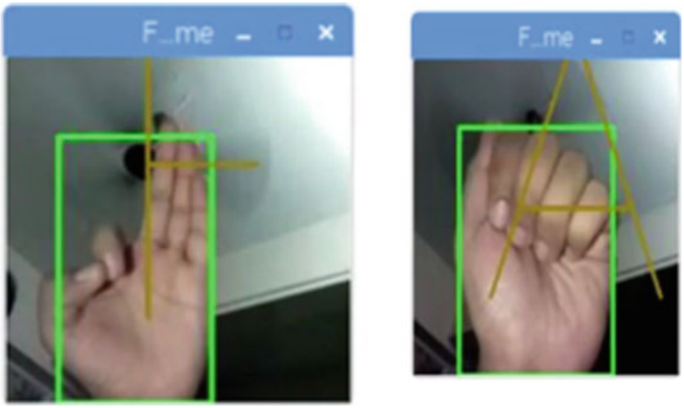


Fig. 4 Signs predicted from interpreter

Table 1 (in page 9) shows the accuracy for the algorithms, excluding ‘J’ and ‘Z’ alphabets as they are not static. For each chosen alphabet, 50 trials were conducted, and the accuracy is computed as the ratio of number of true positives (i.e., the correct classifications) to the total number of classifications (50 in this case). Algorithm 1 achieved an accuracy of 94.83%. And wavelet-based feature vector method (Algorithm 2) has recognized 24 ASL signs are with accuracy of 99%, SNR is always greater than 2 and signs in noisy images are also recognized with good accuracy. Our designed system might fail to analyze and identify the similar signs in certain lighting conditions. However, the results still show a considerable amount of prediction in the case of similar signs. Sometimes there are several factors that contribute to heavy error such as fingers overlapping and gap between the fingers. Overall, the consistent accuracy and robustness of our system makes it ideal to employ it in certain real-world applications.

Table 1 Accuracy (in percentage) of ASL fingerspelling interpreter(s)

Alphabet	Proposed model involving Algorithm 1	Proposed model involving Algorithm 2
A	98	100
B	96	98
C	98	100
D	94	98
E	96	96
F	94	100
G	92	100
H	94	100
I	92	98
K	92	100
L	98	98
M	92	100
N	96	98
O	94	100
P	98	98
Q	96	100
R	94	98
S	92	96
T	96	100
U	94	100
V	94	100
W	94	100
X	96	98
Y	96	100

5 Conclusion

This proposed system captures static image of the hand gesture that is part of american sign language, using Pi camera, interprets the sign intended to be presented, with the help of image processing and machine learning algorithms, and displays the corresponding English alphabet, after referring to the database. This task gets carried out by Raspberry Pi 3 B that runs the bottle-based web app. User authentication and password hashing helps in providing user specific unique experience. All the user data is stored in MySQL database hosted locally in the Raspberry Pi.

Recommendations for Future Work. We suggest including a 3D gesture recognition sensor to capture the orientation part of the gesture and recommend keeping the background static. We also suggest add-ons such as enabling the system to send a text message or an email through a sim card, and extra gestures to communicate with the system.

References

1. Arif, A, Rizvi STH, I. Jawaid MA, Waleed, Shakeel MR (2016) Techno-talk: an American Sign Language (ASL) translator. In: International conference on control, decision and information technologies (CoDIT), 6–8 Apr 2016, pp 665–670
2. Khambaty Y, Quintana R, Shadaram M, Nehal S, Virk MA, Ahmed W, Ahmedani G (2008) Cost effective portable system for sign language gesture recognition. In: IEEE international conference on system of systems engineering, 2–4 June 2008
3. Kosmidou VE, Hadjileontiadis LJ, Panas SM (2006) Evaluation of surface EMG features for the recognition of American Sign Language gestures. In: International conference of the IEEE engineering in medicine and biology society, Aug 30–Sep 3, 2006, pp 6197–6200
4. Siddiqui N, Chan RHM (2017) A wearable hand gesture recognition device based on acoustic measurements at wrist. In: 39th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), 11–15 July 2017, pp 4443–4446
5. Ugale M, Bende S, Haldankar M, Hegde S (2018) ASL Fingerspelling interpretation using SVM and MLP. Int Res J Eng Technol (IRJET) 5(3):1825–1827
6. Islam MM, Siddiqua S, Afnan J (2017) Real time hand gesture recognition using different algorithms based on American sign language. In: IEEE international conference on imaging, vision & pattern recognition (icIVPR), 13–14 Feb 2017
7. Kulkarni M, Butala J, Udpikar V (2014) Static gesture recognition using PMD ToF camera. In: International conference on advances in computing, communications and informatics (ICACCI), 24–27 Sep 24–27, pp 2712–2718
8. Vogler C, Metaxas D (1999) Parallel hidden Markov models for American sign language recognition. In: Proceedings of the seventh IEEE international conference on computer vision, 20–27 Sept 1999
9. Wang S, Song J, Lien J, Poupyrev I, Hilliges O (2016) Interacting with soli. In: Proceedings of the 29th annual symposium on user interface software and technology—UIST 16, 16–19 Oct 2016, pp 851–860
10. Amer GMH, Abushaala AM (2015) Edge detection methods. In: 2nd world symposium on web applications and networking (WSWAN), 21–23 Mar 2015
11. Gao W, Zhang X, Yang L, Liu H (2010) An improved Sobel edge detection. In: 3rd International conference on computer science and information technology, 9–11 July 2010

12. Patil SM (2016) Image fusion using wavelet transform. Int J Eng Adv Technol (IJEAT) 5(4):66–69
13. Pasha S, Kumar A, Hussain SY, Amritansh, Ul Umam MS (2015) Zone based and morphological based feature extraction techniques for english alphabets. Int J Comput Sci Inf Technol (IJCSIT) **6**(3), 2941–2944 (2015)
14. Isaacs J, Foo S (2004) Optimized wavelet hand pose estimation for American sign language recognition. In: Proceedings of the 2004 congress on evolutionary computation (IEEE Cat. No. 04TH8753), 19–23 June 2004, pp 797–802

Security for User Profile Matching in Social Networks

O. Pandithurai, D. Jayashree, E. Kiruthiga, E. Kavya, and A. Ishwarya

Abstract We consider a situation where a client tends to have a client profile database, kept up quite a while ago can handle by corresponding association supplier. In order to perceive clients' organized profile Data it should undergo for Customer investigating. A common example of this functionality is on the social media site. Also starting late, a cloud-based dating site, Internet dating, was compromised, which reveals a lot of online customer accounts to goliaths. This information break has mentioned that bosses investigate viable security insurance for client profiles in a social affiliation. We propose an affirmation protecting response for profile sorting out in easygoing systems by utilizing server. Our solution relies upon homomorphic encryption and enables a client to discover sorting out clients with the help of servers. This is done without uncovering to anybody the request and the tended to client profiles in clear.

Keywords Matching users profiles · Data protection · ElGamal authentication · Paillier authentication · Homomorphic authentication

1 Introduction

Coordinating at least two clients with related relation is a significant and common issue, pertinent to a scope of situations including work chasing, companion finding. Existing online coordinating administrations expect members to confide in an outsider server with inclinations. The coordinating server has full information on the clients' inclinations, which give rise to protection issues, as the server also might spill (either purposefully, or incidentally) client profiles. While pursuing an Internet coordinating assistance, a client makes a "profile" that others can peruse. The client might be approached to uncover subtleties, for example, age, sex, instruction, calling, number of kids, religion, geographic area, sexual proclivities, drinking conduct,

O. Pandithurai (✉) · D. Jayashree · E. Kiruthiga · E. Kavya · A. Ishwarya
Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai, India
e-mail: pandics@ritchennai.edu.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_18

diversions, salary, religion, ethnicity, tranquilize use, home and street numbers, and most loved spots. Much after a record is dropped; most Internet coordinating destinations may hold such data [1]. Clients' own data might be re-revealed not exclusively to imminent matches, yet in addition to publicists and, eventually, to information aggregators who utilize the information for purposes inconsequential to web-based coordinating and without client assent. Likewise, there are dangers, for example, tricksters, sexual stalkers, and reputational harm that join utilizing Internet coordinating administrations.

2 Related Work

Private two-party set-crossing point concern, where was that off opportunity one gathering P_1 inputs $X = \{x_1, x_2, \ldots, x_n\}$ and P_2, the other party inputs $Y = \{y_1, y_2, \ldots, y_n\}$, one gathering learn $X \cap Y$, and the other parties adapts nothing. Their answer depends upon commutative encryption with property: $E_{\text{k1}}(E_{k2}(x)) = E_{k2}(E_{\text{k1}}(x))$, where k_1, k_2 are known to P_1 and P_2, individually. The thought is: P_1 (P_2) encodes its information sources $X(Y)$ with its key $k_1(k_2)$, signified as $E_{\text{k1}}(X)(E_{k2}(\text{Y}))$, and trades them. Next, P_1 sends a couple $(E_{k2}(Y), E_{\text{k1}}(E_{k2}(Y)))$ to P_2, which processes $E_{k2}(E_{\text{k1}}(X)))$, contrasts it and $(E_{k2}(Y), E_{\text{k1}}(E_{k2}(Y)))$, and decodes $E_{k2}(y)$ if $E_{k2}(E_{\text{k1}}(x)) = E_{\text{k1}}(E_{k2}(y))$. At that point, Vaidya et al. stretched out such an answer for n-party setting. Arbitral applied this plan to distinguish companion of—companion in MSN. Freedman et al. gave an answer for the private two gathering set-crossing point issue dependent on polynomial assessment. The thought is P_1 characterizes a polynomial $P(y) = (x_{1-y})(x_{2-y})\ldots(x_{n-y})$ and sends to P_2 homomorphic encryptions of coefficients of the polynomial. P_2 utilizes the homomorphic properties of an encryption [2] framework in order to assess the polynomial at each and every one of his data sources, then this increases each outcome by a new arbitrary number (r) to receive a halfway outcome, and adds this to an encryption of the estimation of information that is P_2 figures $E(rP(y) + y)$. Accordingly, for every one of this components in the two crossing point of gatherings' data sources, the consequence of the calculation is an estimation [3] of related component; while for every other worth, the outcome is an arbitrary. Kissner and Song gave an improved arrangement in the year 2007, that empowers set convergence, set-crossing point, and over-edge set-association tasks on multisets. Afterward, Sangetal, Ye et al. what's more, Dachman-Soled et al. later improved and widened these arrangements. In the year 2007, Li and Wu came up with a genuinely secured convention for the multiple party set convergence. Their thought was like Kissner and Song, where the sources of info are shared among all the gatherings utilizing mystery sharing, and calculations are assassinated on those offers. In the year 2010, Narayanan et al. proposed an improved arrangement. Every one of these arrangements depend on polynomial assessment.

3 Literature Survey

1. Posting of Confidential information through Repositories. Michael Siegenthaler; Ken Birdman 2009. For example, in companies such as medicinal services, there is a need for the electronic sharing of touchy information about security across unmistakable associations. They illustrate how this should be done by requiring organization's to preserve their inheritance records and hold responsibility for the details they actually retain. Without submitting or representing some trustworthy knowledge, combined content, we demonstrate how queries can be addressed in a proper manner that jelly the first privacy security [4]. This report validates our transmitted engine for operation of inquiries and shows how to prettify the machine when only valid markers are used, Of starters, the identity and social security number are identified at first, giving nuances on the price to pay between confidentiality and results.

2. Friend-of-Friend Client server Discovery in Wireless Social media network. Marco von Arb; Roger Wattenhofer; Matthias Bader 2008. Versatile social programming has, as of late, become a functioning area of innovative work. Over the past few years, a huge number of frameworks were proposed that tried to follow an achievement of their Internet bound counterparts. Numerous portable structures aim to extend the utility of current stages with regard to area. The cost of versatility, however, is normally either the absence of the highlights of the famous companionship investigation, or the cost of getting to a focal server needed for this utility.

Attempt to dispatch the issue right now by presenting the decentralized strategy which can identify a client's social neighborhood companions. In contrast to mere misuse of the framework's customer data, the technique depends on genuine companions and addresses the emerging security [4] issues satisfactorily. Moreover, we present VENETA, a portable person to person contact stage that updates our novel companion in the measurement of partner identity between different highlights.

3. Use Secret Ideal Lattice [5], Complete Homorphic Encryption. Thomas Plantard; Willy Susilo; Zhenfei Zhang 2013. All the current fully homomorphic encryption plans depend on three unique issues, specifically the issue of limited separation interpretation over perfect grid, the surmised most prominent normal divisor issue over numbers, and blunder problem learning. Right now, tie the initial two classes of problems together introducing another class of issues that may be affected by both. Bearing in mind this new topic, the minimal division untangling around veiled ideal should be precise panel, and we are introducing another collaborate with absolutely homomorphic coding. Since it is more or less a combination of the two things, our plan's show lies here between ideal panel-based plans and that the whole percentage-based plans. In addition, we are also showing a lower bound and upper bound of the problem on which our strategy rests. Acceptance of this protection assumption remains, we can combine littler criteria, which will bring about a strategy that is more successful than the plans based on cross-section and whole amount. Therefore, in comparison to the state of craftsmanship, our plan makes an ideal choice for learning with error-based plans.

4. Profitable authentication detection in web networking to safeguard security. Jiawei Yuan; Shucheng Yu 2013**.** Biometric recognizable proof is a dependable and advantageous method for distinguishing people. The across the board appropriation of biometric distinguishing proof requires strong security insurance against conceivable abuse, misfortune, or robbery of biometric information. Existing methods for protection saving biometric recognizable proof essentially depend on traditional cryptographic natives, for example, homomorphic encryption and neglectful exchange, which unavoidably acquaint colossal expense with the framework and are not relevant to down to earth enormous scale applications. Right now, propose a novel security protecting biometric recognizable proof plan which accomplishes productivity by abusing the intensity of distributed computing. In our proposed plan, the biometric database is scrambled and re- appropriated to the cloud servers. To play out a biometric recognizable proof, the database proprietor produces a certification for the competitor biometric characteristic and submits it to the cloud.

The cloud servers perform recognizable proof over the encoded database utilizing the accreditation and return the outcome to the proprietor. During the ID, cloud adapts nothing about the first private biometric information. Since the distinguishing proof tasks are safely re-appropriated to the cloud, the constant computational/correspondence costs at the proprietor side are negligible. Exhaustive examination shows that our proposed plan is secure and offers a more significant level of security insurance than related arrangements, for example, kNN search in scrambled databases. Genuine analyses on Amazon cloud, over databases of various sizes, show that our computational/correspondence costs at the proprietor side are a few extents lower than the current biometric distinguishing proof plans.

5. Protection of Account Matching for Online Micro blogging in Close vicinity Rui Zhang; Jinxue Zhang; Yanchao Zhang; Jinyuan Sun; Guanhua Yan. Nearness-based versatile long range interpersonal communication (PMSN) alludes to the social collaboration among truly proximate portable clients. The initial move toward compelling PMSN is for versatile clients to pick whom to connect with.

Profile coordinating alludes two customers take a quick look on their own statuses and are pledging in PMSN for buyer decision. It, in any case, disputes with the creation of protection stresses by consumers over disclosure on their own statuses to fulfill misfits. This document performs this active test by coordinating private preparation reveals with the great-grained novelties. Our shows allow two buyers to process profile planning without discovering any details concerning about their user profiles beyond the results of the assessment, structuring novel fine-grained private coordinating conventions. Our conventions empower two clients to perform the profile coordinating without uncovering any of the data about their personal profiles. As opposed to existing coarse-grained private coordinating plans, our conventions permit better desperation between the PMSN clients and that can bolster a scope of coordinating measurements at the various protection levels. The presentation of our conventions is altogether examined and assessed through genuine advanced mobile [6] phone tests.

4 Existing System

In existing framework, where one client demand a client user profile database, kept and maintained by long social correspondence association supplier, to see clients whose profiles orchestrate the profile constrained by the examining client. Kept up by a long range interpersonal communication specialist co-op, to recognize clients whose profiles coordinate with the profile determined by the questioning client. An average cloud-based dating is a case of this method. Most recently, Ashley Madison, a cloud-based dating site, was deactivated, which brings about revelation of an enormous number of dating client profiles [7]. This information break has asked analysts to investigate viable security assurance for client profiles in an interpersonal organization. Client subtleties would not be scrambled so programmer effectively hacking clients subtleties and messages [6] something like pictures, recordings, and content.

5 Proposed System

In proposed structure, we propose an insurance sparing response for profile organizing in casual networks using separate servers. Our outcome depends on homomorphic cryptography and motivates a consumer to locate coordinating buyers with the aid of different computers without exposing them to everyone the inquiry and the addressed customer profiles in clear. Utilizing two calculations, one is El Gamal encryption algorithm, and second is homomorphic encryption algorithm. We propose a security protecting answer for profile coordinating in interpersonal organizations. Our solution is based on homomorphic encryption and permits a client to discover coordinating clients, without uncovering to anybody the question and the questioned client profiles in clear. Our solution accomplishes client profile security and client inquiry protection as long as at any rate one of the numerous servers is straightforward. Client subtleties will be scramble while register and message meeting utilizing calculation.

6 Modules Description

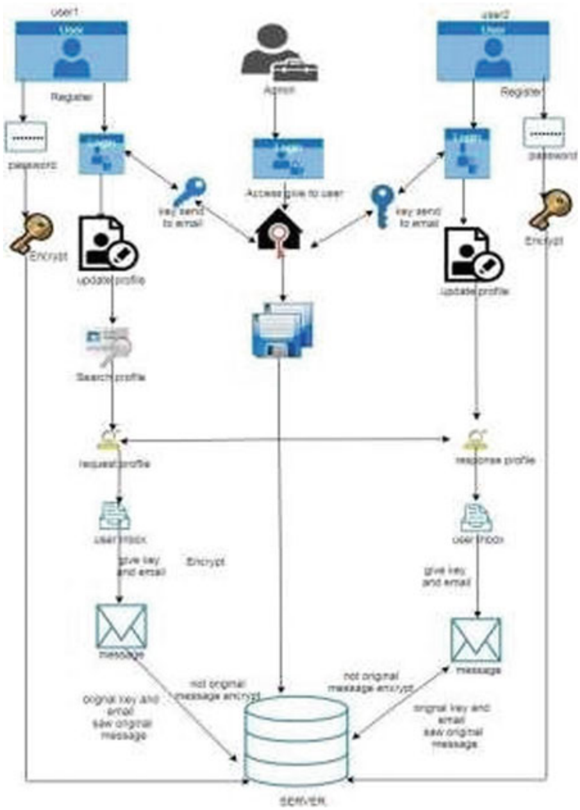
Interface design is the most primary module for our endeavor. The huge activity for a customer is to drag login window to the customer window. This authentic customer id which made for the client security purpose. Right now, we have to step into the login customer id. It will check username, and mystery express is facilitate or not (authentic customer id and considerable mystery Illustration). On the inconceivable plausibility, we enter any invalid username or mystery key; we cannot get into the login window to customer window. It will project botch message. So, we are protecting from unapproved customer getting into the login window to customer window. It will provide a better security to our endeavor. So server contains customer id and mystery word server moreover which check the affirmation of every customer. It well improves the security and protection from unapproved customer getting into the framework. In this endeavor, we are using JSP for preparing structure. Added we support the login customer details and server checkings. User login is the second module in our project [8], symbolizing a work unit performed against a database within just a content management program (or different system) and managed in a consistent and reliable way, inclusive of other structures operations.

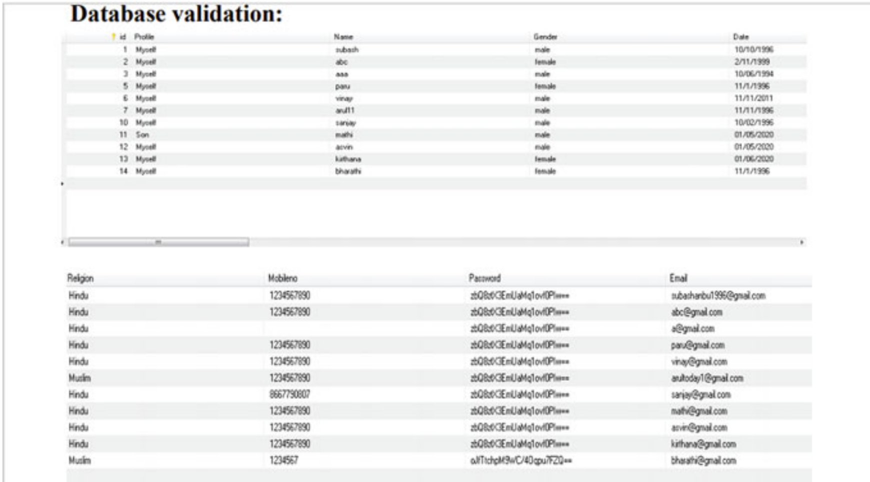
A transaction usually reflects any change in the amount passed to the supplier by the recipient of the database. User upload details are the module in which the imported information with personality will be saved in the database to allow the user display all info to another user. Admin login is the section in our project, and a research unit conducted against a database [6] within a database management framework (or similar system) is symbolized here and handled in a consistent and reliable manner separate of other transactions. A contract usually reflects any change in the amount that the user of the database must pass to the supplier. Admin checks the user information [9] file which implies what we should be currently performing. Admin client access, if no account is logging into admin access that is not needed, the user must revoke the file. And clients submit notice to warn. The client will also regard the significant number of user accounts in this device, if you want to send user requests to friend list. You would be liable for the database contained in your profile. Throughout this device, one consumer is securing [10] convocation to another client message, and user specifics and mails are maintained by admin.

7 System Architecture

The frameworks engineer builds up the fundamental structure of the framework, and we propose a cumulative sum (CUSUM) calculation, and we can place a little piece of information in neighborhood machine and haze server so as to ensure the security. Also, in view of computational insight, this calculation can figure the circulation extent put away in cloud, mist, and nearby machine, separately. Through the hypothetical security [11] investigation and exploratory assessment, the achievability of our plan has been approved, which is actually an incredible enhancement to existing distributed storage plot.

8 Result Analysis





Database validation:

id	Profile	Name	Gender	Date
1	Myself	subash	male	10/10/1996
2	Myself	abc	female	2/11/1999
3	Myself	aaa	male	10/06/1994
5	Myself	paru	female	11/1/1996
6	Myself	vinay	male	11/11/2011
7	Myself	arul11	male	11/11/1996
10	Myself	sanjay	male	10/02/1996
11	Son	mathi	male	01/05/2020
12	Myself	acvin	male	01/05/2020
13	Myself	kirthana	female	01/06/2020
14	Myself	bharathi	female	11/1/1996

Religion	Mobileno	Password	Email
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	subashanbu1996@gmail.com
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	abc@gmail.com
Hindu		zbQ8z6G3EmUaMq1ovf0Pl===	a@gmail.com
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	paru@gmail.com
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	vinay@gmail.com
Muslim	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	arultoday1@gmail.com
Hindu	8667790807	zbQ8z6G3EmUaMq1ovf0Pl===	sanjay@gmail.com
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	mathi@gmail.com
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	acvin@gmail.com
Hindu	1234567890	zbQ8z6G3EmUaMq1ovf0Pl===	kirthana@gmail.com
Muslim	1234567	oJlTtchpM9wC/40qpu7FZQ==	bharathi@gmail.com

9 Conclusion

Right now, proposed another answer for protection safeguarding coordinating client profile with the help of homomorphic encryption has been proposed. Our solution permits the client to discover the coordinating clients without uncovering the inquiry and the client profiles. Security [12] examinations have demonstrated that the new convention accomplishes client profile protection and client question security. The exploratory outcomes have demonstrated that the new convention is common sense and plausible.

References

1. Srikant AR, Evfimievski A (2003) Information sharing across private databases. In: SIGMOD 2003, pp 86–97
2. Nissim BK, Goh EJ (2006) Evaluating 2-DNF formulas on ciphertexts. In: TCC 2006, pp 325-341
3. Chaum (1982) Blind signatures for untraceable payments. In: Crypto 1982, pp 199–203
4. Lindell Y, Hazay C (2008) Efficient protocols for set intersection and pattern matching with security against malicious and covert adversaries. In: TCC in the year 2008
5. Gentry C (2009) Fully homomorphic encryption using the ideal lattices. In: STOC in the year 2009, pp 169–178
6. von Arb M, Kuhn M, Wattenhofer R, Veneta (2008) Detecting app engine contacts in smartphone instant messaging. In: IEEE WIMOB 2008, pp 184-189
7. Freedman M, Nissim K, Pinkas B (2004) Efficient private matching and the set intersection. In: EUROCRYPT in the year 2004, pp 1–19
8. Bloom BH. Space/time trade-offs in the hash coding with the help of an allowable error. Commun ACM 13(7)

9. Micali S, Goldwasser S (1982) Probabilistic encryption and how to play mental poker keeping secret all partial information, In: Proceedings of 14th symposium on theory of computing, in the year 1982
10. Cristofaro D, Tsudik G (2010) Pleasant, intuitive uniqueness private specified crosswalk guidelines. In: Financial Cryptography and Data Security 2010
11. Dachman-Soled D, Raykova M, Malkinand T, Yung M (2009) Efficient robust for private set intersection. In: ACNS in the year 2009
12. ElGamal T (1985) A cryptosystem with public key and a verification structure focused on separate algebra. IEEE Trans Inf Theory 31(4):469–472

AI and ML-Based Crop Detection Application

M. Malathi, K. K. Faiyaz, R. M. Naveen, C. Nithish, and R. Induja

Abstract Crops play the most important role in our everyday lives. Thus, the assessment of crop is vital. The growing technology plays an integral part in crop identification and in techniques like machine learning, deep learning. Nowadays, there is a tremendous growth of mobile devices. But with machine learning systems, there is an increasing demand for smartphone apps. This paper focus on identification of crop using field image of the crop captured along with its geo-location using mobile camera in the absence of Internet connection. The captured image is recognized using image processing technique. Based on the results, the users will get to know about the crop. The photograph and detailed description of the crop with the geo-location will be stored in the local storage of the device which can be exported/e-mailed.

Keywords Computer vision · Artificial intelligent · On-device machine learning · TensorFlow

1 Introduction

Machine learning is the field of research that enables the computer to learn itself without explanatory programming. Machine learning is the application of artificial intelligent (AI). The process is divided into two types; training and testing. The user

M. Malathi (✉) · K. K. Faiyaz · R. M. Naveen · C. Nithish · R. Induja
Sri Krishna College of Technology, Coimbatore, India
e-mail: m.malathi@skct.edu.in

K. K. Faiyaz
e-mail: 18tuit029@skct.edu.in

R. M. Naveen
e-mail: 18tuit058@skct.edu.in

C. Nithish
e-mail: 18tuit061@skct.edu.in

R. Induja
e-mail: 18tuit037@skct.edu.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_19

provides the training data to the system which helps it to get trained. The trained system is tested using test data. A large amount of training data gives good accuracy to the results. When the system is given images as the input, the images have to be processed with the help of image processing techniques. The identification of crops requires images of the plants in real-time conditions. These images can be either laboratory-based or field-based images [1]. Machine learning has radically changed how the world could process, model, and interprets data [2]. The machine learning systems run in the cloud and are computationally intensive and require significant memory. Mobile devices are projected to hit 5.6 billion by 2020 [3]. Machine learning (ML) is now widely used in mobile applications for object detection, health monitoring, malware detection, language translation, and so on [4–6]. Machine learning is thought to play an important part in the development of mobile devices [7]. This on-device machine learning system developed by the engineering team expander, allows technologies to be used, without being linked to the cloud for any use. A large selection of training data is used during the training process to fine-tune the hyperparameters in the deep neural network (DNN) using gradient descent algorithms [8]. The wide DNN models surpass the memory of mobile devices, and thus the off-chip memory has to accommodate them, which significantly consumes more energy. To reduce the intricacy in computing, the fast Fourier transform-based multiplication is used [9, 6, 10]. This paper focuses on the identification of the crop using the field image of the crop. The mobile camera is used to capture the field image of the crop. The captured image is then processed using machine learning algorithms (CNN). A mobile application that automatically recognizes the plant is important for real-world ecological surveillance [11–13], invasive exotic plant control [14, 15], popularization of ecological science, and so forth. To upgrade the efficiency of these models, plant identification draws the attention of engineers and scholars.

2 Literature Survey

There are various studies and research done on crop identification using machine learning. Deshmukh and Manjusha Deshmukh detected paddy leaf using the K-means algorithm. This paper deals with the image acquisition, image preprocessing and segmentation, selection, and description of the features using ANN [16]. Segmentation of images is built with the K-means algorithm. These papers result in the fast detection of paddy leaves. Images of paddy leaf captured by using cameras are handled via the remote server. Yu Sun, Yuan Liu, et al. have proposed "Deep Learning for Natural Plant Recognition" [11]. In this system, the dataset which consists of 10,000 pictures of 100 ornamental plants is collected from the University of Beijing Forestry campus by cell phone in a natural scene. For plant classification, a deep learning model is built consisting of 26 layers and 8 residual building blocks. This model achieves a 91.78 percent recognition rate. H. Gosau, P. Bonnet, and A. Joly have proposed "Plant identification in an open-world (life clef 2016) [17]." The main aim of this system is to evaluate plant identification methods and very large scale schemes, similar to real-world biodiversity monitoring scenario conditions. The main novelty is that the identification task is strong enough to handle the unseen categories. N. Kumar et al. ha proposed "A Computer Vision Device for Automated recognition of Plant Species" [18]. This system explains the detection of plant species using automated object detection in a mobile app. Removing non-leaf photos, segmenting the leaf from the context, extraction characteristics reflecting the curvature of the contour of the leaf over multiple scales, and description of the species from a dataset of the 184 trees are done using computer vision components [19]. S. H. Lee, Y. L. Chang, C. S. Chan, and P. Remagnino have proposed an approach for plant recognition which uses neural convolution network [20]. This system identified 1000 forms of plants photograph referring to 7 different plant organs and even invasive species automatically identified from unknown groups. They used the CNN model to blend the habitat and organ characteristics for the plant classification task. The results of this system are limited when matched with VGGNet. While working, automatic plant taxonomy has yielded fruitful results. On observations, these trained models meet the requirements of completely automated environmental monitoring scenarios [14]. The deep convolutionary neural networks proposed in [21] had shown exceptional conduct in ILSVRC-2012's large scale image classification task [22]. The conventional classification models depend on preprocessing to remove highly complicated backgrounds to boost desirable functionalities and features.

3 Proposed Model

The following are the major components in this system: (1) dataset preparation, (2) implementation of CNN using TensorFlow, (3) on-device machine learning, and (4) working on mobile application (Fig. 1).

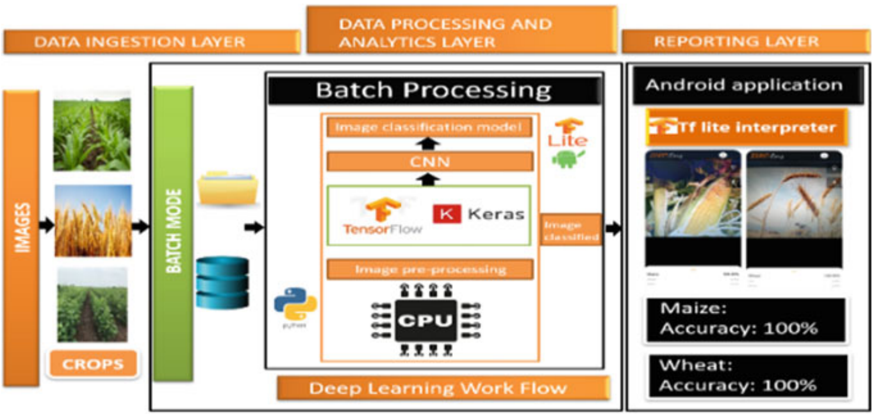


Fig. 1 Technology stack of this system

3.1 Dataset Preparation

Data preparation is the first significant step in a machine learning project. It is the act of manipulating raw data in a manner that can be analyzed readily and precisely by data scientists and analysts, who can run it through machine learning algorithms to make predictions. The dataset of 30 different classes of crops, each of which 200 images have been collected from the natural scene. The most special feature of this dataset is that it contains different images for various cycles of growth for a given crop.

3.2 Implementation of CNN Architecture

Convolution neural network (CNN), a deep learning model used for classification and detection, has been increasingly developed with the advancement in deep learning and large machine hardware, showing good accuracy on different crops [23]. A neural convolutionary network (CNN) is a deep learning algorithm, which is specifically used for analyzing images. Multilayer perceptrons are regularized copies of CNNs. Multilayer perceptrons are fully connected network, that is, every neurons in the previous layer is mapped to each neuron in that layer. This multilayered CNN is built using TensorFlow (Fig. 2).

CNN architecture used in this project is MobileNet V2. MobileNet V1 is an architecture used mainly to run a trained model embedded in the mobile application. This architecture optimizes the trained model and makes it effective to run on the mobile application. It uses very few parameters and addition, multiplication of convolution to

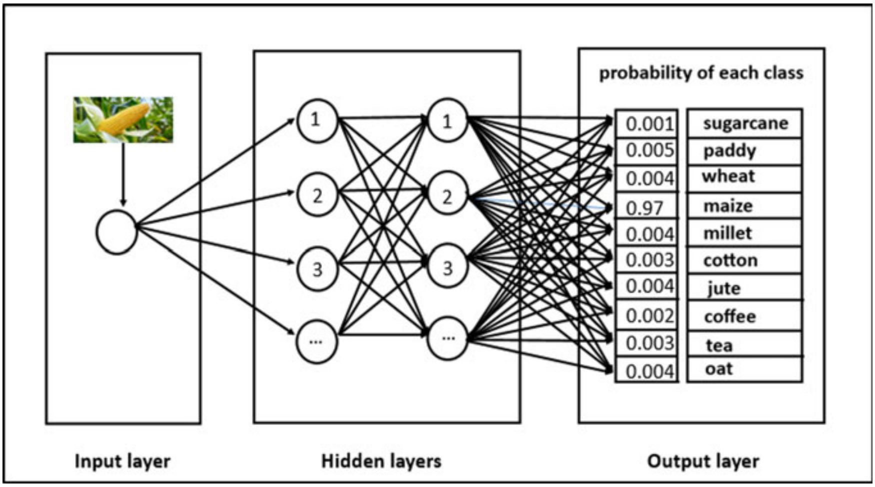


Fig. 2 Deep learning neural network

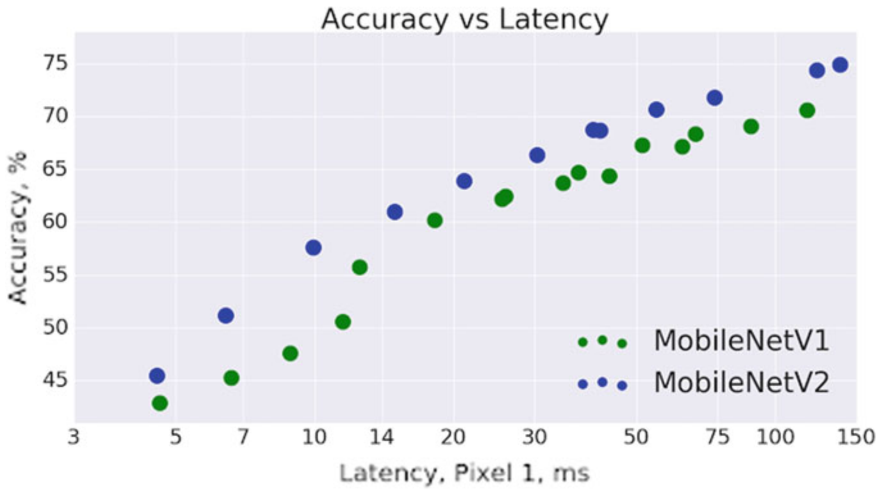


Fig. 3 MobileNet V1 versus MobileNet V2

optimize the size and complicacy. MobileNet V2 uses depthwise convolution which is fast and efficient, whereas MobileNet V1 uses pointwise convolution which uses the same channels or uses double of it, and that makes it more costly. MobileNet V2 uses advanced features like linear bottlenecks between the layer and shortcut connections between the bottlenecks. But these advanced features are not found in MobileNet V1.

Figure 3: It explains the comparison of accuracy and latency between MobileNet V1 and MobileNet V2. This proves that MobileNet V2 is more efficient when compared to MobileNet V1.

Steps involved in creating a neural network using TensorFlow are:

A. Importing Datasets: The crops dataset along with its labels are imported into our model.
B. Resizing and recoloring the images: The images in the dataset are resized and recolored in the RGB format.
C. Creating layers: A stack of convoluted layers along with dense and flatten layers are created.
D. Training the dataset.
E. Obtain a trained model.

On-Device Machine Learning:

On-device machine learning enables the machine learning attributes to execute on the user's personal devices. The TensorFlow lite makes inference on-device machine learning with small latency and limited binary size. The above-trained model is converted into a mobile compactable form using TensorFlow lite converter. The interpreter on TensorFlow lite runs the specially optimized models on mobile phones.

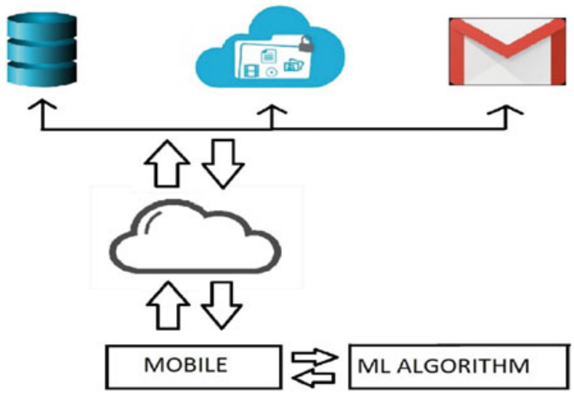


Fig. 4 Overall architecture of our method

Working on Mobile Application:

In the mobile application, first, authentication and approval are done. After this process, the user is navigated to the home page, where the user can capture the image of the crop along with its geo-location. Now, the process of image recognition is efficiently and effectively performed. As a result, the bottom sheet displays the recognized image in text format along with the accuracy rate and inference time. The main advantage of this app is that it operates even without the presence of an Internet connection, which is achieved by learning machines on-device. By default, the recognition is carried out by the CPU. We have given an option to change the processor from CPU to GPU. After recognition, the details of the crops are displayed to the user based on the label with the help of Wikipedia. Wikipedia API is used to set up a connection between Wikipedia and the application. The text-to-speech (TTS) option is provided for user convenience. This application will also provide weather data of the place using the geo-location. The recognized crop image along with its information and geo-location is mailed to the admin in the presence of the Internet.

Figure 4: Whenever the user captures the crop image using the mobile camera, automatically, the deep learning algorithm recognizes the image even in the absence of the Internet and renders the output to the user. Once the Internet connection is established, the crop's image with its geo-location and the information about the crop will be stored in the cloud and also e-mailed to the respective person (Table 1 and Fig. 5).

4 Results and Discussion

The dataset is prepared from the live field images and is not taken from the Internet which therefore helps in attaining high accuracy. The dataset also consists of mobile captured images so as to ease the process of recognizing the images from end user's device. The dataset is prepared at three stages of crops, the early stage, the ripened

Table. 1 Training accuracy rate of different class which has been collected during training the model

Class	Accuracy (%)
Paddy	90
Wheat	89
Oat	91
Maize	89
Millet	87
Jute	85
Sugarcane	86
Tea	99.98
Coffee	74
Cotton	71

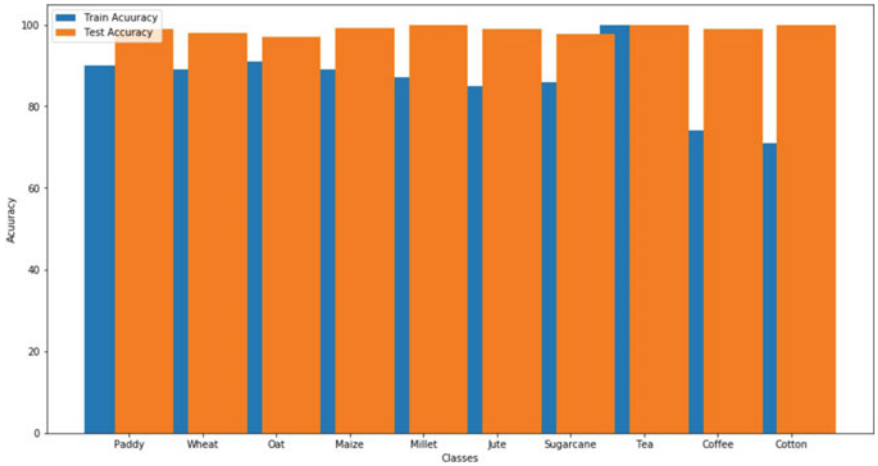


Fig. 5 Comparison between training and testing accuracy rate of our deep learning model

stage, and the parched stage, which in turn helps in providing a consistent accuracy for any crop. Successions of experiments were conducted to check the accuracy of our model. The testing accuracy rate is increased by multiplying the number of neural layers in our model. During different experiments, we have tuned the important parameters of the training model to improve the accuracy of our model.

Figure 6 is taken during testing of our application. This image shows the accuracy rate attained and information displayed in the application during the detection of maize.

Figure 7 is taken during the testing of our application. This image shows the accuracy and the information of the wheat that are displayed in the application. The accuracy rate is 100%.

Fig. 6 Output of maize prediction

The accuracy rate varies from crop to crop and device to device. This trained model produces an average accuracy rate of about 97% on a series of testing under different situations and circumstances.

Fig. 7 Output of wheat prediction

5 Conclusion

Agriculture plays an important role in global economy of the country. Owing to the population growth, the farm industry is constantly being forced to increase the productivity and produce more crops. Nowadays, artificial intelligence and machine learning systems are used in various number applications related to agriculture. The usage of machine intelligence in agriculture resulted in more healthy crops and also increased the overall productivity. Many systems used images dataset for identification of the crops. Our dataset is not collected from Google or other Web sites; instead, it is collected from the fields at various stages of the growth period of each crop, which leads to the development of more accurate image analysis tool. The model is trained under different methodologies and analyzed, and the best methodology is chosen to achieve high accuracy. In this system, the trained machine learning algorithms provide good and better results. Most of the currently used agricultural-based machine learning application used Internet to process the results. This application does not rely on the Internet connection, which allows explorers to identify any crops irrespective of the region. Pests and diseases in crops lead to the decrease in the productivity of crops. In future, this application will also include the pest detection.

References

1. Poonguzhali R, Vijayabhanu A (Apr 2019) "Machine learning crop condition evaluation". Int J Recent Technol Eng (IJRTE), ISSN: 2277–3878, 7(6S2)
2. Good fellow I, Bengio Y, Courville A (2016) Deep learning, Cambridge, MA, USA, The MIT Press
3. Cisco (2016) "Cisco visual networking index: forecast and methodology 2015–2020", White Paper
4. Wang P, Cao B, Yu P, Sun L, Bao W, Zhu Zhu X (2018) "Deep learning towards mobile applications." 2018 IEEE 38th, international conference on distributed computing systems (ICDCS), Vienna, pp 1385–1393
5. He K, Gkioxari G, Dollr P, Girshick R (2017) "Mask r-cnn", Proceedings of the IEEE international conference on computer vision (ICCV)
6. Ding C, Liao S, Wang Y, Li Z, Liu N, Zhuo Y, Wang C, Qian X, Bai Y, Yuan G, Ma X, Zhang Y, Tang J, Qiu Q, Lin X, Yuan B (2017) "Circnn: accelerating and compressing deep neural networks using block-circulant weight matrices", Proceedings of the 50th annual IEEE/ACM international symposium on microarchitecture (MICRO). ACM, pp 395–408
7. Lane ND, Bhattacharya S, Mathur A, Georgiev P, Forlivesi C, Kawsar F (2017) Squeezing deep learning into mobile and embedded devices. IEEE Pervasive Comput 16(3):82–88
8. Kingma DP, Ba J (2015) "Adam: a method for stochastic optimization". Int Conf Learn Represent (ICLR)
9. Han S, Pool J, Tran J, Dally WJ (2015) "Learning both weights and connections for efficient neural networks", Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS), pp. 1135–1143
10. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale Image recognition. CoRR, abs/1409.1556
11. Sun Y, Liu Y, Wang G, Zhang H (2017)" Deep learning for plant identification in natural environment", School of Information Science and Technology, Beijing Forestry University
12. Joly A, Goëau H, Bonnet P et al (2014) Interactive plant identification based on social image data. Eco Inform 23:22–34
13. Goëau H, Bonnet P, Joly A (2015) "LifeCLEF plant identification task 2015." In Proceedings of the conference and labs of the evaluation forum (CLEF '15)
14. Goëau H, Bonnet P, Joly A "Plant identification in an open-world (life clef 2016)." In Proceedings of the CLEF working notes, 2016
15. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S et al (2015) Going deeper with convolutions, 2015 IEEE conference on computer vision and pattern recognition (CVPR), pp 1–9

16. Deshmukh R, Deshmukh M (2015) "Detection of paddy leaf diseases. Int Conf Adv Sci Technol
17. Goëau H, Bonnet P, Joly A (2016) "Plant identification in an open-world (lifeclef 2016)." In Proceedings of the CLEF working notes, 2016
18. Kumar N, Belhumeur PN, Biswas A, Jacobs DW, Kress WJ et al. (2012) Leafsnap: a computer vision system for automatic plant species identification, European Conference on Computer Vision, pp 502–516
19. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. In: Computer vision–ECCV 2014, pp 818–833. Springer
20. Lee SH, Chang YL, Chan CS, Remagnino P (2016) Plant identification system based on a convolutional neural network for the life clef 2016 plant classification task
21. Krizhevsky A, Sutskever I, Hinton GE (2012) "Imagenet classification with deep convolutional neural networks". In Advances in Neural Information Processing Systems, pp 1097–1105
22. http://www.image-net.org/challenges/LSVRC/2012/
23. Krizhevsky A, Sutskever I, Hinton GE Advances in neural information processing systems. MIT Press; Cambridge, MA, USA, 2012. Imagenet Classification with Deep Convolutional Neural Networks; pp 1097–1105.

Modified Pre-emptive Priority-Based Round Robin Scheduling Algorithms with Dynamic Time Quantum: MPPBRRACBDQ and MPPBRRALBTDQ

Sudharshan Nagarajan, R. Akhil Krishna, M. Sivagami, and N. Maheswari

Abstract Scheduling is an essential part of an operating system. The main aim of scheduling is to schedule jobs and allot them to the CPU in a balanced way. It follows a set of rules to give prioritization to various processes. This would ensure that a system serves all requests and would achieve a certain quality of service with the existing hardware. In this paper, analysis of two Round Robin Scheduling Algorithms is carried out and modifications are made to improve the two algorithms which result in a quicker completion of the processes. The performance of the proposed algorithms, MPPBRRACBDQ and MPPBRRALBTDQ, is analyzed by comparing with existing well-known algorithms such as FCFS, SJF, PS and RR and also OMDRRS (Goel et al. in Int J Adv Comput Sci Appl 7:216–221, 2016) and improved Round Robin (H. B. Parekh, S. Chaudhari (2016) Improved round robin CPU scheduling algorithm. International Conference on Global Trends in Signal Processing, Information Computing and Communication, 184–187.). In MPPBRRACBDQ, the time quantum is varied dynamically based on the remaining burst time and is multiplied by two, while in MPPBRRALBTDQ, the time quantum is dynamically assigned based on the burst time, i.e., the least burst time among the processes is chosen to be the time quantum. The results prove that the new proposed algorithms perform better than the existing algorithms. Notable improvements are observed in waiting time (WT), turnaround time (TAT) and the number of context switches. The TAT, WT and context switches are compared and graphically represented. This research is an attempt to enhance the efficiency of the scheduling by developing and implementing new scheduling algorithms.

S. Nagarajan (✉) · R. Akhil Krishna · M. Sivagami · N. Maheswari
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, India

M. Sivagami
e-mail: msivagami@vit.ac.in

N. Maheswari
e-mail: maheswari.n@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_20

Keywords Scheduling algorithms · Round Robin · Pre-emptive · Priority · Dynamic time quantum · CPU scheduling

1 Introduction

CPU scheduling is the process of assigning a single resource among many jobs. Scheduling monitors the method of allocation and time span of each task. The main aim of scheduling is to optimize the performance of the CPU in a system. It is essential to optimize the system and to make full use of the resources. Many algorithms are available for CPU scheduling, each with its own advantages and disadvantages.

Process scheduling is a term used to portray the activities of the process manager. The process manager handles the selection of various processes based on a particular strategy and also the removal of the processes from the CPU.

One of the essential parts of multiprogramming operating systems is process scheduling. More than one process can be loaded into the executable memory at a time in multiprogramming operating systems, and the processes which are loaded into the executable memory share the CPU using time multiplexing.

Round Robin algorithm is a pre-emptive process scheduling algorithm that works based on the concept of time quantum. The time defined in the time quantum specifies the duration for which a process can run. The processes are pre-empted once they get executed for the time specified by the time quantum by context switching.

The main aim here is to survey existing modifications of Round Robin algorithms and proposes new algorithms to improve scheduling further. The proposed algorithms are modified pre-emptive priority-based Round Robin algorithm with.

i) Conditional-Based Dynamic Time Quantum
ii) Least Burst Time as Dynamic Time Quantum.

2 Related Work

There has been a lot of algorithms which have been used to schedule the processes. One of the most widely used scheduling algorithms is Round Robin and many research papers have improved upon the original Round Robin algorithm. Many of these Round Robin algorithms mainly focus on the time quantum. A dynamic Round Robin algorithm was developed using a variant time quantum [7] approach and later a modified median Round Robin algorithm (MMRRA) [3], which used a variable quantum time was proposed to eliminate the shortcomings of dynamic Round Robin. This included the average response time as a criterion for comparison. In another research [9], an optimal time quantum was chosen such that the time allotted to the processes were equal. Some proposed algorithms [4] have chosen the time quantum such that their main area of focus remained on stability with respect to mean turnaround time and waiting time. In a few proposed algorithms [10], the time quantum was calculated by a fixed mathematical formula which increased the performance of the system. Many other algorithms [6, 8, 11] have increased the CPU utilization and throughput while decreasing the response time, waiting time, turnaround time and the number of context switches.

One such improved Round Robin algorithm [2] was proposed by Harshal Bharatkumar Parekh and Sheetal Chaudhari in 2016. This algorithm combined three existing scheduling algorithms and it resulted in a new algorithm which reduced the time a process spent in the waiting state. The number of context switches was also reduced in this algorithm.

The OMDRRS algorithm [1] was proposed by Neetu Goel and Dr. R. B. Garg in 2016. This algorithm combined the concept of fundamental scheduling algorithms. Dynamic time slice (DTS) was estimated and it was used to allocate each process with a different time quantum with respect to context switching avoidance time, shortest CPU burst time and priority.

An algorithm which combined Round Robin with priority scheduling [5] was also developed. This algorithm performed better than both the Round Robin and priority scheduling algorithms. A novel CPU scheduling algorithm [13] was proposed which acted as both pre-emptive and non-pre-emptive based on the time of arrival. This algorithm improved efficiency in real-time uni-processor multi-programming operating system.

There could be more efficient algorithms which are capable of reducing the number of context switches, turnaround time, waiting time, and in turn, the entire execution time of all the processes. This paper proposes two such algorithms.

3 Proposed Algorithms

This paper proposes two pre-emptive priority-based Round Robin algorithms, MPPBRRACBDQ and MPPBRRALBTDQ, for single processor CPU scheduling. New scheduling algorithms MPPBRRACBDQ and MPPBRRALBTDQ have been proposed to improve the performance of RR and PS algorithms.

3.1 Modified Pre-Emptive Priority-Based Round Robin Algorithm with Conditional-Based Dynamic Time Quantum (MPPBRRACBDQ)

MPPBRRACBDQ combines the concept of fundamental scheduling algorithms. Dynamic time slice (DTS) is estimated using shortest CPU burst time, context switching avoidance time, and priority and is used to allocate different time quantum for each of the processes. This algorithm is based on the OMDRRS algorithm [1] and it has been modified to increase the efficiency and performance of the algorithm.

The proposed algorithm executes the processes as pre-emptive factor-based scheduling till all the processes have arrived and then continues its execution as a non-pre-emptive-based factor scheduling algorithm. The factor analysis value has been modified to give more importance to the burst time than the arrival time and the priority. The multiplication constants for the factor analysis value were chosen using trial and error in such a way that they resulted in the least TAT and WT for all the problems. 50% of the burst time 25% each of the arrival time and priority are considered for the factor analysis value. The time quantum is varied dynamically based on the remaining burst time and is multiplied by two.

Algorithm:
Step 1. Compute the Factor Analysis (F) to be equal to the sum of 50% burst time, 25% arrival time and 25% priority of the process.

Step 2. Reorder the processes in the ready queue such that the processes are sorted in ascending order of their factor values, i.e., the head of the ready queue should contain the least factor value.

Step 3–1. Set 'low' to be equal to the burst value of the first process in the RQ.

Step 3–2. Set 'high' to be equal to the burst value of the last process in the RQ.

Step 3–3. Calculate the average of low and high and assign it to TQ.

Step 4–1. Implement the pre-emptive-based approach till all the process are available and then continue the non-pre-emptive-based algorithm.

Step 4–2. Assign the time quantum and for each process assign k to be equal to TQ.

Step 5–1. If the burst time of the process is less than k, then allocate the CPU to that process till it is completed.

Step 5–2. Else if the remaining burst time of the process is less than half of k, then allocate the CPU again to that process till it is completed.

Step 5–3. Else, the process occupies the CPU till the time quantum and it is then added to the RQ for the next round of execution based on the remaining burst time in ascending order. The time quantum is changed dynamically using the formula, TQ = TQ *2.

Step 5–4. Assign TQ to K and continue the execution from Step 4.

3.2 *Modified Pre-Emptive Priority-Based Round Robin Algorithm with Least Burst Time as Dynamic Time Quantum (MPPBRRALBTDQ)*

MPPBRRALBTDQ is implemented as an enhanced version of the Round Robin algorithm. A priority level of low, medium or high is given to each of the processes and the time quantum is decided based on the priority level for that process and it is executed. A dynamic time quantum is selected that allows the complete execution of a process provided its remaining execution time is lesser than or equal to 0.2th of its actual time. The time quantum is taken as 0.8th fraction of the maximum burst time from the available processes in the ready queue. Processes in the ready queue with burst time greater than the time quantum are kept on hold, while the CPU is assigned to the other smaller processes by the scheduler. Now, after the execution of the smaller processes is completed, the maximum burst time is assigned to the time quantum. This algorithm is based on the improved Round Robin algorithm [2] and it has been modified to increase the efficiency and performance of the algorithm.

The proposed algorithm assigns processes to the CPU based on its burst (execution) time, that is, using shortest job first (SJF) algorithm and also assigns the time quantum dynamically based on burst time, i.e., the least burst time among the processes is chosen to be the time quantum.

Algorithm:
Step 1–1. For every process in the Job Queue, initialize the flag to FALSE.

Step 1–2. Assign the time quantum dynamically based on burst time, i.e., the least burst time among the processes is chosen to be the time quantum.

Step 2. Set Priority to be equal to 0.8 times the time quantum and set the High Priority to be equal to 1.2 times the time quantum.

Step 3–1. Find all the processes in the Ready Queue (RQ) having very low BT and assign those processes to the CPU to complete the execution, that is, using Shortest Job First (SJF) algorithm.

Step 3–2. Set flag to TRUE and remove the process from the RQ.

Step 4. Else, allot the CPU in Round Robin to the next process in the Ready Queue.

Step 5–1. If BT is lesser than or equal to the TQ, then complete the entire execution of the process for its remaining burst time.

Step 5–2. Set flag to TRUE and remove the process from the RQ.

Step 6–1. If BT is greater than the TQ and BT is lesser than or equal to 1.3 times the TQ, and if Priority is HIGH, then complete the entire execution of the process for its remaining burst time.

Step 6–2. Set flag to TRUE and remove the process from the RQ.

Step 7–1. If BT is greater than the TQ and BT is lesser than or equal to 1.2 times the TQ, and if Priority is LOW or MEDIUM, then complete the entire execution of the process for its remaining burst time.

Step 7–2. Set flag to TRUE and remove the process from the RQ.

Step 8. Else, execute the process till its assigned Time Quantum and switch to the next process in the Ready Queue until the Ready Queue is empty.

4 Result and Analysis

The modified algorithms have been implemented using C Language on Linux environment. The metrics WT, TAT and the number of context switches have been used to analyze the performance of the proposed algorithms and the results of the proposed

PID	BT	AT	Priority	PID	BT	AT	Priority
1	23	0	3	26	16	8	3
2	34	5	1	27	33	5	8
3	34	3	3	28	12	3	7
4	12	6	4	29	22	6	2
5	8	8	2	30	19	9	9
6	10	4	5	31	34	5	5
7	31	1	1	32	40	9	4
8	23	2	4	33	12	1	8
9	9	3	5	34	15	3	3
10	16	6	1	35	7	5	6
11	1	5	2	36	14	3	4
12	12	8	3	37	16	5	6
13	15	9	9	38	28	2	8
14	6	6	1	39	35	6	9
15	7	2	5	40	16	8	7
16	9	3	4	41	20	2	4
17	11	5	8	42	15	3	2
18	7	8	9	43	4	2	3
19	4	9	6	44	6	3	1
20	15	5	2	45	23	5	5
21	20	6	4	46	16	8	7
22	14	3	5	47	11	2	8
23	7	2	6	48	20	5	9
24	24	1	2	49	30	2	2
25	22	5	5	50	3	5	4

Fig. 1 Problem—PID, BT, AT, Priority of 50 processes

algorithms for the three datasets have been compared with the results of FCFS, PS, RR, SJF(NP), SJF(P) and also other algorithms [1, 2] for the same datasets.

The results from comparison of algorithms are tabulated and shown in the graphs (Figs. 1, 2, 3, 4, 5 and 6).

4.1 Example 1

(See Figs. 1, 2, 3, 4 and 5).

4.2 Example 2

(See Table 1 and Fig. 6).

PID	X	1	33	43	43	43	11	43	50	50	50	44	14	15	23	19
End Time	0	1	2	3	4	5	6	7	8	9	10	16	22	29	36	40

PID	16	35	5	9	6	18	47	33	4	28	12	17	36	42
End Time	49	56	64	73	83	90	101	112	124	136	148	159	173	188

PID	22	34	20	10	26	37	41	40	46	13	1	21	24
End Time	202	217	232	248	264	280	300	316	332	347	369	389	413

PID	8	29	25	48	30	45	7	49	38	2	3	31	27
End Time	436	458	480	500	519	542	563	593	621	655	689	723	756

PID	39	32	7
End Time	791	831	841

Each process (i.e. PID 1 to 50) is represented in a different shad

Fig. 2 Gantt chart for the MPPBRRACBDQ algorithm

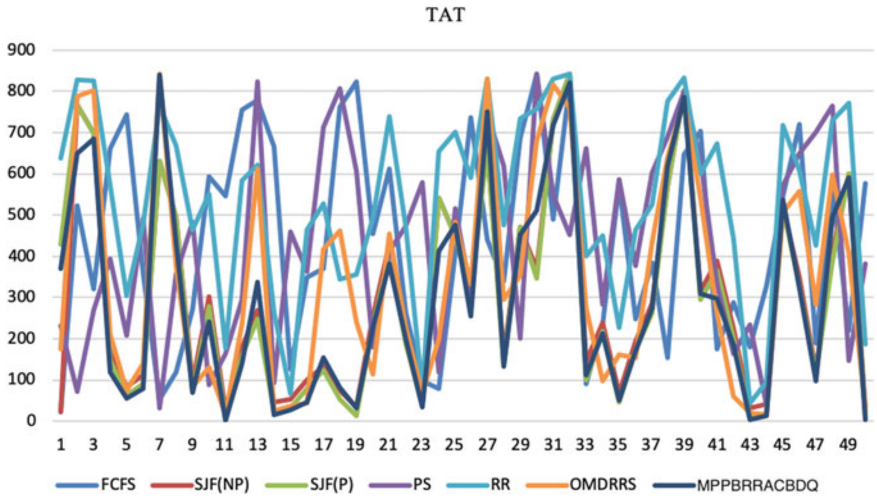


Fig. 3 Turnaround time for different algorithms

The proposed algorithms give us a better average TAT and average WT than FCFS, SJF, PS, RR and the other algorithms [1, 2]. The number of context switches has also been greatly reduced, which as a result reduces the overall execution time. The proposed algorithms greatly improve upon the existing algorithms, as seen from the results.

The proposed MPPBRRACBDQ and MPPBRRALBTDQ algorithms performed better than the popular RR and PS algorithms. Drastic improvements are noticed

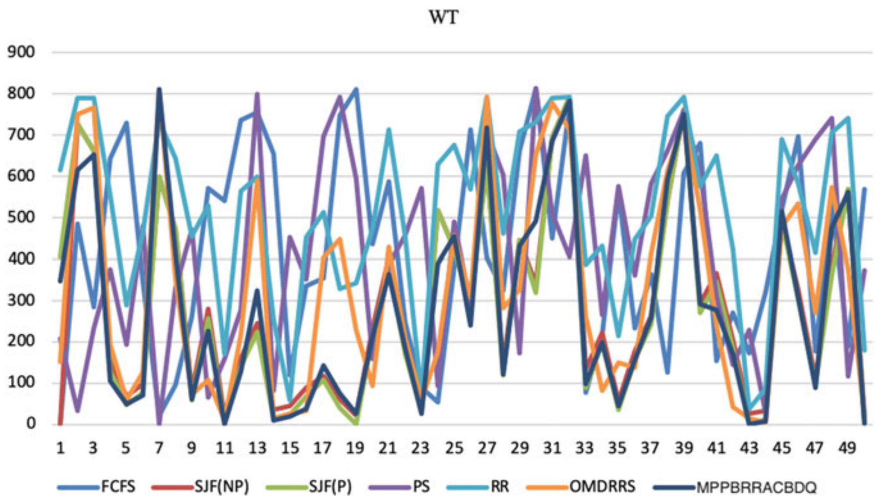
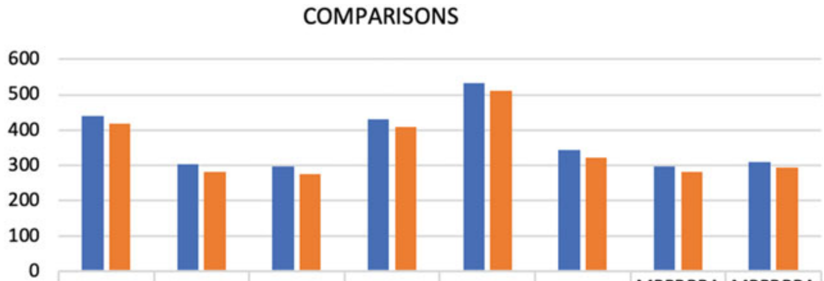


Fig. 4 Waiting time for different algorithms



	FCFS	SJF(NP)	SJF(P)	PS	RR	OMDRRS	MPPBRRA CBDQ	MPPBRRA LBTDQ
Avg. TAT	438.72	302.54	295.68	431.36	532.98	342.92	296.64	310.5
Avg. WT	417.3	281.12	274.26	409.94	511.56	321.5	279.82	293.68

Fig. 5 Comparison between Avg. TAT and Avg. WT for different algorithms

with respect to response time, turnaround time, waiting time and context switching in this algorithm when compared to RR and PS. The time complexity of the MPPBRRACBDQ algorithm is equal is O(n) which is similar to that of a linear sorting algorithm. The throughput of the MPPBRRALBTDQ algorithm is increased by executing processes which have smaller remaining burst time or smaller burst time with priority by assigning the process to the CPU when it becomes available but not pre-empting the process which is currently running. This algorithm completes the execution of the processes when the remaining burst time is very less and hence reduces the number of context switches.

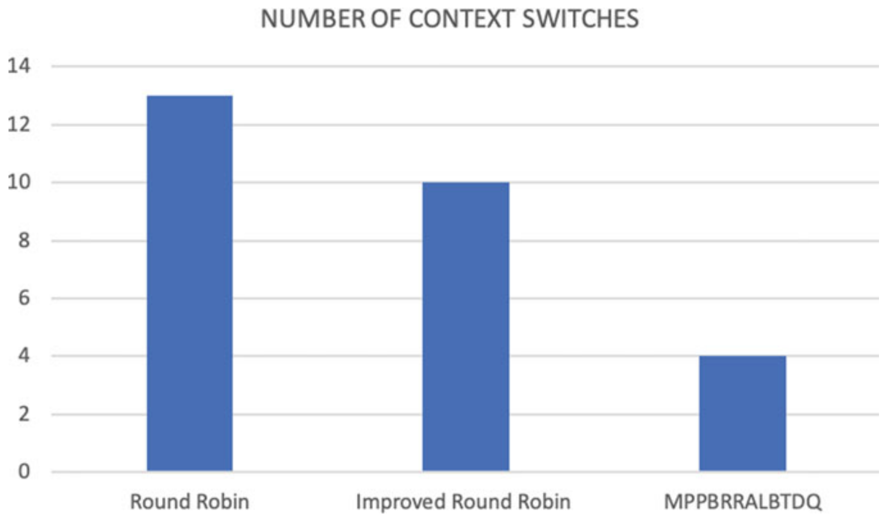


Fig. 6 Comparison between the number of context switches for the different algorithms

Table 1 Problem—PID, BT, AT, Priority of 50 processes

PID	BT	AT	Priority
1	550	0	3
2	800	200	1
3	200	100	3
4	2600	400	2
5	1600	0	2

Figure 5 gives a complete comparison of average TAT and average WT for all the scheduling algorithms such as FCFS, SJF, PS, RR, OMDRRS, and the two algorithms proposed in this paper, MPPBRRACBDQ and MPPBRRALBTDQ. It can be inferred that the average WT and TAT for the two algorithms are relatively low compared to the other algorithms. Figure 6 shows a comparison of the number of context switches between RR, improved RR and MPPBRRALBTDQ. The number of context switches in the proposed algorithm is lesser than half of improved RR and lesser than one-thirds of RR. This shows that the number of context switches has been improved greatly in the proposed algorithm compared to the existing variations of Round Robin algorithms.

Now, comparing the two algorithms proposed in this paper, i.e., MPPBRRACBDQ and MPPBRRALBTDQ, we see that the proposed MPPBRRACBDQ is better in most cases than the proposed MPPBRRALBTDQ. This inference is based on the average TAT and average WT of both the algorithms.

5 Conclusion

This paper analyses and compares two CPU scheduling Round Robin-based algorithms with other algorithms and improves those two algorithms significantly. The proposed algorithms and the other scheduling algorithms are executed on the simulator. The results indicate that the improvement to algorithms [1, 2] proposed in this paper, perform better than the existing algorithms based on mean TAT, mean WT and number of context switches which reduces overhead, and concludes that the proposed algorithms MPPBRRACBDQ and MPPBRRALBTDQ are superior then commonly used algorithms. On comparing the two algorithms proposed in this paper, we conclude that MPPBRRACBDQ performs better than MPPBRRALBTDQ. These two modified algorithms have been implemented and tested with single CPU and can further be extended and improved for multicore systems.

References

1. Goel N, Dr RB, Garg (2016) Performance analysis of CPU scheduling algorithms with novel OMDRRS algorithm. Int J Adv Comput Sci Appl 7(1):216–221
2. H. B. Parekh, S. Chaudhari (2016) Improved round robin CPU scheduling algorithm. *International Conference on Global Trends in Signal Processing*, Information Computing and Communication, 184–187.
3. H. Mora, S. E. Abdullahi, S. B. Junaidu (2017) Modified median round robin algorithm (MMRRA). *International Conference on Electronics Computer and Computation*, 1–7.
4. Eldahshan K, Elkader AA, Ghazy N (2017) Achieving stability in the round robin rlgorithm. International Journal of Computer Applications 172(6):15–20
5. Harshita J, Chowdhury S, Ramya G (2017) Survey on various scheduling algorithms. Imp J Interdiscip Res 3(4):1749–1752
6. Mishra KM, Khan AK (2012) An improved round robin CPU scheduling algorithm. Journal of Global Research in Computer Science 3(6):64–69
7. Khatri J (2016) An improved dynamic round robin CPU scheduling algorithm based on variant time quantum. IOSR Journal of Computer Engineering 18(6):35–40
8. Rahul Joshi, S. B. Tyagi (2015) Smart optimized round robin (SORR) CPU scheduling algorithm. International Journal of Advanced Research in Computer Science and Software Engineering, 5(7), 568-574.
9. Mohan S, Rajnesh S (2018) Optimized time quantum for dynamic round robin algorithm. International Journal for Research in Engineering Application & Management 4(3):318–321
10. Gull H, Iqbal SZ, Saeed S, Alqatani MA, Bamarouf Y (2018) Design and evaluation of CPU scheduling algorithms based on relative time quantum. J Comput Theor Nanosci 15(8):2483–2488
11. Rashid, Raihan, Zakaria, Hossain, Momin (2017) Improved performance of round robin CPU scheduling algorithm using non-preemptive SJF. International Journal of Scientific & Engineering Research, 8(12), 1734-1738.
12. Choudhary N, Geetika G, Yukti G, Soumil K (2018) A comparative study of various CPU scheduling algorithms using MOOS simulator. Journal of Advances in Shell Programming 5(3):1–5
13. Bandarupalli SB, Nutulapati NP (2012) A novel CPU scheduling algorithm-preemptive and non-preemptive. International Journal of Modern Engineering Research 2(6):4484–4490

14. Pradeep KJ, Rahul S (2018) Comparative study of CPU scheduling algorithms based on Markov Chain. International Journal for Research in Engineering Application & Management 4(3):668–672
15. Andrew S. Tanenbaum, Herbert Bos (2015) Modern Operating Systems, Fourth Edition.

Decentralized Classroom Using Blockchain

V. Sivasubramanian, V. Brindha Devi, M. Meenaloshini, and T. Ahath Khan

Abstract Education is the greatest and the most essential asset in this world that any human can attain. But the current education system has lots of unwanted intermediaries, which makes it difficult to deliver the concept to each and every individual. Current system is unable to bring out a particular student's potential to the fullest and is not able to make them showcase their capabilities and talents. Courses provided in the system lack flexibility and are not updated regularly. This leads to circulation of outdated and inappropriate information. To counter all these problems, our proposed system uses blockchain technology to establish trust between parties like teachers, students, and parents by eliminating intermediaries to provide seamless and up to date info with the power of distributed ledger. Our proposed system fragments the cities into small cells which will be governed by an education director who is responsible for recruiting teachers to tutor the student's in their cell. These teachers are provided with a web portal which allows them to invoke a smart contract where students can enroll themselves in any courses. Our proposed system provides two ways of attending the course one is the present online video tutorial with improved transparency. The next way provides a traditional classroom where the teacher rents a building or a hall which can hold all the students and provide the essential needs. Students can pay the fees of the course in the beginning or pay a daily fee. The money paid gets locked in the smart contract, and after completion of each class, the rent for that particular day is transferred to the building owner, and the wage for the teacher is also transferred instantaneously. In our proposed system, various new courses can be added in the go and will also abide by the central government's syllabus. Students just need to complete certain mandatory courses for their grade completion in the given period of time along with the topics of their choice.

Keywords Blockchain · Ethereum · Smart contract · Ether · Distributed ledger · Decentralized

---

V. Sivasubramanian · V. Brindha Devi (✉) · M. Meenaloshini · T. Ahath Khan
Information Technology, Sri Sairam Institute of Technology, Chennai, India
e-mail: hodit@sairamit.edu.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_21

1 Introduction

Population explosion, pollution, traffic, corruption, disease outbreaks, poverty, protest, and climatic conditions always hinder the education system in the present over the past and will continue in the future also. There is no accountability for the amount of fees paid in the current educational system. This leads to a lack of trust between the institute and the students. There are numerous unnecessary intermediaries who lead to unaccountable expenses, which are literally high for a normal student. There is no traceability of the student–teacher relationship, and trust is generalized in an institute rather than a teacher as a person. Area per person in Hong Kong is just 40 ft^2, and under search conditions, they are not able to build schools and colleges. This concludes that the current centralized education system cannot survive in the near future, and a revolutionary change is necessary. Our solution will decentralize and disrupt the education system with the help of blockchain technology.

Our proposed system using ethereum [1] network uses blockchain [2] technology. Ethereum provides a distributed ledger where everyone in the network can view all the transactions between two parties, and the data stored over the network is immutable. Ethereum consists of smart contracts which regulate all the transactions without the presence of third parties which increases trust between the participants of the network. To change any data over the ethereum [1] network certain amount of ether is required.

Using blockchain technology, we are planning to make the education system decentralized and accessible to everyone. Our ultimate aim is to bring quality education to everyone and anyone irrespective of their social or financial status. Education is fundamental to the growth of any society. We plan to divide the given region into sectors, and several sectors put together will be governed and administered by an education administrator of that region. He will be the one responsible for all the sectors that lie within his region. Each sector will have a set of educators, who will be appointed by the education administrator himself. All the details of the educator will be uploaded in blockchain which is publically [3] to all. Each educator will handle sets of students. Students will be linked to the educator by the education administrator, and it is his responsibility to maintain a correct student to teacher ratio in all the sectors under his region. Salary to educators will be credited to them each day over the entire duration of the course. Based on the feedback from the students, educators will be ranked. There is no central entity called schools or colleges in this sector. Teachers can rent a place which is accessible by all students in that sector. Scenarios where teachers cannot take class can be converted into online class. After completion of a particular course, students will get their certificate which is uploaded in blockchain. Students, as they keep on completing courses in each level, will keep moving to higher levels of courses. Apart from this system, private tutoring will also be covered under our scheme, but it would not be monitored by the education administrator. By implementing this scheme, we intend to bring transparency in the education system, removing all the middle men and profit seeking organizations who have made education into a business. Our goal is to bring education directly to the students. It should all be kept between the student and the teacher.

2 Literature Survey

A. L. Franzoni et al., C. Cárdenas et al., and A. Almazan et al. tested their system with 1518 teachers who were undergoing teacher training and stored their certification details in blockchain. Their system only stores their certificate and their performance in the course but failed to record the outcomes. Our proposed system records feedback and results of each and every batch of students who were tutored by the staff along with the interview details and selection criteria of the teacher [4].

W. Zeng et al., Q. Meng et al., and R. Li et al. have used face recognition along with RFID reader, in order to record attendance in the classroom. We have also used the facial recognition system to mark the attendance [5].

In 2017, N. Songkram et al., researched classrooms hosted virtually to improve the current skills of high school students. His learning process made students to understand new concepts and get a good grip of it; then the students would create something which would be related to the concept, and finally, it is evaluated. In our system, all these steps are included from the staffs end to tutor the students in the well-organized way [6].

In 2016, Mahesh et al. [7] proposed a system to monitor whether a student is using a smartphone for non-academic activity during study time using their application. We have used this idea to stop students' other focus.

C. L. Gunn et al. and J. Raven et al. proposed that the students and faculty should change to the new smart learning system, and his results were predominantly positive. We have included smart and active learning in our smart classroom system [8].

In 2011, Yang and Chen [9] presented a novel auto feedback system. In this system, they used PTZ cameras to detect whether a particular student is paying attention to the class or not. We have also used this technique in our system to detect the behavior of the student and to conclude whether the class is interesting or not.

S. Liu et al., Y. Chen et al., H. Huang et al., L. Xiao et al., and X. Hei et al. developed a cyber-physical-social system that uses reinforcement learning techniques to give better learning guidance to students in smart classrooms. This system failed to hold record of the data, and our system have recorded all the minute transactions [10].

Yuan [11] has proposed the concept of cloud classroom, and we have opted for the Google glass in case of the quarantine period and any other emergency period.

Jayahari et al. [12] proposed a system where the instructors can upload the material in their system ahead of time and make learning fun, but they have failed to explain whether the students are benefited with it, and the end results of the video are not stored.

L. Chamba et al. and J. Aguilar et al. proposed a system with a special design of the agent that provides services of augmented reality to display, design, etc. We have also used the concept of augmented reality wherever there is practical need in students curriculum [13].

R. Guntha et al., B. Hariharan et al., and P. V. Rangan et al. proposed certain ways to remove echo, and these ways were tested on the classrooms in order to improve the audio communication of the students. In our proposed system, the faculty can perform some of these methodologies to improve the communication of the students [14].

Q. Liu et al., Q. Guan et al., X. Yang et al., H. Zhu et al., G. Green et al., and S. Yin et al. in 2018 came up with an idea "Education-Industry Cooperative System Based on Blockchain." They have proposed a system where knowledge, skills, and certificates of the students are shared with the industries with higher transparency. Our system also uses this ideology of sharing information not only with industries, but also with the public as a whole [3].

Kumar et al., M. P. Gupta et al., M. K. Singh et al., and J. Madaan et al. failed precisely as to how they would tackle the classroom during emergency. Our proposed system uses the power of blockchain and image processing to empower online classes where the student has to sit in front of their system for the scheduled course time, and their attendance is recorded in blockchain [15].

3 Proposed Implementation

Every cell in an area will have an education director who is responsible for the teachers tutoring the students in his cell. Cells are nothing but fragments of areas which will be represented by an education director along with a group of teachers for different courses. Teachers are selected by conducting examinations, demos conducted in front of students, and the selection criteria is uploaded in the blockchain network along with their degree certificates and a list of skills they possess. These skills will be reviewed by the students participating in the course. Our proposed system provides two options—Online classroom and offline classroom. Online classrooms will only be opted in case riots, quarantine period [11], or any other calamity.

The teachers present in the cell should propose a schedule for the course along with a fixed timing. Students can opt for the required course via the web portal built over the ethereum [1] network. The teacher must report in the web portal, and the students must also report their attendance in the teacher's web portal, and the absentees must give a statement through their web portal, and it is mandatory. Students' attendance will be identified by a unique id which is mapped to their biometrics. Students have to submit their assignments and projects in the blockchain network, and anyone can view these works and evaluate. This is done to improve the quality of the teacher, and anyone can rectify the students' errors and appreciate their work. When the student completes the course, they have to attend a competitive exam, and these results [4] would be stored in the blockchain, and a feedback system backed up by these results will help the education director and parents to evaluate the course tutor. For schools, each grade will have a certain time period, and the student has to fulfill the requirements in order to take a higher grade course. Flexibility of course will increase gradually over different grades, and the student can start specializing in a particular but should complete the mandatory course which is required to attend the next grade. Our proposed system plans to bring forth every student's potential by analyzing. Teacher's other main job is to analyze each and every student and update it over the blockchain network. Teachers will also comment on what will be best for the student to pursue as a carrier in future. Teachers can rent a suitable place in the cell with the opinions from the student, and the rental for the place will go through the smart contract to the building owner whose private ethereum [1] wallet address will be stored in the smart contract. Daily wage of the teacher is released from the smart contract after the class completion. If the student does not like the course, they can talk with the education director in the presence of the course teacher and can come to a conclusion. If the student decides to leave the course, his fees are transferred to his wallet with minimal deductions. It is not necessary that a student needs to complete a course in his present cell, and they can enroll in any cell. We have placed a PTZ [9] camera to monitor the behavior of the student with his permission.

Students after taking their classroom courses must opt for a sport or any events for two hours which will be overlooked by the physical directors who are recruited again by the education director. They can recommend able students to participate in district tournaments and provide training for a variety of competitions present outside the world but to do so certain courses must be completed.

In the case of an online classroom, the staff will intimate all the students 10 min before the class starts. The students must have a web camera installed in their system. With image processing, we can ascertain whether the student is watching the livestream until the end.

Our proposed system will have a separate space for building renter's where they can upload the address and images of their property along with certain quality checks like amenities present, sanitary standards, noise over the area, etc. To help the students during different scenarios helper's can be enlisted by the rental company, but the helper's wage will be directly transmitted from the smart contract to the helper's address.

Apart from the school system provided by our proposed solution, anyone can take any subject, and it is not necessary for the education director to check them. People will be given a platform to build their own reputation in teaching what they are good at. They can teach in their homes or the student's home; its all up to them.

4 Proposed Block Diagram

Our proposed architecture is built over the blockchain network, and the participants of the network are students, education directors, teachers, rental companies, and students. When a certain event is completed, the smart contract will update the details over the distributed network. To access the blockchain, network interfacing is done using react js along with node js (Fig. 1).

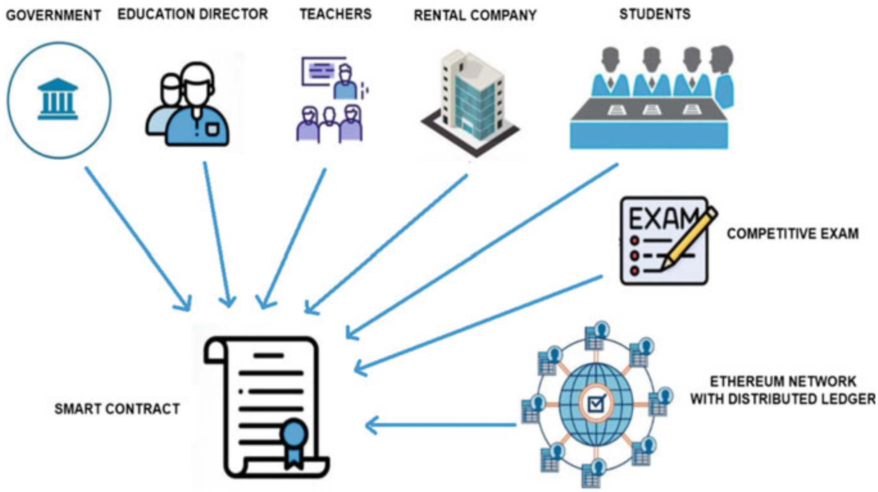


Fig. 1 Proposed system architecture

5 Proposed Algorithm

Step 1: The government district body must split the area into cells in proportion with the population.
Step 2: Certain exams and qualifications must be tested for the post of education directors, and these must be done by the district government body and finalize the education directors for the cells.
Step 3: Education director will do a transaction in the blockchain network stating that he is the right official, and his address can be viewed by the public.
Step 4: The education director lists the vacancies for different courses. The teachers can apply via the blockchain [2] network, and they have to upload their digital verified certificates. The education director will shortlist the candidates for demo and interview via an offline exam.
Step 5: The selected teacher's first take's a demo in front of students, and the students' opinions are recorded in the blockchain. Then, based on the feedback, the education director calls the teachers for an interview.
Step 6: Selected teachers have to provide their ethereum [1] wallet address along with their personal details which will be uploaded in the blockchain. Only these teachers can take the courses.
Step 7: The teacher publishes course details on the web portal. Parents can view his profile and past accomplishments and enroll their child under him for the particular course. While enrolling, they pay the full fees for the course or they can pay on a daily basis.
Step 8: Once the count has filled up to the limit, the teacher releases the schedule and the place where the class is going to happen, and if the majority accepts the proposition, the class will commence as per the schedule. If the students want the

proposition to be changed, they have to negotiate with the teacher, and this would be monitored by the education director.

Step 9: In case of offline class the students biometric, retina can be scanned, and if they are present, their attendance is reported and stored in the blockchain. In case of online class, face recognition and image processing techniques are used for marking attendance.

Step 10: The way teachers take class, give assignments, take field trips, sudden tests, and other events are stored as logs in the blockchain. Teachers give the feedback on each and every student on a regular basis which are updated in the students profile, and these details are only accessible by the teacher's, education director, and parents.

Step 11: Class ends when the staff and the students mark their leaving attendance. As soon as the class ends, the money is transferred to the teachers account and the rental building's owner and the helpers.

Step 12: When the student completes the course, he or she is given a certain period of time, and a competitive exam is held common for different cells for every course. The question papers are generated by different teachers in a given format, and these papers are fed to machine learning algorithms, and a randomized question paper will be generated.

Step 13: The results [4] of the examination will be updated over the student's portal, and the teacher's portal will be updated also. After the results, the students will be allowed to fill a feedback form with multiple questions regarding the course and its tutor. This feedback is again fed to the blockchain and is permanently viewable in the teacher's log.

Step 14: In case if the staff do not produce satisfactory results [4] and the majority of the students are against the way a teacher acts, this issue will be taken by the education director, and a meeting will be held in presence of student teacher and education director. Based on the conclusion, the staff will be given warning or the staff can be changed for that particular course resulting in blackmarks in the teacher's profile.

Step 15: In case the student is unable to attend the course, he or she has enrolled, the money is transferred to their account if paid, and they can take the course anytime within the given year.

6 Performance Evaluation

We have compared our system, and the results have been portrayed in our graph. Our system had starting issues, and people had some issues to get used to it. As months passed, people adapted to our system, and they were satisfied due to the flexibility our system provides. They were able to schedule their day in amor efficient way (Fig. 2).

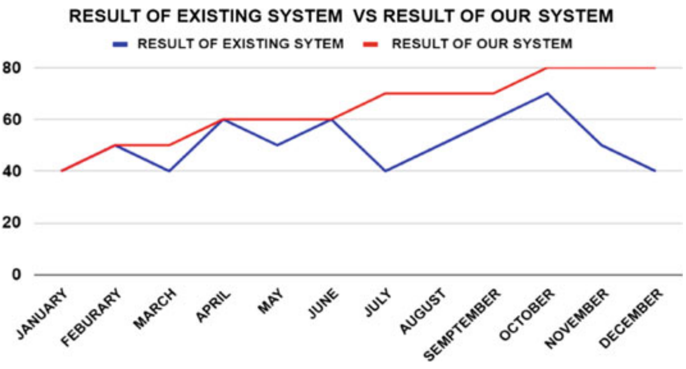


Fig. 2 Comparison using Graph

7 Result Analysis

Schools these days provide less flexibility on the variety of courses, and the syllabi are outdated. Our proposed system solution provides a great range of subjects to learn along with the basic knowledge to be qualified to the next grade. There are lots of intermediaries present in the current education system which is unable to build a strong trust between the parent and the teachers. Trust is kept over the centralized figure which are the schools, and trust is not built over the teacher. Our system cuts the intermediaries which will also provide a better pay for teachers, and the payment split up for the course is revealed crystal clear to the parents which increases the parents trust over the teacher. These days there is no place to construct enough school, but there is a steady increase in population, and our current education system cannot facilitate all the students. A merit system is built in our proposed solution which weighs every student based on their accomplishments, skills, and their scores, and the students wallet is provided with a certain number of credits which helps them to pay for their courses. In the previous education system, there was no clear view or awareness of how merit was given and who received it. In Hong Kong, the average living space per person is 50 square feet, in such a condition, there is no enough place to build a proper school. Our proposed solution reuses empty spaces and public playgrounds which reduces the cost of maintaining education in the country. Since our proposed system splits the area into cells, the number of jobs present in the education field increases. Since all the results [4] and feedback of a course are stored in the blockchain [2] along with the material and assignments provided by the teacher, the teacher needs to be precise and up to date with information as there is a healthy competition born out of our proposed solution. During riots or quarantine period, schools and colleges are shut down for a long period of time, and our proposed system effectively utilizes the present technologies available and continues the flow of educating the children. In the current education system, we have to go under a lot of procedures to transfer to a different school. But by our solution, any student can take any course from any cell as long he has a unique identity card of the country.

Students will have lots of free time as they do not need to attend schools or colleges from morning to evening, and they can use their free time effectively to learn new things. Our proposed system will reduce the stress on students and promote sports and physical activities as part of everyday's curriculum. Since the staff analyze every student's potential, when they take the course it is easy to bring the full potential of the child out, and it will be easy for the students to know their pro's and con's while deciding their career or job. The faculty can provide practical training with the help of technologies like augmented reality and virtual reality.

8 Conclusion

This paper mainly focuses on shifting the traditional education system to a dynamic, flexible, and efficient system where trust is not centralized by an organization anymore. There are certain obstacles in the course of implementing our proposed solution. In the current fast paced world, everything happens in a matter of seconds, but ethereum [1] networks are unable to cope up with it as the hashing and changing the distributed ledger takes a certain amount of time. Technology development in the field of blockchain [2] is rapid, and the above obstacle would whittle down as the ethereum [1] network is under constant development. During the initial stage of our proposed solution, it would be difficult over the participants of the network would find it to get accustomed to the system, awareness must be created among the public [3], teachers must be trained to use such a system, and teachers must create a comfortable atmosphere for the student. There would be some hassle in the initial stage to select the correct administration, tutors and set up the requirements of the course like lab, etc. If initial capital is applied correctly, then the system can be scaled, and changes can be done easier over the future run.

References

1. Buterin V (2013) A next generation smart contract & decentralized application platform
2. Satoshi Nakamoto satoshin@gmx.com www.bitcoin.org, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008
3. Liu Q, Guan Q, Yang X, Zhu H, Green G, Yin S (2018) Education-industry cooperative system based on blockchain. In: 2018 1st IEEE international conference on hot information-centric networking (HotICN), Shenzhen, 2018, pp 207–211.
4. Franzoni AL, Cárdenas C, Almazan A (2019) Using blockchain to store teachers' certification in basic education in Mexico. In: 2019 IEEE 19th international conference on advanced learning technologies (ICALT), Maceió, Brazil, 2019, pp. 217–218.
5. Zeng W, Meng Q, Li R (2019) Design of intelligent classroom attendance system based on face recognition. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, 2019, pp 611–615.

6. Songkram N (2017) Virtual smart classroom to enhance 21st century skills in learning and innovation for higher education learners. In: 2017 Tenth international conference on mobile computing and ubiquitous network (ICMU), Toyama, 2017, pp 1–4.
7. Mahesh G, Jayahari KR, Bijlani K (2016) A smart phone integrated smart classroom. In: 2016 10th International conference on next generation mobile applications, security and technologies (NGMAST), Cardiff, 2016, pp 88–93
8. GunnCL, Raven J (2017) Smart education: introducing active learning engineering classrooms in the Middle East. In: 2017 Fourth HCT Information Technology Trends (ITT), Al Ain, 2017, pp 1–4.
9. Yang S, Chen L (2011) A face and eye detection based feedback system for smart classroom. In: Proceedings of 2011 international conference on electronic & mechanical engineering and information technology, Harbin, 2011, pp 571–574
10. Liu S,Chen Y, Huang H, Xiao L, Hei X (2018) Towards smart educational recommendations with reinforcement learning in classroom. In: 2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE), Wollongong, NSW, 2018, pp 1079–1084.
11. Yuan Y (2016) Cloud classroom boost online learning and educational resources sharing. In: 2016 International symposium on educational technology (ISET), Beijing, 2016, pp 80–83.
12. Jayahari KR, Nair SP, Bijlani K (2017) Solution for automating teaching environment setup. In: 2017 IEEE international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM), Chennai, 2017, pp 122–127
13. Chamba L, Aguilar J (2016) Design of an augmented reality component from the theory of agents for smart classrooms. IEEE Lat Am Trans 14(8):3826–3837
14. Guntha R, Hariharan B, Rangan PV (2016) Analysis of echo cancellation techniques in multi-perspective smart classroom. In: 2016 international conference on advances in computing, communications and informatics (ICACCI), Jaipur, 2016, pp 1135–1140.
15. Kumar H, Gupta MP, Singh MK, Madaan J (2017) Education in emergencies: smart learning solutions and role of governance to mitigate the challenges. In: 2017 3rd international conference on advances in computing, communication & automation (ICACCA) (Fall), Dehradun, 2017, pp 1–6.

A Novel Algorithm for Automatic Essay Grading Using Natural Language Processing Techniques

D. Jayashree, O. Pandithurai, G. Vishwak Senan, K. Pranav Kumar, and T. Sundaramoorthy

Abstract Evaluation of an English essay is one of the important and complex tasks which is done manually by skilled and efficient professors and faculties till date. The growth of science and technologies enables to automatic evaluation of an English essay using natural language processing (NLP) techniques (Rogovschi et al. in 2017 International joint conference on neural networks (IJCNN), Anchorage, AK, pp. 1628–1632, 2017 [1]; Banga and Mehndiratta in 2017 International conference on information technology (ICIT), Bhubaneswar, pp. 264–267, 2017 [2]; Roy and Cheung in 2018 28th international telecommunication networks and applications conference (ITNAC), Sydney, NSW, pp. 1–6, 2018 [3]). The intelligent system—built upon NLP multiple neural network model—gives out generic evaluation and the topic/question correlation for any given English essay.

Keywords Automatic evaluation · Natural language processing · Generic evaluation · Topic/question correlation

1 Introduction

International examinations like Graduate Record Examination (GRE), International English Language Testing Systems (IELTS), etc., is gaining popularity day-by-day

D. Jayashree (✉) · O. Pandithurai · G. Vishwak Senan · K. Pranav Kumar · T. Sundaramoorthy
Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India
e-mail: jayashree.d@ritchennai.edu.in

O. Pandithurai
e-mail: pandics@ritchennai.edu.in

G. Vishwak Senan
e-mail: vishwaksenang.2016.cse@ritchennai.edu.in

K. Pranav Kumar
e-mail: pranavkumark.2016.cse@ritchennai.edu.in

T. Sundaramoorthy
e-mail: sundaramoorthyt.2016.cse@ritchennai.edu.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_22

as this examination's results are considered as criteria for various universities and companies. Therefore, the number of international students across various counties is taking the exams, increasing the count day-by-day, and there is a huge time buffer to evaluate their English essays and publish the results. To reduce the stress on the organization who are hosting these examinations and students to practice their writing skills, our project aims to evaluate the English essay so that the organization can focus their work in other aspects of examinations, and students can practice at free will.

2 Literature Survey

There are various existing NLP techniques which are considered to be best for text classification which is essential for evaluation of the essay. The initial text classification techniques were to classify parts-of-speech of a sentence.

2.1 N-Gram Model

The n-gram model otherwise known as Shannon's Markov Chain [4–6] estimates the probability of the occurrence of the next word given the n-1 words. Therefore, this is a probabilistic model with a good accuracy if trained from a sophisticated dataset.

2.2 Perceptron Layer

This model of finding the parts-of-speech tagger of a sentence uses 3-layer perceptron layer with n-inputs, n representing the total number of words in the dataset [7–9]. While the computational cost of training the network is drastically higher than the n-gram model but according to the previously conducted experiment [10], the accuracy was 99.4% without overfitting into the data. The accuracy was obtained by using the elastic hidden perceptron layer.

2.3 Long Short-Term Memory (LSTM)

Long short-term memory is a type of recurrent neural network (RNN) which is a very famous model to predict a series of a vector. This model has high computational requirement but yields a high accuracy for a large dataset. In a research, the low-cost hardware implementation of LSTM [11, 12] made the neural network to be more optimized and efficient for any given low-cost machine. This invokes the possibility of training the neural network locally rather than renting GPUs online. The low-cost LSTM is achieved by applying stochastic function rather than hyperbolic activation function.

3 System Architecture

The main goal of the project is achieving the five crucial factors for evaluating an English essay. The factors are

1. Grammar and spell check.
2. Sentence complexity.
3. Style continuity.
4. Usage of advanced lexical resources.
5. Coherence and cohesion.

The above-mentioned factors are checked by developing a dedicated model for each factor having the input of the whole essay. The primary services of the project (evaluation engine which evaluates the essay) are hosted in any public clouds, so that organizations can utilize our services too. Since there are five generic factors involved in evaluation of an essay, we will develop five neural network model which evaluates on its each of the factor.

Model 1—Grammar and spell check.
Model 2—Sentence complexity.
Model 3—Style continuity.
Model 4—Lexical resources.
Model 5—Coherence and cohesion.

Full architecture is implemented using python as the programming language. The architecture of the project is as given in Fig. 1.

4 Novel Algorithm

4.1 Model 1 Algorithm

Training Algorithm:

Step 1: The corpus dataset is scraped from oxford dictionary website.
Step 2: Separate the nouns, verbs, pronouns, proverb, adjective, and adverb into separate text file from the scarped dataset. Then the words are encoded using vectorization method.
Step 3: Load it into the python class train_model using pandas and allocate corresponding label coming from corresponding text file.

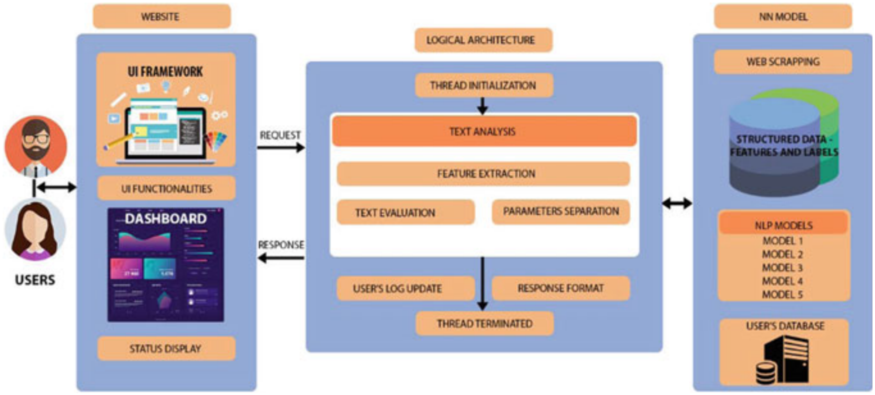


Fig. 1 System architecture

Step 4: Separate the labeled dataset into training and testing dataset in 7:3 ratio.
Step 5: Apply the feature extraction "Word2vec" model on the dataset (Ignore Noun for spelling correction).
Step 6: Save the trained model.

Consuming Algorithm:

Step 1: Get the sentences from the essay.
Step 2: Apply the trained model 1 on each sentence.
Step 3: If the sentence has any grammatical or spelling error, assign 0. Else 1
Step 4: Repeat Step 3 for all the sentences and count the total number of sentences.
Step 5: Count the total number of 1's and divide it by the total number of sentences.
Step 6: Multiply the value done on Step 5 by 10, and this score is projected to the user.

4.2 Model 2 Algorithm

Training Algorithm

Step 1: Download the lexical database from Shakespeare play and from famous English movie subtitles.
Step 2: Separate the lexicons into 5 buckets according to their difficulty level.
Step 3: Assign the vector value to each bucket randomly initially using embedding algorithm.
Step 4: Apply bidirectional LSTM on the word and train the words according to the bucket value.
Step 5: Save the model.

Consumption Algorithm

Step 1: Get the whole essay as the input.
Step 2: Tokenize the sentences from the essay.
Step 3: Apply the built model on the sentences and get the bucket value for each sentence.
Step 4: Average the bucket vector value from each sentence.
Step 5: Round off and provide the mark for the parameter—sentence complexity.

4.3 Model 3 Algorithm

Training Algorithm

Step 1: Download any english lexicon database from scraping the Internet for different accent words.
Step 2: Separate the words into various buckets according the style (American, British, Australian English variations).
Step 3: Apply multinomial Gaussian mixture model (GMM) [4] on the bucket to get probability distribution model ready.
Step 4: Save the model.

Consumption Algorithm

Step 1: Get the essay as the input.
Step 2: Tokenize the essay into sentences.
Step 3: Give the numbers

1—American English
2—British English
3—Australian English and so on for each sentence.

Step 4: Count the number of 1s, 2s, 3s on the number
Step 5: Assign the mark according to the table.
Step 6: Repeat step 5 for each number.
Step 7: Average out the mark for each category.
Step 8: Assign the value from step 7 and project it as a mark in style continuity (Table 1).

4.4 Model 4 Algorithm

Training Algorithm

Step 1: Download the lyrics of famous authors—Shakespeare—as our training dataset.

Table 1 Style continuity score sheet

Continuity percentage (%)	Mark assignment
0–20	2
21–40	4
41–60	6
61–80	8
81–100	10

Step 2: Separate into 5 buckets naming from primitive to advanced.
Step 3: Apply bag of words feature extraction algorithm into these separated databases.
Step 4: Save the model.

Consumption Algorithm

Step 1: Get the essay as the input.
Step 2: Tokenize the essay into list of sentences.
Step 3: For each sentence, find the most advanced word in that sentence
Step 4: Assign the vector value of the most advanced word for that sentence.
Step 5: Repeat step 3 and step 4 for all the sentence.
Step 6: Average out all the vector value of sentences.
Step 7: Project the step 6 score for the parameter—lexical resource.

4.5 Model 5 Algorithm

Training Algorithm

Step 1: Get the questions database by scrapping ETS GRE sample analytical writing assessment questions.
Step 2: Find the subject word in the question and initiate a random vector value for all the subject word from the question.
Step 3: Find all the relative words according to the initial vector and build a sequence for the root vector which the subjective word from question.
Step 4: Build the sequence by repeating step 3 considering the current vector value to be root node. Thus, a tree is build depth wise.
Step 5: Save the built model along with its corresponding root node which the question.

Consumption Algorithm

Step 1: Get the essay as the input from the user.
Step 2: Tokenize the essay into list of sentences.
Step 3: Get the initial vector value of the question assigned to the user.
Step 4: For each sentence, use the model to get the vector value for each sentence.

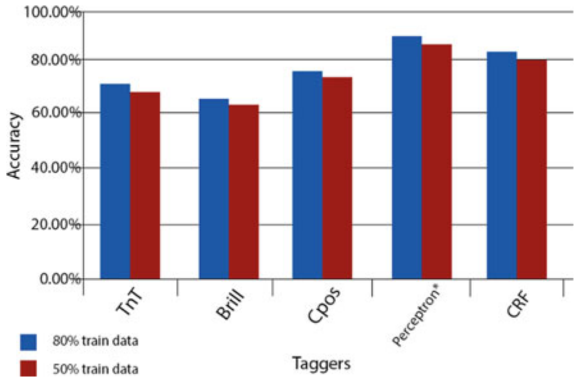


Fig. 2 POS tagger comparison

Step 5: From the node, traverse into the built model (tree) according the vector value of each sentence.
Step 6: If the path has minimum vector difference, then the sentences are most related to each other.
Step 7: Normalize the difference in the range of 0–10 and assign the end normalized value for the parameter coherence and cohesion.

5 Model Evaluation

There were variety of POS tagger produced various results, yet perceptron POS tagger yields out higher accuracy (Fig. 2).

Multinomial hidden Markov model (HMM) is a probability-based classification model which yields out grammatical mistakes in the sentence. Dictionary-based n-gram model accuracy depends solely on the dataset we used. The dataset to find the correct spelling is scrapped from several open dataset and movie scripts.

Guassian mixture model is a probabilistic clustering algorithm which takes random centroid and clusters the datapoint with the probability regarding the centroid. This classifies the level the lexicon used in the sentence.

Since bidirectional neural network is a generative neural network, it works for dynamic datapoints. The need for the dynamic datapoints is due to the complexity of the sentence varies from style to style.

6 Conclusion and Future Works

Evaluators take time in correcting the answers, and a lot of manpower is used for it. Also, the testing centers must allot places for evaluating the answers. All these steps take time, and more money needs to be spent. Thus, by the use of our essay grader, the test centers can organize these tests in a convenient way. This software solves the problem of manually correcting the answers and grading it. We aim to develop this software so that the English language testing centers such as IDP, British council can make use of this and evaluate the candidates writing skills, the way he has logically connected the sentences, vocabulary, etc. We have developed the software by keeping in mind the different constraints which the evaluators keep in mind while evaluating. The main working part is with the help of neural networks [13], and we have created individual network for each and every task being performed. When a candidate appears for the IELTS/TOEFL test, he will be asked to appear for four sections. They are listening, speaking, reading, and writing. In future, we can do the essay grading for multiple languages such as Tamil, Telugu, Malayalam, etc., which will be helpful for school education. As of now, speaking section is being evaluated when the candidate speaks. The evaluator will be asking few questions, and the candidate will be answering. This can be automated in future using speech processing. We can record the audio and apply ML algorithms to it to find the eloquence of the candidate. So, these are several future works which can be applied to enhance the product.

References

1. Rogovschi N, Kitazono J, Grozavu N, Omori T, Ozawa S (2017) t-distributed stochastic neighbor embedding spectral clustering. In: 2017 International joint conference on neural networks (IJCNN), Anchorage, AK, 2017, pp 1628–1632.
2. Banga R, Mehndiratta P (2017) Tagging efficiency analysis on part of speech taggers. In: 2017 International conference on information technology (ICIT), Bhubaneswar, 2017, pp 264–267.
3. Roy B, Cheung H (2018) A deep learning approach for intrusion detection in internet of things using bi-directional long short-term memory recurrent neural network. In: 2018 28th international telecommunication networks and applications conference (ITNAC), Sydney, NSW, 2018, pp 1–6
4. O'Sullivan JA, Mark K, Miller MI (1994) Markov random fields on graphs for natural languages. In: Proceedings of 1994 workshop on information theory and statistics, Alexandria, VA, USA, 1994, p 47
5. Mandita F, Abdullah HM, Anwar T, Assawinjaiptech P (2018) A novel algorithm for a grammar model checking using statistical Markov model. In: 2018 Seventh ICT international student project conference (ICT-ISPC), Nakhonpathom, 2018, pp 1–6.
6. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed representations of words and phrases and their compositionality. Adv Neural Inf Process Syst 26:3111–3119
7. Singh SP, Kumar A, Singh L, Bhargava M, Goyal K, Sharma B (2016) Frequency based spell checking and rule based grammar checking. In: 2016 International conference on electrical, electronics, and optimization techniques (ICEEOT), Chennai, 2016, pp 4435–4439.
8. Hassan A, Mahmood A (2017) Efficient deep learning model for text classification based on recurrent and convolutional layers. In: 2017 16th IEEE international conference on machine learning and applications (ICMLA), Cancun, 2017, pp 1108–1113.
9. Singh OP, Sinha R (2017) Sparse representation classification over discriminatively learned dictionary for language recognition. In: TENCON 2017—2017 IEEE Region 10 conference, Penang, 2017, pp 2632–2636.
10. Ma Q, Uchimoto K, Murata M, Isahara H (1999) Elastic neural networks for part of speech tagging. In: IJCNN'99. International joint conference on neural networks. Proceedings (Cat. No.99CH36339), Washington, DC, USA, vol 5, pp 2991–2996

11. Li S, Wang Q, Liu X, Chen J (2018) Low cost LSTM implementation based on stochastic computing for channel state information prediction. In: 2018 IEEE Asia Pacific conference on circuits and systems (APCCAS), Chengdu, 2018, pp 231–234.
12. Wang R, Pan W, Ma J, Wang H (2019) A synonym extraction method based on Intimacy*. In: 2019 IEEE international conference on agents (ICA), Jinan, China, 2019, pp 57–60.
13. Majewski M, Zurada JM (2008) Sentence recognition using artificial neural networks

Challenges in Adoption of Industry 4.0 in Apparel Supply Chain: A Cross Case Study

U. C. Amal Krishnan, Amal T. Sasidharan, Timal Prakash, and C. P. Vivek

Abstract The research mainly revolves around Industry 4.0 (I4.0) and challenges of its adoption in apparel supply chain in Indian textile industry. Research questions were formulated to uncover the issues faced by the small and medium enterprises (SMEs) of Indian textile industry in adopting I4.0. Attaining maximum efficiency in the bottom line operations is the most desired advantage of I4.0 in textile industry. Industry 4.0 is at the lowest priority for ready-made and garment industry because it is predominantly man power driven, and specialization is in a very less need of automation. It is found that cost of bringing latest technologies into the industry is detrimental at this point of time due to lower labour cost in the industry when compared to the available alternative technologies. There is a huge opportunity for future research in the area of ready-made and garment industry regarding the increase in productivity which is possible with I4.0 adoption.

Keywords Industry 4.0 · Indian textile industry · Operational efficiency · Internet of Things

1 Introduction

The research mainly revolves around Industry 4.0 (I4.0) and barriers to overcome for its adoption in apparel supply chain in Indian textile industry. As the name suggests, the fourth industrial revolution refers to the concept of machines being augmented with wireless connectivity and sensors, interconnected with a system that can visualize the entire production line and have an assistive intelligence to take decisions on its own in factories.

The two important factors that are related to I4.0 as mentioned in Iyer [1] are:

I. Data Integration—How SMEs integrate the process data, product data, quality data, plant data and logistics data with the newly adopting communication system in the manufacturing units.

U. C. Amal Krishnan (✉) · A. T. Sasidharan · T. Prakash · C. P. Vivek
Amrita School of Business, Amrita Vishwa Vidyapeetham, Coimbatore, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_23

II. Predictive Maintenance—How SMEs monitor the operating hours, number of products processed and time required for each unit of products to be produced.

Apart from the above factors, during the adoption of Industry 4.0, manufacturing SMEs should also examine the software platforms to be considered for collecting data and selecting right data for suitable decision making. As [1] countries, unlike the developed economies such as Germany, USA, Japan, etc., that are advanced in combining digital technologies and sustainable manufacturing processes, emerging powers like India should concentrate on comparative advantages and technologies that are critical in enhancing their growth. The effectiveness of addressing challenges in developing a manufacturing ecosystem by amalgamating latest trends of digital manufacturing and upskilled or reskilled human resources will also have a far-reaching influence on how they progress in the world market.

According to related studies, following are the improvements achieved by the eight value drivers of any business organization:

1. Resource/Process—Productivity increased by 3–5%.
2. Asset utilization—Total machine down time has been reduced by 30–50%.
3. Labour—Productivity has been improved by 45–55% through automation.
4. Inventories—Costs of inventory holding got reduced by 20–50%.
5. Quality—Costs incurred in quality testing has been reduced by 10–20%.
6. Supply/Demand match—Accuracy of forecasting supply and demand increased to more than 85%.
7. Time to market—Reduced by 20–50%.
8. Services/Aftersales—Maintenance costs have been managed to bring down by 10–40%.

As mentioned in Khan and Turowsk [2], while adopting I4.0 in an organization, routine processes should continue as before with tweaks and adjustments wherever required. Also, employees should be motivated to participate in Industry 4.0 projects. Khan and Turowsk [2] recommend a step-wise approach that involves solving current challenges at first and then to improve the current situation. This will enable the production to obtain benefits quickly and make it ready for future scenarios through gaining advantages if I4.0. According to Ganzarain and Errasti [3], the barriers in adoption of Industry 4.0 are employment disruption, high implementation cost, organizational and process changes, need for enhanced skills, lack of knowledge management system, security and privacy issues, lack of internal coverage and IT/IoT facilities, lack of standards and reference architecture. The findings of Ganzarain and Errasti [3] also shows a current need of guided support in developing an organization specific I4.0 vision for which they have recommended a three stage maturity model by incorporating envision, enable and enact as the key methodologies of their framework.

After referring different articles, the research gap that we found out was the consequences of I4.0 adoption in SME's and justification of non-adoption of the same. Research questions were formulated to uncover the issues faced by the SMEs of Indian textile industry in adopting I4.0. Through in depth interviews and observation, this study tries to find answers to the following questions with respect to textile industry.

- Why do adopters adopt I4.0 and what are the drivers?
- Why do non-adopters don't?
- What are the areas of business that need to be prepared?
- What are the costs of adoption?
- What are the consequences of adoption?
- What changes do adopters make in their operations?

2 Industry 4.0 in Textile Industry

Similar to any manufacturing sector, final product of the textile industry is also a combination of raw materials, processes and marketing and distribution. In order to attain maximum efficiency in the bottom line operations, every industry has to forecast each step coming under 3Ms—men, material and method and integrate those into their existing supply chain. This is not possible in every industry due to the difference in end market requirements and economic viability. For example, the technological advancements preferred in food industry may not be applicable in pharmaceutical industry. Same is the case with textile industry which is the area of research where this particular study revolves around.

In case of ready-made and garment industry even though the Industry 4.0 may be helpful, it is found that cost of bringing those technologies into the industry is detrimental at this point of time. The main reason for this is the lower labour cost in the industry when compared to alternative technologies available. At present, the industry 4.0 is at the lowest priority for ready-made and garment industry because it is predominantly man power driven as specialization is in a very less need of automation, and the number of people required to convert raw material into final product is high. This also explains the reason for the business to be dominant in countries with relatively less labour cost like Bangladesh and India which otherwise could have been moved to developed economies. Man power in the textile industry has been facilitated through partial automation in terms of stitching where some of the delicate processes are still done manually. One of such processes is the placement of cloth into the stitching machine which if automated would incur high cost to the operation. Adoption of economically viable technologies for the same would open up more opportunities for Industry 4.0 in near future.

The overall factory operation of Indian textile industry can be considered very close to Industry 4.0:

- Use of infrared rays in stitching (machine operation is automated with the detection of fabrics through IR rays which reduces the processing time).
- Automated pedal operation (Enables the semi-skilled labours to work with maximum efficiency).
- Absenteeism monitoring of employees.

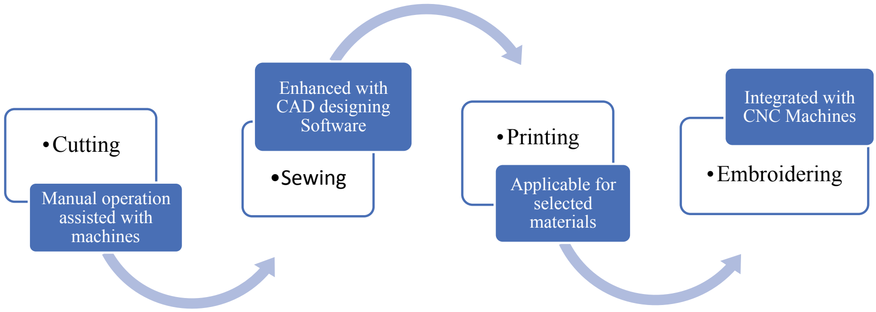


Fig. 1 Process flow in ready-made and garment industry

- Real-time efficiency monitoring of employees on hourly basis.

Figure 1 shows the process flow followed in ready-made and garment industry.

Considering the textile industry of India, operational efficiency is the one which is to be improved in a large extend especially in terms sewing which is done at the later stages of manufacturing. Few of the foreign companies like Adidas has adopted Industry 4.0 for sewing through sewbotics where man made fabrics are melted and joined using vacuum suction instead of stitching. It is also to be noted that this type of product is applicable only for bonding special fabrics for scuba diving, finswimming, spearfishing, etc., whereas viable technologies are still unavailable for commercial production. Embroidery designing in Indian textile industry is happening in an Industry 4.0 environment by integrating CNC machines with automated software which has increased the productivity. Connecting machines with similar operations together by reducing the ideal time for improving the plant efficiency is another kind of automation happening with respect to Industry 4.0 in the industry.

Spinning mills are identified as one of the biggest beneficiary of I4.0 among all the verticals in the textile industry, owing to its adaptability to automation. Compared to a traditional spinning mill, mills that adopts I4.0 are estimated to spend over 60% less in terms of manpower. During last 3–5 years, there has seen significant increase in the adoption of I4.0 in spinning mills to compensate for lack of manpower. The main reason for the I4.0 adaptability of spinning mills is the low commodity/product variety and the capital intensive nature.

2.1 Impact of I4.0 on Value Creation Aspect of Textile Industry

Industry 4.0 integration can help the manufacturer to gain in depth knowledge of the customer requirement. Products can be customized to a much greater extent than possible currently, when there is seamless flow of data from the customer directly to

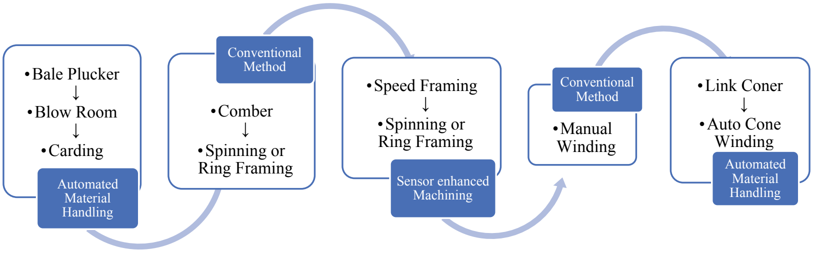


Fig. 2 Spinning mill operations in Indian textile industry

the manufacturer. But this would also require the manufacturer to possess technologies that are flexible enough to handle such highly varied demand. But such technology at the moment is at research level, and not commercially available/affordable, especially to SMEs.

Figure 2 shows the operations that are executed through I4.0 adoption as well as manually in spinning mills of Indian textile industry.

Being the initial phase of operation in the textile industry, spinning mill obtains cotton as the raw material which is in a compressed form called bale. This bale (compact and tough cotton) will be loosened through pre-opening method. Loosened cotton may consists of seed and other foreign materials which would be removed by the blow-room machine. Carding process is carried out after blowing to make the cotton more finer by removing the short fibres along with any unwanted materials if present. In large-scale spinning mills, an advanced technology called bale plucker can be used which automatically feds the bale to the pre-opening machine from which the by-products will be sent to blow room and then to carding machines with the help of I4.0 connectivity.

Next phase of operations start with combing in which still more short fibres that are present will be removed for further processes. After combing, the draw frame machine does the paralleling of fibre and makes the yarn in uniform thickness. Speed framing or simplex machine is where the first form of thickness reduction of bundled fibres is done. This is followed by spinning or ring framing where the fibre is transferred to cylindrical tubes called cops. These cops are finally fed to winding machine where the fibres are winded around tapered paper rolls which makes it easy for the textile manufacturers (customers of spinning mill product) to unwind the fibres for further production processes. In large-scale spinning mills which consist of around 100 spinning machines with a range of 1,00,000 spindles, bobbin transfer system can be implemented to transfer the fibres from speed frame machines to ring frame machine along with auto cone winding which involves the untwisting of winded cops as well as winding the tapered paper rolls with the processed fibre. Link coner is another automated operation that is done in the winding operation where the cops from the ring frame is automatically fed through the conveyer to auto coner. Both these machines are a part of I4.0 technology which would be economically unviable to be adopted in medium or small scale spinning mills which is consisting of machines with just around thousand spindles.

Combing and draw framing are the only processes that are not yet connected with I4.0 technology since it demands the fibres to be transported in large bundles of cans which would be impractical to do so with the existing technology. So higher capacity plants are also in need of a new technology to be invented which makes the automated transfer of fibres from carding machine to comber or directly to draw frame machine economically viable as well as operationally productive. In small and medium capacity production, single labour can monitor different machines of blow room, carding as well as draw framing which is the main reason for small and medium capacity spinning mills to be still dependent on manual labour operations that is economically viable and efficient when compared to high cost automatic machines.

Management of manufacturing SMEs in textile industry lacks the data visibility that is required to know about the actual cost that is incurred in the daily operation of the plant. This includes number of working hours, delay time, number of units produced, rejection rate, etc. Here is where the importance of Internet of Things (IoT) comes into picture which helps to collect these data and put it together to interpret meaningful information from the same. Another useful information that can be obtained by the intervention of IoT platform is the variation of part to part manufacturing. These data can be drawn and analysed every few seconds which enables the management to get the visibility about all the essential and vital piece of information required. This helps them to take intelligent decisions in the operation of machines.

3 Results

From the primary qualitative research conducted followed by the cross-case analysis among small- and medium-scale manufacturers in textile and garment industry as well as consultancy firms in the field of automation, the interpretations we managed to draw for the formulated questions are:

(1) **Why do adopters adopt I4.0 and what are the drivers?**

- Attaining maximum efficiency in the bottom line operations is the most desired advantage of I4.0 in textile industry. With the help of I4.0 adoption, machines with similar operations can be connected together by reducing the ideal time for improving the plant efficiency.
- Through the use of infrared rays in stitching, machine operation is automated with the detection of fabrics through IR rays which reduces the processing time. I4.0 also enables the semi-skilled labours of the industry to work with maximum efficiency by using automated pedal operation.

- From the management point of view of the industry, absenteeism as well as real-time efficiency of employees can be done on hourly basis which further expands the productivity.

(2) **Why do non-adopters don't?**

- Ready-made and garment industry is at the lowest priority of Industry 4.0 because it is predominantly man power driven, and specialization is in a very less need of automation and number of people required to convert raw material into final product is high.
- When comparing the operational cost, labour intensive processes are not preferable to be automated due to high volume of intricate work and large variety of materials involved that demands human touch.
- There is unavailability of appropriate automated machines for certain tasks like cutting and sewing which makes the technology adoption expensive in the industry.
- In case of spinning, small and medium capacity mills are still dependent on manual labour operations that is economically viable and efficient when compared to high cost automatic machines because single labour can monitor different machines of blow room, carding as well as draw framing which are the main operations carried out in spinning mill.

(3) **What are the areas of business that need to be prepared?**

- Considering the textile industry of India, operational efficiency is the one that needs to be improved in a large extend especially in sewing which is done at the later stages of manufacturing for which I4.0 will have a huge impact upon adoption.
- Integration of data into the conventional machine operations need to be done in order to gain more in depth knowledge about the customer requirements so as to customize the products to the most extend possible.
- Importance of the same should also be conveyed to people who work closely with manufacturers including employee, raw material suppliers as well as the machine manufacturers who play a major role in preparing the business towards I4.0 adoption.

(4) **What are the costs of adoption?**

- It is found that cost of bringing those technologies into the industry is detrimental at this point of time, and the main reason for this is the lower labour cost in the industry when compared to alternative technologies available.
- Material property is another barrier that demands sophisticated machines that are expensive and rare in the country for handling stretchable types of apparel clothes.
- Major overhauling of the labour force is required in terms of training as well as job replacement while adopting I4.0 technologies which eventually

distorts the current harmony of optimal cost of operation and return on investment (ROI).

(5) **What are the consequences of adoption?**

- The main issue manufacturing SMEs, specifically textile industry considered face is the lack of business development which restricts them from effectively utilizing the I4.0 technology.
- Another challenge faced by them in adopting I4.0 at an enterprise level is that they operate at very low margins, and the reason for this is the increasing demand of customers in terms of quality.

(6) **What changes do adopters make in their operations?**

- While adopting I4.0 technologies, sensors and gauges can be integrated into the CNC machines used in the plants to understand the amount of power consumed, quantity of parts produced, speed with which these machines are running (optimal speed or sub-optimal speed) and also the number of parts that met and does not met the standards.
- With the use of data-driven technologies, any non-value adding operation including supervision will be eliminated.
- Business model of an SME starting from suppliers to various value adding processes that are followed in the day-to-day operation has to be altered in such a way that enhances the manufacturer to effectively utilize the machines with the help of data and data analysing technologies.

4 Discussions and Conclusion

In case of apparel and garment industry, feasibility is the main issue faced in terms of I4.0 adoption. Raw materials are passed through various sewing machines and different labours. These operations are not synchronous in the sense that they do not run at a linear speed. For example, bottom and top stitching might take more time than elastic stitching. This kind of asynchronous operation makes specific workstation to remain idle in between the process. IoT platforms will help in the balancing of these workstations in real-time basis. One of the main component used for this efficiency enhancement is the average number of stitches required for a particular apparel or garment to meet the preferred quality standards. Another area which can be benefited through I4.0 is the plants with interconnected machines where the malfunctioning of one machine can affect the productivity of whole plant. In order to reduce the manufacturing cost, either the wastage should be reduced or the productivity should be increased, and information is required to achieve both which forces the recent manufacturers to become software or data driven. Sewbotics (as mentioned before) is one of the futuristic technologies that could be adopted in Indian garment industry which puts out a lot of information for IoT enhancement.

These type of technological advancements will even enable the manufacturers to get information about the depth, length and weight of the stitch involved in a particular garment production. Even though the current Indian textile industry is not in a position to afford these technologies, over the time once the sensor costs drop and technology matures, I4.0 adoption can be seen more extensively in the country. Once the data visibility regarding cycle time variation is made possible in any manufacturing SMEs, productivity is expected to increase by twenty percentage. There is a huge opportunity for future research in the area of ready-made and garment industry regarding the increase in productivity which is possible with I4.0 adoption especially in terms of stitching and embroidery works. This is because having a real-time information on a daily basis is expected to help any manufacturer to plan and react to the situations more abruptly and efficiently.

References

1. Iyer A (2018) Moving from Industry 2.0 to Industry 4.0: A case study from India on leapfrogging in smart manufacturing. Procedia Manufacturing 21:663–670
2. Khan, A., & Turowski, K. (2016). A survey of current challenges in manufacturing industry and preparation for industry 4.0. In *Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry"(IITI'16)* (pp. 15–26). Springer, Cham.
3. Ganzarain, J., & Errasti, N. (2016). Three stage maturity model in SME's toward industry 4.0. *Journal of Industrial Engineering and Management (JIEM)*, *9*(5), 1119–1128.

A Comprehensive Review on Reputation-Based Trust Management Framework for Cyber Physical Systems

V. Kanchana Devi

Abstract Wireless sensor networks (WSNs) act as base for all cyber physical systems (CPSs). CPS is composed of physical components like sensors and actuators, software influenced components like controllers. In which, largely used cryptographic mechanism is unsatisfactory to protect sensors which are exposed to physical world. Since sensors are commonly employed in hostile environment for long period of time, where there are possibilities for advisory to harm the sensors physically and acquire cryptographic keys which are in memory of sensors. Hence, a novel reputation trust management mechanism-based framework (NRTF) is designed to monitor the unwanted activities and develop trustworthiness of sensors by examining merits and demerits of various reputation-based trust models. Sadly, existing reputation-based systems for WSN and CPS do not explore healthiness against attacks related to reputation-based mechanism. This paper provides a framework for highly critical CPS by examining state-of-the-art with novelty.

Keywords Cyber physical system · Reputation · Sensor networks · Watchdog mechanism · Trust management

1 Introduction

Cyber physical system is an integration happened between physical components such as sensors, actuators with the software components in order to respond with dynamic changes in real-world scenario. Due to technology limitations and environmental influences, sensors are inherently lack in trustworthiness. A novel reputation-based trust framework is intended to serve the foundation for security of sensors in CPS. This paper presents many challenges that are not addressed by existing reputation-based trust management mechanisms. The sensors in CPS are assumed to be highly distributed network of lightweight, small wireless sensor nodes usually deployed in huge number to observe the environment parameters like pressure, temperature, and

V. Kanchana Devi (✉)
Vellore Institute of Technology, Chennai, Tamil Nadu, India
e-mail: kanchanadevi@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_24

humidity. Generally, sensor nodes are equipped with batteries. Thus, adding more complex security features influence the energy of sensors, sometimes it drains the battery. In TinyOS, data authentication consumes energy by 3%, also data authentication along with encryption increases the consumption of energy approximately 14%. Also, it is unfeasible to embed complex cryptographic mechanisms in such small device.

To address these kinds of challenges in sensor nodes, a novel reputation-based trust management framework is presented in this paper by analyzing various trust computation techniques. It is composed as four stages: (i) data gathering stage (ii) data modeling stage (iii) decision-making stage (iv) distribution stage. The purpose here is to build a trust among the nodes by observing the neighbors with the help of proposed framework. Each stage in the framework is composed of technologies to observe and calculate the trust among the neighbors.

The reputation and trust are usually two different words which are tied together in sociology. The integration of reputation-based trust management into security of CPS strengthens the performance. Prevention of attacks is not possible with this kind of reputation-based trust management framework but this helps in detection of advisory behavior. The WSN undergoes various attacks specified in [1] such as (i) spoofed-data attack (ii) replay attack (iii) selective-forwarding attack (iv) Sybil attack.

1.1 Attacks in WSN

Spoofed-Data Attack. This spoofed-data attack is performed after replay attack. In replay attack, the adversary eavesdrops the data transfer between the sender and receiver and captures the required information. With captured information, the adversary performs required changes and reinjects packet again later into the network. This kind of attack affects the reputation value of a sensor node. Such kind of spoofing of data is known as "Spoofed-Data Attack." It is highly risky while transmitting any sensitive information over network in CPS.

Replay Attack. Simply the attacker captures data from network without understanding content exchanged. This attack is launched without much computation spent for decryption of data. This replay attack influences the CPS network easily and shuts system down easily.

Selective-Forwarding Attack. A sensor node is said to be compromised if its cryptographic key is captured by the adversary. Such compromised nodes reject to forward the incoming packets it meant to be. Some compromised nodes forward the incoming packets at certain times. This typical attack in compromised sensor nodes are known as "Selective-Forwarding Attack." This attack affects the entire CPS to collapse within period of time.

Sybil Attack. More than one identity is created for a particular node mainly for the purpose of degrading the reputation value of a legal node. The neighboring nodes falsely increase the reputation count of attacker. This type of attack is highly dangerous and difficult to identify since the attackers enters the network with fake identification. All the above specified attacks drastically affect the overall performance of nodes in the network of CPS.

1.2 Attacks in Trust Reputation

In [1], the summary of various attacks which targets the reputation of different trust models is provided below.

On/Off Attack. Overall system performance is affected by this kind of attack. This attack is difficult to identify since the adversary alternatively changes its behavior. This is similar to that of selective-forwarding attack in WSN.

Bad-mouthing Attack. As the name specifies, a node provides false information about other neighboring nodes in the network. Some reputation systems use recommendations as second-hand information for building the trust. Such systems collapse or become disjoint network soon. Simply, the well-behaved nodes are considered as bad node by such bad-mouthing attack.

New-Comer Attack. This type of attack is quite easy to establish. Since if the reputation value of a particular node becomes low, then that node withdraw itself from the network and join again as a new-comer. This kind of attack takes much time to identify.

Ballot-Stuffing Attack. It is opposite to the bad-mouthing attack. In this attack, a bad nodes reputation is improvised by providing false positive information. This is also called as "False-Praising Attack." This type of attack is avoided when the reputation building depends only on the first-hand or direct information.

2 Literature Survey

Shaikh et al. [2] exclusively designed a lightweight group-based trust management scheme (GTMS) for clustered nodes in WSNs. The trust value is computed by frequent interaction among adjacent nodes. The author has proposed three levels to compute trust at node, cluster head, and base station. This particular scheme adopts a new trust computation technique to defy the attack of malicious nodes in network. GTMS does not demand any kind of enormous data storage and complicated calculations. Nevertheless, only perceiving the number of both successful and unsuccessful communication does not influence the computed trust value. In re-trust [3], authors defined an attack-resistant and lightweight trust management scheme (LTMS). This type of system is designed commonly for medical sensor network and reflects hierarchical structure, with the inclusion of master and slave nodes. This paper practices sliding-time window embedded with aging-factor to detect and isolate attacks like on/off attack and bad-mouthing attack. Bad-mouthing attack is excluded by isolating the outliers after gathering recommendations about the neighbors. Conversely, shortcomings of this system are that master nodes must have ample storage for storing the computed data and energy for complex calculations.

In [4], authors introduced an agent-based trust mechanism for WSNs (ATSN). Here, behaviors of sensor nodes are monitored with the help of agent-nodes and categorize as good or bad nodes based on the observed behavior. An agent node is designed mainly to count good and malicious behaviors, and results are stored in the form of three-tuple. Use of such agent node in ATSN reduces the computational resources and energy consumed. This ATSN ignores all the second-hand information (recommendations) and consider only the direct information. A reputation-based framework for sensor networks (RFSN) is proposed by Ganeriwal and Srivastava [5] in which, the observing nodes use the reputation of the other nodes to compute trust. This system incorporates watchdog mechanism to observe interaction of adjacent nodes, and beta distribution is used for representing the observed data. The limitation of this particular system is that it does not address the compromised node attacks. Li et al. [6] designed a lightweight trust system with decision-making for the clustered WSNs (LDTS). The cancelation of feedback among cluster members helps in improving the efficiency. The self-adaptive weighted method is adopted here for trust computation in the cluster head communication level. This trust system exceeds the general disadvantage of usual weighting trust models, in which weights are allocated individually.

Song et al. [7] present an effective trust evaluation model (ETEM) using multi-factor. Each nodes' reliability is computed by integrating first-hand information with second-hand trust. Furthermore, the classification methods and weight assignment are reliant on communication among nodes, the Hoeffding's-Inequality used in this trust system. The projected result shows that this particular model resists against all kind of attacks in WSN. Zhan et al. [8] proposed an innovative approach with resilient trust model (RST) and focused mainly on data integrity in communication and special trust model for the hierarchical WSNs. The trust level of each sensor node in the network is computed with the history and latest risks. Later, it involves Gaussian trust model in order to provide higher data integrity. This trust model resists against faulty data communication and malicious data operation. The main disadvantage here is that the proposed model failed to consider the energy consumption for the trust calculation. Feng et al. [9] present Bayesian-based trust management mechanism (BTM) which involves direct observation trust and indirect observation trust. Formal is computed with modified Bayesian equation by including punishment factor and a sliding window method for updating trust value along with a new term forgetting factor. Additionally, the indirect trust calculation is triggered by neighbors. This Bayesian method resists against all sort of attacks invoked by adversary. In [10] RAHIM, authors presented a reactive defense technique for the clustered WSN. This work targets the compromised nodes in the network. The accuracy in data aggregation strengthens the network by this model.

Marti [11] introduced routing misbehavior in mobile ad hoc network. The concept of *watchdog* and *pathrater* is used in this paper to minimize the routing misbehavior. Throughput, routing overhead, and accuracy are the parameters taken into consideration. Fenye et al. [12] used multidimensional attributes of trust for intrusion detection and geographic location with probability for heterogeneous sensor nodes. This model aims at message delivery ratio and delayed delivery. In [13], authors proposed a lightweight trust model for Fog-CPS (LTMF) based on multiple factor and multiple dimensional parameters which are formulated with statistical regression. Wang [14] proposes trust qualification for networked CPS (TQN). The paper described about highly integrated hardware and software devices which are tightly coupled with latest computational algorithm for the purpose of gathering the sensed data, processing, communication, and controlling. A specially designed CPS with graph model is used to demonstrate ability and genuineness of the trust system. In paper [15], authors proposed a blockchain-based CPS. This paper focuses on the properties data integrity, authenticity, and data identity and implemented using a testbed. Kanchana et al. [16] have presented a trust management framework for CPS by identifying the anonymous packet droppers in a heterogeneous environment.

Alzaid et al. [1] proposed a comprehensive review on reputation-based trust models and proposed an innovative framework for sensor network. This model has been designed in such a way to resist against all kind of reputation and sensor attacks. In paper [17], DTEM is proposed which includes sliding window with lightweight trust to improvise security in WSN. The results depict that this model is dynamic to resist against attacks which targets the direct and indirect communications. Momani et al. [18] have surveyed several trust models of e-commerce, peer-to-peer network, WSN, and ad hoc network. Author presented comparisons with different methodologies such as rating, weight, probability, game theory, neural network, fuzzy logic, and swarm intelligence. Jayasinghe et al. [19] used machine learning in security of Internet of Things (IoT) and proposed a quantifiable trust assessment model. Machine learning algorithm is used for classifying trust factors to make appropriate decision based on location, activity, and data interest. Sivagurunathan et al. [20] have proposed trust-related security mechanism. Different methodologies are presented by addressing real-time issues like optimization, planning, healing, and configuration. In paper [21], authors have introduced a cooperative technique for enhancing the lifetime of the network. The performance of proposed model has been compared and analyzed with energy and routing parameters.

3 Comparative Analysis of Existing Trust Models

Different trust models are classified in existing models based on information collection and storage are shown in Fig. 1.

There are certain merits and demerits of these types of trust model which are listed in Tables 1 and 2.

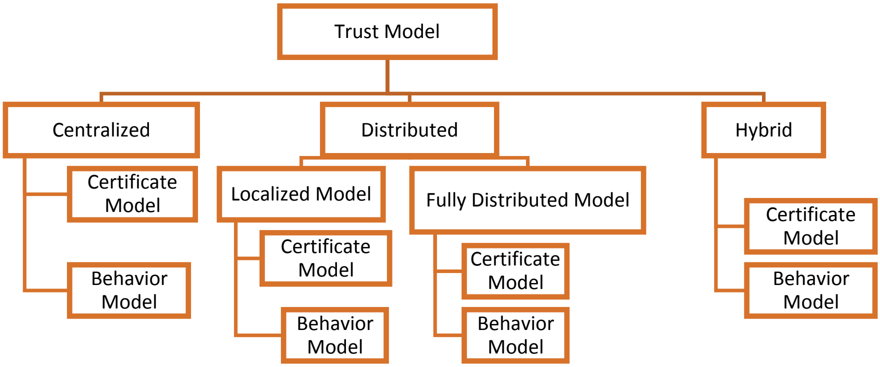


Fig. 1 Types of trust model—information collection and storage

Table 1 Merits and demerits of trust model

Structure	Merits	Demerits
Centralized structure	Less memory usage and less computation	More communication overhead, less reliable, and lack of scalability
Distributed structure	Reliable and scalable	Large computation required
Hybrid Structure	Less communication and less memory usage compared to centralized structure	Large computation and memory required compared to centralized

4 Suggestion Proposed: Novel Reputation-Based Trust Framework

The novel reputation system framework is composed of four different stages as shown in Fig. 2. This framework has been designed as suggestion for improving the lifetime of the network by identifying the malicious nodes earlier.

4.1 Stage 1: Data Gathering Stage

This is the first stage in building the security framework in such broad cyber physical system. This stage comprises of four levels in data gathering. They are source of data, the type of data to be collected, the scope, and the watchdog technique to monitor the neighbors to collect data for trust calculation [22].

Source of Data. Trust of each node in the network is constructed by purely relying on the data collected from the node. An observing node must decide on the source as manual or automatic data collection. Here, in CPS, the source is embedded in the physical entity as sensor. These sensors collect data in a self-adopting mode.

Table 2 Analysis of trust models

Trust models	Data exchange structure/	Description	Computation technique	WSN attacks	Reputation attacks	Detection	Decision metric
GTMS [2]	Hierarchical structure	Group	First-hand and second-hand data gathering	Yes	Yes	Malicious nodes	Discrete
LTMS [3]	Hierarchical structure	Weight	Aging-factor, recommendation	No	Yes	On/off, bad-mouthing	Discrete
ATSN [4]	Distributed structure	Agent	Certificate technique	Yes	Yes	Compromised attacks	Probability
RFSN [5]	Distributed structure	Reputation	Beta probability	No	Yes	Malicious attack/packet forwarding	Binary
LDTS [6]	Hierarchical structure	Weight	Self-adaptive weight method	Yes	No	Malicious attack/packet forwarding	Probability
ETEM [7]	Distributed structure	Multifactor, Hoeffding's-Inequality	First-hand and Second-hand data gathering	Yes	Yes	Selfish node	Discrete
RTS [8]	Hierarchical structure	Resilient	History and Gaussian trust model	Yes	No	Misbehavior attack	Probability
BTM [9]	Distributed structure	Bayesian	Bayesian equation with punishment, sliding window, forgetting factor	Yes	Yes	Malicious node	Discrete

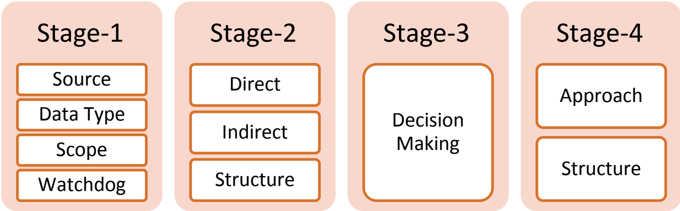


Fig. 2 Structure of reputation system framework in stages

In some software requirement, the data is collected manually. For example, the feedback/survey/options are collected from the customers to take appropriate actions.

Type of Data. Collected data from end device is supposed to be categorized depending on the application of CPS since the data type requirement usually vary from one application to another. The different data types available in such end devices are binary or unary, discrete, or continuous. Typically, if the datatype is unary either only negative or only negative. Similarly, if the datatype is discrete, then the data collected is negative, positive, or neutral, and if the datatype falls under the category of continuous, then the data collected is real values.

Scope of Data. One of limitation of trust management is that it has to be designed according to the requirement of the application. It should be unique from system to system. For example, scope of aggregation differs from scope of monitoring. This particular requirement must be considered while designing the system. The scope monitoring requires watchdog technique to observe the neighbors.

Watchdog Technique. This watchdog technique is embedded in most of the existing system for the purpose of monitoring the neighbor node behavior such as packet forwarding, buffering, and energy consumption. This technique has to be tailored according to the requirement of the application it has been designed for. The verification of packet forwarding by a suspected node is combined with the timeout and buffering concepts. A counter is usually maintained to capture the activities of compromised nodes.

4.2 Stage 2: Data Modeling Stage

The second stage in the proposed novel reputation-based trust management framework is data modeling stage. In this stage, the data collected from the end devices using the watchdog technique is handed over as input to this stage. The collected data undergoes into two categories such as (i) Approach (ii) Structure.

Data Modeling Approach. The approach used for data collection is of two types, namely direct approach and indirect approach. In direct approach, the data collected by an end node without using an agent or from neighboring. The observing node simply observes the packet forwarding task of the neighbor node by watchdog. The data collected by the observed node using this direct approach is known as "*first-hand* *information.*" For certain applications, same directly observed data exchanged as recommendation with the adjacent nodes on demand. This recommended data about a particular node is the "*second-hand information.*" There are certain disadvantages in using such recommendations such as increase in network traffic due to number of data exchanges, sometimes leads to bottleneck condition in limited storage devices.

Data Modeling Structure. The suitable structure needs to be decided while designing reputation-based trust management system. There are three structures available for computing the reputation for collected values. Firstly, distributed structure, every node in the network independently computes the trust reputation value from the previous stage which are gathered from the adjacent sensor nodes in CPS. This distributed structure consumes less energy compared to other structures. Secondly, centralized structure, every collected data is propagated to centralized server node. Sometimes, centralized nodes are cluster heads. These cluster heads compute reputation value for the trust management system. Along with general communications, updated reputation values are embedded and transmitted to the subscribed sensor nodes in CPS. The limitation of this kind of centralized structure is that there is possibility of single point failure at cluster head. Thirdly, hybrid structure, the combination of distributed and centralized structure.

4.3 Stage 3: Decision-Making Stage

The third stage in reputation system framework is the decision-making stage. This stage helps in identifying whether the observed node is trusted node or not. With the decision, observer decides whether further communications to be proceeded or not. Commonly used decision metrics differ from application to application. Some metrics used are discrete values, continuous values, binary, threshold values, ranking, rating, scale, etc. Most commonly used metric depends on threshold.

4.4 Stage 4: Dissemination Stage

Finalized decision is communicated to the subscribed or adjacent nodes using different structure and approaches. The dissemination process is carried over in a scheduled manner, periodically, or instantly. These approaches help in reduction of traffic and congestion in the network. Carefully designed dissemination systems resist against targeted both WSN attacks and reputation attacks in CPS. Effective design in structuring (distributed or centralized) helps in decreasing the energy consumption. The distributed structure and proactive approach suit appropriate for highly distributed CPS.

4.5 Technical Workflow

Technically, the data from hostile environment is collected automatically by the sensors, and the watchdog mechanism is triggered by enabling the promiscuous mode. This helps in monitoring the neighboring node. After forwarding the packet, the observing node ensures that whether the forwarded packet has been transmitted by the next node. If not, the counter maintained for failure is incremented else the success counter is incremented. Beta distribution Eq. 1 used in modeling phase intakes two discrete parameters as input and produces a single continuous output. The two counter values obtained from watchdog mechanism are given as input to the beta distribution to check the trust level of a particular node. This output is called the trust value which is checked with the neutral value 0.5 to verify the neighboring node can be trusted or not in decision-making stage. If the node is trusted, then further packets are transmitted via the same path. Otherwise, the node identity is intimated via TCP connect to the sink node in dissemination stage.

$$f(x;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{1}$$

Here, x varies between lower bound (0) and upper bound (1). The two parameters α and β are positive numbers which are greater than 0.

5 Result Discussion

The experiment is performed using MATLAB for analyzing performance of network by increasing the sensor nodes from 10 to 50. Here, for every ten sensors, 2% of nodes are made malicious nodes and observed the network lifetime with various trust models. The simulation scenario is set as 100 m $\times$ 100 m, with 10–50 sensor nodes. Certain parameters like number of sensor nodes, number of selfish nodes, network lifetime, and throughput are varied with the scenarios and for the purposes of experimentation. In this setup, network becomes disjoint, if the number of malicious nodes increases. The malicious nodes simply drop the incoming packets which is assumed to be forwarded. This makes the network disjoint.

The initial energy of each sensor has been kept as 0.5 J, initial trust value 0.5, packet length per bit as 2000, and trust estimation period as 10 s. By comparing various trust models in Fig. 3, it has been proved that the proposed trust model (NRTF) improves the lifetime of the network by identifying the malicious nodes. Figure 4 shows the comparison of throughput before and after identifying the compromised node in the network.

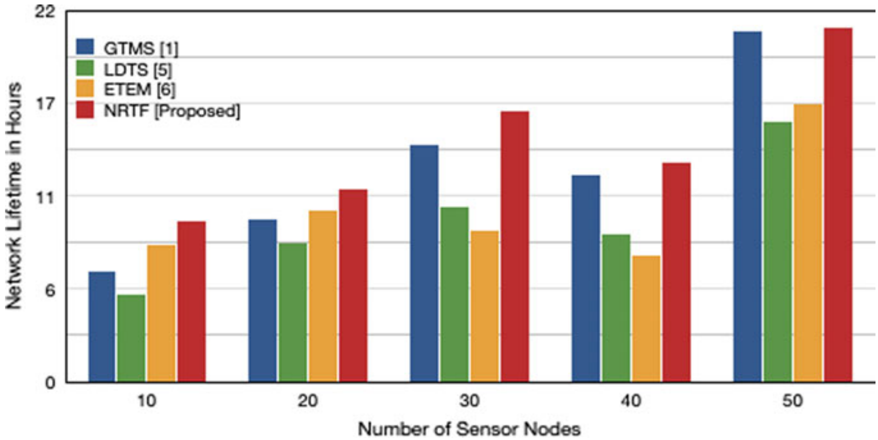


Fig. 3 Comparison—trust model (number of sensors and network lifetime)

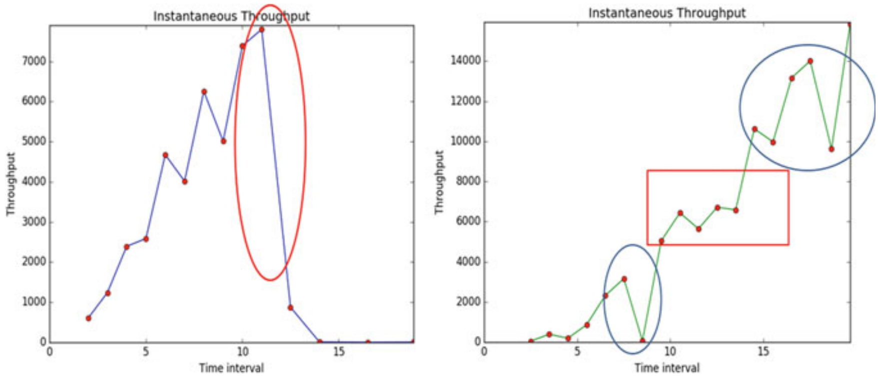


Fig. 4 Comparison of throughput before and after identification of compromised node

6 Conclusion

This paper presents a complete review on various reputation-based trust management models in cyber physical system and wireless sensor networks. Firstly, it explains the different existing trust models, compared and categorized different structure of trust models along with the merits and demerits. Secondly, it explains the various attacks on WSN and reputation system which affects the performance of trust management system. Thirdly, by examining various trust models, a carefully analyzed novel framework has been presented in this paper for resisting against various attacks.

References

1. Alzaid, Hani, Manal Alfaraj, Sebastian Ries, Audun Jøsang, Muneera Albabtain, and Alhanof Abuhaimed. "Reputation-based trust systems for wireless sensor networks: A comprehensive review." *In IFIP International Conference on Trust Management*, pp. 66–82. Springer, Berlin, Heidelberg, (2013).
2. Shaikh RA, Jameel H, d'Auriol BJ, Lee H, Lee S, Song Y (2009) Group-based trust management scheme for clustered wireless sensor networks. IEEE Trans Parallel Distrib Syst 20(11):1698–1712
3. He D, Chen C, Chan S, Bu J, Vasilakos AV (2012) ReTrust: attack-resistant and lightweight trust management for medical sensor networks. IEEE Trans Inf Technol Biomed 16(4):623–632
4. H. Chen, H. Wu, X. Zhou, and C. Gao, "Agent-based trust model in wireless sensor networks," in *Proceedings of the 8thACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD '07)*, pp. 119–124, IEEE, Qingdao, China, August (2007).
5. S. Ganeriwal and M. B. Srivastava, "Reputation-based frame- work for high integrity sensor networks," in *Proceedings of the* 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks *(SASN '04)*, pp. 66–77, October (2004).
6. Li X, Zhou F, Du J (2013) LDTS: a lightweight and dependable trust system for clustered wireless sensor networks. IEEE Trans Inf Forensics Secur 8(6):924–935
7. J. Song, X. Li, J. Hu, G. Xu, and Z. Feng, "Dynamic trust evaluation of wireless sensor networks based on multi-factor," in Proceedings *of the 14th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2015*, pp. 33–40, fin, August (2015).
8. Zhan G, Shi W, Deng J (2011) SensorTrust: a resilient trust model for wireless sensing systems. Pervasive Mob Comput 7(4):509–522
9. R. Feng, X. Han, Q. Liu, and N. Yu, "A credible bayesian- based trust management scheme for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2015, Article ID 678926, (2015).
10. Labraoui N, Gueroui M, Aliouat M, Petit J (2011) RAHIM: Robust Adaptive Approach Based on Hierarchical Monitoring Providing Trust Aggregation for Wireless Sensor Networks. J Univ Comput Sci 17(11):1550–1571
11. S. Marti, T. Giuli, K. Lai and M. Baker, "Mitigating Routing Misbehavior in Mobile Ad-Hoc Networks" in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, ACM Press, Boston, (6–11), pp. 255–265, (2000).
12. Fenye Bao, Ing-Ray Chen, Moon Jeong Chang, and Jin-Hee Cho, "Hierarchical Trust Management for Wireless Sensor Networks and its Applications to Trust-Based Routing and Intrusion Detection" in *IEEE Transactions on network and service management*, vol. 9, no. 2, pp.169-183, (2012).
13. A. K. Junejo, N. Komninos, M. Sathiyanarayanan and B. S. Chowdhry, "Trustee: A Trust Management System for Fog-enabled Cyber Physical Systems," in *IEEE Transactions on Emerging Topics in Computing* (2019).
14. Wang Y (2018) Trust quantification for networked cyber-physical systems. IEEE Internet Things J
15. Beckmann, A., Milne, A., Razafindrakoto, J.J., Kumar, P., Breach, M. and Preining, N. "Blockchain-Based Cyber Physical Trust Systems". *IoT Security: Advances in Authentication*, pp.265–277 (2020).
16. V, K. D. and R, G. (2017) "A Generic Trust Management Framework for Heterogeneous Sensors in Cyber Physical Systems", *Asian Journal of Pharmaceutical and Clinical Research*, 10(13), pp. 176–179, (2017).
17. Ye, Zhengwang, Tao Wen, Zhenyu Liu, Xiaoying Song, and Chongguo Fu. "An efficient dynamic trust evaluation model for wireless sensor networks." *Journal of Sensors* 2017 (2017).
18. Momani, Mohammad, and Subhash Challa (2010) Survey of trust models in different network domains. arXiv preprint arXiv:1010.0168

19. Jayasinghe, Upul, Gyu Myoung Lee, Tai-Won Um, and Qi Shi. "Machine learning based trust computational model for IoT services." *IEEE Transactions on Sustainable Computing* 4, no. 1 (2018): 39–52.
20. Sivagurunathan, S., and K. Prathapchandran. "Trust-Based Security Mechanisms for Self-Organized Networks (SONs)." *In Securing the Internet of Things: Concepts, Methodologies, Tools, and Applications*, pp. 1782–1805. IGI Global, 2020.
21. V, K., R, G. "Semi Markov process inspired selfish aware co-operative scheme for wireless sensor networks (SMPISCS)". *Cybersecurity* 2, 4 (2019). https://doi.org/10.1186/s42400-019-0021-3.
22. V, Kanchana Devi; R, Ganesan. Trust-based Selfish Node Detection Mechanism using Beta Distribution in Wireless Sensor Network. *Journal of ICT Research and Applications*, [S. l.], v. 13, n. 1, p. 79–92, Apr. 2019.

A Robust Statistical CNNCTC-Based AI Model for Tracking and Monitoring COVID-19

Balika J. Chelliah, S. Arunkumar, R. Prabavathi, and B. Aarthi

Abstract The world today battles the new coronavirus (COVID-19) spread that can lead to a major endemic or pandemic outbreak with devastating outcomes which has overturned all our usual calculus seemingly. Because of the unavailability of a vaccine and the serious extent of infectiousness, social distancing is perhaps the only way left to fight the infection. The tracking system is still a challenge for monitoring the COVID-19 affected person and the community spread. This leads to an epidemic situation, where the number of secondary infections keeps going higher and higher which intern leads to the difficulty of tracking and treating each infected person. An AI-based mathematical modelling is proposed for tracking and monitoring the pandemic outbreak. A GPS tracking system, geolocation using Fit bit and Wi-Fi tracking modules are developed to locate the secondary infections and the carriers. Calculation of real-time coordinates and the conditional probabilities with respect to time is done followed by the classification based on risk using CNN connectionist temporal classification (CTC) model, a type of neural network which is used to train the recurrent process. Monitoring the affected person using sensors like thermal sensors and a tracking chip inbuilt in a wristband for real-time updates. A comparative study is done on various classification models. Therefore, the system provides a robust model with 96.5% accuracy to combat this pandemic situation.

Keywords COVID-19 · GPS · Fit bit · CNN · Connectionist temporal classification (CTC) · Thermal sensors

---

B. J. Chelliah · S. Arunkumar · B. Aarthi
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, India
e-mail: ballikaj@srmist.edu.in

B. Aarthi
e-mail: aarthib@srmist.edu.in

R. Prabavathi (✉)
Assistant Professor, Sri Sairam Institute of Technology, West Tambaram, Chennai, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_25

1 Introduction

The pandemic of coronavirus malady 2019 (COVID-19) is spreading everywhere throughout the world, facing the day to day challenge and threat posed by the virus which originated in Wuhan. The Centres for Disease Control and Prevention (CDC) have reported that SARS-CoV-2 is a respiratory infection and is transmitted primarily through "respiratory beads" between people when indicative individuals sniff or hack. In case of a pandemic outbreak, emergency management units must arrange a viable alleviation technique to stop the ailment spread utilizing constrained assets. To create a fruitful response, a detailed mathematical model of how the disease can spread and needs to be developed. Previously developed models are predominantly based on homogeneous blending models which treat each member of the population as having an indistinguishable chance of illness. Such an inference is intuitively unrealistic. Many demographic groups (e.g. health workers, infants and the elderly) have higher infection risk rates. In addition, lifestyle habits such as the use of public transport may have an effect on the risk of infection. Using communication networks to reflect the interaction level between the people and census data to estimate geographical location and habits of travellers. The outcomes are intermittently shown on zone maps utilizing GIS for visualization and planning purposes. The proposed system provides an AI-based tracking and monitoring system in order to avoid the social community spread. A statistical model is developed based on the classification on the levels of risk into three zones. CNNCTC-based classification model is proposed to classify and train the recurrent process of identification based on variety of parameters. Monitoring the health of the individual is done using thermal sensors which tracks the body temperature and tracking chip located in the wristband to indicate the proximity.

The system developed provides a robust and effective framework for tracking the contacts made by the COVID influenced individual. The classification based on timestamp is done using a mathematical model. Finally, a monitoring system for the isolated individual using sensors is done.

2 Related Work

Simulation studies play a key role in promoting and understanding of how infectious diseases spread and help the prediction of pandemic outbreak scenario. Disease-spread simulation models are also used to explain the consequences of changes in individual actions or government policies, or to research parameters and prevention plan features for disease outbreaks. The pandemic outbreak demonstrated that increased density and mobility of the population would play important roles in the spread of emerging infectious diseases and could potentially contribute to future pandemic. Nonetheless, monitoring the detailed essential behavioural patterns of each person for a whole population can be challenging and needs considerable computational power [1].

The crowdedness level of the existing and proposed layout was assessed using an agent-based pedestrian flow simulation model. Often included in the model were the pedestrian movement habits and commuter activities such as queuing up at a facility or standing at the waiting room. The simulation results could be used to collect data on crowdedness level indicators such as density distributions which are useful in assessing the flow of passengers [2].

We could quantify the severity and likelihood of infection in some urban settings by understanding the spatiotemporal characteristics of the contacts in mass transportation activities between infectious and susceptible persons. There were several models based on microscopic agents, which attempt to simulate the movement of passenger flows in daily public travel activities [3, 4].

A cognitive science approach was suggested, focused on heuristics of behaviour. It had been suggested for pedestrians to use two types of cognitive approaches to change their walking directions and speeds, driven by visual awareness, namely the range of obstructions in the candidate lines of sight. This model also predicted the development of phenomena of self-organization, such as the spontaneous creation of uni-directional waves of stop-and-go. However, the combination of pedestrian heuristics with body strikes results in extreme density crowd turbulence—a phenomenon found in recent crowd disasters [5].

Form of stochastic influenza simulation to examine the effectiveness of quarantine, and pre-vaccination in containing an infectious viral at source [6].

The two widely used models for understanding the spread of pandemic infectious diseases are deterministic models and stochastic models [7, 8].

A distributed model of a non-homogeneous agent-based simulation of pandemic disease to integrate a public healthcare system into a simulation of pandemic influenza. Results indicate that the provision of health care substantially improves the spread of influenza due to increased interactions within the population and have found that the establishment of flu centres reduced deaths due to flu and decreased hospital admissions [9].

For pandemic influenza, stochastic model was used to analyse practical approaches that can be used in reaction to emerging outbreaks. The model is designed to record disease attack rates and basic estimates of re-productive numbers. This model projected an average incidence of 34.1% disease attack in the absence of intervention [10].

The contact tracing app has been established by the government named, Aarogya Setu uses Bluetooth and GPS systems from a smartphone to alert users when they arrive within six feet of patients with COVID-19. The alert notification is generated by scanning through location-specific, government-owned databases. They are followed by Ministry of Health guidance on how to isolate oneself and the course of action required when people experience coronavirus symptoms [11].

The Quarantine Monitor, developed by the Tamil Nadu e-Governance Agency, tracks individuals who are quarantined according to the official state database. This assists the Department of Health and the Tamil Nadu Police in the last two months to efficiently monitor and handle information about COVID-19 cases and people with international travel history. The software allows for live position monitoring once activated and produces warnings and information that are sent to state authorities [11].

Facilitating prompt and sufficient emergency medical services in the event of a disaster or an emergency is a big obstacle in the safety scenario [12].

The development of mobile application in wireless network improves the medical system to operate throughout the world. E-healthcare system of remote patient monitoring [13] where the system is to mitigate the problem occurring in remote area.

3 Methodology

The proposed system uses GPS tracking system to locate the secondary infections and the carrier followed by the classification based on risk into three zones. The stability of a dataset is maintained using coefficient of variation (CV), a statistical tool. Monitoring is done using sensors in wrist band for the complete health status.

The system provides the complete solution through these two modules.

(i) Tracking the nearest node who travelled along with the carrier node

   1. Using the GPS location of the user, geolocation using Fit bit and Wi-Fi tracking modules the location of the carrier node and the nearest node along with the carrier node can be identified through network access.
   2. Once the nearest node is identified then this group of nodes are classified in three categories based on the timestamp using the connectionist temporal classification (CTC).

      - First it classifies timestamp of the nearest node and always stays along with carrier node.
      - Second it classifies timestamp of the nearest node and stays greater than 3 h and lesser than 10 h.

      Finally, timestamp of the nearest node stays in lesser time.

(ii) Wrist band monitoring system for the quarantined person

   1. A thermal sensor will be deployed in the wrist band which triggers when the temperature goes high.
   2. When the node tries to move over the proximity area, tracking chip in the band gives the alert message (Fig. 1).

4 Mathematical Modelling

As we are trying to deal with the position and time spent at that position with respect to the position of the affected over the time, in order to analyse the level of the spread.

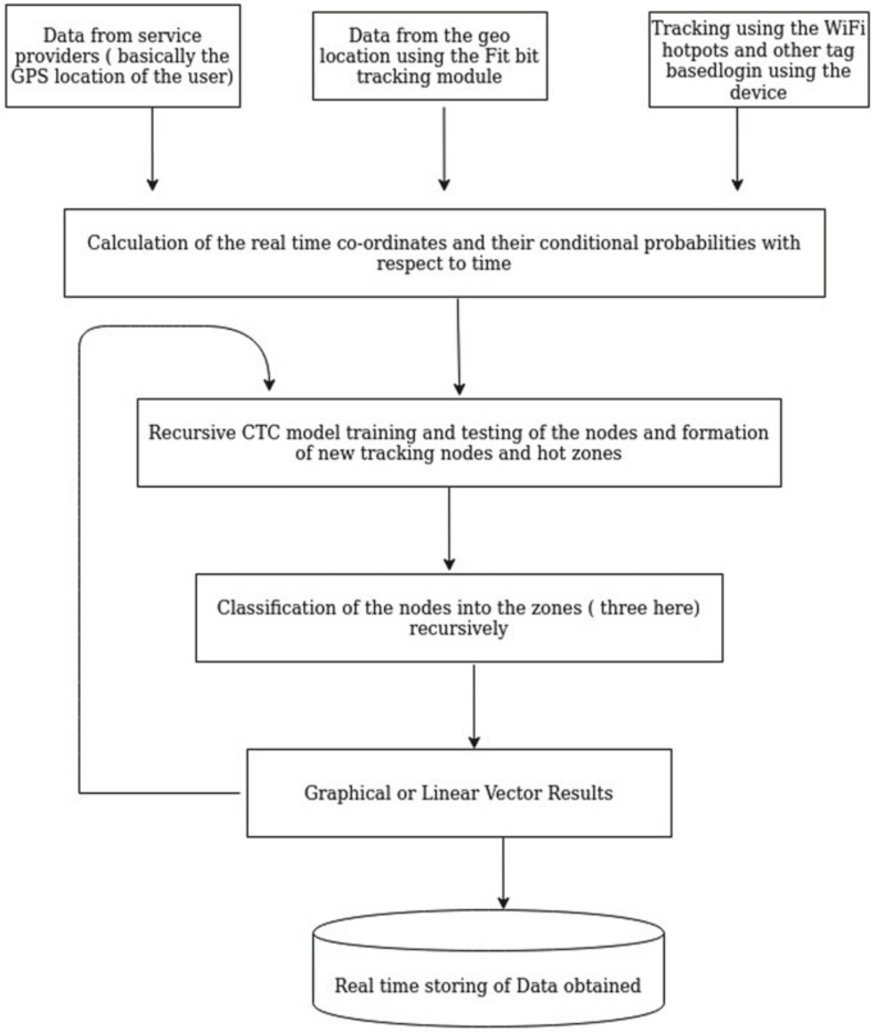


Fig. 1 CNNCTC-based AI model for tracking and monitoring COVID-19

This will basically lead to a series of recursive checks of position and time with respect to time continuously and at the same time avoid the uncertainty of position and time interval identification at the same time. Here, we cannot use the simple network of classification algorithm to analyse the position-based time intervals; hence, we use the theory of connectionist temporal classification (CTC) which is a type of neural network used to train the recurrent process of identification of subject based on variety of parameters at the same time which are derived from their position and time gap.

Let us take the two parameters along any axis of this directed graph based on time to be:

1. Position from the infected/active person, so let us map the input positional sequences as $X = [x1, x2, \ldots, xN]$ and $Y = [y1, y2, \ldots, yN]$ and we have to find an accurate mapping of this location as a result of their corresponding conditional probabilities:

$$Y* = Y \text{ argmax } p(Y|X)$$
$$X* = X \text{ argmax } p(X|Y)$$

2. Time spent very close to the active person would be a series of data that we will get again the corresponding conditional probability with respect to time from each time frames, hence, will lead to a recursive calculation of the dependence of the factor:

$$p(\pi|x) = \pi(t = 0 \rightarrow t = y)y(\pi t \rightarrow y),\ \forall\pi \in L0$$
$$p(\pi|y) = \pi(t = 0 \rightarrow t = x)y(\pi t \rightarrow x),\ \forall\pi \in L0.$$

Now, the new directed graph will have the time delays or the feedback loops to run the recursive classification algorithm to analyse these two broad parameters with new values in every time frame and the interact with other nodes (other people coming in contact) in the target or the risk zone.

To represent a single frame CTC network, we use B to define the conditional probability of a given labelling $l \in L \leq T$ as the sum of the probabilities of all the paths corresponding to it:

$$p(l|x) = \sum\{\Pi\varepsilon B - 1\}p(\pi|x)$$

Now deriving it for a single set of data with both the parameter at the same time:

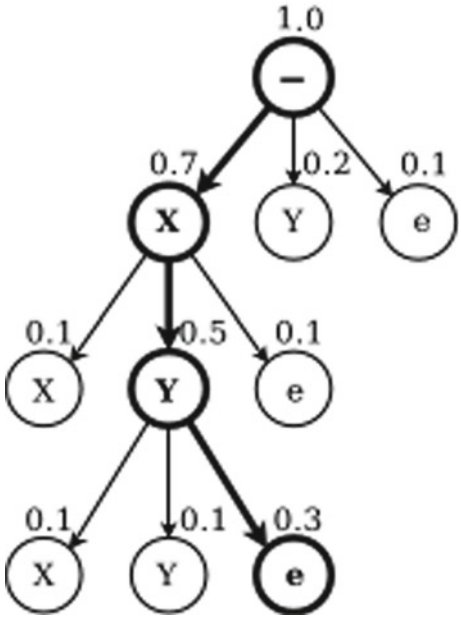
$$p(Y|X) = \sum\{A\varepsilon\ AxAy\} \prod\{t - 1 \rightarrow T\}pt(At|X)$$

Given the above formulation, the output of the classifier should be the most probable labelling for the given input pattern (Fig. 2):

$$h(x) = \text{arg max } p(l|x).$$
$$h(x) \approx B(\pi*), \text{ where } \pi* = \text{arg max } t\ p(\pi|x)$$

Now, we fix the levels of riskiness based on their position (average vector length from the main subject) using their coordinates' conditional probability with respect to the time intervals such as less than 3 h, 3–7 h and more than 7 h to determine the stage of the spread. Then, we train the model to continue performing this recursively every time frame to closely monitor without losing any data. The motto of the function is

Fig. 2 Directed graph based on risk zone



derived from the principle of maximum likelihood. That is, minimizing it maximizes the log likelihoods of the target labels (Figs. 3, 4 and 5).

After the maximum likelihood training is done with the data (there is no much loss as this is CTC compared to LSTM), we can get the result in the form of graphical depiction where we will get a classified vector of various points in different levels grouped together (here, we will have three levels as we have discussed above) and mostly highly affected levels' nodes will again undergo the same CTC training in a recursive manner which leads to grouping again and again, or we can also have the same graphical results in the form of a vector consisting of the conditional probability

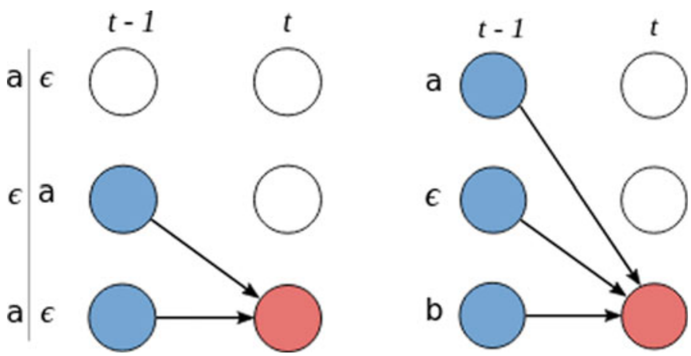


Fig. 3 Graph based on likelihood of target labels

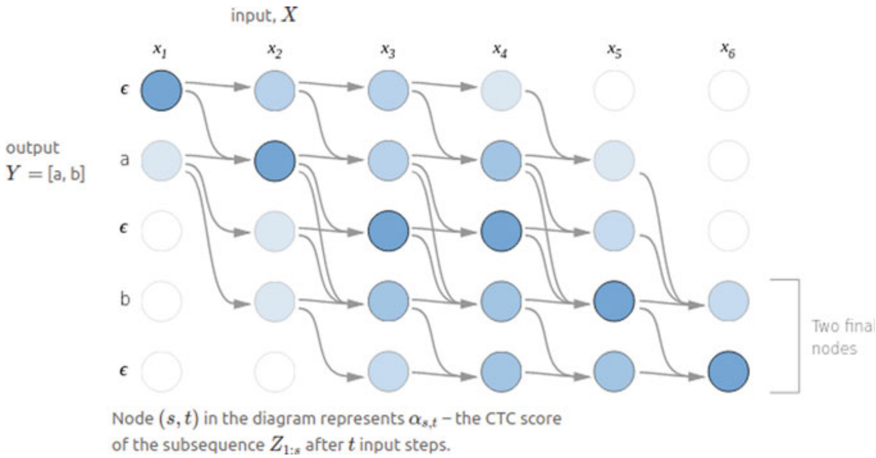


Fig. 4 Graph based on CTC score

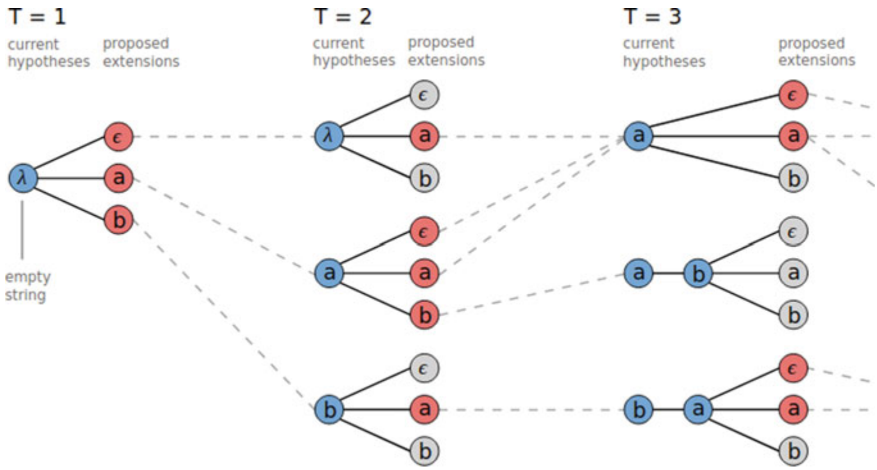


Fig. 5 CTC beam search algorithm with an output alphabet {Ɛ, *a*, *b*} and a beam size of three

value (aggregated to a floating value) whose rise will mark the start of tracking using CTC for that particular node. And the advantage here is, since there is almost negligible loss in CTC, we can go to a higher-level tracking and analysis of the spread in real time.

5 Implementation and Results

In this proposed framework, we analysed the performance of classification models based on various CNN models for monitoring and detecting COVID-19 carriers. The experimental studies were carried out on the models, and implementation was carried out using the Python IDE CNNCTC model. Results showed that our proposed model is reliable relative to the other CNN models tested. Every classifier's output measurement is calculated in terms of accuracy, sensitivity, specificity, false positive rate (FPR) and F1 score. In addition, this experimentation used carrier position and time spent close to that carrier node as the CTC classifier. The results in Table 1 shows the accuracy of different classification models with mean minimum and maximum achievable models and Table 2 shows the performance measure in terms of sensitivity, false positive ratio (FPR) and F1 score. The testing and training are adopted with randomized selection with the ratio of 80:20.

The three metrics in addition to classification accuracy are commonly used in CNN models for performance evaluation. The CTC model used in the proposed system gives good accuracy compared to other CNN models.

The research proposed is using qualified CNN modelling to get the highest accurate COVID-19 detection efficiency. We assessed the performance results of deep-function characteristics based on different classification models and compared with the proposed mathematical model CNNCTC and its result shows that the model performs well compared to other models. According to the results, CNN CTC model achieved the highest classification accuracy of 96.5%.

Table 1 Accuracy (%) of different CTC-based classification models, which uses deep characteristics of different CNN modelling algorithms

Classification models	Mean	Minimum	Maximum
AlexNet	94.86	90.66	98.66
GoogleNet	90.73	84.00	98.66
InceptionV3	90.26	85.33	96.00
ShuffleNet	65.26	57.00	70.66
ResNet50	95.33	92.00	98.66
CTC	96.5	93.2	98.5

Table 2 Performance measures (%)

Classification models	FPR	Sensitivity	F1 score
AlexNet	3.56	94.86	94.85
GoogleNet	5.13	91.73	91.74
InceptionV3	4.86	90.26	90.28
ShuffleNet	15.36	65.26	62.79
ResNet50	3.23	94.33	93.54
CTC	2.34	96.5	94.32

6 Conclusion and Future Work

A robust and effective implementation to combat the pandemic situation is done using AI-based mathematical modelling. The challenges faced in tracking and monitoring COVID-19. A GPS tracking system, geolocation using Fit bit and Wi-Fi tracking modules are developed to locate the secondary infections. The calculation of real-time coordinates their conditional probability with respect to time is done followed by recursive training using CTC model and classification of nodes based on risk. The proposed system CNNCTC-based mathematical AI achieves an accuracy of 96.5%. In addition to this, a monitoring tracking chip and thermal sensor are attached in the wrist band for complete monitoring of the COVID-19 quarantined person. Therefore, the system provides a significant solution to the pandemic outbreaks of coronavirus. The system can be future enhanced for tracking asymptomatic person but results are positive.

References

1. Tsui K-L, Wong ZS-Y, Goldsman D, Edesess M (2013) Tracking infectious disease spread for global pandemic containment. IEEE Intell Syst 28(6)
2. Liu et al SB (2013) Crowdedness in metro stations: passenger flow analysis and simulation. In: Transportation Research Board 92nd annual meeting compendium of papers, 2013. http://trid.trb.org/view.aspx?id=1241991
3. Lo SM (2011) Report on the passenger flow in Kowloon Tong Station, report submitted to Mass Transit Railway Corporation (MTRC), unpublished.
4. Asano M, Iryo T, Kuwahara M (2010) Microscopic pedestrian simulation model combined with a tactical model for route choice behaviour. Transp Res Part C: Emerg Technol 18(6):842–855
5. Moussaïd M, Helbing D, Theraulaz G (2011) How simple rules determine pedestrian behavior and crowd disasters. Proc. Nat'l Acad Sci 108(17):6884–6888
6. Longini IM et al (2005) Containing pandemic influenza at the source. Science 309(5737):1083–1087
7. Sander B et al (2009) Economic evaluation of influenza pandemic mitigation strategies in the United States using a stochastic microsimulation transmission model. Value Health 12(2):226–233
8. van den Dool C et al (2007) Modeling the effects of influenza vaccination of health care workers in hospital departments. Vaccine 27(44):6261–6267
9. Lizon NE, Aleman DM, Schwartz B (2010) Incorporating healthcare systems in pandemic models. In: Proceedings 2010 winter simulation conference, 2010, pp 2230–2236
10. Andradóttir S et al (2011) Reactive strategies for containing developing outbreaks of pandemic influenza. BMC Public Health 11(supplement 1)
11. https://www.mygov.in/aarogya-setu-app/
12. Dutta SR, Roy M (2016) Providing context-aware healthcare services using circular geofencing technique. In: 2016 International conference on computing for sustainable global development (INDIACom) @ IEEE xplore
13. Boric-Lubecke O, Gao X, Baboli M, Singh A, Lubecke VM (2014) E-healthcare: remote monitoring, privacy, and security. In: MTT-S International on IMS, 2014 IEEE, 1–6 June 2014

Application of Neuroevolution in Autonomous Cars

G. Sainath, S. Vignesh, S. Siddarth, and G. Suganya

Abstract With the onset of electric vehicles, and them becoming more and more popular, autonomous cars are the future in the travel/driving experience. The barrier to reaching level 5 autonomy is the difficulty in the collection of data that incorporates good driving habits and the lack thereof. The problem with current implementations of self-driving cars is the need for massively large datasets and the need to evaluate the driving in the dataset. We propose a system that requires no data for its training. An evolutionary model would have the capability to optimize itself towards the fitness function. We have implemented neuroevolution, a form of genetic algorithm, to train/evolve self-driving cars in a simulated virtual environment with the help of Unreal Engine 4, which utilizes Nvidia's PhysX Physics Engine to portray real-world vehicle dynamics accurately. We were able to observe the serendipitous nature of evolution and have exploited it to reach our optimal solution. We also demonstrate the ease in generalizing attributes brought about by genetic algorithms and how they may be used as a boilerplate upon which other machine learning techniques may be used to improve the overall driving experience.

Keywords Neuroevolution · Neural networks · Genetic algorithm · Generation · Fitness · Selection · Crossover · Mutation

1 Introduction

The society of automobile engineers (SAE) has coined six different levels of autonomy beginning at level 0, the absence of any autonomy, to level 5, complete autonomy requiring no human intervention whatsoever. Currently, many luxury vehicles possess level 3 autonomy in terms of cruise control and active lane control, and a handful of vehicles possess level 4 autonomy. Level 5 autonomy in cars is still under research and development. The main barrier to attain this level of autonomy is the task of collecting data and the lack thereof. Although a deep model can be extremely

G. Sainath · S. Vignesh · S. Siddarth · G. Suganya (✉)
Vellore Institute of Technology, Chennai, India
e-mail: suganya.g@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_26

adept at learning and generalizing features, it can only learn what it sees. Humans can learn from different scenarios. Essentially, even if it learns to navigate through a busy street, it may not be able to correct oversteer or understeer due to several factors such as poor roads and tyre wear causing a loss of traction, which may not have been accounted for in the training dataset. That is why an evolutionary approach would solve these issues. What if the car could learn to drive on its own, via trial and error, over countless generations? It would have trained, evolved to overcome such edge cases and scenarios, and would know exactly what to do once it detects wheel spin or any form of loss of traction/grip.

2 Overview of Existing Systems

2.1 Neural Networks (Supervised Learning)

A neural network is an interconnected network of neurons, also called nodes. Each neuron has a set of output edges that activate based on the resultant value obtained from the weighted inputs it received from the previous layer (Fig. 1).

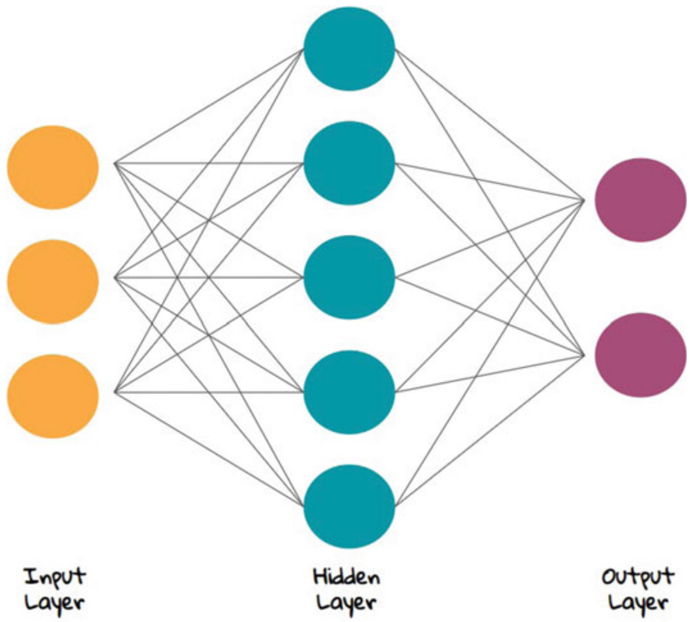


Fig. 1 Topology of an artificial neural network

In a supervised learning approach, we would have a list of attributes or features as our inputs and a list of targets as our outputs. We would then have to use back-propagation to train our neural network to correct its weight to suit our target and increase its accuracy.

So, in a situation where it is difficult to obtain a dataset large enough to train the neural network to a certain degree of accuracy, we will face problems arriving at our optimal solution. This is especially true in the scenario of self-driving cars, where large corporations like Nvidia use 1000 h of driving data to train their vehicle to navigate the roads. In such scenarios, we could adopt an evolutionary technique that requires no datasets and train our model in a simulated environment.

Although deep neural networks utilizing convolutional layers have performed extraordinarily in several scenarios, the problem arises after the fact that it can only learn what it is shown or taught. Since there are countless more possibilities of things that can happen on the road, which cannot be accounted for in the driving data we gather. Moreover, such networks are notoriously prone to over-fitting.

2.2 Reinforcement Learning

Reinforcement learning is another critical area of research in autonomous vehicles, where an agent learns to accomplish a task by gathering experience by itself, rather than through a supervised dataset. The basic gist of the algorithm is that an agent granted a reward when it performs an action that is desirable in the current scenario and gets punished if it does something undesirable. Although this form of carrot and stick approach seems to be how we, as individuals learn, the key drawback of this algorithm is that the agent has no prior experience whatsoever. We humans learn pretty quickly through this approach due to the generalization of a multitude of experiences that we have gathered from birth till date. This is not the case for the agent, and so it takes quite a while, depending on the complexity of the problem, for the agent to gather enough experiences in order for it to determine whether a certain action is desirable or not [1, 2] (Fig. 2).

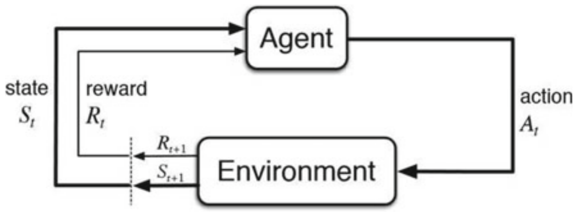


Fig. 2 Basic flow of reinforcement learning

3 Proposed System

Neuroevolution is a genetic algorithm that is used to evolve artificial neural networks. In this model, each species of a generation has a brain (the neural network) that has a set of genes (weights). In the beginning, all species of the population have random weights and hence perform random actions. It is through serendipitous discovery that a certain species gets closer to our solution. We select this species based on a fitness function and pick similarly performing species to perform crossover. After crossover, we mutate this gene and pass it on to the next generation. Owing to the nature of this sort of evolution, genetic algorithms are easily parallelizable as the actors (neural networks and their respective vehicle) of the population are individual entities independent of each other [3].

So the entire genetic algorithm can be summarized to three key processes:

- Selection: We select the best species of the generation based on the fitness function.
- Crossover: We crossover the genes of the population to converge onto our solution.
- Mutation: We mutate the genes, in the hope of a better solution, of the selected species following crossover (Fig. 3).

We can see that mutation and crossover seem a little opposite to one another. Mutation randomizes the weights of a certain percentage of neurons while crossover tries to converge them. There is a trade-off here between exploration and exploitation. Exploration via mutation is exploring new gene sets out of a hope that something new can lead to promising results whereas exploitation via crossover is taking what you

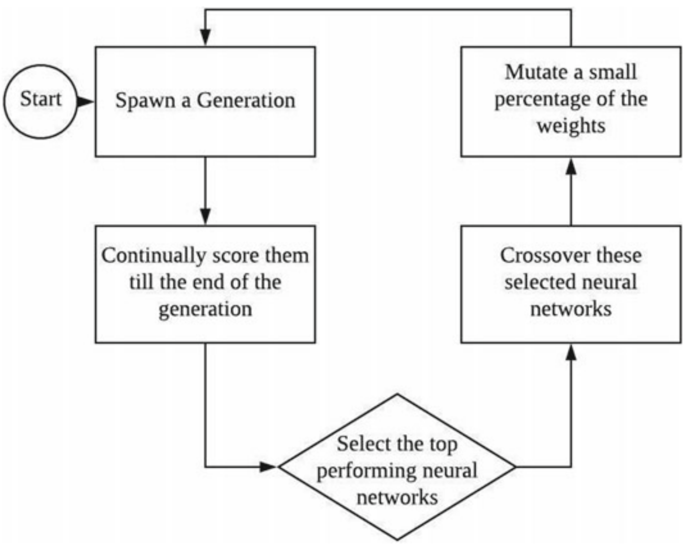


Fig. 3 Simple pipeline of neuroevolution

learned and using that information, combining the best, to inform newer decision-making processes.

4 Implementation

We chose to simulate neuroevolution using Unreal Engine 4, which is a game engine that utilizes Nvidia's PhysX Physics Engine to replicate real-world like vehicle dynamics, which is essential if we plan to transfer the learning that has happened in this environment.

Compared to using simulators such as CARLA, which was also built on the same engine, we have a lot more freedom when we build the whole environment from the ground up, in terms of level design, vehicle physics, frame times (time dilations) and overall gives more power to the user.

4.1 Vehicle Dynamics

FWD Layout—Since most vehicles these days in the low to mid-tier range are front-engine, front-wheel drives (FF), we chose this as our vehicle layout, and for the differential, we went with a limited-slip differential (LSD) that prevents wheel spin, which is getting more and more common these days. The transmission of the vehicle is set to automatic. The suspension settings have also been altered so that it favours understeer rather than it oversteer, as most manufacturers do these days, as it is easier to correct understeer. Weight transfer and tyre traction are also essential aspects that dictate the vehicle's physical handling and are simulated accurately.

RWD Layout—We also wanted to observe how this approach would fare on a more difficult layout which is harder to control, which is the front-engine, rear wheel drive (FR), also the typical sports car layout, as they more prone to over-steering and sliding through corners without proper throttle control and adequate countersteering. The suspension of this layout has also been altered so that it favours oversteering behaviour rather than understeer.

4.2 Neural Net

Each vehicle that we simulate has a brain that controls the values for the throttle pedal, the brake pedal, and the steering angles directly. This brain is our deep neural network which outputs a value from -1 to 1 for all the above inputs of the vehicle. The inputs to the neural network are the distances (normalized to 0–1) we obtained by tracing a point cloud around the vehicle. We also feed in the current speed of the vehicle (normalized to its maximum speed) and the angle between the velocity of the vehicle and its forward vector, which provides the neural net information about in which direction the car is sliding towards, if or when it does.

4.3 Genetic Algorithm

The genetic algorithm in this simulation is a higher-level entity that oversees the processes responsible for selection, crossover and mutation. It controls the mutation and crossover rates and is responsible for spawning and tracking all the features of the entire vehicle species for each generation of the population. Initially, in the first generation, all the weights of all the neural networks are initialized to random values, and it is through serendipitous discovery aided by the fitness function that we converge on to a solution through selection and crossover and also search for a better solution through mutation.

4.4 Working

In the first generations, all the weights of all the neural networks in the vehicles are initialized randomly, so they have no clue what to do when they are spawned and hence move randomly. To remove poorly performing agents, which is a crucial part of Darwinian evolution, "survival of the fittest", we de-spawn vehicles that crash into obstacles and guard rails or those that do not reach a certain threshold score within a predetermined period of time.

Pipeline:

1. Genetic algorithm entity spawns a vehicle population
2. Select top-performing vehicles based on the fitness function
3. Perform crossover by the weighted average of their weights with respect to their relative fitness in terms of the population
4. Mutate a small percentage of the weights by setting them to random values
5. Spawn the next generation of vehicles
6. Repeat for *n* generations till satisfactory behaviour is observed (Fig. 4).

Selection For each vehicle, the distance travelled (in the direction of the course) each frame, which we call the score of the neural net, is calculated as:

$$\Delta d = v \times \Delta t \tag{1}$$

$\Delta d =$ distance travelled that frame
$v =$ instantaneous speed
$\Delta t =$ frame time.

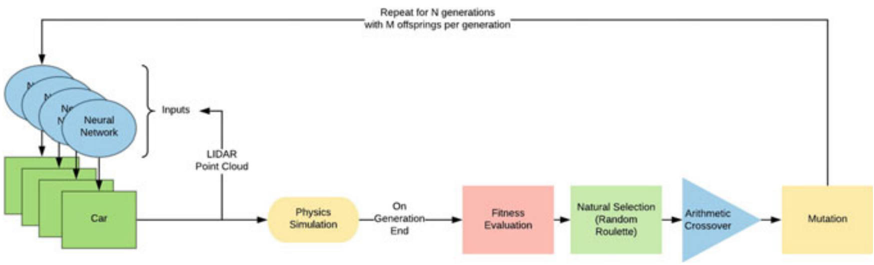


Fig. 4 Neuroevolution architecture

In order to prevent over-correcting behaviour and that it does not game the fitness score, we increment its score only when the angle between the velocity vector and the car's forward vector is less than a threshold value which we set as 10°.

From this, we calculate the net score each frame, which is the total distance travelled (until it de-spawns) and is calculated as:

$$\text{score} = \sum \Delta d \tag{2}$$

At the end of each generation, the relative fitness of each neural network is calculated as:

$$\text{fitness}_i = \frac{\text{score}_i}{\sum_{j=1}^{p} \text{score}_i} \tag{3}$$

where

fitness_i = relative fitness of the current neural network
score_i = total distance travelled by the vehicle
p = total population of the generation.

Now, for spawning the next generation of vehicles, we pick the top n vehicles with the greatest fitness (n could be selected arbitrarily, we chose it to be 1/10th of the population, p).

Crossover We now perform an arithmetic crossover of these n species by weighted addition of their weights with respect to their fitness. For each connection in the neural network, the weight of the connection after crossover is calculated as:

$$\text{new } w = \sum_{i=1}^{n} (w_i \times \text{fitnees}_i) \tag{4}$$

where

fitness_i = relative fitness of the current neural network
w_i = weight of a certain connection in the neural net

n = the top selected species of the generation.

Mutation Once we perform crossover for about 80% of the weights in the neural network, we then move on to mutation which is typically done to about 20% of the weights by using a random function on the weights.

5 Results

Within the simulated environment, we have observed that the population over several generations has evolved to not crash into obstacles, and also through sheer randomness have decided to stick to one side of a lane in certain simulations. They also perform advanced traffic management techniques such as zipper merges. A zipper merge is when a car continues to stay on the lane even after a blockade is located. They merge into the free lane only once they are close to the blockade in order to prevent traffic congestion that would occur if everyone stopped using that lane entirely (Figs. 5, 6, and 7; Table 1).

Modeling using various simulations, iterating and variating different parameters, are that when the fitness function is simple, it generalizes the course pretty quickly and is able to navigate it well over very few generations. But, when we alter it so that it favours a certain style of cornering or maintaining a certain amount of speed, it takes drastically more generations for it achieve this sort of specialization. This interpretation is backed by the discrepancies seen in the number of generations it took for the front-engine, rear-wheel drive (FR) layout compared to the front-engine, front-wheel drive (FF) layout (Fig. 8).

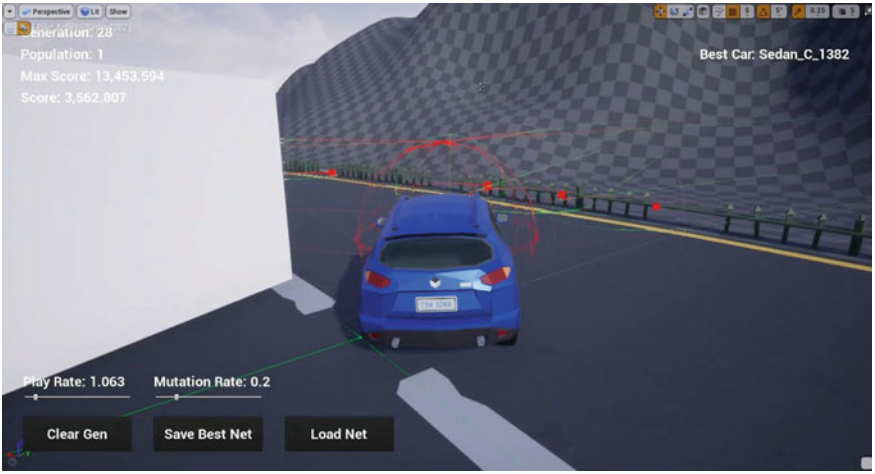


Fig. 5 Neural net switches lane only at the very verge of colliding on to the obstacle (white wall to left)

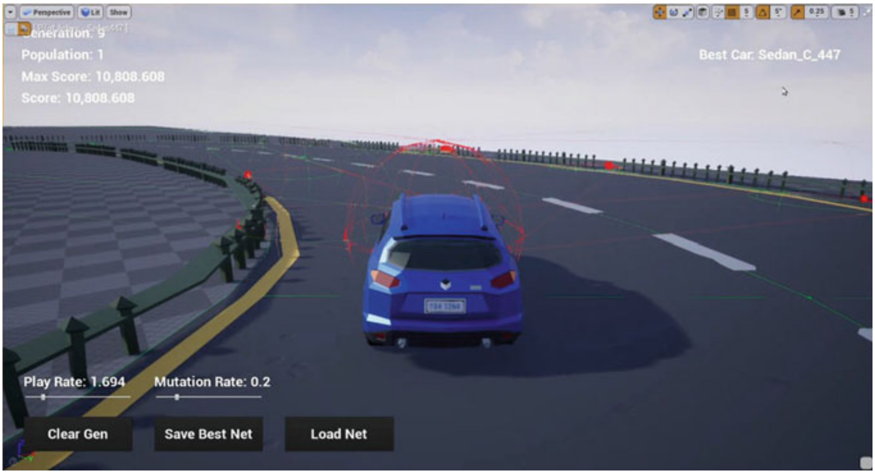


Fig. 6 Neural net has decided to stick to the left lane on the road through sheer randomness which can be nurtured by altering the fitness function

Fig. 7 Neural net has learned how to countersteer and control the car on the onset of oversteer

6 Conclusion and Future Work

Based on the results we have observed above, we can come to the conclusion that genetic algorithms such as neuroevolution can speed up the initial phase of generalizing features several fold compared to traditional techniques such as back-propagation which place the prerequisite of procuring a massive dataset for training so as to not overfit the solution. But once the network attains the basic cognitive

Table 1 Generations taken by the neural nets to evolve enough to navigate the entire course without crashing

Layout	Crossover rate (%)	Mutation rate (%)	Generation	Population
FR	80	20	97	4850
FR	80	10	225	11,250
FR	90	20	145	7250
FR	90	10	171	8550
FF	80	20	24	1200
FF	80	10	26	1300
FF	90	20	38	1900
FF	90	10	12	600

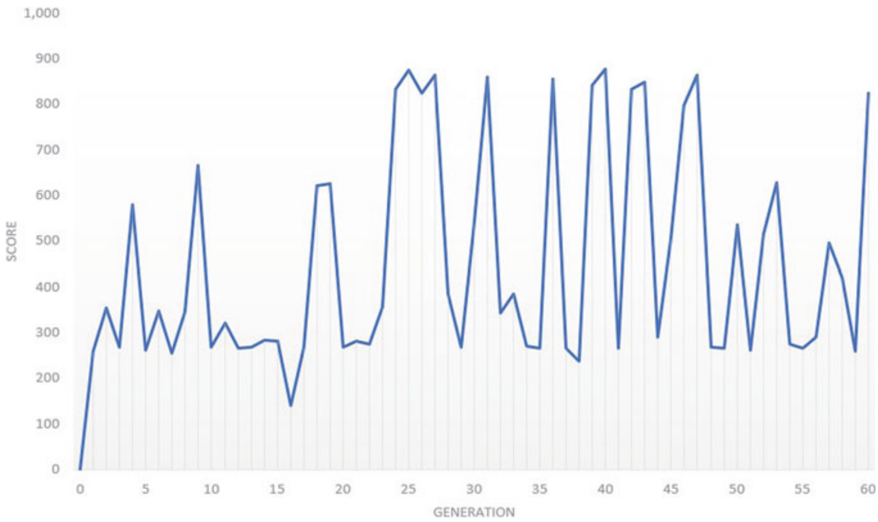


Fig. 8 Once one the neural nets hit a peak, it is able to constantly replicate the peaks which is an indication of evolution

abilities for driving, it can be further improved upon through reinforcement learning techniques such as deep Q-learning [4, 5], since it has already gathered a plethora of experiences over several generations, which is one of the key barriers slowing down reinforcement, since now we can quickly jump to the phase where the focus is more on obtaining as many rewards as possible rather than the initial phase of gathering experience where the agent primarily tries to just not get punished for its actions.

References

1. Saez Y, Perez D, Sanjuan O, Isasi P (2008) Driving cars by means of genetic algorithms. In: Rudolph G, Jansen T, Beume N, Lucas S, Poloni C (eds) Parallel problem solving from nature – PPSN X. PPSN 2008. Lecture Notes in Computer Science, vol 5199. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-87700-4_109
2. Such FP, Madhavan V, Conti E, Lehman J, Stanley KO, Clune J (2017) Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. ArXiv, abs/1712.06567
3. Stanley KO, Clune J, Lehman J et al (2019) Designing neural networks through neuroevolution. Nat Mach Intell 1:24–35. https://doi.org/10.1038/s42256-018-0006-z
4. Stanley KO, Miikkulainen R (2002) Efficient reinforcement learning through evolving neural network topologies. In: Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation (GECCO'02). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 569–577
5. Iglesias Rodriguez R, Rodríguez M, Regueiro C, Correa J, Barro S (2006) ICINCO 2006, Proceedings of the Third International Conference on Informatics in Control, Automation and Robotics, Robotics and Automation, Setúbal, Portugal, August 1-5, 2006, pp 188–195

Discrepancy Histopathological Breast Cancer Image Classification with Deep Convolutional Network

Hari Krishna Timmana and C. Rajabhushanam

Abstract In the world, bosom cancer is one of the main reasons that lead to death and has a huge death rate compared to other cancer types. Early detection may increase the chances of survival by proper treatment. The detection of cancer from images is very complex, difficult, and based on the pathologist's knowledge and experience. To overcome these issues, we need a robust system to classify the images into cancerous and normal automatically. To design the system, we have experimented with two image datasets that have class discrepancy: namely IDC cancer dataset and BreakHis dataset. Existing research reveals that the deep convolutional network-based solutions are advanced techniques to solve the complex issues in medical tomographic diagnosis, especially in analyzing complex histopathology images. The class discrepancy of the dataset impacts deep network performance during the training and validation. To balance the dataset classes, we applied various sampling techniques including oversampling and undersampling on the imbalanced dataset to achieve the good performance of classifiers that are developed from the deep neural network. The experimental analysis of the paper shows that synthetic sampling results the effective performance during system training.

Keywords Class discrepancy · Breast cancer classification · Deep learning

1 Introduction

In recent years, bosom cancer is the second-highest disease that leads to death after lung cancer [1]. This is the common and primary cause of deaths in middle-aged women. As per breast cancer survey [2], introspective ductal carcinoma (IDC) is the most commonly occurring disease and occupied 70–80% among all other breast cancer sub-types. In the US, more than 0.18 million women are newly affected by breast cancer every year [3]. However, early detection of cancer and medical treatment is highly required for the survival of the cancer patients [4].

H. K. Timmana (✉) · C. Rajabhushanam
Bharath Institute of Higher Education and Research, Chennai 600073, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_27

Bosom cancer analysis consists set of procedures that include physical assessment and medical tomography tests, namely: computerized mammograms, magnetic resonance imaging (MRI), ultrasound, biopsy [5]. The biopsy is the common and most reliable technique among others that identifies the malignancy in a predictable area. In this process, the sample cells are collected and extracted for the procedure which involves in collection and extraction of tissue samples to examine the malignancy using a microscope. Finally, pathologist classifies and presents histopathological instances by inspecting hematoxylin–eosin (H&E) stain under the microscope. In this process, it inspects the related areas of complete tissue with tissue architecture to assess the impact of cancer [6].

Even for experienced pathologists, manual analysis of bosom cancer is a complex and particular task due to inter-observer differences. The machine learning-based computer-aided system is required to help physicians during cancer analysis to reduce the load [7–9]. Automation of cancer analysis is the latest research area, and the deep network-based system provided more performance compared to other machine learning models [8, 10]. Andrew Janowczyk et al. [7] surveyed available deep learning methods using seven different histopathological datasets, and the results show that the deep learning methods provided high performance among other techniques.

Nuclei segmentation, mitosis detection, and lymphoma classification are the major applications that use the deep learning models and concepts in the medical tomography domain. It is difficult to develop an optimal deep learning model and challenging in designing classifying algorithms if the dataset contains class discrepancy. In recent years, medical pathology becomes one of the most interesting domains for researchers [8, 11–13]. Even though some of the machine learning algorithms [14, 15] result in a good performance but there is no further research available related to class discrepancy datasets for classification using deep network models. We espouse a set of sampling techniques to address the impact of the class discrepancy. The experimental results show that the class discrepancy impacts the model performance negatively and gives deprived results compared to the balanced dataset.

2 Dataset and Protocol

We used imbalanced datasets including ductal carcinoma malignant (IDC) dataset and BreakHis dataset to assess the CNN model performance. A brief note on experimental datasets is given as follows.

2.1 IDC Dataset

The first dataset [7] comprised of 162 slide images of bosom malignant growth examples examined at 40×. The dataset is easily accessible and includes 277,524

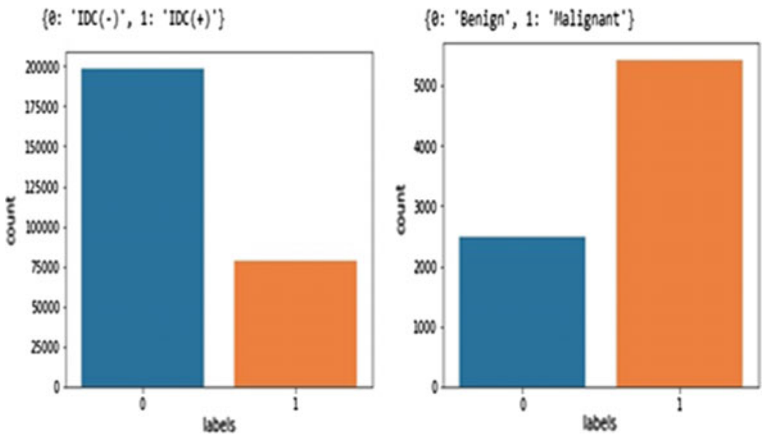


Fig. 1 Overview of experimental datasets

RGB computerized image patches 50 × 50 [16]. Figure 1a shows that the dataset is imbalanced as information classes not equally distributed. It depicts that IDC negative images are more than three times of IDC positive images. The main objective is to separate IDC positive images from IDC negative images.

2.2 BreakHis Dataset

BreakHis [17] is a complex dataset that has 7909 pictures and eight subclasses of bosom disease. It is collected from 82 mysterious patients at the Pathological Com-position and Cytopathology Laboratory, Brazil. BreakHis dataset is of four magnification factors: 40×, 100×, 200×, and 400×. Under the microscope, both the cancerous and noncancerous tumors are stored into a variety of pathologists which depend on the tumor cell.

The image is a three-channel RGB of 8 bit, and each channel is of size 700 × 460. These datasets are highly imbalanced: malignant images are of 68.64%, and benign images are of 29.36%. In this dataset, malignant image volume is two times than benign image volume. The main objective for this dataset is to classify the malignant (cancer) images and benign (normal) images.

Figure 1a, b shows the overview of the class discrepancy for experimental datasets. IDC has cancer images three times than normal, whereas BreakHis dataset has malignant images two times than benign images. Even this problem is considered as binary classification, it is challenging due to class discrepancy. Most of the AI techniques, including machine learning and deep learning, consider evenly distributed datasets for training and validation. So, the class discrepancy impacts the prediction capacity of the network model. The model achieves high performance for the majority of highly distributed classes and does not predict minority class.

2.3 Dataset Skewness

Training the model with a class discrepancy dataset is the main issue that impacts the performance of the model negatively in both training and validation. The classification techniques that are commonly used are designed to deal with balanced class datasets and provide improved performance. Most of the real-world applications such as fraud detection, medical image analysis, and satellite image oil spill recognition are considered as datasets with the class discrepancy. In medical tomography, minor samples are more appreciated than others.

Several methods are proposed to address the issue of the class discrepancy of the dataset. The main goal of them is to improve the classification performance of the experimental model. Machine learning models resolved the issue based on sampling [18] and cost [19]. The sampling techniques are classified into oversampling, undersampling, and hybrid sampling that contains both of the techniques.

The undersampling technique resamples the dataset until the minority class reaches near to the majority class by removing the samples. More samples will be discarded if the class discrepancy is high, and it has more risk to lose valuable instances from the dataset. In the oversampling technique, the model may lead to overfitting by adding duplicate samples to the dataset. This issue can be addressed by adding original or augmented images. Hybrid techniques are implemented to address the issues of oversampling and undersampling that generates artificial samples from minority class rather than duplication. These generated samples are added to feature space, and there is no loss of valuable information related to the original dataset. Synthetic minority oversampling technique (SMOTE) [18] and adaptive synthetic (ADASYN) [14] are common and mainly used hybrid algorithms for classification on deep neural networks. The SMOTE has three main steps to address the class discrepancy

- Choose a sample i from minority class
- Choose another sample j from nearest neighbors of i
- New sample s is generated by interpolating and b randomly.

Figure 2 shows six images of cancer and normal from the IDC dataset along with two new IDC cancer images created using the SMOTE technique from IDC positive (cancer) images. In this work, we experiment with several oversampling and undersampling techniques to address the class discrepancy of both datasets and evaluate them in terms of accuracy, sensitivity, specificity, f1-score, and BAC using deep neural network.

3 Models and Methods

In this section, we define the experiments to show the impact of the class discrepancy of the dataset on the classification performance of the deep network. The CNN

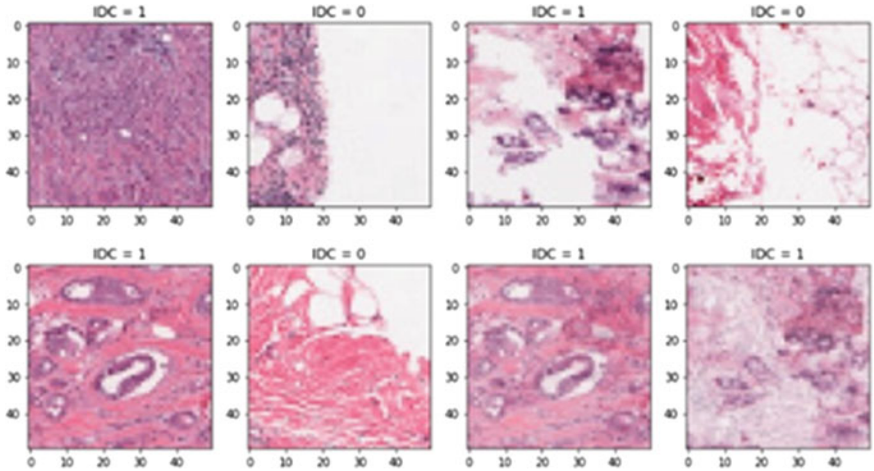


Fig. 2 Images generated through SMOT technique

contains several layers followed by the classifier to classify the images. And it converts the image data into feature map arrays at each stage. It takes the input image, trains the model, and provides the output after each block. The output of the final layers shows classification information.

3.1 System Architecture

Keras provides a high-level abstraction for TensorFlow back-end to implement deep neural networks. It provides rich application programming interface (API) to implement deep learning networks easily and quickly. For our experiments, we have implemented a simple convolutional neural network using Keras that contains three convolution layers and two output classes. We trained and validated the proposed model for 30 epochs with a batch size of 16. The summary of the experimental CNN model architecture is shown in Table 1.

Table 1 Tabular view of experimental CNN architecture

Layer	Kernels	Activation	Kernel size
Input	3	–	$50 \times 50 \times 3$
Conv2D	32	ReLu	3×3
Conv2D	64	ReLu	3×3
MaxPool2D	–	–	2×2
FC	128	Dropout (0.25)	–
FC	2	Dropout (0.25) + ReLu	–
Sigmoid	–	–	–

The projected CNN model consists of two convolution layers trailed by a max-pooling layer that has a kernel size of 2×2. After that, we have two fully connected layers that contain a dropout of 0.25 and 0.5, respectively. In this paper, we are considering the scenario as the binary classification. Hence, we applied the sigmoid as activation function and binary cross-entropy as the loss function. We used Adam optimizer to fine-tune the CNN model.

3.2 Experiments and Evaluation

To experiment with IDC dataset on the convolutional neural network, images are divided into training and test dataset. Training data has 70% of images where test data has 30% images. Test dataset is used to evaluate CNN after training.

BreakHis dataset has 7909 images with a height of 700, a width of 400, and magnification factors of $40\times$, $100\times$, $200\times$, $400\times$. Usually, CNN extracts more features from large images and becomes a complex system. To address that issue, reduced the image sized to 50×50. So that the calculation complexity is reduced compared to large-size images. Images of magnification factors including $40\times$, $100\times$, $200\times$, $400 \times$ become 50×50 images for the testing and validation of the CNN model. The dataset is randomly divided into 70 and 30% for model training and testing, respectively. We assessed the model for the images, not for the number of patients.

Spanhol et al. [20] and Robertson et al. [7] have provided the experimental results for image classification using convolutional neural network. We are comparing our results with their results in terms of specificity, accuracy, BAC, and f-score. As indicated [21], BAC is obtained using (1)

$$\mathrm{BAC} = \frac{\text{sensitivity} + \text{specificity}}{2} \tag{1}$$

In medical tomography, accuracy indicates that the capability of recognizing and classifying the image. F1-score and BAC are also keys to measure the model performance.

4 Results and Discussion

The performance metrics of the experimental datasets on convolutional neural network are as shown in Tables 2 and 3. The experimental results display that the class discrepancies impact the CNN model performance negatively. And also show that the oversampling technique results in better performance metrics in terms of accuracy over other mentioned techniques. Both table results presented that sensitivity and specificity of the dataset are highly improved compared to imbalanced one.

Table 2 Performance metrics of breakhis dataset

Sampling	Sensitivity	Specificity	F-score	Accuracy	BAC
Imbalanced	90.05	73.96	89.17	85.01	82.01
Undersampling	72.52	86.67	78.04	79.59	79.6
Oversampling	82.23	91.27	86.13	86.13	86.75
ADASYN	77.53	84.92	80	81.37	81.23
SMOTE	90.9	80.34	86.34	85.62	85.62

Table 3 Performance metrics of IDC cancer dataset

Sampling	Sensitivity	Specificity	F-score	Accuracy	BAC
Imbalanced	71.41	90.98	73.59	85.4	81.19
Undersampling	87.89	73.36	81.94	80.62	80.63
Oversampling	84.35	84.54	84.43	84.44	84.49
ADASYN	82.97	85.1	84.02	84.02	84.05
SMOTE	80.85	90.12	84.78	85.48	85.48

The class discrepancy issue resolved using the undersampling technique that results are improved sensitivity and drip in other metrics including specificity, accuracy, BAC, and f-score.

Table 2 shows that the oversampling approach improves the precision score by 3–4% over the imbalance dataset. The sensitivity and specificity of the discrepancy dataset have resulted in 71.45 and 90.98%. But in Table 2, the CNN system results show better efficacy, specificity, F-score, precision, and adjusted accuracy and rather improved around F-measure and BAC to 84.78% and 85.48%, respectively. Earlier results of the datasets are 71.8%, 84.23% [7] and 76.48%, 84.68% [22]. Advanced apps absorbed IDC and non-IDC district recognition throughout the slide pictures, although we are thinking of grouping into two classes of picture patches and reversing the issue of segmentation in [22].

Table 3 shows the experimental results of binary classification (i.e., Benign vs. malignant) on the BreakHis dataset. It also depicts that the balanced dataset gives more performance compared to the imbalanced dataset. The experimental values show the improvement after balancing the discrepancy dataset in terms of performance metrics, including accuracy, sensitivity, F-score, BAC, and specificity. From this, we can observe that the accuracy of the system is high during training, whereas low while testing phase due to the class discrepancy.

Accuracy is assessed on scaled images. The best performance (accuracy) of CNN at the image level is 86.13% with the oversampling method where unbalanced dataset results from 85.01%. The performance metrics shown in Table 2 are obtained from 40× images of the BreakHis dataset. Metrics of other magnification factors including 100×, 200×, 400× are reported in [10].

5 Conclusion

In this paper, we have analyzed the class imbalance impact on convolutional neural network performance and experimented with multiple different approaches to address the discrepancies. Class discrepancy impacting the classification performance of the model negatively. In most of the scenarios, oversampling results in better performance over other methods to address class discrepancy issues. For both experimental datasets, the oversampling method improves performance compared to other approaches.

So, we can conclude that the SMOTE results in a better and robust performance to address the class discrepancy issue in breast cancer classification. As a feature of future work, we might want to play out our examinations on various datasets from the medical domain to see how and which technique is the most significant for arrangement from these datasets.

References

1. World Health Statistics 2020. https://apps.who.int/iris/bitstream/handle/10665/332070/9789240005105-eng.pdf?ua=1. Last accessed 2020/08/16
2. Understanding Breast Biopsy What is a Breast Biopsy? (2020)
3. GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 v1.0. (2012)
4. Saving Women's Lives. National Academies Press (2005). https://doi.org/10.17226/11156.
5. Elston CW, Ellis IO: Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Elston CW, Ellis IO (1991) Histopathology 1991; 19:403–410. AUTHOR COMMENTARY. Histopathology. 41, 151–151 (2002). https://doi.org/10.1046/j.1365-2559.2002.14691.x
6. Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson ANA, Nelson HD, Pepe MS, Allison KH, Schnitt SJ, O'Malley FP, Weaver DL (2015) Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA—J Am Med Assoc 313:1122–1132. https://doi.org/10.1001/jama.2015.1405
7. Janowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 7. https://doi.org/10.4103/2153-3539.186902.
8. Cruz-Roa A, Basavanhally 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: Gurcan MN, Madabhushi A (eds) Medical imaging 2014: digital pathology, p 904103. SPIE. https://doi.org/10.1117/12.2043872.
9. Mac Namee B, Cunningham P, Byrne S, Corrigan OI (2002) The problem of bias in training data in regression problems in medical decision support. Artif Intell Med 24:51–70. https://doi.org/10.1016/S0933-3657(01)00092-6
10. Kowal M (2014) Computer-Aided diagnosis for breast tumor classification using microscopic images of fine needle biopsy. In: Advances in intelligent systems and computing. Springer, pp 213–224. https://doi.org/10.1007/978-3-642-39881-0_17
11. Filipczuk P, Fevens T, Krzyzak A, Monczak R (2013) Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. IEEE Trans Med Imaging 32:2169–2178. https://doi.org/10.1109/TMI.2013.2275151

12. Kubat M, Holte RC, Matwin S (1998) Machine learning for the detection of oil spills in satellite radar images. Mach Learn 30:195–215
13. Chen C, Liaw A Using random forest to learn imbalanced data
14. He H, Bai Y, Garcia EA, Li S (2008): ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In: Proceedings of the international joint conference on neural networks, pp 1322–1328. https://doi.org/10.1109/IJCNN.2008.4633969.
15. He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 21:1263–1284. https://doi.org/10.1109/TKDE.2008.239
16. Provost F (2000) Machine learning from imbalanced data sets 101
17. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng 63:1455–1462. https://doi.org/10.1109/TBME.2015.2496264
18. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority oversampling technique. J Artif Intell Res 16:321–357. https://doi.org/10.1613/jair.953
19. Hong X, Member S, Chen S, Harris CJ (2007) A Kernel-based two-class classifier for imbalanced data sets. IEEE Trans Neural Netw 18. https://doi.org/10.1109/TNN.2006.882812
20. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) Breast cancer histopathological image classification using convolutional neural networks. In: Proceedings of the international joint conference on neural networks. Institute of Electrical and Electronics Engineers Inc., pp 2560–2567. https://doi.org/10.1109/IJCNN.2016.7727519.
21. Sun Y, Wong AKC, Kamel MS (2009) Classification of imbalanced data: a review. Int J Pattern Recognit Artif Intell 23:687–719. https://doi.org/10.1142/S0218001409007326
22. Veta M, Pluim JPW, Van Diest PJ, Viergever MA (2014). Breast cancer histopathology image analysis: a review. https://doi.org/10.1109/TBME.2014.2303852

Detection and Prevention of Cyberbullying Using Ensemble Classifier

S. Thanigaivel, S. Harshan, M. Syed Shahul Hameed, and K. S. Umadevi

Abstract Social media has an irrational behavior due to harassment by cyberbullies. Firstly, most of the cyberbully prevention systems concentrate on a particular social media platform (SMP). Secondly, they address a particular category of cyberbullying. Finally, the efficiency and accuracy of the detection may vary depending on the classifier. The proposed work aims to overcome the above-mentioned issues using machine learning algorithms. Knowledge gained from the datasets is used to train the highly efficient and training machine learning algorithm. The implementation details provide many useful insights about cyberbullying prevention and systematically analyze cyberbullying detection on various topics with many social media platform analysis with machine learning-based models.

Keywords Cyberbully · Ensemble classifier · Machine learning · Twitter · Stemming · Vectorization

1 Introduction

Picture speaks more than a word likewise the words create more impact when being said online than in person. Online communities use these comments and share it to create sensational influences like bullying, stalking, trolling, or other harassment. Cyberbullying is a kind of bullying leads to harassment through electronic media [1]. There are many witnesses and/or victims who got offended by cyberbullying which leaves an impact on victims from psychological disorder to threat/loss of life [2]. The issue raises since the individuals related to the victim are also part of these online

S. Thanigaivel (✉)
Jeppiaar Maamallan Engineering College, Sriperumbudur, India

M. Syed Shahul Hameed
College of Applied Science, Salalah, Sultanate of Oman

S. Harshan · K. S. Umadevi
Vellore Institute of Technology, Vellore, India
e-mail: umadeviks@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_28

communities, and thus, they play role in it, thereby the impact increases. Though the Information Technology Act 2000 has no law to protect from cyberbullying, the offences like bullying, stalking, defamation, etc., are punishable under penal code sections [3]. Due to the freedom, the online communities have in accessing/sharing thoughts in social media for, the effect of cyberbullying is increasing every day. Even, using abusive words in social media is likely to be trending these days. Social media platforms (SMP) is taking initiatives to eliminate these issues on their respective platform [4–6]. Anyone having an Internet connection or owning a smartphone can cyberbully another person without exposing their real identity. In this situation, cyberbullying is characterized by the difference in the strength of the victim and the aggressor. This type of bully can be anyone: large or small, celebrity or an outcast, an honor student or a challenged learner and can take place anywhere, even in your own house.

According to Center for Disease Control, those kids/adults who witnessed bullying and not being part of it have more feelings of less connectedness and need of advice from the elder people [7]. Recent surveys stated that the most common form of bullying is name-calling and spreading of rumors. Also, the same survey stated that about 59% of teenagers in America experienced some form of cyberbullying through their phones, and nearly about 4400 individual cases have committed suicide because of cyberbullying [8, 9]. Of Course, the main intention of using social media websites to be connected with people from all over the world, to share their feelings and spread the word about positivity and reducing negative things. But due to cyberbullies, the positivity vanishes to a point where we can only see negative criticism and causing damage to personal extent. Cyberbullying as a type of bullying has the same basic features as (1) intentional aggressive behavior; (2) repetitiveness or regularity of act; (3) inequality of victim and aggressor; (4) consequences in the form of mental or physical harm [1]. A special feature is that cyberbullying occurs through electronic technologies. Trusting relationships with parents is an important element of alternate cyberbullying. The child should know that he or she will not be scolded, that phone or computer will not be taken away. To young people today, using their technology is akin to being banished from their social world, and they often would rather risk from their being the target of cyberbullying than losing their technology.

Statistics show that cyberbullying is always in the way of social media development, despite opposition from the state and the introduction of various anti-bullying programs [10]. A vivid example is the USA, where each state has a statutory regulation of anti-bullying and cyberbullying events at school. According to symantec, about eight out of ten people were offended due the various sorts of cyberbullying in India. Out of these, around 63% suffered due to online abuses and affronts and 59% were underwent rumors and gossip for degrading their images. Detecting cyberbullying has always been a challenge, and several problems have to be faced with the dataset, algorithm, accuracy level of the results, and social media's security terms and policies. The aim of the proposed work is surpassing those challenges with a better model than the existing systems. Most of the contributions in detecting/preventing cyberbullying use either Naive Bayes or K-NN models to detect the cyberbullying. Hence, the proposed work uses an ensemble classifier with all those models embeds in them as a voting classifier which increases efficiency of our model. Moreover, it gives better results in terms of high detection rate, reduced unusual acts, getting direct results by reporting, false alarming of bully words, and reduced performance and computation time.

2 Related Works

Biswas et al. explain the concept of cyberbullying and its detections by using few machine learning algorithm techniques [5]. The author used K-NN, Naive Bayes, decision tree and SVM algorithms for the classification of bullying text. This simply gave the results and analysis over the comparison over these algorithms, and accuracy is pretty good. Then, Dadvar et al. present appropriate method to detect bullies based on the contents of the post or tweets which are being posted on social media [11]. It explains the usage of cyberbully features and user-based features to detect cyberbullying. This paper is about using incremental features for training SVM to classify bullying. There is no explanation for the scale for measuring the accuracy. Accuracy is comparatively low.

The automatic detection and prevention of cyberbullying for web 2.0, a framework like feature is presented for cyberbullying detection [12]. Authors contributed toward detection of cyberbullying in Dutch dataset containing social media messages and proposed a methodology to evaluate the acceptable comment of this data. They conclude that exploring the feasibility of automatic recognition could reduce the manual monitoring for cyberbullying detection. A shadow error analysis revealed that the implicit realizations of cyberbullying are fairly hard to recognize, as lexical cucs are such profanity. A comparative study of stemming algorithms helped a lot to understand the stemming feature in addition to text mining and information retrieval from the dataset [13–15]. Porter stemmer features made the stemming feature extraction process easier to convert the words to their alphas or stems.

The spam detection on twitter using traditional classifiers elaborates about the usage of an API-based detection technique for fake spam [16, 17]. Differentiating a legitimate user and spammer on twitter can also be an effect for cyberbullying. It compares four different classifier's results, namely SVM, K-NN, random forest, and Naive Bayesian and classifies the spams from the legitimate tweets. F-measure of this paper results in 95%. Integrating cyberbullying detection with this API will result in a great improvement for social media. To classify cyberbullying comments on Instagram social media for Indonesian selebgram using support vector machine, the authors used R library [18]. As a dataset, they have chosen comments from accounts of Indonesian selebgram/celebrities after surveying that people get bullied in Instagram more (42%) than other social media (Fig. 1).

It works on cyberbully detection with deep learning models [19]. They used three different datasets and four different DNN models by coupling with transfer learning beat state of the art result for three datasets. Severity of the bullying used over the

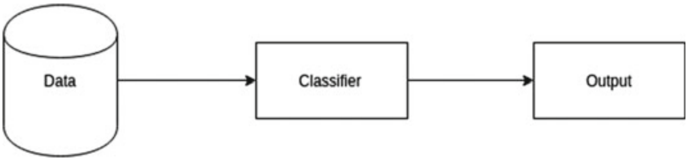


Fig. 1 Existing model

learning and testing improves the quality of the prevention. Actions are not taken detection only is not enough for prevention of cyberbullying.

Zhao, R., et al. explain the new method of detection in Twitter using bag-of-words [20]. As a dataset, they used texts or posts from Twitter. At first authors identified the list of insulting words using professional's familiarity and linguistic resources for implementing EBoW and then extended the features to define bullying features. Cosine similarity was used to assign different weights to the bullying features between word embedding. After that based on weight, they classified the intensity of cyberbullying. They trained and tested with fivefold comparison with BoW, EBoW, LDA, and LSA. The result of EBoW came out best of all. Precision was 76.8%, recall 79.4%, and F1 score 78.0%. After this representation, they intended to use SVM linear to classify and detect the texts. Sine, they have not classified yet, the result of detecting bully words is found to be ambiguous.

3 Proposed Work

3.1 Dataset Preparation

As an initial stage, the proposed work involves in collecting the dataset followed by which it loads the dataset and then performs pre-preprocessing. Since supervised learning is used, the data is split into two—training and testing with the ratio of 80%–20%, respectively. Furthermore the data is split into X and Y as data and labels, respectively. The dataset is collected from formspring website, and it is raw with various unnecessary characters, spaces, and punctuations. Preprocessing involves the process of sanitization of the dataset by removing these unnecessary punctuations and spaces and decomposing complex data into simpler and multiple parts which helps for better working of the model. The collected dataset is of CSV format (comma separated value) file which is most popular in use. We used pandas for the manipulation of the dataset.

3.2 Feature Engineering

Second step of our work is feature engineering. In this step, the fresh dataset is used to extract basic features which can be used directly in the machine learning model. This includes the creation of new features from the existing data.

3.2.1 Tokenization

The process of cutting and separating every word forms a document or sentence for conversation. Basically, each word is split from another word by inserting a symbol that includes space as a character, or single quote or dot or colon or semicolon or any user-defined character. So, tokenization is the process of splitting the texts into meaning pieces called tokens. Based on the requirements, we can define our own meaning of the token for using it in a meaningful way, as the non-letter or symbols perform the separation. The difficulty of tokenization is understanding the language types.

3.2.2 Stemming

It is the process of amalgamating the different forms of a word into a basic regular representation called stem. For example, the procedure used for information retrieval is based on the interest implied by the particular word, given a word "cutting" is changed into "cut." We have used two types of stemming Porter stemmer algorithm and Snowball stemmer algorithm.

3.2.3 Stop Word Removal

In this process, useless words are removed from the text conversation according to the vocabulary of the language. In our case, we used natural language toolkit English stop word package. For example, stop words are very frequently used common words like "are", "and", "that", "is", "was", etc. Since they are not used in the classification of the texts and return useless information, they can be removed.

3.2.4 Transform Case

In this step, the words in the dataset are changed to common cases either upper or lower for neutralizing the next process with process not to differentiate between uppercase and lowercase letters.

3.2.5 Lemmatization

It is one of the important steps of data preprocessing. This process is not like stemming and eliminates the improper words ensuring that the language of the root word. In this process, the root word is known as lemma. A lemma or lemmata is the canonical form, citation form of a set of words.

3.3 Feature Extraction

Feature extraction in machine learning is essential to reduce the data into easily manageable variables. Feature extraction often includes processes like selecting, merging variables for reducing the data that needs to be processed. Feature extraction needs to be carried without losing any data. This is also called vectorization. Here we used count vectorization, a form of feature extraction carried out mainly in text-based data.

3.3.1 Count Vectorization

The idea behind the count vectorization is very simple. We will be creating a matrix with words and its count as rows and columns. At first, the count is filled with 0's and increased with one each time we hit a corresponding word.

3.4 Model Training

Here in this step, we train the model using fit command. The model we use is the ensemble classifier model. In the classifier of ensemble, the results of sets of classifiers are combined in some way either weight or un-weighted voting to classify a new output. For the set of classifiers, we took four classifiers under weighted voting. They are Naïve Bayes classifier, logistic regression classifier, K-NN classifier, and SVM classifier model.

3.4.1 Naive Bayes

Naive Bayes algorithm is one of the probabilistic algorithms. It is based on Bayes theorem with simple assumption between pair of features. Most of the NLP problems use Naive Bayes algorithm. In Bayes theorem, probability $P(A|B)$ where A is the possible outcomes and B is the instance which needs to be classified is found.

$$P(A|B) = P(A|B) * P(B)/P(A) \tag{1}$$

When a Bayes algorithm is applied to the multinomial distributed data, then it is called as multinomial Naïve Bayes, and it is one of the Naive Bayes variants used in text classification.

3.4.2 Logistic Regression

Logistic regression is a classification algorithm when a decision threshold is considered. The important step of logistic regression is fixing threshold value, and it depends on the problem.

$$g(z) = 1/1 + e^{-z} \tag{2}$$

3.4.3 SVM

The SVM using a linear kernel is best with many features [21, 22]. Because mapping the information to the next dimensional space does not really improve the performance. In cyberbully detection, both the numbers of comments and features are large. Training a SVM with a linear kernel is quicker. The function of linear SVM is

$$K\left(x_i, x_j\right) = x_i * x_j \tag{3}$$

Here, x_i and x_j are the vectors within the input space.

3.4.4 K-NN

K nearest neighbor (K-NN) is a simple classification algorithm. K-NN generates similarities in input data and tries to put the new data in one of the available groups (Fig. 2). K-NN is also called a lazy learning algorithm since it performs actions during classification. Some of the formulas used for distance calculation are

$$\text{Euclidean} = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$

$$\text{Manhattan} = \sum_{i=1}^{k} |x_i - y_i|$$

$$\text{Minkowski} = \left(\sum_{i=1}^{k} (|x_i - y_i|)^q\right)^{i/q}$$

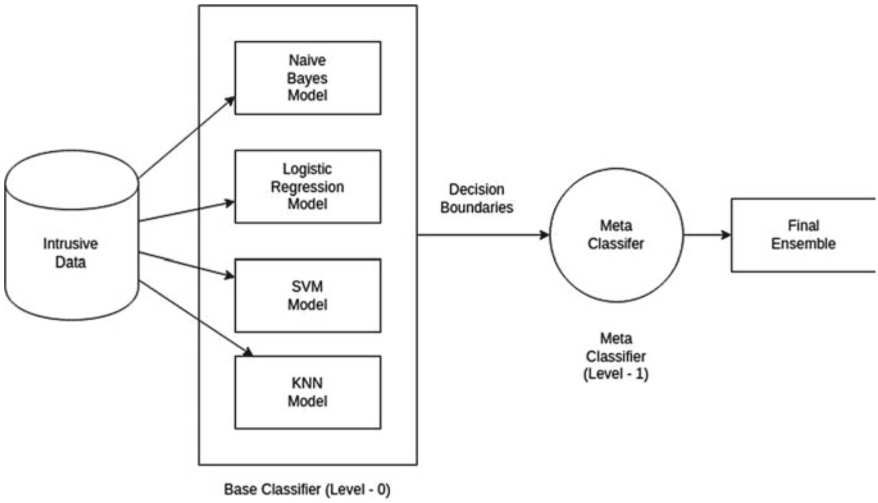


Fig. 2 Working model

4 Result Analysis

4.1 Precision

Precision is the set of found result that are relevant to the prediction:

$$\text{Precision} = \text{true positive}/(\text{true positive} + \text{true negative}) \tag{4}$$

For example, for cyberbully detection on a set of comments, precision is the number of true positive predictions divided by the total number of true positive and true negative predictions. The precision and recall are combined to provide a single measurement for a system, called F-score.

4.2 Recall

Recall is the set of the relevant results that are successfully retrieved.

$$\text{Recall} = \text{true positive}/(\text{true positive} + \text{false negative}) \tag{5}$$

For example, for cyberbully detection on a set of comments, recall is the number of true positive predictions divided by the total number of true positive and false negative predictions. As we mentioned earlier, it is combined with precision to obtain a single measurement, called F-score.

4.3 F-Score

F-score is the combination of both precision and recall. Recall is obtained by dividing the product of precision and recall with sum of precision and recall and multiplying with 2.

$$\text{F-score} = 2 \times (\text{precision} \times \text{recall}/\text{precision} + \text{recall}) \tag{6}$$

4.4 Results Observed

- *Naïve Bayes:* We expected multinomial Naïve Bayes as one of our best classifiers, and the results are promising. It gave an accuracy of 85.6%.
- *Logistic Regression:* Logistic regression is the best fitted classifier for our dataset. Using logistic regression, we got an accuracy of 87.1%
- *K Nearest Neighbors (K-NN):* We tried K-NN classifiers with different "K" values. $K > 2$ yields better results for our dataset. With this set-up, we got an accuracy of 84.2%.
- *Support Vector Machine (SVM):* There are several kernels used in SVM, but we have implemented a "Linear" Kernel. Linear SVM showed an accuracy of 84.3%.
- Ensemble Classifier: In ensemble, we used majority voting for decision making. Accuracy observed is >86%.

5 Live-System Implementation

Python library named tweepy was used to retrieve data from twitter. At first, applied used Twitter developer account to acquire elevated privileges like accessing twitter data outside of twitter website and app. Once an application is granted, a secret codes are generated which are used in Twitter API to access data. In the analysis, 20 user's home timeline tweets were extracted and preprocessed it and fed it for prediction. If the prediction was said to be "Yes," then it reports the tweet and gets the tweet's user ID for blocking the user. The data received from the twitter is in JavaScript Object Notation (JSON) which is key-value pair format. We used a function called api.home_timeline() to get tweets from the user's timeline and api.report_spam() for reporting. So when the tweets are posted, our API detects if any cyberbully comments are present, reports them spam, and the results are reported in Table 1.

Table 1 Result observed

Classifiers	Accuracy	F-score	Recall
K-NN	84	81	84
Naive Bayes	86	82	87
Logistic regression	87	83	87
SVM	84	82	84
Ensemble	86	82	87

6 Conclusion and Future Work

This work to detect and prevent cyberbullying in English language using machine learning algorithms with balanced data is clean. This work can process the cyberbully text and report them spam and also block them on twitter. It can also be converted into API and can be added to any text-based posting platforms like Facebook, reddit, etc. Efficiency of our work is calculated to be 86% and can be improvised by adding a few more classifiers to the ensemble or by improvising dataset. Since there is always a shadow line in the detection of cyberbullying due to the language and the way of understanding a normal word, few bullying is still left out of the box and can be improvised with emotional understanding of words with classifier, which can be our future work.

As what we saw from analyzing the dataset, it is true that people also use smileys in non-bully text which make those text cyberbully, understanding punctuation smileys from the dataset is required. Since most people tend to use their native language in their comments, the ability to predict cyberbullies in various languages other than English is necessary. Our method is based on supervised learning which differs from real-world data; we will be generating a dataset from real-world SMPs. Since there are no dedicated functions provided in the twitter API for selecting replies of particular tweets and direct messages, ability to search direct messages and tweet replies should be implemented.

References

1. Feinberg T, Robey N (2009) Cyberbullying. The Educ Dig 74(7):26
2. Kim YS, Leventhal B (2008) Bullying and suicide. A review. Int J Adolesc Med Health 20(2)
3. Rao S (2001) Information Technology Act: consumers' perspective. Econ Polit Wkly, pp 3501–3503
4. Whittaker E, Kowalski RM (2015) Cyberbullying via social media. J Sch Violence 14(1):11–29
5. Biswas S, Ayon RH, Rozario RE, Rahman FMM Cyberbullying classification and vulnerability detection on social media. Unpublished paper, Ahsanullah University of Science and Technology, Bangladesh, 2018.
6. Hamm MP, Newton AS, Chisholm A, Shulhan J, Milne A, Sundar P, Hartling L (2015) Prevalence and effect of cyberbullying on children and young people: a scoping review of social media studies. JAMA Pediatr 169(8):770–777
7. Centers for Disease Control (2019) National Center for Injury Prevention and Control. Preventing bullying

8. Machackova H, Dedkova L, Mezulanikova K (2015) Brief report: the bystander effect in cyberbullying incidents. J Adolesc 43:96–99.
9. Hinduja S, Patchin JW (2010) Bullying, cyberbullying, and suicide. Arch Suicide Res 14(3):206–221
10. Donegan R (2012) Bullying and cyberbullying: history, statistics, law, prevention and analysis. The Elon J Undergraduate Res Commun 3(1):33–42
11. Dadvar M, Trieschnigg D, Ordelman R, de Jong F (2013) Improving cyberbullying detection with user context. In: Serdyukov P et al (eds) Advances in information retrieval. ECIR 2013. Lecture Notes in Computer Science, vol 7814. Springer, Berlin, Heidelberg
12. Van Hee C et al Automatic detection and prevention of cyberbullying. In: International conference on human and social analytics, proceedings, Saint Julians, Malta, 2015, pp 13–18.
13. Jivani AG A comparative study of stemming algorithms. Int J Comp Tech Appl 2(6):1930–1938
14. Umadevi KS, Gaonka A, Kulkarni R, Kannan RJ Analysis of stock market using streaming data framework. In: Proceedings of international IEEE conference on advances in computing, communications and informatics, pp 1388–1390, 2018
15. Poddar K, Amali GB, Umadevi KS Comparison of various machine learning models for accurate detection of fake news. In Proceedings of IEEE conference on innovations in power and advanced computing technologies, vol 1, pp 1–5, 2019
16. McCord M, Chuah M (2011) Spam detection on twitter using traditional classifiers. In: Calero JMA, Yang LT, Mármol FG, García Villalba LJ, Li AX, Wang Y (eds) Autonomic and trusted computing. ATC 2011. Lecture Notes in Computer Science, vol 6906. Springer, Berlin, Heidelberg
17. Saleena N (2018) An ensemble classification system for twitter sentiment analysis. Procedia Comput Sci 132:937–946
18. Andriansyah M et al Cyberbullying comment classification on Indonesian Selebgram using support vector machine method. In: 2017 Second international conference on informatics and computing (ICIC), Jayapura, 2017, pp 1–5.
19. Agrawal S, Awekar A (2018) Deep learning for detecting cyberbullying across multiple social media platforms. In: Pasi G, Piwowarski B, Azzopardi L, Hanbury A (eds) Advances in information retrieval. ECIR 2018. Lecture Notes in Computer Science, vol 10772. Springer, Cham.
20. Zhao R, Zhou A, Mao K: Automatic detection of cyberbullying on social networks based on bullying features. School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (2016)
21. Suykens JAK, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett, pp 293–300
22. Mavroforakis ME, Theodoridis S (2006) A geometric approach to support vector machine (SVM) classification. IEEE Trans Neural Netw 17(3):671–682

Obstacle Avoidance for Aerial Vehicles in Autonomous Navigation

Abhiraj Chaudhary and A. Karmel

Abstract Obstacle avoidance is the back bone of autonomous navigation as it enables vehicles to reach desired location avoiding hurdles in the path. It is one of the ongoing challenging researches in the arena of cyber physical systems. In this article, comparison of various obstacle avoidance algorithms such as Artificial Potential Field (APF) approach, Vector Field Histogram (VFH) approach, Bubble Band approach, Mounted Sonar approach, Dist-Bug approach, bug-1 and bug-2 approach and Tangent Bug approach has been addressed. The main objective is obstacle avoidance, where obstacle can be in the form of radars or any other type of equipment. A novel obstacle avoidance procedure for low-altitude flying vehicles (Static Obstacle Avoidance) and high-altitude flying vehicle (Dynamic Obstacle Avoidance) has been proposed using A-Star and Deep Q-Network Reinforcement Learning techniques, respectively. Test bed has been created considering object model, sensor model, obstacle environment and controller. Implementation of the same has been done through visualization using Pygame for A-Star approach and Director for Deep Q-Network Reinforcement Learning.

Keywords Reinforcement learning · Unmanned aerial vehicles · Deep learning · Q learning · Deep Q-network · A-star · Artificial intelligence

1 Introduction

Nowadays, use of drones, robots or aerial vehicles have increased by the armed forces and other intelligence services for gathering information for the purpose of monitoring various situations in combat along the line of control. There are times, when a few moves should be made, without getting in reconnaissance of the foe, or neighboring countries utilizing unmanned aerial vehicle (UAV) [1–5] or aerial vehicles. Escaping from various obstacles put in the path needs to be avoided, so that the target is reached and desired action to be performed without raising any alarms,

A. Chaudhary · A. Karmel (✉)
Vellore Institute of Technology, Chennai, Tamil Nadu 600127, India
e-mail: karmel.a@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_29

regardless of whether its avoiding enemies' radars, or various infrared equipped gadgets or any other type of obstacle. If a system is in place to find the best path [6, 7] or getting information about location of all obstacles beforehand, it will help the armed forces in executing their actions with more precision and higher efficiency.

Currently only fifth generation aircrafts, also known as stealth planes, are able to evade radar that too using hardware assistance not software assistance.

Lot of researches was done on various obstruction avoidance algorithms, but still nothing concrete has been found supporting use of a particular algorithm in a specific place. Moreover, there have been findings and experiments pertaining to use of various technologies ranging from image detection to Artificial Intelligence (AI) for obstruction avoidance. This line of research is still wide open with various possible outputs.

The proposed system is distributed into two sections which is implementation and visualization using A-Star and Pygame and the other is Deep Q-Network Reinforcement Learning [2], respectively. In the first section, obstacle avoidance and determination of the best path to a target location and visualization of the obstacles has been dealt. Latter section deals with learning about obstacle avoidance [8, 9] using dynamic obstacles, after training, object being able to successfully avoid all the dynamic obstacles and reach the target location.

2 Literature Review

In the field of self-maneuverable vehicles, dodging obstructions is a highlighted issue. Setting or forming the environment in which the object would move is not feasible at all. Able to dodge obstructions which can be static / non-static, whose details are unknown, object must be provided with an obstruction avoidance algorithm. Process of avoiding obstructions is a dual stage process which includes, ability to perceive the obstruction in path of the object and ability to change its course so as to avoid the obstruction in the environment.

Artificial potential field (APF) approach [10, 11] is based on physics concepts, i.e., potential of object (+ve or −ve). Objects with same polarity always repel each other and opposite is true. In APF, hypothetical polarity is assigned to the object, destination and the hurdles. Key point is object and hurdles in its path are given the same charge (+ve) and destination is given opposite charge (−ve). Same charges attract which implies object moves toward the destination and with opposite charges moves away. Vector field histogram approach (VHF) [12, 13] is an approach which uses the concept of histograms from statistics to avoid obstructions. VFH uses some kind of technology embedded with object which scans the neighboring environment and sends it to backend for analysis. A polar histogram is plotted with knowledge acquired through external technology set up on object. Histogram is plotted and further analyzed by the system, which highlights the areas which are free of obstacles in the field of view. Choosing which course to tread upon is based on simple logic of least distance. This process keeps on going until object reaches the destination.

In supervised learning, the drawback is the massive data set labeling is heavy and time consumption too. Reinforcement learning [14] overcomes these issues by letting an agent learn with the data from its environment. Bubble band technology [15] merges path planning and disruption. The bubble band represents the maximum distance between the agent for a given position and obstacles in every direction. In route planning, the agent follows a previously calculated path without allowing too much distance between the bubbles of continuous positions and thus laying out smooth paths.

In tangent bug algorithm (TBA) [16], hierarchical data are used to construct a local tangent graph to select the right direction locally when it is close to the target position. As the sensor's maximum detection range increases, the TBA paths are produced in a simpler environment that reaches the right path globally.

One of problems associated with Bug-1 algorithm [17] is time taken, on the other hand object reaches its desired destination with high trust. Bug-2 algorithm [17] computes slope of the line connecting starting and final destination when it encounters an obstruction it tries to find the coordinate from where slope of line joining coordinate and destination is same as initial calculated slope by moving on borders of the object

A-Star approach [18], finds an obstacle free path considering obstacles in path finding procedure. A-Star algorithm uses weighted graph which represents that the factors of distance and time effects the method in finding the most suitable and appropriate path (Table 1).

3 Proposed Approaches for Obstacle Avoidance

3.1 Proposed Approach for Obstacle Avoidance in Low-Altitude Flying Vehicles

The object in question is tasked to achieve some goal by reaching a destination. During the travel, many static obstacles come in place like trees, hills, branches, buildings. Hence during flying at low altitude only static obstacles come in place. The environment where the aerial vehicle enters is certain and obstacles are landscape. The proposed system is distributed into two sections which is implementation of approach and a working prototype displaying obstruction avoidance using A-Star and Pygame.

The proposed approach is a combination of tangent bug and artificial potential method. For low-altitude flying vehicles or the case of non-moving obstacles this algorithm helps in removing disadvantages and couple the pros of tangent bug and APF approach. In the scenario, the object in question, the obstruction and the destination are given some charges. Object and obstruction are given same charges so that they go far from each other and on the other hand destination is given complete opposite of what object has been given. When an obstruction is perceived it applies the

Table 1 Comparison of various obstacle avoidance algorithms

Algorithm	Pros	Cons
APF approach	Understandable and simple to execute	1. Object comes to rest when net sum of charges becomes zero even though it may not have reached the location 2. Gets choked up in situation where very less gap present in multiple obstructions
VFH approach	Noise due to mounted sensors on object is removed	1. Not able to reach destination sometimes 2. Gets choked up in scenario of closely obstructions 3. Turns a blind eye toward attributes of the object
Bubble band approach	1. Better than the histogram approach in choked up scenario 2. Less memory and power consumption 3. Can avoid both obstructions at rest and in-motion	1. Hiccups during process 2. In depth path finding required 3. Noise from tech stack mounted object may cause an issue
Bug-1 approach	In a set duration of time, it will find a path, else will stop the process of path finding	1. Reduced efficiency is due to lot of memory and calculating power being consumed 2. It is a process which is very long and calculates each and everything
Tangent bug algorithm	Covers least path as compared to others as heuristic value is reduced	It loses identity and its process if length of path between position and goal starts to spike up

concept of tangent bug which is finding the slope from the point around obstruction to destination which is equivalent to slope from starting to beginning.

Figure 1 shows the system architecture for low-altitude flying vehicles. The following steps described the working procedure of low-attitude flying vehicle prototype.

1. It starts with defining a map with obstacles, start position and the target.
2. Locations of above are assigned to various nodes, like on a map, following which these nodes are mapped to coordinates on screen.
3. Object starts from a starting position and starts exploring various nodes, using A-Star algorithm.
4. Exploration is represented using green color and optimal path is displayed using yellow color between start and target, avoiding all the obstacles. For the prototype of scenario, Pygame (library of Python) is used.

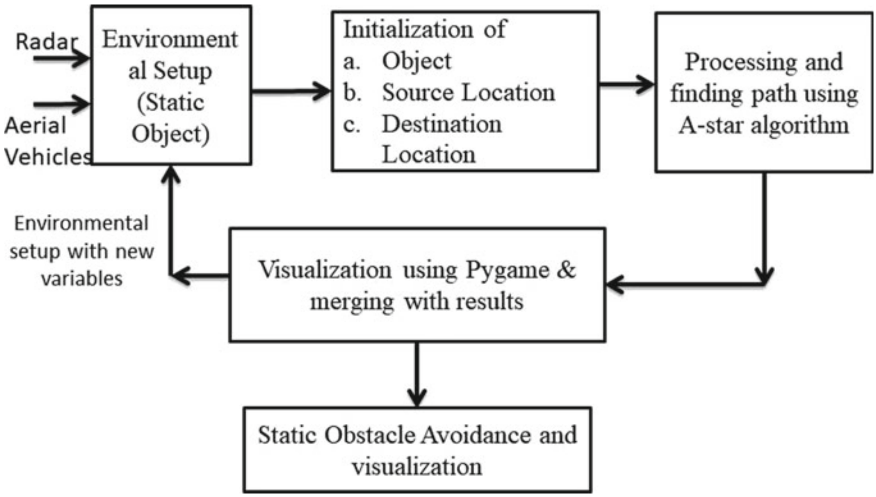


Fig. 1 System architecture for low-altitude flying vehicles

3.2 Proposed System for Obstacle Avoidance for High-Altitude Flying Vehicles

In this object in question is tasked to achieve some goal by reaching a destination. But in such scenarios, obstacles during travel are radars. Radars are of two types static and dynamic, which paint the scenario with dynamic obstacles. Hence during flying at high-altitude dynamic obstacles come in place. Basically, the environment aerial vehicle is going to enter is unknown and obstacles are radars.

The proposed system is distributed into two parts which is implementation of approach and a working prototype displaying obstruction avoidance. Deep Q-Network Reinforcing Learning [2] Agent and MITs Director will be used. Continuous scanning of moving objects and making decisions in its own path and the prototype to display all the above-mentioned functions in real time is done.

The proposed approach is a combination of bubble band technique and fixed sonar method. For high-altitude flying vehicles or the case of moving obstacles former mentioned combination helps in removing disadvantages of each and other and couple the pros of each approach. The object in the scenario has a set boundary like a bubble and is able to be in-motion. A sonar is mounted which is able to perceive obstructions and that bubble around the object defines from how far the perceiving starts.

Figure 2 shows the system architecture for high altitude flying vehicles. The following steps described the working procedure of high attitude flying vehicle prototype.

1. Deep Q-Network (DQN) reinforcement learning agent is set which navigates an object in a simulator to a target waypoint, while avoiding obstacles.

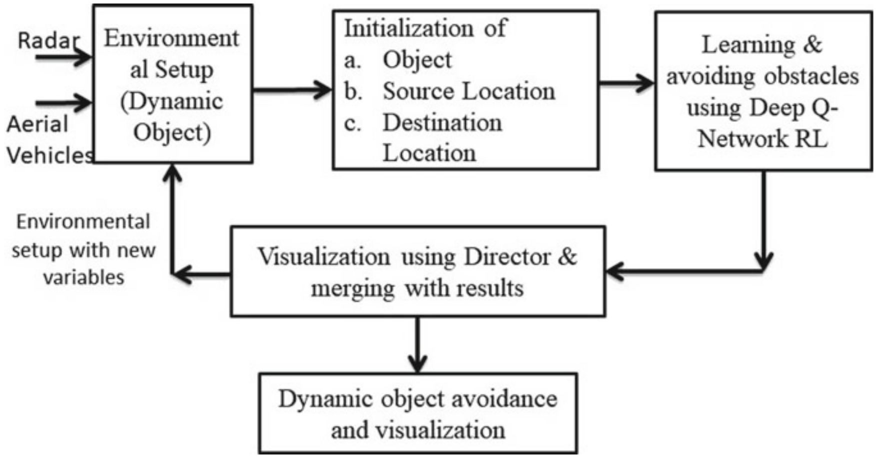


Fig. 2 System architecture for high-altitude flying vehicles

2. For this project assumption a map of the environment a priority is not taken but rather rely on sensor (in this case a LIDAR) to inform control decisions.
3. Software stack is integrated with a visualization tool called Director.
4. Director code is used for prototyping and ray-casting. The simulation environment is written in Python.

4 Results and Discussions

4.1 Test Bed for Implementation

Considering the object model, with a fixed width of the wings. Placement of the object is detailed by (x, y, θ) in which Cartesian position is denoted by x, y and θ represents the orientation of the object. In the test bed, the ability to control derivative of orientation is enabled defining a static speed 'v'.

Considering sensor model, object uses a small arrangement of point lidar. To be specific, numbers of rays (N) are spread across uniformly over field of view as $[\theta_{\min}, \theta_{\max}]$. In the test bed, N=20 and $[\theta_{\min} = -\pi, \theta_{\max} = \pi]$. The entire lasers have a $d_{\max}$ as the maximum range. 'nth' lasers replies x_N, which is represented as distance to the first obstruction discovered or $d_{\max}$ if no obstacle is detected. Now a single laser measurement can be summarized as $X = (x_1, \ldots, x_N)$. Table 2 shows the sensor parameter values.

Choosing of right obstruction setting for my project needed some trial & error. Too thin obstructions were not appropriate as they could be hidden among array of point lidars and in addition slender and long obstacles do not replicate to real life obstructions such as radar. Hence circular obstructions were picked, as they

Table 2 Sensor parameters

Sensor Parameter	Values
No. of points LIDAR	20
Maximum laser range, $d_{\max}$	10 m
Field of view $[\theta_{\min}, \theta_{\max}]$	$[-\pi, \pi]$

Table 3 Obstacle environment parameters

Obstacle environment parameters	Values
Area of square world	250 m^2
Obstacle density	0.18 m^2 (45 Obstacles in 250 m^2)
Circular obstacle radius	1.75 m

removed this error. Choosing the apt density of obstructions and their corresponding size was trial-and-error process. A square border stopped the vehicle from leaving. Table 3 represents obstacle environmental parameters. Braitenberg styled controller is equipped, which requires the squared inversing, each and every of the sensor measurements as features:

$$\varphi B(x_i) = 1/x^2 \tag{1}$$

The whole sum of features is calculated for the left and right, and once again the action is chosen to turn to the direction $\min(n_{\text{left}}, n_{\text{right}})$.

4.2 A-Star Algorithm for Obstacle Avoidance

Implementation starts with defining a map with obstacles, start position and the target. Locations of above are assigned to various nodes, like on a map, following which these nodes are mapped to coordinates on screen. Object starts from a starting position and starts exploring various nodes, using A-Star algorithm. Exploration is represented using green color and optimal path is displayed using yellow color between start and target, avoiding all the obstacles. For all visualization of the process, Pygame (library of Python) is used. Result from obstacle avoidance using A-Star algorithm is a starting position and a target for an object is set (Fig. 3).

Following which obstacles (which can be radars, high mountains and etc.) are determined in path between target and starting position. It can be clearly seen that algorithm has tried all possible nodes, but has rejected those with obstacles or rather a high cost. Then, algorithm starts working and finds the most optimal path to the target, escaping from all the obstacles one by one. In the end, a yellow line is shown to represent the final path, i.e., without colliding into any obstacles and shortest possible distance from start to target, hence obstacle avoidance is achieved.

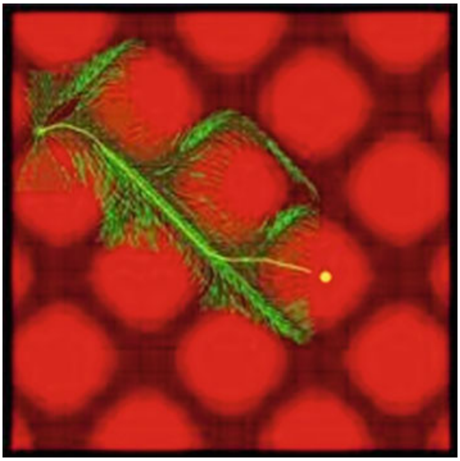


Fig. 3 After all nodes and possible paths are checked, final path is chosen (obstacle free)

4.3 Deep Q-Network Reinforcement Learning Agent for Obstacle Avoidance

Deep Q-Network (DQN) reinforcement learning agent is set which navigates an object in a simulator to a target waypoint, while avoiding obstacles. For this project assumption, a map of the environment a priori is not taken but rather rely on sensor (in this case a lidar) to inform control decisions. Software stack is integrated with a visualization tool called Director. Director code is used for visualization and ray-casting. The simulation environment is written in Python. This implementation has a main class called simulator which combines together several other modules, e.g., moving_object, sensor, controller, net and etc., to construct a working simulation class. Highlighting factor is MITs Director uses visualization toolkit for graphics engine and as a result ray-casts generated are efficient as being in C++. A time step of dt = is used, and scipy.integrate is used to for integrating forward the dynamics. The code includes various python files, which have specific functionality:

1. Moving_Object.py—This python file, makes a moving object, which gets a random start and target location. In this the object is made to move randomly and learn when collide and start again, iteratively
2. Net.py—This python file is responsible for setting up of the deep Q-Network and implement the learning feature for the object
3. Sensor.py—This executes the sensor attached to object to detect obstacles and change movement/path accordingly
4. Simulator.py—This python file, simulates all the visualization into one screen
5. World.py—This python file, sets up the environment which has number of obstacles and their dynamic movement.

Result from using Deep Q-Network (DQN) Reinforcement learning agent for obstacle avoidance is a moving object is seen with rays protruding from the object, which are actually to detect obstacles and learn from. Then, a starting position and target is defined, following which object starts moving through clutter of obstacles. Then, training and learning of object starts, when it collides with an obstacle it learns from it, so that next time it doesn't collide. After a few iterations of learning, object starts detecting obstacle (rays turn red when an obstacle is in course) and it diverts from the course and keeps looking forward for the target (Figs. 4 and 5).

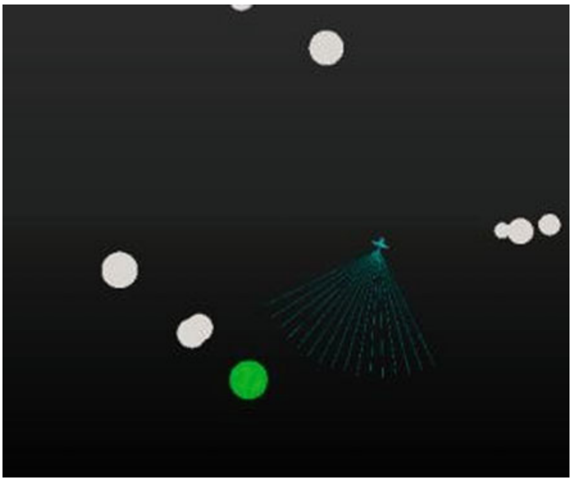


Fig. 4 Object with LIDAR sensor for detecting obstacle

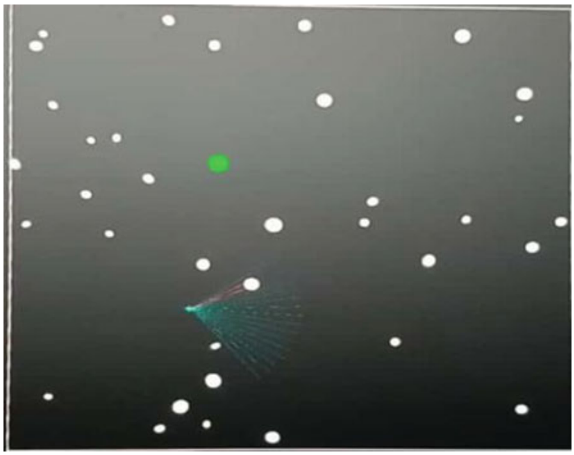


Fig. 5 Rays turn red, when object detects obstacle and changes

5 Conclusion

Various obstacle avoidance algorithms were thoroughly researched and analyzed. The research on obstacle avoidance algorithms with a comparative study of all algorithms, stating merits and demerits of each are cumulated. Comparative study helped in choosing Tangent Bud and Artificial Potential Method algorithms for Low-Altitude Flying Objects (Static Obstacle Avoidance) and Bubble band Technique and Fixed Sonar Method for High Altitude Flying Objects (Dynamic Obstacle Avoidance). Literature study assisted in choosing the appropriate technologies and technical requirements; A-Star algorithm for Low Altitude Flying Objects (Static Obstacle Avoidance) and Deep Q-Network (DQN) Reinforcement learning agent for High Altitude Flying Objects (Dynamic Obstacle Avoidance). Visualization of obstacle avoidance was done in both A-Star and Deep Q-learning using Pygame and Director respectively. Hence successful obstacle avoidance has been achieved. This work can be further extended by implementation of same in real-time environments using real objects.

References

1. Gandhi D, Pinto L, Gupta A (2017) Learning to fly by crashing. In: 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, pp 3948–3955
2. Tai L, Li S, Liu M (2016) A deep-network solution toward model-less obstacle avoidance. In: 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, pp 2759–2764
3. Zsedrovits T, Zarándy Á, Vanek B, Péni T, Bokor J, Roska T (2011)Collision avoidance for UAV using visual detection. In: 2011 IEEE international symposium of circuits and systems (ISCAS). IEEE, pp 2173–2176
4. Bodin WK, Redman J, Thorson DC (2007) U.S. Patent No. 7,228,232. Washington, DC: U.S. Patent and Trademark Office
5. Van Tooren J, Martin HENI, Knoll A, Johannes BECK (2010) U.S. Patent No. 7,737,878. Washington, DC: U.S. Patent and Trademark Office.
6. Buniyamin N, Ngah WW, Sariff N, Mohamad Z (2011) A simple local path planning algorithm for autonomous mobile robots. Int J Syst Appl, Eng Dev 5(2):151–159
7. Sgorbissa A, Zaccaria R (2012) Planning and obstacle avoidance in mobile robotics. Robot Auton Syst 60(4):628–638
8. Khatib O (1986) Real-time obstacle avoidance for manipulators and mobile robots. In: Autonomous robot vehicles. Springer, New York, NY, pp 396–404
9. Dravidam U, Tosunoglu S (1999) A survey on automobile collision avoidance system. In: Florida conference on recent advances in robotics
10. Rostami SMH, Sangaiah AK, Wang J, Liu X (2019) Obstacle avoidance of mobile robots using modified artificial potential field algorithm. EURASIP J Wirel Commun Netw **70**
11. Zhang Q, Chen D, Chen T (2012) An obstacle avoidance method of soccer robot based on evolutionary artificial potential field. Energy Procedia 16(Part C):1792–1798. https://doi.org/10.1016/j.egypro.2012.01.276
12. Babinec A, Duchoň F, Dekan M, Mikulová Z, Jurišica L, Vector Field Histogram* with look-ahead tree extension dependent on time variable environment. Trans Inst Meas Control 40(4):1250–1264

13. Zaheer S, Gulrez T, Ahamed I, Paramabath T From sensor-space to Eigenspace—a novel real-time obstacle avoidance method for mobile robots. Published online: 08 Sep 2019. https://doi.org/10.1080/03772063.2019.1654937
14. Shin S-Y, Kim Y-G, Kang Y-W (2019) Obstacle avoidance drone by deep reinforcement learning and its racing with human pilot. Appl Sci 9(24):5571. https://doi.org/10.3390/app9245571
15. Susnea I, Minzu V, Vasiliu G Simple, real-time obstacle avoidance algorithm for mobile robots. In: Recent advances in computational intelligence, man-machine systems and cybernetics, pp 24–29. ISBN: 978-960-474-144-1
16. Kamon I, Rimon E, Rivlin E, TangentBug: a range-sensor-based navigation algorithm 17(9):934–953. https://doi.org/10.1177/027836499801700903
17. Quiñonez Y, Barrera F, Bugueño I, Bekios-Calfa J (2018) Simulation and path planning for quadcopter obstacle avoidance in indoor environments using the ROS framework. In: Mejia J, Muñoz M, Rocha Á, Quiñonez Y, Calvo-Manzano J (eds) Trends and applications in software engineering. CIMPS 2017. Advances in intelligent systems and computing, vol 688. Springer
18. Al-Qassar A, Najem A (2019) Obstacle avoidance techniques for robot path planning. https://doi.org/10.26367/DJES/VOL.12/NO.1/7

Modified Particle Swarm Optimization in Wireless Sensor Network for Clustering and Routing

R. Vijaya Prabhu, S. Anslam sibi, and R. Priscilla

Abstract Wireless sensor networks are composed of sensor nodes and are being deployed in many applications, where security is deemed critical. In WSN, the data collected at any node is forwarded to any designated sink node either through single or multiple hops. Due to the weaker nature and intrinsic, they are easily brought down by external security attacks. The routing plays a vital role in forwarding data from one node to another. Optimal route discovery in WSN is often problematic due to several factors like drain of energy in the node and so on. The residual energy of nodes located near the sink area has the high probability of getting drained out sooner than nodes in other areas. The proposed research addresses this issue through an energy-efficient clustering technique and trustable routing protocol specifically designed for WSNs that minimize the energy consumed by the nodes at the sink area. In this work, a novel clustering algorithm based on modified particle swarm optimization (MPSO) technique had been proposed to form clusters that engage in selecting the cluster heads at the sink coverage area and account for devising a solution to the energy hole problem. Also an energy trust system (ETS) for WSNs had been formulated in the proposed research for effectively detecting the Sybil attacks. Multi-level detection based on identity and position verification is carried by the proposed ETS. Cluster-based trust-aware secure routing had been performed in the proposed research that successfully detects the available active nodes prior to forwarding any data and establishes alternate routes, if need arises. The main aim of the proposed research is to preserve energy of nodes and improvise the security of data while routing from source to sink nodes.

Keywords WSN · Energy conservation · Energy hole problem · Sybil attacks · Energy trust system · Clustering · Modified particle swarm optimization algorithm

R. Vijaya Prabhu (✉) · S. Anslam sibi · R. Priscilla
St.Joseph's Institute of Technology, Chennai, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_30

1 Introduction

The wireless sensor network is advancing in recent years. The sensor devices are mostly used as input devices to detect physical or environment conditions. In this, node to node data transmission is done by calculating base station distance using various algorithms.

A wireless sensor network is composed of numerous tiny sensors that monitor and track information about the environment in which they are deployed. The aggregated information is transmitted through wireless links to any designated sink node. Data is relayed through several intermediate nodes across a gateway towards the destination node like a wireless Ethernet.

2 Literature Survey

2.1 Routing Techniques

Souihli et al. [1] had proposed a novel a load balancing mechanism that could be applied in MANET using shortest path routing protocols. Distribution of maximum load is observed at the centre of the network, and accordingly, the proposed load balancing schemes drive the traffic or the observed load from the centre of the network. Based on the characteristics of routing protocol, the central node is designated as proactive or reactive. In reactive types, the nodes are characterized according to their centrality that is based on the size of their routing tables. In proactive types, the centrality of a node is fixed based on its multi-point relay selector list. The proposed load balancing schemes provide efficient load distribution, end-to-end delay and packet delivery fraction. The proactive feature in the proposed schemes effectively identifies the nodes that are congested and overloaded and relieves them. The authors had not taken into account of the performance of their schemes under non-uniform node distributions.

Protecting the privacy of the sink nodes in WSN is highly challenging. The sink nodes are highly vulnerable as they can easily fall prey to adversaries who can eavesdrop on the packets to identify its destination. The sink node privacy issue had been addressed by Long et al. [2] who had proposed a ring-based routing (RBR) technique. In the RBR technique, the authors had introduced the concept of multiple rings and multiple routing lines. Data is not directly transmitted to the sink node instead to the nearest routing ring. Data transmission takes place through routing rings that are relayed by further rings through routing lines till it reaches the destination. The count of anonymous sink nodes is kept equal to the number of network nodes. The proposed RBR technique is highly scalable and increases the network lifetime and accounts for high energy efficiency.

Lee et al. [3] had analysed the adhoc networks and had proposed the split multipath routing (SMR) protocol. The proposed SMR protocol functions as on on-demand protocol that creates maximum number of disjoint routes. Each session of data transmission comprises of two routes, namely the shortest delay route and maximally disjoint shortest delay route. These maximum disjoint routes are designed with the idea of avoiding routes that may be congested to efficiently utilize the network resources. In adhoc networks, the presence of multiple paths helps in overcoming disconnected routes where source can choose any available route without the necessity of route recovery. The proposed protocol introduces end–to-end delay whenever routes are disconnected and consumes more energy during such occurrences.

Hu et al. [4] had proposed a technique that brings a balance between the quality of detecting the target and the lifetime of WSN. The authors had proposed an intelligent adjustable sensing frequency for mobile target detection based on the monitor quality optimization. In the proposed technique, the target is monitored using two schemes, namely target detection with sensing frequency K (TDSFK) and target detection with adjustable sensing frequency (TDASF). The former technique senses the frequency increase from 1 to K, and the latter technique adjusts the sensing frequency on nodes having residual energy. The two techniques are adaptable for networks containing static sinks and are not suited for networks with mobile sinks.

Zhang et al. [5] had proposed a dynamic sensing technique that addresses the routing issues for maximizing the utilization of rechargeable sensor networks. The authors had proposed a balanced efficient energy allocation scheme (BEAS) that manages the energy utilization in sensor nodes and makes the entire process as smooth as possible. In the proposed technique, the optimal sensing rate and routing control had been accounted through a distributed sensing rate and routing control (DSR2C) algorithm. The cost expended in managing the energy allocation and controlling the topology is very high in the proposed technique.

2.2 Cluster Formation Techniques

Li et al. [6] had analysed the WSN and proposed that resource efficiency and dependability of a trust system are required for securing data. Present trust systems are generating too much overhead and possess low dependability, hence not suitable for enforcing high security. The authors had proposed a lightweight and dependable trust system (LDTS) for WSNs that employ clustering algorithms. Energy preservation is brought in using a lightweight trust decision-making technique that operates based on identities of nodes in the WSN clusters. The effect of malicious nodes is negated, and system efficiency is improved by cancelling of feedback between the individual cluster members or cluster heads. Since huge data is forwarded by the cluster heads, a dependability-enhanced trust evaluating approach had been implemented at the cluster heads for cooperation purposes. It is huge improvement when compared to traditional techniques that fix weights subjectively. When compared with existing trust systems, the proposed LDTS generates less overhead and consumes less memory. The dependability-enhanced trust evaluating approach effectively detects and prevents malicious nodes and faulty cluster heads.

Naruephiphat et al. [7] had proposed an innovative clustering algorithm named limiting member node clustering (LmC) algorithm that limits the node members included under each cluster head based on a threshold value. A new cost function is used for selecting the cluster head by taking into account of battery level, energy consumption and its distance to the base station. The transmission range of base stations had been considered for improving the performance of clusters in WSN. The proposed approach provides efficient delivery ratio, increased network lifetime, low latency and less energy consumption when compared with other existing techniques.

2.3 Trust-Based Routing Techniques

Zhou et al. [8] had analysed the WSN based on trust and had proposed a trust behaviour collection technique that behaves like a watchdog. The watchdog transmits the data in multiple hops based on the location of the sensor nodes and the trustworthiness of the target nodes. The proposed technique calculates the trust of nodes and minimizes the cost of energy expended by the watchdog. The proposed technique reduces the life span of the network due to high energy consumption. Also consistent security is not always afforded.

Zhan et al. [9] had designed and implemented a robust trust-aware routing framework for WSNs named trust-aware routing framework (TARF) for securing the multi-hop routing in the dynamic WSN from adversaries. In the proposed technique, each node tracks the trustworthiness of its neighbours and accordingly selects any for transmitting its data. The proposed technique focuses on trustworthiness and energy efficiency. Protection is provided against the replay attacks. TARF is scalable and highly resilient, and the same is proved by comparing it with other existing trust techniques for large-scale WSNs.

Yu et al. [10] had reviewed the trust mechanisms and attacks taking place in the WSNs. The authors had categorized all attacks that take place based on trust in WSN. An intelligent behaviour attack model is implemented to identify inconsistencies in behaviour in the content domain. Various methodologies of trusting techniques had been analysed to provide proper emphasis on trust schemes in WSNs. The proposed approach makes use of Bayesian trust model, entropy trust model, fuzzy trust model, game theory trust model and subjective logic trust model. Behaviour attacks at MAC layer are successfully identified by the proposed approach. Also, efficient protection is not afforded against other security attacks like black hole attack, worm whole attack, selective forwarding and hello flood attack.

Danyang in et al. [11] had studied the critical attacks that could be launched on data transmission in WSNs due to limited energy resource and improper deployment of the sensor nodes. The authors had proposed a trust sensing-based secure routing mechanism (TSSRM) that is lightweight in nature and possessed the ability to thwart several simultaneous attacks. Performance analysis and evaluation of the proposed TSSRM had shown that it affords effective security against a host of attacks. Secure routing algorithms based on trust sensing routing protocols had been proposed for improving security and overcome common network attacks in the WSN environment. The proposed technique improves the reliability of data transmission in the network when compared to other existing trust-based techniques.

Maarouf et al. [12] had analysed the constrained nature and the prevailing insecurity in WSN and had proposed a trust-aware routing for wireless sensor networks (WSNs) named efficient monitoring procedure in a reputation system (EMPIRE). Reputation-based solutions had been suggested to enforce trust-aware routing where a node needs to continuously monitor its environment for detecting any misbehaviour events that are taking place. Given the resource scarcity present in the WSN, this seems to be a costly affair. The proposed EMPIRE technique is a probabilistic, distributed monitoring methodology that reduces the monitoring activities that are performed by any node and at the same time maintains the attack detecting capability to a satisfactory level. The authors had carried out the simulation using the Monte Carlo simulation technique.

WSNs are generally deployed in security-critical applications, and due to the inherent restrictions in sensor nodes and deployment environment, security can be easily compromised. Present techniques do not efficiently evaluate the trust among nodes and fail to establish energy-efficient routing to enforce security in the WSNs. In order to overcome these issues, energy-efficient clustering and trust-based secure routing scheme are necessary.

3 System Architecture

The proposed approach mainly focuses on energy-efficient clustering and trust-based Sybil attack detection scheme with secure routing in WSNs. The initial process of the secure routing is to form clusters based on location of the nodes and energy. A modified PSO-based clustering algorithm is implemented for cluster formation around the sink coverage area, a process that is initiated by the sink node. The trust and energy of the node calculation as well as Sybil node detection are done by using energy trust system (ETS). ETS consists of two levels of estimation such as CH level estimation and BS level estimation. In CH level estimation, CH receives a control packet from a CM and initiates the verification process. If the CM is legal energy and trust values are estimated based on successful, unsuccessful interaction and timing window, then accepts the packet. Else the CM is illegal that CM is declared as a Sybil node and then deletes the received packet. In BS level estimation, BS receives a packet from a CH and initiates the verification process. If the CH is legal energy and trust values are estimated based on successful, unsuccessful interaction and timing window, then accepts the packet. Else the CH is illegal, that CH is declared as a Sybil node and then deletes the received packet. BS sends the feedback packet via the trusted route to the source CM. Source node transfers the data packet via trusted route to the CH, and CH transfers the data to BS. If the data packet reaches the BS, the BS sends the feedback to its source node. The system architecture had been represented in Fig. 1.

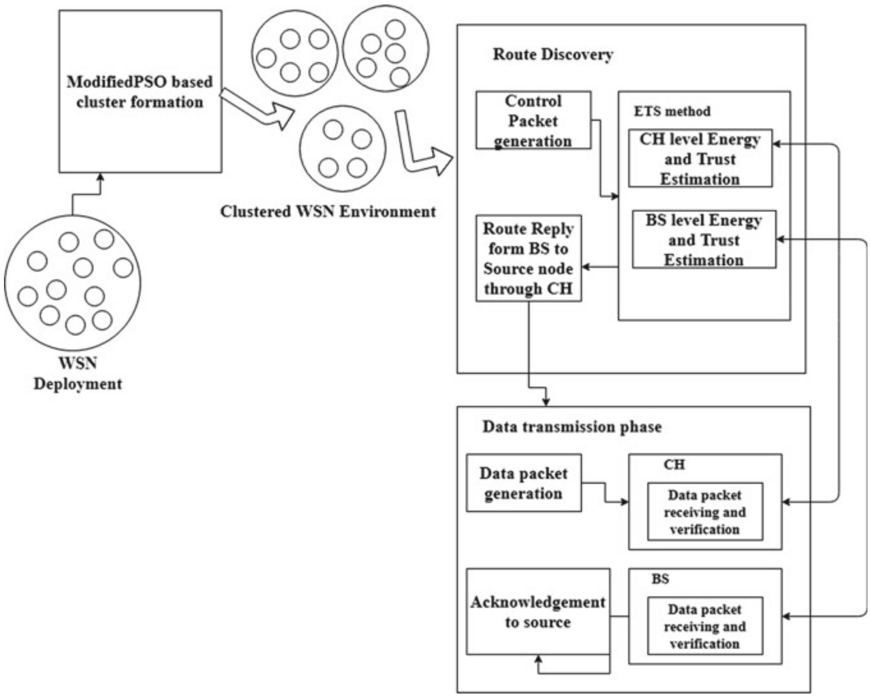


Fig. 1 Energy-efficient clustering and trustable routing

3.1 Cluster Formation

Sensors that are in the vicinity of sink are mostly utilized for delivering data to the sink. Since they are constantly engaged in data transfers, their energy level gets drained up quickly than other nodes. This leads to the formation of hotspots at the sink surrounding area resulting in network isolation. This could be termed as hotspot or energy hole issue. The proposed modified PSO-based clustering algorithm improves the lifetime of the network by reducing the formation of hotspots.

In the modified PSO-based clustering, cluster formation is initiated by the sink node around its coverage area. An info-req-msg is transmitted by the sink to all nodes present in the coverage area. Whenever a message is received by the nodes around sink, they transmit the received message along with the info-reply-msg that contains the sensor-id, position, velocity and current residual energy. The sink node receives the information and updates itself.

Consider a problem space composed of N number of particles in the sink coverage area. The fitness function formulated deduces the particle that has the best value in the swarm and also identifies the best position of each particle periodically. The fitness function for the modified PSO-based clustering is calculated for each particle by equation given below:

$$f_p = a_1x_1 + a_2x_2 + a_3x_3$$

where a_1 and a_2 are weighing parameters that take value between 0 and 1 and $a_3 = 1 - a_1 - a_2$.

$$x_1 = \sum_{i=1}^{n} \frac{d(\text{current particle member})_i}{c_n}$$

$$x_1 = \sum_{i=1}^{n} \frac{E(\text{member})_i}{E(\text{current particle})}$$

3.2 Energy Trust System

The proposed work focuses on eliminating Sybil attacks in WSN by developing a lightweight trust-based system that has the capability to carry out multi-level detection mechanism in clustered WSNs. Establishing trust detection system in a clustered topology inside WSN reduces the communication overhead and energy consumed and improves the network scalability and throughput factors. In a heterogeneous WSN with N sensors, there are three entities, namely a base station (BS), cluster heads (CHs) and cluster members (CMs).

In the proposed approach, each sensor node has a unique identity, and positions of every sensor including the BS and CHs are recorded. The sensor nodes are identified by the monitoring node (either BS or CH) based on the information they send that contains their ID, position and energy level along with the sensed data. There are two levels of detection in the proposed ETS, namely cluster head level detection and base station level detection.

3.2.1 Cluster Head Level Detection

Detection is initially carried out at the cluster head (CH) level. Clustering of sensor nodes provides effective topology control. Responsibility of nodes within a cluster is assigned to the respective CH. Whenever a CH receives a message from one of its sensors, it applies the ETS scheme, using verification and trust mechanisms to check the source of that message. It then forwards the data to the BS by performing the following:

1. **Checking**
   The cluster head verifies the ID and position of the sender. Once it confirms that the sender of the message is from its own cluster region, it initiates the next

step, namely the trust calculation, else it drops the received message. By doing so, Sybil attacks with forged ID and position can be successfully thwarted.

2. **Trust calculation**

In this step, the worst-case scenario happening through Sybil attack where the attacker succeeds in impersonating the ID and position of legitimate node is checked. The trust factor $T_{\text{ch},x}$ is deduced as shown in the below equation:

$$T_{\text{ch},x} = -\sum \frac{S_{\text{ch}x}}{S_{\text{ch},x} + U_{\text{ch},x}} - \sum \frac{1}{\sqrt{U_{\text{ch},x}}}$$

where $S_{\text{ch},x}$ indicates the total number of successful interactions between CH with node x, and $U_{\text{ch},x}$ indicates the total number of unsuccessful interactions between CH with node x. If the deduced value is found to be more than 0.3, then the node x is regarded as a trusted node. On the contrary, if the deduced trust value is found to be less than or equal to 0.3, node x is regarded as a Sybil node.

3.2.2 Base Station Level Detection

Second level of detection is carried out at the base station level. In the proposed system, the base station BS is designated as the central command authority. Due to the formation of clusters in the network, the amount of data transmitted to the BS gets reduced. The BS tracks the working of CHs in the network. Whenever a BS receives a message from any CH, it applies the proposed ETS methodology as follows:

1. **Checking**

Upon receiving a message, the BS verifies the ID and position of the sender. When it ascertains that the sender is an authentic CH, it begins the trust calculation step, else the BS drops the message.

2. **Trust Calculation**

The trust value $R_{bs,\text{ch}}$ is deduced by the BS as follows:

$$R_{bs,\text{ch}} = -\sum \frac{S_{bs,\text{ch}}}{S_{bs,\text{ch}} + U_{bs,\text{ch}}} - \sum \frac{1}{\sqrt{U_{bs,\text{ch}}}}$$

where $S_{bs,\text{ch}}$ indicates the total number of successful interactions between BS with cluster head ch, and $U_{bs,\text{ch}}$ indicates the total number of unsuccessful interactions between BS with cluster head ch. If the deduced trust value is found to be more than 0.3, then the cluster head ch is regarded as a trusted cluster node. On the contrary, if the deduced trust value is found to be less than or equal to 0.3, the cluster head ch is regarded as a Sybil node.

3.3 Data Transmission Phase

In the proposed algorithm, both single and multi-path routing are implemented for selecting a near optimal route. A near optimal route with minimum cost, less hop count and maximum residual energy is selected. When an activity happens, the source sensor node transmits data in the shortest path in its corresponding time slice. During its next subsequent time slice, it transmits the next data packet to an alternate neighbour present in its coverage area, regarding it as the best-case multi-path route to traverse towards the sink.

4 Implementation and Results

4.1 Cluster Formation Phase

The sensor nodes are simulated with the help of Network Simulator2. The modified PSO-based clustering algorithm implemented at the sink coverage area by the sink node to form clusters. Once clusters are created, the sink transmits the info-req-msg to the sensors present in its coverage area. Upon receiving the info-req-msg, the sensor nodes send their respective information through the info-reply-msg. This info-reply-msg contains their ID, position, velocity and residual energy. The sink node maintains and updates the received information. The modified PSO-algorithm employed in this work is shown in Algorithm 4.1.

Algorithm 4.1 Modified PSO-based clustering algorithm

1: Procedure modified PSO-based clustering ()
2: Begin
3: Population of N particles are initialized with random positions and velocities
4: If target fitness or maximum iteration is not attained
5: **for** each particle p in N **do**
6: Calculatefitnessvalue(f_p) of each particle
7: **if** f_p>f (pBest) **then**
8: pBest =f_p
9: **endif**
10: **end for**
11: gBest = max(pBestinP)
12: **for** each particle p in N **do**
13: Calculatevelocity
14: Calculateposition
15: **end for**
16: End while
17: End Modified PSO-based clustering

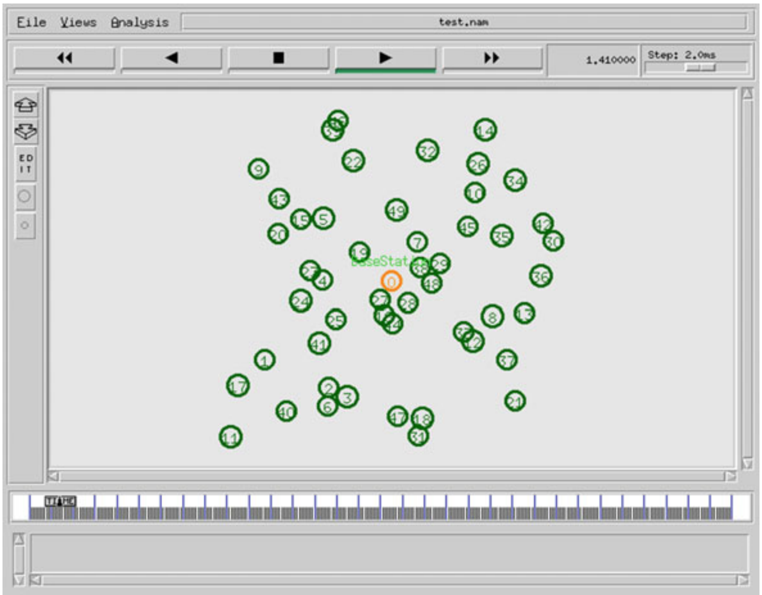


Fig. 2 Network deployment

Figure 2 illustrates the deployment of nodes and network formation. The output of the cluster formation in the network is shown in Fig. 3.

4.2 Energy Trustsystem

Every sensor nodes in the network including the cluster head CHs and base station BS are assigned unique ID and position values. The sensor nodes are identified by the monitoring node which could be either the CH or BS, based on the information they transmit that includes their ID, position, energy level and the actual sensed data. The proposed ETS has two levels of detection, namely cluster head level detection and base station level detection.

Upon receiving a message from any of its members, the cluster CH applies the ETS technique to ascertain the authenticity of the sender using the verification and trust mechanisms. The cluster head level detection algorithm employed in this work is shown in Algorithm 4.2 (Figs. 4 and 5).

Algorithm 4.2 Cluster head level detection algorithm

1: Input: Check whether the message was sent from a legitimate or sybil node.
2: Output: Identification of a sybil attack in WSN by applying the ETS.
3: Initialize the sensor nodes
4: CH receives a message from a CM

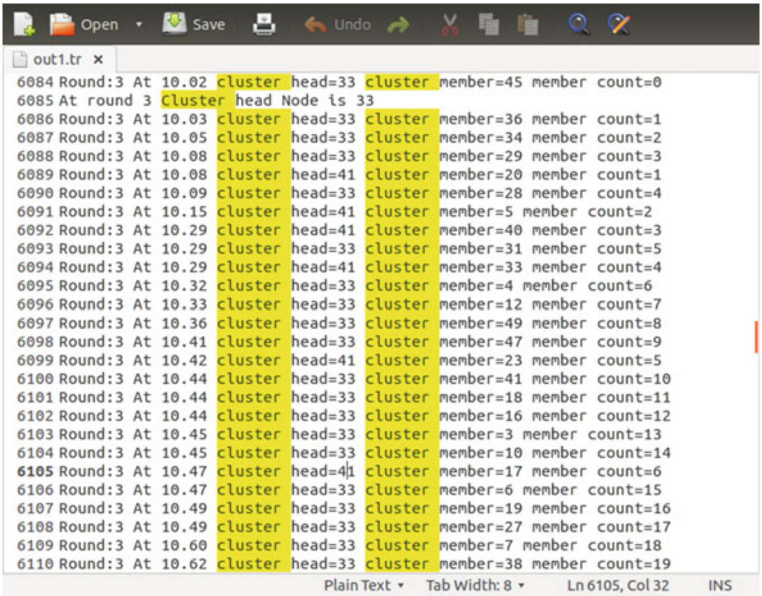


Fig. 3 Cluster formation

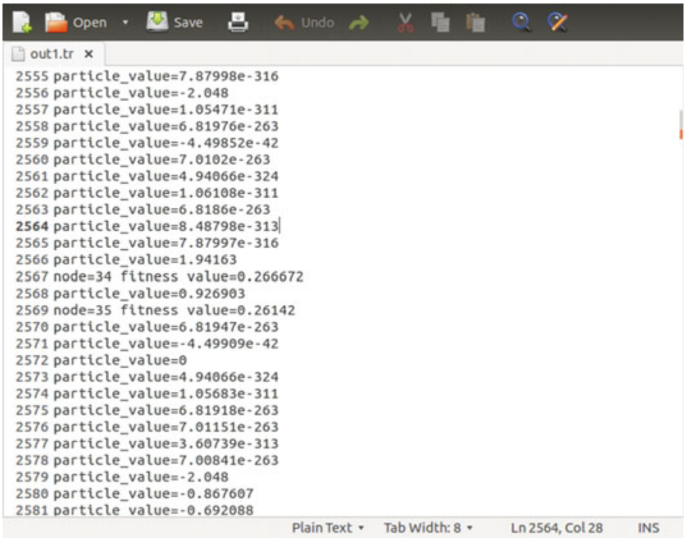


Fig. 4 Fitness value calculation

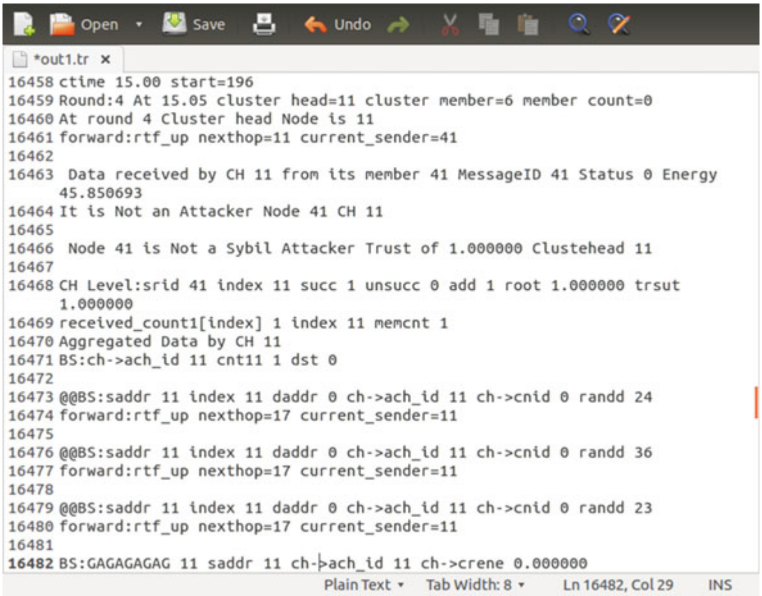


Fig. 5 Cluster head level verification

5: CH initializes verification (ID,position,energy) of the CM
6: **if** legal ID and position **then**
7: Calculate a trust

$$R_{bs,\mathrm{ch}} = -\sum \frac{S_{bs,\mathrm{ch}}}{S_{bs,\mathrm{ch}} + U_{bs,\mathrm{ch}}} - \sum \frac{1}{\sqrt{U_{bs,\mathrm{ch}}}}$$

8: **if** T >0.3**then**
9: Trusted node and send the message toBS
10: **endif**
11: **else**
12: Delete received message from CM(sybilnode)
13: **endif**

In a similar manner, the BS, upon receiving a message from any of the cluster heads CHs, applies the ETS technique to ascertain the authenticity of the sender using the verification and trust mechanisms. The base station level detection algorithm employed in this work is shown in Algorithm 4.3 (Figs. 6, 7 and 8).

Algorithm 4.3 Base station level detection algorithm

1: Input: Check whether the message was sent from a legitimate or sybil node.
2: Output: Identification of a sybil attack in WSN by applying the ETS.
3: Initialize the sensor nodes
4: BS receives a message from a CH

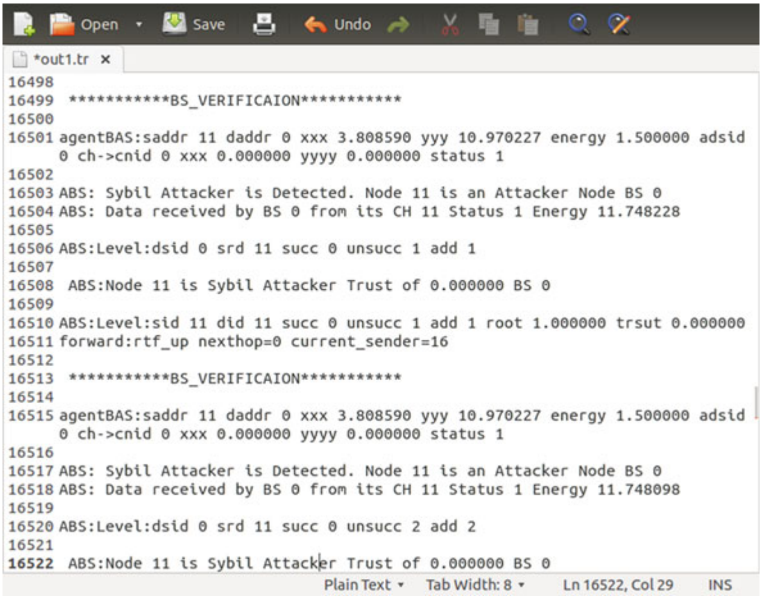


Fig. 6 Base station level verification

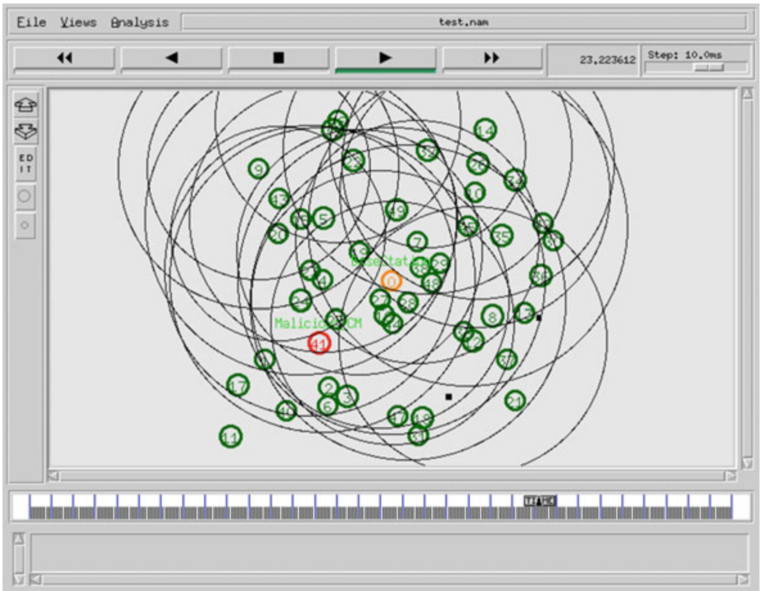


Fig. 7 Sybil attacker node detection

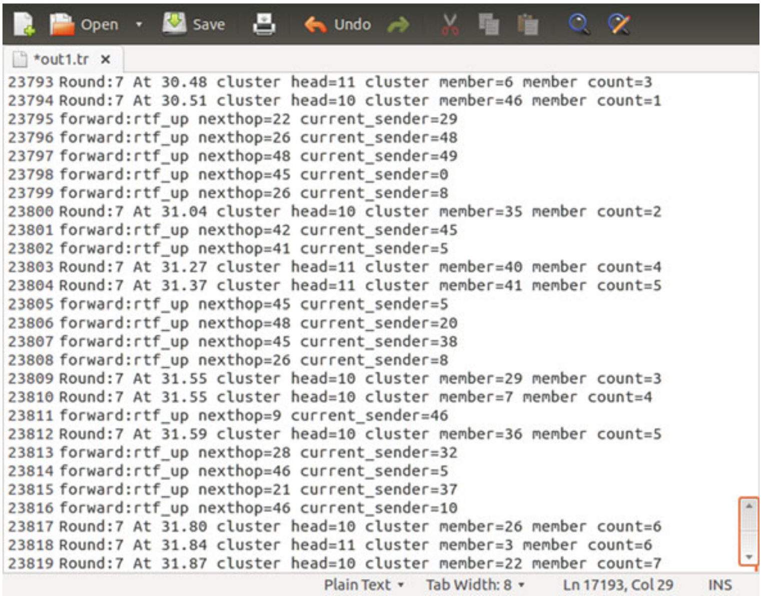


Fig. 8 Data transmission to BS

5: BS initializes verification (ID,position,energy) of the CH
6: **if** legal ID and position **then**
7: Calculate a trust

$$R_{bs,ch} = -\sum \frac{s_{bs.ch}}{s_{bs,ch} + U_{bs,ch}} - \sum \frac{1}{\sqrt{U_{bs,ch}}}$$

8: **if** T >0.3**then**
9: Trusted node and accepts themessage
10: **endif**
11: **else**
12: Delete received message from Ch(sybilnode)
13: **end**

5 Conclusion

In this paper, data is transferred over multiple nodes in a secured manner. The clustering algorithm is implemented to form cluster nodes. ETS maintains unique ID for each node, and fitness level is calculated using cluster level head detection algorithm. Sybil attacker node is verified before data transmission to BS. Base station level detection reduces attackers. Algorithm like modified PSO cluster algorithm, cluster level head detection and BS level algorithm are implemented to reduce the attackers' entry during data transmission from node to node.

References

1. Souihli O, Frikha M, Hamouda M (2009) Load-balancing in MANET shortest-path routing protocols. Ad Hoc Netw 7:431–442
2. Long J, Dong M, Ota K, Liu A (2014) Achieving source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks. IEEE Trans 2
3. Lee SJ, Gerla M (2001) Split multipath routing with maximally disjoint paths in ad hoc networks. Split Multipath
4. Hu Y, Dong M, Ota K, Liu A, Guo M (2016) IEEE Syst J 10:1160–1171
5. Zhang Y, He S, Chen J (2016) Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks. ACM Trans Netw 24(3), 1632–1646
6. Li X, Zhou F, Du J (2013) LDTS: a lightweight and dependable trust system for clustered wireless sensor networks. IEEE Trans Inf Forensics Secur 8(6)
7. Maarouf I, Baroudi U, Naseer AR (2009) Efficient monitoring approach for reputation system-based trust-aware routing in wireless sensor networks. IET Commun 3(5):846–858
8. Zhou P, Jiang S, Aravazhi Irissappane A, Zhang J(2015) Toward energy-efficient trust system through watchdog optimization for WSNs. IEEE Trans Inf Forensics Secur 10(3):613–625
9. Zhan G, Shi W, Deng J (2010) TARF: a trust-aware routing framework for wireless sensor networks. In: European conference on wireless sensor networks wireless sensor networks pp 65–80
10. Yu Y, Li K, Zhou W (2012) Trust mechanisms in wireless sensor networks: attack analysis and countermeasures. J Netw Comput Appl 35(3):867–880
11. Qin D, Yang S, Jia S, Zhang Y, Ma J, Ding Q (2016) Research on trust sensing based secure routing mechanism for wireless sensor network. IEEE Access
12. Naruephiphat W, Charnsripinyo C (2010) An energy-aware clustering technique for wireless sensor networks. Sustain Wirel Sens Netw

A Hyperspectral Image Classification Method-Based Auxiliary Generative Adversarial Networks with Probabilistic Graph Model

Tvisha Trivedi, S. Geetha, and P. Punithavathi

Abstract Hyperspectral image classification takes up a significant role in domains like mineral industry, agriculture and military purposes. Nevertheless, due to the availability of few labeled samples of hyperspectral images, it is fairly challenging while performing high-dimensional hyperspectral image classification. Presently, generative adversarial networks (GANs) have been employed widely for creating more synthetic yet realistic image samples. Still, getting high-quality image samples devoid of uncontrolled divergences and unwanted noise is a major challenge. To cope up with such challenges and to generate high-quality hyperspectral image samples, an auxiliary generative adversarial network with probabilistic graph (AGAN-PG) is proposed in this work. Specifically, the model has (1) a spatial-spectral generator unit that exploits the distinguishing spectral-spatial characteristics of the hyperspectral image data and (2) a discriminator unit that identifies the area categories of hyperspectral image cubes. Further, to make advantage of the relatively large amount of unlabeled data available, a conditional random field that refines the preliminary classification results generated by GANs is proposed. Eventually, the model increases the number and quality of training image samples generated, wherein the impact of overfitting is also avoided. Experimental results obtained using two well-known hyperspectral datasets—Pavia University and Indian Pines—demonstrate that the proposed framework AGAN-PG achieved promising classification accuracy even with a small number of initial labeled hyperspectral image samples for training.

T. Trivedi · S. Geetha (✉) · P. Punithavathi
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India
e-mail: geetha.s@vit.ac.in

T. Trivedi
e-mail: tvisha.trivedi2019@vitstudent.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_31

1 Introduction

With the rapid advancement of remote sensing technology, hyperspectral imaging sensors capture images of high spatial resolution even with hundreds of spectral bands, like those on the recently launched satellites, namely Gaofen-5 and Zhuhai. The imaging spectrometer thus captures images in the approximately continuous spectrum in the range of the visible to infrared wavelength, thus, the acquired hyperspectral images (HSIs) comprises of hundreds of spectral bands over the same perceived scene, which have brought challenges and opportunities for the subsequent image processing and analysis tasks [1]. From the contiguous and narrow spatial-spectral bands, the land cover targets can be identified at high accuracy. Hence, the HSI have been extensively employed for mineral exploration, crop monitoring, military surveillance and urban planning. The primitive task involved for such applications is the HSI classification which aims to assign each pixel in the HSI to one particular category and is the most pulsating field in hyperspectral remote sensing field [2]. However, the imbalance prevailing between the high-dimensional nonlinear characteristics in spatial-spectral domain and the inadequate annotated samples poses to be the toughest problem in HSI classification.

2 Literature Survey

The huge dimensionality of the HSI data makes it challenging to find demonstrative features that discriminates one class from the other classes. A common adversity in HSI classification task is that the annotated (labeled) samples are limited and inadequate, which leads to severe overfitting problem when we use deep learning models for classification. Generative adversarial network (GAN) is a framework which encompasses generative models. These generative models employ backpropagation optimization technique. It also comprises a generator unit to generate synthetic samples; and a discriminator unit to determine whether input samples are from real class or synthetic distribution class [3].

However, the GANs employ deep convolution neural network (CNN) as the discriminator unit during the HSI classification process. There is a heavy loss in spatial relationships among features while using such approaches. Recently proposed CapsNet [4] encoded the complex characteristics of the images and coded their properties as vector outputs, they still use length of the vector to denote the possibility of the entity existence [5].

The existing GAN-based HIS classification models are incapable of using the spatial and spectral features for joint classification [6, 7]. Probabilistic graphical models have been broadly utilized in post-processing image segmentations and in investigating the huge amount of un-annotated image to improve the classification performance [8]. For example, the conditional random fields (CRF) which operate as a higher-order term are capable of considering more complex relationships among different spectral bands [9]. Hence, the use of probabilistic graphical models is a functional way to consider the un-annotated image samples for HSI classification task, since this method needs only an initialized pixel-wise prediction and raw unlabeled HIS image data.

Thus, this paper proposes a novel HSI classification technique which is constructed on auxiliary generative adversarial network with probability graph (AGAN-PG). The auxiliary GANs have been combined with conditional random fields—probabilistic graphs to effectively use the spatial-spectral features and huge unlabeled image datasets of HSIs.

The main contributions of the paper are summarized as follows:

- The proposed method uses auxiliary GAN framework which employs 3D-CNN as the generator unit and 3D capsule network (3D-CapsNet) as the discriminator unit for the classification of HIS. It performs better than the existing classification methodologies, for instance, 1D-CNN; even when the availability of the training samples is limited.
- The GAN framework is comprised of conditional random fields (CRF). These CRF are capable of considering more complex relationships among different spectral bands. Hence, they can operate as a higher-order term.
- The proposed method can train the discriminator by operating on the generated images and limited annotated samples simultaneously. The model generalization ability and classification performance is thereby improved effectively. Two widely used HSI datasets—Pavia University and Indian Pines have been deployed in the experiments to test the proposed model.

The rest of this paper is organized as follows. Section 3 discusses the proposed AGAN-PG in detail. Section 4 presents the experimental results which have been compared to the two popular existing approaches like one-dimensional CNN (1D-CNN) and adaptive CNN (Ada-CNN). Finally, Sect. 5 presents the conclusion and future work.

3 Proposed Method

The proposed method conjoins AGAN and 3D-CapsNet along with the PG model, which has been largely employed in the feature learning process in the recent times.

3.1 Auxiliary Generative Adversarial Network Model

If the distribution of the real training images is $p(x)$, then the distribution of the input noise η becomes $p(\eta)$. The input noise 'η' is mapped into a new space $G(\eta)$ by the generator unit $G(x)$. The probability of the difference of an input image sample 'x' from the real training image, say $G(\eta)$, is measured by the discriminator unit, $D(x)$.

During the adversarial training process, $\ln(G(\eta))$ is maximized so as to consign correct class label to the corresponding synthetic image samples. The G is trained simultaneously to minimize the value of ln(1-D(G(x))). Then, eventually the value of the function V(D, G) of is given by Eq. (1) as below:

$$\min(G)\max(D)V(D,G) = E_{x\sim p(x)}[\ln(D(X))] + E_{\eta\sim p_\eta(\eta)}[\ln(1 - D(G(\eta)))] \tag{1}$$

where E is the expected value of probability's empirical estimate.

Assume that a deep network has been designed in such a way that its final layer contains 'K' capsules corresponding to the number of classes, i.e., 'K', it can classify during classification. Hence, the total length of each capsule is the probability of the image belonging to the specific class 'K'. The intersection of generated synthetic and true image samples is considerably small during training phase, and results in uninformative gradients in top of the low dimensional variations in the base GAN. Actually, this is a variant of the deep convolutional GAN [10]. Only difference is that it uses 3D-CapsNet as the discriminator unit in the place of CNN.

3.2 HSI Spatial-Spectral Classification Architecture

The proposed approach employs AGAN-PG for spatial-spectral classification of HIS samples, as shown in Fig. 1. Extended multi-attribute profile (EMAP) is employed to blend the spatial and the spectral features, initially. This step reserves spatial and spectral information as well as lowers the computational complexity considerably so as to stabilize the GAN training process [11]. The generator generates synthetic samples by mapping the input noise together with the auxiliary class labels. The dimensions of the synthetic samples remain same as that of the HSI patches. The discriminator provides class labels to the input sample images based on the input real training images and the generated synthetic images.

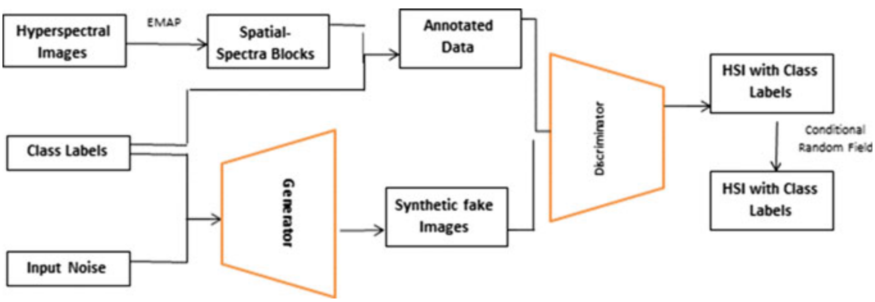


Fig. 1 Architecture of the proposed AGAN-PG model for HSI classification

The proposed approach uses the 3D-CNN as the generator unit. It takes noise 'η' and the respective class 'c' and generates the synthetic image samples with suitable class labels. A simple linear matrix multiplication process forms the first layer of the generator unit and the subsequent 4 fractionally staged convolution operations later map this high-dimensional notation into 3D cube image patches. Batch normalization procedure is applied to stabilize the learning phase which normalizes the input of both the units ensuring a zero mean value with unit variance. Rectified linear unit, i.e., 'ReLU', activation is applied in all but the last layer to facilitate the model to learn fast. The last layer uses 'Tan H' activation function to achieve the optimal classification performance.

The discriminator unit of the proposed approach is a 3D-CapsNet framework. It validates the input with corresponding class labels based on both annotated real training dataset and generated synthetic images samples. The discriminator unit is comprised of a convolutional layer, primary capsule layer and secondary capsule layer. The data flows between the primary capsule layer and the secondary capsule layers. The secondary capsule layer is comprised of primary critic layer and secondary critic layer which are connected to the primary capsule layer. The former provides a score to the input sample indicating it as real or synthetic, wherein the latter classifies the spatial-spectral image sample patch into its respective class label. Further, the generator unit attempts to play away the discriminator unit by learning discriminative features implicitly and thereby generating synthesized fake images.

4 Experimental Results and Discussion

4.1 Performance Metrics

The main accuracy evaluation metrics as listed below are calculated to evaluate the classification performance of the proposed AGAN-PG.

- Overall Accuracy (OA)—a ratio of the number of pixels classified correctly in the test image to the total number of image samples in the entire test image dataset as shown in Eq. (2).
- Average Accuracy (AA)—average classification accuracy of every category as in Eq. (3).
- Kappa Coefficient (κ)—evaluates the classification accuracy using the consistency between the real markers and the obtained classification results, as shown in Eq. (4).

The Kappa score above 80% means that the classification images and real information of objects is highly consistent, in turn denotes that the detection accuracy is considerably high. A score in the range of 61–80% indicates good consistency, in the range of 40–60% reveals medium consistency and below 40% denotes rather a poor consistency.

$$\mathrm{OA} = \frac{\sum_{i=1}^{c} n_{ii}}{\sum_{i,j=1}^{c} n_{ij}} frac \sum_{i=1}^{c} n_{ii} \sum_{i,j=1}^{c} n_{ij} \tag{2}$$

$$\mathrm{AA} = \frac{\sum_{i=1}^{c} A_i}{c} frac \sum_{i=1}^{c} A_i c \tag{3}$$

$$\mathrm{Kappa} = \frac{\mathrm{OA} - P_e}{1 - P_e} frac \mathrm{OA} - P_e 1 - P_e \tag{4}$$

The value of κ is calculated using Eqs. (5)–(7)

$$P_e = \frac{1}{N^2} \sum_{i=1}^{c} a_i b_i \tag{5}$$

$$a_i = \sum_{j}^{c} n_{ij} \tag{6}$$

$$b_i = \sum_{j}^{c} n_{ji} \tag{7}$$

where A_i indicates the proportion of correctly classified image samples of class 'i' in the total number of samples of this particular class, N denotes the total number of samples that are to be classified, a_i indicates the true samples of class 'i' ground objects and b_i indicates the predicted image samples in class 'i' ground objects.

The experiments in the proposed approach have been implemented using the Keras open-source library which is built on top of TensorFlow, in a system with Windows 10.

4.2 *Hyperspectral Image Datasets*

To demonstrate the ability of the proposed AGAN-PG model, two popular HSI datasets were used—Indian Pines dataset (collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)) and Pavia University dataset collected by Reflective Optics System Imaging Spectrometer (ROSIS). Due to the challenges like high dimensionality and inadequate training samples in these datasets, their interpretation had been proved to be difficult. A brief description of these two image datasets are as follows.

Indian Pines database was collected in 1992 using the AVIRIS imaging sensor. This database covers Northwest Indiana Agricultural Region which is mainly used for agricultural research. The specifications of the dataset have been listed as follows:

- Number of spectral bands: 220
- Spatial resolution: 20 m

Table 1 Training and testing image samples—Indian pines dataset

No.	Class	Training	Testing
1	Alfalfai	14	32
2	Corn-notilli	428	1000
3	Corn-mintilli	249	581
4	Corni	71	166
5	Grass-pasturei	145	338
6	Grass-treesi	219	511
7	Grass-pastured-mowedi	8	20
8	Hay-windrowedi	143	335
9	Oatsi	6	14
10	Soybean-notilli	292	680
11	Soybean-mintilli	737	1718
12	Soybean-cleani	178	415
13	Wheati	62	143
14	Woodsi	380	885
15	Buildings-Grass-Trees-Drivesi	116	270
16	Stone-Steel-Towersi	28	65
Total		3076	7173

- Dimensions: 145 × 145
- Spectral range: 0.4–2.45 μm
- Types of ground objects: 16
- Total labeled pixels: 10,249.

Tables 1 and 2 show the number of samples from various types of ground objects in the Indian Pines dataset and Pavia University dataset.

Table 2 Training and testing image samples—Pavia University dataset

No.	Class	Training	Testing
1	Asphalti	1989	4642
2	Meadowsi	5595	13,054
3	Gravelsi	630	1469
4	Treesi	919	2145
5	Painted metal sheetsi	404	941
6	Bare Soili	1509	3520
7	Bitumeni	399	931
8	Self-Blocking Bricksi	1105	2577
9	Shadowsi	284	663
Total		12,834	29,942

Pavia University dataset: This HSI dataset was collected by ROSIS imaging sensor and the area of coverage is through Pavia University in northern region of Italy. The specifications of the dataset have been listed as follows:

- Number of spectral bands: 103
- Spatial resolution: 1.3 m
- Dimensions: 610 × 340
- Types of ground objects: 09
- Total labeled pixels: 42,776.

The proposed AGAN-PG model has been compared with the two other image classification strategies, viz., 1D-CNN and the Ada-CNN models. Figures 2 and 3 depict the false color image with the respective ground truth maps distribution of that HSI and the classification map produced by the proposed AGAN-PG model for Indian Pines and Pavia University datasets, respectively.

The classification accuracies of Indian Pine dataset have been presented in Table 3. From the classification accuracy values reported in the table, it is concluded that

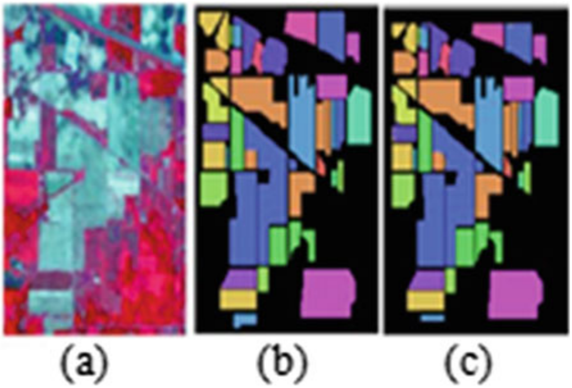


Fig. 2 Indian Pines dataset—classification map. **a** iOriginal dataset, **b** iGround-truth Classification Map; **c** classification map using AGAN-PG

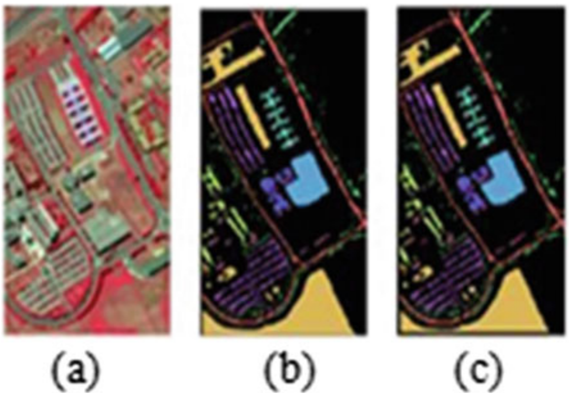


Fig. 3 Pavia University dataset—classification map. **a** iOriginal dataset, **b** iGround-truth Classification Map; **c** classification map using AGAN-PG

Table 3 Classification accuracy for each class—Indian Pines dataset

No.	Class	1D-CNN	Ada-CNN	Proposed AGAN-PG
1	Alfalfai	60.62	57.97	83.93
2	Corn-notilli	74.93	73.03	81.02
3	Corn-mintilli	74.81	82.47	80.84
4	Corni	56.61	66.31	69.92
5	Grass-pasturei	83.40	91.38	86.97
6	Grass-treesi	83.53	86.58	89.98
7	Grass-pastured-mowedi	80.22	69.35	66.63
8	Hay-windrowedi	94.84	94.07	95.27
9	Oatsi	40.70	50.08	64.22
10	Soybean-notilli	75.86	77.16	79.14
11	Soybean-mintilli	73.44	76.90	76.63
12	Soybean-cleani	69.84	78.51	77.87
13	Wheati	90.02	95.59	92.47
14	Woodsi	88.09	90.99	90.63
15	Buildings-Grass-Trees-Drivesi	73.86	73.20	70.99
16	Stone-Steel-Towersi	98.42	98.73	100.00
OA		78.03	81.03	82.16
AA		73.26	74.92	79.18
Kappa		74.73	78.07	79.40

the 1D-CNN performs poor due to the limitation of training image samples. It is evident that the classification accuracy for Class 1 and Class 4 images are 60.62% and 56.61%, respectively, while using 1D-CNN. The overall accuracy is still 78.03% which is very low when compared to the Ada-CNN and the proposed AGAN-PG. The Ada-CNN model performs better than the 1D-CNN with an overall accuracy of 81.03%. Still it performs poor in case of Class 1 and Class 9 images where the number of samples is limited. The proposed AGAN-PG has shown the overall accuracy of about 82.16% though the classification accuracy of each class is unbalanced.

The classification accuracies of Pavia University dataset are reported in Table 4. The 1D-CNN produces a classification accuracy of about 91.88%. Nevertheless, owing to the training sample inadequacy, Class 7 shows an accuracy of around 79.36% only. The Ada-CNN executes a domain adaptation in the basic GAN framework. Both the generated image samples and the original image samples are input to the CNN during classification. Hence, the overall accuracy becomes 92.50%, which is better than that of the 1D-CNN. While the proposed AGAN-PG has an overall accuracy of about 93.38%; this is better than the other two techniques.

The proposed AGAN-PG considers both domain adaptation and conditional random field concepts and improved the quality of generated image samples. This is

Table 4 Classification accuracy for each class—Pavia University dataset

No.	Class	1D-CNN	Ada-CNN	Proposed AGAN-PG
1	Asphalti	94.10	94.97	94.27
2	Meadowsi	93.93	95.21	97.88
3	Gravelsi	76.60	76.10	76.63
4	Treesi	97.40	97.00	99.21
5	Painted metal sheetsi	100.00	100.00	100.00
6	Bare Soili	90.19	93.31	86.25
7	Bitumeni	79.36	85.41	88.75
8	Self-blocking Bricksi	84.11	79.43	85.07
9	Shadowsi	99.30	100.00	99.94
OA		91.88	92.50	93.38
AA		89.51	89.74	91.40
Kappa		89.03	89.86	90.95

seen as the major reason for the improvement of the overall accuracy in case of both the datasets.

5 Conclusion

The proposed auxiliary generative adversarial network model with probability graph model (AGAN-PG) for HSI spatial-spectral classification is operating in two aspects, (1) both the generator unit and discriminator unit employ a 3D deep learning model on the cube HSIs for maintaining the details among the features and (2) the limited annotated image samples and generated synthetic image samples are used to increase the classification performance, especially when the labeled HSI samples are limited practically. The performance of the proposed AGAN-PG has been evaluated over two popular HSIs against the traditional CNN and the adaptive CNN model-based classification methods. The overall accuracies for two groups of datasets are around 82.16% for Indian Pines dataset; and 93:38%, for Pavia University dataset. The integrated AGAN and conditional random field in the form of probability graph model has exhibited promising performance and thus is a new method of semi-supervised learning in the HSI classification task.

References:

1. Ghamisi P, Yokoya N, Li J, Liao W, Liu S, Plaza J et al (2017) Advances in hyperspectral image and signal processing: a comprehensive overview of the state of the art. IEEE Geosci Remote Sens Mag 5(4):37–78. https://doi.org/10.1109/mgrs.2017.2762087

2. Yue J, Zhao W, Mao S, Liu H (2015) Spectral–spatial classification of hyperspectral images using deep convolutional neural networks. Remote Sens Lett 6(6):468–477. https://doi.org/10.1080/2150704x.2015.1047045
3. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Proceedings of the 27th international conference on neural information processing systems, Montreal, Canada, vol 2, pp 2672–2680
4. Xue Z (2019) A general generative adversarial capsule network for hyperspectral image spectral-spatial classification. Remote Sens Lett 11(1):19–28. https://doi.org/10.1080/2150704x.2019.1681598
5. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. In: Guyon I, von Luxburg U, Bengio S, Wallach HM, Fergus R, Vishwanathan SVN, Garnett R (eds) Advances in neural information processing systems (NIPS 2017). Long Beach, CA, USA, pp 3859–3869
6. Gao H, Yao D, Wang M, Li C, Liu H, Hua Z, Wang J (2019) A Hyperspectral image classification method based on multi-discriminator generative adversarial networks. Sensors 19(15):3269. https://doi.org/10.3390/s19153269
7. Zhao W, Chen X, Chen J, Qu Y (2020) Sample generation with self-attention generative adversarial adaptation network (SaGAAN) for hyperspectral image classification. Remote Sens 12(5):843. https://doi.org/10.3390/rs12050843
8. Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I, Abbeel P (2016) Infogan: interpretable representation learning by information maximizing generative adversarial nets. In: Advances in neural information processing systems. MIT Press, Barcelona, Spain, pp 2172–2180
9. Zhong Z, Li J, Clausi D, Wong A (2020) Generative adversarial networks and conditional random fields for hyperspectral image classification. IEEE Trans Cybern 50(7):3318–3329. https://doi.org/10.1109/tcyb.2019.2915094
10. Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434
11. Dalla Mura M, Benediktsson J, Waske B, Bruzzone L (2010) Morphological attribute profiles for the analysis of very high resolution images. IEEE Trans Geosci Remote Sens 48(10):3747–3762. https://doi.org/10.1109/tgrs.2010.2048116

VESDNet: Deep Vessel Segmentation (U) Network for the Early Diagnosis of Diabetic Retinopathy

G. Kumaran, Balika J. Chelliah, and S. Arun Kumar

Abstract Eye screening is the first level diagnosis method among the other medical diagnosis in finding the anomalies in human body. Diagnosing those diseases in the initial stage of the problem can save the people from losing visual and life. Deep analyzing of retinal anatomical components like as blood vessels, exudates, hemorrhages, fovea, microaneurysms and optical disk are the major tasks that contribute to the pre-screening of the anomaly detection using digital fundus images. Among them, accurate segmentation of blood vessels is the most widely accepted method of diabetic retinopathy diseases analysis as damage in the blood vessels is the sign for the diabetic retinopathy and other health issues. In this presented paper, we have addressed an approach based on deep convolutional neural network for the accurate blood vessel segmentation in the context of diabetic retinopathy analysis. In our approach, the retinal image is first pre-processed using computer vision algorithms and vessel segmentation is performed using multi-label deep learning segmentation model corresponding to thick and thin vessel pixels. We performed comparison of our approach with existing methods for both accuracy and speed using the open-source DRIVE and MESSIDOR fundus image databases.

Keywords Blood vessel · Diabetic retinopathy (DR) · Color fundus image · Medical image segmentation · Deep learning

1 Introduction

Precise segmentation of retinal morphology is always be the highly significant need in ophthalmologic pathologies such as diabetic retinopathy and glaucoma. It has been identified globally that the most common cause for visual damage in human eye is due to diabetic retinopathy (DR) and the common approach of assessing the severity of DR is analyzing vascular structure of the retina. Diabetes minimizes the life span

G. Kumaran (✉) · B. J. Chelliah · S. Arun Kumar
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_32

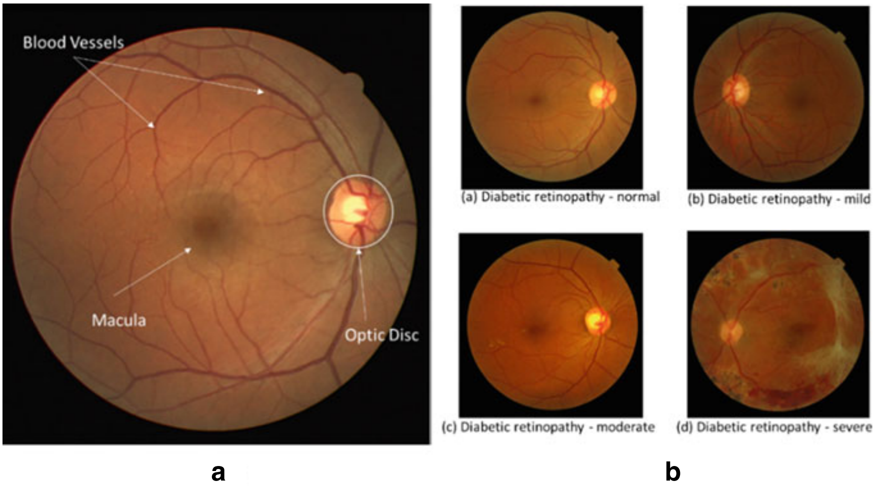


Fig. 1 Fundus image structural components and various DR conditions: **a** anatomical components of fundus images, **b** various diabetic retinopathy conditions

of the human being up to ten years and which is the most cause of blindness and non-traumatic amputation [1, 2]. Similarly, glaucoma is another major cause of blindness in vision system of human and the assessment of glaucoma severity accomplished by analyzing the optic nerve head of retina. Optic nerve head (ONH) analysis requires accurate elimination of vessel structure. Also, the analysis of vascular structure itself directly will leads into identification of up normal structures such as microaneurysms, executes and other lesions. Evaluation of such an animalities can help us diagnosis of numerous health conditions such as DR, hypertension and arteriosclerosis. Figure 1 shows the various anatomical structures of retinal image.

In early days, vascular structure segmentation was performed manually by ophthalmologist experts who have extensive knowledge about the vascular structures. The major challenges involve while doing manual process is delineating thin vascular structures due to its low brightness variations with background. This always tends to inconsistent in accuracy and prone to errors. So, the automatic segmentation of vascular structure is always demanding one for the accurate separation of blood vessels and early diagnosis of DR and other health conditions. So, numerous computer vision approaches have been proposed so far which are varying from traditional approaches, mid-level to high level machine learning and deep learning-based approaches.

As part of making artificial intelligent (AI) systems in real, image processing and computer vision fields play significant roles. Machine learning (ML)-based approaches are very common in making intelligent computer vision systems in which features are extracted and system is trained to learn and accomplish the specific tasks. This approach is called supervised method because it uses prior data knowledge and uses ANN, SVM and other learning algorithms for progressively learning handcrafted object features. Among the many proposed algorithms, ANN-based approaches are giving promising results on the accuracy perspective comparing to the other algorithms due to its behaviors like how our brain learning system works. However, its computational cost is very high and due to the inadequate computational hardware systems not much research was done using it during the early days. Now there have been a great advancement in the computational hardware (GPU) systems which attracted the researches to revisit the ANN-based approaches and proposed the new set of algorithms called deep neural networks or simply called the filed as deep learning (DL).

In this paper, extracting blood vessel structure from the fundus scan images using the emerging deep learning approach has been addressed and compared with prior existing methods. The main objective of our approach is to show the improvement in accuracy in detecting both thick and thin vessels effectively by enhancing the U-Net-based deep learning segmentation network model.

We have organized this paper into many sections as follows: Various existing works on blood vessel removal methods are discussed in Sect. 2. Our method is discussed in Sect. 3. Section 4 talks about experimentation and system performance evaluations. Conclusion and future scope details are given in Sect. 5 finally.

2 Related Works

Medical image segmentation field is the hot topic in the past few decades, and many numbers of researches and algorithms have been proposed and in particularly retinal image analysis using fundus images. At the high level, vessel segmentation techniques are classified under two broad categories, viz., unsupervised and supervised segmentation methods. Methods that are belong to supervised category use two set of datasets: training and testing dataset. Both datasets consist of ground truth data labeled as vessel pixels and non-vessel pixels by expert ophthalmologists. Unsupervised methods do not need any labeled data for classification. It will be used for extracting hidden pattern of vessel pixels by applying various filter techniques.

In the early stages of this field, blood vessel segmentation approaches were often using the traditional methods, wherein edge detection, template matching and morphological analysis-based techniques have used. In [3], the authors addressed segmenting retinal blood vessel using combination of 2D Gabor filter and matched filters. In their work, the enhancement of blood vessels was achieved by Gabor filters and vessel extraction was performed using matched filters. The authors showed the significant results using their proposed system. In [4], 2D matched filters were proposed by the authors. In their proposed methods, vessel segmentation was achieved by applying CLAHE method as pre-processing technique to improve the vessel lines and followed with Laplacian of gradient (LoG). The main advantages of their system are less computation time, easy implementation, effective and simple.

In [5], Kumar et al. discussed vessel segmentation using region growing approach. The authors used morphological methods to highlight the vessel centerlines. Then, vessels connectivity was identified using directional filters. Final region growing algorithm was performed to join the all vessel pixels to detected vessel centerlines. Wang et al. [6] presented another unsupervised approach based on morphological processing of retinal images. In their method, they used corrected morphological transformation and fractal dimension and segmented the vessel pixels in retinal images. Enhancement of contrast between vessels and background was done using linear structured morphological operators. The presented approach was applied on publicly available dataset and shown great achievement in the accuracy. You et al. [7] presented a novel radial projection method scheme to locate vessel center linear lines presented in the low intensity levels. Also, a modified steerable complex wavelet transform-based method was used to enhance the vessel lines. Finally, they have used semiautomatic method to combine both to segment the vessel lines.

Xu et al. [8] proposed a system based on SVM classifier and used wavelet features. In their method, wavelet features were extracted from the fundus images and used line detector to detect the reedy vessel pixels. These features were trained using SVM and image pixels are classified into vessel pixels and non-vessel pixels according the features trained. Another SVM-based method was worked and proposed by Osareh and Shadgar [9] using features extracted from Gabor transform from fundus images. In combined with Gaussian mixture model (GMM), SVM was used to classify the vessel and non-vessel pixels. This method shows advantage of in detecting thin vessel pixels and has limitation in detecting reediest pixels due to the presence of noise in the images.

Neural network-based methods are based on inspired by natural human learning system. It is simulated to behave as our human learning system. Thangaraj et al. [10] proposed neural network-based novel system for effective segmentation of vessel pixels in retinal images. In their approach, 13 handcrafted features were derived such as Gabor filter responses, Hu moment invariants, GLCM features, local binary pattern features and Frangi's features to train NN classifier. Marín et al. [11] proposed a method based on neural network approach, wherein the authors used GLCM and moment invariant techniques as feature extractors. The handcrafted features that represent the vessel and non-vessel pixels are trained using NN classifier and vessel segmentation was achieved. Fraz et al. [12] proposed another supervised approach using boosted decision trees. In their proposed system, the authors have used feature vectors obtained using gradient vector field orientation, line strength measures, morphology transformation along with Gabor filter features. The authors shown performance in terms of accuracy and speed along with other parameters such as robustness and simplicity using open dataset.

With the recent advancements in computational systems and convolutional neural network algorithms, automatic features extraction using CNN have become latest emerged algorithms in medical image segmentation which is becoming higher due to its promising accuracy results. Cross-modality data transformation-based approach was discussed in Li et al. [13]. In this paper, the authors proposed a novel method, wherein retinal image to vessel map transformation is performed using deep neural network training for the transformation to achieve. The accuracy of the proposed procedure attained good score. In [14] proposed a bottom-top short connection, a multi-scale supervised deep neural network approach. In this method, the authors employed short connections along the side layers in order to transfer semantic information. ResNet-101 and VGGNet models were used for experimentation. The major limitation of this method was failed to accurately segment the thin vessel pixel. A combined approach using convolutional neural network and structured prediction was proposed by Dasgupta et al. [15]. A multi-label inferencing kind of segmentation task was formulated by the authors of this paper. And they have utilized the implicit advantages of CNN and structured prediction combination to achieve the segmentation. The proposed method achieved a good accuracy. A three-level network model learning for vessel extraction was proposed by Yan and Yang [16]. The authors of this paper created three stages of training network for the wide and thin vessel segmentation as well combined fusion network. The proposed method showed the good discrimination ability to segment thick and thin pixel segmentation as well. Ronneberger et al. [17] proposed U-Net architecture for the bio-medical image data segmentation. The proposed method framework is an end to end fully convolutional network (FCN)-based approach, and many proposals have been addressed based on this U-Net architecture and its variants for blood vessel segmentation [18–23]. All the above deep learning methods use data augmentation techniques to increase the dataset and pre-processing methods for the improvement of model accuracy. Hybrid methods use both computer vision and deep learning methods combined to get the desired accuracy as well as to increase the system performance.

3 Proposed Method

Our approach is named as VESDNet (Deep Vessel Segmentation Net), a deep learning-based vessel segmentation approach having U-Net as its base architecture. The reason for adapting U-Net as base architecture is due to its success in pixel level segmentation problems and in particularly medical image analysis. The base U-Net architecture is reformed with additional layers at both contradiction and expansion side and parameters fine-tuning was performed to achieve the increased accuracy. The overall functional flow of our implemented work is illustrated in Figs. 2 and 3 and illustrated the architecture of the given system and segmentation network used in our approach for segmentation of vessel pixels in fundus retinal image.

As shown in the network diagram, the VESDNet consists of two paths: contradiction and expansive. The flow in the left is called contradiction path and the right side given flow is called expansive path. The contradicting path in the left is the typical convolutional network architecture which consists of 3×3 convolutions repeated for three times. The feature map created from convolution process is passed to rectified linear activation function (ReLU) layer is applied and followed by 2×2 max pooling operation is applied to reduce the size of the feature map. At finishing each max pooling operation, down sampling is achieved and in parallel feature channels are doubled.

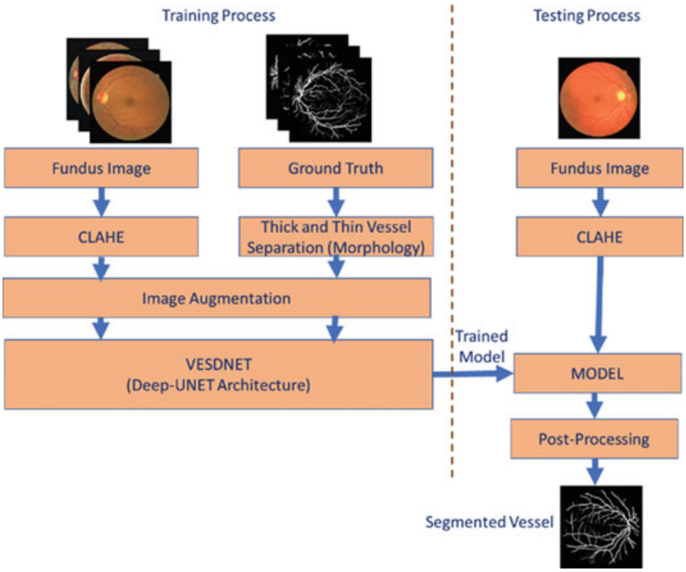


Fig. 2 Architecture of the implemented system

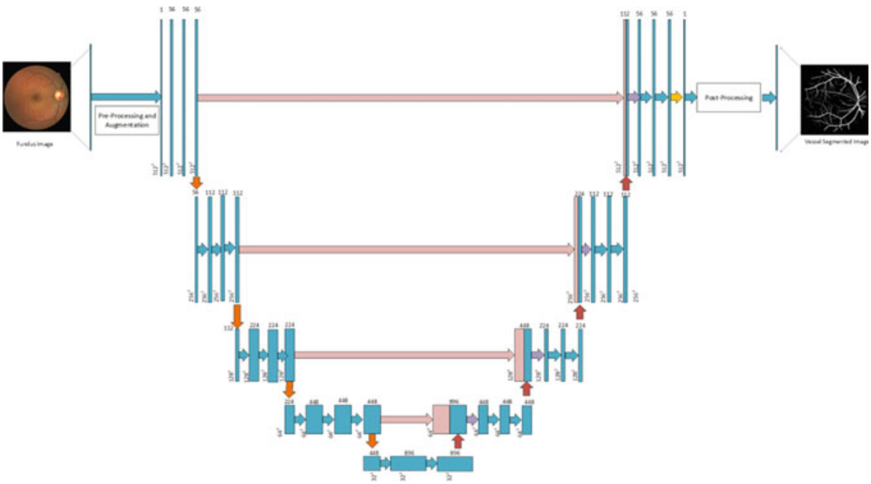


Fig. 3 VESDNet deep learning network architecture

In contradictory to encoder flow, decoder flow comprises steps for performing up sampling of feature maps followed by convolution process with kernel size of 2 × 2. As a result of convolution process, feature channels got reduced by two in contrast to encoder counterpart. So, in order to keep the feature channels as same as encoder path, encoder feature channels are concatenated with respective feature channels from contradiction path by cropping it to match the size same as expansive path size. Followed to process of concatenation, regular convolutions operation is applied further with two kernel of size 3×3 followed by ReLU nonlinear activation. Finally, in the last layer, a 1×1 convolution operation is applied on each 56-component vector that is mapped to desired class.

4 Experiment

4.1 Dataset

We have used open-source available fundus dataset: DRIVE and MESSIDOR for our experiments. Number of color fundus images contained in the DRIVE dataset is 40 that were captured using Canon CR5 3-CCD camera with a maximum FOV of 45^{o} and the resolution of 565×584. The MESSIDOR dataset contains 1200 color fundus images acquired using 3-CCD camera. Images were captures by mounting the camera on a Topcon TRC NW6 retinograph with a 45° FOV and in three resolutions: 1440×960, 2240×1488 or 2304×1536. In our experiment, we used one set of images for testing our model.

4.2 Fundus Images Pre-processing

Preparing the Original Data Before training the model, pre-processing need to perform on each image in dataset to enhance the vessel to background contrast. Enhancement of the fundus is accomplished by applying contrast limited adaptive histogram equalization (CLAHE) method. This process enhances the vessels to background contrast and hence improves the quality of the retinal image.

Data Augmentation We have applied augmentation techniques to increase the training samples size. We have used simple augmentation algorithms such as rotation and horizontal and vertical flipping.

Preparing the Ground Truth Data Both thick pixels and thin pixels are separated out from the ground truth binary map. Morphological skeletonization and open/close operations are performed to do this process. After this step, we would have two ground truth binary maps for thick and thin pixels, respectively.

4.3 Training Parameters

Regarding the network parameters, Adam optimizer was used in place of optimizer and we have set learning rate of 0.0001 to train the network. At the final layer, softmax function was used to perform the classification of vessel pixel. Binary image of vessel is obtained by thresholding probability map with optimum threshold of 0.5 and 0.3 for both thick and thin vessels, respectively. Keras and TensorFlow were used as the deep learning APIs for our implementation using Python. Training was done for 200 epochs.

4.4 Development Platform Setup

The hardware environment of desktop includes NVIDIA RTX 2080Ti GPU, Intel Core i9 9900 K (5.0 GHz Turbo) processor, 11 GiB graphics memory and 32 GB of RAM running on Linux Ubuntu 18.04. We have performed both training and testing on the same hardware setup.

4.5 Performance Evaluation

The performance of our system was evaluated on validation dataset of the DRIVE dataset using the sensitivity (Se) and specificity (Sp) metrics as given below. Given both manually annotated ground truth data and system generated segmented data, then above-mentioned performance metrics are computed as,

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \tag{1}$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \tag{2}$$

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{3}$$

where,

TP True Positive: vessels pixels that are correctly detected
TN True Negative: non-vessels pixels that are correctly detected
FP False Positive: non-vessels pixels that are in-correctly detected
FN False Negative: vessels pixels that are missed to detect.

4.6 Experimental Results

The results obtained using the VESDNet segmentation model on DRIVE test dataset as well as MESSIDOR dataset are shown in Fig. 4. From the figure, it is clearly indicating the effectiveness of our approach on the fundus image blood vessel segmentation. The probability map of the produced results shows how effective the addressed system in classifying vessel and non-vessel pixels with high accuracy rate. Comparison results of our approach with current bench marking results are summarized in Table 1 for DRIVE dataset. Overall, as considering the imbalance in number of vessel and background pixels and number of thin and thick vessel pixels, our system performance considered as better (Fig. 5).

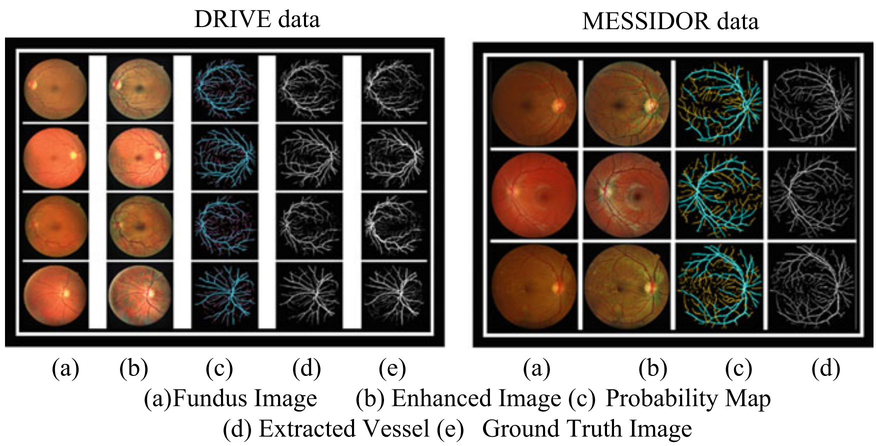


Fig. 4 Vessel segmentation results using our method on DRIVE and MESSIDOR dataset

Table 1 Comparison results on **DRIVE** dataset

Method	Year	Se.	Sp.	Acc.
You et al. [7]	2011	0.7410	0.9751	0.9434
Marin et al. [11]	2011	0.7067	0.9801	0.9452
Fraz et al. [12]	2012	0.7406	0.9807	0.9480
Li et al. [13]	2016	0.7569	0.9816	0.9527
Dasgupta and Singh [15]	2017	0.7691	0.9801	0.9533
Yan and Yang [16]	2019	0.7631	0.9820	0.9538
Our approach	**2020**	**0.7724**	**0.9706**	**0.9614**

The bold values shows the results obtained in the proposed system and was compared with previous proposed systems in terms of sensitivity, specificity and accuracy. As indicates in the Table, propsed system behaviour is better than previous results in terms of vessel segmentation accuracy

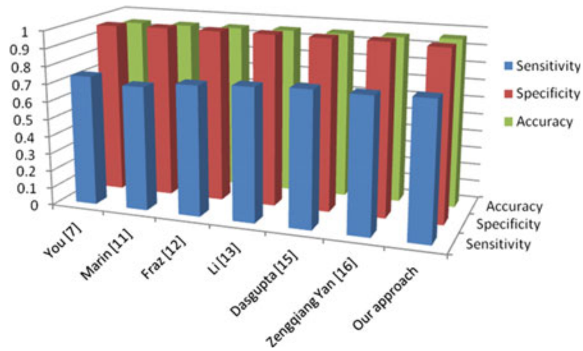


Fig. 5 Comparison of our method with other proposed approaches on **DRIVE** dataset

5 Conclusion and Future Work

In this presented paper, we have addressed the segmentation work for accurate retinal structure of the eye called blood vessel in fundus retinal images using deep learning techniques. Our approach uses separate training set for both thick and thin vessel pixels during the model building. While testing, the resulting probability map is separated out for thin and thick vessel maps and finally merged to get final vessel image. The effectiveness of our method is compared with other proposed systems on DRIVE dataset. Considering the imbalance in the vessel and background pixels and considering the imbalance in the thick and thin vessel pixels, our method considered as best model. The future scope our work is to focus on optimizing the execution time of the system by minimizing number of parameters involved in the network architecture without impacting on accuracy.

References

1. Marshall SM, Flyvbjerg A (2006) Prevention and early detection of vascular complications of diabetes. BMJ 333(7566):475–480
2. Hutchinson A, McIntosh A, Peters J, O'Keeffe C, Khunti K, Baker R, Booth A (2000) Effectiveness of screening and monitoring tests for diabetic retinopathy—A systematic review. Diabet Med 17(7):495–506
3. Yavuz Z, Köse C (2010) Comparing 2D matched filter response and Gabor filter methods for vessel segmentation in retinal images. In: National conference on electrical, electronics and computer engineering (ELECO). IEEE, pp 648–652
4. Kumar D, Pramanik A, Kar SS, Maity SP (2016) Retinal blood vessel segmentation using matched filter and Laplacian of gaussian. In: International conference on signal processing and communications (SPCOM). IEEE, pp 1–5
5. Jerald Jeba Kumar S, Ravichandran CG (2017) Morphological operation detection of retinal image segmentation. In: International conference on intelligent sustainable systems (ICISS). IEEE, pp 1228–1235
6. Wang W, Wang W, Hu Z (2019) Retinal vessel segmentation approach based on corrected morphological transformation and fractal dimension. IET Image Process 13(13):2538–2547

7. You X, Peng Q, Yuan Y, Cheung YM, Lei J (2011) Segmentation of retinal blood vessels using the radial projection and semi-supervised approach. Pattern Recogn Lett 44(10):2314–2324
8. Lili Xu, Luo S (2010) A novel method for blood vessel detection from retinal images. Biomed Eng Online 9(1):14
9. Osareh A, Shadgar B (2009) Automatic blood vessel segmentation in color images of retina. Iran J Sci Technol 33(2):191–206
10. Thangaraj S, Periyasamy V, Balaji R (2017) Retinal vessel segmentation using neural network. IET Image Process 12(5):669–678
11. Marin D, Aquino A, Gegndez-Arias ME, Bravo JM (2011) A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Trans Med Imaging 30(1):146–158
12. Fraz MM, Remagnino P, Hoppe A, Uyyanonvara B, Rudnicka AR, Owen CG, Barman SA (2012) An ensemble classification-based approach applied to retinal blood vessel segmentation. IEEE Trans Biomed Eng 59(9):2538–2548
13. Li Q, Feng B, Xie L, Liang P, Zhang H, Wang T (2016) A cross modality learning approach for vessel segmentation in retinal images. IEEE Trans Med Imaging 35(1):109–118
14. Guo S, Wanga K, Kanga H, Zhang Y, Gao Y, Li T (2019) BTS-DSN: deeply supervised neural network with short connections for retinal vessel segmentation. Int J Med Inf 126:105–113
15. Dasgupta A, Singh S (2017) A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation. In: Proceedings of ISBI, pp 18–21
16. Yan Z, Yang X, Cheng KT (2019) A three-stage deep learning model for accurate retinal vessel segmentation. IEEE J Biomed Health Inf 23(4):1427–1436
17. 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
18. Luo Z, Zhang Y, Zhou L, Zhang B, Luo J, Wu H (2019) Micro-vessel image segmentation based on the AD-UNet model. IEEE Access 7:143402–143411
19. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, Zhang T, Gao S, Liu J (2019) CE-Net: context encoder network for 2D medical image segmentation. IEEE Trans Med Imaging 38(10):2281–2292
20. Son J, Park SJ, Jung KH (2019) Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks. J Digit Imaging 32(3):499–512
21. Jin Q, Meng Z, Pham TD, Chen Q, Wei L, Su R (2019) DUNET: a deformable network for retinal vessel segmentation. Knowl-Based Syst 178:149 –162
22. Wang C, Zhao Z, Ren Q, Xu Y, Yu Y (2019) Dense U-Net based on patch-based learning for retinal vessel segmentation. Entropy 21(2):168
23. Ribeiro A, Lopes AP, Silva CA (2019) Ensemble learning approaches for retinal vessel segmentation. In: 6th Portuguese meeting on bioengineering (ENBENG). IEEE, pp 1–4

'Teaching an Old Dog New Tricks': A Comparative Study on Solutions for Connectivity of Legacy Machinery

Thomas Kendall, Christopher Walsh, R. Jagadeesh Kannan, Judith Apsley, Emad Alsusa, Paulo Bartolo, David Gillen, and Carl Diver

Abstract The Fourth Industrial Revolution (Industry 4.0, I4.0) is the current move in the manufacturing industry to adopt and integrate digital technologies throughout entire company structures with the focus on data and interconnectivity. Although the benefits of these digital technologies are yet to be fully realised, large industrial manufacturing companies are investing heavily into the design and acquisition of digital infrastructure. Small/medium enterprises (SMEs) are set to benefit from this industrial revolution through increased productivity, competitiveness and integration along the supply chain. Barriers for SMEs to adopt the digital technology include the large financial investments in systems and new 'smart' equipment to replace non-connected legacy equipment as well as the skills gap of the workforce to develop, maintain and benefit from the technology. To help overcome these barriers, this study compares three solutions for connecting legacy manufacturing equipment to cloud platforms of varying cost, from hundreds to several thousands of pounds, and capability whilst analysing the equipment and skills required to implement the solutions. This paper shows a range of possible solutions available without large scale financial investment, whilst critically assessing the functionality, security, knowledge required and cost for manufacturers. The solutions are implemented and tested on manufacturing systems at the University of Manchester and wider industrial settings originally lacking in connectivity.

Keywords Industrial Internet of things · SME · Digital technology · Digital manufacturing

---

T. Kendall · C. Walsh · J. Apsley · E. Alsusa · P. Bartolo
The University of Manchester, Manchester, UK

R. J. Kannan
School of Computer Science & Engineering, Vellore Institute of Technology (VIT), Chennai, India

D. Gillen
Blueacre Technology Ltd, Dundalk, Ireland

C. Diver (✉)
Manchester Metropolitan University, Manchester, UK
e-mail: C.Diver@mmu.ac.uk

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_33

1 Introduction

The Fourth Industrial Revolution (Industry 4.0) is the ongoing movement within the manufacturing industry focused on the implementation and development of new digital technologies to transform processes horizontally along the supply chain as well as vertically within individual organisations. First introduced at the Hannover Messe Trade show in 2011 [1], companies are becoming increasingly interested and investing in the developing technologies encompassed by Industry 4.0 to provide data, competitiveness, flexibility and resource efficiency in response to changing consumer expectations and demands [2].

Despite the benefits realised by the Industry 4.0 revolution, there are still specific barriers to adoption, especially in small-to-medium enterprises (SMEs). A lack of digital skills in the workforce, financial resources, cybersecurity concerns and standardisation problems are all identified in the research as barriers to Industry 4.0 implementation [3]. Although implementation for digital technologies has been widely discussed for large multinational enterprises (MNEs) [4], research into the driving forces and barriers to SMEs has been more limited [5].

One major aspect of the Industry 4.0 revolution and the 'smart factory' concept is the Industrial Internet of Things (IIoT). IIoT is the network produced by the connectivity and transfer of data between industrial devices. It is a cornerstone of Industry 4.0 and the mechanism for which the valuable data is collected, transferred and analysed. IIoT incorporates machine learning, deep learning and big data technologies to harness the sensor data with the aim of improving operational equipment efficiency (OEE) of processes [6]. IIoT technologies are helping in ways beyond this, monitoring of processes and events, predictive maintenance, automation, safety monitoring and making data available for use constantly and remotely in real time. Artificial intelligence algorithms processing the data procured through IIoT will also enhance the predictive nature of supply chains and maintenance routines. Devices in the IIoT generate and exchange vast amounts of data, some of which can be safety and security critical or commercially sensitive and as such security of the data is of paramount concern [7].

Internet of things (IoT) solutions consist of three layers for an end-to-end solution. The three layers consist of sensing, networking and application [6, 8]. The sensing layer is a constrained device (the 'thing' in IoT), where the device could be constrained by many factors whether it is storage, processing power or networking capabilities. The networking layer consists of protocols and transfer standards to allow constrained devices to connect to existing networks. The transfer standards dictate how a message is sent between devices. The application layer consists of both front-end and back-end software. Back-end software will allow for messages to be received from the constrained device and stored in a database. The front-end software will allow for visualisation of the data and meaningful outputs.

The capital investment in large-scale changes to manufacturing infrastructure towards the concept of an interconnected 'smart factory' is high. To produce the data from production processes and systems requires interconnected machinery and devices as well as the IT infrastructure to gain value from the collected data, and it is this capital investment and required digital knowledge that acts as a barrier to SME's. As SMEs are more likely than MNE's to have legacy and non-connected equipment throughout the manufacturing system, without embedded sensors and connected systems, the collection of valuable data and connectivity required to realise the benefits of IIoT is currently unobtainable. Enabling the legacy equipment significantly reduces costs by improving OEE and extending the remaining useful life of production machines [9] as well as helping companies effectively sees the utilisation of equipment in a timely manner, rather than relying on 'paper-based' reporting systems. It is important to consider the data required for individual circumstances, as not all data for every aspect of a manufacturing system is required. What data, signal types, connectivity and collection rate all must be considered, as any extra equipment and processing power requirement adds cost and complexity in terms of maintenance.

It can be argued that without the large, embedded company structures of MNE's, SME's themselves can react to and change quickly within the digital manufacturing ecosystem to provide the benefits of the technology. However, this requires some form of investment from SME's in digital infrastructure and upskilling the workforce. Without these developments, SMEs are at risk of falling behind those who invest in digital technologies in terms of competitiveness. To adopt IIoT, a company must make a financial and systematic commitment to innovating company structures with a long-term view on potential benefits and alignment with new company objectives as benefits may not be seen in immediate results [10]. This does not just include physical assets, and there is also the challenge of connecting all digital aspects of a company, from administration to enterprise resource planning (ERP), with the option of investing in complete pre-built company systems or individual customised systems for individual departments or SME's.

Currently, there are no specific guidelines or roadmaps for IIoT development for SME's, with no set individual protocols or data transfer mechanisms. Industry 4.0 is an interdisciplinary approach to developing manufacturing systems, requiring knowledge of both manufacturing systems, digital technologies and software development and implementation. Industry 4.0 requires companies and workers to develop a range of skillsets addressing these areas and working with a collaborative approach.

With an aim to overcome the aforementioned barriers to Industry 4.0 adoption and to showcase the requirement of interdisciplinary approaches to manufacturing, this study compares three solutions for adding IIoT capability to legacy manufacturing equipment. These solutions encompass low-cost open-source solutions, as well as intermediate and industrially capable solutions. The solutions are the outcome of three different projects within the Department of Mechanical, Aerospace and Civil Engineering and Department of Electrical and Electronic Engineering at the University of Manchester. These solutions provide data collection and transfer mechanisms for otherwise non-connected equipment in a range of manufacturing settings as well as providing cloud solutions for the storage, analysis and visualisation of manufacturing data. The solutions themselves vary in cost, capability and knowledge required for implementation. The background and motives of each project are explained, followed by a description of the solution and its implementation. As part of this study, the solutions are compared and critically assessed in terms of cost, functionality, security and the skills required within the workforce.

2 IIoT Solutions

This section introduces the background to the three projects and outlines the solutions for IIoT connectivity of non-connected manufacturing equipment.

2.1 Solution 1

Background The first IIoT solution stems from a project at the Department of Electrical and Electronic Engineering (EEE) at the University of Manchester. The project is motivated by the needs of small businesses to benefit from an Industry 4.0 solution, where an Industry 4.0 solution enables 'smart' manufacturing through data acquisition and analysis. The benefits from an Industry 4.0 solution allow for a wide range of workflow optimisation. From improved resource management to reduced downtime, all of which improve the business' profitability.

The key motivating factor for this project is the need of SME's in the manufacturing sector. Such businesses lack the capital and knowledge to invest in the latest manufacturing equipment with integrated smart technology. This project investigates how a small business may be able to benefit from a low-cost retrofitting Industry 4.0 solution by enabling insights into the manufacturing process for optimised decision-making.

The aim of this project is to demonstrate IIoT in the EEE mechanical workshop through monitoring equipment to enable predictive maintenance. The objectives comprised of analysing the requirements of the mechanical workshop, developing a low-cost wireless sensor network, developing a cloud platform for data storage and analytics, visualisation of the data and developing a business case for the solution.

The mechanical workshop in EEE contains several pieces of legacy and non-connected equipment and as such is an ideal candidate for IIoT developments. The study focusses on five pieces of equipment within the workshop with varying application and data requirements as found from a requirement analysis with the workshop staff. The equipment and the data requirement for each piece of equipment can be seen in Table 1. As none of the equipment had any level of connectivity, sensors for each machine needed to be retrofitted as a primary requirement.

Solution Infrastructure The overview of the solution structure can be seen in Fig. 1. The end device is a microcontroller collecting data from a range of sensors, performing some calculations on the data if appropriate. An XBee (Zigbee) radio allows the microcontroller to communicate within a local area network. Each machine

Table 1 Requirements for machine monitoring for Solution 1

Machine type	Manufacturer	Model	Requirement
Mill	Hardinge Inc.	9BRS	Monitor the speed of the spindle on the Bridgeport to check for wear and tear
3D printer	Tiertime	UP300	Monitor the 3D printer for the humidity of the substrate and the flow of information with the PC
Lathe	Colchester	Student 2500	Monitor the Lathe for the oil level to check for lubrication of the bearings
Laser cutter	HPC Laser Ltd.	LS6090 Pro Laser	Monitor for fire, for the output power of the laser and the level of dust particles in the extractor fan
Bandsaw	Startrite Machine Tool Co Ltd.	20RWF	Monitor current consumption level

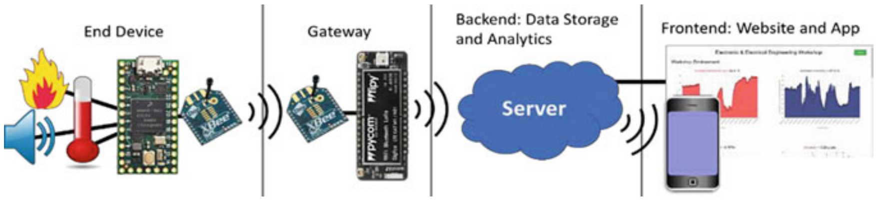


Fig. 1 Solution 1 structure

would have one end device, totalling five. The gateway is a microcontroller (with a range of cloud connectivity options) with an XBee radio to allow for communication with the Zigbee network. The microcontroller processes the data from the end devices and formats the data in an appropriate form to send over the internet, transforming Zigbee packets to HTTP packets. There is only one gateway required in the workshop. The server receives the data from the gateway and stores the data in a MySQL database, and the data is then processed using additional plug-ins to serve the data to the mobile application and website dashboard. The mobile application is used to allow quick and easy access to the information available and to notify the user of immediate issues requiring attention. The website is to act as a storefront to advertise the solution developed and to act as an interface for the user to access more in-depth information not suitable for use in a mobile application.

The sensors were selected from a range of options based on technical features such as improved accuracy, better noise filtration and linearisation, excellent online calibration and diagnosis, low power demand and fast response. However, to ensure that the project objectives are met adequately, a more suitable justification of digital sensors was essential in guaranteeing their safe application, connection compatibility and for a manufacturing environment.

A DHT22 temperature and humidity sensor were selected for monitoring the 3D printer, a HC-SR04 ultrasonic sensor monitors the oil level on the lathe, a CSLA1CD current sensor monitors the current consumption on the bandsaw, a DFROBOT flame sensor monitors the environment in the laser cutter and a GS101201 rotation speed sensor monitors the spindle speed of the milling machine. The sensors selected and purchased for this project are low cost. They are all able to be retrofitted onto the machines without affecting their geometry.

Xbee and FiPy modules were tested for many widely used standards; 802.15.4, Digimesh, Zigbee, Bluetooth, Wi-Fi, Sigfox, LoRa and NB-IoT. The final implementation is a Zigbee local area network transferring data into the cloud over Wi-Fi from a single node acting as a gateway. Additional software for the microcontrollers was developed for each end device whereupon following a request from the coordinator, and the sensor readings are sampled and replied. The Arduino microcontroller of choice is the Teensy 4.0 due to its wide range of connectivity options, small form factor and customisable clock speed.

Due to the open-source nature and low cost of Infrastructure as a Service (IaaS) platforms, two IaaS solutions were developed for this project to highlight two approaches that can be taken. The first uses a MySQL database with a custom chart.js dashboard, the second an XAMPP development environment with Apache distribution containing MariaDB, PHP and Perl as seen in Fig. 2. The Grafana and chart.js solution followed similar development approaches where PHP was used for server-side scripting. Such scripts were written to be able to write data to the database on both systems. The gateway would use a HTTP request to access the desired PHP script and transfer data during the request to be processed by the script, then using

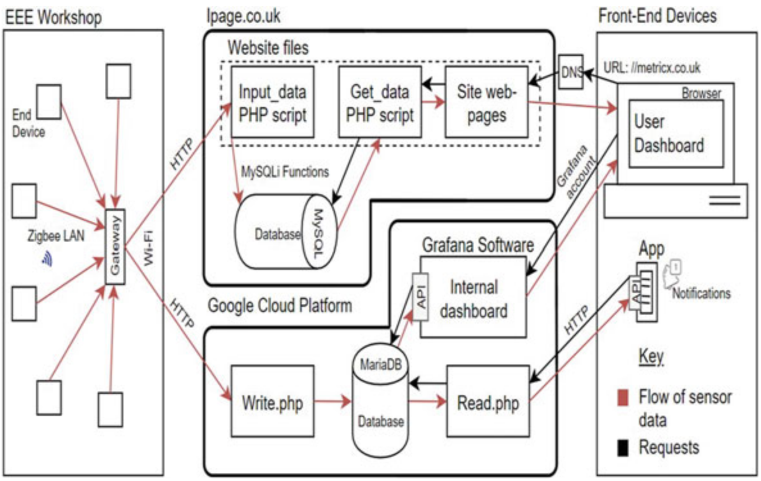


Fig. 2 Solution 1 infrastructure

SQL commands in the script to write the data to the database. The Grafana and chart.js solutions differ when reading the data: Grafana has a direct SQL connection, whereas chart.js required further PHP scripts to read the data and format it for visualisation. The chart.js solution uses web hosting and is receiving identical data to the XAMPP server; the databases are separate for ease of development.

The project website operates as a versatile and simple-to-use home for the business serving several functions. Primarily, the site provides a central platform from which to view collected sensor data. It serves this function though the users individual ‘dashboard’ page where a user may log-in to view their private sensor data in a format which is both secure and customised to suit individual requirements. The website was hosted on iPage with scripts pulling machine metrics from the chart.js and Grafana software.

Finally, an application was developed using *Flutter,* a UI development kit by Google. Early versions used an MQTT broker to call data from a web-hosted database, displaying machine data from the EEE workshop. The final version uses HTTPS requests to read and display last values from the MariaDB (Fig. 2). Alerts are setup on the application to warn and alert the user of any deviation from set parameters on the machines in the EEE workshop allowing action. The development of the solution is ongoing, with all aspects of the solution architecture being continuously improved with an aim for wider functionality of the data.

2.2 Solution 2

Background The second solution stems from a winning entry to the Siemens MindSphere Openspace Challenge competition at the 2018 Hannover Messe. The solution was developed as an answer to a problem put forward by Festo (Festo, Germany) in which a low-cost solution for the monitoring and visualisation of data of pneumatic systems was required. The solution was developed at the Department of Mechanical, Aerospace and Civil Engineering (MACE) at the University of Manchester.

The aim of the project was to develop a low-cost solution built from readily available components and open-source IIoT infrastructure. Building upon the original competition entry, the solution was further developed for industrial capability and expanded to encompass cloud connectivity for online data analysis and visualisation. The purpose of the project was to demonstrate the ability to collect and visualise data for a low cost (less than £1000) as on option for low capital, wide scale implementation across a range of devices. Although the data required for this project is not expansive, only pneumatic supply pressure and ambient temperature data, the solution can be easily adapted for collection of various data from a range of sensors in many industries.

Solution Infrastructure The original solution infrastructure for Solution 2 can be seen in Fig. 3. As it can be seen from following the flow in Fig. 3, the sensor data from

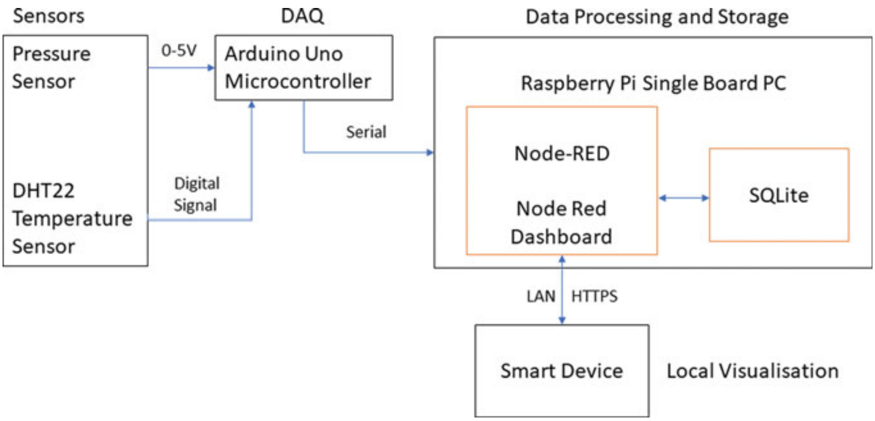


Fig. 3 Solution 2 infrastructure

a Gems pressure sensor and a DHT22 temperature sensor are ultimately visualised on a smart device within the factory location.

The analogue 0–5 V signal from the pressure sensor and the digital signal from the DHT22 sensor are acquired through an Arduino Uno rev 3 microcontroller. The signals are processed and scaled on the microcontroller using a simple C language script and output as a serial communication to the Raspberry Pi single board PC via a USB cable. The Raspberry Pi runs Node-RED flow programming software built on the node.js programming language. Node-RED is an open-source Internet of things (IoT) software solution developed by IBM. It is a flow-based, visual programming tool that allows the building of programs without the requirement of code-based programming.

Data storage is provided by an SQLite database running on the Raspberry Pi. Data is posted to the SQL database through the Node-RED software and can also be queried, edited and deleted by the software. The data is visualised on a local smart device, i.e. on the same network as the Raspberry Pi through the dashboard provided by the Node-RED software in which the pressure data and temperature data can be monitored live, and historical data can be visualised in tables form the SQL database. Different devices around the factory can be accessed by scanning an individual QR code on a smart device to display the relevant device information.

The original solution did not use an industrially certified microcontroller and PC components, nor did the solution have cloud connectivity and as such the solution was further developed to integrate CE certified components and cloud solution. As seen in Fig. 4, the Arduino and Raspberry Pi were replaced by a Siemens IOT2040 intelligent gateway with IO shield. The IOT2040 runs Yocto Linux operating system and has the ability to run Node-RED as in the Raspberry Pi.

For cloud data storage and visualisation, Ubidots is used. Ubidots is a Software as a Service (SaaS) solution that provides data storage, dashboard visualisation and analysis tools. Data can be uploaded to Ubidots by HTTPS and MQTT protocols,

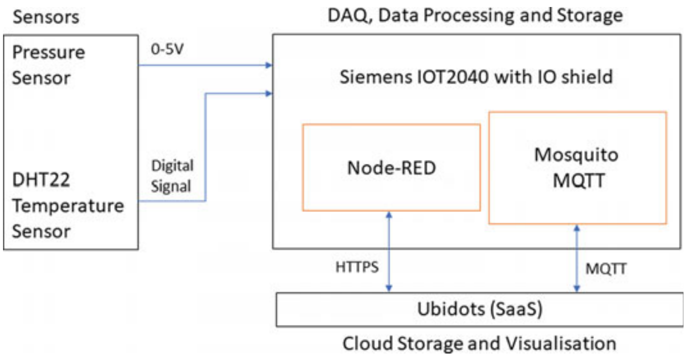


Fig. 4 Solution 2 infrastructure with cloud integration

and in the developed solution both are available. Data can be directed to Ubidots via Node-RED HTTP send nodes, or via MQTT or Ubidots nodes via a Mosquito MQTT broker running on the IOT2040. Ubidots also has the ability to send data to the device for basic remote control applications.

2.3 *Solution 3*

Background The third solution stems from a project at the Department of Mechanical, Aerospace and Civil Engineering in which a novel machining solution is under development with Blueacre Technology Ltd. (Ireland). The original brief for the project outlines the requirement for connectivity in line with the current movements of Industry 4.0, and as such connectivity must be considered from the development outset in a concept called ‘Design for Connectivity’. However, the machine itself is predominantly built from off the shelf components without integrated connectivity or smart sensors. Therefore, an IIoT solution was developed to integrate connectivity into the machining solution. Compared to the previous solutions, components are not retrofitted to the system and so the design of an IIoT solution becomes more difficult, as the components are not designed for connectivity themselves.

Solution Infrastructure As it can be seen from Fig. 5, process data is required for sensors and parts that are not integrated on the same network, with different locations and data protocols. Analogue data from the pressure, temperature and flow data are acquired through a NI CompactRIO DAQ system for simple connectivity to a PC running NI LabVIEW Software. The electrical data from a programmable power supply is acquired through a serial USB connection to the PC and accessed by LabVIEW through specific drivers. The linear motion process data from the drive axes IS acquired through LabVIEW via ActiveX components in the LabVIEW software.

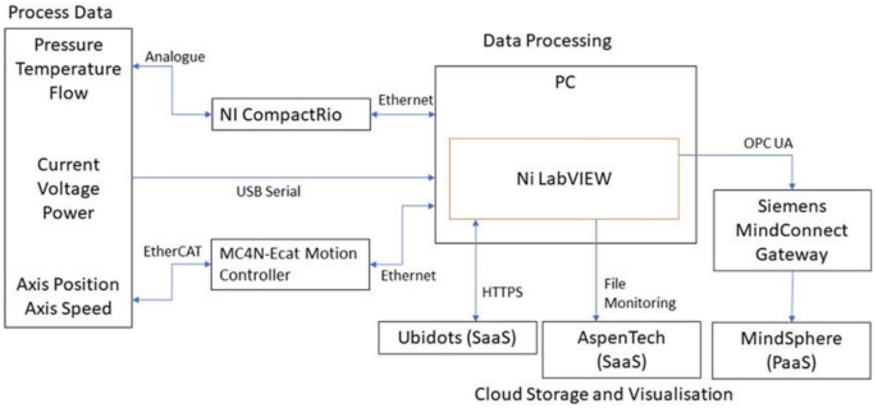


Fig. 5 Solution 3 infrastructure

The LabVIEW software provides all the data processing from all the sensors and components. Three options for cloud connectivity were developed as part of this project, with Ubidots (SaaS), Aspentech (SaaS) and Siemens MindSphere platform as a service (PaaS). Similar to Solution 2, data is uploaded to Ubidots via simple HTTP post commands in the LabVIEW environment. When the solution was developed, the data transfer mechanism for Aspentech's software was through file location monitoring, in which software hosted on the PC monitors a file location and any data file uploaded to this location is uploaded to Aspentech's cloud. For this data transfer, process data was collected and regularly dumped into the folder location through the LabVIEW program for bulk uploads, and as such was not real time. However, since this solution was developed, data transfer to Aspentech is now possible through MQTT protocols. The third option for cloud services is Siemens MindSphere platform. At the time of development, data transfer to MindSphere was only possible through a Siemens MindConnect gateway using an OPC UA protocol, with the OPC UA server in this case hosted on the NI LabVIEW software. However, since development, data transfer to MindSphere is possible through MQTT and various gateway devices, making implementation less complex.

3 Discussion and Comparison

As outlined in the introduction of this report, a lack of digital skills in the workforce, financial resources and cybersecurity concerns are seen as barriers to the implementation of digital technologies in SME's. This section compares the IIoT solutions in the areas of cost, the technical knowledge required to implement the solutions, cybersecurity aspects of the solutions and a comparison of the functionality.

3.1 Cost

The solutions demonstrated in this report vary in the cost of implementation and complexity. As a general rule, the more an SME relies on off-the-shelf solutions for cloud platforms and the less software that is developed internally on open-source platforms, the higher the capital investment required. Solution 1 is the IIoT solution with the highest requirement for internal development. At the time of writing, the cost of the hardware and software is under £600 (software and webhosting less than £50) for monitoring five workshop machines. However, it must be considered that although the capital cost for equipment and software is low, the cost upskilling the digital skills of current workers or bringing in software development specialists is higher than the other two solutions. The long-term investment in an industrially robust solution also needs consideration, as developing solutions internally require the maintenance and updates of digital infrastructure under the responsibility of the company instead of a supplier when compared to a pre-built solution. The solution is easily scalable without a large increase in capital investment.

Solution 2 with integrated cloud capability is the middle IIoT solution in terms of the cost of implementation. The cost of the hardware is less than £500 lb at the time of writing; however, the cost of the cloud platform starts at roughly £50 per month, taking the cost of implementation higher than Solution 1. The largest capital cost is that of the industrial gateway, which has a limit of 5 input devices so scaling of the solution requires a higher incurred cost than that of Solution 1 for every 5 devices required. However, despite the cost, these industrial gateways do provide the added security that is key for larger organisations where less robust solutions would not be acceptable.

Due to the use of highly capable data acquisition (DAQ) hardware and industrial IIoT cloud platforms, Solution 3 incurs the highest capital cost of implementation of around £10,000 for hardware at the time of writing, with the cost of the Aspentech and MindSphere platforms adding further costs in the thousands of pounds. However, the scalability of the solution is strong due to the expandable nature of the NI CompactRIO hardware and the Industrial IIoT platforms to ultimately lower the cost per device.

3.2 Technical Knowledge Requirements

Solution 1 using open-source components has the highest requirement for the technical knowledge of workers. Due to the 'do it yourself' nature of the programming of the software throughout the system and the internal development of the web platform, the digital skills required are high. For example, the microcontroller requires Python scripts, the databases require skills in PHP requests, SQL commands, JavaScript code programming and knowledge of HTTPS and MQTT protocols. The web and application development require knowledge of CSS code, HTML, JavaScript, Python, HTTP and MQTT. All of this must be considered when implementing a solution of this kind as the digital skills required are high and require investment. The provenance of the information for developing these skills is often unknown and due to the reliance on a wider open-source community to not have the ownership and accountability afforded by a company providing set solutions. However, with the low-cost nature of the system, Solution 1 provides a good introduction to the digital technology and skills development for workers of SME's without high capital investment in the technology itself and can act as a gateway to providing a company with the digital skills required in an Industry 4.0-driven manufacturing sector.

Due to the use of the Node-RED flow programming tool and the intuitive Ubidots platform with technical support available from the company and manufacturing community, the technical skills required to implement Solution 2 are much lower than Solution 1. This provides an simple gateway into IIoT solutions for SME's without the requirement for software and web development specialists, providing basic IIoT skills for the current workforce.

Solution 3 requires a strong knowledge of NI LabVIEW software; however. this is a widely used software in the manufacturing industry. As with Node-RED, the intuitive flow programming nature of NI LabVIEW does not require the need for specific knowledge of programming. The expansive nature of LabVIEW allows configuration of several data transfer mechanisms as seen in Fig. 5. The cloud platforms used offer strong functionality without the need for specific digital skills knowledge. Information provided for the developing of digital skills is reliable and verifiable under the ownership of companies when compared to the open-source solutions.

3.3 Cybersecurity

All the solutions consider cybersecurity as of paramount importance. For Solution 1 at gateway level, security is handled through the password protection of local Wi-Fi networks. Any solution using MQTT and HTTPS protocols use encrypted messages. Security issues in the databases were addressed by detailing firewall, password and encryption methods for a custom-built IoT cloud service. For web and application development although some web hosting providers do provide security systems, the responsibility for the security of data and user is the responsibility of the developer and must be considered when internally developing platforms.

When using pre-built commercial cloud hosted platforms as in Solutions 2 and 3, the security is provided by the platform itself and requires consideration only on the gateway and data transfer where encryption and password security on devices and Wi-Fi networks should be used. Although procuring higher cost than internal development, one of the strongest benefits of using IIoT solutions such as MindSphere, Ubidots and Aspentech is the security knowledge and systems they provide over open-source solutions. Although incurring a higher cost, the industrial gateways required for MindSphere with ISO 27001 standards built in do provide greater security over lower-cost IoT gateways relying on application programming interface (APIs).

3.4 Functionality

Although used in the specific applications as outlined in this study, all three IIoT solutions can be configured for wider use cases and the structures themselves are transferrable. In this sense, all three solutions provide strong functionality. For the cloud-based solutions, in Solution 1, the functionality of the platform is dependent on the company and the level of development invested. Data processing for monitoring, predictive maintenance, safety alerts, etc., is all possible but require development, and the technical knowledge is the limiting factor in the functionality of the solution. These developments are useful for introduction to IIoT solutions, and for upskilling the digital skills of the workforce and assessing the benefit of digitalisation; however, there will likely be a point where the systems become to complex and the capital and labour cost of scaling solutions becomes to high for SME’s and a movement to industrial solutions is required, where the digital skills will continue to be useful.

For the SaaS and PaaS solutions used in Solutions 2 and 3, the functionality is dependant on that offered by the provider. For the solutions developed in this study, MindSphere provides the highest functionality at the highest cost, with Ubidots providing the lowest functionality at the lowest cost. With an application-based store, the MindSphere platform provides an ever expanding array of solutions for data visualisation and processing for predictive maintenance, error detection, machine learning amongst other tools. Ubidots provides a highly functional general user interface (GUI) or dashboard for data visualisation as well as systems for alerts and event monitoring, but the data processing offered is low.

Aspentech’s solutions offer a highly functional GUI and a wide range of tools for asset performance, system optimisation, AI and actionable insights amongst others. The functionality of these platforms is offset against the cost when compared to self-developed solutions, but the functionality is more easily accessed and benefited from and should be considered when developing an IIoT solution. All solutions have time series storage, with the MindSphere platform using a hierarchical asset structure compared to the flat asset structures of the other solutions.

4 Conclusions

This study provides an overview of three IIoT solutions developed for various manufacturing settings at the University of Manchester and within industry. As it has been seen, these solutions vary in cost, functionality, cybersecurity and required knowledge for technical implementation, raising considerations for SME’s when developing IIoT solutions. It can be seen that there is no simple direct IIoT solution requiring an individual skillset found within traditional manufacturing environments. In all solutions, new skills were required and developed, with understanding of the manufacturing application, machines, sensors, design and fitting and all the software and digital skills outlined in this study. This falls in line with the barrier on skills in the workforce and needs addressing in any SME when considering implementing digital solutions.

Solution 1, with the highest amount of internal development, offers the lowest cost solution, but requires the highest investment in terms of upskilling the workforce with digital skills and the largest investment in security and functionality. Solutions 2 and 3 offer strong functionality at increased cost; however, the security, functionality and reduced technical knowledge requirements are strong benefits of using pre-built hardware and cloud platform solutions. It must also be noted that the solution relies on a constant skillset within the workforce, and these skills or workers must be retained to maintain a reliable solution, an issue that is of less risk to the pre-built industrial solutions.

In all cases, it can be seen that IIoT development for SME environments is possible on several levels and easily scalable to various environments. This study provides considerations for SME's when developing IIoT solutions in line with the developing Industry 4.0 environment without the need of large capital investment of new manufacturing equipment and systems. Solutions can be integrated and retrofitted into pre-existing industrial environments without the need for vast investment in digital upskilling. This study offers just three solutions but offers insight into the wide range of open-source and industrial options and solutions possible to realise the benefits of IIoT.

Any solutions developed should consider the security and the robustness of components for use in industrial settings.

References

1. Kagermann H, Wahlster W, Helbig J (2013) Securing the future of German manufacturing industry: recommendations for implementing the strategic initiative INDUSTRIE 4.0. In: Industry-science research alliance
2. Horváth D, Szabó RZ (2019) Driving forces and barriers of Industry 4.0: do multinational and small and medium-sized companies have equal opportunities? Technol Forecast Soc Change 146:119–132. https://doi.org/10.1016/j.techfore.2019.05.021
3. Kiel D, Arnold C, Voigt K-I (2017) The influence of the Industrial Internet of Things on business models of established manufacturing companies—A business level perspective. Technovation 68:4–19. https://doi.org/10.1016/j.technovation.2017.09.003
4. Müller JM, Buliga O, Voigt K-I (2018) Fortune favors the prepared: how SMEs approach business model innovations in Industry 4.0. Technol Forecast Soc Change 132:2–17. https://doi.org/10.1016/j.techfore.2017.12.019
5. Schmidt R, Möhring M, Härting R-C, Reichstein C, Neumaier P, Jozinović P (2015) Industry 4.0-potentials for creating smart products: empirical research results. In: International conference on business information systems. Springer, pp 16–27

6. Madakam S, Uchiya T (2019) Industrial Internet of Things (IIoT): principles, processes and protocols. In: Mahmood Z (ed) The Internet of Things in the industrial sector: security and device connectivity, smart environments, and Industry 40. Springer, Cham, pp 35–53
7. Sadeghi A, Wachsmann C, Waidner M (2015) Security and privacy challenges in industrial Internet of Things. In: 2015 52nd ACM/EDAC/IEEE design automation conference (DAC), pp 1–6
8. Sultanow E, Chircu A (2018) Bringing clarity to the java IoT jungle. 19:26–34
9. Strauß P, Schmitz M, Wöstmann R, Deuse J (2018) Enabling of predictive maintenance in the brownfield through low-cost sensors, an IIoT-architecture and machine learning. In: 2018 IEEE international conference on big data (big data). pp 1474–83
10. Gilchrist A (2016) Industry 4.0: the industrial Internet of Things. Springer

Intelligent Self-Protection Solution Against COVID-19

Andrew Moses, R. Jagadeesh Kannan, Tulasi Prasad Sariki, and G. Bharadwaja Kumar

Abstract With the ongoing crisis, it is important for every human to protect themselves from the virus infection. Even though the enemy is invisible and looks terrifying, it is very well possible to defend it with simple techniques. One among those techniques is to maintain a safe distance when we encounter another human. But, how do we always keep track of this distance of 6 feet apart can always seem like an illusion and there is a high possibility that we may lose track of our position. In this paper, we propose a system which will help in detecting when a human comes in contact with another person less than the recommended distance. As the solution demands advanced methodologies, the present paper proposes mask RCNN-based deep learning architecture in the computer vision domain. This algorithm is able to draw bounding boxes effectively whenever the person comes closer than the preferred social distance.

Keywords Deep learning · Computer vision · Image identification · Object detection · Convolutional neural networks · Region-based CNN

1 Introduction

Due to COVID-19 pandemic, life has become tough and a lot of constraints have been imposed for one's own safety. It is also observed that COVID-19 has become the fastest developing pandemic which the universe has ever witnessed. According to the reports of WHO, as of August 16th, 2020, a total of 21.2 million people have been infected leading to more than seven hundred thousand deaths. More than two hundred countries around the world have been affected by this deadly virus. Research is still being carried on how the virus spread is happening. For now, it is asserted that the virus spreads from person to person via physical contact like touch or handshake. Also, there were incidents reported where the individual was affected by coming in close contact less than one meter distance or in a closed room where the

A. Moses (✉) · R. J. Kannan · T. P. Sariki · G. Bharadwaja Kumar
Vellore Institute of Technology, Chennai, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_34

air circulates in a humid environment. Physical distancing has been imposed almost by every nation to combat this pandemic situation.

Situations were worst in certain zones where many people are living together especially in countries where a joint family system is followed as well as especially where the population density is very high. It was noted that the spread of the virus was very rapid if people did not follow social distancing [1]. Given such knowledge about the spread of the virus, physical distancing is the only counter-response that humanity can follow to control the spread of the virus. It is the only defense mechanism, at least until the scientific consortium finds a vaccine to handle the problem. According to Eksin et al. [2] susceptible-infected-recovered (SIR) model, the social distance parameter can be determined with the help of infected and recovered persons represented as I and R using the following model.

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta S \frac{1}{N} a(I, R), \tag{1}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = -\delta I + \beta S \frac{1}{N} a(I, R), \tag{2}$$

$$\frac{dR}{dt} = \delta I, \tag{3}$$

where β and δ are the disease infection and healing rates, respectively, and $N = S + I + R$ is the population size. They have studied two types of social distancing methods named as long-term awareness and short-term awareness and found that social distancing plays a vital role in determining the individuals becoming sensitive to the disease prevalence. In accordance with the previous study, Courtemanche et al. [3] also stated that strong social distancing has influenced the spread of COVID-19 virus in the USA. In this juncture, the present work tried to come up with an automated solution using deep learning-based techniques to alert the people when any person comes closer within one meter distance so that the spread of the virus can be minimized upto some extent.

2 Related Works

Numerous clinical experiments are taking place for finding the vaccine and medicine for COVID-19. Finding a better solution in a short span of time is not possible, so social distancing is one of the best possible solutions for minimizing the spread of virus without affecting the economy of the nation. Social distancing is most prominent in movement of pedestrians in public gatherings. Observing the suspicious activities of pedestrians in public places is possible with the help of object detection techniques.

Various works have been carried out in the literature to predict the people coming closer than the preferred social distance in public places. Yang et al. [4] proposed an active surveillance system capable of detecting distances between pedestrians and generating a warning if they are within two meters of distance. To trigger the alarm for the safe movement of pedestrians, they represented all the pedestrians into a pose vector in real-world coordinates on the ground plane and used a corresponding list of inter-pedestrian distances. They experimented with faster RCNN and YOLO V4 on three datasets to find out the social density. Nguyen et al. [5] done an extensive survey on emerging technologies for social distancing. They detailed various technologies with respect to artificial intelligence, computer vision, and other sensor-based techniques. Various technologies based on Wi-Fi, RFID, cellular for social distancing are discussed. Finally, they mentioned issues and challenges like privacy, real-time scheduling and optimization in social distancing.

Ramadass et al. [6] proposed a drone-based real-time monitoring system for observing the social distance between the people in the public places. The drone camera utilizes the yolov3 deep learning framework to find the social distance is strictly maintained or not and whether the people in the public are moving with the masks or not. The complete process is automated by means of deep learning and drones. Few works are carried out on social distancing by means of cellular phones and wearable activity trackers [7–9]. They used mobility of the mobile devices with GPS to track the movement of people in the different workplaces. Activity trackers are upgraded by augmenting the social distancing feature in their apps to generate alerts in the common places.

Prior to this pandemic various works are designated for movement of objects. Its scope includes defense, humanitarian services, insights over any given sport in the real time and even in extra-terrestrial activities—80% of their work depends on computer vision techs. The deep convolutional neural networks (CNN) nets have a requirement where the input has to be fixed size or dimension which is a hurdle when predictions are made and reduces the performance, In Spatial pyramid pooling, this hurdle is addressed [10]. In the later work, the object's search in a given image is made selective and the problem of various possible regions in recognition is discussed [11]. Later, proposals were made by increasing the efficiency of objects by the edge boxes technique originating from the Microsoft research team [12]. Another system trained portions of image for objects' existence which is a mixture of varied scales parted in different models [13]. Dataset was another issue, and also computations made from the scratch for various similar outputs were time-consuming and redundant processes were followed. To overcome such challenges, Microsoft Common Objects in Context [14] is proven to be the starting point for various researches from that time onwards. Similarly, two other researches [15, 16] focused their outcomes in various other architectures to fulfill the challenges involved in computer vision tasks.

Article tracking and object detection are comparative regarding usefulness. These two errands include recognizing the item and its area. However, the main contrast between them is the sort of information that you are utilizing. Article detection manages pictures while object tracking manages recordings.

3 Design and Architecture

The end goal of this work is to identify human objects who did not wear the mask but come closer within the specified distance and draw bounding boxes around them in a given image. The preprocessing involves slicing individual frames from a given video clip and storing them as images. Similarly, the post-processing involves stitching the processed output images into a video clip.

The present work would like to develop a model using the image database containing images that are used for pedestrian detection which is open sourced and also labeled dataset. Penn Fudan dataset is a collection of 170 images with annotations such as bounding box as well as pixel level masks. Even though the dataset is very small, it is one of the few quality datasets which are currently available in the open forum which helps in human object detection with segmented annotations.

Mask RCNN is the algorithm proposed for object detection in this work. Mask RCNN is a simple extension of faster RCNN. Faster RCNN has two outputs for each candidate object, i.e., a class label and a bounding box offset [16]. Along with these two outputs, mask RCNN adds another output called the object mask. Mask RCNN can be built with varying flavors. The proposed method utilizes residual networks as its backbone. If we look deeper into mask RCNN, it is a block which combines faster RCNN and a fully convolutional network. Faster RCNN has two stages: First stage is the region proposal network (RPN) where it gives the model to choose the regions in a given image selectively without iterating through all of them, and this process saves a lot of time and resource. Earlier versions of the fast RCNN approach brought in approximately 2000 regions for a given image and would take the model 5–10 s on an average computing machine to classify the bounding box with its respective labels. This shortcoming was overcome by the RPN method. The second stage of faster RCNN is the label classification of proposed bounding box regions generated by regression techniques. The proposed regions are generated by the RPN. As discussed, the faster RCNN gives output a bounding box and a corresponding label to it, but the pixel level segmentation is achieved by the fully convolutional network (FCN). Thus, the proposed model had its backbone as residual network (ResNet) with feature pyramid network (FPN) pre-trained on common objects in context (COCO) dataset which had 91 classes. But the objective which we solve here has only two classes: background and the human object. Hence, the model while loading into memory was modified by removing the predictor heads with a new modified faster RCNN predictor head which has only two channels as output. Similarly, the mask region of interest (ROI) predictor was also updated accordingly. During the training phase, the standard stochastic gradient descent algorithm was utilized with a learning rate (LR) scheduler. The LR scheduler helps in reducing the learning rate as and when the number of epochs increases with a step size measure. This helps the model to take smaller steps to converge at the global minima without overshooting. The residual network has two types of connections. Scenarios where the input dimensions are same as the output dimensions, a typical connection block is followed, while when the input is different that the output dimension, the entries are padded with extra zeros using a 1×1 conv layer to match the output dimension using the formula given in Eqs. (4) and (5).

$$y = F(x, \{W_i\}) + x \tag{4}$$

$$y = F(x, \{W_i\}) + W_s x \tag{5}$$

x: input vectors of a considered layer
y: output vectors
W_i, W_s: weights, linear projection
F: residual function.

Now, we shall also need to understand the underlying intermediate results of the above techniques. It is evident that certain regions might be overlapped in different sizes and may correspond to different classes in the same given image. This characteristic of the result is because of the way in which the problem is structured. Deliberate steps can be taken to overcome this behavior, if it results in certain problems over our objective getting fulfilled.

Take an image and slice it into sections (let us say 10×10 section) and then take each slice and give it as input to the convolutional neural network (CNN). The neural network has multiple filters having the capability to extract the key features from the image which was passed into it. Then, the extracted features are fed into the regressor which can help in classifying the objects. The drawback here is if a large number of regions are proposed, all the regions have to be considered for the classification purpose. This is a very time-consuming task. The paper will dwell on addressing this drawback by mainly focusing on effective steps which can be taken to reduce the burden by reducing the large number of region slices.

So, how would we make this a non-tedious assignment? Here, the key is to discard the slices which probably might not have the target item. Those slices must be disposed of. Now, the step which involves segregating the slices that have an item or not is called as region proposals. The proposition is having a higher order of interest of holding an article. Numerous region proposal calculations were explored by researchers to choose a section of interest in a given image, popularly known as region of interest (ROI). RCNN represents region-based convolutional neural network. It utilizes one of the outer area proposition calculations to choose the ROI. The following steps are involved in this method (1) Take a picture. (2) Select ROI utilizing outside locale proposition calculation (3) For every locale: Provide the slice to the CNN, then the network would extract the highlights from the network. (4) Highlighted features are sent to the classifier. The anticipated slices may be covering in certain places and also maybe shifting in its area across the image. Hence, a suppression technique is utilized to overcome the markers over the sliced area. This relies on the technique known as maximum non-suppression. The bounding boxes are exhibited by the highlights according to a score. The score name is called intersection over union (IOU). Region-based CNN's engineering was one among the best performing models until recently. But as highlighted earlier, the time taken for each test picture during the derivation expends about 50 s; this is because of the resulting quality of feed forwarding of CNN which also includes the extraction step. The later upgraded model was named fast RCNN. The key upgrades were done in the feature extraction method in the sliced regions of interest. When it comes to region-based CNN, extraction of features happens for every distinct proposition while, in fast RCNN extraction happens just a single time for a unique picture. At that point, the significant ROI highlights are picked depending on the area of the locale recommendations. These area recommendations are developed before passing a picture to CNN. Then each ROI is passed along with the features to a classification module. During this surmising process, the architecture consumes almost two seconds for every test sample and this process to be exponential quicker than the region-based CNN. The detailed reason behind the quickening would be the adjustment in the extraction component of ROI. Faster RCNN has this new component known as region proposal network (RPN), i.e., the neural network component which comes in the place of discarding the slices of area proposition calculations. The newly introduced network finds a way out to choose the slices of area which may be of interest. This selection process saves a lot of computation time. Thus, it is indeed a great deal when compared to the previous model fast region-based CNN. Faster region-based CNN utilizes approximately 0.2 s for each sample image when passed through the induction step and is also proven to be exponentially quicker than the previously discussed model that is fast region-based CNN. The Pytorch version of mask RCNN is utilized in this proposed work. Figure 1 shows the underlying mask RCNN architecture.

Physical distancing is our objective and that is to maintain distance between individuals. We have used the OpenCV technique as shown in Fig. 2. The tool utilizes the perspectives of computer vision techniques. It is a library for recognizing objects in a given image.

4 Implementation Details

As the objective of this project was to find the precise location of the person in a given image, a dataset with annotations such as bounding box and the mask dataset was needed. The experiment utilized Penn Fudan dataset which consists of pedestrians' images sliced from CCTV footage from two universities. The dataset consists of a very small number of images that is 170 images, but it has more than three hundred instances of pedestrians. Each image had a bounding box info and a pixel level mask segmented info attached to it. The mask had 0 as the background and starting from 1 as the masked instance of a person in the image. If more than one person is found in the image the second person would be masked with 2 and so on. For a detailed workflow of the proposed solution, refer Fig. 3.

The proposed implementation was initially fine tuned with a big deeper net model that is ResNet50 (where 50 stands for the number of layers in the network) [17]. As the network was deep, it consumed all the available resources which we had. We had 6 gig of NVIDIA graphic processing unit (GPU) with 1920 CUDA cores and

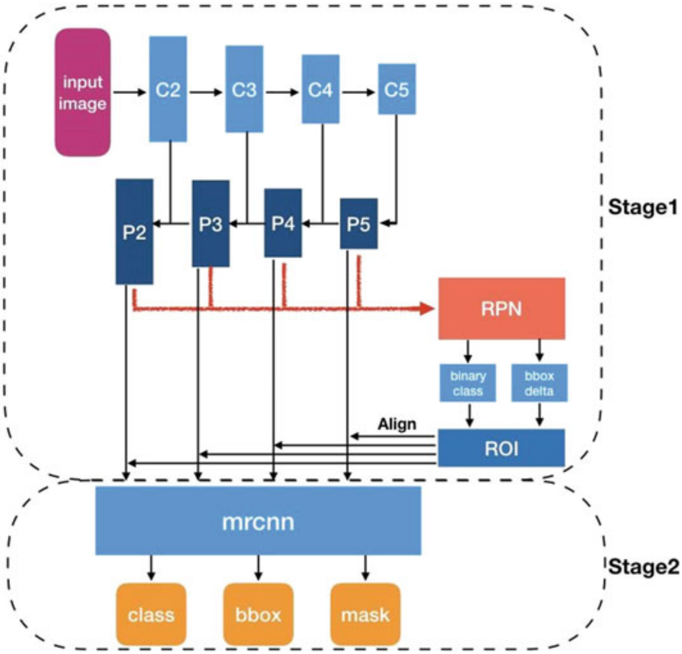


Fig. 1 Two-stage architecture of mask RCNN. *Source* https://medium.com/@tibastar/mask-r-cnn-d69aa596761f

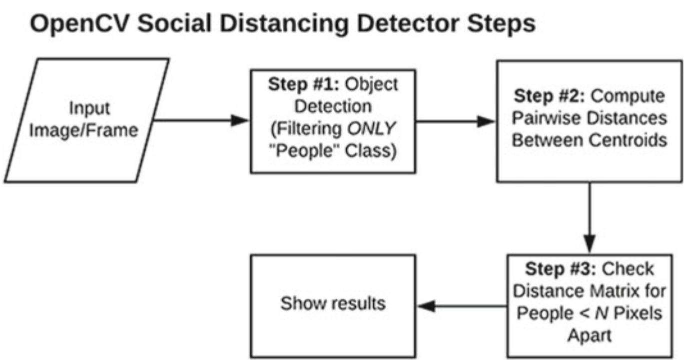
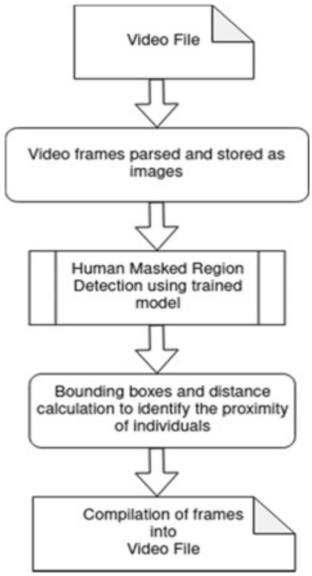


Fig. 2 Steps involved in an OpenCV-based social distancing application. *Source* https://www.pyimagesearch.com/2020/06/01/opencv-social-distancing-detector/

240 tensor cores. The experiment took 10–15 min of runtime for 10 epochs. Later, steps were taken to modify the backbone from ResNet to a different network. But due to mismatches in stacking the mask annotations dataset to the different backbone network it failed. The alternate backbone network chosen was mobilenetv2 keeping in mind to execute the trained solution on personal devices like mobile where computing

Fig. 3 Workflow of the proposed solution



resources are minimal. Then, the experiment was continued with ResNet18 and ResNet34 again where the network is not very deep and requires lesser resources to compute.

Figure 4 and Table 1 show the results procured by the experiments. It is noted that the deeper network has exponential increase in performance when compared

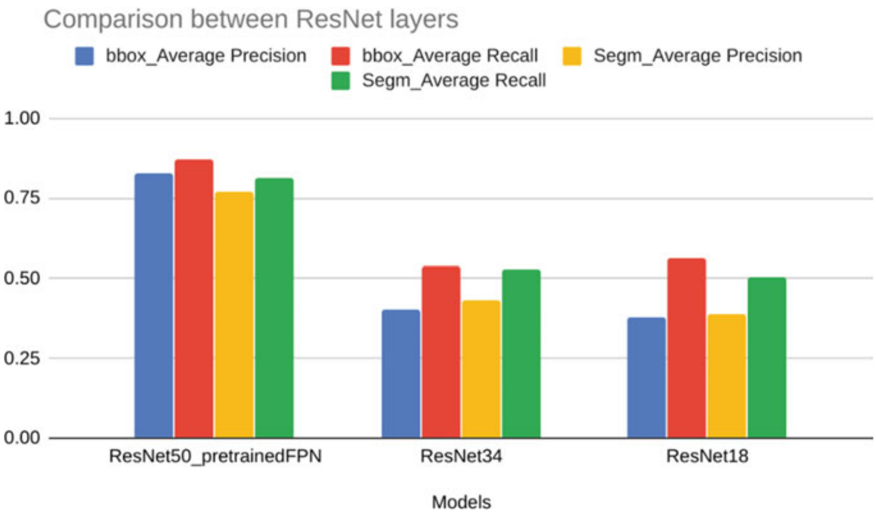


Fig. 4 Performance of various ResNet models

Table 1 Detailed results across various models

Model	BBox Avg. precision	BBox Avg. recall	Segm Avg. precision	Segm Avg. recall
ResNet50	0.829	0.873	0.769	0.811
ResNet34	0.403	0.539	0.431	0.527
ResNet18	0.381	0.56	0.387	0.502

with its shallow counterparts. Even though ResNet34 is nearly double to the size of ResNet18, it did not show a considerable increase in performance. Of course in certain parameters a slight increase was seen. But it is more equivalent to ResNet18. Hence, if the use case involves minimizing the network due to resource constraints, ResNet18 would be a good choice over ResNet34. The result metrics being average precision of instances where only Intersection over union (IoU) is more than 50% is considered a correct prediction. Hence, there were instances which are predicted correctly, but since the IoU value is less than 50% it won't be considered a correctly predicted case. Average recall gives the insight of how the model performs while classifying the positive cases. In certain models, the performance of the model with mask region over the bounding box region has better results, and this might be because of the small size of the test dataset. Therefore, on taking an overview the model tends to perform better in bounding boxes when compared to masked region predictions. Figure 5 shows two frames of the experiment video clip, with blue boxes highlighting an unsafe distance between individuals, whereas the red box marks the distance as safe.

Fig. 5 Experimental results—blue boxes being high risk while red have been marked as safe

5 Conclusion

The proposed method paves a way to address the social distancing rule to be implemented in an automated fashion. The research also compares and gives more informed recommendations to choose the appropriate state-of-the-art model when it comes to solve human object detection which is the core computer vision problem while building an automated system which helps in social distancing norms in public. In the future, we would like to develop an IOT-based solution in our campus to alert the students whenever they come closer than the preferred social distance. This helps in preventing spread of viruses and other infections in our campus.

References

1. Lewnard JA, Nathan CL (2020) Scientific and ethical basis for social-distancing interventions against COVID-19. Lancet Infect Dis 20(6):631
2. Eksin C, Paarporn K, Weitz JS (2019) Systematic biases in disease forecasting–the role of behavior change. Epidemics 27:96–105
3. Courtemanche C et al (2020) Strong social distancing measures in the United States reduced the COVID-19 growth rate: study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States. Health Affairs, 10–1377
4. Yang D et al (2020) A vision-based social distancing and critical density detection system for COVID-19. arXiv preprint arXiv:2007.03578
5. He K, Zhang X, Ren S, Sun J (2014) Spatial pyramid pooling in deep convolutional networks for visual recognition. In: European conference on computer vision (ECCV)
6. Ramadass L, Arunachalam S, Sagayasree Z (2020) Applying deep learning algorithm to maintain social distance in public place through drone technology. Int J Pervasive Comput Commun
7. Pépin JL et al (2020) Wearable activity trackers for monitoring adherence to home confinement during the COVID-19 pandemic worldwide: data aggregation and analysis. J Med Internet Res 22(6):e19787
8. Sun S et al (2020) Using smartphones and wearable devices to monitor behavioural changes during COVID-19. arXiv preprint arXiv:2004.14331
9. Tripathy AK et al (2020) EasyBand: a wearable for safety-aware mobility during pandemic outbreak. IEEE Consumer Electron Mag
10. Uijlings JR, van de Sande KE, Gevers T, Smeulders AW (2013) Selective search for object recognition. Int J Comput Vis (IJCV)
11. Zitnick CL, Dollar P (2014) Edge boxes: locating object proposals from edges. In: European conference on computer vision (ECCV)
12. Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D (2010) Object detection with discriminatively trained partbased models. IEEE Trans Pattern Anal Mach Intell (TPAMI)
13. Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollar P, Zitnick CL (2014) Microsoft COCO: common objects in context. In: European conference on computer vision (ECCV)
14. Zhu LNJ, Chen X, Yuille AL (2015) DeePM: a deep part-based model for object detection and semantic part localization. arXiv:1511.07131
15. Lenc K, Vedaldi A (2015) R-CNN minus R. In: British machine vision conference (BMVC)
16. He K, Gkioxari G, Dollár P, Girshick R (2018) Mask R-CNN. arXiv:1703.06870 [cs.CV]
17. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. arXiv:1512.03385v1 [cs.CV]

Detecting Suicidal Ideation from Online Texts

Snigdha Ramkumar, Tulasi Prasad Sariki, G. Bharadwaja Kumar, and R. Jagadeesh Kannan

Abstract According to the World Health Organization (WHO), suicide is a global issue and around 800,000 people commit suicide every year. In addition to that, WHO also reports that for every person who dies as a result of suicide, there may be more than 20 who survive an attempt. Early intervention is the most efficacious method to prevent suicides, and this involves understanding and detecting suicidal ideation. With the growth and popularity of the Internet, people have started to discuss suicidal tendencies on online platforms. By analyzing the differences between the way suicidal and non-suicidal people text online, it is possible to detect texts that contain suicidal ideation. In the present work, various natural language features of online posts, along with predicted gender and age of the person, have been leveraged to detect suicidal ideation in the text. The dataset was collected from a popular social platform called Reddit. Various supervised learning algorithms have been applied and compared. The result analysis of the present study substantiates that the models proposed are very promising in predicting suicidal ideation from the text

Keywords Suicide · Ideation · Prevention · Supervised learning

1 Introduction

Day by day, the number of people falling victim to suicide is increasing, and the need to tackle this problem is vital. According to the World Health Organization, nearly one person dies every 40 s, a rate which is set to increase in the coming years

S. Ramkumar (✉) · T. P. Sariki · G. B. Kumar · R. J. Kannan
Vellore Institute of Technology, Chennai, India

T. P. Sariki
e-mail: tulasiprasad.sariki@vit.ac.in

G. B. Kumar
e-mail: bharadwaja.kumar@vit.ac.in

R. J. Kannan
e-mail: jagadeeshkannan.r@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_35

[1]. In addition to that, WHO also reports that for every person who dies as a result of suicide, there may be more than 20 who survive in such an attempt. Suicides are usually preventable with timely, evidence-based, and often low-cost interventions. Some countries have tried to prevent suicide by deeming it to be illegal and by restricting access to firearms, toxic chemicals, and other materials that may be used to commit the act. However, these measures are not able to address the underlying issues.

Suicide is a taboo topic in many societies. Despite being a major public health concern, many people are unwilling to openly discuss it, and hence, prevention has not been adequately addressed. People commit suicide due to a variety of factors: feelings of isolation and hopelessness, mental illnesses, substance abuse, trauma, personal loss, unemployment, physical illness, or previous suicide attempts. These factors can cause people to think about killing themselves and having these suicidal thoughts is termed suicidal ideation. Suicidal ideation includes fleeting considerations about killing oneself to detailed plans on how to go about it. Apart from those who have committed suicide, many more have attempted it, with an even greater number thinking about it. According to the data from the 2017 National Survey on Drug Use and Health, 4.3% of adults in the USA had thoughts about committing suicide, with 0.6% of all adults attempting the act [2]. This is already a huge number of people, despite the survey not taking into account minors and other less-developed countries. Since suicide occurs concurrently with suicidal ideation, detecting these thoughts and addressing the underlying factors would be an effective way to prevent suicide.

Intervention, even by just reaching out, actively listening, and offering support can help in this regard and is what workers for suicide crisis hotlines are trained to do [3]. Therapy, counseling, and medication can also help people with suicidal ideation. However, these are only possible when either the suicidal person actively seeks help, or if another concerned person brings the issue to light.

Data obtained from therapy has helped understand these thoughts and work toward preventing further suicides. Questionnaires like the patient health questionnaire and the general health questionnaire were administered to people who had undergone psychiatric treatment. Studies have been conducted to analyze these responses using natural language processing [4] and classifying these responses in an attempt to predict suicidal tendencies using supervised learning techniques, which rely on already labeled data [5]. However, collecting and digitizing such data is an expensive and time-consuming process; furthermore, it relies on established relationships between the patient and the test administrator. Also, the data obtained cannot be used to generalize suicidal ideation on a large scale.

Additionally, according to the National Institute of Mental Health, around 50% of mentally ill Americans do not receive treatment, and this percentage only increases in less-developed countries [6]. While therapy helps, it can be inaccessible to many, due to cost, location, or social stigma [7]. In addition to that, suicide is deemed illegal in some countries, and hence, some patients may not seek help or may not be entirely truthful when responding to questions [8]. There is an increasing need to detect suicidal ideation as early as possible using more effective ways to prevent suicide. This is where online content would be useful.

It was found that people tend to be more open about their thoughts on the Internet when compared to other more direct forms of communication [9]. The growing popularity of the Internet has also allowed multiple communities for like-minded people to gather together and share their voices. With the increasing usage of smartphones and the Internet, there is a tendency for people to talk about suicide, suicidal ideation, and suicide attempts online. Users also can remain anonymous on some of the social platforms. Anonymity increases candidness, and people may write about why they want to commit suicide, discuss their plans for attempting this, or may even ask other users for tips. Using online content to identify patterns and differences between how suicidal and non-suicidal people write can help detect suicidal ideation. In the present, we used the data collected from one such website called Reddit and applied several machine learning algorithms to detect patterns in the texts containing suicidal ideation.

For the current work, we created a new novel dataset with more than 100,000 entries of suicidal and non-suicidal posts, forming a roughly balanced dataset taken from Reddit. Apart from using just the title and text, we make use of additional metadata to examine further links between suicidal ideation and online interaction. According to research, males are more likely to commit suicide, while females have higher rates of suicidal ideation. Age is also a factor, with most suicides occurring in the age group 15–29 [10]. Our research involves predicting the age and gender of the user to analyze the risk. We wish to examine whether these demographic factors allow us to make better predictions and whether these features online reflect real-world trends.

The user's number of posts and their overall impact on the community (positive + negative) is also extracted. Additionally, the time of posting is also noted. Studies have shown that depression peaks in the colder months [11] and that the rates of suicides increase in the summer. Links between this, along with the time of posting, and suicidal ideation are examined. Categorical features are also obtained, and supervised learning is used to detect suicidal ideation.

2 Related Works

According to the World Health Organization, a person commits suicide every 40 s, and it is the second leading cause of death in young people [1]. The American psychological association defines suicidal ideation as the thought of serving as the agent of one's own death [12]. Detecting suicidal ideation at an early stage would be useful in preventing these deaths. Multiple factors may cause a person to commit suicide, common ones include mental disorders, family conflict, physical illnesses, social isolation, unemployment, previous attempts, and seasonal variations.

Age and gender may also play a role, with men more likely to commit suicide. However, females are more likely to attempt suicide and have higher rates of suicidal ideation [10, 13]. Clinical settings were initially studied to provide accurate information to identify underlying patterns. Studies were conducted to examine the link between mental illnesses and suicidal ideation [4]. Online questionnaires and telephonic interviews were conducted to detect suicidal intent as well as underlying mental issues, and the associations between these results were analyzed using logistic regression. It was found that comorbidity disorders greatly increased the risk of suicidal ideation. With an understanding of prevalent factors, classical methods of preventing suicide in psychiatric settings using questionnaires (like PHQ-9 and GHQ-12) are used to prevent suicides [14]. Using text-based features to identify a link between written language and users' suicidal thoughts has proven to be effective. In a study conducted by Cook et al. [5], the GHQ-12 questionnaire was administered to recently discharged psychiatric patients. Structured questions with Likert scales and choice-based answers were followed by an open-ended question regarding the patients' feelings. *N*-grams were used to analyze this and were compared against the structured questions to identify underlying patterns. These studies relied on previously established relationships in a clinical setting. To expand the sample size for detecting ideation, text data available online was used.

Mental treatment sometimes involves asking patients to use journals and diaries to keep track of their moods, and these can be a rich source of information into understanding how a suicidal user may articulate their thoughts. For automatic detection, it is more convenient to use social media for raw data as opposed to digitizing these writing samples. Coppersmith et al. used publicly available data from Twitter for this purpose [15]. Users who had previously attempted suicide were manually identified based on certain keywords and filtered according to their history. Their previous tweets were retrieved and analyzed using character n-grams and LIWC for common word analysis [16]. They also made use of the World Well Being Project [https://wwbp.org/] to predict the age and gender of the users [17]. This, along with geographic location, was sometimes publicly available.

As online content can be anonymous, methods for predicting age and gender can be used. Data from a Belgian social networking platform with posts in Flemish-Dutch, which included user data such as their age, gender, and the location was used to predict age and gender [18]. Predictions would help ensure online safety, such as protection against impostors or pedophiles. Gender and age were considered together, and word and character unigrams, bigrams, and trigrams were used for the prediction. Slang was found to be a good indicator and worked well with SVM. Miller et al. worked with manually labeled tweets to predict the gender using character n-grams with naïve-Bayes and perceptron [19]. Cheng et al. also predicted gender using online content, more specifically, emails [20]. Statistical features such as the number of words, characters, alphabets, digits, and special characters were considered, along with the vocabulary and parts-of-speech. LIWC was also used, and with Naïve-Bayes, Logistic regression, decision trees, and SVM models, the gender could be predicted.

Shing et al. created the University of Maryland Reddit Suicidality dataset and used this data to assess suicide risk [21]. To enhance suicidal ideation detection in user-generated content, Shaoxiong Ji et al. also used data from Reddit [22]. Data could be automatically classified using the communities ('subreddits'), and texts and titles of anonymous users were extracted and classified. These texts were analyzed using multiple features: statistical, parts-of-speech, bag-of-words, TF-IDF, topic identification using LDA [23], LIWC, and word embedding. The machine learning models included support vector machines, random forest, gradient boosting decision trees, XGBoost, and LSTM. This was compared against publicly available, manually annotated Twitter data. This was expanded on by Tadesse et al. by building an LSTM-CNN model with tenfold cross-validation for hyperparameter tuning [24]. Top 200 occurring unigrams and bigrams from both suicidal and non-suicidal posts were also compared. Making use of non-clinical data allowed researchers to identify broader patterns and identify ideation among a larger population.

Sentiment analysis has also been evolving over the years, with a number of powerful tools and packages coming into existence. Along with Latent Dirichlet Allocation and LIWC, Tadesse et al. made use of this in detecting depression in the online forum, Reddit [25]. Sentiment analysis has been used in detecting and analyzing changes in mental health, especially on data from social media [26]. Hybrid approaches to sentiment analysis have been studied, with the possibility of using fuzzy sets as well [27]. Sentiment analysis has evolved from predicting positive and negative reviews from labelled datasets to understanding more complexities from sentences used in everyday language.

3 Design and Architecture

The overall architecture of the proposed model is depicted in Fig. 1.

3.1 Data Collection

Suicidal Ideation Detection

For our research, we aim to expand on non-clinical methods to detect suicidal ideation. We introduce a new dataset from the social news site and platform, Reddit. Users sign up with an email address and are identified only by their username; most users choose to stay anonymous. The site comprises of different communities, ranging from news, politics, books, and movies to memes, dating, and legal advice. These communities are termed “subreddits.” The platform allows its users to generate and share content such as text, images, gifs, videos, and links.

One such subreddit is “r/SuicideWatch” [https://www.reddit.com/r/SuicideWatch/], a community where users generally post about their suicidal thoughts. Common themes involve discussing why they want to commit suicide and talking about previous attempts. The content is completely user-generated and exhibits suicidal intentions. This was extracted using Pushshift.io, which is a RESTful API that allows for comments and submissions from Reddit to be obtained [28]. The following details

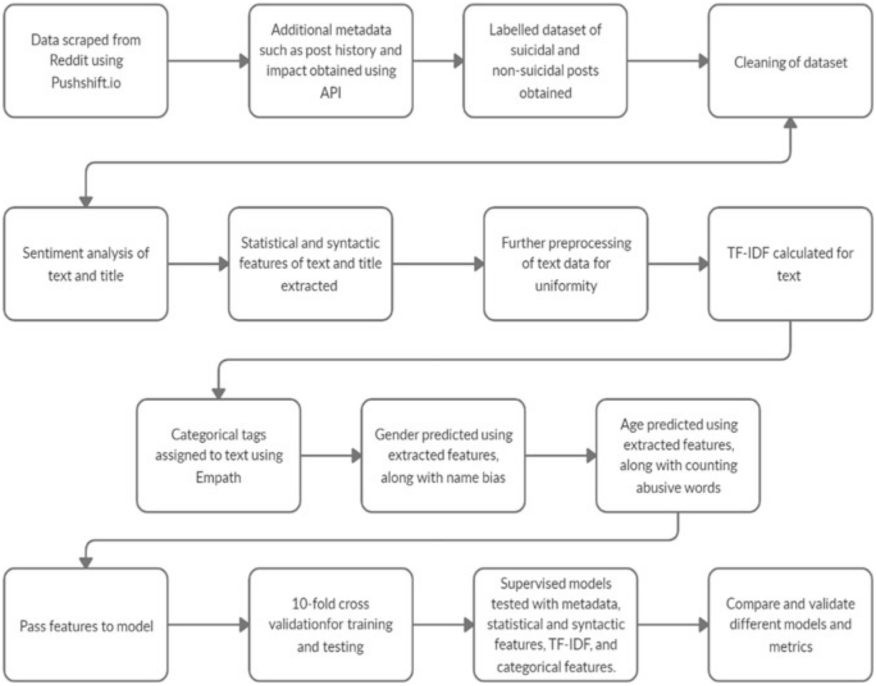


Fig. 1 Overall architecture of the proposed system

were obtained, for the year of 2019: *Post ID, Title, Text, URL, Author, Score, Publish Date and Time, and Total Number of Comments.*

Similarly, posts with no suicidal tendencies were extracted from other popular subreddits, r/books, r/jokes, r/legal advice, r/casual conversation, and r/college. All the content was user-generated and posts containing links to external websites, photos, and videos were excluded. A balanced dataset of 57,900 suicidal and non-suicidal posts each was obtained. Few sample sentences for each suicidal and non-suicidal is shown in Table 1.

Table 1 Sample sentences from the prepared dataset (suicidal/non-suicidal texts)

Suicidal posts	(i) "I screwed up my life. I can only think of killing myself. Can I just end my life?" (ii) "There is no end to this suffering. So much medication and therapy, I cannot take it anymore"
Non-suicidal post	(i) "I have found that a lot of my hobbies and interests are developed via books" (ii) "What are your thoughts on the latest episode?"

This paper also takes into account the user's history on the website as well as their overall impact on the platform, measured by their "karma." This was extracted from the username (Author) using the Pushshift,io API.

Gender
To predict the binary gender, comment data was first collected from two subreddits: r/AskMen and r/AskWomen. The subreddit comprised of answers to various questions, and the responses to these were extracted. Comments on Reddit sometimes consist of a "Flair," which is sort of an identifying tag for the comment. Flairs on these subreddits included the genders "Male" and "Female," and comments with this flair were retained while others were discarded. The flair formed the label for this collection of texts.

Age
Ages were classified according to the generations: Millennials, Generation X, and Baby Boomers. Data for these labels were extracted from age-based subreddits like r/YoungAdults, r/college, r/Over30Reddit, and r/AskOldPeople.

3.2 Preprocessing and Feature Extraction:

To enhance the quality of the model and subsequent classification, preprocessing is a vital step. Missing values, noise, and outliers can negatively impact predictions and hence needs to be cleaned. After forming our initial dataset, the data was first cleaned, with null values being imputed using average values or having the entire row dropped. These rows included posts with non-text features, like images and videos.

Statistical Features
Using simple splitting of the texts and by examining for certain words and values, the number of words, characters, sentences, stop-words, special characters, alphabets, and numbers is obtained for both the title and the text following it. These features would be helpful in identifying patterns distinguishing suicidal and non-suicidal users.

Syntactic Features
Parts-of-speech [POS] tagging was done on the text body to capture grammatical variations. The text was converted to lowercase tokens, and each token was tagged. The total number of these tags were counted. Some of the features included nouns, verbs, adjectives, adverbs, conjunctions, and determiners.

Preprocessing
For conducting further analysis of the text, it needs to be further preprocessed to ensure uniformity and reduce any unnecessary complexities. The text is converted to

lowercase, and punctuation and symbols are removed. The text is then lemmatized [29], and finally, all stop-words are removed.

Sentiment
The polarity and subjectivity of the title and the accompanying text body are obtained. Subjectivity is the measure of how factual or opinionated the text is, with -1 being highly objective and $+1$ being highly subjective. Polarity captures the intensity of emotion, with -1 being negative and $+1$ being positive.

Category
The text is classified according to different categories such as "family," "health," "death," "disgust," and "joy," and the differences between suicidal and non-suicidal posts are noted. For instance, posts containing suicidal ideation tend to contain more negative emotions, with posts three times more likely to contain words relating to fear, death, violence, and shame. On the other hand, posts not containing suicidal ideation contain relate to more mundane, everyday occurrences such as work, pets, vehicles, politics, science, and food. Non-suicidal texts also included more positive or artistic content, such as art, fashion, wealth, and superheroes. These categories are obtained using Empath [30]. Empath generates lexical categories from the seed word, using neural embedding with a skip-gram network to identify related words. 196 categories are obtained for this dataset. The frequencies of various category occurred in both suicide, and non-suicide contexts are shown in Table 2.

TF-IDF
Term frequency and inverse document frequency of the text bodies are measured, to identify how relevant a word is to a document in a collection of documents. The 1000 most common terms along with their weights are used for classification.

For term i in document j:

$$w_{i,j} = tf_{i,j} \times \log(N/df_i)$$

$tf_{i,j}$ = occurrences of i in j
df_i = number of documents containing i
N = total number of documents.

Gender
The gender (binary) of the original poster is predicted using statistical, POS, and TF-IDF features. Usernames are also taken into account, with commonly occurring feminine themes (such as "Princess," "Girl," or "Mrs") and masculine themes (such as "King," "Master," or "Sir") being assigned a bias. For instance, "Queen_GHive" is more likely to be a female user while "SirC0rv0Attan0" is more likely to be male. Using supervised learning, the user's gender could be approximated with a 76% accuracy.

Table 2 Comparison of category frequency between suicidal and non-suicidal posts

Category	Suicidal (S)	Non-suicidal (NS)
Aggression	422.5114515	137.3291595
Anger	281.6296718	53.94599622
Animal	82.96879976	218.2746191
Attractive	138.8616971	221.4848071
Body	890.1018551	464.2089851
Children	1026.711287	981.3833341
Crime	407.402506	288.989833
Death	1684.902677	401.3388266
Emotional	754.5121286	259.409868
Envy	217.9078767	48.19206531
Fight	406.4309799	175.7817617
Friends	988.9213271	752.718057
Hate	1213.433456	287.3940133
Help	720.3910234	439.4044486
Injury	702.8472431	271.2126657
Kill	862.2858965	135.638744
Love	974.7530245	440.7328187
Messaging	239.2098075	509.0716925
Musical	103.3208845	237.0139734
Philosophy	37.12768108	103.1571624
Reading	265.5422294	1241.972947
Sadness	1013.695174	205.0424478
School	346.2648777	1256.273138
Science	82.91227208	429.3559075
Shame	1490.857814	399.0663526
Suffering	1609.475045	308.8747392
Swearing_terms	513.4920365	220.9371358
Technology	80.55812567	254.3235745
Torment	360.6838217	63.60695699
Violence	1773.5791	598.1974871

Age

The generation to which the user belongs to is predicted using statistical, POS, and TF-IDF features, along with calculating the number of emojis, exclamation marks, and other special characters. Abusive language is taken into account, with the number of swear words being counted. The multiclass classification was done with 74% accuracy.

4 Implementation

A combination of different features is tested with different machine learning algorithms, and the results are analyzed with the existence of suicidal ideation being used as the dependent variable. Here, tenfold cross-validation has been considered for different metrics. The machine learning models considered in the present work include simple algorithms like Naïve-Bayes and logistic regression, as well as models implementing ensemble methods like random forest, extreme gradient boosting (XGBoost), and light gradient boosting (LightGBM).

Naïve Bayes, logistic regression, and random forest methods have been used as benchmarking algorithms because they have been proven to work well with text data in the literature [29]. Even though, random forest works well, bagging algorithms like random forest can only reduce the variance. Hence, boosting algorithms like XGBoost & LightGBM have been proposed. In spite of the fact that XGB can reduce the bias, it is well known that the training generally takes a long time because of the fact that trees are built sequentially. However, LightGBM is capable of performing equally good with a significant reduction in training time when there are large number of features as compared to XGBoost.

5 Conclusion

The work currently focused on individual posts, and the suicidal tendencies detected from language preferences and text features. By using a very large dataset of completely user-generated content, the model yielded effective results that can be expanded to other social platforms. All the experimented models along with their accuracies are depicted in Table 3.

Apart from the additional metadata such as the time of posting and the overall impact that the user has made on the platform, we also introduced age and binary gender into our dataset. According to WHO, age and gender greatly impact the rates of suicide, with a larger number of males committing suicide than females, and with people in the age group of 15–29 being the most prone to committing this act. However, we observed that these do not reflect real-world statistics on suicide when dealing with suicidal ideation on social platforms where anonymity is preserved. There are only marginal differences in the number of suicidal and non-suicidal posts concerning age and gender. On excluding them from our model, we observe a slight increase across all the scores.

With the features obtained, our state-of-the-art model has been able to predict suicidal ideation accurately. The results obtained are better than those that do not take these additional features into account. Using light gradient boosting, the model effectively handles large instances of data with high speed and accuracy. False negatives would be especially detrimental, and the model minimizes this with an AUC score of 96.10. On identifying texts with suicidal ideation, further analysis based on

Table 3 Comparison of different models with different features

Method	Features	K-fold cross (tenfold)				
		Accuracy	Precision	Recall	F-score	AUC
Naïve Bayes	Statistics	65.18	60.17	94.44	73.51	64.48
	Statistics + POS	66.87	61.64	93.29	74.23	66.24
	TF-IDF (unigram)	84.37	80.54	91.56	85.71	84.21
	TF-IDF (bigram)	66.87	61.64	93.29	74.23	66.24
	Statistics + POS + TF-IDF	67.44	61.97	94.07	74.72	66.81
	Statistics + POS + TF-IDF + empath	67.37	61.87	94.36	74.74	66.73
Random forest	Statistics	90.26	92.69	87.89	90.23	90.32
	Statistics + POS	91.23	93.33	89.24	91.24	91.28
	TF-IDF (unigram)	89.03	87.54	91.59	89.52	88.96
	TF-IDF (bigram)	91.23	93.33	89.24	91.24	91.28
	Statistics + POS + TF-IDF	95.62	95.85	95.58	95.71	94.42
	Statistics + POS + TF-IDF + empath	95.53	95.40	95.86	95.64	95.52
XGBoost	Statistics	89.69	92.45	86.95	89.62	89.76
	Statistics + POS	90.53	93.16	87.95	90.48	90.59
	TF-IDF (unigram)	85.29	87.91	82.62	85.17	85.35
	TF-IDF (bigram)	90.53	93.16	87.95	90.48	90.59
	Statistics + POS + TF-IDF	94.37	96.56	92.29	94.37	94.42
	Statistics + POS + TF-IDF + empath	95.15	96.84	93.56	95.18	95.19
Logistic regression	Statistics	85.23	83.75	88.25	85.94	85.16
	Statistics + POS	86.55	85.55	88.69	87.09	86.5
	TF-IDF (unigram)	90.05	89.64	91.07	90.35	90.02
	TF-IDF (bigram)	86.55	85.56	88.69	87.09	86.51
	Statistics + POS + TF-IDF	89.78	89.27	90.97	90.11	89.75
	Statistics + POS + TF-IDF + empath	90.38	90.05	91.28	90.66	90.36
LGBM	Statistics	90.76	92.58	89.02	90.76	90.79
	Statistics + POS	91.8	93.23	90.5	91.84	91.83
	TF-IDF (unigram)	88.3	89.09	87.82	88.45	88.31
	TF-IDF (bigram)	77.09	88.15	63.65	73.92	77.37

(continued)

Table 3 (continued)

Method	Features	K-fold cross (tenfold)				
		Accuracy	Precision	Recall	F-score	AUC
	Statistics + POS + TF-IDF	95.86	96.93	95.92	94.93	95.91
	Statistics + POS + TF-IDF + empath	96.08	96.94	96.13	95.33	96.10
LGBM	Excluding age and gender	96.29	97.03	96.34	95.65	96.30

that user could be identified: the communities they are a part of, their history, and their usage trends. User profiling on the detection of suicidal tendencies would yield more information needed to prevent suicide.

References

1. World Health Organization (WHO) (2020) Age-standardized suicide rates (per 100 000 population), 2020. Global Health Observatory (GHO) data. http://www.who.int/gho/mental_health/suicide_rates/en/. Accessed 17 Aug 2020
2. McCance-Katz EF (2019) The national survey on drug use and health: 2017. Substance abuse and mental health services administration. Accessed 7 May 2019
3. Mokkenstorm JK, Eikelenboom M, Huisman A, Wiebenga J, Gilissen R, Kerkhof AJ, Smit JH (2017) Evaluation of the 113Online suicide prevention crisis chat service: outcomes, helper behaviors and comparison to telephone hotlines. Suicide Life-Threat Behav 47(3):282–296
4. Cook BL, Progovac AM, Chen P, Mullin B, Hou S, Baca-Garcia E (2016) Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. Computational and mathematical methods in medicine 2016
5. Alpaydin E (2020) Introduction to machine learning. MIT Press
6. National Institute of Mental Health (NIMH), https://www.nimh.nih.gov/index.shtml
7. Chekroud AM, Foster D, Zheutlin AB, Gerhard DM, Roy B, Koutsouleris N, Chandra A et al (2018) Predicting barriers to treatment for depression in a US national sample: a cross-sectional, proof-of-concept study. Psychiatr Serv 69(8): 927–934
8. Mishara BL, Weisstub DN (2016) The legal status of suicide: a global review. Int J Law Psychiatry 44:54–74
9. Hemelrijk E, van Ballegooijen W, Donker T, van Straten A, Kerkhof A (2012) Internet-based screening for suicidal ideation in common mental disorders. Crisis
10. Turecki G, Brent DA (2016) Suicide and suicidal behaviour. Lancet 387(10024):1227–1239
11. Meesters Y, Gordijn M (2016) Seasonal affective disorder, winter type: current insights and treatment options. Psychol Res Behav Management
12. Jacobs DG, Baldessarini RJ, Conwell Y, Fawcett JA, Horton L, Meltzer H, Pfeffer CR, Simon RI (2010) Assessment and treatment of patients with suicidal behaviors. APA Pract Guidel, 1–183
13. Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, Joiner TE (2010) The interpersonal theory of suicide. Psychol Rev 117(2): 575
14. Simon GE, Rutter CM, Peterson Do, Oliver M, Whiteside U, Operskalski B, Ludman EJ (2013) Does response on the PHQ-9 Depression questionnaire predict subsequent suicide attempt or suicide death? Psychiatr Serv 64(12):1195–1202

15. Coppersmith G, Leary R, Whyne E, Wood T (2015) Quantifying suicidal ideation via language usage on social media. In: Joint statistics meetings proceedings, statistical computing section, JSM
16. Pennebaker JW, Boyd RL, Jordan K, Blackburn K (2015) The development and psychometric properties of LIWC2015
17. Sap M et al (2014) Developing age and gender predictive lexica over social media. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)
18. van de Loo J, De Pauw G, Daelemans W (2016) Text-based age and gender prediction for online safety monitoring. Int J Cyber-Secur Dig Forens (IJCSDF) 5(1):46–60
19. Miller Z, Dickinson B, Hu W (2012) Gender prediction on twitter using stream algorithms with n-gram character features
20. Cheng N, Chandramouli R, Subbalakshmi KP (2011) Author gender identification from text. Dig Inv 8(1):78–88
21. Shing H-C, Suraj N, Ayah Z, Meir F, Hal D III, Philip R (2018) Expert, crowd sourced, and machine assessment of suicide risk via online postings. In: Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic, pp 25–36
22. Ji S, Yu CP, Fung S-f, Pan S, Long G (2018) Supervised learning for suicidal ideation detection in online user content. Complexity
23. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. J Mach Learn Res 3:993–1022
24. Tadesse MM, Lin H, Xu B, Yang L (2020) Detection of suicide ideation in social media forums using deep learning. Algorithms 13(1):7
25. Tadesse MM et al (2019) Detection of depression-related posts in reddit social media forum. IEEE Access 7: 44883–44893
26. Pruksachatkun Y, Pendse SR, Sharma A (2019) Moments of change: analyzing peer-based cognitive support in online mental health forums. In: Proceedings of the 2019 CHI conference on human factors in computing systems
27. Benlahbib A, Nfaoui EH (2020) A hybrid approach for generating reputation based on opinions fusion and sentiment analysis. J Organ Comput Electron Comm 30(1):9–27
28. Baumgartner J (2019) Reddit May 2019 Submissions. https://doi.org/10.7910/DVN/JVI8CT, Harvard Dataverse, V1
29. Perkins J (2010) Python text processing with NLTK 2.0 cookbook. Packt Publishing Ltd.
30. Fast E, Chen B, Bernstein MS (2016) Empath: understanding topic signals in large-scale text. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 4647–4657

Novel Drug Discovery Using Generative Model for COVID-19

V. Krishna, R. Jagadeesh Kannan, Tulasi Prasad Sariki, and G. Bharadwaja Kumar

Abstract Drug discoveries often need expertise knowledge and insanely complex biochemical tests for the discovery of molecular chemical properties. In recent years, there has been an increasing trend of using AI and deep learning-based tools which aid the domain expect to speed up the process of drug design. The use of dynamically un-supervised deep learning systems is used to identify certain properties of atoms of molecules that could have aided the pharmaceutical scientist at many times. The drug discovery comes under sixth goal of millennium development goals which is used as a standard procedure to combat diseases such as HIV, AIDS, dengue, malaria, and another global pandemic. In this paper, we propose a new un-supervised molecular embedding procedure, which provides a continuous vector of molecule in their latent space. These molecules in their latent space can aide in generation of new atoms or a combination of atoms which have relevant chemical nature and can be used as a direct and effective replacement for the existing molecules. The model proposed in this paper is an LSTM autoencoder (long short-term memory) to model the sequence to sequence approach since the molecule input is taken a string format called simplified molecular-input line-entry system (SMILES).

Keywords Deep learning · Computer vision · Autoencoder · LSTM · SMILES · Millennium development goals

V. Krishna (✉) · R. J. Kannan · T. P. Sariki · G. Bharadwaja Kumar
Vellore Institute of Technology, Chennai, India

R. J. Kannan
e-mail: jagadeeshkannan.r@vit.ac.in

T. P. Sariki
e-mail: tulasiprasad.sariki@vit.ac.in

G. Bharadwaja Kumar
e-mail: bharadwaja.kumar@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_36

1 Introduction

Understanding the causes or conditions of the diseases is the first step in drug design. In the case of Covid-19, the virus "sabotage" cells use its distinctive spike proteins. Making copies of itself is the primary genetic programming of any virus, and COVID-19 is no exception. When the RNA of the virus has entered a cell, new copies are made, and in this phase, the cell is destroyed, allowing new viruses to attack neighbouring cells. Infected cells show how SARS-CoV-2 glycoprotein comes in contact with the ACE2 to the cell membrane protein to enter human cells. Hence, in order to prevent the spread of the disease, all we have to do is to prevent the disease from binding. It can be done with the help of antibody, which is a large, protein structure represented in a Y like molecular structure that is primarily generated by the immune system to fight against the foreign cells or organism that enters human body. If an immune system cannot provide antibodies, small molecule drugs are used which act as an antibody and binds with the virus. Identifying the suitable inhibitor from the existing small molecule compounds in the drug database is done in lab's either by manually binding the small molecule inhibitor with the virus and finding the docking score for that molecule or by using binding software to find the binding score using virtual docking. Instead of using available molecules in the drug database, the scientism uses their domain knowledge to generate new molecule drugs often called as novel drugs or novel compounds. A novel drug or a new molecular entity (NME) is an active, complex, and molecular compound not previously approved by FDA/EMA. They are usually generated using domain knowledge of scientist, but the process of novel drug generation can be replaced with machine learning and AI-based approaches. Basic idea is to create a new molecular embedding technique using deep learning, which can be used to generate a combination of novel atomic structure which can be used to generate new molecular entity. In the present work, an autoencoder model with LSTM layer is proposed which uses a lower dimensional vector conversion unit to convert molecules into a lower dimensional entity and then a similar component to reverse the process at the output. The present work trained the autoencoder model with LSTM layer, on a large collection of unlabelled data obtained from GDB databases (hosted by Raymond research group) at university of Bern [1]. The dataset contains 26.4 million of molecules made up of 17 different atoms from the periodic table.

2 Related Work

This section describes a few related works that have been carried out for drug discovery using deep learning methods in the literature. Bjerrum and Threlfall [2] have explored how a standard neural network along with long short-term memory can detect chemical laws and produce molecules that can be compared by training on existing elements encoded as SMILES. These networks can produce novel quantities, but they are chemically sensitive cells. The features of the molecules are organized by training in a variety of new data that includes fragments such as molecules and drugs such as molecules. The molecules produced and the training data have the same distribution of molar mass, the number of logP, the number of hydrogen acceptors, the number of exchange bonds, and the surface area compared to their training services. Withnall [3] have used the attention and edge memory in the neural network system to obtain graphs, and they have set up their own arguments against the eight different chemical data used in the literature and experiments. The method used to eliminate the need to present a priori information of the function and calculation of the chemical description using only the basic structures is derived from the graph. The results remain in line with other state-of-the-art machine learning methods and set a new standard for the target screening task target.

Qi [4] proposed a self-correcting autoencoder which uses a combination compression and decompression of molecular vector structure. Graphs are omnipresent data structures which reflect interactions between entities. Emphasizing the use of graphs to depict chemical molecules, they have discussed in this paper the process to modulate these structures that correspond to a function derived from the dataset. We also proposed a model of self-correcting autoencoder in which both the encoder and the decoder are organized in a sequence of structures. Their demodulator suggests that network extension phases are arranged sequentially, and they explore and evaluate options available that minimize this linearization possible downsides. Studies equate our methodology to a broad variety of starting points on the role of producing molecules and prove that their process is far more useful in achieving the results of the initial data on qualitative importance parameters. In addition, they have established that their design helps us to build molecules that are (locally) efficient in desirable properties using accurate latent space shaping.

Warmuth [5] has used effective modelling for the drug development issue with help of SVM. They consider those from a wide range of molecules, which in as few docking and processing variations as possible bind themselves to a specific molecule. For binding action against this molecule, a comparatively small set of proteins was filtered in each iteration. To choose successive collections, they employed the so-called "successful learning paradigm" of machine learning. The primary assessment plan concentrates on the "support vector machines" highest likelihood provided by the hyperplane. This straight line separates the active current collection from the protease inhibitors and also has the highest distance from each labelled compound. Systematic review of numerous other DuPont pharmaceuticals selection approaches on datasets and noticed that the approaches centred on the overall kernel function margin easily exceed the finer versions. Zernov [6] has used support vector machines for the process of drug discovery. In contrast with the deep neural network techniques, they have tested the ability of algorithms for determining narcotic-like and biopharmaceutical similarities for large compound collections. SVM utilizes the same collection of descriptors for all forms of data to outperform individual neural networks. They also used SVM to estimate the behaviour of enzymatic inhibitors for inorganic salts and noticed that our SVM model is predictive efficiency and better than the one stated earlier for conventional QSAR. Ekins [7] have used multiple AI approaches like support vector machines have shown their effectiveness in drug discovery with an end-to-end approach. They leverage the usually larger datasets which allowed calculation of objectives and intermolecular interactions with improved precision levels for bioactivity. The integrated implementation of these end-to-end (E2E) machine learning models is widely applicable and has significant implications for the development of future therapies and their target identification.

3 Design and Architecture

The long short-term memory (LSTM) is a form of deep neural network built to acquire deep-term characters, which is useful for certain types of sequence prediction tasks that require the network to retain information over longer time periods. In chemin-formatics, we have the SMILES portrayal of atoms. Essentially, they are portrayal of particles in string design. With the help of SMILES, proteins are structured as strings, and recurrent neural networks are equipped to determine one more SMILE provided a series of previous characters as knowledge [4]. After training the network, it is possible to generate novel drug like molecules. This can be done by selecting one from the most probable and then feeding them back to the network, predict the next molecule in the chain, feedback network allows the molecule chain sequence back to the network and so forth. The systems learn the structure of SMILES series atoms which results in the process of making new protein adhering to similar principles by repeating through the characters. It takes two or three thousand atoms to prepare for overfitting and compound guidelines fundamental the SMILES molecule bonds. Another methodology is to use an autoencoder. Essential thought is to derive the contribution of string to a dormant dimension (a latent space where the molecule is represented in lover dimension) and afterwards reproduce the original input from this space. In the wake of preparing, the autoencoder could be used in a generator and a converter configuration for converting molecules to lower dimension or generation molecules from latent space. The generative part may be utilized to test the region surrounding an atom with intrigue that is the lead compound and produce comparable, however, not equivalent, molecules. Conditional aautoencoder model offers same comparable quality to the proposed model, that uses a converter to encrypt the actual vector or scalar representation, and then a demodulator to undo the vector to the spatial domain. That counts is that the VAE model places the presumption that any specific Euclidean distance fits the embedded domain. So, there is no proof nor test findings to assist the probability distributions on the associated spatial domain, in any situation.

3.1 SMILES Molecular Representation

The molecules are originally presented via the SMILES [8] and that is a screen format for representing morphological and chemical arrangement via regular expressions.

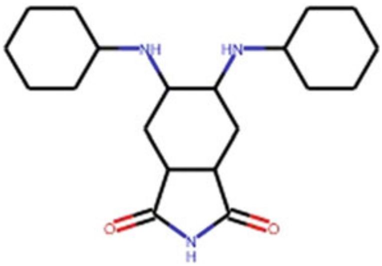


Fig. 1 Molecule 4,5-dianilinophthalimide

The SMILE method describes the organic compounds in a model description, in which the electrons, links and loops are embedded in a matrix. An examples of SMILE is 4,5-dianilinophthalimide which looks like in Fig. 1 can be represented in SMILES format as O=C1 NC (=O)C2CC(NC3CCCC C3)C(NC3CCCCC3)CC21.

3.2 Model Design and Training

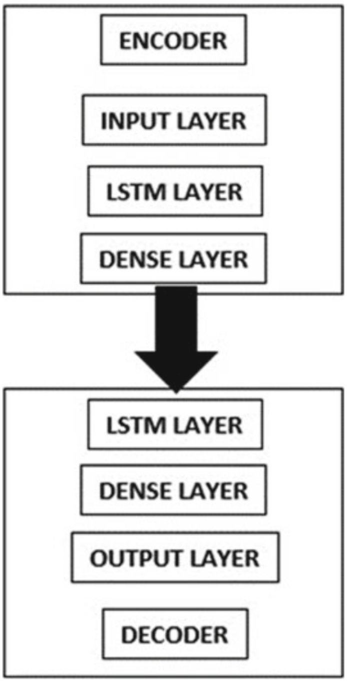
For encoding as one-hot arrays, the SMILES can be vectorized. To achieve so, all symbols contained in the SMILES series are comprised of a collection of characters. Furthermore, such characters beginning and end are added and utilized as trigger to decoder but will indicate that the SMILES flow has stopped. The halt symbol often serves as a wrapping to achieve the same duration for all vectors, so that the system could be programmed in continuous mode. The set of symbols is used to define two databases, for translating back and forth between index and text. The maximum length of the SMILES strings is required and set to the limit encountered with some extra. Such additional parts are used to make the model more flexible and allows the model to run on a large number of SMILES strings. Next the character collection and the database are used in the NumPy arrays to configure the required bits. The effect would be a "harmonica-roll" with any SMILES series of molecules. The X dataset terminates as"!", however, dataset Y is phase shifted by one character, begins with the real SMILES letter. Its time to develop the autoencoder with the vectorized data ready. First, we will import Keras objects and calculate the input and output dimensions from the vectorized data. Defines also the amount of LSTM cells to be used for the converter and integrator and determines the dormant element. LSTM layers form various forms of artificial neural units containing the necessary system for a hidden representation over longer iterations using a recurrent neural network. They were designed to monitor what to do with the identity C cells, with input, output, and forget windows. The condition H is passed here as duplicate of the repeated iterations and tends to decide what will be done in the next phase of identity C. The gating systems guarantee that the system can maintain a hidden representation effectively by locking the entry and ignoring barriers before a new situation or feedback opens them up.

It serves as a substitute for normal recursive systems, which typically improve the RNN's efficiency in detecting long-term connections. The model is kept rather simple overall. Keras's working API is used, which would make it easy to later reconstruct the various layers into encoder and decoder versions. The model is kept rather simple overall. The input SMILES string is read using a single layer of 64 LSTM cells. The output from the sheet is skipped but the bottleneck's later identity C and H is fused and recombined into a "back" row. These include the encoder. Keras's working API is used, which would make it easy to later reconstruct the various layers into encoder and decoder versions. Then, the middle output vector is placed across two different dense layers to decipher the states to be set on the converter layer LSTM. Such replication of the internal states from the sequencer to the converter makes a dormant area of a specific scale than that provided by the layers of the LSTM in itself. The LSTM layer will then collect the input again and will be charged with determining every next identity in the series. So, from the implicit expression of the molecule and the individual! "Allows creation of the next character in a series, describing the first atom, e.g. "C", "N", etc. Then the *Y* values are matched with the! "For the dimensions *X*.

Similarly with the approach where the value of one repeat stage is used as reference to the next, this way of training is called instructor implementation. However, if the network makes a mistake at the start of the series, the preparation of the later sections of the series will not be affected. This trains the converter far better efficiently than merely pushing it into the critical network and instead seeing if it can build the entire sequence. At each point, the output of LSTM neurons is positioned in a complex channel with increasing neuron feature. Therefore, the ultimate training framework on the input neurons and output neurons is built from the artifacts. The overall design from the model being suggested is seen in Figs. 2 and 3.

The present work has used TensorFlow and Keras packages for implementation [5]. Keras "Callbacks" library is used to record the history, and "ReduceLROn-Plateau" library is used to gradually decrease the rate of learning every time when a flat error rate is absorbed. The ML network is optimized with Adam optimization technique and categorical cross entropy as the loss function. X-train is connected to the platform twice to provide the feedback at both encoder and the decoder. We can also use data collected in the background item to map the training cycle. Loss and loss of validity have certain variations, but in the end, the loss and loss rates of validity are more or less gone. You may now use the eligible autoencoder parts to create the different encoder and decoder versions. Implementation of a new layer from the hidden layers into an output layer inside a specific model is quite easy. Once the design is built, the input layer specified it, and the output of the neck was replicated as the source. This model should take a vectorized grin to and store the implicit region in. The next model to be desired is a model that can decipher the latent space at the LSTM cells into the states to be placed. A new input matching the latent space is defined but the levels well before may be reused to obtain the h and c states. And the values are derived from the studied layout. There has to be a little of research on the decoder model. It can be used to predict one character at a time in stateful mode. Therefore, the layers are correctly described. After the

Fig. 2 Autoencoder model



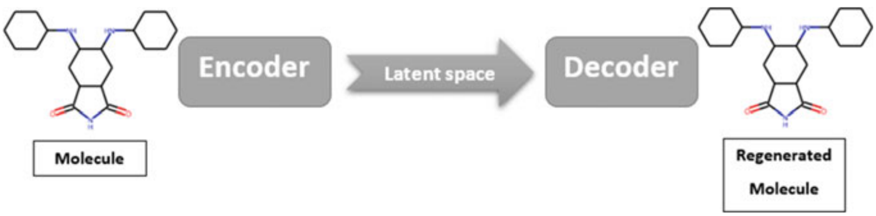


Fig. 3 Molecular conversion process

decoder model has been established, the corresponding weights are transferred from the trained autoencoder model (Figs. 4 and 5).

4 Implementation Details

The SMILES to an implicit design may be used in their valued signature to reflect the SMILES as latent space. A valuable aspect of a signature is that identical compounds create common signatures. A basic analysis can be performed for associated proteins to see whether specific proteins achieve significant generators in the latent space. Here the absolute discrepancy between latent vectors is used as a statistical similitude.

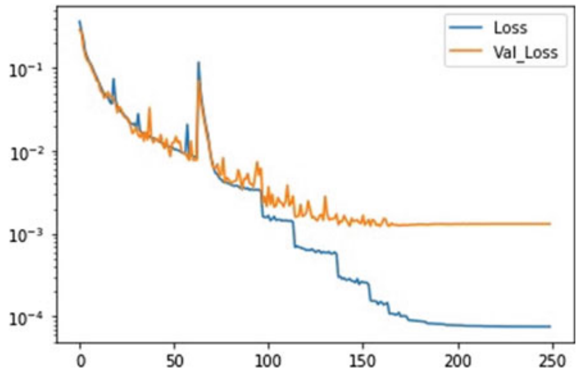


Fig. 4 Change in loss with respected to epoch

```
Model: "model"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
input_1 (InputLayer)            [(None, 49, 25)]     0
__________________________________________________________________________________________________
lstm (LSTM)                     [(None, 64), (None,  23040       input_1[0][0]
__________________________________________________________________________________________________
concatenate (Concatenate)       (None, 128)          0           lstm[0][1]
                                                                 lstm[0][2]
__________________________________________________________________________________________________
dense (Dense)                   (None, 64)           8256        concatenate[0][0]
__________________________________________________________________________________________________
input_2 (InputLayer)            [(None, 49, 25)]     0
__________________________________________________________________________________________________
dense_1 (Dense)                 (None, 64)           4160        dense[0][0]
__________________________________________________________________________________________________
dense_2 (Dense)                 (None, 64)           4160        dense[0][0]
__________________________________________________________________________________________________
lstm_1 (LSTM)                   (None, 49, 64)       23040       input_2[0][0]
                                                                 dense_1[0][0]
                                                                 dense_2[0][0]
__________________________________________________________________________________________________
dense_3 (Dense)                 (None, 49, 25)       1625        lstm_1[0][0]
==================================================================================================
Total params: 64,281
Trainable params: 64,281
Non-trainable params: 0
__________________________________________________________________________________________________
```

Fig. 5 Model summary

This analysis does not find out whether related proteins can have dissimilar implicit dimensions, but is a basic neighbourhood search. We will need two steps to sample the latent space. First use the latent to states model to calculate the c and h states and then set the initial LSTM network decoder states. The input char vector will be fed to the LSTM network and recursively search the next letter until "E" last letter is identified. The output SMILES are docked with MERS protein sequence with the help of PyRx virtual docking software; the resultant docking score for the eight generated molecules is shown in Figs. 6 and 7.

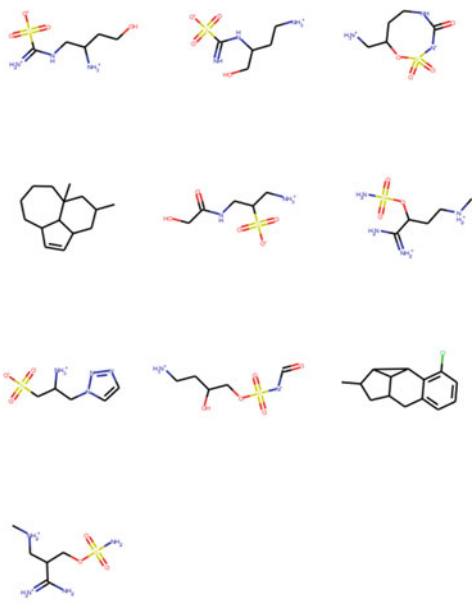
```
[25086 27353 19769 27289 23266 34017  4203 22832 73462 52009]
257499    [NH3+]C(CCO)CNC(=[NH2+])S([O-])(=O)=O
257327         [NH3+]CCC(CO)NC(=N)S([O-])(=O)=O
259439         [NH3+]CC1CCNC(=O)[N-]S(=O)(=O)O1
218020               CC1CC2C=CC3CCCCC(C)(C1)C23
260336         [NH3+]CC(CNC(=O)CO)S([O-])(=O)=O
256857       C[NH2+]CCC(OS(N)(=O)=O)C(N)=[NH2+]
252373        [NH3+]C(CN1C=CN=N1)CS([O-])(=O)=O
257334           [NH3+]CCC(O)COS(=O)(=O)[N-]C=O
217852            CC1CC2CC3=CC=CC(Cl)=C3C3C1C23
257804       C[NH2+]CC(COS(N)(=O)=O)C(N)=[NH2+]
```

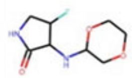
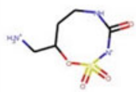
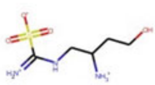
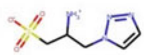
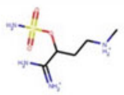
Fig. 6 Generated Novel molecules

5 Conclusion

LSTM-based autoencoder is used in this paper to generate novel drug molecules. The encoder and decoder part of the autoencoder can be used separately to generate molecules and find the similarities between the molecules. The implicit field can define the SMILES framework, but it also seems useful for a number of cheminformatics-related tasks. When it comes to infringing regions amongst proteins or incorporating unwanted noise, the precision of SMILES reconstruction tends to go down. The latent space cannot, however, be completely infinite, reflecting the distinct life of every proteins. The generated novel molecules are docked with MERS protein structure which is a previous version of Covid-19. The docking scores clearly provides that the generated molecule although similar in chemical property due to different structure shows varying docking score. Hence, this autoencoder model can be used as a tool in computational drug design and can be used to generate novel molecules given a starter molecule, and this can help speed up the process of drug research in Covid-19.

Fig. 7 Docking score of different Novel molecules



Molecule	Docking Score
	10.2
	10.2
	4.6
	11.2
	4.8
	4.8

References

1. Pearson PL (1991) The genome data base (GDB)—a human gene mapping repository. Nucleic Acids Res 19(suppl):2237–2239
2. Bjerrum EJ, Threlfall R (2017) Molecular generation with recurrent neural networks (RNNs). arXiv preprint arXiv:1705.04612.
3. Withnall M, Lindelöf E, Engkvist O, Chen H (2019) Building attention and edge convolution neural networks for bioactivity and physical-chemical property prediction, chemrxiv.org.
4. Qi L, Miltiadis A, Marc B, Gaunt AL (2018) Constrained graph variational autoencoders for molecule design. arXiv:1805.09076
5. Warmuth MK, Liao J et al (2003) Active learning with support vector machines in the drug discovery process. J Chem Inf Comput Sci 43(2):667–673
6. Zernov VV, Balakin KV et al (2003) Drug discovery using support vector machines. The case studies of drug-likeness, agrochemical-likeness, and enzyme inhibition predictions. J Chem Inf Comput Sci 43:2048–2056
7. Ekins S, Puhl AC, Zorn KM et al (2019) Exploiting machine learning for end-to-end drug discovery and development. Nat Mater 18:435–441
8. Abadi M et al (2015) TensorFlow: large-scale machine learning on heterogeneous distributed systems.

Structural Strength Improvement of 3D Printing Parts from Topology Optimised Design Using Anisotropic Material Modelling

Thomas P. Morley, Sravanthi Sashikumar, Julfikar Haider, and Weizhuo Wang

Abstract Additive manufacturing (AM) offers diversity, customisability and creativity, making it an important tool to lead Industry 4.0. Reducing the cost of prototyping, bespoke and small-scale production is some of the key advantages of 3D printing. Parts created by topology optimisation and generative design are usually easier to make by AM. Industrial sectors such as automotive, biomedical and manufacturing have begun to see AM as a cost-effective process for complex components. AM is not without its drawbacks, print failures, distortion, rough surfaces and anisotropic properties, and lack of material data is limiting the quality assurance of this technology. It is known that the mechanical properties of the printed parts are sensitive to specific AM process parameters, e.g. the printing direction to the anisotropic property. In this study, the improvement of the prospective mechanical property of a topologically optimised design of a wheel was carried out. As the mechanical load of a wheel is crucial in its application, the prospective mechanical properties of the wheel to be made by AM are investigated by using simulation so that better outcome from AM may be predicted with the customised process planning.

Keywords Topology optimisation · Additive manufacturing · 3D printing simulation · Mechanical properties

1 Introduction

Formula Student [1] is a worldwide competition where students are tasked with designing and building a formula-style car to compete against other universities. It is an important and exciting curriculum for undergraduate engineering education. One of the objectives is to allow engineering students to learn and apply the latest technologies to design, build and test the components and systems of Formula Student cars. Smart design, smart manufacturing and smart operation are important themes in the context of Industrial 4.0 [2, 3]. Within the themes, computational-aided topology

T. P. Morley · S. Sashikumar · J. Haider · W. Wang (✉)
Department of Engineering, Manchester Metropolitan University, Manchester, UK
e-mail: w.wang@mmu.ac.uk

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_37

optimisation [4], generative design (GD) [5] and additive manufacturing (AM) [6] play a crucial role.

1.1 Topology Optimisation and Generative Design

Topology optimisation in structural design has been in constant development since it was first introduced in 1988 [4]. Since then, topology optimisation has been used to minimise the amount of material for a component whilst maintaining the required strength. Topology optimisation produces complex shapes which traditional manufacturing techniques would either take too long to manufacture or even would not be possible. Since the increase in the use of additive manufacturing, topology optimisation has been used to produce the optimal strength to weight ratio designs. Additive manufacturing can produce these complicated geometries but lightweight designs with little or no alterations required.

The initial computer-aided design (CAD) model taken into a topology study is the most basic design with the most material possible. With the proper definition of constraints (e.g. geometric, material strength, resonant frequencies, etc.) and boundary conditions, topology optimisation can be applied to re-distribute the material layout in order to remove the regions with low stress and add more material to the high-stress regions. Popular algorithms include solid isotropic material penalisation (SIMP) and evolutionary structural optimisation (ESO) [5].

With the availability of modern computer power, reliable computational mechanics codes, machine intelligence algorithms and advanced manufacturing techniques, an integrated design framework, known as *generative design,* has become practical. One of the widely used concepts was defined by Shea et al. [7] "*generative design systems are aimed at creating new design processes that produce spatially novel yet efficient and buildable designs through exploitation of current computing and manufacturing capabilities*".

Generative design is usually implemented in iterative optimisation approaches that have been implemented in a couple of commercial software such as AUTODESK® Fusion 360. The process takes user input boundaries, constraints, requirements and manufacturing capabilities. This results in a range of optimal designs which can be manufactured. The software uses cloud computing to run complex simulations and calculations, creating thousands of compatible designs. Components designed using generative design are not limited to additive manufacturing.

With the structural strength as the main objective, these are designs which minimise mass or increase stiffness. This can be done with the implementation of a factor of safety (FOS). The FOS is a multiplication of the maximum amount of force, and a component will be subjected to in its life cycle. Designing off the FOS gives engineers peace of mind when it comes to components which are safety–critical and are required to have a much higher FOS.

Combining the designs being produced with the best strength to weight ratio and additive manufacturing is hugely economical. This allows for manufacturing

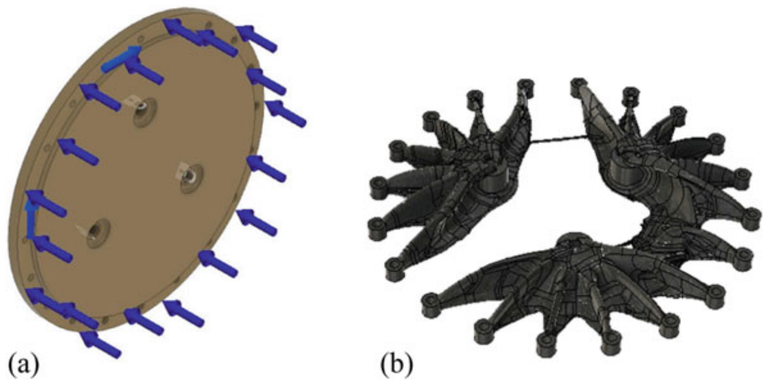


Fig. 1 **a** Original wheel design; **b** a generative design

components with very little waste materials. Less material will result in shorter production times and less material costs, increasing the money saved per component.

For example, in the wheel design of a Formula Student vehicle, as shown in Fig. 1a, the input forces and obstacle regions may be used to produce a coarse convex hull as an initial volume. A level-set method topology optimisation then uses the volume to produce a shape. A hexahedral element-based finite element analysis (FEA) is run to ensure the geometry produced is meeting the constraint conditions set up at the start of the study. The topology optimisation runs through multiple iterations to produce a minimised volume and to produce the beam shapes seen in many generative design studies. A beam network optimisation was used to create and connect nodes, all whilst varying thicknesses, as shown in Fig. 1b. The algorithms [8] produce iterations to check whether the model is fully constrained.

1.2 Additive Manufacturing

3D printing, also known as additive manufacturing(AM) [9], is the "fabrication of a physical, three-dimensional part of arbitrary shape directly from a numerical description, typically a computer-aided design (CAD) model by a quick, totally automated, and highly flexible working process without any tooling" [6]. AM was first developed in the 1980s by Charles W. Hull. A stereolithography technique which focussed ultraviolet (UV) light to cure a photopolymer resin was used. This may be done in successive layers to produce a 3D model. Within this study, fused deposition manufacturing (FDM) was the chosen AM technique. This is one of the most used processes in 3D printing. A filament is melted and extruded through a nozzle, building up layers to make the required shape.

The use of 3D printing within the industrial sectors has allowed more companies to conduct rapid prototyping [10]. This process allows for designers who have produced a 3D model using CAD to ensure that it will fit without conflicting with any of the other components. Rapid prototyping permits the production of a full-scale real-life part quickly and economically. The cost of a print varies greatly. One main factor affecting price is infill density. It is the percentage of the interior parts of the 3D printed model that contains material as opposed to space. This can be used to reduce print time and material usage. Another condition that can be altered is the material choice. During rapid prototyping, the geometric shape of the component is usually more important than the material used. For this reason, a "cheap" material such as acrylonitrile butadiene styrene (ABS) can be used as an alternative. This material is widely used for AM and is recyclable [11, 12].

There are a number of issues of additive manufacturing including print failures, distortion, rough surfaces, anisotropic properties and lack of material data that is limiting the quality assurance of this technology [13–16]. For example, it is known that the mechanical properties [17] of the printed parts are sensitive to specific AM process parameters [18, 19], e.g. the printing direction to the anisotropic property.

Without stringent testing of printed parts, the uses of AM are present in only a handful of real-life applications. Print failures are a common occurrence with AM processes and can be caused by a large range of factors. There are many failures across each AM method or material, but very few common across all [20].

Print predictive technologies [21–23] may be achieved by using multi-physics simulation, e.g. distortion prediction and compensation due to thermal stress effect [24, 25]. Printing simulation may also help to prevent print failures by analysing a computer-aided design (CAD) model and correct issues with the mesh. Automated process monitoring [26–29] can be used to detect defects produced during the print, allowing the user to repair or restart it depending upon the severity of the defect. However, this technology is currently limited and requires many print failures to gather the required data. This must be done for each different model, printer and material [30].

In the following sections, the improvement of the prospective mechanical property of a topologically optimised design of a wheel was carried out. As the mechanical load of a wheel is crucial in its application, the prospective mechanical properties of the wheel to be made by AM are investigated by using simulation so that better outcome from AM may be predicted with the customised process planning.

2 Design of a Formula Student Car Wheel

The design, optimisation and production of the wheel are for a Formula Student vehicle. As shown in Fig. 1a, forces acting upon the wheels during the competition events were used to provide the boundary conditions for simulations. The maximum lateral force (acting perpendicular to the car) and longitudinal force (acting around the circumference of the wheel) were 2000 N and 1777 N, respectively. To produce the forces acting upon the wheel, some assumptions were made. These were the assumption of optimal tyre operating temperature, thus producing maximum grip. A maximum mass of the vehicle and driver is 300 kg and the effects of camber acting on the car being negligible.

As shown in Fig. 1a, the lateral force was applied to the area where the rim of the wheel would be attached. This is where the force acting upon the tyre during cornering would transfer to through the rim. At the same time, a moment was calculated using the maximum longitudinal force and the radius of the tyre. This provided a simulation of a "worse-case scenario" which could occur during a spin-out when racing or testing. The constrained points of the simulation were the three bolt locations.

3 Topology Optimisation of Wheel Design

Four stages of the topology optimisation process of the wheel are shown in Fig. 2. The first image (a) shows a basic wheel design which was used in the study. Image (b) shows the mesh results of the topology study obtained from a SIMP algorithm. It

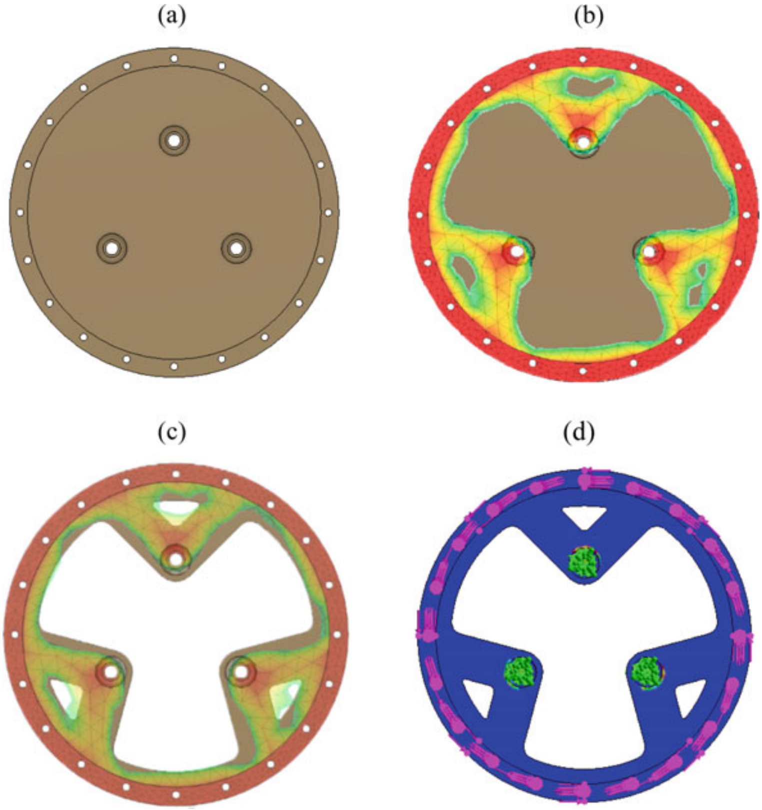


Fig. 2 Topology optimisation process of the wheel

may be seen that the boundaries are jagged because of the meshes. It may be better to smooth the edges, e.g. to reduce stress concentration, as shown in image (c). The FOS of the final wheel design, as shown in image (d), has more material on than required from the topology study. The initial mass of design (a) was 1.08 kg, design (d) has a mass of 0.570 kg.

The topology optimisation study was run to have a factor of safety (FOS) of 1.5. As the wheel is designed to be used in a motorsport application and therefore should be as light as possible, the smaller the FOS, the lighter the wheel. When creating the final wheel design, additional material was placed slightly external to the topology mesh. This resulted in a minimum FOS of 1.9 in the areas where the wheel would be bolted to the hub.

The topology optimisation was carried out using the solid isotropic microstructure with penalisation [31, 32] algorithm in which the material assumption of isotropy and homogeneity was made. However, it is known that in the physically printed parts, such assumptions are usually invalid. A number of experimental studies [24, 33–37] have shown the effect of the process parameters to the outcome and quality of the printed parts in the perspective of mechanical characteristics. The specimens used in these studies are in the standard shapes. In practice, it is crucial that the results from the standard testing can be applied to the real applications of part design. In the next section, the printing process planning of the topologically optimised wheel is investigated. The main factor we considered is the anisotropic properties due to the raster deposition directions [10, 13, 19]. It is aiming to improve the strength of the printed wheel by customised the deposition directions based on the internal loads of the structure.

4 DFM Process Planning to Improve Strength

As shown in the literature [17, 38], the strength of a 3D printed part is significantly affected by the process parameters such as build orientation and raster angle in FDM. The anisotropic properties in the printed parts may be modelled by the transversely isotropic material constitutive behaviour [12], as expressed in Eq. (1).

$$\begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \gamma_{xy} \\ \gamma_{yz} \\ \gamma_{xz} \end{Bmatrix} = \begin{bmatrix} 1/E_p & -v_p/E_p & -v_{pz}/E_p & 0 & 0 & 0 \\ -v_p/E_p & 1/E_p & -v_{pz}/E_p & 0 & 0 & 0 \\ -v_{pz}/E_p & -v_{pz}/E_p & 1/E_z & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+v_p/E_p & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2G_{pz} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2G_{pz} \end{bmatrix} \begin{Bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \tau_{xy} \\ \tau_{yz} \\ \tau_{xz} \end{Bmatrix} \tag{1}$$

where σ, τ and ε, γ are stress and strain components, respectively, E_p and E_z are the elastic moduli of the planar directions (x-y plane in Fig 3) and the principal direction

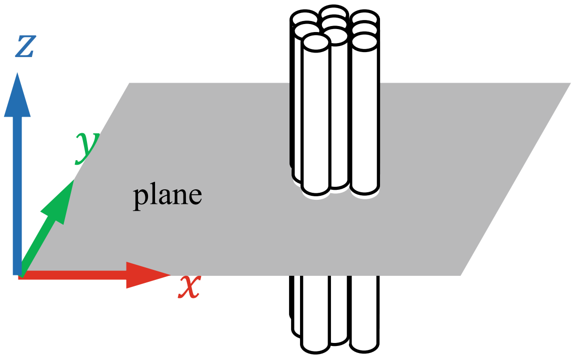


Fig. 3 Schematic diagram of a transversely isotropic material model

(z-direction in Fig 3), respectively, ν_p is the Poisson's ratio in the planar directions; ν_{pz} and G_{pz} are the Poisson' ratio and shear modulus between the principal direction and the plane. A set of experimental values of these constitutive parameters of ABS using FDM were estimated by Zou et al. [12] and will be applied in this study. The strength in the principal direction is slightly stronger than the transverse directions [12, 33].

Thus, this would help to plan the deposition directions during the printing process.

The initial study of structural analysis showed that the applied moment/torque mainly contributes to the stress along the circumferential direction, and the applied lateral force mainly contributes to the stress along the radial direction. The two main stress directions may help in planning the deposition directions. It may be helpful to align the principal direction of the transversely isotropic material model to the main stress directions. It is proposed, as shown in Fig. 4, that the wheel is sectioned into the front and back plates. For the front plate, the principal direction of the transversely isotropic material model is aligned to the circumferential direction of the wheel as shown in Fig. 4a, which may improve the structural strength to resist the applied moment. For the back plate, the principal direction of the transversely isotropic material model is aligned along with the radial directions, as shown in Fig. 4b, so that the structural strength may be improved to resist the applied lateral

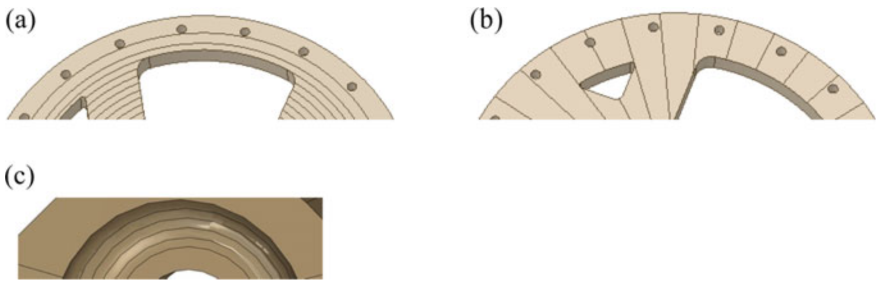


Fig. 4 " Customised" printing process path for the deposition to potentially improve the strength; the black lines illustrate the deposition paths.

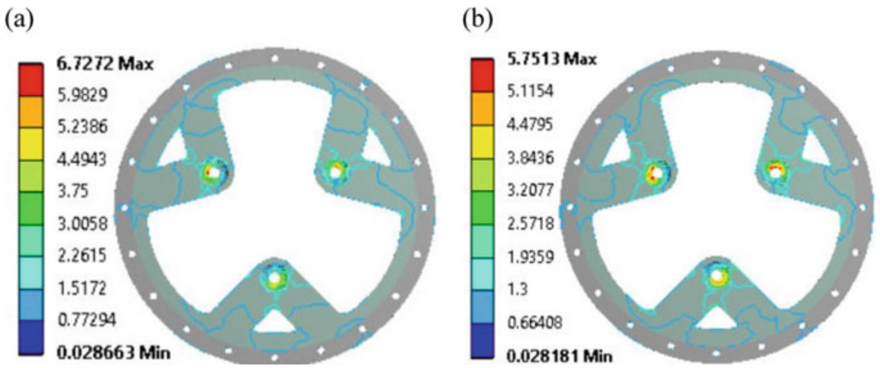


Fig. 5 Equivalent stress distribution (MPa) of the wheel subjected to the design loads; **a** simulation from a rectangular deposition pattern; **b** simulation from a customised deposition pattern as illustrated in Fig. 4

force. Furthermore, the circular deposition pattern may be applied to the three holes that are closed to the centre of the wheel, as shown in Fig. 4c.

5 Results and Discussion

Results of stress analysis by ANSYS from different deposition patterns are shown in Fig. 5. Two deposition patterns are considered with identical boundary and loading conditions. The equivalent stress of the first pattern from a rectangular deposition pattern is shown in Fig. 5a, while the results from the customised deposition pattern, as described in Fig. 4, is shown in Fig. 5b.

It may be seen that the stress is more concentrated in Fig. 5a, whilst the high stresses are more spread around the holes in Fig. 5b. The maximum value of the equivalent stress in the wheel from the customised deposition pattern in Fig. 5b is $\frac{6.7272-5.7513}{0.5\times(6.7272+5.7513)} = 15\%$ less than the other one in Fig. 5a. Therefore, it may be possible to improve the structural strength of a printed wheel by customising the deposition pattern based on the alignment between the principal direction of the transversely isotropic material model and the principal stress directions.

6 Conclusions

In this study, a wheel design of a Formula Student car with computational-aided topology optimisation was carried out for additive manufacturing. The uncertainty in additive manufacturing was briefly surveyed. The process influence on the mechanical properties of the wheel design was simulated by customising deposition patterns of the DFM process. It was found that aligning the principal direction of the transversely isotropic material model for DFM to the principal direction of stress may improve the strength of the printed part. Further work will be crucial to develop a mathematical formulation of this strength improvement approach with experimental validation.

References

1. IMechE, Formula student—institution of mechanical engineers. [Online]. Available: https://www.imeche.org/events/formula-student
2. Qi Q, Tao F (2018) Digital twin and big data towards smart manufacturing and Industry 4.0: 360 degree comparison. IEEE Access 6:3585–3593
3. Liu J, Ma Y, Fu J, Duke K (2015) A novel CACD/CAD/CAE integrated design framework for fiber-reinforced plastic parts. Adv Eng Softw 87(9):13–29
4. Sigmund O, Maute K (2013) Topology optimization approaches. Struct Multi Optimisation 48(6):1031–1055
5. Weinberg D, Michael S (2016) Generative design with autodesk Nastran topology optimisation. Autodesk, Las Vegas
6. Noorani R (2018) 3D printing: technology, applications, and selections. CRC Press, London, p 3
7. Shea K, Aish R, Gourtovaia M (2005) Towards integrated performance-driven generative design tools. Autom Constr 14(2):253–264
8. Nourbakhsh M, Morris N, Bergin M, Iorio F, Grandi D (2016) Embedded sensors and feedback loops for iterative improvement. In: ASME 2016 international design engineering technical conferences and computers and information in engineering conference, North Carolina
9. Guo N, Leu MC (2013) Additive manufacturing: technology, applications and research needs. Front Mech Eng 8(5):215–243
10. Gardan N (2014) Knowledge management for topological optimization integration in additive manufacturing. Int J Manuf Eng 2014:1–9
11. Zhao P, Rao C, Gu F, Sharmin N, Fu J (2018) Close-looped recycling of polylactic acid used in 3D printing: an experimental investigation and life cycle assessment. J Clean Prod 197:1046–1055
12. Zou R, Xia Y, Liu S, Hu P, Hou W, Hu Q, Shan C (2016) Isotropic and anisotropic elasticity and yielding of 3D printed material. Compos Part B Eng 99(8):506–513
13. Domingo-Espin M, Puigoriol-Forcada JM, Garcia-Granada A-A, Llumà J, Borros S, Reyes G (2015) Mechanical property characterisation and simulation of fused deposition modeling polycarbonate parts. Mater Des 83(10):670–677
14. Popescu D, Zapciu A, Amza C, Baciu F, Marinescu R (2018) FDM process parameters influence over the mechanical properties of polymer specimens: a review. Polym Test 69(8):157–166
15. Morettini G, Razavi SMJ, Zucca G (2019) Effects of build orientation on fatigue behavior of Ti-6Al-4V as-built specimens produced by direct metal laser sintering. Procedia Structural Integrity 24: 349–359. ISSN 2452-3216 https://doi.org/10.1016/j-prostr.2020.02.032
16. Liverani E, Toschi S, Ceschini L, Fortunato A (2017) Effect of selective laser melting (SLM) process parameters on microstructure and mechanical properties of 316L austenitic stainless steel. J Mater Process Technol 249(11):255–263
17. Yao T, Ye J, Deng Z, Zhang K, Ma Y, Ouyang H (2020) Tensile failure strength and separation angle of FDM 3D printing PLA material: Experimental and theoretical analyses. Compos Part B Eng 188(5):107894
18. Martínez J, Diéguez JL, Ares E, Pereira A, Hernández P, Pérez JA (2013) Comparative between FEM models for FDM parts and their approach to a real mechanical behaviour. Procedia Eng 63:878–884

19. Galeja M, Hejna A, Kosmela P, Kulawik A (2020) Static and dynamic mechanical properties of 3D printed ABS as a function of raster angle. Materials 13(1):297
20. Sood AK, Ohdar RK, Mahapatra SS (2012) Experimental investigation and empirical modelling of FDM process for compressive strength improvement. J Adv Res 3(1):81–90
21. Bikas H, Stavropoulos P, Chryssolouris G (2015) Additive manufacturing methods and modelling approaches: a critical review. Int J Adv Manuf Technol 83(7):389–405
22. Abbot DW, Kallon DVV, Anghel C, Dube P (2019) Finite element analysis of 3D printed model via compression tests. Procedia Manuf 35:164–173
23. Koeppe A, Padilla CAH, Voshage M, Schleifenbaum JH, Markert B (2018) Efficient numerical modeling of 3D-printed lattice-cell structures using neural networks. Manuf Lett 15(1):147–150
24. Wu W, Ye W, Wu Z, Geng P, Wang Y, Zhao J (2017) Influence of layer thickness, raster angle, deformation temperature and recovery temperature on the shape-memory effect of 3D-printed polylactic acid samples. Materials 10(8):970
25. Li Y, Zhou K, Tan P, Tor SB, Chua CK, Leong KF (2018) Modeling temperature and residual stress fields in selective laser melting. Int J Mech Sci 136(2):24–35
26. Zhang S, Sun Z, Long J, Li C, Bai Y (2019) Dynamic condition monitoring for 3D printers by using error fusion of multiple sparse auto-encoders. Comput Ind 105(2):164–176
27. Tlegenov Y, Hong GS, Lu WF (2018) Nozzle condition monitoring in 3D printing. Robo Comput Integr Manuf 54(12):45–55
28. Scime L, Beuth J (2018) A multi-scale convolutional neural network for autonomous anomaly detection and classification in a laser powder bed fusion additive manufacturing process. Additive Manuf 24(12):273–286
29. Delli U, Chang S (2018) Automated process monitoring in 3D printing using supervised machine learning. Procedia Manuf 26:865–870
30. Dizon JRC, Espera AH, Chen Q, Advincula RC (2018) Mechanical characterisation of 3D-printed polymers. Additive Manuf 20(3):44–67
31. Liu J, Gaynor AT, Chen S, Kang Z, Suresh K, Takezawa A, Li L, Kato J, Tang J, Wang CCL, Cheng L, Liang X, To AC (2018) Current and future trends in topology optimisation for additive manufacturing. Struct Multi Optimisation 57(5):2457–2483
32. Pereira J, Fancello E, Barcellos C (2004) Topology optimisation of continuum structures with material failure constraints. Struct Multi Optimisation 26:50–66
33. Zhang P, Liu J, To AC (2017) Role of anisotropic properties on topology optimisation of additive manufactured load bearing structures. Scripta Materialia 135(7):148–152
34. Sood AK, Ohdar RK, Mahapatra SS (2009) Parametric appraisal of fused deposition modelling process using the grey Taguchi method. Proc Inst Mech Eng Part B J Eng Manuf 224(7):135–145
35. Sood AK, Ohdar RK, Mahapatra SS (2010) Parametric appraisal of mechanical property of fused deposition modelling processed parts. Mater Des 31(1):287–295
36. Singh S, Ramakrishna S, Singh R (2017) Material issues in additive manufacturing: a review. J Manuf Processes 25(1):185–200
37. Hofland EC, Baran I, Wismeijer DA (2017) Correlation of process parameters with mechanical properties of laser sintered PA12 Parts. Adv Mater Sci Eng 2017:1–11
38. Liu J, Ma Y (2016) A survey of manufacturing oriented topology optimisation methods. Adv Eng Softw 100(10):161–175

Uncovering of Depression Through User Profiling

Sruthi Shiva, Tulasi Prasad Sariki, G. Bharadwaja Kumar, and R. Jagadeesh Kannan

Abstract This paper aims to understand how people suffering from depression interact with the online world and how it is different from those not afflicted with this mental illness. It focuses on how people interact, react with various communities, and how they in turn convey information. These texts and its authors are used to extract strategic and novel set of features to be fed into machine learning models. Various studies have tackled uncovering depression symptoms from a single post, using various features extracted from the text. This paper aims to understand if it is possible to uncover if a person suffers from depression from their history of posts that need not be the posts dealing explicitly with sadness, or such stereotypical depressive topics. The analysis of results in this paper shows promising results with users being classified with depression with a high accuracy.

Keywords Machine learning · Depression · User profiling

1 Introduction

With the increase in the popularity of social networking websites as a tool to express a person's creativity and thoughts, it can be used to accurately detect the thought process of a person to a certain extent. This content generated in these social platforms reflect updates not only on the day to the events of their lives, but also on their mental states. With the stigma associated with mental illness, it becomes harder for people to come forward to be diagnosed or accept that they have the illness. Reddit, being

S. Shiva (✉) · T. P. Sariki · G. B. Kumar · R. J. Kannan
Vellore Institute of Technology, Chennai, India

T. P. Sariki
e-mail: tulasiprasad.sariki@vit.ac.in

G. B. Kumar
e-mail: bharadwaja.kumar@vit.ac.in

R. J. Kannan
e-mail: jagadeeshkannan.r@vit.ac.in

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_38

a completely anonymous website, where users are not known by who they actually are, but rather by usernames and personas they don upon, they tend to be more honest in expressing how they feel, than to somebody face to face making it more reliable in detecting mental illnesses [1].

Usually, texts taken at a single point of time may not be able to differentiate between sadness and depression. Sadness can be triggered by several events that occur in the life of a person and is linked to a particular event and usually passes. However, if it does not pass, or if the individual becomes unfit to continue to their capacity, this could be an indication that they are suffering depression. Social, occupational, and other functioning are impaired. Symptoms of depression last for a long time.

The need for this hour would be that of how such large volumes of data can be collected and analyzed and studied to give invaluable insights on dealing with sensitive topics like mental illness, bullying, etc.

The motivation of this paper is to use the online resources around us to help identify those who are suffering. A key ingredient of effective health policy and planning in every country is a reliable estimate of the ratio of the country's population who are affected by diseases or health conditions. Identification of these disorders is extremely important as the consequences are high of lost health which is taken into consideration. Depression is one of the biggest contributors to global disability (7.5% 2015) as ranked by WHO.

A major contributor to suicide death is depression, which totals up to 800,000 per year. Common mental disorders consist of two important indicative categories: depressive and anxiety disorders. These disorders are highly predominant in humans and have a severe impact on the mood or feelings of those affected; symptoms can be detected under different ranges concerning their severity (mild to severe) and duration (from days to months to years). Being diagnosable health conditions, depression is distinct from sadness, stress, or fear that experienced from time to time in their usual lives.

The World Health Organization reports that at least 30% of the adult population suffer or have suffered from a mental disorder. These mental illnesses have however been left undiscovered and untreated by the people suffering from the illness itself, who may have dismissed these feelings as somewhat normal and not as a symptom to something serious as depression. This paper aims to tackle the need for depression detection of users pertaining to one of the most popular social news aggregations, Web content rating, and discussion website Reddit.

Reddit is one of the biggest online networking websites on the planet, with an excess of 85 million unique visitors for every month. Reddit is apt for our study as a result of its huge volume, yet in addition since it offers visitors anonymity and ranges a wide scope. Enlisted users can anonymously examine different posts on more than 1 million subpages, called subreddits. In addition, an extensive number of subreddits are devoted to emotional wellness.

2 Related Works

Studies in understanding depression have been in place even before the advent of the internet. Multiple studies have been conducted to study how language is used by patients with different types of mental disorders, including schizophrenia [2], suicidal tendencies [3], and depression [4].

Using surveys and user studies, many widely accepted scales and criteria were developed like Beck's Depression Inventory [5] which comprised of 21 questions about users' mental and physiological state. CES-D scale [6] consists of questions about the mental conditions aligned with different scores or the degree of emotion experienced. The score is then calculated and then matched up with a metric to diagnose the level of depression.

In some works, participants were used to fill questionnaires or give interviews from which behaviors were analyzed. These methods may seem compelling, but it is extremely time-consuming and hard to assure if the data is robust and generalizable. Moreover, these questionnaires are targeted over characteristics that are defined and fail to keep up with new findings.

Most of the current work, however, has started using social media as the source of analysis [7–9]. Most studies use Twitter data for predicting mental illness such as depression [10], PTSD [11], and bipolar disorders [12]. Moreover, various social media postings of individuals suffering from depression were used in model building to anticipate the possibility of depression [13].

Lately, computers have been used to analyze words and language with tools such as Linguistic Inquiry and Word Count. The tool is used to gather features for multiple works dealing with mental health [14]. In recent studies, first person singular pronouns are considered to signify distress or negative sentiment rather than as a specific indication of depression [15]. Studies have indicated that updates on social media especially Facebook can divulge indications of major depressive episodes [16], and a substantial difference in how social media is used by depressed and non-depressed users [17].

Reddit has only been for analysis of mental illness very recently. Studies have been conducted to analyze the language in different subreddits related to mental health [18] to indicate that its complexity varies across different subreddits. A lot of work has been done in the prediction of anxiety and bipolar [19] disorders from posts on Reddit.

Previous work has explored the impact of people's profiles on how they interact with social media. Studies have inferred the detection of gender from blogs, reviews, email, and tweets [20]. Other attributes explored include the user's location [21], age, political orientation [22].

3 Design and Architecture

Described as the front page of the Internet, Reddit is a social news aggregator that is a mix of a discussion platform as well as a link distributor. It is divided into forums called subreddits, where users create, consume, and curate content in this medium. The platform is largely self-driven, with moderators acting as forum guardians.

Registered members can contribute to the site with various multimedia contents. The community then determines which content and discussions are significant using a points system of upvotes and downvotes and then displayed at the top of the feed. Every community has their own rules, etiquette, subscribers, and posts that may pertain exclusively to their members or may be shared among various other subreddits.

To extract the data, Reddit's official API was used to collect posts and associated metadata from several mental health subreddits: specifically using a Python wrapper PRAW.

Since Reddit is anonymous, there is no direct way by which users suffering from depression can be identified to be the source of analysis. However, there are multiple subreddits that are dedicated forums in discussing various mental illnesses. A comprehensive list of subreddits was generated with Reddit's native subreddit search feature. The subreddits hence crawled were "depression" and "depression help" were crawled to collect posts in the subreddit in the past six months. Users were only considered to be part of the depressed data only if their posts in these subreddit satisfied the pattern "(I'm/I was/I am/I've been) (diagnosed/have/had/am diagnosed) (depression/depressed)." Since these subreddits exclusively deal with users suffering from depression, most of the users are inherently part of the dataset. This pattern is applied to omit people who have posted in this subreddit seeking help for someone they know. These users are then compiled and then serve as the base users from which their post history is derived.

To assemble a list of users who are not suffering from any mental illness, posts from Reddit's r/all page were collected for the past six months. Reddit's r/all is not a subreddit but rather a page that pulls highly voted posts from all subs randomly, according to the websites confidential algorithm. This gives us a comprehensive list of users of a wide range of users from which users need to be filtered out. A comprehensive list of subreddits related to mental health was generated with Reddit's native subreddit search feature. The list compiled consists of more than a hundred subreddits dealing with a variety of mental illnesses. The user's history of posts were generated and if a user is either part of these subreddits or any of their posts satisfies the pattern "(I'm/ I was/ I am/ I've been) (diagnosed/have/had/am diagnosed) (depression/depressed)," the username is removed. This was done to make sure that users suffering from any mental illness are removed from the list of healthy users to maintain the integrity of the dataset to be created.

An episode of depression varies widely, with durations between a quarter of year to a full year [23]. With a maximum number of people only having depressive episodes for a year, each user's posts over a year were crawled and extracted. For every post scraped information regarding the title of the post, the body or its textual content, id, timestamp when the post was made, author id, was further obtained. In this paper, upvotes and downvotes of a post were not considered as other people's reaction to a post does not ascertain or discount the value or importance of the person posting.

Dataset 1: This dataset consists of the history of posts from users who have self-reported to having suffered from depression in the past one year.

Dataset 2: This dataset consists of the history of posts from users who show no sign of having suffered any mental illnesses in the past one year.

These datasets were then marked and then combined and randomized to make the resulting final dataset balanced. The number of users was matched to make sure that both fractions of data were equally represented.

4 Implementation Details

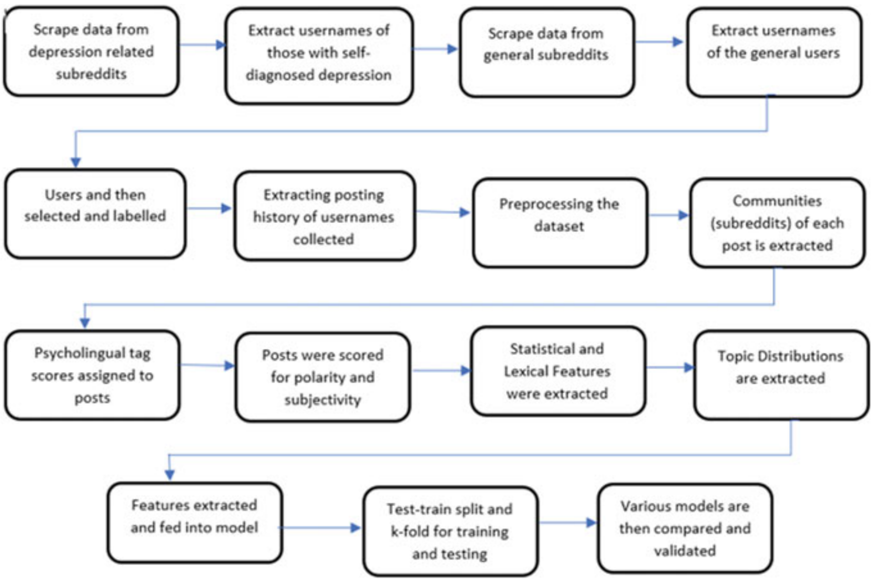


Fig. 1 Architecture of the proposed system

4.1 Preprocessing

Texts in social media tend to be flexible and use a variety of symbols and characters that are not compatible for textual and semantic analysis. Such an example would be emojis, which were removed using a predefined library of these symbols. Moreover, extra non-alphanumeric characters such as special characters, punctuation marks, and white spaces were removed. Inflectional endings of words are removed using morphological analysis of words and vocabulary, and the dictionary form of a word is returned which is referred to as the lemma, in a process called lemmatization. Hence, deriving the base of the world can help remove inconsistencies in spelling due to inflection, part of speech and so on and enhance the efficiency of the algorithm. Lastly, stop words were removed from the text to reduce the dimensionality of the text. Since Reddit is fast changing social media, many users and posts may get removed due to rule violations, being posted in the wrong format or being posted multiple times. Hence, when these posts or users are removed, it is also removed from the dataset to prevent inaccuracies.

4.2 Topic Modeling

To understand the nature of interaction of different types of people in Reddit posts, general linguistic attributes in the content. Various models were used to extract topic distribution, to determine the differences in what depressed and non-depressed users usually post about. Optimally ten-dimensional topic features were chosen to obtain the hidden topics in the post. Latent Dirichlet Allocation (LDA) model, which is an unsupervised probabilistic and non-negative matrix factorization (NMF), a linear-algebraic model was employed.

4.3 Feature Extraction

For each post, multiple features were extracted such as psycholinguistic features using the Empath tool [24], which classifies words into two hundred predefined and manually curated categories, with similarities based on neural embeddings. Lexical features extracted were that of the different parts of speech using the Penn Treebank tag set to observe how posts differed semantically between the two groups of people. Moreover, polarity and subjectivity of texts were also extracted. Polarity is expressed in the float range of values within the range of $[-1,1]$, where 1 is positive and -1 is a negative statement. Subjective is similarly expressed within the range of $[0,1]$ and refers to personal opinion, emotion or judgment expressed in the post.

4.4 Modeling

The posts were then grouped by the username and then analyzed a whole to study trends through user profiling. The psycholinguistic features along with the lexical features and thus extracted, along with their measure of interactivity, that is the count of posts within the time period. A key feature extracted from the user is the communities (or subreddits) they choose to be part of and contribute to. This aims to see if a user's mental health can be predicted by how they interact with this social media.

Since the task of depression disorder prediction is a binary classification task, thus features and classifiers such as Gaussian Naive Bayes (NB), support vector machine and logistic regression with stochastic gradient descent (SGD), random forest ensemble (RF), bagging, AdaBoost, and XGBoost were used. The models were evaluated, and the hyperparameters were tuned using test–training split and k-cross-validation. As baselines, a majority class classifier was used to validate the quality of the trained models.

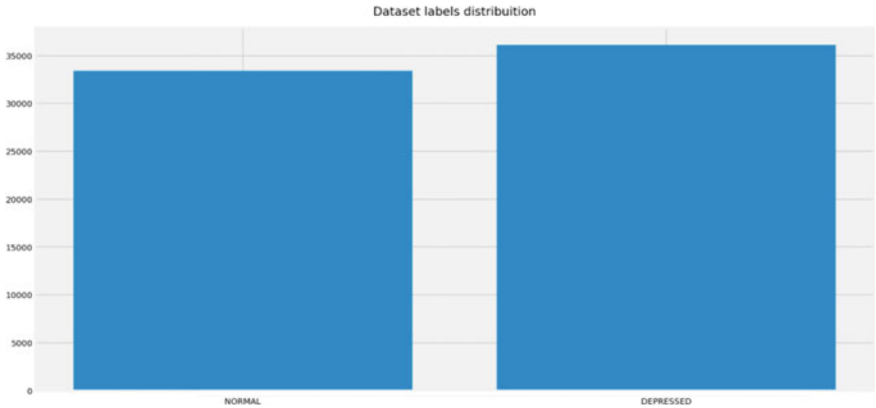
4.5 Word Embedding Approach

For a word embedding approach, a Word2Vec model was trained using the genism implementation of Word2Vec in Python. Stop words are removed and words appearing with frequency of less than 10 counts are ignored. Once these embeddings are learnt for words across the corpus, the embeddings for all words within a document, ignoring words excluded from the Word2Vec model vocabulary are combined by averaging the word vectors. A sequential model is derived with an embedding layer, and a LSTM layer over which a dense layer with a sigmoid activation function is stacked. An embedding layer with a 300-dimensional vector space to embed was defined and input documents with 300 words each. These weights increase the vectors corresponding to words that are more uniquely identifying document while decreasing words relative to their global frequency across all documents in the corpus.

5 Result Analysis and Conclusion

From the various models created and tested, all the tests pass the baseline accuracy of around 50%. We can also observe that a user posts through time are consistent and do not show much of a change. The values of features extracted as the mean from posts of a particular user (Table 1) are almost similar to almost every individual post of this user (Table 2). The entire dataset consists of 69,544 posts from all types of user, with 8572 unique Reddit forums, and 4479 users. Predicting using either from a single post or as an average from multiple posts almost gives a similar result, with the user profiling doing slightly better.

The post-distribution of submissions of users based on their mental health is:



Results with Word2Vec and a sequential model gave similar results with about 70% accuracy and a loss of 55% over 10 epochs. However, when combined with the POS tags, subjectivity, polarity and post count, the models performed better with an accuracy of 75%. The best individual results are observed when tracking all the subreddits or forums the user has contributed to, giving results with an accuracy of 94% (Table 2), which can be labeled as the key feature to distinguish between the users of the two groups. The model performs slightly better when the Subreddit History feature is combined with other features, together giving an accuracy of 96% (Table 1).

The least effective features extracted from the texts were the parts of speech which gave only a slight advantage to any randomized model, which can be used to conclude that parts of speech alone cannot be effectively used to categorize people into depressive or not. The various Empath feature extracted perform better with an accuracy of 71% when combined with user profiling (Table 1).

From the various models used to analyze the data, Naïve Bayes performs the worst, performing only slightly better than baseline classifier at times, whereas the best results have been obtained by using Boost, with Bagging and AdaBoost giving similar results.

From the topics generated from the topic modeling, it is observed that depressed people usually talk about how bad they feel most of the time, moreover in various contexts and not only when they talk about how they are depressed. Topics either deal with relationships, how nothing is going write and also deals with them asking for help and also a few following up to deal with their issue.

On the other hand, non-depressed people seem to be talking about a wide variety of topics that deals with work, video games, songs, books, looking for a place to talk or someone to connect with, issues dealing with money, property, etc. Though a few

Table 1 Results when the features are analyzed with various machine learning models with user profiling

		Baseline	NB	RF	SVM	Log Reg	AdaBoost	Bagging	XGB
POS tag	Accuracy	0.49	0.52	0.60	0.60	0.61	0.61	0.61	0.62
	Precision	0.51	0.70	0.63	0.65	0.63	0.64	0.63	0.65
	Recall	0.51	0.14	0.53	0.51	0.61	0.58	0.57	0.57
	F1	0.51	0.23	0.58	0.57	0.62	0.61	0.60	0.61
	AUC	0.49	0.54	0.60	0.61	0.61	0.61	0.61	0.62
Empath	Accuracy	0.50	0.68	0.69	0.69	0.71	0.70	0.71	0.71
	Precision	0.53	0.89	0.80	0.89	0.83	0.79	0.77	0.79
	Recall	0.53	0.39	0.50	0.42	0.52	0.56	0.60	0.56
	F1	0.53	0.53	0.61	0.56	0.63	0.65	0.68	0.65
	AUC	0.50	0.69	0.70	0.70	0.72	0.71	0.71	0.71
Empath + POS tag	Accuracy	0.50	0.71	0.72	0.75	0.75	0.74	0.75	0.75
	Precision	0.52	0.90	0.84	0.83	0.80	0.81	0.80	0.83
	Recall	0.51	0.44	0.52	0.64	0.71	0.63	0.69	0.61
	F1	0.52	0.58	0.64	0.72	0.75	0.71	0.74	0.70
	AUC	0.50	0.72	0.73	0.76	0.76	0.74	0.75	0.75
Subreddit History + POS tag + Empath	Accuracy	0.52	0.75	0.84	0.86	0.87	0.95	0.95	0.96
	Precision	0.56	0.76	0.84	0.94	0.92	0.94	0.95	0.96
	Recall	0.57	0.80	0.86	0.79	0.83	0.96	0.96	0.97
	F1	0.56	0.78	0.85	0.86	0.87	0.95	0.96	0.96
	AUC	0.52	0.75	0.84	0.87	0.87	0.95	0.95	0.96

Table 2 Results when the features are analyzed for every single posts of a user individually

		Baseline	NB	RF	SVM	Log Reg	AdaBoost	Bagging	XGB
POS tag	Accuracy	0.51	0.57	0.59	0.60	0.59	0.61	0.62	0.61
	Precision	0.52	0.61	0.63	0.60	0.61	0.64	0.62	0.65
	Recall	0.53	0.46	0.52	0.70	0.57	0.56	0.69	0.56
	F1	0.52	0.52	0.57	0.64	0.59	0.59	0.65	0.60
	AUC	0.50	0.57	0.59	0.60	0.59	0.61	0.61	0.62
Empath	Accuracy	0.50	0.66	0.67	0.67	0.68	0.69	0.69	0.69
	Precision	0.52	0.85	0.79	0.77	0.78	0.73	0.73	0.77
	Recall	0.52	0.36	0.47	0.47	0.49	0.62	0.62	0.54
	F1	0.52	0.49	0.58	0.58	0.60	0.67	0.67	0.63
	AUC	0.50	0.66	0.68	0.68	0.68	0.69	0.69	0.69
Empath + POS tag	Accuracy	0.51	0.72	0.73	0.58	0.74	0.74	0.75	0.76
	Precision	0.53	0.92	0.84	0.86	0.75	0.78	0.75	0.84
	Recall	0.52	0.43	0.54	0.10	0.80	0.70	0.81	0.64
	F1	0.52	0.57	0.65	0.10	0.77	0.74	0.78	0.72
	AUC	0.51	0.73	0.74	0.60	0.74	0.75	0.75	0.77
Subreddit History	Accuracy	0.49	0.75	0.63	0.94	0.94	0.94	0.94	0.94
	Precision	0.55	0.74	0.61	0.92	0.94	0.93	0.94	0.93
	Recall	0.54	0.85	1.00	0.97	0.96	0.97	0.95	0.98
	F1	0.55	0.79	0.76	0.95	0.95	0.95	0.94	0.95
	AUC	0.49	0.73	0.58	0.93	0.94	0.94	0.93	0.94

mention sadness, it is only a small factor and not a major area of focus in any of these topics.

Which the increase of dependence of technology in our lives, social media can become an important tool, and in the right, hands can be used for our wellbeing. Previous works deals with scouring the net to predict whether a single post gives signs of depression. This may not be the perfect marker as it could correctly predict sadness and may not be able to pinpoint depression.

Moreover, even though post may not show signs of depression at first sight, constant learning and training of machine learning algorithms can detect signs to depression in posts not explicitly dealing with sadness with about 70% accuracy.

The following word cloud is that of posts submitted by depressed people, where the words with the highest frequency are the biggest. Significant words are time, work, people, feel, want, still, one, friend, etc.

The following word cloud is that of posts submitted by non-depressed people, where the words with the highest frequency are the biggest. Significant words know, game, go, make, till, thing, etc.

However, when extracting certain features from the texts posted, we can see that the prediction rate is about the same, which can also be used to train models instead of the texts directly. However, the best results come up when the forums posted on is collected and analyzed.

From the above topics generated, we can see that depressed people usually talk about how bad they feel most of the time, moreover in various contexts and not only when they talk about how they are depressed. Topics either deal with relationships, how nothing is going write and also deals with them asking for help and also a few following up to deal with their issue.

On the other hand, non-depressed people seem to be talking about a wide variety of topics that deals with work, video games, songs, books, looking for a place to talk or someone to connect with, issues dealing with money, property, etc. Though a few mention sadness, it is only a small factor and not a major area of focus in any of these topics.

As rightly said that you can judge a person by their company, it can be shown the subreddits browsed by the user is the most important feature to identify those affected by this mental illness.

References

1. Johnson GJ, Ambrose PJ (2006) Neo-tribes: the power and potential of online communities in health care. Commun ACM 49(1):107–113
2. Taylor MA, Reed R, Berenbaum SA (1994) Patterns of speech disorders in schizophrenia and mania. J Nerv Ment Dis
3. Thomas CB, Duszynski KR (1985) Are words of the Rorschach predictors of disease and death? The case of "whirling." Psychosom Med
4. Schnurr PP, Rosenberg SD, Oxman TE (1992) Comparison of TAT and free speech techniques for eliciting source material in computerized content analysis. J Pers Assess 58(2):311–325
5. Reynolds WM, Gould JW (1981) A psychometric investigation of the standard and short form beck depression inventory. J Consult Clin Psychol 49(2):306
6. Radloff LS (1977) The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1(3):385–401
7. Coppersmith G, Dredze M, Harman C (2014, June) Quantifying mental health signals in Twitter. In: Proceedings of the workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, pp 51–60
8. Lin H, Jia J, Nie L, Shen G, Chua TS (2016, July) What does social media say about your stress? In: IJCAI, pp 3775–3781
9. Akbari M, Hu X, Liqiang N, Chua TS (2016, February) From tweets to wellness: wellness event detection from twitter streams. In Thirtieth AAAI conference on artificial intelligence
10. De Choudhury M, Gamon M, Counts S, Horvitz E (2013) June). Predicting depression via social media, In: Seventh international AAAI conference on weblogs and social media
11. Coppersmith G, Dredze M, Harman C, Hollingshead K, Mitchell M (2015) CLPsych 2015 shared task: depression and PTSD on Twitter. In Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, pp 31–39
12. Coppersmith G, Dredze M, Harman C, Hollingshead K (2015) From ADHD to SAD: analyzing the language of mental health on Twitter through self-reported diagnoses. In: Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality, pp 1–10
13. De Choudhury M, Counts S, Horvitz E (2013, May).Social media as a measurement tool of depression in populations. In Proceedings of the 5th annual ACM web science conference, pp 47–56
14. Pennebaker JW, King LA (1999) Linguistic styles: language use as an individual difference. J Pers Soc Psychol 77(6):1296
15. Tackman AM, Sbarra DA., Carey AL, Donnellan MB, Horn AB, Holtzman NS, ... Mehl MR (2019) Depression, negative emotionality, and self-referential language: a multi-lab, multi-measure, and multi-language-task research synthesis. J Pers Soc Psychol 116(5):817
16. Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, Becker T (2011) Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depress Anxiety 28(6):447–455

17. Park M, McDonald D, Cha M (2013, June) Perception differences between the depressed and non-depressed users in twitter. In: Proceedings of the International AAAI Conference on Web and Social Media (vol 7, No. 1)
18. Gkotsis G, Oellrich A, Hubbard T, Dobson R, Liakata M, Velupillai S, Dutta R (2016, June) The language of mental health problems in social media. In: Proceedings of the third workshop on computational linguistics and clinical psychology, pp 63–73
19. Sekulić I, Gjurković M, Šnajder J (2018) Not just depressed: Bipolar disorder prediction on reddit. arXiv preprint arXiv:1811.04655
20. Weber I, Castillo C (2010, July) The demographics of web search. In Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval, pp 523–530
21. Cheng Z, Caverlee J, Lee K (2010, October) You are where you tweet: a content-based approach to geo-locating twitter users. In: Proceedings of the 19th ACM international conference on Information and knowledge management, pp 759–768
22. Rao D, Yarowsky D, Shreevats A, Gupta M (2010, October) Classifying latent user attributes in twitter. In: Proceedings of the 2nd International Workshop on Search and Mining User-Generated Contents (pp 37–44)
23. Spijker JAN, De Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA (2002) Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 181(3):208–213
24. Fast E, Chen B, Bernstein MS (2016, May) Empath: understanding topic signals in large-scale text. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 4647–4657

A Study on Work–Life Balance and Marital Satisfaction of Faculty Members

V. Kandaswamy Sharma and Beulah Suresh

Abstract This paper manages the outcome of conjugal fulfillment on the work–life parity of faculty members by implementing statistical techniques. Each individual's fantasy is to accomplish the correct harmony between work and life. This is an immense test for employees because of the dynamic changing patterns in training, advancing innovation for assessment and instructing, online training modalities, and elevated expectations of all the stakeholders. Added to this are the increased class strength, availability to students on a 24/7 basis, the race for ranking, and ratings of the Institutions that make them mentally tired and physically weak. Faculty members are struggling to balance their career demands with their personal lives and more often than not, personal lives take a huge hit. Added to this turmoil is the current economic scenario that has put them in such a financial situation that their very survival is at stake. The satisfaction of life is the sum of satisfaction of life at work and at home. Faculty members when they spend more time at work or for work have lesser interaction with their spouses, and marital satisfaction is reduced due to an increase in work–spouse conflict.

Keywords Marital satisfaction (MS) · Work–life conflict · Work–life balance (WLB) · Enhancement · Happiness

1 Marital Satisfaction (MS)

For most cultures, the practice of marriage is a sacred tradition. In India, marriage is a spiritually bound family partnership that binds two families together and is also considered a sacred obligation, thus a spiritual protection and family requirement. Many marriages are arranged by the parents and other members of the family.

V. K. Sharma (✉)
Thiruvalluvar University, Vellore 632115, India

B. Suresh
Department of Business Administration, Auxilium College (Autonomous), Vellore 632006, India

R. J. Kannan et al. (eds.), *International Virtual Conference on Industry 4.0*,
Lecture Notes in Electrical Engineering 355,
https://doi.org/10.1007/978-981-16-1244-2_39

Marital satisfaction (MS) is how much an individual's needs, wishes, and wants are satisfied in their marriage. Marital happiness is a subjective state that "only the individual partner can define, an individual's overall evaluation of their marriage." Marital satisfaction (MS) refers to a global degree of advantage reported by individual spouses in their marital partnership [7]. Marital satisfaction (MS) is also seen as the perception of the overall quality of the marriage by an entity, or the happiness of the person with the marriage.

In the marital satisfaction dynamic target theory [5], it argues that marital ambitions, in particular those prioritized, are the most significant determinant of marital happiness in marriage. People have many goals in their marriage which they want to achieve. Second, the expectations of the various marital goals continuously shift during adulthood. Second, whether the marital goals prioritized in the marriage are achieved at a certain point of growth determines marital fulfillment. Third, other factors may also influence marital satisfaction either by adjusting the expectations of the different marital goals or by encouraging the achievement of the marital goals prioritized.

Working hours, work commitments, career engagement, and managing power are among some of the factors affecting marital discontent [8]. Those with a good family life had more problem-solving communication, more and greater quality time together, more productive contact, less child-rearing tension, less personal discontent, and less economic disputes. Those are all variables that are likely to be connected to requirements and aspirations in the workplace.

2 Work–Life Balance (WLB)

The employment landscape has undergone lots of changes. Business demands immediate resolution and bigger profits. With the arrival of innovations in technologies and faster processing, the work and life which were two separate hemispheres earlier have blended into one, thereby affecting each other. WLB is a significant worry for all the employed. Individuals are finding it hard to manage life and work issues and handle them separately. Work has gradually occupied the living space, thereby the tension between roles of career–life reduces productivity and adversely affects family relationships. Harmony between work and life has of course been compromised. Individuals worry about losing their jobs and having to work harder and longer hours.

The WLB mirrors a socio-mental measurement that produces worth and individual difficulties. WLB [4] is balancing the internal burden from one's aspirations and set reasonable targets. Work–life balance is being constantly segmented from the viewpoint of a life-stage.

WLB is usually associated with the balance between the amounts of time spent and resources devoted to work and other personal interests, to maintain the overall quality of life [1].

Faculty members are not used to the pace remarks [6]. The pace demands attention, longer hours at work, and the need to satisfy the institutions', parents', results, career growth, and family demands. With reduced research grants and students' expectations to be available almost immediately to clarify the doubts, evaluate, and publish the results within 24 h also puts enormous pressure between work and life.

3 Objectives of the Study

Analyze the significance of WLB and MS of faculty members.

Predictive and corrective measures along with suggestions are highlighted.

4 Methodology

A well-structured questionnaire was prepared and circulated to the faculty members of higher educational institutions.

5 Sampling Method

Snowball sampling—a non-probability sampling method. Though the sample population was faculty members from the Vellore District, faculty members were requested to forward to their friends. This move from the faculty members helped the questionnaire to circulate widely. The data was received from faculty members all over India and was not limited to the Vellore District, Tamil Nadu.

The questionnaires were constructed using a five-point Likert scale and were closed-ended.

The questionnaire was circulated through Google forms and had three sections.

(a) Demographic details containing 10 items.
(b) **Work–life balance** containing 15 items. This was based on the scale developed by Fisher-McAuley [2]. The three-dimensional scale measures the WLB on.
(c) Obstruction of work with personal life (OWPL),
(d) Obstruction of personal life with work (OPLW), and
(e) Upgrade of work/personal life (UWPL).
(f) Index of Marital Satisfaction (IMS) Scale containing 25 items

To assess the respondent's feelings about various elements, behaviors, attitudes, and occurrences to describe the degree of dissatisfaction, IMS scale is used [3].

The questions measure similar content like feelings, marriage status, and behaviors and either positively or negatively stated to reduce or eliminate the response set by the respondent. Questions 1,3,5,8,9,11,13,16,17,19,20,21, and 23 were scored reversed.

Based on the scorings of the assessment, the results revealed that the correlation coefficient is very low. This leads to the effect that people who are healthy and fulfilling marital relationships are always under 30 and if not it is above 30 as per the IMS.

6 Statistical Techniques

The data was suitably tabulated after due cleansing and was explained using SPSS 20 software. The statistical tools used to analyze the data were the chi-square test, correlation analysis, ANOVA, factor analysis, and multiple regression analysis.

7 Testing of Adequacy

To test the measure of the proportion of variance among various variables Kaiser–Meyer–Olkin (KMO) factor analysis method is implemented to find the measure of sampling adequacy which helps the researcher to find the causes by the underlying factors. The value of adequacy is 0.927. Based on the independent attribute chi-square test, it is normality which is endorsed by using the significant value (Table 1).

The loading factor "obstruction of work with personal life (OWPL)" has six components. These represent the effect on close to home carries on with because of the impedance of work on close to home life. The loading factor is from 0.820 to 0.678.

The loading factor "obstruction of personal life with work (OPLW)" has four components. These represent the effect on work because of the obstruction of individual life at work. The loading factor is from 0.845 to 0.620. The loading factor "upgrade of work/personal life (UWPL)" has five components. These represent the enhancement of life and work. The loading factor is from 0.715 to 0.503. The variable "my personal life drains strength from me for work" is loaded with the highest value of 0.845.

8 Reliability of WLB and MS

The scale reliability is evaluated to calculate the internal accuracy of the variables and below is given Cronbach's alpha values obtained from the extracted variables.

From Table 2, the Cronbach's alpha of all the factors is more than a lower limit of 0.6, which indicates high internal consistency of the data.

The highest reliability factor is OWPL, and corrective measures have to be taken for WPLE.

Table 1 Factors of work–life balance

Factor name & code	Code	Statement	Factor loading
Obstruction of work with personal life (OWPL)	W2	I miss out on personal events due to work	0.820
	W3	I ignore personal needs due to work	0.796
	W4	I am putting my existence on hold for work	0.776
	W1	My personal life is dominated by work	0.743
	W6	My job hampers my personal life	0.706
	W5	I have difficulty juggling work and non-work	0.678
Obstruction of personal life with work (OPLW)	W12	My personal life drains strength from me for work	0.845
	W13	My work is affecting my personal life	0.827
	W11	I am having trouble working because of my private matters	0.808
	W10	I am exhausted to be working effectively	0.620
Upgrade of work/personal life (UWPL)	W14	I am in a great working mood due to my personal life	0.715
	W15	Personal life gives me vigor in my work	0.690
	W8	My job allows me to go on with personnel activities	0.629
	W9	My job keeps me in a better mood	0.531
	W7	I am content with the proportion of time available for non-business activities	0.503

Table 2 Reliability Analysis between WLB and MS

Factor	Cronbach's alpha	No. of items
OWPL	0.904	6
OPLW	0.848	4
UWPL	0.666	6
MS	0.963	25

9 Relationship between WLB and MS

There is no significant difference between the WLB and M.

To find out the relationship between work–life balance factors and marital satisfaction, Pearson correlation is applied. There is a strong link between obstruction

Table 3 Correlation between WLB and MS

Dimension	Marital satisfaction
OWPL	0.32
OPLW	0.372
UWPL	− 0.327

of work with personal life and obstruction of personal life with work, as well as marital satisfaction (Table 3). There is a negative association between upgrade of work/personal life and marital satisfaction, which means that these two factors move on the same side. For obstruction of work with personal life and obstruction of personal life with work and marital satisfaction, the null hypothesis is rejected and is not rejected for the upgrade of work/personal life and marital satisfaction, and α is statistically confirmed at a level of significance of 1 percent.

10 ANOVA

ANOVA test is applied to WLB and MS with the age gathering of the respondents. The age bunches are grouped into a time of under thirty years, between thirty years to forty years, from forty years to fifty years, or more than fifty years. To analyze the variance, the null hypothesis is set with the end goal that there is no critical contrast between WLB and MS among the different age gatherings of respondents.

The P-Values for both balances between work and life and marital fulfillment are given in Table 4. The WLB parameter variable is considered to be the best predictor and is extremely significant, based on the ANOVA tests. Since the P-value is less than 0.05 for WLB and MS, for the related variables of WLB and MS, the null hypothesis is dismissed at a level of 1 percent significance.

Table 4 Age group and WLB—ANOVA test

		Sum of squares	df	Mean square	*F*	Sig
Work–life balance	Between groups	4.972	3	1.657	7.387	0
	Within groups	45.319	202	0.224		
	Total	50.291	205			
Marital satisfaction	Between groups	7.175	3	2.392	4.006	0.008
	Within groups	120.614	202	0.597		
	Total	127.789	205			

11 Effect of Demographic Variables on Marital Satisfaction

Testing of Hypothesis: There is no significant difference between the demographic and the factors on the fulfillment of wedded life.

From Table 5, it can be established that except for the demographic variable average monthly income, for all the other demographic variables have significant values that are greater than 0.05, and hence, the null hypothesis for these demographic variables is rejected.

12 Multiple Regression Analysis

To test the mathematical model, multiple regression analysis is conducted for the estimation of more than one variable with the independent variable balancing function with existence.

From Table 6, it can be inferred that the R^2 value is 0.077 which proves that 7.7% varies in marital satisfaction can be elucidated with the three independent variables, obstruction of work with personal life, obstruction of personal life with work upgrade of work/personal life (Table 7).

From the above analysis of variance (ANOVA) table, it is to be construed that there is no critical distinction between the factors at a 5% level of significance.

13 Composite Reliability

Composite reliability (also known as construct validity) is a measure of the reliability of scale items (Table 8).

The composite reliability score has to be higher than 0.7, and the values are 0.942, 0.914, and 0.949, respectively, for the three factors of WLB. Based on the above composite reliability measures, there is a high internal consistency between the latent variables.

14 Convergent Validity

Convergent validity, a parameter to check the association between the two measurements of constructs that are highly related.

Only those variables with convergent validity should form part of the study. The average variance extracted (AVE) for each of the latent variables has to be higher than the tolerable threshold of 0.5 (Table 9).

Table 5 Relationship between demographic factors and fulfillment of wedded life

		Sum of squares	df	Mean square	F value	P-value	Status
Age group	Between groups	46.025	64	0.719	1.223	0.164	NS
	Within groups	82.907	141	0.588			
	Total	128.932	205				
Married since	Between groups	44.026	63	0.699	0.896	0.683	NS
	Within groups	107.583	138	0.78			
	Total	151.609	201				
Number of children to be taken care of	Between groups	36.693	64	0.573	1.05	0.399	NS
	Within groups	75.895	139	0.546			
	Total	112.588	203				
Number of elderly relatives to be taken care of	Between groups	90.025	64	1.407	1.092	0.33	NS
	Within groups	179.014	139	1.288			
	Total	269.039	203				
Employment status of spouse	Between groups	47.614	63	0.756	0.756	0.893	NS
	Within groups	136.874	137	0.999			
	Total	184.488	200				
Average hours spent in Institution daily	Between groups	73.413	64	1.147	1.124	0.282	NS
	Within groups	143.912	141	1.021			
	Total	217.325	205				
Average monthly income (In Rs.)	Between groups	159.588	64	2.494	1.645	0.008	S
	Within groups	213.79	141	1.516			
	Total	373.379	205				
The average time taken to reach the workplace	Between groups	50.694	64	0.792	0.656	0.97	NS
	Within groups	168.945	140	1.207			

(continued)

Table 5 (continued)

		Sum of squares	df	Mean square	*F* value	*P*-value	Status
	Total	219.639	204				
Workdays in a week	Between groups	24.134	64	0.377	1.367	0.065	NS
	Within groups	38.881	141	0.276			
	Total	63.015	205				

Table 6 Regression analysis

Model	*R*	*R* square	Adjusted *R* square	Standard error of the estimate
1	0.278	0.077	0.058	0.843

S Significant. *NS* Not significant

Table 7 ANOVA test

Model	Sum of squares	Df	Mean square	*F*	Sig
Regression	11.697	4	2.924	4.117	0.003
Residual	139.912	197	0.71		
Total	151.609	201			

Table 8 Composite reliability

Latent variable	AVE
OWPL	0.942
OPLW	0.914
UWPL	0.949

The above statistical calculations lead to the fact that the AVE values are higher than the threshold value.

Table 9 Convergent validity

Latent variable	AVE
OWPL	0.569969
OPLW	0.608951
UWPL	0.332878

15 Major Findings of the Study

- There is a solid relationship between the upgrade of work/individual life and conjugal fulfillment.
- There is no noteworthy contrast among age groups concerning work–life balance and conjugal fulfillment.
- There is no significant contrast among sexual orientation and conjugal fulfillment in the work–life equalization of employees.
- There is no substantial difference between faculty members' demographic variables and marital satisfaction except for the average monthly salary that does not impact demographic factors.
- The R^2 value is 0.077 which proves that 7.7% varies in marital satisfaction can be elucidated with the three independent variables WIPL, PLIW, and WPLE.
- The value of Cronbach reveals that it is highly significant, consistent, and reliable concerning the latent variables.
- The convergent validity is achieved based on the average variance extracted of each variable, and the values are above the threshold of 0.5.
- In view of the index of marital satisfaction scale (IMS) created by Walter [3], the normal of the IMS size of the 205 respondents was 27.22 which is well beneath the cutting point score of 30. This demonstrates employees have a steady and fulfilling conjugal relationship.

16 Suggestions

Even though there is a positive connection between WLB and MS, it is extremely moderate. Endeavors must be made by businesses, families, and people to help in overseeing work and individual life.

As per the results of the research, unity does not exist if the pressure is transferred to the family members. If there is no peace at home, then it is not possible to work in the organization. The different factors such as mental wellbeing, general happiness, career success, and involvement in social activities lead to marital satisfaction. Consequently, adjusting work and family assumes a fundamental job in conjugal fulfillment.

Communication between life partners is the way into a cheerful life. This is significant as it will assist with overseeing the two contentions and differences very well without harming the texture of conjugal fulfillment.

Mental wellness programs, adequate rest between classes, balanced workload, realistic time management, family support in managing the kids at home, and helping the children with their studies, adequate financial planning, saving, and security will also help to maintain and achieve work–life balance and improve the marital satisfaction.

Counseling sessions are needed to mitigate stress. Pursuing hobbies and activities that give a break from the routine are most welcome.

Exercising together is being advocated as it takes of the health of the couples as well as granting private time. This helps to bring the couples closer.

Having a good married life makes life better.

17 Conclusion

This investigation has endeavored to comprehend two significant ideas of WLB and MS in one endeavor. The key to happiness at work and home is balance and satisfaction. There is a relation between the two, and this study attempts to establish that claim.

The index of the marital scale has been used to measure the marital satisfaction of the faculty members in the Vellore District. This scale would measure and reflect the magnitude of the relations between partners. The minimum score of 0 on the IMS scale is possible and is not an indication that there are no problems between couples. The problems could be insignificant or outside the purview of this questionnaire.

This study has covered the faculty members and can be expanded to include other professions and in other geographical domains. Also, the sample size must be expanded to include a larger set of audience.

Acknowledgements The researcher acknowledges the support of all those who participated in the study and helped to facilitate the research process.

References

1. Clarke MC, Koch LC, Hill EJ (2004) The work-family interface: differentiating balance and fit. Fam Consum Sci Res J 33(2):121–140. https://doi.org/10.1177/1077727X04269610
2. Fisher-McAuley G (2003) Modeling the relationship between work/life balance and organizational outcomes. In: Annual conference of the society for industrial-organisational psychology. Orlando, pp 1–26
3. Hudson WW, Glisson DH (1976) Assessment of marital discord in social work practice. Soc Serv Rev 50(2):293–311. https://doi.org/10.1086/643372
4. Khallash S, Kruse M (2012) The future of work and work-life balance 2025. Futures 44(7):678–686. https://doi.org/10.1016/j.futures.2012.04.007
5. Li T, Fung HH (2011) The dynamic goal theory of marital satisfaction. Rev Gen Psychol 15(3):246–254. https://doi.org/10.1037/a0024694
6. Berg M, Seeber BK (2016) The slow professor: Challenging the culture of speed in the academy. In The slow professor: Challenging cult speed acad. Univ Toronto Press
7. Roach AJ, Frazier LP, Bowden SR (1981) The marital satisfaction scale: development of a measure for intervention research. J Marriage Fam 43(3):537–546. https://doi.org/10.2307/351755
8. Shree RM (2011) Worklife balance & career satisfaction of critical care nurses in private hospitals at Coimbatore. Indian J Appl Res 2(2):113–114. https://doi.org/10.15373/2249555x/nov2012/42