

Volume 47 Number 4 December 2023

ISSN 0350-5596

Informatica

**An International Journal of Computing
and Informatics**



1977

Editorial Boards

Informatika is a journal primarily covering intelligent systems in the European computer science, informatics and cognitive community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor from the Editorial Board can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the list of referees. Each paper bears the name of the editor who appointed the referees. Each editor can propose new members for the Editorial Board or referees. Editors and referees inactive for a longer period can be automatically replaced. Changes in the Editorial Board are confirmed by the Executive Editors.

The coordination necessary is made through the Executive Editors who examine the reviews, sort the accepted articles and maintain appropriate international distribution. The Executive Board is appointed by the Society Informatika. Informatika is partially supported by the Slovenian Ministry of Higher Education, Science and Technology.

Each author is guaranteed to receive the reviews of his article. When accepted, publication in Informatika is guaranteed in less than one year after the Executive Editors receive the corrected version of the article.

Executive Editor – Editor in Chief

Matjaž Gams

Jamova 39, 1000 Ljubljana, Slovenia

Phone: +386 1 4773 900, Fax: +386 1 251 93 85

matjaz.gams@ijs.si

<http://dis.ijs.si/mezi>

Editor Emeritus

Anton P. Železnikar

Volaričeva 8, Ljubljana, Slovenia s51em@lea.hamradio.si

<http://lea.hamradio.si/~s51em/>

Executive Associate Editor - Deputy Managing Editor Mitja

Luštrek, Jožef Stefan Institute

mitja.lustrek@ijs.si

Executive Associate Editor - Technical Editor

Drago Torkar, Jožef Stefan Institute Jamova

39, 1000 Ljubljana, Slovenia

Phone: +386 1 4773 900, Fax: +386 1 251 93 85

drago.torkar@ijs.si

Executive Associate Editor - Deputy Technical Editor Tine

Kolenik, Jožef Stefan Institute

tine.kolenik@ijs.si

Editorial Board

Juan Carlos Augusto (Argentina)

Vladimir Batagelj (Slovenia)

Francesco Bergadano (Italy) Marco

Botta (Italy)

Pavel Brazdil (Portugal)

Andrej Brodnik (Slovenia)

Ivan Bruha (Canada) Wray

Buntine (Finland)

Zihua Cui (China)

Aleksander Denisiuk (Poland)

Hubert L. Dreyfus (USA) Jozo

Dujmović (USA)

Johann Eder (Austria) George

Eleftherakis (Greece)

Ling Feng (China)

Vladimir A. Fomichov (Russia)

Maria Ganzha (Poland)

Sumit Goyal (India) Marjan

Gušev (Macedonia)

N. Jaisankar (India)

Dariusz Jacek Jakóbczak (Poland)

Dimitris Kanellopoulos (Greece)

Dimitris Karagiannis (Austria)

Samee Ullah Khan (USA)

Hiroaki Kitano (Japan)

Igor Kononenko (Slovenia)

Miroslav Kubat (USA) Ante

Lauc (Croatia)

Jadran Lenarčič (Slovenia)

Shiguo Lian (China)

Suzana Loskovska (Macedonia)

Ramon L. de Mantaras (Spain)

Natividad Martínez Madrid (Germany)

Sanda Martinčić-Ipišić (Croatia)

Angelo Montanari (Italy)

Pavol Návrat (Slovakia)

Jerzy R. Nawrocki (Poland)

Nadia Nedjah (Brasil)

Franc Novak (Slovenia)

Marcin Paprzycki (USA/Poland)

Wiesław Pawłowski (Poland)

Ivana Podnar Žarko (Croatia)

Karl H. Pribram (USA)

Luc De Raedt (Belgium)

Shahram Rahimi (USA)

Dejan Raković (Serbia)

Jean Ramaekers (Belgium)

Wilhelm Rossak (Germany)

Ivan Rozman (Slovenia)

Sugata Sanyal (India)

Walter Schempp (Germany)

Johannes Schwinn (Germany)

Zhongzhi Shi (China) Oliviero

Stock (Italy)

Robert Trappl (Austria)

Terry Winograd (USA)

Stefan Wrobel (Germany)

Konrad Wrona (France)

Xindong Wu (USA)

Yudong Zhang (China)

Rushan Ziatdinov (Russia & Turkey)

Honorary Editors

Hubert L. Dreyfus (United States)

An Efficient Meta-Platform for Providing Expert Medical Help to Italian and Slovenian Users

¹Primož Kocuvan, ²Samo Eržen, ³Ivana Truccolo, ⁴Flavio Rizzolio, ¹Matjaž Gams, ¹Erik Dovgan

¹Institut Jožef Stefan, Ljubljana, Slovenija

²ARCTUR Računalniški inženiring d.o.o. Nova Gorica, Slovenija

³ANGOLO Odv, Italian Association of Long-term Cancer Survivors, Italy

⁴Università Ca' Foscari di Venezia, Italy

E-mail: primoz.kocuvan@ijs.si, samo.erzen@arctur.si, ivana.truccolo@gmail.com, flavio.rizzolio@unive.it, matjaz.gams@ijs.si, erik.dovgan@ijs.si

Keywords: EMH, electronic and mobile health, open community, aggregation platform, search bot, recommendation system, Insieme

Received: December 25, 2021

When dealing with medical issues, Google and similar search engines can be of great help. A user inputs a couple of words about a particular medical issue and receives a list of replies that hopefully provide also the desired answer. The problem is that it is far from easy and fast to find the most proper answer in the huge list of related web pages, with some of them even misleading or harmful. Experience shows that a large proportion of the population falls for inadequate medical advices, some alternative and some even dangerous. This paper presents a novel meta-platform for electronic and mobile health-related services with its main purpose to provide proper quality information to a user with a medical issue, kind of an expert medical “doctor Google”. The platform is centered around “services” and human help. A “service” is an intermediate medical concept between a user and the expert of related feed-backs which redirects a user to another website, video, or other multimedia item that contains important more detailed information of proper quality for a user. A user can also search for the related medical issues, e.g., further-links through a local search upon services, and also ask a human or an AI chatbot/assistant. The platform is dedicated to Italian and Slovenian users in Italian, Slovenian and English language.

Povzetek: Članek predstavlja učinkovito meta-platfomo za zdravstveno pomoč italijanskih in slovenskih uporabnikov.

1 Introduction

Digital health-related platforms are transforming the way humans deal with health issues and in this way also society. Patients and physicians are adapting to the new digital mechanisms, which are becoming a prevalent way of providing information in the COVID-19 times. At the same time, it is desired that the platforms are lean, lightweight, and efficient. “Efficient” means that an average person who seeks crucial information about relevant health-related topic receives the correct and medically accurate information in the smallest amount of time. By this, the person who might be unfamiliar with the medical topic should still grasp the new information/knowledge on how to deal with the related issue. Because the system informs the patients (average users or advanced users), and provide advice and general directions thus unloading part of the current overload from health-related workers (doctors, nurses, caregivers, and also pharmacists), they can work in a non-stressful environment where they can dedicate more time to quality treatment of patients. In this article we explore new ways of delivering the correct information to the users by describing our Insieme platform, which stores and

aggregates contributions from companies, associations, organizations, and also individuals.

The article is organized as follows. In Section 2 we present related work. Section 3 describes the Insieme project that aims to develop the Insieme platform. Section 4 concludes the paper with the summary of our work.

2 Related work

The rise of health platforms is an emerging activity that become particularly widespread in the time of the COVID-19 pandemic [2], [4]. People have got information by the world wide web, radio, television, and also by phone, but not as much by physical contact. The pandemic sped up the ICT revolution in healthcare sector [13]. The key feature of the platforms or ecosystems is that the information is available online, where everybody can access it by computer, laptop, tablet, or smartphone. One of the drawbacks of health platforms today is that they are mostly operated by private companies for profit, which raises concerns about privacy and integrity of data [1], [2]. In this article, we first describe some alternative health-related ecosystems, after that, we describe our Insieme platform, which is open to everybody, be it hospitals, health centers, companies, or individuals. The Insieme platform is one of the new Electronic and Mobile Health

platforms (EMH) aiming at providing quality information free of charge to everybody without storing any private data of any kind more than what is needed for providing advices to the users and for personalizing chatbot conversation.

Private and public platform providers are trying novel approaches for the development of the next (future) health-based platforms [3]. For example, they are adopting methods for telemedicine and remote monitoring of patients. Many platforms are collecting data from different sensors and wearable devices. They can detect abnormal cholesterol and glucose blood levels, heart rhythm and falls [11]. In this section, we shortly describe two platforms similar to Insieme. First is the EkoSmart smart city platform with the EMH module, developed within the same project name in Slovenia. It was developed with the cooperation of Jožef Stefan Institute, Faculty of computer science, Faculty of electric and computer science, and 12 private companies from Slovenia. The second platform is Arcadia.

2.1 EkoSmart

The purpose of the EkoSmart project was to develop an ecosystem for smart cities with all the support mechanisms. The result of the project was a distributed ecosystem of tools and systems that one could use on its own, or combine them into a larger system from the existing blocks. EkoSmart originally consisted of six major blocks, with one of them being EMH, i.e., the health ecosystem focusing on flexible architecture for connecting people, based on open data, semantic connection, self-adaptiveness, and self-regulation [7], [9]. Within the project development, the idea for the next-generation platform emerged, resulting in the Insieme project. The idea was to start designing and programming a new only the EMH module of the EkoSmart system, all along consistently building on the expertise obtained through the 3-year project.

2.2 Arcadia

Arcadia is a private company that develops a similar data aggregation ecosystem for health with several interesting functionalities, but unfortunately, it is not open [8]. Also, the functionality of Arcadia is somehow different since they are using and extracting data from Electronic Health Records (EHRs). EHRs are stored on different servers in different formats. They developed very fast procedures for extracting these data from servers (in total around 2 billion records [8]) and displaying it in their application. As a consequence, there is a permanent issue with privacy since the algorithms for searching, recommending, and connecting different data are of the company [10]. Nevertheless, the platform demonstrates several modern concepts and orientations that are somehow shared with those in the Insieme platform.

2.3 Summary

In the Table 1, we show a summary of features for other more popular health platforms which we did not describe

in previous subsections. We also included the Insieme platform for better comparison. We can see that Insieme platform has all three features, while other have at most two of them.

Web platform	Services	Videos	Products
Mayo Clinic	X	X	
Orphanet	X		
WebMD	X		
MedScape	X	X	
Insieme	X	X	X

Table 1: Summary table.

3 Description of the Insieme platform

To understand the “philosophy” of the Insieme platform, the following rule provides the main utility function:

“The golden rule of the platform is that it provides short, precise, and easy to understand information for an average person with a low health literacy level.”

All other design and application choices are heavily influenced by the golden Insieme rule.

The second major concept is that it is an open aggregation platform which means that an open professional community (crowd community) of doctors, nurses, caregivers, companies, and individuals can use and contribute to the platform by adding new services and other modules, as long as they are verified by the editorial board. Service is probably one of the essential building blocks of the system. It is an independent piece of information with the title, description, and other accompanying data, which point to other websites or multimedia information where the user can get more detailed information about products, applications, and user stories of that specific service. Also, in the upgraded version of the platform (version 2.0), one can add other elements in addition to services, such as videos, articles, prototypes, domains, and datasets. By this, the community helps develop and improve people’ health literacy.

One of the key motivations of the Insieme platform is to help the users find the relevant information in much less time and of more quality than if they use Google, Bing, Duckduckgo, or any other search engines. As a consequence, it is efficient in the sense of consuming time to find specific quality information. To guarantee quality information the partners pay attention to the basic health quality information criteria: accuracy, accessibility, relevancy, interpretability [14]. It is a fairly new concept that connects people in all possible ways and not only in the classical doctor-patient framework, while at the same time providing only the trustworthy information doctors and health information professionals find relevant for an average user. The information provided on INSIEME platform is designed to complement, not replace, the

relationship between a patient and his/her own physician (HONcode principles [15]). The platform does not store any personal information on the platform.

The Insieme platform consists of two basic screens. On the first screen, there are:

- Services, presenting basic information about a specific health issue and potential solutions and useful opportunities to deal with it
- Partners, consisting of partners on the project and any Italian or Slovenian partner that provides an EMH service of any kind, be it free of charge or commercial
- Search, for all information in the Insieme platform
- Chatbot assistants, providing natural communication and search through natural language queries
- Recommendation module, recommending future actions
- Contacts with (online) experts who provide expert and trustworthy help in real-time

3.1 Search bot – virtual assistant

The main purpose of the search assistants is that they provide information through human communication, similar to Siri or Google assistant. The developed assistant performs similarly to the search module, i.e., finds proper connections to the Insieme modules, and can also execute some additional communication in the assistant way. For example, it can call external assistants, such as the assistant for waiting queues, JSI assistant, and in the near future will have access to the assistant for stress, anxiety, and depression [12]. Figures 1, 2, and 3 show an example where a user asks “X-ray of lungs” written in the Slovenian language. The query is in Slovene because it fetches the data from the Slovenian National institute for public health website, containing the list of all medical procedures available from all the hospitals and medical centers in the country. Figure 2 shows the buttons where the user can select a Slovenian region. Also, the user selects the level of urgency (not shown).

At the end (Figure 3) the bot answers with a list of medical centers that provide the searched medical procedure. An example is MEDILAB – diagnostic radiological center and UKC Ljubljana.

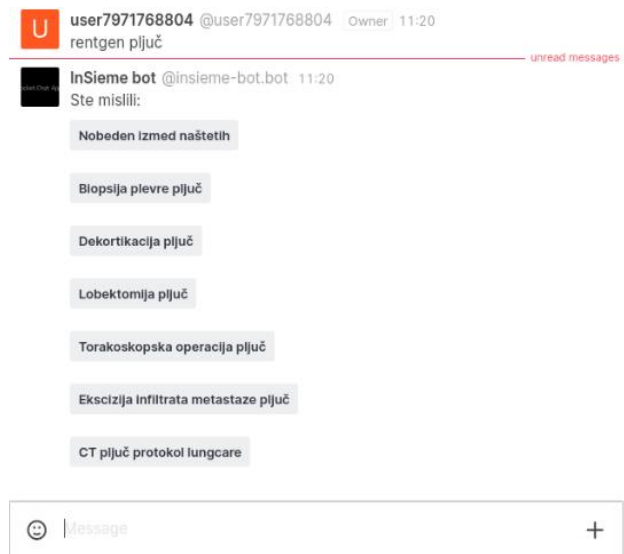


Figure 1: Search string: “X-ray of lungs” in Slovenian. The answer consists of a list of possible replies since the assistant is not sure what the user wants, and therefore lists potential answers that somehow fit the question.

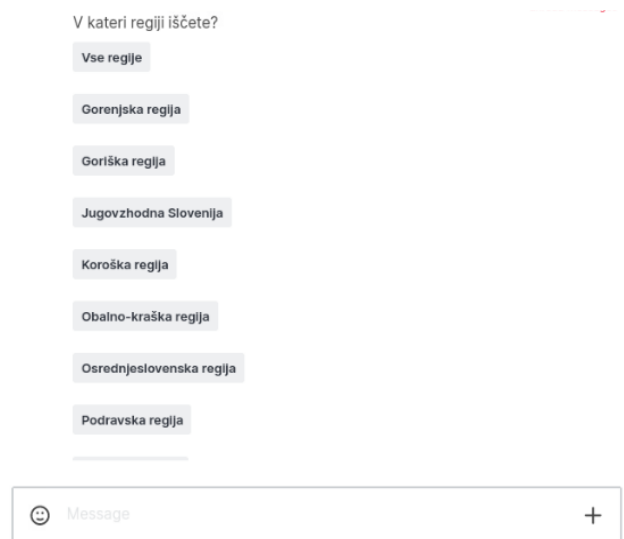


Figure 2: Selection of the region in Slovenia.

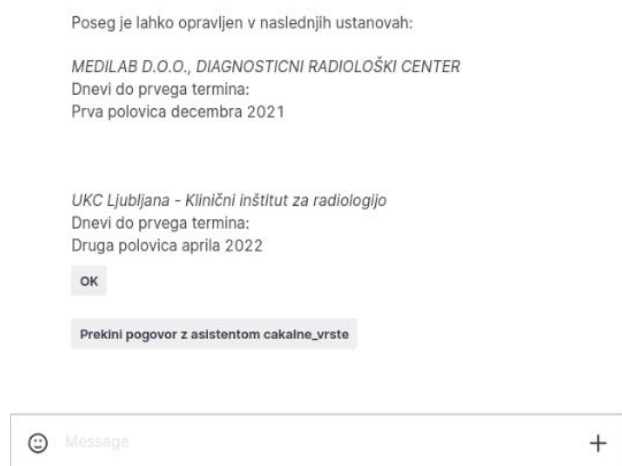


Figure 3: The final result from the bot: best proposed institutions for that particular medical help needed.

3.2 Services

A service presents the basic information about a specific medical problem and potential solutions. Services are connected to medical fields and subfields (such as dermatology and oncology). The structure of a service is the following:

- 1) Title
- 2) Description
- 3) Professional help
- 4) Web and smartphone applications
- 5) Associations
- 6) Articles
- 7) Products
- 8) Forums/User stories
- 9) Videos
- 10) Additional information

This is the advised structure where not all items are included for each service. Also, additional information can be included. The services follow the same core rule of the Insieme platform:

“Provide the most relevant information in the shortest possible time.”

As a consequence, the service structure that was empirically observed to be most efficient as a generalization of previous attempts was summarized in the structure proposed.

The title, description, and professional help as the most relevant information are obligatory data, while other sections can be omitted.

3.3 Example of a service for Rosacea

This is an example of a service Rosacea saved into the field of dermatology. All examples are in English, providing some information for Slovenian and some for Italian Resources.

Description:

Rosacea is a common chronic inflammatory disease that presents with recurrent flushing, erythema, telangiectasia, papules, or pustules on the nose, chin, cheeks, and forehead. Both a prompt diagnosis and appropriate therapeutic interventions are required to prevent permanent scarring, persistent erythema, and ocular sequelae. The incidence of rosacea is estimated to be greater than 5% in the general population, shows a female predominance, and generally occurs in adults between 30 and 50 years of age.

It is advised to omit the triggers (heat, sun exposure, stress, alcohol, etc.). The use of SPF and moisturizing skincare is recommended.

Professional help:

- (SLO) Dermatovenerološka ambulanta, UKC Ljubljana, T: 01 522 37 44, e-pošta: derma.narocanje@kclj.si
- (SLO) Splošna dermatološka ambulanta, UKC Maribor, T: 02 321 27 18, e-pošta: derma.narocanje@ukc-mb.si
- (SLO) Medicinski center Cardial, Ljubljana, T: 01 548 40 80, e-pošta: info@cardial.net
- (ITA) Azienda sanitaria universitaria Giuliano Isontina: SC (UCO) Clinica di Chirurgia Plastica, Trieste, Ospedale di Cattinara, Strada di Fiume 447, Torre Chirurgica, 9° piano. T. 040 3994258.
- (ITA) Istituto Dermatologico Europeo
- (ITA) Skin Doctors' Center

Associations:

- (ITA) Faccia a Faccia con la Rosacea
- (ITA) AITER onlus
- (ENG) National Rosacea Society

Application:

- Diary for disease monitoring (DE)

Additional information:

- Rosacea
- (ITA) Rosacea

3.4 Example of a service for Testicular cancer

This is the second example of a service “testicular cancer”, in the field of oncology.

Description:

Testicular is one of the most common types of cancer in men aged 15 to 35. In previous years the incidence of testicular cancer in developed countries has been rising. It is thought that changes that happen in the early developmental stages of the embryo may cause testicular cancer to develop. The risk factors for developing testicular cancer are undescended testicles at birth, infertility, smaller testicular volume, anomalies of the urinary tract, and previous testicular cancer. In the beginning stages, it manifests as a hard painless lump in the testicles. As the disease progresses symptoms like bone pain, especially back pain, fatigue, weight loss.

Early detection and diagnosis are important for a better prognosis. Men should perform testicular self-examination once a month. It is advised to do the exam after a warm shower or bath. If a hard lump is detected, it is important to visit your family doctor.

Professional help:

- (SLO) KO za urologijo, Univerzitetni klinični center Ljubljana, T: 01 522 26 89, e-pošta: urologija.ambulanta@kclj.si;
- (SLO) Oddelek za urologijo, Univerzitetni klinični center Maribor; T: 02 321 14 47, e-pošta: marjeta.deicman@ukc-mb.si;
- (SLO) Onkološki inštitut Ljubljana, T: 01 587 91 63 ali klicni center OI: T: 080 29 00, e-pošta: info@onko-i.si, triaza@onko-i.si;
- Onkofon za pogovor in svetovanje onkološkim bolnikom, T: 080 23 55;

Applications:

- Ball Checker

Associations:

- testicularcancersociety.org

Video:

- Dr. Oz Teaches Testicular Cancer Self-Check At Home In 3 Easy Steps | TODAY
- What is NOT Testicular Cancer?

Additional information:

- clevelandclinic.org: Testicular Cancer

- testicularcancersociety.org: Testicular self-exam
- Video: How To Check Yourself For Testicular Cancer. Testicular Cancer Society

3.5 Recommendation system

A recommendation system is a mechanism that filters information that is stored in the database and predicts for an individual user the best-rated recommendation. For example, someone who is searching on the booking.com website for a hotel in Ljubljana should have similar recommendations as somebody performing a similar task in the same city. When a person looks for a hotel only in Ljubljana and is not interested in finding one in Koper, Maribor, or any other city in Slovenia, this narrows the space of potential answers. The recommendation system uses attributes of each item, previous searching behaviors, and the collection of data from many users to predict the next item. It utilizes machine learning algorithms and/or hybrid methods. Machine learning is very successful for such tasks, proven in thousands of similar applications. We used the popular Python library “LightFM” [4]. LightFM implements many variations of methods for recommendation systems. The implementation problem with the recommendation system is at the start of its operation, known as “cold start” [6]. In other words, in the beginning, there are no data stored in the database when deploying a fresh application. There are some solutions to this problem, e.g., described in [6], [7]. For the services on the Insieme platform, there are two tables that are a subset of the entire system. Table 2 shows services and their attributes. For simulation purposes, we created a table with users (Table 3). Here one can see that “User 1” searched and visited services that are connected with infections. “User 2” visited services that are connected with cancer and “User 3” with services connected with lungs. By this, it is possible to conclude that for “User 1” we would recommend more services/diseases which are infections, e.g., flu, hepatitis, tetanus. For “User 2” one would recommend cancers, e.g., skin cancer, liver cancer, brain cancer, while for “User 3” lung diseases e.g. asthma, bronchitis, should be recommended.

Table 2: Services and their attributes.

Service/Attributes	Cancer	Infections	Lung
Pneumonia		X	X
Lung cancer	X		
Covid-19		X	X
Skin cancer	X		

Table 3: Services and user choices.

Service/User choice	User 1	User 2	User 3
Pneumonia	X		X
Lung cancer		X	X
Covid-19	X		X
Skin cancer		X	

As presented in this example, the machine learning algorithms are connecting attributes and user choices by using conditional reasoning and cross-sections of sets.

4 Discussion

By comparing Insieme platform to other meta medical platforms available online (gathered in Table 1) there is one crucial difference. The Insieme platform offers direct connection between medical practitioners or medical data and users (patients). Essentially it helps people to get the correct medical information in very short time. It is an aggregating platform with products, videos, services, faculties, sellers and manufacturers. This is a novelty in the electronic and mobile health domain.

5 Conclusion

This article described a new way of making connections of users with medical problems to their solutions via the Insieme (EMH) platform. All the functionality of the platform is dedicated to this issue using a variety of approaches to accommodate the user as much as possible. For example, Insieme stories a large and heterogeneous set of replies and connections provided to the users, provides human help from online experts, presents lists of institutions providing relevant services or help, and shows further links, videos, and other multimedia functions.

Insieme also provides search and conversation with the bot in natural language regarding the services, partners, or appointments for specialists. We provided the text of two randomly chosen services as an example. Each service aggregates all the carefully chosen information a medical expert would provide as advice to a general audience with a particular problem. Data aggregation websites are not new, but we used this concept in a new way in a particularly demanding field, which is medicine. The platform is open, meaning that everyone who has an internet connection and has some professional medical knowledge can contribute to the platform. In addition, the platform has administrators and moderators with the main task of evaluating and accepting quality new services.

The essential question is whether the two countries or at least one of them will be able to find proper use of the Insieme platform. The best would be if a particular national or municipality institution would support the transition of the designed system into actual massive use.

Based on the major efforts taken to provide as fast, reliable, and precise information to general health issues

and the years-long experience with similar platforms and services the team expects that such a system would be of great help to citizens.

In our initial tests, the Insieme platform on average enables finding a proper information in a couple of minutes, while it took about ten to a couple of ten minutes for an average user to find similar information through a general search engine.

Acknowledgement

The paper was supported by the ISE-EMH project funded by the program Interreg V-A Italy-Slovenia 2014-2020. The authors also acknowledge the financial support from the Slovenian Research Agency (research core funding No. P2-0209). We also thank to all medical students who contributed to the Insieme platform with adding services.

References

- [1] Lam, Lawrence G (2001) Digital Health-Data platforms: biometric data aggregation and their potential impact to centralize Digital Health-Data, *MIT thesis*
- [2] Dorsey, E.R. The new platforms of health care. *npj Digit. Med.* **4**, 112 (2021). <https://doi.org/10.1038/s41746-021-00478-5>
- [3] Samad, Zainab & Mahmood, Sana & Siddiqui, Sameen. (2020). Imagining a lean and agile digital health ecosystem – a measure of pandemic responsiveness. *Journal of Global Health*. 10. 10.7189/jogh.10.020391.
- [4] Alex H Krist, Robert Phillips, Luci Leykum, Benjamin Olmedo, Digital health needs for implementing high-quality primary care: recommendations from the National Academies of Sciences, Engineering, and Medicine, *Journal of the American Medical Informatics Association*, Volume 28, Issue 12, December 2021, Pages 2738–2742, <https://doi.org/10.1093/jamia/ocab190>
- [5] Primož Kocuvan, Erik Dovgan, Matjaž Gams (2021). An Analytical and Empirical Comparison of Electronic and Mobile Health Platforms. *Insieme project Workshop*, Jožef Stefan Institute, Ljubljana.
- [6] Gjorgji Noveski, Jakob Valič (2021). Analysis of a recommendation system used for predicting medical services. *Insieme project Workshop*, Jožef Stefan Institute, Ljubljana.
- [7] Short description – EkoSmart, Accessible at: <https://feri.um.si/raziskovanje/mednarodni-projekti-in-projekti-strukturnih-skladov/ekosistem-pametnega-mesta/> [1.12.2021]
- [8] Arcadia aggregating EHR Accessible at: <https://arcadia.io/products/aggregate-your-data/> [1.12.2021]
- [9] Official website of the EkoSmart project – health related Accessible at: <https://ekosmart-agregator-asistentov.docker-e9.ijs.si/> [1.12.2021]
- [10] Melissa J. Harvey, Michael G. Harvey (2014), Privacy and security issues for mobile health platforms, <https://doi.org/10.1002/asi.23066>

- [11] J. J. Rodrigues Barata, R. Munoz, R. D. De Carvalho Silva, J. J. P. C. Rodrigues and V. H. C. De Albuquerque, "Internet of Things Based on Electronic and Mobile Health Systems for Blood Glucose Continuous Monitoring and Management," in *IEEE Access*, vol. 7, pp. 175116-175125, 2019, doi: 10.1109/ACCESS.2019.2956745.
- [12] Kolenik, T.; Gams, M. PerMEASS—Personal Mental Health Virtual Assistant with Novel Ambient Intelligence Integration. CEUR-WS. 2020. pp. 8–12. Available online: <http://ceur-ws.org/Vol-2820/AAI4H-2.pdf> (accessed on 23 April 2022)
- [13] Weiner, J.P. Doctor-patient communication in the e-health era. *Isr J Health Policy Res* 1, 33 (2012). <https://doi.org/10.1186/2045-4015-1-33>
- [14] Al-Jefri M, Evans R, Uchyigit G and Ghezzi P (2018) What Is Health Information Quality? Ethical Dimension and Perception by Users. *Front. Med.* 5:260. doi: 10.3389/fmed.2018.00260
- [15] HONcode principles, Accessible at: https://www.hon.ch/HONcode/Guidelines/hc_p2.html [5.1.2022]

Real Time QoS in WSN-based Network Coding and Reinforcement Learning

Amra Sghaier and Aref Meddeb

ISITCom, NOCCS Lab, University of Sousse, Tunisia

E-mail: sghaier_amra@hotmail.fr, aref.meddeb@eniso.u-sousse.tn

Keywords: Real time QoS, Reliability, end-to-end delay, energy consumption, packet delivery rate, throughput, packet error rate, packet loss rate.

Received: March 29, 2021

In recent years, wireless sensor networks have experienced significant advancements, driven by a reduction in development costs. This rapid growth in WSNs has led to the emergence of various potential and emerging applications, including real-time applications, which pose challenges due to their substantial requirements. As the number of applications continues to increase, ensuring both reliable and real-time Quality of Service (QoS) communication in resource-constrained WSNs becomes a paramount concern. To address this challenge, we propose the use of network coding (NC), a novel research area applicable in diverse environments to overcome several shortcomings within a network. Additionally, we focus on the duty cycle, recognized as one of the most popular techniques for energy conservation. Specifically, we employ the Duty Cycle Learning Algorithm (DCLA) to determine the optimal duty cycle. To guarantee the expected real-time QoS and reliability, we introduce NCDCLA (Network Coding-based Duty Cycle Learning Algorithm). Through simulations in OPNET, our results demonstrate that our approach achieves commendable reliable performance.

Povzetek: Nova metoda NCDCLA za zagotavljanje zanesljivosti QoS v WSN je zasnovana kot kombinacija mrežnega kodiranja in algoritma za učenje delovnega cikla.

1 Introduction

Recent advancements in micro-electro-mechanical systems (MEMS) and wireless communications have garnered attention toward small sensor nodes communicating with each other using radio signals. These sensors are compact, possess limited processing and computing resources, and are cost-effective compared to traditional sensors. These sensor nodes are capable of sensing, measuring, and collecting information from the environment. Following a local decision-making process, they can transmit the sensed data to the user [1].

However, certain characteristics of WSN pose challenges due to limited resources such as energy, bandwidth, memory, processing power, and transmission power. Despite these constraints, WSNs have emerged as one of the most intriguing areas of research, given their diverse applications including military sensing, environmental monitoring, and target tracking. Consequently, some applications present significant challenges due to their extensive requirements in terms of:

- *Real time QoS:* Critical application must support such time bound called deadline. For this, data should be delivered before its deadline.
- *Reliability:* The reliability of a WSN is the probability that end-to-end communication is successfully completed. In other words, reliable data transfer ensures that packets reach their destination.

- *Energy efficient:* Energy consumption must be highly constrained.

For this reason, ensuring communication reliability in resource-constrained wireless sensor networks remains an open area of research to achieve a high degree of real-time Quality of Service (QoS).

To address these challenges, the utilization of network coding and duty cycle learning algorithms has been demonstrated to enhance the performance of wireless sensor networks. Network coding, a novel technique that has garnered significant interest in recent studies, was originally proposed in information theory in 2000 by Ahlswede et al. [2]. The core premise of network coding is that intermediate nodes, referred to as relays, engage in coding operations on the incoming data stream to generate outflows. These nodes recombine incoming data using operations such as the XOR operation. Consequently, network coding offers improvements over traditional routing, where nodes typically perform simple operations like receiving and retransmitting packets.

On the other hand, duty cycling is considered one of the most critical energy conservation techniques. Duty cycling involves periodically placing a node into sleep mode, which is an effective method for reducing energy dissipation in wireless sensor networks (WSNs) [3]. To ensure Quality of Service (QoS) levels, [4] proposed a Duty Cycle Learning Algorithm (DCLA) that adapts the duty cycle during runtime without the need for human intervention, aiming

to minimize power consumption while balancing the probability of successful data delivery and meeting application delay constraints.

DCLA is a mechanism based on reinforcement learning (RL), an area of machine learning that enables machines and software agents to automatically determine ideal behavior within a specific context to maximize performance [5]. It has proven successful in addressing various functional challenges of wireless sensor networks, including energy awareness, real-time routing, query processing, event detection, localization, node clustering, and data aggregation [6].

Our challenge is to implement a new paradigm, NCD-CLA (Network Coding-based Duty Cycle Learning Algorithm) [7][8][9][10], which holds the potential to offer significant benefits across various communication network metrics, such as throughput, delay, energy efficiency, wireless resources, security, complexity, and resilience to link failures.

The remainder of the paper is organized as follows: In the next section, we summarize related work. Section 3 surveys several key technologies fundamental to our study of network coding and DCLA in WSN. We evaluate the performance of our approach in section 4 and conclude by outlining directions for our future work in section 5.

2 Related works

In recent years, network coding has been investigated as a method to achieve improvements in wireless networks [11]. Additionally, many researchers in the field of network coding have underscored the importance of this technique. Reference [12] outlines the two main benefits of this approach: potential throughput improvements and a high degree of robustness.

The effectiveness of a network-coding strategy depends on the context, and in the case of Wireless Sensor Networks (WSNs), it should leverage the broadcast nature of the medium while considering the capacity limitations of the nodes [13]. Furthermore, [14] proposed an enhanced AdapCode schema. This schema allows for the reduction of power consumption for the entire network and prolongs the lifetime of the network by minimizing packet communications throughout the code dissemination process.

The ultimate goal of the work described in [15] is to improve network efficiency and extend its lifetime. This solution reduces the overall volume of data transfer. Moreover, in [16], the authors described some drawbacks of applying network coding in real-world sensor network scenarios. However, authors [17] proposed and investigated the use of network coding to improve real-time performance in IEEE 802.15.4-based wireless sensor networks. They developed a performance model that analytically characterizes the real-time performance of a single M/M/1 node with network coding.

According to the network coding technique for packet

encoding, [18] proposed the NCQ-DD routing protocol, which can efficiently conserve bandwidth and node energy to improve the efficiency and accuracy of data transmission. The delivery rate of packet groups is also enhanced. [19] analyzed two robust implementations of network coding for transmission in sensor networks. Wang X. et al. [20] suggested a network coding-based approach in data dissemination to achieve rapid dissemination, thereby reducing energy consumption and decreasing delay.

Additionally, [21] introduced CodeDrip, a data dissemination protocol for Wireless Sensor Networks. The main concept behind this protocol is to apply Network Coding to the dissemination process, reducing the number of transmitted messages and consequently saving energy consumption. CodeDrip requires additional space in the packet to store message IDs and buffers to store combined messages. These overheads can be controlled by specifying the maximum number of messages that can be decoded and the maximum buffer size. Moreover, [22] investigated the concept of network coding in Wireless Sensor Networks (WSN) and presented Re-CoZi, a packet transport mechanism that uses medium-aware advanced acknowledgment mechanisms to provide reliable network coding-based communications over lossy environments. Also, Lie Wang et al. [23] proposed a multirate network coding scheme to improve the energy efficiency of WSNs. This scheme can enhance energy efficiency in three aspects: reducing the number of re-encoding nodes without compromising the performance of network coding, transmitting more data in a transmission period, and working over a very small finite field.

To achieve energy efficiency in WSN, [24] proposed DutyCode, a network coding-friendly MAC protocol that implements packet streaming and allows the application to decide when a node can sleep. Through analysis and real system implementation, it is demonstrated that DutyCode does not incur higher overhead and achieves 20-30% more energy savings compared to network coding-based solutions that do not use duty-cycling.

Duty cycle has demonstrated efficiency in numerous studies aimed at balancing objectives to ultimately extend the lifetime. Several works aim to adapt the service cycle mechanism to enhance performance. Euhana et al. [25] proposed OWR, a practical opportunistic routing scheme based on duty cycle, which exhibits significant improvements in terms of energy efficiency, delay, and resilience to sink dynamics. Sukumar and Aditya [26] introduced a QoS-aware MAC protocol in which the MAC layer utilizes the network layer's next-hop information for better adaptation of the duty cycle based on DSS delay. Smita and Prabha [27] suggested a MAC protocol with adaptive duty cycle that gradually adjusts the contention window, offering very high throughput and low delay characteristics. Pangun et al. [28] proposed an adaptive optimal duty cycle algorithm running on top of the IEEE 802.15.4 medium access control to minimize power consumption while meeting reliability and delay requirements. However, the adaptation of the duty cycle introduces the possibility of two

cases: a small duty cycle increasing the delay and a higher duty cycle reducing energy efficiency. For this reason, the adaptation of the duty cycle becomes crucial.

3 NCDCLA protocol design

In this section, the system model of NCDCLA is described. NCDCLA aims to enhance real-time communication in IEEE 802.15.4 beacon-enabled mode. NCDCLA is integrated into the MAC sublayer and the application layer. The duty cycle adaptation algorithm is incorporated into the MAC sublayer, while network coding is integrated into the application layer.

3.1 Markov decision process and reinforcement learning

Markov decision processes (MDPs) provide a mathematical framework for modeling decision-making in situations where outcomes are partly random and partly under the control of a decision maker. MDPs are an intuitive and fundamental formalism for reinforcement learning (RL).

This algorithm consists of a sequence of numbered steps (S, A, T, R):

- S : is a discrete set of environment states $S = \{s_1, s_2, \dots, s_n\}$
- A : is a set of actions from each state $A = \{a_1, a_2, \dots, a_n\}$
- T : is the transition probability from state s to a successor state s'
- $R(a, s)$: is the reward function

In the process of reinforcement learning, an agent interacts with its environment through rewards. At each step t , the agent chooses an action a from the set of actions available and a reward r . The agent seeks to maximize the total reward it accumulates in the long run. The agent is guided by a coordinator that, at first glance, selects a set of states such as the energy-saving level l .

The DCLA agent employs a reinforcement learning technique known as Q-learning to find the optimal policy Π^* which is represented by the value function in a two-dimensional table indexed by state-action pairs. Mathematically, the optimal policy is defined as:

$$\Pi^*(s) = \operatorname{argmax}(Q^*(s, a)) \quad (1)$$

After each step t , the Q value is updated as follows :

$$Q_{(t+1)}(s, a) = Q_t(s, a) + \alpha[R(s, a) - Q_t(s, a)] \quad (2)$$

Every new Q value is computed as the sum of the old value and a correction term α . $R(s, a)$ is the reward function and is defined as follows:

$$R(s, a) = r_t + \gamma \max_{a' \in A(s')} Q(s_{(t+1)}, a_{(t+1)}) \quad (3)$$

The idea is to maximize exploration by selecting random actions during a large number of iterations, allowing for several cycles of reward exploration from the initial state to a goal state. At each passage, the algorithm reinforces the quality of the action that leads to rewards for $nbCycle$. The algorithm stops when all possible states are visited, and the exploration rate $TauxExp$ decreases to a probability of ϵ determining the optimal Q function Q^* . The pseudo-code for the duty cycle learning algorithm is defined as follows:

Algorithm 1 $Q_Learning$

```

Input:  $l, s, \gamma$ 
Output:  $\Pi$ 
Begin
 $Q(s, a) \leftarrow 0$ 
for  $1 < i < nbCycle$  do
   $currentState \leftarrow l$ 
  for  $1 < j < nbAction$  do
     $s \leftarrow currentState$ 
     $TauxExp \leftarrow random(0, 1)$ 
    if  $TauxExp < \epsilon$  then
       $a \leftarrow randomAction(s)$ 
    else
       $a \leftarrow \operatorname{argmax}_{a'}(Q(s, a'))$ 
    end if
     $s' \leftarrow a(s)$ 
    Rewards_Evaluation ( $\Pi$ )
     $Q(s_t, a_t) = Q(s_t, a_t) + \alpha[r_t + \gamma \cdot \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$ 
    if  $s' = desiredState$  then
       $Exit$ 
    end if
  end for
end for
End

```

$$r_t = r_e + r_u + r_d + r_o \quad (4)$$

The reward r_t in reinforcement learning is computed as the sum of four components: energy r_e , super frame utilization r_u , delay r_d , and queue occupation r_o rewards. This computation enables the DCLA agent to learn the optimal duty cycle. The pseudo-code for the evaluation of rewards is defined as follows:

3.2 Network coding

The fundamental assumption of network coding is that intermediate nodes, referred to as relays, are utilized to perform coding operations on the incoming stream, resulting in outflows. These nodes recombine incoming data using the XOR operation. Network coding offers advancements beyond traditional routing, where nodes typically perform only simple operations such as receiving and retransmitting packets.

fig. 1 depicts the system model of the network coding. There are two sources A, B , and one destination (sink).

Algorithm 2 Rewards_Evaluation

```

Input:  $r_t$ 
Output:  $s_t$ 
Begin
 $sum \leftarrow 0$ 
for  $s \in S$  do
  for  $i = 0$  to  $a = |s|$  do
     $sum \leftarrow sum + r_t(s_t, \Pi(s))$ 
     $s_t \leftarrow Q(s, \Pi(s))$ 
  end for
end for
End

```

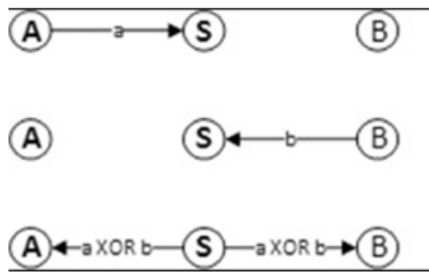


Figure 1: System model of network coding

Both sources broadcast their data messages to the relay S and to the destination. The relay combines the incoming data a and b to produce the data message $a \oplus b$. Next, the relay sends $a \oplus b$ to the sink.

Following the principle of traditional routing, we have four transmission units. However, with the model presented in Figure 1, we have three transmission units. Therefore, the gain is 3/4. Network coding provides benefits in terms of delay because the data will be transmitted after three transmissions instead of four transmissions. It also offers benefits in energy consumption because the relay broadcasts the input data after combining only once. Additionally, it provides benefits in bandwidth because the channel will be occupied for a shorter duration.

In our scheme, the packet coding operation is performed by the PAN coordinator. The network coding model consists of two interfaces, Sender and Receiver, representing the PAN coordinator operating the concepts of network coding. The transmitting interface sends the packet that has undergone network coding to the sink. The receiving interface sends the incoming packets to the nc_proc function precisely via the nc_value attribute to code them, and then the combined packet will be relayed to the sink as an output stream.

In addition, the transmitter interface contains other parameters such as $Bitsize_{of_N}C_value$ and $Output_{N}C_value$. The first parameter defines the code size used in network coding operations (in our case, equal to 2), and the second parameter indicates how the nc_value is calculated. The latter is randomly taken as a zero value (in other cases, it can be a specified number). Thus, there

is another parameter, $Timeout_value$, that is responsible for the waiting time in the queue before the packets are coded. If the incoming packets are directly processed, then $Timeout_value$ is set to zero.

The pseudo-code of network coding is defined as follows:

Algorithm 3 nc_proc

```

Input:  $P_i$ 
Output:  $P_e$ 
Begin
if  $Bitsize\_of\_NC\_value = 1$  then
  send packet without coding
   $Timeout\_value \leftarrow 0$ 
end if
for each packet  $P_i$  in queue do
  if  $Bitsize\_of\_NC\_value = 2$  then
    coding packet with XOR operator
     $P_e \leftarrow P_i \otimes P_{i+1}$ 
     $Timeout\_value \leftarrow 0$ 
  else
    coding the two first packet with XOR operator
     $P_e \leftarrow P_i \otimes P_{i+1}$ 
     $Timeout\_value \leftarrow Timeout\_value + 1$ 
  end if
  send coding packet to sink
end for
End

```

In this scheme, the principle of network coding involves the agent coordinator encoding all received packets. Subsequently, it relays the encoded packet to the sink.

3.3 NCDCLA

A system model of NCDCLA is considered with N sensor nodes scattered uniformly in an area. The nodes are named based on their role in the network. The nodes are differentiated into three groups:

- Sink: receives and decodes data
- Agent Coordinator (AC): encodes data received and retransmit the generated data to the sink. AC are duty cycle learning algorithm enabled
- Node (N): Sensor node senses, gathers and transmits data to the AC.

The processing flowchart of NCDCLA is defined as follows in Figure 2.

4 Experimental results

4.1 Simulation models

In this section, we assess our proposed scheme using OPNET simulator v14.5 [29]. For the simulation, we examine a small network comprising 10 Micaz nodes randomly

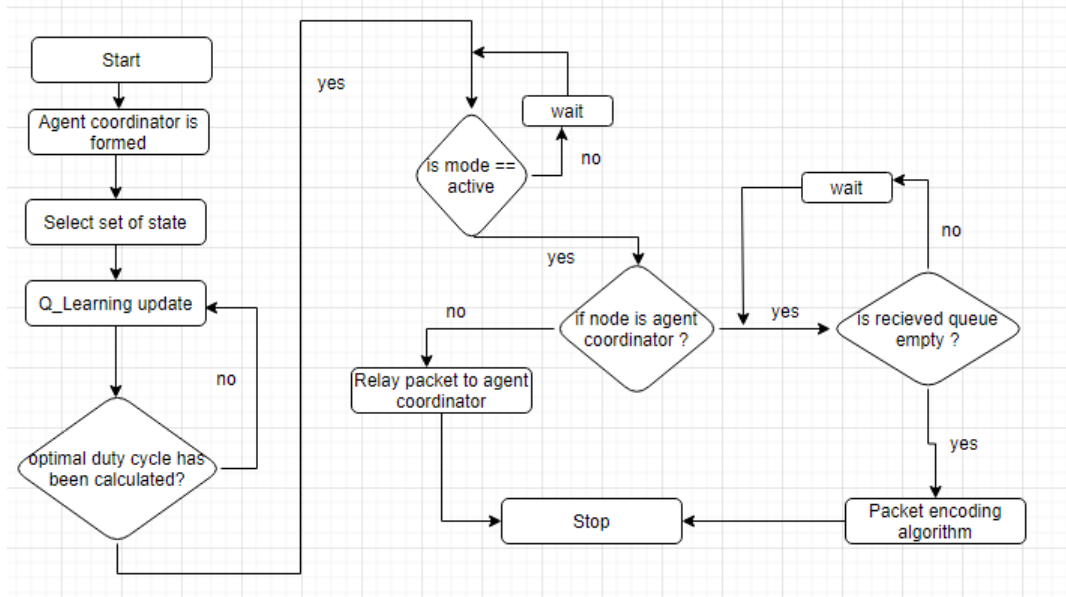


Figure 2: Flow chart of NCDCLA

placed within a 1000x1000m area. The primary parameters utilized in the simulations are provided in Table 1.

Parameter	Value
Data Rate (Kbps)	250
Packet size (bits)	120
Number of node	10
Initial energy (mAh)	16
Learning rate α	0.1
Discount factor γ	0.5
Bitsize_NC_value	2
Output_NC_value	0
Timeout_value (s)	0

Table 1: Parameters used by the NCDCLA simulations

A set of performance metrics is considered, including energy consumed, energy remaining, end-to-end delay, packet delivery rate, throughput, bit error rate, and signal-to-noise ratio. The definitions of these metrics are given below:

- Energy Consumed: The total energy used by all nodes in the network.
- Energy Remaining: The total energy remaining for all nodes in the network.
- End-to-End Delay: The total sum of transferred packets across a network from the source to the sink node.
- Throughput: The rate of successfully received data packets by the node per unit time.
- Packet Delivery Ratio: The ratio of packets successfully delivered to the sink node compared to the total

number of packets sent by all sensor nodes in the network.

- Packet Loss Rate: Corresponds to the acceptance or rejection of a packet, respectively.
- Retransmission Attempts: The number of retransmissions due to collision or channel errors.
- Transmission Success: Indicates the success of transmitting a packet.

4.2 Performance of NCDCLA

The simulation results provide a comparison among IEEE 802.15.4, network coding with duty cycle, and NCDCLA.

Ensuring end-to-end delay is a crucial Quality of Service (QoS) parameter for forwarding data in a time-constrained Wireless Sensor Networks (WSNs) environment. This parameter is defined as the total delay, including MAC delay and queuing delay, between the sending and reception of a packet.

Scheme	Queuing delay(s)	MAC delay(s)
IEEE 802.15.4	0,013	0,029
NC+ duty cycle	0,00016	0,00093
NCDCLA	0,000077	0,000064

Table 2: Comparisons of performance

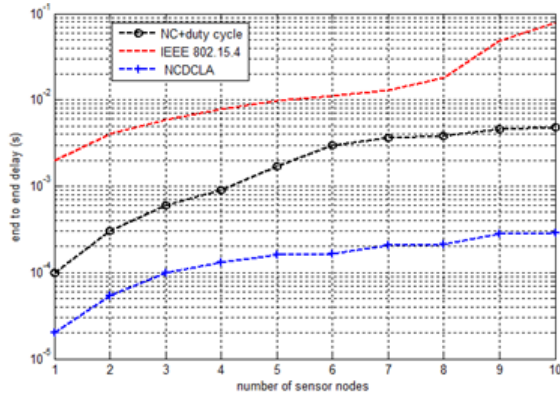


Figure 3: The average end-to-end delay

fig. 3 shows the average end-to-end delay. Based on the results, we observed that NCDCLA exhibits slightly lower end-to-end delay compared to other schemes. Similarly, table 2 presents the MAC delay and queuing delay of our scenario with 10 nodes. We observed that the queuing delay and MAC delay of NCDCLA are lower than those of other schemes because NCDCLA adapts the duty cycle, taking into account the parameter of end-to-end delay. Additionally, network coding (NC) reduces the average end-to-end delay.

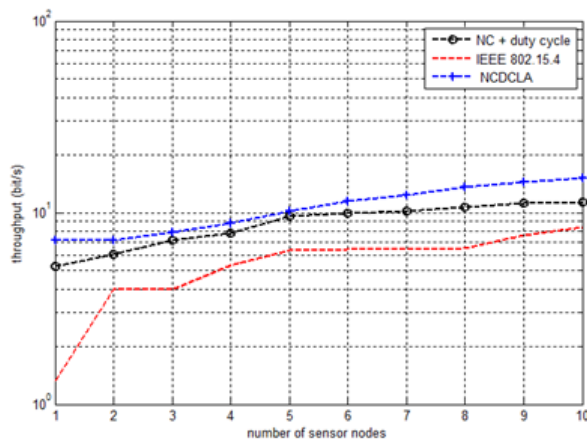


Figure 4: The average throughput

In this scenario, we calculate the average throughput. As depicted in Figure 4, we can observe that as the number of sensor nodes increases, the average throughput also increases. Thus, we confirm that throughput is directly proportional to the increased number of sensor nodes. Additionally, we observed that NCDCLA exhibits slightly

higher throughput compared to other schemes because the utilization of network coding could potentially double the throughput.

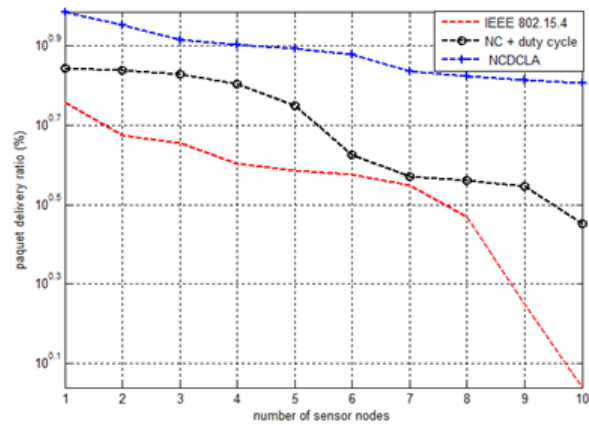


Figure 5: The average packet delivery ratio

Primarily, the packet delivery ratio is another significant parameter for evaluating the performance of QoS. Figure 5 illustrates the average packet delivery ratio. According to these results, we observed that NCDCLA has a higher packet delivery ratio compared to other schemes. NCDCLA achieves 80% when the number of sensor nodes is 10, whereas NC with duty cycle and IEEE 802.15.4 achieve 39% and 28%, respectively.

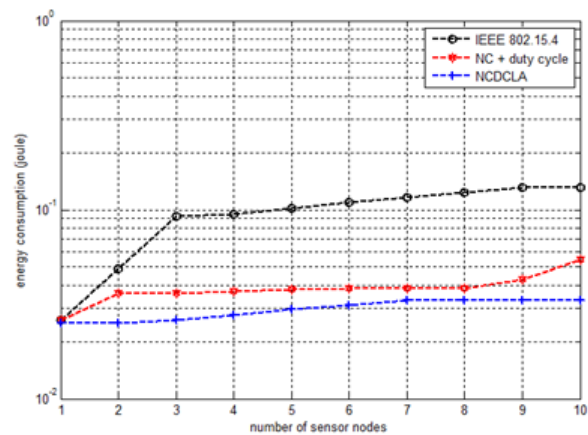


Figure 6: The average energy consumption

The energy consumption can significantly impact Quality of Service (QoS), making real-time applications particularly challenging due to their demanding requirements in terms of end-to-end delay, throughput, packet delivery ratio, and, concurrently, energy consumption and network lifetime.

fig. 6 displays the average energy consumption, where IEEE 802.15.4 exhibits slightly higher energy consumption compared to NC with duty cycle and NCDCLA. The lower energy consumption for NCDCLA and NC with duty cycle is attributed to the use of the duty cycle mechanism. Furthermore, NCDCLA demonstrates slightly lower energy

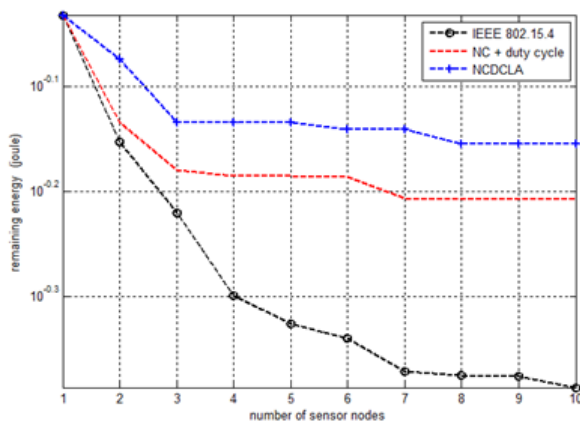


Figure 7: The average remaining energy

consumption than NC with duty cycle because network coding reduces the number of transmissions, and NCDCLA adapts the duty cycle to minimize energy consumption, thereby extending the network’s lifetime, as illustrated in Figure 7.

Critical applications are sensitive to packet loss. Figure 8 shows the average packet loss ratio. It is evident that our scheme exhibits a slightly lower packet loss rate compared to other schemes. The lower packet loss rate for NCDCLA is attributed to the optimal duty cycle, which relies on packet loss rewards. Additionally, the concept of network coding contributes to reducing packet loss.

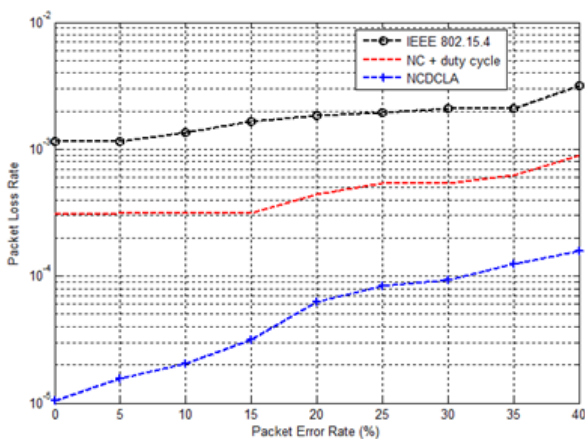


Figure 8: The average packet loss rate

fig. 9 displays the average retransmission attempts for packet error rate. As depicted in this figure, NCDCLA exhibits a slightly lower number of transmission attempts compared to other schemes. The lower number of retransmissions can be attributed to the use of network coding, which reduces the packet loss rate and increases the transmission success, and the optimal duty cycle, which aims to decrease the packet loss rate.

Furthermore, retransmission attempts can impact energy consumption. As illustrated in Figure 10, NCDCLA demonstrates slightly lower energy consumption. How-

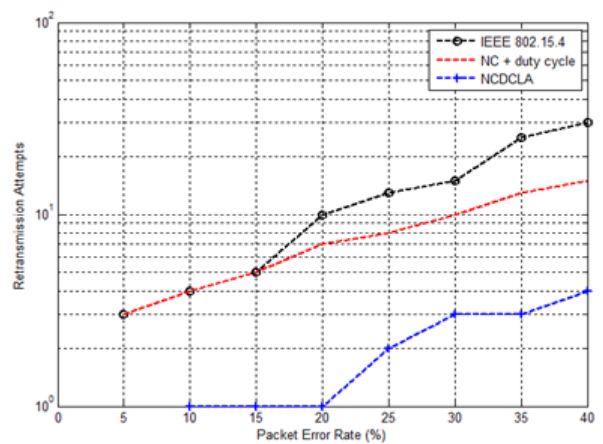


Figure 9: The average retransmission attempts

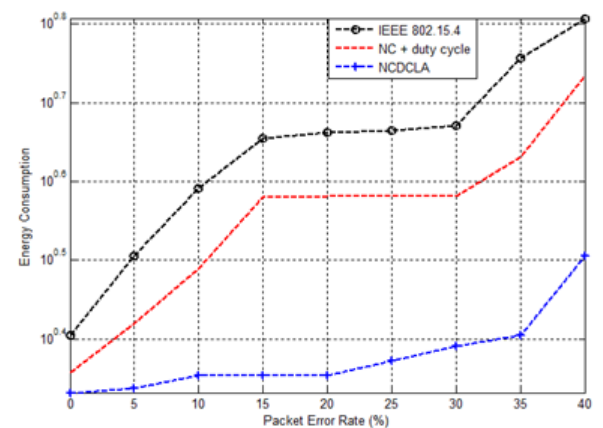


Figure 10: The average energy consumption

ever, IEEE 802.15.4 and NC with duty cycle exhibit a higher level of energy consumption.

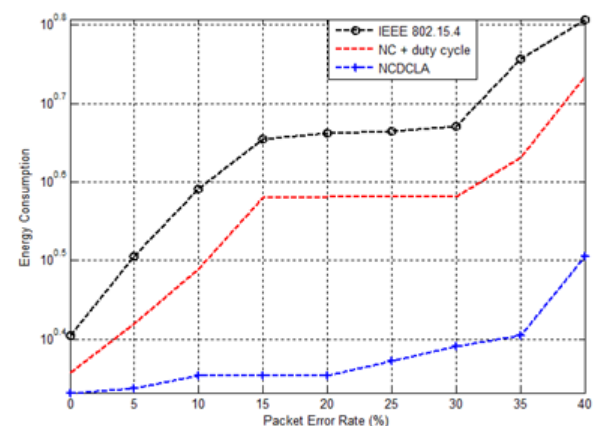


Figure 11: The average transmission success

fig. 11 illustrates the average transmission success. Based on the results, it is observed that as the packet error rate increases, the transmission success decreases. However, NCDCLA exhibits a higher transmission suc-

cess compared to other schemes. When the packet error rate is 40%, NCDCLA achieves a transmission success rate of 55%, whereas IEEE 802.15.4 and NC with duty cycle achieve 24% and 30%, respectively.

5 Conclusion

Evolving technology has spurred a growing demand for real-time applications in wireless sensor networks. Consequently, supporting Quality of Service (QoS) has become a pivotal challenge. In this paper, we introduce a novel scheme, Network Coding and Duty Cycle Learning Algorithm (NCDCLA), designed to address QoS concerns in Wireless Sensor Networks (WSN). The simulation results demonstrate that NCDCLA significantly enhances performance across various metrics, including energy efficiency, delay, throughput, packet delivery ratio, packet loss rate, and transmission success.

Acknowledgement

We appreciate the time and effort invested by the editor in reviewing this manuscript. Our sincere thanks go to the esteemed reviewers for their valuable comments and suggestions, which have contributed to enhancing the quality of this paper.

References

- [1] Pandey, A. K., & Srivastava, S. (2016). *Survey on wireless sensor routing protocols*. Int J Sci Res, 5(5), 1145-1149. <https://doi.org/10.21275/v5i5.nov163558>
- [2] Ahlswede, R., Cai, N., Li, S. Y., & Yeung, R. W. (2000). *Network information flow*. IEEE Transactions on information theory, 46(4), 1204-1216. <https://doi.org/10.1109/18.850663>
- [3] Saraswat, J., & Bhattacharya, P. P. (2013, February). *A Study on Effect of Duty Cycle in Energy Consumption for Wireless Sensor Networks*. In IJCA Proceedings on Mobile and Embedded Technology International Conference 2013 (No. 1, pp. 43-48). Foundation of Computer Science (FCS). <https://doi.org/10.5121/ijcnc.2013.5109>
- [4] de Paz, R., & Pesch, D. (2010, August). *Dcla: A duty-cycle learning algorithm for ieee 802.15. 4 beacon-enabled wsns*. In International Conference on Ad Hoc Networks (pp. 217-232). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-17994-5_15
- [5] Mahajan, S. (2014). *Reinforcement Learning: A Review from a Machine Learning Perspective*. International Journal, 4(8). <https://doi.org/10.1201/b17476-11>
- [6] Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H. P. (2014). *Machine learning in wireless sensor networks: Algorithms, strategies, and applications*. IEEE Communications Surveys & Tutorials, 16(4), 1996-2018. <https://doi.org/10.1109/comst.2014.2320099>
- [7] Sghaier, A., & Meddeb, A. (2018, March). *NCDCLA: QoS aware Network Coding based Duty Cycle Learning Algorithm for Real time and Reliable Wireless Sensors Networks*. SSS'18, Hammamat-Tunisia
- [8] Sghaier, A., & Meddeb, A. (2018, November). *Enhancing Guaranteed end-to-end delay in IEEE 802.15. 4*. In 2018 International Conference on Smart Communications and Networking (SmartNets) (pp. 1-5). IEEE. <https://doi.org/10.1109/smartnets.2018.8707395>
- [9] Sghaier, A., & Meddeb, A. (2018, December). *Model Based Validation of Real Time QoS for NCDCLA Protocol in Wireless Sensor Networks*. In International conference on the Sciences of Electronics, Technologies of Information and Telecommunications (pp. 361-372). Springer, Cham. https://doi.org/10.1007/978-3-030-21009-0_35
- [10] Sghaier, A., & Meddeb, A. (2022, December). *Comparative Study of QoS-aware Network Coding Protocols in WSNs*. In 2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA) (pp. 1-5). IEEE. <https://doi.org/10.1109/aiccsa56895.2022.10017702>
- [11] Yang, S., & Koetter, R. (2007, June). *Network coding over a noisy relay: a belief propagation approach*. In Information Theory, 2007. ISIT 2007. IEEE International Symposium on (pp. 801-804). IEEE. <https://doi.org/10.1109/isit.2007.4557108>
- [12] Fragouli, C., Le Boudec, J. Y., & Widmer, J. (2006). *Network coding: an instant primer*. ACM SIGCOMM Computer Communication Review, 36(1), 63-68. <https://doi.org/10.1145/1111322.1111337>
- [13] Miao, L., Djouani, K., Kurien, A., & Noel, G. (2012). *Network coding and competitive approach for gradient based routing in wireless sensor networks*. Ad Hoc Networks, 10(6), 990-1008. <https://doi.org/10.1016/j.adhoc.2012.01.001>
- [14] Hou, I. H., Tsai, Y. E., Abdelzaher, T. F., & Gupta, I. (2008, April). *Adapcode: Adaptive network coding for code updates in wireless sensor networks*. In INFOCOM 2008. The 27th Conference on Computer Communications. IEEE (pp. 1517-1525). IEEE. <https://doi.org/10.1109/infocom.2008.211>
- [15] Skulic, J., & Leung, K. K. (2012, September). *Application of network coding in wireless*

- sensor networks for bridge monitoring*. In Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on (pp. 789-795). IEEE.
<https://doi.org/10.1109/pimrc.2012.6362891>
- [16] Voigt, T., Roedig, U., Landsiedel, O., Samarasinghe, K., & Prasad, M. B. S. (2012). *On the applicability of network coding in wireless sensor networks*. ACM SIGBED Review, 9(3), 46-48.
<https://doi.org/10.1145/2367580.2367588>
- [17] Aoun, M., Argyriou, A., & van der Stok, P. (2011, February). *Performance evaluation of network coding and packet skipping in ieee 802.15. 4-based real-time wireless sensor networks*. In European Conference on Wireless Sensor Networks (pp. 98-113). Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-642-19186-2_7
- [18] Zhu, M., Zhang, D., Ye, Z., Wang, X., & Wang, J. (2015). *A network coding based routing protocol in wireless sensor networks*. International Journal of Future Generation Communication and Networking, 8(2), 365-372.
<https://doi.org/10.14257/ijfgcn.2015.8.2.30>
- [19] Sanson, J. B., Gomes, N. R., & Machado, R. (2013, January) *Optimization of wireless sensor network using network coding algorithm*. In The Twelfth International Conference on Networks (ICN) (p. 21).
<https://doi.org/10.1049/cp:20070330>
- [20] Wang, X., Wang, J., & Xu, Y. (2010). *Data dissemination in wireless sensor networks with network coding*. EURASIP Journal on Wireless Communications and Networking, 2010(1), 465915.
<https://doi.org/10.1155/2010/465915>
- [21] Júnior, N. D. S. R., Vieira, M. A., Vieira, L. F., & Gnawali, O. (2014, February). *CodeDrip: Data dissemination protocol with network coding for wireless sensor networks*. In European Conference on Wireless Sensor Networks (pp. 34-49). Springer International Publishing.
https://doi.org/10.1007/978-3-319-04651-8_3
- [22] Salhi, I., Ghamri-Doudane, Y., Lohier, S., & Rousel, G. (2011, October). *Reliable network coding for zigbee wireless sensor networks*. In Mobile Adhoc and Sensor Systems (MASS), 2011 IEEE 8th International Conference on (pp. 135-137). IEEE.
<https://doi.org/10.1109/mass.2011.122>
- [23] Wang, L., Yang, Y., Zhao, W., Xu, L., & Lan, S. (2014). *Network-coding-based energy-efficient data fusion and transmission for wireless sensor networks with heterogeneous receivers*. International Journal of Distributed Sensor Networks, 10(3), 351707.
<https://doi.org/10.1155/2014/351707>
- [24] Chandanala, R., & Stoleru, R. (2010, June). *Network coding in duty-cycled sensor networks*. In Networked Sensing Systems (INSS), 2010 Seventh International Conference on (pp. 203-210). IEEE.
<https://doi.org/10.1109/inss.2010.5572223>
- [25] Ghadimi, E., Landsiedel, O., Soldati, P., Duquenooy, S., & Johansson, M. (2014). *Opportunistic routing in low duty-cycle wireless sensor networks*. ACM Transactions on Sensor Networks (TOSN), 10(4), 67.
<https://doi.org/10.1145/2533686>
- [26] Nandi, S., & Yadav, A. (2011). *Cross layer adaptation for QoS in WSN*. arXiv preprint arXiv:1110.1496.
<https://doi.org/10.5121/ijcnc.2011.3520>
- [27] Pawar, S., & Kasliwal, P. (2012). *A QoS Based Mac Protocol for Wireless Multimedia Sensor Network*. IOSR Journal of Electronics and Communication Engineering (IOSRJECE), 1(5), 30-35.
<https://doi.org/10.9790/2834-0153035>
- [28] Park, P., Ergen, S. C., Fischione, C., & Sangiovanni-Vincentelli, A. (2013). *Duty-cycle optimization for IEEE 802.15. 4 wireless sensor networks*. ACM Transactions on Sensor Networks (TOSN), 10(1), 12.
<https://doi.org/10.1145/2529979>
- [29] Dunaytsev, R. (2010). *Network Simulators: OPNET Overview and Examples*. Lecture Slides, Department of Communications Engineering, Tampere University of Technology, 2-69
<https://doi.org/10.1201/b12515-5>

A Framework for Evaluating Distance Learning of Environmental Science in Higher Education Using Multi-Criteria Group Decision Making

Katerina Kabassi

Department of Environment, Ionian University, Minotou Giannopoulou 26, Panagoula, 29100 Zakynthos, Greece
E-mail: kkabassi@ionio.gr

Keywords: e-Learning, environmental science, evaluation criteria, Fuzzy AHP

Received: January 2, 2021

Purpose: Due to Covid-19, big changes took place in Universities around the world. Universities were asked in March 2020 within a short while to provide the whole of the available lessons using e-learning methods. Since the health crisis continues, e-learning has expanded on a variety of contexts and simultaneously has created an urgent need for designing, developing, and implementing a valid and reliable distance learning procedure. The validity and efficiency of the aforementioned procedure is a major issue that has to be tested at the end of the semester. Therefore, developing a valid framework to evaluate the e-learning process has become more important now than in the past due to the ongoing pandemic. *Design/methodology/approach:* The evaluation of the educational process is a multi-criteria problem that is based on the points of view of both instructors and students. In order to solve this multi-criteria problem of e-learning evaluation, a new framework for evaluating e-learning in Higher Education has been developed. This framework uses group decision-making with multiple criteria and is called ENVEL. This paper defines the set of evaluation criteria and uses the Fuzzy Analytic Hierarchy Process to prioritize criteria and help the decision-makers draw conclusions on the effectiveness and success of e-learning. *Findings:* The framework takes into account heterogeneous groups of students and professors, makes different calculations for these groups, and can extract useful conclusions by comparing the results of the different groups. The framework has been applied in the Department of Environmental Science at the Ionian University and conclusions have been made on its effectiveness and usage. *Originality:* Trying to focus on the evaluation of e-learning in a whole study program in Higher Education, and not only on single courses, the paper describes a novel framework for e-learning evaluation using multi-criteria decision-making with heterogeneous groups of users. This framework provides a formal way of combining different aspects of the evaluation of e-learning and collecting summative results.

Povzetek: Raziskava uvaja okvir ENVEL za ocenjevanje e-učenja v visokem šolstvu z uporabo večkriterijskega skupinskega odločanja in metode Fuzzy AHP.

1 Introduction

E-Learning has garnered increasing attention in Higher Education in the last few decades (Martín-Lara & Rico 2020, Njenga 2017, Otto & Becker 2018, Schieffer 2016). Several case studies for the application of e-learning in higher education have been reported (e.g. Sulčić & Lesjak 2009, Al-Fadhli & Khalfan 2009, Bhadauria 2016; Sheikh & Gupta 2018). However, a lack of usage at the university level was clear (Mercader & Gairin 2020). Indeed, before the COVID-19 pandemic, e-learning was growing by approximately 15.4% yearly in educational institutions around the world without pressure on teachers, students, or institutions (Alqahtani

& Rajkhan 2020). Since the health crisis continues, e-learning has expanded on a variety of contexts and simultaneously has created an urgent need for designing, developing, and implementing a valid and reliable distance learning procedure. The validity and efficiency

of the aforementioned procedure is a major issue that has to be tested at the end of the semester.

However, as Stöhr et al. (2020) report, previous studies have mainly focused on asynchronous online learning, rather than synchronous or mixed modes of online learning (Hrastinski 2008, Siemens et al. 2015). Furthermore, as Barteit et al. (2020) point out, the effectiveness of e-learning was mainly evaluated by comparing e-learning to other learning approaches such as traditional teaching methods or evaluating students' and teachers' attitudes (Frehywot et al. 2013).

Several systematic reviews and meta-studies on the effectiveness of e-Learning on single courses have been conducted (Wu & Hwang 2010, Noesgaard & Ørngreen 2015, Abdalla 2007, Liaw 2008, Haverila & Barkhi 2009), but there is a lack of similar experiments that would evaluate e-learning adoption in a whole study program. Some of the studies on e-learning system evaluation focused on the technology-based components

(Islas et al. 2007), others focused on the human factors (Liaw et al. 2007), and others are meta-reviews (Salter et al. 2014).

Taking into account that most reports on e-learning mainly focused on evaluating single courses and not a whole study program and the fact that the effective adoption of e-learning can only be confirmed by evaluating the educational process, we notice that there is a great need for a tool that would evaluate e-learning adoption of a whole study program and not a single course. The effectiveness of e-learning depends on several factors and criteria (Harrati et al. 2016, Abuhlfaia & Quincey 2019, Alqahtani & Rajkhan 2020) and Jeong & González-Gómez (2020) highlight the necessity of determining those.

As the evaluation of e-learning is affected by several factors and criteria that try to combine the points of view of different decision-makers, multi-criteria group decision-making may be found effective for designing a formal framework for e-learning evaluation. Indeed, MCDM has been used in the past for evaluating e-learning systems and applications (Mahdavi et al. 2008, Stecyk 2019, Çelikkilek & Adıgüzel Tüylü 2019, Alqahtani & Rajkhan 2020, Jeong & Gonzalez-Gomez 2020). However, these approaches did not focus on evaluating the e-learning of a whole study program.

Therefore, this paper focuses on presenting a framework for evaluating e-learning of a study program in Higher education that is called ENVEL. Its name originates from the first application of the framework in the Department of Environment (ENVironment E-Learning evaluation). The framework defines the groups of decision-makers, the set of criteria, and the weights of their importance in the reasoning of the decision-makers while evaluating e-learning. The framework considers the instructors and students participating in the educational process as decision-makers and provides a formal way of combining different aspects of the evaluation of e-learning using Multi-Criteria Decision Making (MCDM) and collecting summative results.

MCDM has evolved rapidly over the last decades (Zopounidis 2009) and different decision approaches have been proposed. These approaches differ in the way the objectives and alternative weights are determined (Mohamadali & Garibaldi 2011). The Analytic Hierarchy Process (Saaty 1980) is one of the most popular MCDM theories and has been used before for combining criteria for e-learning success, but for single courses or systems (Anis & Islam 2015, Vinogradova & Kliubas 2015, Jasim et al. 2018, Alqahtani & Rajkhan 2020). The AHP is chosen amongst other MCDM theories because it presents a formal way of quantifying the qualitative criteria of the alternatives, in this way removing the subjectivity of the result (Tiwari 2006).

As Erensal et al. (2006) point out, the conventional AHP may not fully reflect a style of human thinking as users usually feel more confident in giving interval judgments rather than expressing their judgments in the form of single numeric values. The theory's combination with the fuzzy theory resulted in Fuzzy AHP (FAHP) (Buckley 1985), which in comparison with other MCDM

methods is considered by many researchers (e.g. Ramanayaka et al. 2019) as a more effective solution to solve MCDM related problems because of its powerful ability to deal with imprecise and uncertain data. Furthermore, the method's ability to make decisions by making a pairwise comparison of uncertain, qualitative, and quantitative factors and also its ability to model expert opinion (Mulubrhan et al. 2014) is another important reason for its selection against other alternatives. As a result, FAHP has been used before for combining and prioritizing criteria in e-learning systems' evaluation (Tai et al. 2011, Anggrainingsih et al. 2018, Lin 2010, Altun Turker et al. 2019, Naveed et al. 2020).

Given the above advantages of FAHP, ENVEL uses the particular theory to prioritize criteria. The framework has been applied in the Department of Environmental Science at Ionian University for evaluating the e-learning conducted in the special circumstances that occurred during the spring semester of 2019-2020 due to the Coronavirus emergency. 14 professors and 98 students of the Department that took part in the e-learning participated in the evaluation experiment.

The paper is organized as follows: Sections 2, 3, and 4 describe the framework ENVEL. More specifically, section 2 focuses on the criteria used in the evaluation process, section 3 on the prioritization of the criteria, and section 4 on the evaluation of the e-learning aspects. Section 5 presents how the Department of Environment at Ionian University turned to e-Learning during the spring semester of 2019-2020 and section 6 describes a case study, which involves the application of ENVEL in the specific department for the evaluation of the whole study program provided by e-learning methods. Section 7 includes a discussion of the results of the evaluation conducted using ENVEL and proposals that could improve the whole e-learning process. Finally, in the last section, the conclusions regarding the ENVEL framework are presented.

2 ENVEL: Defining criteria for e-learning evaluation

Different MCDM theories and criteria have been used for evaluating e-learning systems. The most common approaches to evaluations of e-learning systems that use MCDM are presented in Table 1. Most of these frameworks use two levels of criteria and a combination of two MCDM models. However, the criteria used in these approaches mainly concern the technology used and the way that courses are designed for e-learning. Furthermore, these frameworks mainly focus on the

Table 1: The questions of the questionnaire and their connection with the criteria of ENVEL

	Mahdavi et al. 2008	Stecyk 2019	Çelikkbilek & Adıgüzel 2019	Tüylü & Alqahtani & Rajkhan 2020	Jeong & Gonzalez-Gomez 2020
Levels of criteria	2 levels	1 level	2 levels	1 level	2 levels
No of criteria	1 st level – 4 criteria 2 nd level – 13 criteria	10 criteria	1 st level – 3 criteria 2 nd level – 19 criteria	10 criteria	1 st level – 4 criteria 2 nd level – 16 criteria
What is evaluated	Web-based E-Learning Systems	e-learning course	Components of e-learning systems	e-learning approaches	e-learning systems
MCDM models	AHP, Entropy and TOPSIS Approach	PROME THEE II	fuzzy F-DEMATEL analytic network process	AHP&TOP SIS	F-DEMATEL/MCDA method

Table 2: The questions of the questionnaire and their connection with the criteria of ENVEL

C1Functionality of the system.	
c11: Accessibility	How easy was the access to the e-learning platform?
c12: Response time	The response/upload time was...
c13: Reliability	Was the system reliable (e.g. did you face connection problems, data loss, etc.)?
c14: Easy to use/simplicity	How simple was the usage of the system?
c15: e-Learning Management	The management of the educational material was...
C2Quality of communication.	
c21: Quality of Synchronous Communication student-instructor	The quality of Synchronous Communication student-instructor was...
c22: Quality of Synchronous collaboration between students	The tools for students’ collaboration were...
c23:Students’ participation	How was the students’ participation in comparison to the lessons in the classroom? The active participation of the students in the e-learning lessons was...
C3 e-LearningReadiness.	
c31: e-LearningCulture	Did you have previous experience in e-learning? In the past, were you in favor of e-learning?
c32: e-Learning Support	The quality of the technical support provided by the department was...
c33: e-Learning Infrastructure	The tools for synchronous and asynchronous e-learning provided by the department were...
C4Quality of e-Learning.	
c41: Effectiveness	How would you judge the effectiveness of e-learning?
c42: Acceptability	How do you judge the experience in e-learning during this semester? Would you like to continue e-learning after the end of the COVID-19 era?
c43: Course design	The design of synchronous and asynchronous lessons was...

evaluation of one or two courses and not the evaluation of a whole study program.

Criteria are also used in the cases where no MCDM model is applied. To determine the effectiveness of learner-controlled e-learning, research typically distinguishes between learning processes and learning outcomes (Alavi & Leidner 2001, Gupta & Bostrom 2009). However, Martinez-Caro (2018) argues that there is an absence of a solid and objective measure of learning, and concludes that perceived learning is positively related to satisfaction in e-learning courses. This point of view is in line with Richmond et al. (1987) who proposed the use of a subjective measure: the students' perceived learning, which refers to the extent to which a student believes s/he has acquired specific skills. As a result, the best way of evaluating e-learning is to focus on the students' perceived self-efficacy and perceived satisfaction (Liaw 2008). Another factor that is considered important while evaluating e-learning involves teacher-student interaction (Su et al. 2005, Harasim et al. 1995, Hartman et al. 1995, Elkaseh et al. 2016). Such interaction encourages learners to understand the content better (Su et al. 2005) and students who are shy or uncomfortable about participating in class discussions may prefer participation in online forums (Owston 1997). Additionally, the way the e-learning environment promotes collaboration and generally the interaction between students is of exceptional importance (Benbunan-Fich & Hiltz 2003, Arbaugh 2004) and considered as a critical component of quality education (Anderson 2001). In some cases, students have even expressed a preference for online dialogue over traditional classroom discussion (Clark 2001, Martinez-Caro 2011).

A new way of interaction and learning, which requires a high level of learner control may result in students having negative attitudes or experiencing difficulties (Chou & Liu 2005). Arbaugh's (2008) study confirms that students' prior experience with e-learning can positively affect their implementation. This is in line with the work of Marks et al. (2005), which suggests that students with experience in e-learning courses may perform better in other e-learning courses. As a result, the prior experience may influence the current e-learning experience. Therefore, both of these criteria are to be taken into account while evaluating the e-Learning experience.

Other factors that have been reported to be taken into account in studies that measure the students' perceived satisfaction in e-learning settings involve service quality, system quality, content quality or e-learning quality, learner perspective/attractiveness, instructor attitude, supportive issues, etc. (Aguti et al. 2014, Reiser 2001, Tseng et al. 2011, Salter et al. 2014).

Some researchers claim that the effectiveness of e-learning depends on demographic factors such as age, gender, etc. (Islam et al. 2011), while others argue against these hypotheses (Marks et al. 2005, Martinez-Caro 2018). Therefore, such criteria were not taken into account at all.

Given the above, the criteria used within the ENVEL framework for evaluating the e-Learning process are the following:

C1: Functionality of the system. In this category, all criteria are related to the quality of the system.

c11: Accessibility. The system makes learning materials easily accessible.

c12: Response time. The waiting time for loading learning materials is reasonable.

c13: Reliability. The e-Learning system provides the right solution to learner requests.

c14: Ease of use/simplicity. The user interface should be simple and easy to use.

c15: e-Learning Management. The easiness in designing the e-Learning Process.

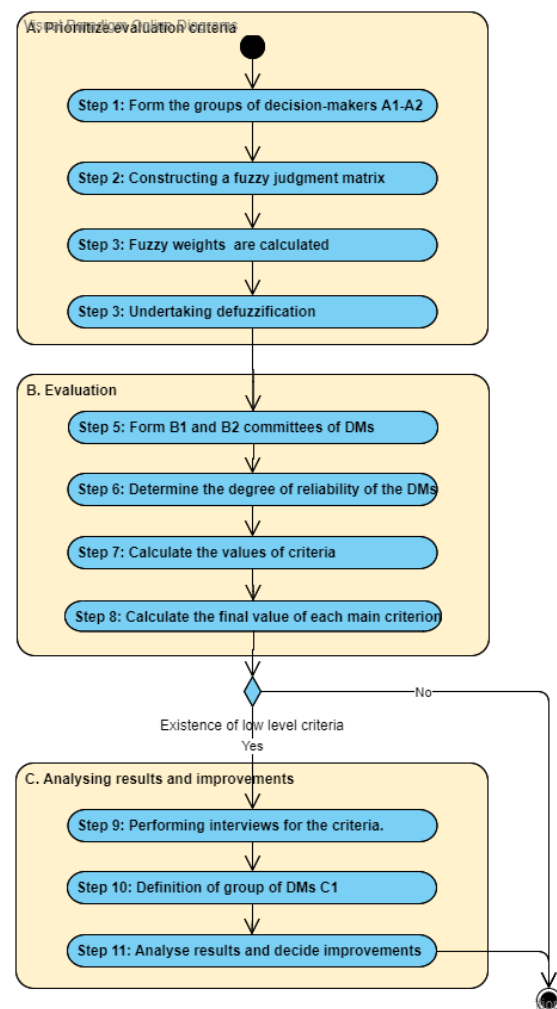


Figure 1: Steps of the ENVEL.

C2: Learner Attractiveness. All the criteria are related to the Learner's Attractiveness.

c21: Quality of Synchronous Communication with the instructor. Quality of synchronous communication with the instructor.

c22: Quality of Synchronous Communication with students. Quality of synchronous communication among students.

c23: Students' participation. Quality of students' participation in the distance lesson.

C3 e-Learning Readiness. In this category, all criteria are related to the way the instructors integrated e-learning.

c3.1: e-learning Culture. Beliefs and attitudes towards e-learning.

c3.2: e-Learning Support. Staff mentoring and support in providing e-learning.

c3.3: e-Learning Infrastructure. Tools provided (recording, blackboard, scheduling, etc.). Especially in the case of a department with laboratories and activities on the field, e-Learning Infrastructure may also involve the tools used for capturing and reproducing the functionality and atmosphere of laboratories and/or activities on the field.

C4: Quality of e-Learning. In this category, all criteria are related to the quality of e-Learning

c41: Effectiveness. The general evaluation of the effectiveness of the current e-Learning experience.

c42: Acceptability. The acceptability of the current e-learning experience.

c43: Course design. The course has been structured correctly.

In order to implement the evaluation experiment and estimate the values of the criteria, a questionnaire was designed. Table 2 presents the questions of the questionnaire and their connection with the criteria for the evaluation of the e-learning experience.

After the criteria have been defined and the questionnaire has been designed, ENVEL consists of 11 main steps that are presented in Figure 1. Steps 1-4 are presented in Section 3 and do not have to be repeated every time ENVEL is implemented. Steps 5-11 are presented in section 4 and the first four are obligatory to run in every evaluation experiment while the last three are only implemented if criteria with low scores occur.

3 ENVEL: prioritize evaluation criteria

According to ENVEL, the e-learning experience is measured using specific evaluation criteria. However, these criteria are not equally important in the evaluation process. For this purpose, ENVEL uses the Fuzzy Analytic Hierarchy Process (FAHP) (Buckley 1985). The steps of the theory for the criteria prioritization are the following:

1. Form the groups of decision-makers A1-A2:

Two sets of decision-makers (DMs) that involve human experts and students are set. All professors and students should have experience in e-learning so that they can make decisions on the importance of the criteria. The appropriate choice of experts is of great importance because only in this way the framework would give reliable and valid results. These groups are called A1 and A2, respectively. As a result, group A1 contained three professors. One was an expert in e-learning, one in pedagogy, and one in education and didactics. Group A2 was comprised of 6 students who had previous experience in e-learning. Both groups of DMs are considered homogeneous and, therefore, no degrees of reliability (or importance) were determined.

2. Construct a fuzzy judgment matrix: To scale the relative importance of the criteria, a fuzzy judgment matrix should be constructed. More specifically, a comparison matrix is formed so that the criteria of the same level are pair-wise compared. Each evaluator is asked to express the relative importance of two criteria at the same level using linguistic terms, which are then transformed into triangular numbers (Table 3). As a result, each evaluator completes a matrix for comparing C1-C4, one for c11-c15, one for c21-c23, one for c31-c33, and finally one for c41-c43. This procedure is done for both professors and students.

Table 3: The linguistic variables and the corresponding triangular fuzzy numbers

Linguistic variables	Triangular fuzzy numbers
Equally important	(1,1,1)
Intermediate 2	(1,2,3)
Weakly important	(2,3,4)
Intermediate 4	(3,4,5)
Strongly more important	(4,5,6)
Intermediate 6	(5,6,7)
Very strongly more important	(6,7,8)
Intermediate 8	(7,8,9)
Absolutely more important	(9,9,9)

According to Buckley (1985), a fuzzy judgment matrix can be defined as: $\bar{R}^k = [\tilde{r}_{ij}^k]^k$, where \bar{R}^k is a fuzzy judgment matrix of evaluator k, \tilde{r}_{ij}^k the fuzzy assessments between criterion i and criterion j of evaluator k, $\tilde{r}_{ij}^k = (l_{ij}^k, m_{ij}^k, u_{ij}^k)$, n is the number of the $\tilde{r}_{ij}^k = (1,1,1)$, when $i = j$ and $\tilde{r}_{ij}^k = 1/\tilde{r}_{ij}^k$, $i, j = 1, 2, \dots, n$.

For example, the matrix for comparing C1-C4 has been completed by an expert in education and didactics as presented in Table 4.

Table 4: The \bar{R}^1 completed by the environmental education expert

	C1		C2		C3		C4		
1							/3	/2	
2	/3	/2			/3	/2	/4	/3	/2
3							/3	/2	
4									

Each DM completes all five matrices and the final values of each matrix are calculated taking into account the geometric mean of the corresponding values of each matrix's cell in the respective matrices. As a result, the final matrices are built. For example, Tables 5-9 present those tables for the professors. Quite similar are the respective tables for the students.

Table 5: Matrix for the pair-wise comparison of the four criteria of the first level.

	C1			C2			C3			C4		
C1	1	1	1	1.26	2.29	3.30	1.00	1.00	1.00	0.30	0.44	0.79
C2	0.30	0.44	0.79	1	1	1	0.33	0.50	1.00	0.22	0.13	0.44
C3	1.00	1.00	1.00	1.00	2.00	3.00	1	1	1	0.30	0.44	0.79
C4	1.26	2.29	3.30	2.29	7.42	4.64	1.26	2.29	3.30	1	1	1

Table 6: Matrix for the pair-wise comparison of the sub-criteria of C1

	c11			c12			c13			c14			c15		
c11	1	1	1	1	2	3	0.25	0.33	0.50	0.26	0.35	0.55	0.33	0.50	1
c12	0.33	0.50	1	1	1	1	0.20	0.25	0.33	0.33	0.50	1.00	0.33	0.50	1
c13	2.00	3	4	3	4	5	1	1	1	1.59	2.62	3.63	2	3	4
c14	1.82	2.88	3.91	1	2	3	0.28	0.38	0.63	1	1	1	1	1	1
c15	1	2	3	1	2	3	0.25	0.33	0.50	1	1	1	1	1	1

Table 7: Matrix for the pair-wise comparison of the sub-criteria of C2

	c21			c22			c23		
c21	1	1	1	1	2	3	0.25	0.33	0.50
c22	0.33	0.50	1	1	1	1	0.20	0.25	0.33
c23	2	3	4	3	4	5	1	1	1

Table 8: Matrix for the pair-wise comparison of the sub-criteria of C3

	c31			c32			c33		
c31	1	1	1	0	0.33	0.50	0.33	0.50	1
c32	2	3	4	1	1	1	1	1	1
c33	1	2	3	1	1	1	1	1	1

Table 9: Matrix for the pair-wise comparison of the sub-criteria of C4

	c41			c42			c43		
c41	1	1	1	2	3	4	1	2	3
c42	0.25	0.33	0.50	1	1	1	1	2	3
c43	0.33	0.50	1	0.33	0.50	1	1	1	1

3. Fuzzy weights \tilde{w}_i are calculated. The geometric mean of the fuzzy comparison value of the attribute i to each attribute can be found as

$$\tilde{r}_i = \left[\prod_{j=1}^n \tilde{p}_{ij} \right]^{\frac{1}{n}}, \text{ for all } i$$

then the fuzzy weight \tilde{w}_i of the i^{th} attribute indicated by a triangular fuzzy number is calculated as

$$\tilde{w}_i = \tilde{r}_i \times \left[\sum_{j=1}^n \tilde{r}_j \right]^{-1} = (w_i^l, w_i^m, w_i^u)$$

4. Undertake defuzzification. Finally, the fuzzy priority weights are converted into crisp values by using the center of area method as follows

$$w_i = \frac{\tilde{w}_i}{\sum_{j=1}^n \tilde{w}_j} = \frac{w_i^l + w_i^m + w_i^u}{\sum_{j=1}^n \tilde{w}_j}$$

Table 10: Weights of the criteria for professors and students in the department of environment

	Weight for professors	Weight for students
c11: Accessibility	0.118	0.170
c12: Response time	0.096	0.126
c13: Reliability	0.412	0.432
c14: Easy to use/simplicity.	0.197	0.220
c15: e-Learning Management	0.177	0.052
c21: Quality of Synchronous Communication student-instructor	0.241	0.241
c22: Quality of Synchronous collaboration between students	0.145	0.145
c23: Students' participation	0.613	0.613
c31: e-Learning Culture	0.226	0.221
c32: e-Learning Support	0.334	0.337
c33: e-Learning Infrastructure	0.440	0.442
c41: Effectiveness	0.528	0.613
c42: Acceptability	0.261	0.241
c43: Course design	0.211	0.145

As a result, for each criterion, the final weights are calculated for the professors and students. These weights of the criteria are presented in Table 10. This process revealed that for both professors and students, the most important criterion of the first level is 'Quality of Learning', followed by the 'Functionality of Learning' and 'Readiness'. Within the sub-criteria of 'Functionality

of the system', the 'Reliability' of the system was considered by far the most important criterion. Regarding 'Quality of communication', the weights of the criteria were the same for professors and students. The sub-criterion 'Students' participation' was considered far more important than the other two. As far as 'e-Learning Readiness', for both groups, the infrastructure was considered important and the support in the e-Learning process was followed. Finally, concerning the 'Quality of e-Learning', 'Effectiveness' is considered much more important than the other two criteria.

Steps 1-4 are not essential to be repeated during the application of ENVEL. Researchers may use the weights presented in Table 10. However, if other researchers feel that the nature of the study program they evaluate may influence the weights of importance of the criteria, then steps 1-4 have to be repeated to calculate new weights.

4 ENVEL: Evaluating e-learning aspects

Steps 5-8 have to be repeated every time ENVEL is implemented, while steps 9-11 may be optionally implemented if low values on criteria have occurred:

5. Form B1 and B2 committees of DMs. The members of the committees are the professors (B1) and the students (B2) participating in the survey. B1 should be a heterogeneous group of professors. This means that the group should contain professors with different perceptions of e-learning and different levels of skills in e-learning and computer usage. Ideally, they should cover different subjects of the study program being evaluated. Similarly, B2 should contain students who have different skills and perceptions of e-learning.

6. Determine the degree of reliability (or importance) of the DMs. Since the evaluation is a problem under group decision-making conditions, the reliability of the DMs should be determined. If the degrees of importance of DMs are equal, then the group of decision-makers is deemed a homogenous group (Chou et al. 2008). Otherwise, the group is deemed a heterogeneous group. If the groups are considered heterogeneous, then it is proposed that the DMs that have previous experience in e-learning can have slightly better reliability than the others as they have experience on the subject. The degrees of importance of DMs are I_t , where t is the DMs, $I_t \in [0,1]$ and $\sum_{t=1}^k I_t \in [0,1]$.

7. Calculate the values of criteria. The calculation of the values of the criteria is made by the heterogeneous groups of DMs. All the questions of the questionnaire used the five Likert scale for their answers, except for the one question of c31 used in step 6 of ENVEL and one of the two questions related to c42 (Would you like to continue e-learning after the end of the COVID-19 era?), in which the answers were three (yes, no, only for the lessons that e-learning seems appropriate). The answer to each question is the value of the criterion corresponding to that question.

The answers to the questionnaires are collected and we make the following estimations:

in the case of the criteria that have only one question assigned to them, the mean of all answers to each question.

in the case of criteria c31 and c42, the values of the criteria are acquired only by the one question that uses five Likert scale answers.

in the case of criterion c23, which involves students' participation and has two possible answers, we calculate the mean of each question and then take the mean of these two values.

The values of the sub-criteria are within the range [1,5]. Those values and the values of I_t are used for calculating the mean which is assigned a value of the criterion c1-c4 and can be further used to conclude about the e-learning application. As a result, if we suppose the k members of the group of DMs had previous experience in e-learning and m DMs hadn't had any experience in e-

$$\text{learning then } r_j = \frac{\sum_{i=1}^k I_i c_{ij}}{k} + \frac{\sum_{i=1}^m I_i c_{ij}}{m}.$$

8. Calculate the final value of each main criterion.

The aggregation of the weights and performance values is calculated as follows:

$$c_i = \sum_{j=1}^n w_j r_{ij}$$

This value is used for characterizing the application of e-learning. Following the study by Linjawi & Alfadda (2018), the scale was set as follows:

- Low score: if the final value ranged from 1 to <3.
- Acceptable/moderate score: if the final value ranged from 3 to <4.
- High score: if the final value ranged from 4 to 5.

9. Definition of the criteria for interviews. This step is implemented for defining the criteria that are characterized as a low score. For the criteria that are characterized as a low score, a set of interviews is performed to find out:

- how severe are the problems related to this criterion
- the exact nature of the problems that were encountered during e-learning implementation.

10. Definition of the group of DMs D1. As soon as the problems are identified a new group of decision-makers is formed. These decision-makers should have specialization in the study program that is evaluated and have experience in e-learning.

11. Analyze results and decide on improvements. The group of decision-makers D1 should decide on the improvements that have to be implemented to ameliorate the e-learning process in the department.

5 The Department of environment turns to e-learning

The Department of Environment runs three different study programs related to the Environment, Conservation, and Technologies of the Environment. These study programs involve several different courses such as physics, chemistry, ecology, protected area management, geographical information systems, databases, waste management, renewable energy sources, etc. The courses are implemented using theoretical

lectures, laboratories, and practice exercises, which in some cases take place on the field.

In Greek Universities, information and communication technologies were mainly used in classrooms or in the form of asynchronous e-learning using Learning Management Systems for uploading additional educational material. The department had been using blended learning for the last decade in some courses but its usage depended only on the professor responsible for each course. More specifically, professors had been using e-class, a Learning Management System for uploading assignments, notes, announcements, etc. However, synchronous e-learning was not allowed in Greek Universities, except for the Hellenic Open University. During the Coronavirus emergency, all Greek Universities were asked to reorganize the educational process and provide all courses remotely. Synchronous e-learning was suddenly not only accepted but was considered mandatory. This is especially challenging for a department of Environmental Science that implements theoretical lectures, laboratories, and practice on the field.

All professors, irrelevant whether they were in favor of e-learning or not, and if they had experience in e-learning or not, were asked to re-organize their courses and provide synchronous e-learning. More specifically, no more than 64% of the professors had previous experience in e-learning, synchronous or asynchronous.

Another obstacle to the implementation was the fact that not all professors were in favor of e-learning. 21% of the professors were not or were very little in favor of e-learning. Another 21% of the professors were moderately in favor of e-learning and only 57% supported techniques of distance education. Despite this fact, all professors successfully transformed their courses and the department managed to provide all courses by distance. After two months of e-Learning implementation, a questionnaire was developed and teachers and students were asked to answer it voluntarily. As a result, 14 professors of various subjects related to the Environment, Conservation, and Environmental Technology, and 98 students of these subjects participated in the study. All of them had been actively attending the e-learning courses.

6 Case Study: Application of ENVEL in the department of environment

Steps 1-4 are implemented once and could be used as-is by other researchers that apply ENVEL. However, steps 5-8 should be implemented in each evaluation experiment. In this section, we present the implementation of steps 5-8 of the ENVEL for the evaluation of e-learning in the Department of Environment at Ionian University.

Table 11: Results of the evaluation of professors and students in the department of environment

	Professor value	Student value	Weighted value for professors	Weighted value for students	
c11: Accessibility	4.456	3.596	0.526	0.611	C1 _{prof} =4.27 C1 _{st} =3.16
c12: Response time	4.484	3.267	0.430	0.412	
c13: Reliability	4.092	2.983	1.686	1.289	
c14: Easy to use/simplicity.	4.516	3.211	0.890	0.706	
c15: e-Learning Management	4.185	2.760	0.741	0.144	
c21: Quality of Synchronous Communication student-instructor	3.936	3.035	0.949	0.731	C2 _{prof} =3.63 C2 _{st} =3.08
c22: Quality of Synchronous collaboration between students	3.723	2.799	0.540	0.406	
c23: Students' participation	3.496	3.162	2.143	1.938	
c31: e-Learning Culture.	3.388	2.629	0.766	0.581	C1 _{prof} =4.22 C1 _{st} =2.88
c32: e-Learning Support	4.484	2.925	1.498	0.986	
c33: e-Learning Infrastructure.	4.452	2.985	1.959	1.319	
c41: Effectiveness	3.724	2.741	1.966	1.680	C1 _{prof} =4.02 C1 _{st} =2.81
c42: Acceptability	4.300	2.876	1.122	0.693	
c43: Course design	4.392	2.993	0.927	0.434	

5. Form a committee of DMs for professors and one committee of DMs for students. During the implementation of ENVEL, the questionnaires were given to the members of B1 (14 professors) and the members of B2(94 students). All members of groups B1 and B2 participated in the e-learning and voluntarily became members of the groups after ensuring that these groups were heterogeneous regarding their perceptions and skills in e-learning.

6. Determine the degree of reliability (or importance) of the DMs. The question of c31 (Did you have previous experience in e-learning?), which the students and the professors could only answer yes or no, can be used to determine the degree of importance of each DM. If we consider that all DMs that have experience have importance $\omega_{\text{exp}} = 1$ and those that don't have experience have importance $\omega_{\text{non-exp}} = 0.85$. Then, the degree of reliability is calculated as $I_k = \frac{\omega_k}{\sum_{i=1}^2 \omega_i}$. As a result, $I_{\text{exp}} = 0.54$ for DMs with experience in e-learning and $I_{\text{non-exp}} = 0.46$ for DMs for users with no experience in e-learning.

7. Calculate the values of the criteria. Taking into account the mean values of criteria and the reliability of the members of the group we calculate the values of all criteria. All these values are presented in Table 10.

8. Calculate the final value of each main criterion. Using the values of sub-criteria that were calculated in step 7 and the weights of the sub-criteria and applying the weighted sum we estimate the values of the criteria. From the analysis of the values of the criteria, one can easily conclude the different aspects of e-learning in a whole study program in Higher Education.

7 Discussion on the results of the case study and proposed improvements

ENVEL has been applied for evaluating e-learning in the Department of Environment at the Ionian University. ENVEL, similar to the frameworks of Mahdavi et al. (2008), Çelikbilek & Adıgüzel Tüylü (2019) and Jeong & Gonzalez-Gomez (2020), has two levels of criteria. Our proposed approach has 4 criteria in the first level and 14 criteria in the second level. ENVEL, unlike Mahdavi et al. (2008) and Alqahtani & Rajkhan 2020 that use AHP, uses Fuzzy AHP, which is considered more friendly to professors and students due to the linguistic terms that are used. However, the main difference between our framework and the frameworks presented in Table 1 (Mahdavi et al. 2008, Stecyk 2019, Çelikbilek & Adıgüzel Tüylü 2019, Alqahtani & Rajkhan 2020, Jeong & Gonzalez-Gomez 2020) is that ENVEL is used for the evaluation of a whole study program implemented by distance and not specific e-learning courses and/or systems. This is the main reason why it was important to develop a new framework instead of using one of the existing ones presented in Table 1. The criteria used in those frameworks do not correspond to the aspects of a whole study program.

The results of the evaluation revealed that the whole e-learning experience was rated as mediocre and, although it was considered satisfactory as a solution to an emergency, it needs improvements if it is to be implemented again. Furthermore, both instructors and students agreed that face-to-face education is more effective, especially for a subject like environmental science that needs laboratories and practical exercises in the field.

One of the main conclusions involves the culture of users in e-Learning. Indeed, the results show that, although 78% of the participants were not in favor of e-learning, 40% considered the current implementation of e-learning very effective and the other 32% characterized it as of medium effectiveness.

The evaluation experiment revealed significant deviations in the views of students and professors. Taking into account how successful the implementation of e-learning was, the values assigned to the sub-criteria were 4.092-4.516 by the professors and much lower by the students (2.760-3.596). These values have occurred by computing the weighted value of the responses. Regarding the factor of communication, again the values of the criteria assigned by the professors were much better than those assigned by students.

However, what seemed rather disappointing was the deviation in the values in the last two main criteria. For example, as far as the readiness for the e-Learning implementation was concerned the value of the criterion for the professors was 4.22 while the corresponding value for the students was 2.88. 83% of the professors and the students stated that they had no previous experience in e-learning. Furthermore, only 28% of the participants were in favor of e-Learning, before that semester. Taking into account the low experience of professors in e-learning and the low acceptability of e-learning in general, before the semester of the evaluation, the support and the infrastructure provided by the university were considered quite satisfactory.

A rather important criterion for evaluating the whole process is the 'Quality of e-Learning'. This criterion is mainly influenced by the sub-criterion 'c41-Effectiveness' and then 'c42: Acceptability' and 'c43: Course design'. The deviation of the values given by students and professors in this criterion is high. For example, the three sub-criteria c41, c42, c43 were rated 3.724-4.392 by professors and 2.741-2.993 by students.

Notably, most of the criteria were rated above 4 by the professor, and therefore, they were considered a high score, except for the criterion 'Learner Attractiveness' which was considered mediocre. However, the values of the criteria provided by the students were much lower and were considered a medium score for the functionality of the system and the learner attractiveness and a low score for the readiness of e-learning and its quality. This shows a significant deviation between the professors' and students' views and shows that there is much more to be done to improve aspects of e-learning before it can be fully and more effectively implemented.

Since criteria, c3 and c4 are rated with low scores, steps 9-11 of the ENVEL should be implemented.

9. Definition of the criteria for interviews. The criteria that have low scores are c3 and c4.

10. Definition of the group of DMs D1. Two human experts from the Department of Environment were selected to perform the interviews with random professors and students who participated in the experiment. More specifically, one expert in e-learning and one in environmental education performed the interviews. Both experts had served as head of the department in the past.

11. Analyze results and decide on improvements. After interviews with the DMs of D1 were performed, the main problems occurred because students were not familiar with the particular method and had problems adjusting to e-learning. Many students also encountered technical problems as they didn't have the equipment to connect from home and had to follow the e-learning courses from their mobile phones. The low value of quality of learning was an expected problem, as professors had designed the courses for on-site learning. The complaints mainly involved laboratories and courses that were supposed to be implemented in situ and affected the effectiveness of e-learning.

The group of decision-makers D1, after performing the interviews, decided on the improvements that have to be implemented to ameliorate the e-learning process and make teaching and learning more effective. The following improvements were proposed:

- Organizing seminars and webinars about e-learning. These seminars would help users get informed on the platforms used and the e-learning process, in general.
- Updating Opencourses. The main platform that was used for uploading the course material, encountered several problems, which could be addressed with a simple update.
- Re-organizing laboratories to be better implemented by distance or postpone them until on-site learning is applicable.
- Purchasing an e-learning platform. In this way, the courses could last longer and have more functionality.
- Giving suggestions to professors about their course design, to cover the e-learning needs in a better and more effective way.

8 Conclusions

In general, face-to-face education may be preferable and more easily carried out than distance education, even if it is synchronous. However, in cases like the corona virus emergency synchronous e-learning was the only solution. This paper presented a framework for e-learning evaluation called ENVEL. ENVEL runs the evaluation experiment under group decision-making conditions using heterogeneous groups of students and professors. Combining the views of heterogeneous groups of people may provide a broader view of the implementation of e-learning. For this purpose, the framework suggests that the sample of users participating in the experiment should involve both students and professors, experienced

and non-experienced users in e-learning, and people with different views on e-learning.

The framework applies multi-criteria decision-making in order to combine the different aspects of e-learning and collect summative results on the e-learning. These results show the effectiveness and success of the e-learning implementation of a whole semester in Higher Education. Furthermore, the particular framework presents an analysis of the criteria that are taken into account in the evaluation of e-learning as well as an estimation of their importance in the evaluation process. The analysis of the values of the criteria provides a useful tool to conclude the specific aspects that need to be addressed to improve e-learning centrally. Therefore, the framework could be used in departments in higher education to draw conclusions and schedule the required changes to improve e-learning application in the whole study program. The weights of the criteria also reveal the priority of the related improvements. For example, the corrections in aspects that are assigned to criteria with higher weight should be addressed in higher priority.

For the estimation of the weights of criteria, FAHP is used. The selection of FAHP over the other MCDM theories lies in the fact that it has a very well-defined process for calculating the weights of criteria in comparison to other theories such as TOPSIS (Hwang & Yoon 1981), VIKOR (Opricovic, 1998), MAUT (Vincke 1992), SAW (Hwang & Yoon 1981; Chou et al. 2008), etc. Moreover, this process instead of asking experts to assign a weight to each criterion, allows them to make pairwise comparisons, and, therefore, was better than other methods. As a result, this process results in capturing better expert reasoning. This is also an advantage of AHP. However, the quantification in numbers may be difficult for some people. Another advantage of FAHP is that it supports decision-makers to assign linguistic variables in the form of numeric values to express their judgment and can incorporate incomplete, unobtainable, and unquantifiable information into the decision model in a fuzzy environment (Ramanayaka et al. 2019, Nagpal et al. 2015, Sivarji et al. 2013, Chang et al. 2008, Sadeghi et al. 2012, Chen et al. 2015).

It is among our plans, to implement this framework for e-Learning evaluation in other similar departments of Higher Education and prove its effectiveness for e-Learning evaluation in Higher Education in general. Furthermore, the same experiment is going to be repeated in the next semester after taking corrective actions in the implementation of e-learning. A second evaluation experiment is going to show if the results of the evaluation are going to change and in what way.

References

- [1] Abdalla, I. (2007). Evaluating effectiveness of e-blackboard system using TAM framework: A structural analysis approach. *AACE Journal*, 15(3), pp. 279-287.
- [2] Abuhlfaia K., de Quincey E. (2019). Evaluating the Usability of an e-Learning Platform within Higher

- Education from a Student Perspective. *ICEEL 2019: Proceedings of the 2019 3rd International Conference on Education and E-Learning*, pp. 1–7 <https://doi.org/10.1145/3371647.3371661>
- [3] Aguti, B., Wills, G.B., Walters, R.J. (2014). An Evaluation of the Factors that Impact on the Effectiveness of Blended E-Learning within Universities. *International Conference on Information Society*, pp. 117-121. <https://doi.org/10.1109/i-society.2014.7009023>
- [4] Alavi, M., & Leidner, D. E. (2001). Technology-mediated learning: A call for greater depth and breadth of research. *Information Systems Research*, 12(1), pp. 1–10. <https://doi.org/10.1287/isre.12.1.1.9720>
- [5] Al-Fadhli, S. & Khalfan, A. (2009). Developing critical thinking in e-learning environment: Kuwait University as a case study. *Assessment & Evaluation in Higher Education*, 34(5), pp. 529-536, DOI: 10.1080/02602930802117032
- [6] Alqahtani A.Y., Rajkhan A.A. (2020) E-Learning Critical Success Factors during the COVID-19 Pandemic: A Comprehensive Analysis of E-Learning Managerial Perspectives. *Educational Sciences*, 10, 216. <https://doi.org/10.3390/educsci10090216>
- [7] Altun Turker Y., Baynal K., Turker T. (2019) The Evaluation of Learning Management Systems by Using Fuzzy AHP, Fuzzy TOPSIS and an Integrated Method: A Case Study, *Turkish Online Journal of Distance Education*, 20(2), pp. 195-218, DOI: 10.17718/tojde.557864
- [8] Anderson, T. (2001). The hidden curriculum in distance education. An update view, *Change* 33, pp. 28-35. <https://doi.org/10.1080/00091380109601824>
- [9] Angrainingsih, R., Umam, M. Z., Setiadi, H. (2018) Determining e-learning success factor in higher education based on user perspective using Fuzzy AHP. *MATEC Web of Conferences* 154, 03011, <https://doi.org/10.1051/mateconf/201815403011>
- [10] Anis A. & Islam R. (2015) The application of analytic hierarchy process in higher-learning institutions: a literature review. *Journal International Business and Entrepreneurship Development*, 8(2). <https://doi.org/10.1504/jibed.2015.070446>
- [11] Arbaugh, J. B. (2004). Learning to learn online: A study of perceptual changes between multiple online course experiences. *The Internet and Higher Education*, 7, pp. 169-182. <https://doi.org/10.1016/j.iheduc.2004.06.001>
- [12] Arbaugh, J. B. (2008). Does the community of inquiry framework predict outcomes in online MBA courses? *International Review of Research in Open and Distance Learning*, 9(2). <https://doi.org/10.19173/irrodl.v9i2.490>
- [13] Barteit S, Guzek D, Jahn A, Bärnighausen T, Jorge MM, Neuhann F. (2020) Evaluation of e-learning for medical education in low- and middle-income countries: A systematic review. *Computer Education*. 145, 103726. doi: 10.1016/j.compedu.2019.103726. PMID: 32565611; PMCID: PMC7291921.
- [14] Benbunan-Fich, R. & Hiltz, S. R. (2003) Mediators of the effectiveness of online courses, *IEEE Transactions on Professional Communication*, 46, pp. 298–312. <https://doi.org/10.1109/tpc.2003.819639>
- [15] Bhadauria R. (2016) E-Learning – A boon for Indian Higher Education System. *International Journal of Engineering Technology, Management and Applied Sciences*. 4(2), pp. 122-128, ISSN 2349-4476
- [16] Buckley, J.J. (1985). Fuzzy hierarchical analysis. *Fuzzy Sets and Systems*, 17(3), pp. 233-247. [https://doi.org/10.1016/0165-0114\(85\)90090-9](https://doi.org/10.1016/0165-0114(85)90090-9)
- [17] Çelikbilek Y. & Adıgüzel Tüylü A.N. (2019) Prioritizing the components of e-learning systems by using fuzzy DEMATEL and ANP, *Interactive Learning Environments*, DOI: 10.1080/10494820.2019.1655065
- [18] Chang, C. W., Wu, C. R., & Lin, H. L. (2008) Integrating fuzzy theory and hierarchy concepts to evaluate software quality, *Software Quality Journal*, 16(2), pp. 263–76. <https://doi.org/10.1007/s11219-007-9035-2>
- [19] Chen, J.-F., Hsieh, H.-N., & Do, Q. H. (2015) Evaluating teaching performance based on fuzzy AHP and comprehensive evaluation approach, *Applied Soft Computing*, 28, pp. 100–108. <https://doi.org/10.1016/j.asoc.2014.11.050>
- [20] Chou, S.Y., Chang, Y.H., Shen, C.Y. (2008) A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes. *European Journal of Operational Research*, 189, pp. 132–145. <https://doi.org/10.1016/j.ejor.2007.05.006>
- [21] Chou, S.-W., & Liu, C.-H. (2005) Learning effectiveness in a web-based virtual learning environment: A learner control perspective, *Journal of Computer Assisted Learning*, 21, pp. 65-76 <https://doi.org/10.1111/j.1365-2729.2005.00114.x>
- [22] Clark, L. J. (2001). Web-based teaching: A new educational paradigm, *Intercom*, 48, pp. 20-23.
- [23] Elkaseh A.M., Wong K.W. and Fung C.C.(2016) Perceived ease of use and perceived usefulness of social media for e-learning in Libyan higher education: a structural equation modeling analysis. *International Journal of Information and Education Technology*, 6(3), pp. 192-199. <https://doi.org/10.7763/ijiet.2016.v6.683>
- [24] Erensal, Y. C., Öncan, T., & Demircan, M. L. (2006). Determining key capabilities in technology management using fuzzy analytic hierarchy process: A case study of Turkey, *Information Sciences*, 176(18), pp. 2755–70. <https://doi.org/10.1016/j.ins.2005.11.004>
- [25] Frehywot, S., Vovides, Y., Talib, Z. et al. (2013) E-learning in medical education in resource constrained low- and middle-income

- countries. *Human Resources Health* 11(4). <https://doi.org/10.1186/1478-4491-11-4>
- [26] Gupta, S., & Bostrom, R.P. (2009). Technology-Mediated Learning: A Comprehensive Theoretical Model. *Journal of the Association for Information Systems*, 10(9), pp. 686-714. <https://doi.org/10.17705/1jais.00207>
- [27] Harasim, L., Hiltz, S. R., Teles, L., & Turoff, M. (1995) *Learning networks: A field guide to teaching and learning online*. The MIT Press, Cambridge, MA.
- [28] Harrati, N., Bouchrika, I., Tari, A., and Ladjailia, A. (2016). Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. *Computer Human Behavior*, 61, pp. 463--471. DOI= <https://doi.org/10.1016/j.chb.2016.03.051>.
- [29] Hartman, K., Neuwirth, C. M., Kiesler, S., Sproull, L., Cochran, C., Palmquist, M., & Zubrow, D. (1995) Patterns of social interaction and learning to write: Some effects of network technologies. In: Z. L. Berge and M. P. Collins, Editors. *Computer mediated communication and the online classroom*, Vol. 2. Hampton Press, Inc, Cresskill, NJ, pp 4778. <https://doi.org/10.1177/0741088391008001005>
- [30] Haverila, M., & Barkhi, R. (2009). The Influence of Experience, Ability and Interest on eLearning Effectiveness. *European Journal of Open, Distance and E-Learning*, https://www.eurodl.org/materials/contrib/2009/Haverila_Barkhi.pdf
- [31] Hrastinski, S. (2008) Asynchronous and synchronous e-learning. *Educause Quarterly*, 31(4), pp. 51-55.
- [32] Hwang, C.L., Yoon, K. (1981) Multiple Attribute Decision Making: Methods and Applications. *Lecture Notes in Economics and Mathematical Systems*, 186.
- [33] Islam, A. Md., Rahim, N.A.N., Liang, T.C., & Momtaz, H. (2011) Effect of Demographic Factors on E-Learning Effectiveness in A Higher Learning Institution in Malaysia. *International Education Studies*. 4(1), pp. 112-121. <https://doi.org/10.5539/ies.v4n1p112>
- [34] Islas, E., Perez, M., Rodriguez, G., Paredes, I., Avila, I., & Mendoza, M. (2007), E-learning tools evaluation and roadmap development for an electrical utility, *Journal of Theoretical and Applied Electronic Commerce Research*, 2(1), pp. 63-75. <https://doi.org/10.3390/jtaer2010006>
- [35] Jasim, M. H., Maznah M. K., Nizal M. S. I. (2018) Evaluation of E-Learning Approaches Using AHPTOPSIS Technique. *Journal of Telecommunication, Electronic and Computer Engineering*, 10 (1-10). pp. 7-10. ISSN 2180-1843.
- [36] Jeong J. & González-Gómez D. (2020). Assessment of sustainability science education criteria in online-learning through fuzzy-operational and multi-decision analysis and professional survey, *Heliyon*, Volume 6, Issue 8, August 2020, e04706, <https://doi.org/10.1016/j.heliyon.2020.e04706>
- [37] Liaw, S.S. (2008). Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system. *Computers & Education*, 51, pp.864–873 <https://doi.org/10.1016/j.compedu.2007.09.005>
- [38] Liaw, S. S., Huang, H. M., & Chen, G. D. (2007). Surveying instructor and learner attitudes toward e-learning, *Computers & Education*. <https://doi.org/10.1016/j.compedu.2006.01.001>
- [39] Lin, H.-F. (2010) An application of fuzzy AHP for evaluating course website quality. *Computers & Education*, 54, 877–888 <https://doi.org/10.1016/j.compedu.2009.09.017>
- [40] Linjawi, AI, Alfadda, LS. (2018) Students' perception, attitudes, and readiness toward online learning in dental education in Saudi Arabia: a cohort study. *Advances in Medical Education and Practice*, 9, pp. 855-863. <https://doi.org/10.2147/AMEP.S175395>
- [41] Mahdavi, I., Fazlollahtabar, H., Heidarzade, A., Mahdavi-Amiri N. and Rooshan, Y.I. (2008). A Heuristic Methodology for Multi-Criteria Evaluation of Web-Based E-Learning Systems Based on User Satisfaction. *Journal of Applied Sciences*, 8, pp. 4603-4609. DOI: 10.3923/jas.2008.4603.4609
- [42] Marks, R. B., Sibley, S. D., & Arbaugh, J. B. (2005). A structural equation model of predictor for effective online learning, *Journal of Management Education*, 29, pp. 531-563. <https://doi.org/10.1177/1052562904271199>
- [43] Martinez-Caro, E. (2011). Factors Affecting Effectiveness in E-Learning: An Analysis in Production Management Courses. *Computer Applications in Engineering Education*, 19(3), pp. 572 – 581. DOI: 10.1002/cae.20337
- [44] Martín-Lara, M.Á., Rico, N. (2020) Education for Sustainable Energy: Comparison of Different Types of E-Learning Activities. *Energies* 2020, 13, 4022. <https://doi.org/10.3390/en13154022>
- [45] Mercader, C., Gairín, J. (2020) University teachers' perception of barriers to the use of digital technologies: The importance of the academic discipline. *International Journal of Educational Technology in Higher Education*, 17, 4. <https://doi.org/10.1186/s41239-020-0182-x>
- [46] Mohamadali, N.A., & Garibaldi, J. (2011). Comparing user acceptance factors between research software and medical software using AHP and Fuzzy AHP. *The 11th Workshop on Computational Intelligence*, 7 - 9 September 2011, Kilburn Building.
- [47] Mulubrhan, F., Akmar, A., Mokhtar, A. & Muhammad, M. (2014). Comparative Analysis between Fuzzy and Traditional Analytical Hierarchy Process. *MATEC Web of Conferences* 13. <https://doi.org/10.1051/mateconf/20141301006>
- [48] Nagpal, R., Mehrotra, D. Bhatia, P. K., & Sharma, A. (2015) FAHP approach to rank educational websites on usability, *International Journal of*

- Computing and Digital Systems*, 4(4), pp. 251–260. <https://doi.org/10.12785/ijcds/040404>
- [49] Naveed, Q.N., Qureshi, M.R.N., Tairan, N., Mohammad, A., Shaikh, A., Alsayed, A.O., et al. (2020) Evaluating critical success factors in implementing E-learning system using multi-criteria decisionmaking. *PLoS ONE* 15(5): e0231465. <https://doi.org/10.1371/journal.pone.0231465>
- [50] Njenga, J.K. (2017) E-Learning in Higher Education. *Encyclopedia of International Higher Education Systems and Institutions*; Springer: Maine, The Netherlands, pp. 1–6. https://doi.org/10.1007/978-94-017-9553-1_323-1
- [51] Noesgaard, S. S., & Ørngreen, R. (2015) The Effectiveness of E-Learning: An Explorative and Integrative Review of the Definitions, Methodologies and Factors that Promote e-Learning Effectiveness. *The Electronic Journal of eLearning*, 13(4), 278-290.
- [52] Opricovic, S. (1998). *Multicriteria optimization of civil engineering systems*. Belgrade: Faculty of Civil Engineering
- [53] Otto, D., Becker, S. (2018) E-Learning and Sustainable Development. *Encyclopedia of Sustainability in Higher Education*; Springer: Cham, Switzerland, p. 8. https://doi.org/10.1007/978-3-319-63951-2_211-1
- [54] Owston, R. (1997) The World Wide Web: A technology to enhance teaching and learning? *Educational Researcher*, 26(2). <https://doi.org/10.2307/1176036>
- [55] Ramanayaka, K. H., Chen, X., & Shi, B. (2019). UNSCALE: A Fuzzy-based Multi-criteria Usability Evaluation Framework for Measuring and Evaluating Library Websites, *IETE Technical Review*, 36(4), 412 - 431, 10.1080/02564602.2018.1498032
- [56] Reiser, R. (2001). A History of Instructional Design & Technology: Part 1: A History of Instructional Media. *Educational Research & Development*, 49 (1), pp. 53-64. <https://doi.org/10.1007/bf02504506>
- [57] Richmond, V. P., McCroskey, J. C., Kearney, P. & Plax, T. G. (1987) Power in classroom VII: Linking behaviour alteration techniques to cognitive learning, *Communication Education*, 36, 1-12. <https://doi.org/10.1080/03634528709378636>
- [58] Saaty, T. (1980). *The analytic hierarchy process*. New York, McGraw-Hill.
- [59] Sadeghi, M., aliRashidzadeh, M., & aliSoukhakian, M. (2012). Using analytic network process in a group decision making for supplier selection, *Informatica*, 23(4), pp. 621–43. <https://doi.org/10.15388/informatica.2012.378>
- [60] Salter, S.M., Karia, A., Sanfilippo, F.M., Clifford, R.M. (2014). Effectiveness of E-learning in Pharmacy Education. *American Journal of Pharmaceutical Education*, 78(4), Article 83.
- [61] Schieffer, L. (2016) The Benefits and Barriers of Virtual Collaboration Among Online Adjuncts. *Journal of Instructional Research*, 5, 109–125. <https://doi.org/10.9743/jir.2016.11>
- [62] Sheikh, T. H., Gupta R. (2018). E-Learning for Higher Education. A Case Study. *Journal of Advanced Research in Dynamical & Control Systems*, 10(4), ISSN 1943-023X
- [63] Siemens, G., Gasevic D. & Dawson S. (2015). *Preparing for the digital university: A review of the history and current state of distance, blended and online learning*. Athabasca: Athabasca University. DOI: 10.13140/RG.2.1.3515.8483
- [64] Sivaji, A., Abdullah, M. R., Downe, A. G. & Ahmad, W. F. W. (2013) Hybrid Usability Methodology: Integrating Heuristic Evaluation with Laboratory Testing across the Software Development Lifecycle, *10th International Conference on Information Technology: New Generations*, pp. 375–83. <https://doi.org/10.1109/itng.2013.60>
- [65] Stecyk A. (2019) Application of PROMETHEE II method for e-learning course selection, *e-mentor*, 1(78), 39-45. DOI: 10.15219/em80.1398
- [66] Stöhr C., Demaziere C., Adawi T. (2020) The polarizing effect of the online flipped classroom. *Computers & Education*, 147, 103789 <https://doi.org/10.1016/j.compedu.2019.103789>
- [67] Su, B., Bonk, C. J., Magjuka, R.J., Liu, X. & Lee, S. (2005) The importance on Interaction in web-based education: A program level case study of online MBA courses, *Journal of Interactive Online Learning*, 4, 119.
- [68] Sulčić, V., & Lesjak, D. (2009). E-Learning and Study Effectiveness, *Journal of Computer Information Systems*, 49(3), 40-47.
- [69] Tai D.W.S., Chen J.L., Wang R. (2011) Research on Fuzzy Linguistic Evaluation in e-Learning Using AHP and TOPSIS Based on Web Resources. In: Lin S., Huang X. (eds) *Advances in Computer Science, Environment, Ecoinformatics, and Education. CSEE 2011. Communications in Computer and Information Science*, 217. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23339-5_69
- [70] Tiwari, N. (2006). Using the Analytic Hierarchy Process (AHP) to identify Performance Scenarios for Enterprise Application, *Computer Measurement Group, Measure It*, 4, 3.
- [71] Tseng, M.-L., & Lin, R.-J., Chen, H.-P. (2011). Evaluating the effectiveness of e-learning system in uncertainty. *Industrial Management & Data Systems*, 111(6), pp. 869-889. <https://doi.org/10.1108/02635571111144955>
- [72] Vincke, P. (1992) *Multi criteria Decision-Aid*. Wiley, Chichester. <https://doi.org/10.1002/mcda.4020030208>
- [73] I. Vinogradova, R. Kliukas (2015) Methodology for Evaluating the Quality of Distance Learning Courses in Consecutive Stages, *Procedia - Social and Behavioral Sciences*, 191, pp. 1583-1589, <https://doi.org/10.1016/j.sbspro.2015.04.364>.

- [74] Wu, W. & Hwang, L.-Y. (2010). The effectiveness of e-learning for blended courses in colleges: A multi-level empirical study. *International Journal of Electronic Business Management*, 8(4), pp. 312-322.
- [75] Zopounidis, C. (2009). Knowledge-based multi-criteria decision support. *European journal of operational research*, 195, pp. 827-828. <https://doi.org/10.1016/j.ejor.2007.11.026>

A Novel Framework Based on Integration of Simulation Modelling and MCDM Methods for Solving FMS Scheduling Problems

¹*Shafi Ahmad, ¹Zahid A. Khan, ²Mohammed Ali, ¹Mohammad Asjad

¹Department of Mechanical Engineering, Jamia Millia Islamia (A Central University), New Delhi, India.

²Department of Mechanical Engineering, Aligarh Muslim University, Aligarh, India.

E-mail: shafiahmad.amu@gmail.com, zahid_jmi@yahoo.com, mohdali234@rediffmail.com, masjid@jmi.ac.in

*Corresponding author

Keywords: job priority rules, simulation modelling, MCDM methods, flexible manufacturing system, scheduling performance measures

Received: March 26, 2021

Scheduling in Flexible Manufacturing Systems (FMSs) is an important area of research as it significantly affects performance of the systems. In scheduling problems, determination of an appropriate order for jobs to be processed on a machine is a difficult task and to solve such problems, job priority rules (JPRs) are used. Several JPRs have been developed with an aim to obtain better performance, measured in terms of one or more scheduling performance measures (SPMs). However, selection of an appropriate rule is still an area of research as no single rule provides better results for all SPMs considered simultaneously. This work proposes a framework which is based on an integration of simulation and multi criteria decision making (MCDM) methods for the selection of an appropriate JPR yielding optimum results for multiple SPMs taken together. The proposed framework includes development of a simulation model to collect values of the SPMs corresponding to different JPRs. Further, five MCDM methods have been used to determine rank of the JPRs. Since different MCDM methods produce different ranking result therefore, the final rank of the JPRs has been determined by comparing the rank derived from these methods using membership degree method. To exemplify the probable application of the proposed framework, it has been implemented on a specific FMS taken from the literature in order to select the best JPR.

Povzetek: Razvit je okvir za reševanje problemov razporejanja v fleksibilnih proizvodnih sistemih, ki združuje simulacijsko modeliranje in metode MCDM.

1 Introduction

High system throughput and customer satisfaction are considered as the most important performance metrics required from a manufacturing system. However, due to conflicting nature of these performance metrics, the concept of flexible manufacturing system (FMS) was evolved which provides flexibility as well as high productivity at the same time. FMS is a centrally or distributed computer control system consisting of automated machines viz. CNC, material handling system (MHS), automatic storage and retrieval system (AS/RS) and other auxiliary devices. Various studies have suggested that a significant amount of improvement in performance can be obtained with installation of FMS over conventional manufacturing systems [1], [2]. Further, the performance of these systems can be additionally enhanced if effective and efficient operational decisions are made [3]. Despite the fact that small setup times, variability of parts, high machine utilization etc. are some of the advantages, there are various operational problems associated with FMSs. The three major operational problems of FMS are classified as scheduling, control problems, and pre-release planning [4].

The focus of this study is to scrutinize scheduling problem associated with FMS. Scheduling is an extensively researched area and it is considered to be an important concern in the management and planning of manufacturing processes [3], [5]. It is the process of assigning available resources to the concerned job so as to enhance productivity, flexibility, profitability, and production of the system [6]. Scheduling in FMS environment is more complex as compared to conventional manufacturing environment due to its versatile capabilities [7]. The FMS scheduling problem consists of two sub-problems [8]. The first one is related to the allocation of the requisite operation to a suitable machine and the second one pertains to the job sequencing of operations on each machine. Over the years, Job Priority Rules (JPRs) are found to be the simplest and most widely accepted means to resolve the second sub problem i.e., sequencing of jobs on each machine. These rules provide precedence to one job over other jobs, based on their performance on a predefined priority function, for processing on the machine. Further, the scheduling problem is described as dynamic or static based on the availability of jobs [9]. A scheduling problem is classified as dynamic if jobs arrive into the system during the scheduling process i.e. all jobs are not available at the

beginning and it is categorized as static if they are available at the starting of the scheduling process [10]. To solve static scheduling problem, many optimization algorithms and heuristics have been developed [11], [12]. However, JPRs are found to be the most appropriate means to resolve the dynamic scheduling problems [13], [14].

JPRs are classified on the basis of processing time, due date, rules neither based on processing time and due date, combinatory rules and rules based on shop floor conditions [9]. The processing time based rules are found to perform well under tight load conditions whereas for light load conditions, due date based rules are preferred [15]. Several JPRs have been developed and proposed in the literature and it has been established that the performance of the FMS is significantly affected by the chosen JPR. Further, selection of an appropriate JPR among the available one is a complex task as no single rule can provide best results for all the performance measures taken together. With this intention, this work attempts to provide an effective framework using a combined approach of simulation modelling and MCDM methods to select the best JPR resulting in the optimum performance of FMSs with dynamic scheduling of parts.

Selection of an appropriate JPR requires various performance measures to be satisfied simultaneously and therefore, the selection problem resembles an MCDM problem. Several MCDM methods such as WSM, WPM, AHP, VIKOR, ELECTRE, TOPSIS etc. are available which can be used to select the best JPR among the available ones [16]. However, due to inherent characteristics of these and many other MCDM methods, the best JPR produced by them may be different. Therefore, the framework proposed in this work examines the priorities of the JPRs on the basis of rank obtained from five MCDM methods viz. Measurement Alternatives and Ranking according to COmpromise Solution (MARCOS), Proximity Index Value (PIV), Multi-Attribute Border Approximation Area Comparison (MABAC), Evaluation based on Distance from Average Solution (EDAS) and Technique of Order Preference Similarity to the Ideal Solution (TOPSIS). Subsequently, it determines the final rank of the JPRs by comparing the rank produced by these methods using membership degree method. To demonstrate the potential application of the proposed framework, it has been employed to select the best JPR for a specific FMS taken from the literature. Rest of the paper is structured as follows: Section 2 discusses the various JPRs, SPMs, simulation modelling, and the MCDM methods employed in the present study. Section 3 describes the steps involved in the development of the proposed framework. Section 4 explains working of the proposed framework through an illustrative example taken from the literature. Finally, section 5 presents conclusion of the present study.

2 JPRs, SPMs, simulation modelling and MCDM methods

2.1. Job priority rules (JPRs) in FMSs

Job priority rules are used to select the next job to be processed on a machine from a set of jobs that are waiting in the queue for processing. Since, these rules are simple and easy to implement, they are most commonly used in FMSs for job sequencing. Consider an FMS with m machines designated as M_i ($i=1, 2, m$) processing n parts say P_j ($j = 1, 2, \dots, n$). If t = Current time of the system, AT_j = Time of arrival of part j in the system, T_{ij} = Time of arrival of part j on machine i , PT_{ij} = Processing time of part j on machine i , DD_j = Due date of part j , TT_j = Total time required to perform all operations on part j , RT_j = Remaining processing time for part j and NR_j = Number of remaining operations to be performed on part j . Some of the most commonly used JPRs along with their priority functions and reference are shown in Table 1.

Table 1: JPRs and their priority functions

JPRs	Symbol	Priority Function	Reference
First Come, First Served	FCFS	$\min (T_{ij})$	[14], [17]
Last Come, First Served	LCFS	$\max (T_{ij})$	[14]
Shortest Processing Time	SPT	$\min (PT_{ij})$	[14], [17]
Longest Processing Time	LPT	$\max (PT_{ij})$	[14], [18]
Earliest Due Date	EDD	$\min (DD_j)$	[18], [19]
First at shop, first out	FASFO	$\min (AT_j)$	[20]
Least Slack Time	LST	$\min (DD_j - t - RT_j)$	[19], [20]
Minimum Critical Ratio	MCR	$\min ((DD_j - t)/RT_j)$	[19]
Maximum Balanced processing time	MBPT	$\max (RT_j)$	[14], [21]
Least Balanced processing time	LBPT	$\min (RT_j)$	[14], [19], [20]
Most Number of Operations Remaining	MNOR	$\max (NR_j)$	[20]
Least Number of Operations Remaining	LNOR	$\min (NR_j)$	[19], [20]

Greatest Total Work	GTW	$max(TT_j)$	[14]
Lowest Total Work	LTW	$min(TT_j)$	[14], [17], [19]
Modified due date	MDD	$min(DD_j - t)$	[19], [22]
Least processing time on Next machine	LPTNM	$min(PT_{(i+1)j})$	[23], [24]
Maximum processing time on Next machine	MPTNM	$max(PT_{(i+1)j})$	[23], [24]
Processing time and due date total	PDT	$min(PT_{ij} + DD_j)$	[25]

2.2. Scheduling performance measures (SPMs)

In a manufacturing system, scheduling performance measures (SPMs) are the attributes used to estimate the performance of a schedule. There are a number of different SPMs which can be used to evaluate the performance of a schedule. However, their consideration may vary depending upon the requirements of a specific industry. The frequently found SPMs are described as follows:

1. Makespan time (MT): It is the amount of time required to complete a set of jobs. It is desirable to schedule the parts in such a way that MT is the minimum. Considering t_o as the time at which first part enters into the system and t_f as the time at which last part exits from the system, MT is defined by Eqn. (1).

$$MT = t_f - t_o \tag{1}$$

2. Average waiting time in the queue (AW): AW is the average waiting time spent by parts on a machine to get processed. Considering a system with m machines designated as MC_i ($i=1,2, \dots, m$) processing n parts labelled as P_j ($j=1,2, \dots, n$). If W_{ji} denotes the waiting time of part j on machine i , the Average waiting time in the queue (AW) on machine i is computed according to Eqn. (2).

$$AW_i = \sum_{j=1}^n W_{ji} \tag{2}$$

3. Machine utilization (MU): It refers to the extent to which the productive capacity of a machine is utilized. Mathematically, it is the ratio of the time a machine is working to the total time it is available for processing as given by Eqn. (3).

$$MU_i = \frac{\text{Time the machine 'i' is working}}{\text{Total time the machine 'i' is available}} \times 100 \tag{3}$$

4. Average lateness (AL): It is the difference between the completion time and the due date of a part. Since, each part has different completion time and may have different due date, the average of the lateness of all the parts is used to measure performance of the system. If DD_j and C_j denote the due date and completion time of a part j respectively, then AL is given by Eqn. (4)

$$AL = \sum_{j=1}^n (DD_j - C_j) \tag{4}$$

5. Number of late parts (NL): Any part completed after its due date is regarded as late. An appropriate schedule is the one which does not result in any late part. Hence, total number of late parts is considered as a SPM and it should be minimized.

2.3. Simulation modelling

A simulation model is a replicate of a real process or system in a virtual space. It has gained importance in the past few years due to its exceptional ability to quantify and observe behavior of complex systems under different scenarios. It helps to examine how an existing system or process might perform if some or all the parameters are altered. In manufacturing environment, it is extensively used to study and compare performance of the system under different designs. Inventory management, scheduling, investigation of different control strategies, are some of the most common issues addressed by simulation modeling.

Simulation modeling finds a wide range application and consequently, several studies based on it have been conducted by researchers for examining and improving performance of the FMSs. For example, Chawla et al., (2018) performed a simulation based investigation to determine optimal utilization of the AGVs [26]. Amoako-Gyampah & Meredith, (1996) conducted a simulation based study and suggested different heuristics for tool allocations in FMS [27]. Mahmood et al., (2017) examined the performance of FMS with the help of modeling and simulation [28]. Hussain & Ali, (2019) examined the impact of control and design factors on different SPMs specifically AW, MU and MT with the help of simulation modeling [29]. A comprehensive review on FMS modeling reported that simulation modeling has been used by different authors to solve problems associated with FMSs [2]. Further, software-based simulation modeling has gained popularity over other methods due to its simplicity. Few examples of prominent software preferred for modeling of FMSs are WITNESS, ARENA, ProModel etc.

2.4. MCDM methods

MCDM methods are techniques that are used to solve decision making problems involving several conflicting attributes/criteria. Over the years, a large number of MCDM methods have been developed and employed to solve decision making problems pertaining to different

knowledge domain. Among the several MCDM methods, TOPSIS is the most widely used [30] and MARCOS [31], PIV [32], MABAC [33] and EDAS [34] are recently developed methods. Therefore, these MCDM methods have been included in the proposed framework. The computational steps of these methods are discussed in the following subsections.

Measurement Alternatives and Ranking according to COmpromise Solution (MARCOS) method

MARCOS is a recently developed MCDM method which can be used to rank alternatives [31]. In this method, rank of the alternatives is determined on the basis of their utility function value which is a connection between reference values and alternatives [35], [36]. The computational steps of this method are discussed as follows [31], [35]:

Step 1: Formulate decision matrix $D = [a_{ij}]_{m \times n}$, where the element a_{ij} represents the value of j th decision attributes for i th alternative and total number of decision attributes and alternatives varies from 1 to m and 1 to n respectively.

Step 2: Insert non-ideal $NI = [a_{Nj}]_{1 \times n}$ and ideal $PI = [a_{Pj}]_{1 \times n}$ alternative at the top and bottom of D to develop extended decision matrix (E). The values a_{Pj} and a_{Nj} for beneficial and non beneficial attributes are computed using Eqn. (5) and Eqn. (6) respectively.

$$a_{Pj} = \max_i (a_{ij}), \quad a_{Nj} = \min_i (a_{ij}) \quad (5)$$

$$a_{Pj} = \min_i (a_{ij}), \quad a_{Nj} = \max_i (a_{ij}) \quad (6)$$

Step 3: Develop normalized decision matrix $N = [s_{ij}]_{(m+2) \times n}$ where, $s_{ij} = a_{ij}/a_{Pj}$ for beneficial attribute and $s_{ij} = a_{Pj}/a_{ij}$ for non beneficial attribute.

Step 4: Determine weighted normalized matrix $W = [v_{ij}]_{(m+2) \times n}$. If w_j is the weight assigned to attribute j , then $v_{ij} = s_{ij} \times w_j$.

Step 5: Compute positive utility degree $PU_i = \sum_{j=1}^n v_{ij} / \sum_{j=1}^n v_{Pj}$ and negative utility degree $NU_i = \sum_{j=1}^n v_{ij} / \sum_{j=1}^n v_{Nj}$ of the alternatives.

Step 6: Determine the utility function (U_i) value of the alternatives using Eqn. (7).

$$U_i = \frac{PU_i + NU_i}{1 + \frac{1 - f(PU_i)}{f(PU_i)} + \frac{1 - f(NU_i)}{f(NU_i)}} \quad (7)$$

where, $f(PU_i) = \frac{PU_i}{PU_i + NU_i}$ and $f(NU_i) = \frac{NU_i}{PU_i + NU_i}$.

Step 7: Rank the alternatives on the basis of their U_i value. Higher the U_i value higher is the rank and vice-versa.

Proximity Index Value (PIV) method

PIV method was developed in 2018 to prioritize different alternatives [32]. This method is popular among researchers due to its advantage of minimizing the rank reversal problem in situations when either more alternatives are added or a few are removed from the existing list of alternatives, as compared to other MCDM

methods specifically, TOPSIS [37], [38]. This method comprises of the following steps [32], [38]:

Step 1: Formulate decision matrix as discussed in step 1 of MARCOS method.

Step 2: Develop normalized decision matrix $N = [s_{ij}]_{m \times n}$

where, $s_{ij} = \frac{a_{ij}}{\sqrt{\sum_i a_{ij}}}$.

Step 3: Determine weighted normalized matrix $W = [v_{ij}]_{m \times n}$. If w_j is the weight assigned to attribute j , then $v_{ij} = s_{ij} \times w_j$.

Step 4: Compute proximity value ($PV_i = \sum_{j=1}^n u_i$) of each alternative, where u_i values are determined using Eqn. (8).

$$u_i = \left\{ \begin{array}{l} \max_i (v_{ij}) - v_{ij}; \quad \text{if } j \in \text{ben} \\ v_{ij} - \min_i (v_{ij}); \quad \text{if } j \in \text{non} \end{array} \right\} \quad (8)$$

Step 5: Rank the alternatives on the basis of their u_i value. A lower u_i value corresponds to higher rank and vice-versa.

Multi-Attribute Border Approximation area Comparison (MABAC) method

MABAC method, developed in 2015, has been effectively used to solve problems pertaining to different knowledge domain [33], [39], [40]. In this method, rank to the alternatives is assigned on the basis of their distance from the border approximation area (BAO). An alternative having highest distance from the BAO is ranked first and rank of other alternatives decreases as their distance from BAO decreases. The computational steps of this method are as under [33], [41]:

Step 1: Formulate decision matrix as discussed in step 1 of MARCOS method.

Step 2: Develop normalized decision matrix $N = [s_{ij}]_{m \times n}$ where, s_{ij} is computed using Eqn. (9).

$$s_{ij} = \left\{ \begin{array}{l} \frac{a_{ij} - \min_i (a_{ij})}{\max_i (a_{ij}) - \min_i (a_{ij})}; \quad \text{if } j \in \text{bene} \\ \frac{\max_i (a_{ij}) - a_{ij}}{\max_i (a_{ij}) - \min_i (a_{ij})}; \quad \text{if } j \in \text{non } b \end{array} \right\} \quad (9)$$

Step 3: Determine weighted normalized matrix $W = [v_{ij}]_{m \times n}$. If w_j is the weight assigned to attribute j , then $v_{ij} = (s_{ij} + 1) \times w_j$.

Step 4: Determine the border approximation area matrix $G = [g_j]_{1 \times n}$ where, $g_j = (\prod_{i=1}^m s_{ij})^{1/m}$

Step 5: Compute total distance of each alternative from the border approximation area $S_i = \sum_{j=1}^n q_{ij}$ where, $q_{ij} = v_{ij} - g_j$

Step 6: Rank the alternatives based on their S_i value. An alternative with the maximum S_i value gets rank 1 and rank decreases as S_i value decreases.

Evaluation based on Distance from Average Solution (EDAS)

EDAS, developed in 2015, is a compensatory method in which the distance of an alternative from the optimal value is used to identify the best alternative [34]. This method has been used to solve air traffic problem [42], personnel selection problem [43], and evaluation of airlines services [44]. The computational steps involved in this method are as follows [34], [44]:

Step 1: Formulate decision matrix as discussed in step 1 of MARCOS method.

Step 2: Compute average solution (AV) corresponding to each attribute $AV = [r_j]1 \times n$ where, $r_j = \frac{1}{n} (\sum_{i=1}^m a_{ij})$

Step 3: Determine positive (PD) and negative (ND) distances from the average solution as defined by Eqn. (10) and Eqn. (11) respectively.

$$PD_{ij} = \left\{ \begin{array}{ll} \frac{\max(0, a_{ij} - r_j)}{r_j} & \text{if } j \in \text{ben} \\ \frac{\max(0, r_j - a_{ij})}{r_j} & \text{if } j \in \text{non ben} \end{array} \right\} \quad (10)$$

$$ND_{ij} = \left\{ \begin{array}{ll} \frac{\max(0, r_j - a_{ij})}{r_j} & \text{if } j \in \text{benef} \\ \frac{\max(0, a_{ij} - r_j)}{r_j} & \text{if } j \in \text{non ben} \end{array} \right\} \quad (11)$$

Step 4: Compute weighted positive distance $WPD_i = \sum_{j=1}^n (PD_{ij} \times w_j)$ and $WND_i = \sum_{j=1}^n (ND_{ij} \times w_j)$ of each alternative. Where, w_j is the weight assigned to attribute j .

Step 5: Determine appraisal score (A) using Eqn. (12).

$$A_i = \frac{PN_i + NN_i}{2} \quad (12)$$

where, $PN_i = \frac{WPD_i}{\max(WPD_i)}$ and $NN_i = \frac{WND_i}{\max(WND_i)}$.

Step 6: Rank the alternatives on the basis of their A_i value. Rank 1 is given to the alternative having maximum A_i value and the rank decreases with decreasing A_i value.

Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)

TOPSIS is the most widely used MCDM method for solving varieties of decision problems belonging to different knowledge domain [45]. This method prioritizes the alternatives on the basis of their distance from positive and negative ideal alternatives. It suggests that an alternative closest to positive ideal alternative and farthest from negative ideal alternative should be ranked first [46]. The steps involved in finding the distances from positive and negative ideal alternatives and thus organizing them as per their performances are as follows [45]–[47]:

Step 1: Formulate decision matrix as discussed in step 1 of MARCOS method.

Step 2: Develop normalized decision matrix $N = [s_{ij}]m \times n$ where, $s_{ij} = \frac{a_{ij}}{\sqrt{\sum_i a_{ij}}}$.

Step 3: Determine weighted normalized matrix $W = [v_{ij}]m \times n$. If w_j is the weight assigned to attribute j , $v_{ij} = s_{ij} \times w_j$.

Step 4: Discover the positive $PI = [a_{Pj}]1 \times n$ and negative $NI = [a_{Nj}]1 \times n$ ideal alternatives. Elements a_{Pj} and a_{Nj} are determined using Eqn. (13) and Eqn. (14) respectively

$$a_{Pj} = \left\{ \begin{array}{ll} \max_i (a_{ij}) & \text{if } j \in \text{benef} \\ \min_i (a_{ij}) & \text{if } j \in \text{non - bene} \end{array} \right\} \quad (13)$$

$$a_{Nj} = \left\{ \begin{array}{ll} \min_i (a_{ij}) & \text{if } j \in \text{benef} \\ \max_i (a_{ij}) & \text{if } j \in \text{non - benef} \end{array} \right\} \quad (14)$$

Step 5: Compute positive distance $PD_i = \sqrt{\sum_{j=1}^n (a_{Pj} - a_{ij})^2}$ and negative distance $ND_i = \sqrt{\sum_{j=1}^n (a_{Nj} - a_{ij})^2}$ of alternatives from the ideal alternative.

Step 6: Compute relative closeness of alternative, $RC_i = \frac{ND_i}{PD_i + ND_i}$.

Step 7: Rank the alternatives based on their RC_i value. A higher RC_i value corresponds to higher rank and vice-versa..

2.5. Method to determine final rank of the alternatives

It has been observed that the computational steps involved in different MCDM methods are different. Hence, it is much likely that rank of the alternatives produced by different MCDM methods may be different. Therefore, membership degree method developed by Yang et al., (2019) is used to determine final rank of the alternatives. Steps of this method are given below (Yang et al., 2019; Wakeel et al., 2020):

Step 1: Arrange alternatives in rows and their rank in columns resulting in the formulation of rank frequency matrix (R) as given by Eqn. (15). Each element c_{pq} of R denotes the frequency of rank q for alternative p over n different MCDM methods.

$$R = \begin{bmatrix} c_{11} & \dots & c_{1q} \\ \dots & \dots & \dots \\ c_{p1} & \dots & c_{pq} \end{bmatrix} \quad (15)$$

where, p and q vary from 1 to m and m denotes total number of alternatives.

Step 2: Formulate membership degree matrix (MD) by dividing each element of R by total number of MCDM methods i.e. n as shown in Eqn. (16)

$$MD = \begin{bmatrix} l_{11} & \dots & l_{1q} \\ \dots & \dots & \dots \\ l_{p1} & \dots & l_{pq} \end{bmatrix} \quad (16)$$

where, $l_{pq} = \frac{r_{pq}}{n}$

Step 3: Determine rank index (RI_p) using Eqn. (17).

$$RI_p = \sum_{q=1}^N (l_{pq}) \times q \quad (17)$$

Step 4: The rank index is used to determine final rank of the alternatives. An alternative with the minimum RI_p value is ranked first and lower rank is assigned to the alternatives with higher RI_p values.

3 Proposed research framework

The jobs received at a workstation undergo a wide range of operations. When job traffic is high, the sequence of job processing becomes very important due to high cost of waiting jobs and the cost of idle workstations. Inefficient scheduling results in the formation of job queues. This situation puts pressure on the managers to develop schedules which handle the job traffic effectively and efficiently. A number of JPRs have been laid for prioritizing the jobs at different workstations. As soon as a workstation finishes a task, the job priority rule decides the succeeding job that enters the workstation. Determining which JPR best suits a particular system is a difficult task as none of the suitable choices give any clear indication on managing a system in the best possible way. Thus, management should select the best option using a systematic approach. The approach must consider multiple contributing factors in order to get a deeper insight into the system and make better decisions. Multi Criteria Decision Making (MCDM) methods can accomplish the above said requirements. MCDM techniques can assist managers and decision-makers in making informed decisions by solving problems involving multiple criteria. This work proposes a simulation-based decision-making framework to select best JPR using MCDM methods. The step-by-step procedure for the proposed research framework is discussed as follows:

1. Identify the potential JPRs used in FMS and SPMs used to examine the performance of the concerned industry.
2. Develop the simulation model of the FMS using appropriate software such as ARENA, WITNESS etc. It is to be noted that while developing the simulation model the modules and corresponding attributes should be provided to collect SPM.
3. After developing the simulation model, collect the SPM values corresponding to each JPR.

4. Arrange JPRs and SPM in rows and columns respectively to formulate decision matrix to be used by the MCDM methods i.e. MARCOS, PIV, MABAC, EDAS and TOPSIS methods to rank the potential JPRs.
 5. Determine the final rank of JPRs by comparing the ranking results of different MCDM methods using membership degree method.
- The proposed research framework comprising of the above steps is shown in Figure 1.

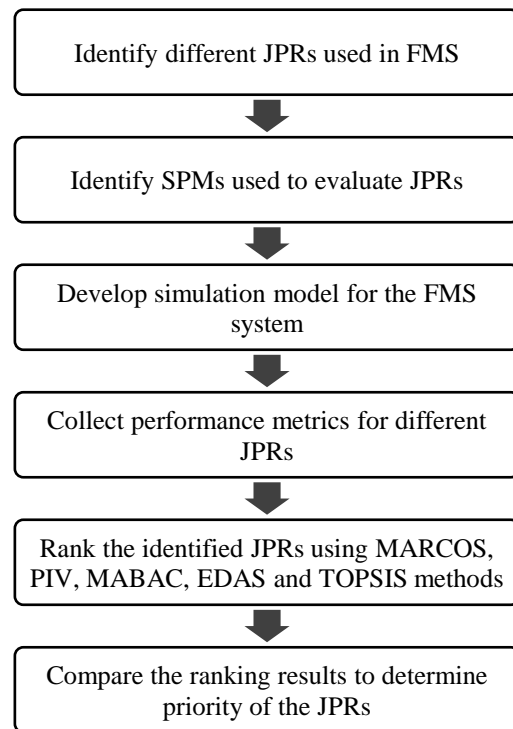


Figure 1: Proposed research framework.

4 Illustrative example

To demonstrate the potential application of the proposed research framework, it has been employed to select the best JPR for an FMS taken from literature [49]. The FMS consisted of five CNC machines named as FMC1, FMC2, FMC3, FMC4, and FMC5 for processing five different part types A2021, B2021, C2021, D2021, and E2021. All five machines had infinite input buffer capacity and three to four operations were essentially required for complete processing of a specific job type. The processing time of various parts was stochastic.

Since inter-arrival time and due date of jobs were not known for the FMS, they were determined using Eqn. (18) [50] and Eqn. (19) [10], [50] respectively.

$$v = \frac{1}{\lambda} = \frac{\mu_p \mu_g}{\eta \omega} \quad (18)$$

$$DD_i = A_i + K \times TT_j \quad (19)$$

where, v and λ are mean inter-arrival and job arrival time respectively, μ_p and μ_g are the mean processing time and number of operations per job respectively, η and ω are

the utilization and number of machines in the shop floor respectively. A_i is the arrival time of job i , TT_j is the total time required to perform of all operations on part j and K is the tightness factor which reflects the amount of expected delay a job will experience and it is taken as 3 in this study.

Further, simulation model for the FMS configuration shown in Table 2 was developed using the student's version of ARENA 16.0 simulation software. While developing the model, following assumptions were made:

- (i) Part transfer, loading and unloading times were all included in the processing time,
- (ii) No rework was allowed,
- (iii) No order cancellation was allowed,
- (iv) Machines never broke down, and
- (v) Pre-emption was not allowed.

Since, the maximum number of parts that can be created in students' version of ARENA is limited to 150, the number of parts for part types A2021, B2021, C2021, D2021 and E2021 was considered as 45, 27, 27, 23 and 16 respectively. Initially, Create module was used to create 5 different parts with inter arrival time computed using Eqn. (18). Further, each part progressed to the Assign module where parameters such as arrival time, processing time, due date, number of operations, and sequence of operations were assigned. In accordance with the sequence of operations, each part was moved to its corresponding machine station. Before slotting a part in the machine for operation, it was passed through Assign Attribute module where parameters such as remaining processing time, remaining number of operations etc. were modified. Subsequently, the parts were processed according to the predefined JPR. The parts were further moved forward to the next corresponding station after passing through another Assign Attribute module. As soon as all the operations were completed on a part, it was moved to the Dispose station where it was checked whether the part was late or not. If the part was late, lateness was stored using the Record module. Finally, if the part under examination was the last part of the system, the current simulation time was collected to measure the make span time. The logic module for the simulation model so developed is shown in Figure A1 of Appendix.

For each of the twenty JPRs defined in section 2.1.1, the simulation model was run to collect the performance values for the five SPM. It needs to be mentioned here that 30 replications of each simulation run were performed and the results of the five SPM for each of the JPRs were collected which are shown in Table A1 of Appendix.

It is observed that MT is minimum (2794.59 minutes) when jobs are processed according to the LBPT rule. Whereas, number of late jobs and average lateness are minimum (75.033 and 664.17 minutes) when SPT and EDD rule are used respectively. Further, mean of AW for all machines is minimum (192.30 minutes) for SPT rule and average MU is maximum (63.73 %) for LTW rule as evident from Table 2.

Table 2: Mean AW and MU

JPR	Mean AW (min)	Mean MU (%)
-----	---------------	-------------

FCFS	256.78	63.47
LCFS	273.09	63.18
SPT	192.30	62.75
LPT	303.41	62.28
EDD	204.37	63.55
FASFO	203.95	63.30
LST	226.11	63.48
MCR	241.53	63.21
MBPT	272.26	63.53
LBPT	228.62	63.70
MNOR	307.75	63.62
LNOR	270.21	63.25
GTW	268.88	63.00
LTW	210.37	63.73
MDD	216.02	63.35
LPTNM	380.75	60.03
MPTNM	250.10	63.02

Thus, based on the results, it is observed that no single JPR provides optimum results for all the SPMs. Hence, MCDM methods were used to find the best JPR that produced the optimal results for all SPMs. Considering Table A1 as the decision matrix, five MCDM methods were employed to determine rank of the JPRs. Equal weight was assigned to the SPMs as they were considered to be equally important. The performance value and corresponding rank of the JPRs derived from different MCDM methods used in this study is shown in Table A2 of the Appendix.

It is found that LTW is ranked first by three MCDM methods i.e. MARCOS, MABAC, and EDAS. However, it ranked second and sixth by PIV and TOPSIS methods. Whereas, PDT and MDD is ranked first by PIV and TOPSIS methods respectively. Thus, it is difficult to suggest which among the considered JPR is should be used as none among them is ranked first by all the methods. Therefore, membership degree method was employed to determine the rank index and final rank of the JPRs which are shown in Table 3.

Table 3: Rank index and final ranks of the JPRs

JPR	Rank Index, RI_p	Final Rank
FCFS	9.6	9
LCFS	11.6	11
SPT	6.8	8
LPT	15.2	16
EDD	3.8	3
FASFO	5.8	6
LST	6.4	7
MCR	13.2	13
MBPT	16	17
LBPT	4.2	4
MNOR	14	14
LNOR	9.6	10
GTW	15	15
LTW	2.2	1
MDD	4.6	5
LPTNM	17.6	18
MPTNM	12.8	12

It is evident from Table 3 that among the considered JPRs, LTW is the top ranked rule and therefore, it is the best one for producing optimum performance of the five SPMs. Further, the next preferred rule is PDT followed by EDD, LBPT, MDD, FASFO, LST, SPT, FCFS, LNOR, LCFS, MPTNM, MCR, MNOR, GTW, LPT, MBPT and LPTNM. Hamidi (2016) proposed the PDT rule to utilize benefits of both SPT and EDD rules and showed that this is an effective and promising rule as compared to FCFS, SPT, EDD, MCR and LST rules [25]. Similar results have been obtained in this study where SPT is found better than other rules except LTW.

5 Conclusion

Flexible manufacturing systems (FMSs) are preferred over conventional manufacturing systems due to their ability to provide flexibility as well as high throughput at the same time. However, there are various operational problems associated with FMSs which need to be resolved to enhance productivity of these systems. Scheduling is one of the operational problems which have attracted attention of the researchers as it significantly affects performance of the FMS. This work intended to provide an effective decision-making framework to resolve one of the scheduling problems i.e. sequencing of jobs on each machine. Job priority rules (JPRs) are used to determine sequence of the jobs to be processed on a machine. These rules provide precedence to a job over other jobs based on a predefined priority function. Several JPRs have been proposed so far to obtain better results for one or two performance measures. However, it is difficult to judge which rule is the best one to produce optimal values for all the performance measures considered simultaneously. The current era of production systems requires better results for more than one performance measures taken together. Hence, selection of an appropriate JPR becomes more difficult as more than one performance measures need to be justified simultaneously. MCDM methods are one of the most powerful decision-making methods used to select the best alternative from among the existing ones on the basis of multiple attributes. Hence, the decision-making framework proposed in this work was based on selection of JPR using MCDM methods combined with simulation modeling. In the proposed framework, initially the potential JPRs and scheduling performance measures (SPM) for the concerned industry were identified. Further, a simulation model of the FMS was developed to collect the performance value for the various SPMs for different JPRs. The SPMs values corresponding to JPRs acted as a decision matrix for MCDM methods. Five MCDM methods were employed to rank the JPRs which produced their different ranks and therefore, it was difficult to decide which JPR is the best one. Consequently, the ranks obtained from different MCDM methods were compared to determine the final rank of the JPRs using membership degree method. The proposed framework was implemented to select the best JPR for a specific FMS taken from the literature. It has been found that for the considered FMS system, LTW rule provides optimum results for the five SPMs.

The proposed framework can be used by the FMS based industries to solve problem related to selection of the best JPR so as to obtain optimum values of their specific SPMs. It needs to be mentioned here that industries may not necessarily employ the same MCDM methods that have been used in the present study to determine rank of JPRs, instead other MCDM methods can also be used. However, steps of the proposed framework listed in the paper are essentially required to be followed for selection of the best JPR leading to the optimal performance of the system.

References

- [1] R. El-Khalil and Z. Darwish, "Flexible manufacturing systems performance in U.S. automotive manufacturing plants: a case study," *Production Planning & Control*, vol. 30, no. 1, pp. 48–59, Jan. 2019, doi: 10.1080/09537287.2018.1520318.
- [2] A. Yadav and S. C. Jayswal, "Modelling of flexible manufacturing system: a review," *International Journal of Production Research*, vol. 56, no. 7, pp. 2464–2487, Apr. 2018, doi: 10.1080/00207543.2017.1387302.
- [3] D.-K. Lee, J.-H. Shin, and D.-H. Lee, "Operations scheduling for an advanced flexible manufacturing system with multi-fixturing pallets," *Computers & Industrial Engineering*, vol. 144, p. 106496, Jun. 2020, doi: 10.1016/j.cie.2020.106496.
- [4] H. Tempelmeier and H. Kuhn, *Flexible manufacturing systems: decision support for design and operation*, vol. 12. John Wiley & Sons, 1993.
- [5] K. E. Stecke, "Design, planning, scheduling, and control problems of flexible manufacturing systems," *Ann Oper Res*, vol. 3, no. 1, pp. 1–12, Jan. 1985, doi: 10.1007/BF02023765.
- [6] A. Prakash, F. T. S. Chan, and S. G. Deshmukh, "FMS scheduling with knowledge based genetic algorithm approach," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3161–3171, Apr. 2011, doi: 10.1016/j.eswa.2010.09.002.
- [7] C. Low, Y. Yip, and T.-H. Wu, "Modelling and heuristics of FMS scheduling with multiple objectives," *Computers & Operations Research*, vol. 33, no. 3, pp. 674–694, Mar. 2006, doi: 10.1016/j.cor.2004.07.013.
- [8] P. Sharma and A. Jain, "Effect of routing flexibility and sequencing rules on performance of stochastic flexible job shop manufacturing system with setup times: Simulation approach," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 231, no. 2, pp. 329–345, Jan. 2017, doi: 10.1177/0954405415576060.
- [9] M. S. Jayamohan and C. Rajendran, "New dispatching rules for shop scheduling: A step forward," *International Journal of Production Research*, vol. 38, no. 3, pp. 563–586, Feb. 2000, doi: 10.1080/002075400189301.

- [10] O. Holthaus and C. Rajendran, “New dispatching rules for scheduling in a job shop — An experimental study,” *Int J Adv Manuf Technol*, vol. 13, no. 2, pp. 148–153, Feb. 1997, doi: 10.1007/BF01225761.
- [11] C. Blum and M. Sampels, “An Ant Colony Optimization Algorithm for Shop Scheduling Problems,” *Journal of Mathematical Modelling and Algorithms*, vol. 3, no. 3, pp. 285–308, Sep. 2004, doi: 10.1023/B:JMMA.0000038614.39977.6f.
- [12] M. K. Marichelvam, M. Geetha, and Ö. Tosun, “An improved particle swarm optimization algorithm to solve hybrid flowshop scheduling problems with the effect of human factors – A case study,” *Computers & Operations Research*, vol. 114, p. 104812, Feb. 2020, doi: 10.1016/j.cor.2019.104812.
- [13] F. T. S. Chan, H. K. Chan, H. C. W. Lau, and R. W. L. Ip, “Analysis of dynamic dispatching rules for a flexible manufacturing system,” *Journal of Materials Processing Technology*, vol. 138, no. 1, pp. 325–331, Jul. 2003, doi: 10.1016/S0924-0136(03)00093-1.
- [14] P. D. D. Dominic, S. Kaliyamoorthy, and M. S. Kumar, “Efficient dispatching rules for dynamic job shop scheduling,” *Int J Adv Manuf Technol*, vol. 24, no. 1, pp. 70–75, Jul. 2004, doi: 10.1007/s00170-002-1534-5.
- [15] M. Krishnan, T. R. Chinnusamy, and T. Karthikeyan, “Performance Study of Flexible Manufacturing System Scheduling Using Dispatching Rules in Dynamic Environment,” *Procedia Engineering*, vol. 38, pp. 2793–2798, Jan. 2012, doi: 10.1016/j.proeng.2012.06.327.
- [16] G.-H. Tzeng and J.-J. Huang, *Multiple Attribute Decision Making: Methods and Applications*. CRC Press, 2011.
- [17] M. Thürer, M. Stevenson, and T. Qu, “Job sequencing and selection within workload control order release: an assessment by simulation,” *International Journal of Production Research*, vol. 54, no. 4, pp. 1061–1075, Feb. 2016, doi: 10.1080/00207543.2015.1047978.
- [18] M. Z. Baharom, W. Nazdah, and W. Hussin, “Scheduling Analysis for Job Sequencing in Veneer Lamination Line,” *Journal of Industrial and Intelligent Information*, vol. 3, no. 3, 2015.
- [19] H.-H. Doh, J.-M. Yu, J.-S. Kim, D.-H. Lee, and S.-H. Nam, “A priority scheduling approach for flexible job shops with multiple process plans,” *International Journal of Production Research*, vol. 51, no. 12, pp. 3748–3764, Jun. 2013, doi: 10.1080/00207543.2013.765074.
- [20] M. Montazeri and L. N. Van Wassenhove, “Analysis of scheduling rules for an FMS,” *The International Journal of Production Research*, vol. 28, no. 4, pp. 785–802, 1990.
- [21] S. Wadhwa, Y. Ducq, M. Ali, and A. Prakash, “Performance analysis of a flexible manufacturing system,” *Global Journal of Flexible Systems Management*, vol. 10, no. 3, pp. 23–34, 2009.
- [22] V. Vinod and R. Sridharan, “Dynamic job-shop scheduling with sequence-dependent setup times: simulation modeling and analysis,” *Int J Adv Manuf Technol*, vol. 36, no. 3, pp. 355–372, Mar. 2008, doi: 10.1007/s00170-006-0836-4.
- [23] R. Haupt, “A survey of priority rule-based scheduling,” *Operations-Research-Spektrum*, vol. 11, no. 1, pp. 3–16, 1989.
- [24] S. S. Panwalkar and W. Iskander, “A Survey of Scheduling Rules,” *Operations Research*, vol. 25, no. 1, pp. 45–61, Feb. 1977, doi: 10.1287/opre.25.1.45.
- [25] M. Hamidi, “Two new sequencing rules for the non-preemptive single machine scheduling problem,” *The Journal of Business Inquiry*, vol. 15, no. 2, pp. 116–127, 2016.
- [26] V. K. Chawla, A. K. Chanda, S. Angra, and S. Rani, “Simultaneous Dispatching and Scheduling of Multi-Load AGVs in FMS-A Simulation Study,” *Materials Today: Proceedings*, vol. 5, no. 11, Part 3, pp. 25358–25367, Jan. 2018, doi: 10.1016/j.matpr.2018.10.339.
- [27] K. Amoako-Gyampah and J. R. Meredith, “A simulation study of FMS tool allocation procedures,” *Journal of Manufacturing Systems*, vol. 15, no. 6, pp. 419–431, Jan. 1996, doi: 10.1016/S0278-6125(97)83055-5.
- [28] K. Mahmood, T. Karaulova, T. Otto, and E. Shevtshenko, “Performance Analysis of a Flexible Manufacturing System (FMS),” *Procedia CIRP*, vol. 63, pp. 424–429, Jan. 2017, doi: 10.1016/j.procir.2017.03.123.
- [29] Mohd. S. Hussain and M. Ali, “A Multi-agent Based Dynamic Scheduling of Flexible Manufacturing Systems,” *Glob J Flex Syst Manag*, vol. 20, no. 3, pp. 267–290, Sep. 2019, doi: 10.1007/s40171-019-00214-9.
- [30] S. K. Yadav, D. Joseph, and N. Jigeesh, “A review on industrial applications of TOPSIS approach,” *International Journal of Services and Operations Management*, vol. 30, no. 1, pp. 23–28, Jan. 2018, doi: 10.1504/IJSOM.2018.091438.
- [31] Ž. Stević, D. Pamučar, A. Puška, and P. Chatterjee, “Sustainable supplier selection in healthcare industries using a new MCDM method: Measurement of alternatives and ranking according to COMPROMISE solution (MARCOS),” *Computers & Industrial Engineering*, vol. 140, p. 106231, Feb. 2020, doi: 10.1016/j.cie.2019.106231.
- [32] S. Mufazzal and S. M. Muzakkir, “A new multi-criterion decision making (MCDM) method based on proximity indexed value for minimizing rank reversals,” *Computers & Industrial Engineering*, vol. 119, pp. 427–438, May 2018, doi: 10.1016/j.cie.2018.03.045.
- [33] D. Pamučar and G. Čirović, “The selection of transport and handling resources in logistics centers using Multi-Attributive Border Approximation area Comparison (MABAC),” *Expert Systems with Applications*, vol. 42, no. 6, pp. 3016–3028, Apr. 2015, doi: 10.1016/j.eswa.2014.11.057.

- [34] M. Keshavarz Ghorabae, E. K. Zavadskas, L. Olfat, and Z. Turskis, “Multi-criteria inventory classification using a new method of evaluation based on distance from average solution (EDAS),” *Informatica*, vol. 26, no. 3, pp. 435–451, 2015.
- [35] Ž. Stević and N. Brković, “A Novel Integrated FUCOM-MARCOS Model for Evaluation of Human Resources in a Transport Company,” *Logistics*, vol. 4, no. 1, Art. no. 1, Mar. 2020, doi: 10.3390/logistics4010004.
- [36] A. Ulutaş, D. Karabasevic, G. Popovic, D. Stanujkic, P. T. Nguyen, and Ç. Karaköy, “Development of a Novel Integrated CCSD-ITARA-MARCOS Decision-Making Approach for Stackers Selection in a Logistics System,” *Mathematics*, vol. 8, no. 10, Art. no. 10, Oct. 2020, doi: 10.3390/math8101672.
- [37] N. Z. Khan, T. S. A. Ansari, A. N. Siddiquee, and Z. A. Khan, “Selection of E-learning websites using a novel Proximity Indexed Value (PIV) MCDM method,” *J. Comput. Educ.*, vol. 6, no. 2, pp. 241–256, Jun. 2019, doi: 10.1007/s40692-019-00135-7.
- [38] S. Wakeel, S. Bingol, M. N. Bashir, and S. Ahmad, “Selection of sustainable material for the manufacturing of complex automotive products using a new hybrid Goal Programming Model for Best Worst Method–Proximity Indexed Value method,” *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, p. 1464420720966347, Oct. 2020, doi: 10.1177/1464420720966347.
- [39] D. Pamučar, Ž. Stević, and E. K. Zavadskas, “Integration of interval rough AHP and interval rough MABAC methods for evaluating university web pages,” *Applied Soft Computing*, vol. 67, pp. 141–163, 2018.
- [40] D. I. Božanić, D. S. Pamučar, and S. M. Karović, “Application the MABAC method in support of decision-making on the use of force in a defensive operation,” *Tehnika*, vol. 71, no. 1, pp. 129–136, 2016.
- [41] D. Bozanic, D. Tešić, and J. Kočić, “Multi-criteria FUCOM – Fuzzy MABAC model for the selection of location for construction of single-span bailey bridge,” *Decision Making: Applications in Management and Engineering*, vol. 2, no. 1, Art. no. 1, Mar. 2019.
- [42] M. Keshavarz Ghorabae, M. Amiri, E. K. Zavadskas, Z. Turskis, and J. Antucheviciene, “Stochastic EDAS method for multi-criteria decision-making with normally distributed data,” *Journal of Intelligent & Fuzzy Systems*, vol. 33, no. 3, pp. 1627–1638, Jan. 2017, doi: 10.3233/JIFS-17184.
- [43] M. K. Kikomba, R. M. Mabela, and D. I. Ntantu, “Applying EDAS method to solve air traffic problems,” *International Journal of Scientific and Innovative Mathematical Research (IJSIMR)*, vol. 4, no. 8, pp. 15–23, 2016.
- [44] D. Stanujkic, G. Popovic, and M. Brzakovic, “An approach to personnel selection in the IT industry based on the EDAS method,” *Transformations in Business & Economics*, vol. 17, no. 2, p. 44, 2018.
- [45] M. Behzadian, S. Khanmohammadi Otaghsara, M. Yazdani, and J. Ignatius, “A state-of-the-art survey of TOPSIS applications,” *Expert Systems with Applications*, vol. 39, no. 17, pp. 13051–13069, Dec. 2012, doi: 10.1016/j.eswa.2012.05.056.
- [46] C.-L. Hwang and K. Yoon, *Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey*. Berlin Heidelberg: Springer-Verlag, 1981.
- [47] Y.-J. Lai, T.-Y. Liu, and C.-L. Hwang, “TOPSIS for MODM,” *European Journal of Operational Research*, vol. 76, no. 3, pp. 486–500, Aug. 1994, doi: 10.1016/0377-2217(94)90282-8.
- [48] W.-C. Yang, S.-H. Chon, C.-M. Choe, and K. Un-Ha, “Materials Selection Method Combined with Different MADM Methods,” *Journal of Artificial Intelligence*, vol. 1, no. 2, p. 89, 2019.
- [49] G. A. Chang and W. R. Peterson, “Modeling and analysis of flexible manufacturing systems: a simulation study,” in *Proceedings of the 2015 ASEE Annual Conference & Exposition, Seattle, WA, USA, 2015*, pp. 14–17.
- [50] R. Rangaritratsamee, W. G. Ferrell Jr, and M. B. Kurz, “Dynamic rescheduling that simultaneously considers efficiency and stability,” *Computers & Industrial Engineering*, vol. 46, no. 1, pp. 1–15, 2004.

Appendix

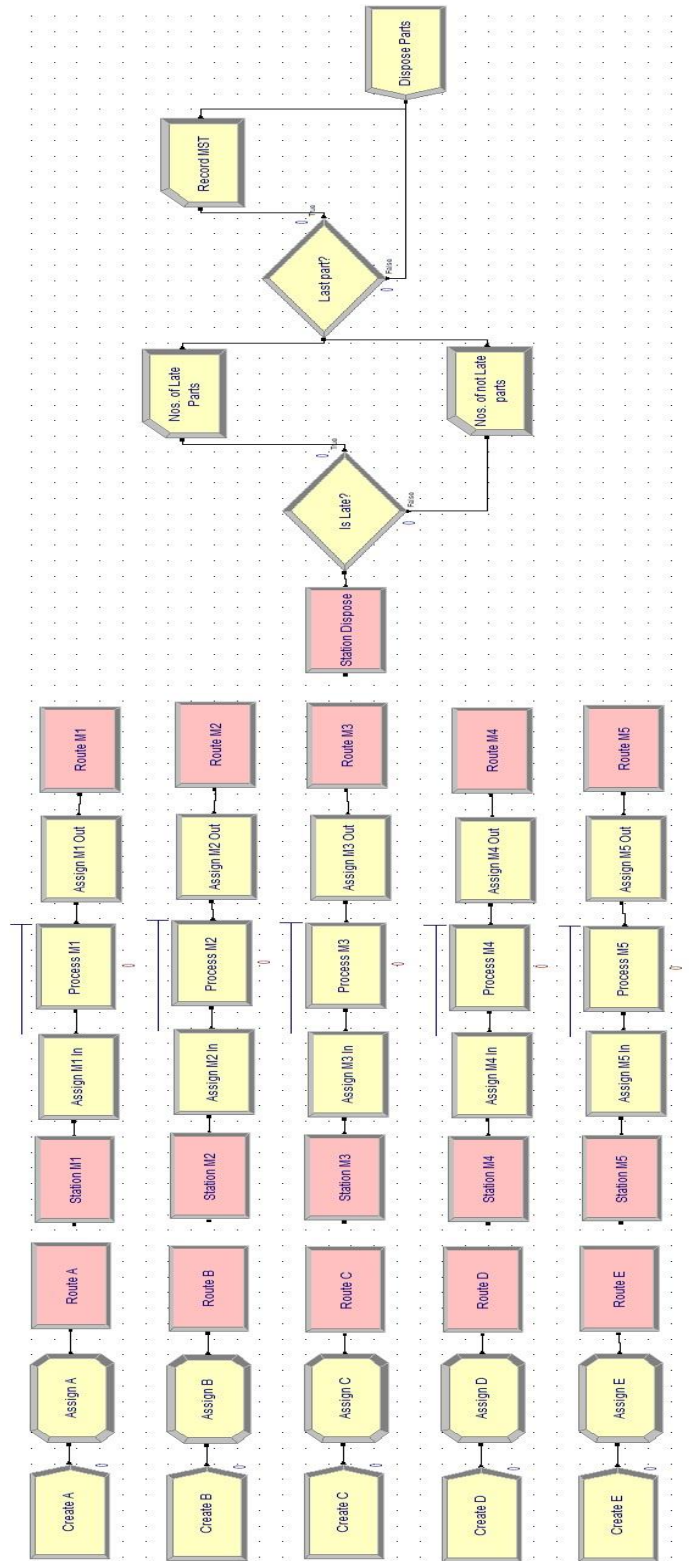


Figure A1: Logic module

Table A1 : Simulation results for SPMs

JPR	MT (min)	Machine Utilization, MU (%)					Average Waiting Time in Queue, AW (min)					Average Lateness (min)	Number of Late Jobs (Nos.)
		FMC5	FMC4	FMC3	FMC2	FMC1	FMC5	FMC4	FMC3	FMC2	FMC1		
FCFS	2837.45	78.62	86.88	99.83	24.6	27.42	199.35	809.98	261.32	8.3953	4.8601	944.94	121.43
LCFS	3050.49	78.58	86.67	99.1	24.36	27.18	158.46	646.99	548.07	8.06	3.883	1630.13	88.133
SPT	2815	78.07	85.81	98.68	24.33	26.88	98.7772	669.79	179.34	4.2176	9.3673	1160.07	75.0333
LPT	3198.25	77.97	85.91	96.44	24.08	26.99	90.2238	554.64	854.89	13.5061	3.7742	1614.5	98.8
EDD	2828.63	79.01	87.09	99.66	24.72	27.29	202.48	512.98	297.58	5.725	3.0891	664.17	115.27
FASFO	2828.17	78.76	86.65	99.41	24.47	27.22	85.7427	587.07	333.54	8.3443	5.0413	750.77	111.23
LST	2937.54	78.88	87.12	99.69	24.55	27.18	141.25	484.55	495.32	5.5276	3.8864	873.77	108.27
MCR	3050.88	78.8	86.6	99.11	24.43	27.12	42.3093	577.36	566.65	12.2729	9.0709	1195.62	98.133
MBPT	2881.25	78.94	86.8	99.99	24.61	27.29	104.33	699.64	534.74	15.8439	6.7646	1849.98	84.5667
LBPT	2794.59	79.39	87.21	99.99	24.54	27.38	233.62	431.28	471.62	4.2936	2.3071	942.12	101.43
MNOR	2908.18	79.1	87.16	99.99	24.53	27.31	126.45	1050.65	345.27	4.0419	12.3513	1748.76	93.7333
LNOR	3070.09	78.56	86.87	99.18	24.45	27.19	280.36	280.54	781.24	6.6853	2.2393	1059.94	115.07
GTW	3184.03	78.5	86.26	98.89	24.34	27.02	27.5716	532.23	760.75	16.4246	7.4244	1840.53	84.4
LTW	2795.86	79.18	87.55	99.99	24.64	27.31	242.76	519.86	281.83	5.4066	1.9688	763.92	108.07
MDD	2918.71	78.7	87.01	99.51	24.55	26.99	165.25	426.85	479.87	5.096	3.0327	813.85	112.1
LPTNM	3018.57	74.63	82.28	94.37	23.03	25.86	466.11	1093.47	336.12	5.7365	2.3219	1396.3	137.4
MPTNM	3089.06	78.41	86.47	98.94	24.29	26.97	99.19	459.92	676.07	9.6503	5.671	1607.35	84.0333
PDT	2832.29	79.08	86.93	99.85	24.76	27.47	208.55	509.68	299.08	4.8158	3.1614	670.79	114.93

Table A2 : Rank of JPRs derived from different MCDM methods

JPR	MARCOS		PIV		MABAC		EDAS		TOPSIS	
	U_i	Rank	u_i	Rank	S_i	Rank	A_i	Rank	RC_i	Rank
FCFS	0.5951	10	0.0649	9	0.0769	9	0.5815	10	0.6305	10
LCFS	0.5777	14	0.0687	11	-0.0036	12	0.5134	12	0.6339	9
SPT	0.6593	4	0.0502	8	0.0792	8	0.8282	6	0.6590	8
LPT	0.5590	17	0.0851	14	-0.1700	17	0.4206	14	0.5726	14
EDD	0.6532	5	0.0419	3	0.1637	4	0.8921	3	0.7291	4
FASFO	0.6193	9	0.0467	6	0.1134	5	0.7989	7	0.7425	2
LST	0.6212	8	0.0471	7	0.1101	6	0.7890	8	0.7365	3
MCR	0.5854	13	0.0769	13	-0.0247	14	0.4857	13	0.5853	13
MBPT	0.5569	18	0.0906	16	-0.0086	13	0.3379	17	0.5277	16
LBPT	0.6612	2	0.0446	5	0.1750	3	0.8561	4	0.6971	7
MNOR	0.5917	12	0.0897	15	0.0031	11	0.4032	15	0.5079	17
LNOR	0.6344	7	0.0665	10	0.0120	10	0.6350	9	0.5934	12
GTW	0.5948	11	0.0924	17	-0.1254	16	0.3858	16	0.5369	15
LTW	0.6749	1	0.0407	2	0.1881	1	0.9238	1	0.7076	6
MDD	0.6398	6	0.0434	4	0.1089	7	0.8455	5	0.7426	1
LPTNM	0.5765	16	0.1027	18	-0.4040	18	0.2867	18	0.4624	18
MPTNM	0.5774	15	0.0707	12	-0.0350	15	0.5401	11	0.6238	11
PDT	0.6605	3	0.0403	1	0.1789	2	0.9140	2	0.7285	5

Blockchain-based Efficient and Secure Peer-to-Peer Distributed IoT Network for Non-Trusting Device-to-Device Communication

Rajesh Kumar Sharma and Ravi Singh Pippal
Department of Computer Science & Engineering,
RKDF University, Bhopal, India
E-mail: rajeshsharma.ercs@gmail.com, ravesingh@gmail.com

Keywords: Security, blockchain, internet of things, avalanche effect, device-to-device communications, peer-to-peer communications

Received: April 5, 2021

The security and privacy issues in the Internet of Things (IoT) are a mandatory process and also a challenging task for researchers. Blockchain technology enhanced and motivated the recent security parameters, and it has been validating various technical sectors since its inception. In this paper, a peer-to-peer distributed IoT network is presented where non-trusting devices can interact with other devices without a trustworthy intermediary using the blockchain technique in automated verifiable mode. Major implementation issues for deploying blockchain in the IoT network are pointed out in this paper. The model presents a modern blockchain technique to surpass the traditional security system for efficient and secure IoT deployment under various conditions. Finally, to validate the signification of blockchain in the IoT network, the avalanche effect is calculated and compared with Triple-DES, AES, and Blowfish cryptographic algorithms for non-trusting device-to-device (D2D) communications and transactions. The result presents significant output changes in hash for the blockchain-IoT integrated model as compared to other cryptographic algorithms. Conclusively, blockchain in the IoT network can make a remarkable impact across various industrial and business applications.

Povzetek: Študija predstavlja varno bločno omrežje za IoT, ki omogoča komunikacijo brez zaupanja med napravami, s primerjavo z obstoječimi kriptografskimi algoritmi.

1 Introduction

The most auspicious technologies in this modern era are the Internet of Things (IoT) and the Blockchain. The perfect solution to incorporate decentralized security in IoT is blockchain technology for peer-to-peer communication, and the integration of these two technologies is certainly required in many applications and domains [1]. There are many possible blockchain applications for IoT cases, including the healthcare industry, supply chain management, public safety, personal data management, finance, education, insurance, notary services, smart homes, and cities [2, 3]. Blockchain is continuously blowing up modern IoT-based industries, and it is expected to motivate IoT onward. The distinguishing features, such as decentralized trust and security provided by blockchain, are crucial characteristics of IoT infrastructure [4]. Many researchers are innovatively working on IoT-blockchain collaboration in several ways due to the prevalence of both technologies and inventing extremely secure and robust systems to address the technical problems [5].

A blockchain framework stores information records in attached blocks, which are connected using a cryptographic algorithm [6]. Basically, it has a continuously extending database list in distributed form to maintain record infor-

mation, where participating nodes in the blockchain validate new and existing records [7]. For establishing a trusted network between nodes, any central third-party authentication or participating node is not required in this decentralized blockchain network [8]. All the completed transactions and their information are always shared, distributed, and updated to each node in the blockchain network, so all the nodes have the exact same information record [9–11].

The most secure and transparent system than centralized transactions can be structured by the blockchain. It can efficiently record transactions and details between two parties using a distributed ledger in a permanent and verifiable way. Ultimately, privacy, security, anonymity, and transparency are the main goals of blockchain technology for all of its users.

Information about other physical conditions in the environment is acquired by Internet of Things (IoT) devices, and they communicate and transmit data with each other using inbuilt software systems. IoT devices generate a huge quantity of information and sensing data, but devices do not inevitably trust others at the time of transactions. Critical privacy issues can be raised because connected IoT devices transmit sensitive personal information and can disclose the preferences and behaviors of their users. When personal, sensitive data is utilized by any centralized or-

ganization, the privacy of IoT users is especially at risk due to the illegitimate usage of data. To solve this problem, blockchain technology can be beneficial in structuring privacy-preserving IoT systems. There are many technical benefits of blockchain, as presented in Figure 1. The foremost benefit is security, and other facilities such as open architecture, timestamped, smart contracts, distributed ledger technology, etc. are significant characteristics of blockchain technology.

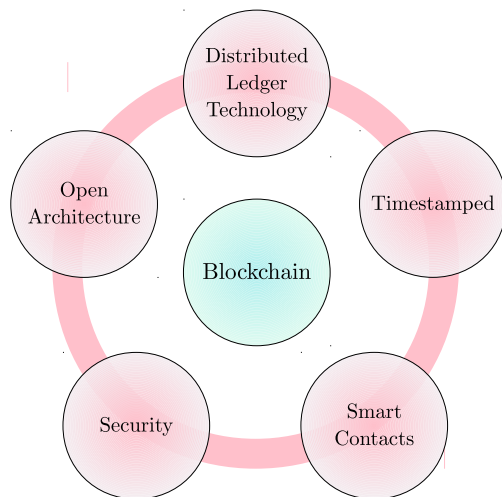


Figure 1: Technological Benefits of Blockchain

Since the blockchain is peer-to-peer (P2P) technology, it is tamper-proof, contains only trustworthy information, and is not dominated by any single centralized entity. The blockchain technology provides secure agreement-based peer-to-peer communications between devices in the IoT network. It also resolves other issues such as trust, privacy, scalability, time-stamping, single points of failure, and reliability challenges of the network for IoT devices. For the transmission of data between devices in a consistent, secured, and time-stamped manner, blockchain technology offers a trustworthy framework for the IoT network. In this infrastructure, IoT devices can take advantage of smart contracts to enable message conversations, which create trust agreements between devices. The characteristics of autonomous activities and intelligent applications for sensing devices can be enabled using these features. To construct a completely distributed, trustworthy blockchain-based digital infrastructure, IoT-based peer-to-peer transactions can be extended to a person-to-device or person-to-person platform. Currently, there are numerous blockchain-based solutions and applications; it has continuously been applied in various cases since its inception.

2 Blockchain technology

Blockchain is a distributed, trusted, shared, and public ledger of transactions that every user can analyze but a single user cannot control. Basically, it is a distributed

database system that preserves the endlessly developing list of all transaction data and records in a cryptographically locked system from any tampering or violation. Since the Blockchain is a distributed system, the concept of a centralized master node is not required, and all the other connected nodes maintain a true copy of the database. An efficient, highly resistant, transparent, and secured digital interaction and transaction storage system is offered by blockchain technology. This system has the potential to enable new industrial and business models.

In blockchain technology, all transaction data and relay information are listed in the connected chain of blocks initiated and generated by anonymous users. It provides advantages such as speed, transparency of transactions, cost efficiency, and high security. Blockchain provides data immutability using cryptographic hashes, where data is stored linearly and the new block keeps a record of the hash value of the previous block. Each transaction is required to undergo validation before being appended to the blockchain by participants in the network. The noteworthy issues constituted by blockchain technology are related mainly to the following:

- Accuracy,
- Trust,
- Intermediaries,
- Decentralization,
- Transparency,
- Transaction freedom.
- Data privacy and security,

3 Beneficial aspect of blockchain-based IoT

The engrossment of blockchain-based IoT research is to gear up an efficient, secure, and trusted peer-to-peer distributed IoT network for communications among non-trusting devices. Blockchain technology offers many beneficial aspects for IoT [12]:

- The blockchain-based shared ledger is immutable and can hold records. Many characteristics of the IoT network, such as the type of device, sensing features, range, embedded software, malfunctions, status, hardware changes, and current position, enhance trust among devices and their data.
- Trust agreement between devices for any sensed and measured value of specific characteristics in the IoT network.
- The smart contracts allow the devices to interact autonomously and independently with other commercial systems and devices.

- Third-party validation, verification, or trust is not required, which provides cost reduction and high transaction speed.
- Easy and efficient recovery of data and information due to distributed ledger recording management during system failure.

The overall important advantage is the transfer of trust in a trustless infrastructure. This is especially applicable in the IoT network, where many types of IoT device manufacturers have a different standard of accuracy, range, and functional characteristics. In the network, if new devices are added, smart contracts can be used to interact and communicate with other IoT devices for transactions, repair services, or replacement. The other devices can assume the responsibilities in case any device malfunctions. It validates and increases the value of data generated by IoT devices.

4 Related work

This section presents existing works, efforts, and literature reviews based on blockchain-integrated IoT systems to target the issues of privacy and security.

Chen *et al.* [14] proposed a data integrity checking system on the Internet of Things (IoT) network using the stochastic blockchain technique. They investigated the potentiality of blockchain to protect data integrity and security in the Internet of Things networks. The conventional decentralized techniques face the issues of network congestion and single-point failure. To overcome these shortcomings, a stochastic-based blockchain method is proposed for verifying data integrity in the IoT network. The proposed method minimizes the quantity of cooperating IoT devices and relies on edge devices to produce the block, which reduces the computational and transmission time. The simulation outcomes present an enhanced success rate against large-area IoT networks with a low number of cooperating devices.

Haseeb *et al.* [13] presented “RTS: a robust and trusted scheme for IoT-based mobile wireless mesh networks” for increasing network coverage with the reliability of the system. Their proposed model for IoT-based mesh networks is presented in Figure 2. In this model, RSA-based cryptography is applied to communication links between gateways and clients of the mesh network with existing malicious devices. The existing data routing method works for static mesh devices and monitors transmission links, which can create a deficient effect on the performance of the network and increase the chance of packet drop ratio. Their proposed infrastructure constitutes a mesh network of mobile clients to perform better network exposure in data transmission lines, considering factors for packet drop reduction and a high ratio of data delivery. This model overcomes the communication cost by using the flooding of distance vector routing over periodic time intervals by mobile mesh clients. Their simulation outcomes present high data rela-

bility as well as a low computational operating expense in different topological networks.

For edge-based devices in the Internet of Things (IoT) network, Pyoung *et al.* [15] proposed “LiTiChain, a blockchain with finite-lifetime blocks”. The LiTiChain handles the difficulty of traditional blockchain as it continuously grows into an information block list. The outdated block of information should be promptly eliminated so that an extended blockchain list can be stored at the end node. To eliminate the information block consistently, a tree-structured end-time ordering graph (EOG) was introduced to arrange block lists according to their endings, and it also maintained the chain connectivity of the blocks.

Mazzei *et al.* [16] designed and implemented secure industrial devices using blockchain-based interfacing techniques. Their proposed system allows interaction devices to be made available to the public as a secure blockchain service. The proposed system can also be easily modified as an independent blockchain-equipped tracking system. The system acts as a connection between blockchain-based security and the industrial Internet of Things, which enable the tokenization of industrial devices.

Lei *et al.* [17] proposed “Groupchain”, an original double-chain structure-based scalable public blockchain system for fog computing in IoT. To address the problem of increased transaction throughput generated by the serialized leader election process, they proposed a double-chain structured-based IoT security system using blockchain. The experimental results of the implemented “Groupchain” prototype present the scalability and transaction efficiency of “Groupchain”.

Muthavhine *et al.* [18] studied concerning cryptographic algorithms applied, especially in the security of the Internet of Things. They collected existing cryptographic methods applied to several IoT devices for encryption and authentication, analyzed the Avalanche effects of cryptographic algorithms for each device, and improved their speed using mathematical methods.

Novo [19] addressed the extensibility problem of accessing, arranging, and managing a large number of secure IoT devices because conventional centralized control to access the system is unable to handle increasing loads effectively. This research introduced a novel access handling control that reduces the management problems of plenty of constrained devices in the Internet of Things network. The proposed technique is a completely decentralized blockchain-based method.

For autonomous cooperative intrusion detection in the devices of large Internet of Things networks, Mirsky *et al.* [20] introduced a blockchain-based security solution. In the IoT network, an agent system is configured for detection of any software exploitation on the devices, which provides regular control for intrusion detection. Any kind of manual activity or update is not required to generate a malware signature in this proposed framework.

Due to the deficient management of the Internet of Things network and devices, particularly the arrangement

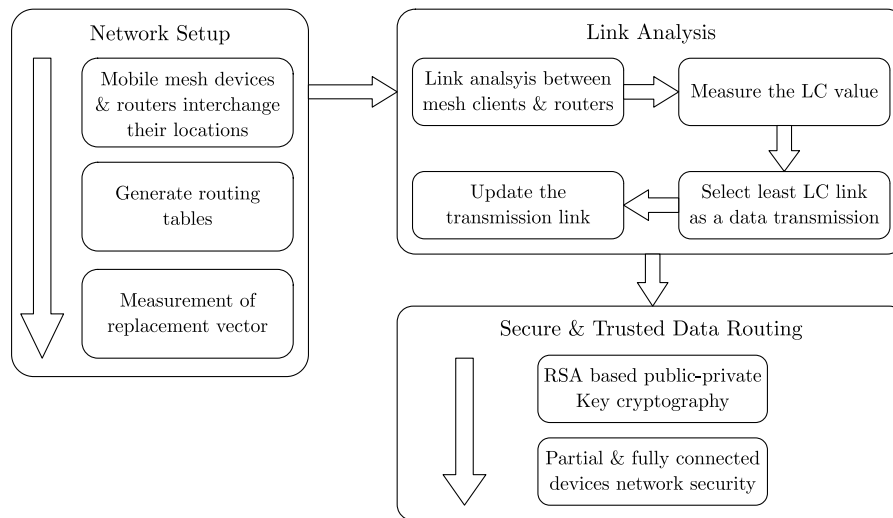


Figure 2: Trusted and Robust Model for IoT-based Mesh Network [13]

and installation method, observable exploitation in the IoT network environment can be detected. To provide a solution to this problem, Yohan *et al.* [21] proposed the “Firmware-Over-the-Blockchain (FOTB) Framework, a secure and efficient blockchain-based firmware update structure between manufacturers of IoT devices and network-deployed devices. In this proposed framework, a peer-to-peer verification technique through a blockchain-based mechanism is applied for the security of firmware distribution activities, which also ensures the integrity of the system in a distributed IoT network.

Li *et al.* [22] designed “Blockchain-based Distributed IoT Data Transaction (BDDT)” a novel data architecture for the IoT network that offers a framework for data producers and users and provides a solution for reliability problems and usage of data facilities offered by the central storage of IoT. The problems of secure data circulation and transaction in IoT network, BDDT system efficiently resolve these problems.

Liu *et al.* [23] combined blockchain technology in the “Attributed-Based Access Control (ABAC) Model”, which carries the benefits of blockchain-based decentralized technology for access control demands into the Internet of Things and solves the problem of conventional access control techniques. The ABAC model provides a device authentication system, implements the management policy of ABAC, and verifies device security access using the smart application. A Hyperledger Fabric-based opensource access control system “Fabric-IoT” is developed and applied in the network. The final steps involved in this model are network deployment using blockchain, installation of chaincodes, and invoking smart contracts.

Rathee *et al.* [24] proposed a blockchain technology-based secure hybrid “Industrial Internet of Things (IIoT)” framework. In this framework, the activities of employees are collected and stored by blockchain-based industrial IoT devices to maintain the security, transparency, and tracing

of all work by producing the hash of each record for IoT devices. The proposed secure framework expressively minimizes the loss ratio of the product and falsification problems in network devices.

Chatterjee *et al.* [25] developed a lightweight authentication network protocol for secure key swapping, text messaging, and supporting the hierarchically architectural framework of the IoT network. To protect against adversarial active and passive threats, the Physically Unclonable Function (PUF) [26] cryptography method is used. The limitations of previous PUF-based security methods can be eliminated using their proposed protocol, which is strong against many threats.

Wang *et al.* [27] proposed a hierarchically structured storage system for storing the blockchain in a cloud network, and it maintains newly added blocks in the blockchain network. They presented a blockchain-integrated Internet of Things architecture to protect and maintain blocks and transactions produced in IoT networks. The blockchain and cloud connection as a software interface is designed to build block synchronization for storage in the cloud.

For the security of supply chain management applications, Malik *et al.* [28] proposed a blockchain-based “Trust Management Framework (TrustChain)” to solve the issues related to the quality trust of commodities and the entity of logging data. Basically, the “TrustChain” framework utilizes blockchain technology to monitor all interactions. The agent- and asset-based reputation model is also provided by this framework; it reaches efficiency and automation through smart contracts with the reputations provided to the same participant for the particular product.

Biswas *et al.* [29] designed and implemented a novel “lightweight block cipher named LRBC” to constrict Internet of Things resources. By integrating Feistel and substitution-permutation networks (SPN), the proposed structure has been implemented, which takes advantage of

both techniques. To resist linear and differential attacks, LRBS produces a high quantity of extinct S -boxes. Guruprakash and Koppu [30] carried out an empirical investigation to showcase that the “Edwards curve digital signature algorithm (EdDSA)” can serve as a performance-enhancing alternative to the “elliptic curve digital signature algorithm (ECDSA)” in the context of Blockchain and IoT. Tanweer Alam [31] introduced “IBchain”, an IoT and blockchain-integrated system that can be used for secure communications in a smart city network.

Various state-of-the-art methods for securing the IoT network have been provided in the literature, with their advantages and limitations. Many researchers addressed blockchain-based solutions in their unique and specific proposed methods. However, these studies lack validation and comparative analysis with other applicable modern cryptographic algorithms. To address this research gap, a comparative analysis of blockchain with modern cryptographic techniques is provided in this research, using the Avalanche effect as a parameter to observe the significant changes in the hash value.

5 IoT management in blockchain

The management of IoT devices in blockchain technique offers essential information to the significant structured layer and works on the controllable components of the systems. Figure 3 illustrates the structured layer of blockchain-based IoT management. The four major components of

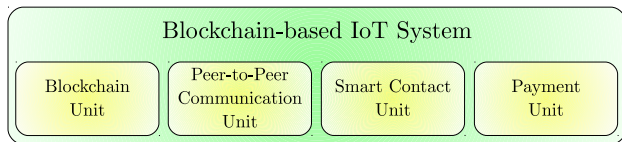


Figure 3: IoT Management in Blockchain

this structured layer are represented as:

- **Blockchain Unit:** It manages useful information and blockchain activities. All kinds of ledger data and information are used by this part. It contains three sub-units:

1. Blockchain of things,
2. Microservice for blockchain,
3. Smart contracts.

The microservice index contains information about smart contracts.

- **Peer-to-peer Communication Unit:** Data communication, exchange, and transfer are facilitated by this component with the help of peer-to-peer network technology. The management process related to the blockchain of things is also controlled by this component.

- **Smart Contract Unit:** The function and process required for smart contracts defined by the system are provided by this component. The blockchain unit is responsible for storing smart contract codes and information. The smart contracts are capable of processing the mechanism of the system.
- **Payment Unit:** All the payment processes and transactions are supported by this unit in cooperation with the structured function layer and other components of the IoT device management layer. The wallet information of the user and IoT devices is also managed by this unit.

6 Proposed method

In cryptographic analysis, the avalanche effect is basically a mathematical function applied to the encryption technique, and it is evaluated as the most preferable attribute of the encryption algorithm. The avalanche effect presents a considerable change in the ciphertext if a few changes are made in either the plaintext or key. This basic property is known as the avalanche effect in cryptography. Generally, it measures the quantity of effect on the ciphertext concerning minor alterations made to the key or the plaintext.

The method is to consider fixed-length inputs to the hash function f . Otherwise, it is problematic what probability distribution it wants to impose on the input set $\{0, 1\}^*$ which is the collection of all finite input strings. In practice, hash functions do have an upper limit on the input string, but that’s astronomical, in terms of testing all input strings. The simple explanation of the Avalanche Effect is that “A small change in the plaintext (or key) should create a significant change in the ciphertext”. Concerning these characteristics, “Data encryption standard (DES)” has been proven to be significantly strong.

So, let’s assume the hash function has a security parameter of k bits. This corresponds to the function acting like a random function with outputs of length $n = 2k$ bits.

The testing would generate numerous random values from a uniform distribution on $\{0, 1\}^m$, thus treating the hash function as a random function $f : \{0, 1\}^m \rightarrow \{0, 1\}^n$. Let this random set of inputs be denoted by X . Now define

$$a_{ij} = \#\{x \in X : [f(x \oplus e_i)]_j \neq [f(x)]_j\} \tag{1}$$

for $1 \leq i \leq n, 1 \leq j \leq m$, where e_i is the vector with a one in the i^{th} position and zeroes everywhere, and $[u]_j$ denotes the j^{th} component of vector u . a_{ij} counts the number of inputs from X that differ in the j^{th} output bit when the i^{th} input bit is flipped.

It can now be defined as a degree of strict avalanche criterion, $D_{\text{SAC}}(f)$ as

$$D_{\text{SAC}}(f) := 1 - \sum_{i=1}^n \sum_{j=1}^n \left\| \frac{2a_{ij}}{\#X} - 1 \right\| \tag{2}$$

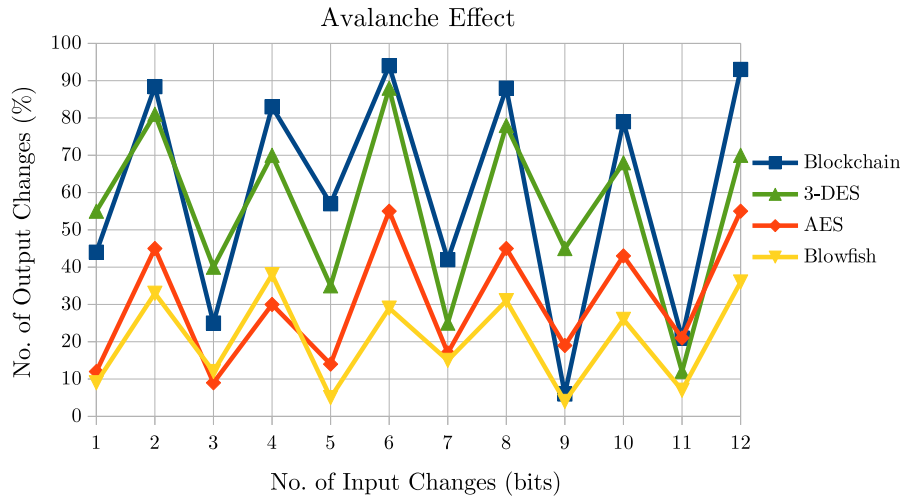


Figure 4: Avalanche Effect on Cryptographic Algorithms

with the expectation that $D_{SAC}(f)$ should be approximately 1, i.e., the sum of the absolute differences. A modification in a single bit of plaintext or key generates a significant modification in many bits at the ciphertext outcome; this is well known as the Avalanche effect.

$$\text{Avalanche Effect} = \frac{N_c}{N_{p|k}} \quad (3)$$

where, N_c is the number of bit changes in ciphertext and hash, whereas $N_{p|k}$ is the number of bit changes in plaintext or key in the IoT-generated data. The Avalanche effect hash function is presented in Figure 5.

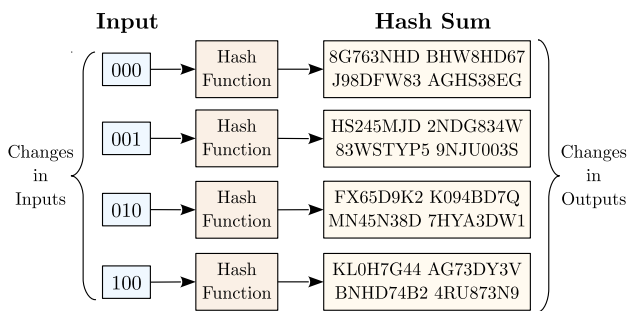


Figure 5: Avalanche Effect Hash Function

7 Result analysis

After a successful simulation, the test result shows the outcome of the Avalanche effect. Similar and uniformly small changes were made to the input value of plaintext or key (IoT-generated data), and the hash functions generated a hash sum as the output. For testing purposes, three cryptographic encryption methods are considered "Triple Data Encryption Standard (3-DES), Advanced Encryption Standard (AES), and Blowfish", along with the blockchain method. A desirable attribute of any encryption technique

is that some minor modification in either plaintext generated from IoT devices or the key must generate a considerable difference in the ciphertext output and its associated hash value. The simulation result is presented in Figure 4. The Avalanche effect as a significant change in hash value is presented in the Table 1, which shows the percentage of output changes in hash with respect to the number of input bits changes for Blowfish, AES, 3-DES, and Blockchain. Table 1 is basically a numeric representation of the simulation result (Figure 4). The quantity of changes in the percentage of hash output depends on the cryptographic algorithms and the changes in the number of input bits. Table 1 definitely shows that the blockchain presents high percentages of variations in hash output as compared to other cryptographic methods, as 1–2 bits changes in input affect 9%–33% changes in a hash using Blowfish, 12%–45% for AES, 55%–81% for 3-DES, and 44%–88.4% for Blockchain. Similarly, when changing 11–12 bits in input, the percentage of changes in hash output for Blowfish is 7%–36%, for AES it is 21%–55%, for 3-DES it is 12%–70%, and for Blockchain it is 21%–93% as the highest changes. The blockchain percentage may not vary uniformly with the increment in the number of inputs because the avalanche effect is basically a ratio of the number of bit changes in ciphertext or hash to the number of bit changes in plaintext or key. The significance of the result is considered based on the maximum percentage of variation in the hash for the blockchain. The result shows the high avalanche effect of blockchain as compared to Triple-DES, AES, and Blowfish. Significantly, the maximum changes in the bits of ciphertext can be observed in the blockchain method as compared to other cryptographic techniques. A blockchain-based IoT network is more secure for non-trusting device-to-device communications and transactions.

Our objective is to present a method for efficient and trustworthy P2P communications and transactions in a blockchain-based distributed IoT network for non-trusting D2D communication without a centralized 3rd party.

Table 1: Avalanche Effect on Cryptographic Algorithm

No. of Input Bits Changes	Output Changes (%)			
	Blowfish	AES	3-DES	Blockchain
1–2	9%–33%	12%–45%	55%–81%	44%–88.4%
3–4	12%–38%	9%–30%	40%–70%	25%–83%
5–6	5%–29%	14%–55%	35%–88%	57%–94%
7–8	15%–31%	17%–45%	25%–78%	42%–88%
9–10	4%–26%	19%–43%	45%–68%	6%–79%
11–12	7%–36%	21%–55%	12%–70%	21%–93%

Changes made by intruders in IoT-generated data can be validated using the Avalanche effect, and blockchain-based integrity can be provided by using the proposed model.

8 Conclusion and future work

This research presents a significant peer-to-peer distributed IoT network based on blockchain technology for secure and efficient non-trusting device-to-device communication and transaction. Manipulating and integrating the IoT network with blockchain modeled a secure system successfully. The model presents a modern blockchain technique to surpass the traditional security system for efficient and secure IoT deployment under various conditions. Finally, to validate the signification of blockchain in the IoT network of non-trusting device-to-device communication, the avalanche effect is calculated and compared with Triple-DES, AES, and Blowfish cryptographic algorithms using IoT-generated data. The result presents significant output changes in hash for the blockchain IoT integrated model as compared to other cryptographic algorithms. Using the Avalanche effect calculation, the hash function with the encryption technique of blockchain can significantly provide strength to the IoT network, as proven by the security level validation.

The proposed work can be applied to applications based on intrusion detection techniques. This work can be extended by comparing blockchain with other hybrid cryptographic methods. In the future, using the Avalanche effect, the validation assessment can be done in the fog/edge computing model as well as on a cloud server. The avalanche effect assessment is an engrossing process to notice the IoT-generated data processing techniques of blockchain that can be enhanced for validity checking in other applications such as banking, financial, business, and industrial transactions.

References

- [1] P. Cui, U. Guin, A. Skjellum, and D. Umphress, “Blockchain in iot: Current trends, challenges, and future roadmap,” *Journal of Hardware and Systems Security*, vol. 3, no. 4, pp. 338–364, Dec 2019. doi: <https://doi.org/10.1007/s41635-019-00079-5>
- [2] P. Rathee, *Introduction to Blockchain and IoT*. Singapore: Springer Singapore, 2020, pp. 1–14. doi: https://doi.org/10.1007/978-981-13-8775-3_1
- [3] V. Gatteschi, F. Lamberti, and C. Demartini, *Blockchain Technology Use Cases*. Singapore: Springer Singapore, 2020, pp. 91–114. doi: https://doi.org/10.1007/978-981-13-8775-3_4
- [4] C. Davila and J. Tarnow, *The Blockchain in IoT*. Cham: Springer International Publishing, 2019, pp. 269–296. doi: https://doi.org/10.1007/978-3-319-99516-8_10
- [5] M. H. Miraz, *Blockchain of Things (BCoT): The Fusion of Blockchain and IoT Technologies*. Singapore: Springer Singapore, 2020, pp. 141–159. doi: https://doi.org/10.1007/978-981-13-8775-3_7
- [6] G. Pulkkis, J. Karlsson, and M. Westerlund, *Blockchain-Based Security Solutions for IoT Systems*. John Wiley & Sons, Ltd, 2018, ch. 9, pp. 255–274. doi: <https://doi.org/10.1002/9781119456735.ch9>
- [7] A. Erdem, S. Ö. Yildirim, and P. Angin, *Blockchain for Ensuring Security, Privacy, and Trust in IoT Environments: The State of the Art*. Cham: Springer International Publishing, 2019, pp. 97–122. doi: https://doi.org/10.1007/978-3-030-18075-1_6
- [8] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” <https://www.bitcoin.org>, 03 2009.
- [9] S. Van Hijfte, *Blockchain Platforms: A Look at the Underbelly of Distributed Platforms*. Morgan & Claypool, 2020. doi: <https://doi.org/10.2200/S01022ED1V01Y202006CSL011>

- [10] S. S. Shetty, C. A. Kamhoua, and L. L. Njilla, *Blockchain for Distributed Systems Security*. Wiley-IEEE Press, 2019, pp. 205–232. doi: <https://doi.org/10.1002/9781119519621.ch10>
- [11] M. Swan, *Blockchain: Blueprint for a New Economy*, 1st ed. O'Reilly Media, 2015.
- [12] G. Fiorentino, C. Occhipinti, A. Corsi, E. Moro, J. Davies, and A. Duke, *Blockchain: Enabling Trust on the Internet of Things*. John Wiley & Sons, Ltd, 2020, ch. 11, pp. 141–157. doi: <https://doi.org/10.1002/9781119545293.ch11>
- [13] K. Haseeb, I. Ud Din, A. Almogren, N. Islam, and A. Altameem, “Rts: A robust and trusted scheme for iot-based mobile wireless mesh networks,” *IEEE Access*, vol. 8, pp. 68 379–68 390, 2020. doi: <https://doi.org/10.1109/ACCESS.2020.2985851>
- [14] Y. Chen, L. Wang, and S. Wang, “Stochastic blockchain for iot data integrity,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 373–384, 2020. doi: <https://doi.org/10.1109/TNSE.2018.2887236>
- [15] C. K. Pyoung and S. J. Baek, “Blockchain of finite-lifetime blocks with applications to edge-based iot,” *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2102–2116, 2020. doi: <https://doi.org/10.1109/JIOT.2019.2959599>
- [16] D. Mazzei, G. Baldi, G. Fantoni, G. Montelisciani, A. Pitasi, L. Ricci, and L. Rizzello, “A blockchain tokenizer for industrial iot trustless applications,” *Future Generation Computer Systems*, vol. 105, pp. 432–445, 2020. doi: <https://doi.org/10.1016/j.future.2019.12.020>
- [17] K. Lei, M. Du, J. Huang, and T. Jin, “Groupchain: Towards a scalable public blockchain in fog computing of iot services computing,” *IEEE Transactions on Services Computing*, vol. 13, no. 2, pp. 252–262, 2020. doi: <https://doi.org/10.1109/TSC.2019.2949801>
- [18] K. D. Muthavhine and M. Sumbwanyambe, “An analysis and a comparative study of cryptographic algorithms used on the internet of things (iot) based on avalanche effect,” in *2018 International Conference on Information and Communications Technology (ICOIACT)*, 2018, pp. 114–119. doi: <https://doi.org/10.1109/ICOIACT.2018.8350759>
- [19] O. Novo, “Blockchain meets iot: An architecture for scalable access management in iot,” *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1184–1195, 2018. doi: <https://doi.org/10.1109/JIOT.2018.2812239>
- [20] Y. Mirsky, T. Golomb, and Y. Elovici, “Lightweight collaborative anomaly detection for the iot using blockchain,” *Journal of Parallel and Distributed Computing*, vol. 145, pp. 75–97, 2020. doi: <https://doi.org/10.1016/j.jpdc.2020.06.008>
- [21] A. Yohan and N.-W. Lo, “Fotb: a secure blockchain-based firmware update framework for iot environment,” *International Journal of Information Security*, vol. 19, no. 3, pp. 257–278, Jun 2020. doi: <https://doi.org/10.1007/s10207-019-00467-6>
- [22] H. Li, L. Pei, D. Liao, X. Wang, D. Xu, and J. Sun, “Bddt: use blockchain to facilitate iot data transactions,” *Cluster Computing*, May 2020. doi: <https://doi.org/10.1007/s10586-020-03119-w>
- [23] H. Liu, D. Han, and D. Li, “Fabric-iot: A blockchain-based access control system in iot,” *IEEE Access*, vol. 8, 2020. doi: <https://doi.org/10.1109/ACCESS.2020.2968492>
- [24] G. Rathee, A. Sharma, R. Kumar, and R. Iqbal, “A secure communicating things network framework for industrial iot using blockchain technology,” *Ad Hoc Networks*, vol. 94, p. 101933, 2019. doi: <https://doi.org/10.1016/j.adhoc.2019.101933>
- [25] U. Chatterjee, R. S. Chakraborty, and D. Mukhopadhyay, “A puf-based secure communication protocol for iot,” *ACM Trans. Embed. Comput. Syst.*, vol. 16, no. 3, Apr. 2017. doi: <https://doi.org/10.1145/3005715>
- [26] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, “Physical one-way functions,” *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002. doi: <https://doi.org/10.1126/science.1074376>
- [27] G. Wang, Z. Shi, M. Nixon, and S. Han, “Chainsplitter: Towards blockchain-based industrial iot architecture for supporting hierarchical storage,” in *2019 IEEE International Conference on Blockchain (Blockchain)*, 2019, pp. 166–175. doi: <https://doi.org/10.1109/Blockchain.2019.00030>
- [28] S. Malik, V. Dedeoglu, S. S. Kanhere, and R. Jurdak, “Trustchain: Trust management in blockchain and iot supported supply chains,” in *2019 IEEE International Conference on Blockchain (Blockchain)*, 2019, pp. 184–193. doi: <https://doi.org/10.1109/Blockchain.2019.00032>
- [29] A. Biswas, A. Majumdar, S. Nath, A. Dutta, and K. L. Baishnab, “Lrbc: a lightweight block cipher design for resource constrained iot devices,” *Journal of Ambient Intelligence and Humanized Computing*, Jan 2020. doi: <https://doi.org/10.1007/s12652-020-01694-9>
- [30] J. Guruprakash and S. Koppu, “An empirical study to demonstrate that eddsa can be used as a performance improvement alternative to ECDSA in blockchain and iot,” *Informatica (Slovenia)*, vol. 46, no. 2, pp. 277–290.
- [31] T. Alam, “Tbchain: Internet of things and blockchain integration approach for secure communication in smart cities,” *Informatica (Slovenia)*, vol. 45, no. 3, pp. 477–486.

CoBiAt: A Sentiment Classification Model Using Hybrid ConvNet-Dual-LSTM with Attention Techniques

Roop Ranjan and A K Daniel

Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP, India

Keywords: deep learning, Dual-LSTM, keras, FasText, attention, emotion analysis

Received: Januar 14, 2022

Many researchers have recently turned their attention to emotion analysis as a resultant to the number of social reviews of various services. User behaviour may be better understood with the plethora of data, which makes it possible to work toward enhancing QoS. The critical areas of research in language processing are text categorization, which places unorganised data into relevant categories. In several natural language processing (NLP) applications, LSTM and CNN are utilised for text classification. CNN models use techniques to obtain top-level features. In this study, an attention-based model using Dual-LSTM and ConvNet has been proposed. For effectiveness verification of the model, it is trained using two unique datasets. The proposed hybrid model has demonstrated a significant performance gain when compared to previous deep learning techniques. In comparison to other traditional machine learning models, the suggested approach yields outcomes with a higher level of accuracy.

Povzetek: Predstavljena je metoda CoBiAt za klasifikacijo čustev z uporabo Dual-LSTM in ConvNet .

1 Introduction

The widespread usage of social media allows people to provide comments on events, situations, services, and products qualities [1]. These comments are frequently based on user experiences, which may include good or bad thoughts about items or services. These suggestions will assist firms in improving their services, allowing them to generate enhanced profit. As a result, it is critical to assess user input gathered from social networking websites. Analysis of Sentiments is useful in expressing users' opinions (positive, neutral or negative) about their services through text data [2]. It has also been observed that scholars are becoming more interested in social media platforms such as Facebook and Twitter. Businesses embrace public opinion analysis because it describes human activity and behaviour, as well as how individuals are influenced by the viewpoints of others.

In recent era, the widespread applications using deep learning has led to advancements in image processing, natural language processing, and speech recognition. Deep learning is superior to machine learning in classification of sentiment analysis problems due to the availability of large datasets and the inexpensive mass fabrication of capable Graphics Processing Unit (GPU) units [5]. Deep learning is extensively utilized in the models based on natural language processing (NLP) [6], including emotion analysis, because to its autonomous learning characteristics. The two most often used deep learning algorithms in sentiment analysis of reviews are CNN and Recurrent Neural Network (RNN). Gradient disappearing and exploding problems plague RNN to a large extent. RNN is challenging to train for long distance

correlations for a given series because of these difficulties. The RNN model underpins BiLSTM, which has shown promising results in text-based sentiment analysis. An LSTM model like this features channels for two-way communication to help the network comprehend its environment. The forward and backward layers of BiLSTM allow the network to access the sequence's prior and subsequent contexts [7].

In the past few years, many CNN- or RNN-based methods for classifying text have been proposed [8, 9]. CNNs can learn the local behaviour from temporal data, but not sequential correlations. RNNs are specialised for sequential modelling, as opposed to CNNs, but are unable to extract features in parallel. Traditional RNNs, however, result in an exploding and vanishing state versus its gradient for extended data sequences. Vanishing gradient and gradient explosion problems are successfully solved by long short-term memory (LSTM) [10], a type of RNNs architecture with LSTM units as hidden units.

The attention-based mechanism aids in enhancing the performance of deep learning models, sentence summarization, and reading comprehension, as demonstrated in the machine learning translation process [11, 12]. The majority of deep learning implementations for text analysis use word embedding techniques to produce feature vectors from the dataset. BiLSTM is unable to prioritise the critical information from the data and just collects contextual information from the features. In contrast to BiLSTM, CNN has a convolutional layer that extracts and shrinks vector features.

In this research we aim to construct a novel text classification utilising a hybrid deep learning attention-oriented optimal model that integrates the strong

characteristics of CNN with BiLSTM employing an attention mechanism in order to handle the aforementioned problem. By adding a convolutional layer to the CNN model with attention features, the attention-based Conv-BiLSTM mechanism, a unique approach, aims to solve the shortcomings of BiLSTM. The suggested model's overall methodology involves training the input data using the Keras skip-gram model, which is then passed to the Convolution Layer, which draws out the data's basic semantic information. The BiLSTM layer, which combines the attention-based approach to identify which characteristics are significantly associated with semantics and should be employed for final classification, receives the feature vectors obtained by the Conv layer.

The training and performance evaluation of the proposed model is performed on two datasets; the first set of data is the collection of tweets for Indian Railways (hereafter IR Dataset) starting from 01st October 2019 and 31st October 2019 [13] and the second dataset is IMDB reviews. Word embedding is performed using two prominent word embedding algorithms, FasText and Keras embedding. Experiments demonstrated that the Self Attention Based Conv-BiLSTM model outperformed other deep learning models and conventional machine learning methodologies.

The following are the key contributions of the research:

(1) Two distinct Word embedding approaches FasText and Keras embedding were used to render tweets as word vectors. Both of the strategies for word embedding make use of pre-trained, supervised word vectors that are able to capture the semantics of individual words and are taught using a large corpus of words. The effectiveness of the proposed CoBiAt model will be evaluated through the utilisation of these two-word vector models as the objective.

(2) ConvNet that has been coupled with BiLSTM and the Self Attention methodology has been offered for classifying the reviews. The ConvNet module collects local features via word embedding, the Self Attention-based BiLSTM module extracts long-distance associations, and the selected features are then categorised in the classification result.

(3) The experimental results are compared with other popular deep learning-based techniques and standard machine learning approaches to prove the potential of the proposed optimised model.

There are several drawbacks in deep learning models like CNN, RNN and LSTM etc. The main issue with CNN is that it does not provide clear encoding the orientation and position of content. Whereas RNN suffers from exploding problems and gradient vanishing issues. The LSTM technique takes longer time for training the dataset also implementing dropout in LSTM is a very tedious task. All the above issues with deep learning methods motivated us to design an approach that overcomes the drawbacks of deep learning methods. Therefore, we performed regressive experiments and proposed a hybridized model that combines the strong features of CNN with BiLSTM using an efficient attention-

mechanism to further optimise the performance and accuracy of the classification provided by the model.

The following is the structure of this research paper: The background and important research for text sentiment categorization using ConvNet and attention-based Dual-LSTM are discussed in Section 2. Section 3 describes in detail how the proposed model works. Section 4 describes the environment setup for executing the CoBiAt model. Section 5 discusses in depth the experimental outcomes for the research and comparison with other models. Section 6 finishes with a conclusion and suggestions for future research.

2 Related work

In recent times, public opinion analysis on network Deep learning techniques have shown remarkable accomplishments in the area of natural language processing in recent years. Deep learning is subfield of machine learning that seeks to represent high-level of abstractions in the given data. This is accomplished through the use of model architectures with complex structures or those built of several nonlinear transformations [14]. Convolutional Neural Network (CNN) learns complicated, high-dimensional, and non-linear mapping relationships by fully utilising the structure of multi-layer perceptron. It has been frequently utilised and has produced good results in image identification and speech recognition applications [15].

CNN was proposed for use in natural language processing by [16], who also constructed a dynamic Convolution Neural Network (DCNN) technique to analyse dataset of varying lengths. In [17], the authors developed a method for analysing opinions regarding health services. They amassed 2,026 tweets using Twitter hashtags to compile their dataset. They compared DNN and CNN with word2vec embeddings as two DL models. CNN's model was the most accurate. The model was trained on a fairly limited dataset in this work, and neither model addressed the negation problem. In [18], five different combinations of LSTM models were utilised to analyse tweets. To train the models, they utilised both dynamic and static CBOW and a word embedding. The results demonstrated that the integrated LSTM model trained using dynamic CBOW performed better than the other models.

Many neural network-based techniques have recently been demonstrated to be effective in a number of sentimental analysis applications. Given the attention-based mechanism's tremendous success in natural language processing, its application in Sentiment Analysis tasks has gained in prominence. The paper [19] proposed using the deep attention mechanism to analyse user evaluations and produced better results than a recurrent neural network (RNN). The model [20] suggested a bidirectional gated recurrent unit-based position-aware bidirectional attention network (PBAN) (bi-GRU). To tackle this difficulty, the LSTM model [20] was suggested, which had the capacity to maintain sequence information and produced good results on several sequence modeling tasks.

A powerful tool for capturing the relationship between context and aspect at the next level is an attention mechanism. The authors in [21] made a multi-grained attention network (MGAN) that uses both fine-grained and coarse-grained attention mechanisms to collect information about how aspect and context interact with one another. In another research [22] used the concept of feed-forward networks and multi-head attention (MHA) to effectively extract the hidden context presentation of and embeddings of aspects. In the paper [23], movie reviews from IMDb were analysed to see how people felt about them. Some pre-processing steps were taken to get rid of characters, symbols, words that were repeated, and "stop words." Then, CountVectorizer was used to extract features. The authors proposed CNN model and compared with several traditional models. The proposed CNN model achieved 99.33% accuracy. The dataset has 3000 reviews that are either good or bad. Meng et al. [24] used a CNN to get the higher-level feature representation from a better layer of word embedding. After a BiLSTM takes in local and global semantic information, an attention layer is used to focus on important term characteristics of aspects. Furthermore, Ma et al. created an attention-based LSTM that uses the common-sense knowledge of sentiment-related concepts proposed in SenticNet for incorporating external knowledge [25].

In contrast to conventional machine learning classification algorithms, LSTM has demonstrated its effectiveness in achieving high classification accuracy [26]. This article [27] implemented deep CNN and BiLSTM. Deep CNN was implemented on character-level embeddings for improving word embeddings' information. Then, bidirectional LSTM is used for classifying the sentences according to their sentiment. This article emphasises data standardisation in order to obtain high performance, the researchers have created a tweet processor model to remove unnecessary terms from tweets. In their research [28], offer Hybrid two Convolutional Neural Networks and Bidirectional LSTM, a further variant of hybridized deep learning architecture for sentiment classification. Two CNN layers and a bidirectional LSTM layer are employed here. Seven datasets are evaluated using three pre-trained word vectors, GloVe, Word2Vec, and FasText. Word2Vec has been observed to be more efficient than the other two-word vectors.

In another research [29], the authors suggested a network of deep memory for sentiment classification. The proposed method may simultaneously record both user and product information. Inference-based memory components and a large long-term memory that also serves as a knowledge basis compose this memory network. Model architecture has two parts. Each document is represented by an LSTM, and its rating is predicted using a deep memory network with several levels (hops), all of which are content-based attention models.

The researchers in [30] used Dual LSTM and Keras word embedding to classify traveller attitudes. The authors' proposed model performed word embedding using the two distinct words embedding algorithms

word2Vec and Keras embedding. Keras embedding provided better results as compared to that of word2Vec. The authors in [31] devised hierarchy-based attention (HA) technique for capturing the hierarchical structure of documents at the sentence and text levels, where information of varying significance was given special treatment while generating document representations. Because most prior approaches focused just on localized information of contents and ignored preferences of global users and quality of product, [32] proposed a model for classification of sentiment using attention method for product information for global users. For text sentiment analysis, [33] recommended combining CNN with three distinct attention mechanisms: LSTM attention, vector attention, and pooling attention. [34] used a self-attention with a sparse technique for determining text emotion polarity by capturing the significance of each word. In another work, [35] looked at the challenge of classifying personality traits based on textual content. The authors used a hybrid CNN+ LSTM model to classify the text in a different personality trait. The research performed by the author's shows that CNN is a powerful method for selecting the best characteristics that improve prediction accuracy and the LSTM model keeps earlier context information, which makes it easier to use important context information at the start of a phrase. Their proposed model proved to be significantly superior to other models. However, the lack of attention model was observed in their results.

In innovative research in field of bio-medical engineering in [36], researchers presented a data analytical system for EEG utilising a multi-layer Gated Recurrent Unit (GRU) for anomaly identification. The proposed model consists of four stages from data collection to model performance evaluation. Using a publicly accessible EEG dataset, the suggested model achieved accuracy of 96.91 percent, sensitivity of 97.95 percent, specificity of 96.16 percent, and 96.39 percent F1 score. In [37] Wang et al., demonstrated the concept of weak tagging for sentiment classification. The authors presented a BiLSTM emotion categorization model with multi stages of training and a emotion classification model with weak tagging data denoising. The model for emotion classification based on weak type tagging information denoising had the best classification performance of all the experimental groups, but the model was also observed to be the most time-consuming.

In the field of medical imaging deep learning has shown promising results. In [38], authors have demonstrated a model based on Transfer Learning (TL). Using a state-of-the-art CNN for fundus image processing, we customised TL for mild and multi-class DED (diabetic eye disease) cases in this study. Using fine-tune, optimization the researchers achieved 88.3% accuracy. In another article [39], a unique hybrid CNN-BiLSTM deep learning strategy for four-level facial pain recognition is proposed. To achieve the results of pain intensity estimation satisfactorily, the fully connected layer of the VGG-Face was optimized for this task by addition of a fully connected layer, and the dimensions of the extracted features was reduced using PCA (Principal Component

Analysis) to improve the algorithm's overall computational efficiency. The improved algorithm achieved an AUC of 98.4% and a test precision of 90%. By training a large number of sentiment text corpora, Using algorithms convex hull and convolution neural network, [40] proposes a model for fault detection in wireless sensor networks. The authors performed several experiments and found that CNN with Naïve Byes is proving to be better and more efficient. The authors in [41], presented a model using block-chain with deep learning. Described all the fundamental ideas involved in the management and security of such data, and offered a novel solution to handle the hospital's big data utilising Deep Learning and Block-Chain technology to ensure their safety.

Since RNN can preserve a sequence of information over time, it is a helpful supplement to CNN; nonetheless, RNN is severely affected by the Gradient explosion problem described by [42]. Because of these issues, distance correlation in a sequence is difficult to train with RNN. Bi-LSTM is an RNN model that has reportedly shown promising outcomes in the analysis of sentiment in the given text. It contains two LSTMs to allow the network for better understanding of the contexts provided. Backward LSTM's and forward layers grant the for accessing the sequence's preceding and subsequent context. Text sentiment classification, on the other hand, uses vectors to represent the text, which is often performed in a large-dimension space. When the Bi-LSTM retrieves relevant knowledge from several obtained features, it is unable to place a premium on the most important data [43].

Considering the above issues in deep learning methods for text classification, the CoBiAt model has been proposed here that combines the strong feature of CNN and Bi-LSTM with attention model. The performance of the model indicates that the proposed model has the potential to tackle above issues in deep learning models.

3 Proposed method

The proposed attention-based method has the following layers:

- 3.1 Input and Pre-Processing
- 3.2 Word Vector Matrix
- 3.3 ConvNet Layer
- 3.4 Dual-LSTM Layer
- 3.5 Attention Layer

3.1 Input and Pre-Processing Layer

In this part, the words are standardized and cleansed by changing them from a human language to text format in order to eliminate superfluous elements. This phase assists classifiers in achieving high performance and rapid sentiment classification. In this work, steps included tokenization of sentences into related words by using

NLTK (Natural Language Toolkit), a popular python library: convert upper case to lower case; duplicate text removal, removal of special characters, reduction of several spaces, hashtags, URL, mention, punctuations and other stopwords; and lemmatization of texts in dataset.

3.2 Word vector matrix layer

The Word Vector Matrix layer embeds pre-processed input tokens. Token representation reveals hidden ties between words that often appear together. The input dataset was trained using the Keras Skip-gram model. Here the text from the input dataset has N-words, vector of the t^{th} words is represented by w_t with $t \in [1, T]$ in the given text. Here the text is given with words w_t , the words are embedded by layers for vectors using a layer of embedding W_p . x_t is depiction of vector for w_t , which is represented using Eq.1.

$$x_t = W_p w_t \quad (1)$$

The model was trained with skip-gram approach by optimising the average likelihood log. The method trains semantic embedding by prediction of the target word based on context and detecting semantic links between words.

3.3 ConvNet layer

With the ConvNet layer, the task of selecting features from input data is accomplished. In view of reducing the number of dimensions in the source text, ConvNet layers are employed. Several 1-Dimensional Convnet kernels are employed for conducting convolution for the input vectors in this study. Equation (2) creates a sequential vector of text by integrating the component vectors of word embedding:

$$X_{1:V} = [x_1, x_2, x_3, x_4, \dots, x_V], \quad (2)$$

where V represents the total size of tokens in the given text. Convolution kernels of different sizes are implemented on $X_{1:V}$ for capturing the n-gram characteristics (where $n=1,2,3$) of the given text using a one-dimensional Convolution to capture the intrinsic features. When a window of r words spanning from $v: v+r$ is input during the v^{th} convolution, the convnet process creates features set for such window as below:

$$h_{r,v} = \tan h(W_r x_{v:v+r-1} + d_r), \quad (3)$$

Where $X_{v:v+r-1}$, are the embedded vectors obtained of the given word in a given window, W_r is the weights matrix that can learn, and d_r is the bias used in embedding. Since every filter primarily has to be implemented to different parts of the word, therefore the filter's feature maps with the size of convolution r are:

$$h_r = [h_{r1}, h_{r2}, h_{r3}, h_{r4}, \dots, x_{V-r+1}]. \quad (4)$$

It is advantageous to utilise Convnet kernels of varying sizes in order to capture hidden connections between adjacent words. The most essential objective of

using a CNN for text-based feature retrieval is that it minimises the number of learnable parameters throughout the max-pooling-based process of feature learning. Multiple channels of convolution act on the input, with each channel holding data at distinct timestamps. Consequently, the output of each ConvNet channel during the max-pooling operation is the largest value of all timestamps for that channel. Max pooling is implemented here to the maps of features with convolution size r for each convolution kernel to produce:

$$s_r = \text{Max}^v(h_{r1}, h_{r2}, h_{r3}, h_{r4}, \dots, x_{V-r+1}) \quad (5)$$

For obtaining the final feature map of the window, p_r is combined for every filter size $r = 1, 2, 3$ and for extracting the n -gram (where $n=1, 2, 3$) hidden features:

$$h_r = [s_1, s_2, s_3]. \quad (6)$$

3.4 BiLSTM layer

The Dual-LSTM layer receives attributes as input from the ConvNet layer and extracts the final hidden state to produce features. The prior and subsequent context information is accessible to the Dual-LSTM module, and the data collected by the BiLSTM can be seen in two distinct text-based formats. The ConvNet feature sets are input into the Dual-LSTM model, which generates a sequential representation. By aggregating information for words in both directions (ahead and backward), Dual-LSTM obtains word annotations and therefore the annotations include contextual information. The forward LSTM represented as (\vec{L}_1) reads the sequences of features from first to last, whereas the reverse LSTM represented as (\vec{L}_2) reads from last to first. The word annotation received from \vec{L}_1 is represented by $L_{forward}$ and from \vec{L}_2 the annotation is represented by $L_{backward}$ the bidirectional process is iterated and the final feature representation output (L) is obtained as follows:

$$L = \{L_{forward}, L_{backward}\} \quad (7)$$

3.5 Attention layer

This output representation from the Dual-LSTM layer is delivered to the attention layer, which assesses which features are highly interconnected and should be used for final categorization. The attention mechanism, which is a fully linked layer with a softmax function that focuses on qualities of selected words to reduce the impact of less significant words on the text's sentiment, is a completely connected layer. The process of the attention layer is as follows:

The word annotation $L_{forward}$ is initially supplied to get $\vec{U}_{forward}$ by single perceptron as an uncovered format of $L_{forward}$. $\vec{U}_{forward}$ is represented as follows:

$$\vec{U}_{forward} = \tanh(f * L_{forward} + b) \quad (8)$$

Where weight and bias in the neuron is represented by f and b respectively. Hyperbolic tangent function is represented by $\tanh()$. The layer calculates the significance of each word based on the similarity between $\vec{U}_{forward}$ and a text level context vector $\vec{C}_{forward}$ for measuring the importance of every text. The softmax function is then utilized by the layer for calculating the normalized weight \tilde{Z}_{fwd} of each word as follows:

$$\tilde{Z}_{fwd} = \frac{\exp(\vec{U}_{forward} * \vec{C}_{forward})}{\sum_{i=1}^M \exp(\vec{U}_{forward} * \vec{C}_{forward})} \quad (9)$$

Here, the total number of texts in a particular set of texts is denoted by M .

The text level context vector $\vec{C}_{forward}$ is an illustration of high level for the descriptive words for the set of word sequences that is initialized on a random basis and learned together throughout the phase of training.

The forward context representation Fc is then produced as a weight-based addition of the read word annotations in the forward direction on the basis of weight parameter \tilde{Z}_{fwd} . Fc is a component of the attention layer's output that may be represented as:

$$Fc = \sum(\tilde{Z}_{fwd} * L_{forward}) \quad (10)$$

Similar to \tilde{Z}_{fwd} , \tilde{Z}_{bwd} is calculated with the help of the backward direction hidden state $L_{backward}$. Hc , like Fc , is a backward context representation that is part of the attention layer's output, and it is represented as:

$$Hc = \sum(\tilde{Z}_{bwd} * L_{backward}) \quad (11)$$

The forward context representation Fc is concatenated with the backward context representation denoted as Hc , BiLSTM gets an interpretation for a particular sequence of features, and finally delivers the classification results. The attention layer improves the accuracy of prediction and decreases the size of learned weights required for predicting by using this method.

4 Experiments

4.1 Dataset

Experiments are carried out in this part to evaluate the performance of the proposed model for text categorization on two unique datasets. Using Roop

Ranjan et al. [30] as a starting point, we have 25000 tweets from people who used Indian Railways services on various days in October 2019. Table 1 breaks down these tweets into three different categories: positive, neutral, and negative. The binary-labeled set of IMDB movie reviews is the second dataset used. There are 14641 review tweets collected in the dataset as shown

in Table 2. The dataset from IMDB movie was also utilised to compare the model to prior sentiment categorization research. The IMDB movie dataset can be obtained from <https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews/metadata>.

In this section, experiments are conducted for

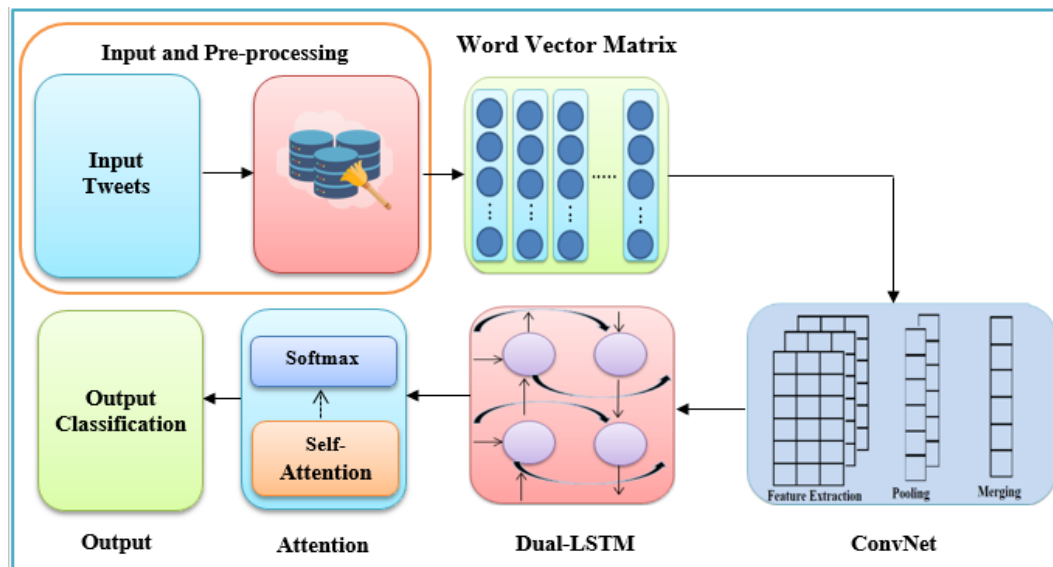


Figure 1: Proposed architecture

Table 1: Categorization details of tweets dataset

Tweets Dataset			
Size	Positive	Negative	Neutral
25000	10695	8953	5352

Table 2: Categorization details of IMDB movie reviews dataset

IMDB Movies Review Dataset			
Size	Positive	Negative	Neutral
50000	18324	20625	11051

Further both dataset are divided into Training, Validation and Testing Set using Gaussian distribution. The ratio of the total dataset is 60:20:20 for training, validation and testing dataset.

4.2. Experimental setup

Because of the GPU environment's support, Google Colab with Keras is being used with backend such as Tensorflow for Keras. Computing-intensive machine learning methods can be trained in shorter amounts of time when running on GPUs. In a GPU context, greater computational power is available, allowing for more training iterations while fine-tuning the machine learning models.

4.3 Setting of hyper-parameters

Implementing hyper-parameter tuning is crucial for High model performance can only be achieved if hyper-

parameter adjustment is implemented. The randomised search method is utilized to optimise hyper-parameters and improve the accuracy. Using a random combination of hyperparameters, randomised search determines the optimal answer for developing the model. Due to grid search's inability to perform well when there are a large number of dimensions, random search is preferred over grid search. Table 3 represents the hyper-parameters values using randomised search in the proposed model.

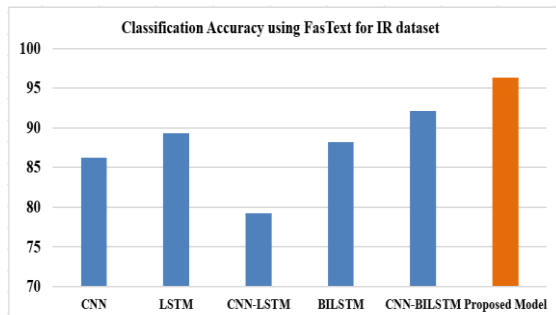
Table 3: Setting of hyper-parameters

Parameters	Values
Dimension(Embedding)	Keras(300)
Size of Kernel	5
Output Size(Dual-LSTM)	32
Filter Size	32
Function (Regularization)	L2
Activation	SoftMax
Weight Constraints	Kernel Constraints (max norm is 3)
Epochs Count	100
Batch Size	32
Batch Normalization	Yes
Learning Rate (LR)	0.03
Optimization	Adam

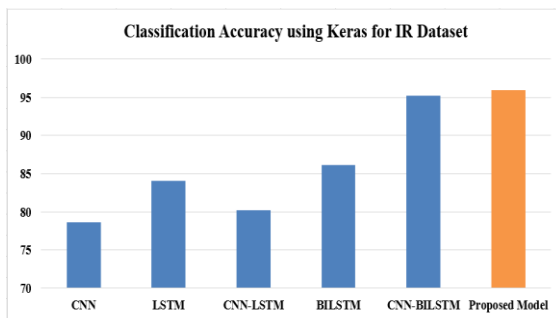
5 Results and discussion

5.1 Performance comparison

Using optimal hyper parameter values, the proposed model was compared to CNN, LSTM, CNN-LSTM, and BiLSTM, which are all deep learning-based models. The IR dataset and the IMDB Movies review dataset were used to make the comparison. Figure 2a shows the results of comparing the proposed model's overall accuracy with other deep learning models using FasText Embedding. On the IR dataset, CNN, LSTM, and BiLSTM all did better with FasText embedding than with Keras embedding. However, CNN-LSTM and CNN-BiLSTM with Keras embedding (Fig. 2b) gave more accurate results. The fact that the observation was made shows that the proposed model is much better than other methods. For the IR dataset, the proposed model is 96.32 percent accurate with FasText and 95.98 percent accurate with Keras.

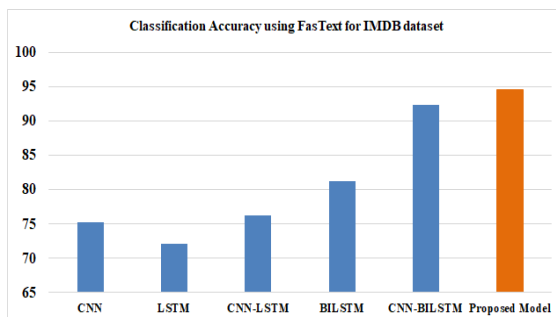


(a)

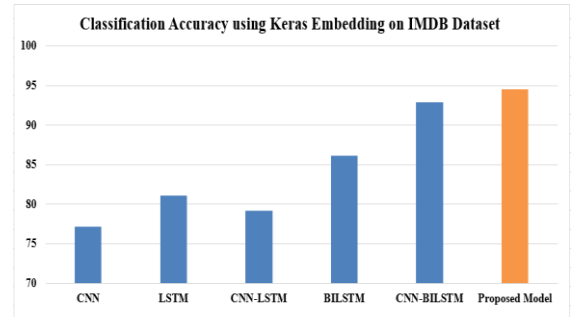


(b)

Figure 2: Accuracy of different models for IR dataset



(a)



(b)

Figure 3: Accuracy of different models based for IMDB dataset

The figure 3a and 3b provides comparison of the overall accuracy of the proposed architecture with other deep learning methods using Keras Embedding. The observation depicts that FasText embedding performs better than Keras embedding for IMDB dataset with an improvement of 1.12%.

5.2 Evaluation of performance

The proposed system's performance is assessed using the standard evaluation matrix illustrated in Figure 4.

		Predicted Class	
		No	Yes
Observed Class	No	TN	FP
	Yes	FN	TP

Figure 4: Standard evaluation parameters

The standard validation parameters are described as below:

- True Negative (TN) - These are accurately forecasted negative outcomes, demonstrating the value of actual class is zero and the outcome of the anticipated class is zero i.e. correct prediction of negative classes.
- True Positive (TP) - TP are observed positives that are accurately predicted and indicate that the outcome of the actual class and the outcomes of the expected class are positive i.e. correct prediction of positive classes.

False negative and false positives happen if actual class is different from the anticipated class.

- False Positive (FP) – observations of the anticipated class is positive and actual class is negative i.e. incorrect prediction of positive classes.
- False Negative (FN) - When the actual class is positive while the projected class is negative i.e. incorrect prediction of negative classes.

Using these standard parameters following rules are implemented for evaluation of effectiveness of the proposed hybrid model:

Precision (P) = Precision is termed as the proportion of correct anticipated positive outcomes to total projected positive outcomes.

$$P = \frac{TP}{TP + FP} \tag{12}$$

Recall(R) = the proportion of properly forecasted positive outcomes to the overall observations in the positive class.

$$R = \frac{TP}{TP + FN} \tag{13}$$

F-Measure (F) = the average of Recall and Precision is termed as F-Measure. Resulting in score takes into account both false negatives and false positives.

$$F = \frac{2 * (P * R)}{(P + R)} \tag{14}$$

Accuracy (A) = the most important performance parameter is accuracy; this is just the proportion of predicted observations which are correct to all observations.

$$A = \frac{TN + TP}{TN + TP + FN + FP} \tag{15}$$

Figure 5 illustrates the overall performance of two independent datasets and two distinct word embedding procedures utilising a variety of deep learning techniques. The suggested model outperformed competing strategies for both the IR and IMDB datasets. The overall precision of IR dataset with FasText embedding was observed 96.32% which is 3.07% more than the CNN-BiLSTM Model and outperformed three other models with a huge improvement. The recall value is improved by 2.12% than the nearest best performing model BiLSTM. The overall performance of the proposed model having F-measure was observed to be 96.16% and accuracy of 96.32% for FasText embedding. The performance of the proposed model has shown lesser performance when Keras Word Embedding is implemented on the different deep learning techniques and the proposed model. The model proposed here with Keras embedding displays the promising improvement for the classification with 96.01% precision, 96.32% recall, 96.16% F-measure and 95.98% accuracy. This was observed because CNN and LSTM lacked proper information about the forthcoming context of the network's huge corpus of words.

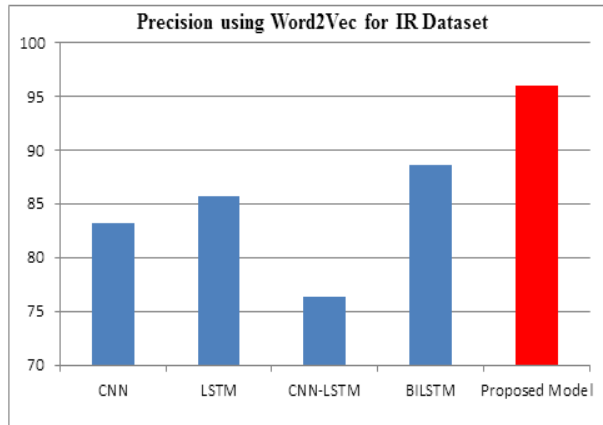
For IMDB dataset the study presented that the proposed model has shown improvement on other techniques. The proposed model with FasText shows the improvement of 3.03% over the CNN-BiLSTM model in predicting positive class prediction. As far as recall is concerned the improvement is 3.06%. For F-measure it shows improvement by 2.56%. The performance of proposed model over other studied techniques with Keras Embedding is also impressive with improvement of 94.63% in precision and 94.56% in accuracy.

Models	Dataset	FasText(%)				Keras(%)			
		P	R	F	A	P	R	F	A
CNN	IR	84.25	82.35	83.29	86.21	79.58	77.63	78.59	78.62
LSTM		86.98	85.25	86.11	89.32	81.52	82.32	81.92	84.12
CNN-LSTM		81.25	78.63	79.92	79.25	79.56	79.65	79.60	80.18
BILSTM		89.92	86.25	88.05	88.25	86.62	85.14	85.87	86.14
CNN-BILSTM		93.25	93.89	93.57	92.15	94.34	96.7	95.51	95.19
Proposed Model		96.32	96.01	96.16	96.32	96.01	96.32	96.16	95.98
CNN	IMDB Movie Reviews	74.21	75.21	76.32	75.21	76.33	77.1	76.71	77.17
LSTM		73.25	72.25	74.26	72.11	81.52	82.32	81.92	81.13
CNN-LSTM		76.32	79.35	79.31	76.32	79.56	79.65	79.60	79.16
BILSTM		82.26	77.65	82.15	81.25	86.62	85.14	85.87	86.14
CNN-BILSTM		90.36	91.62	92.65	92.35	92.63	93.42	93.02	92.9
Proposed Model		95.12	94.68	95.21	95.68	94.63	94.12	94.37	94.56

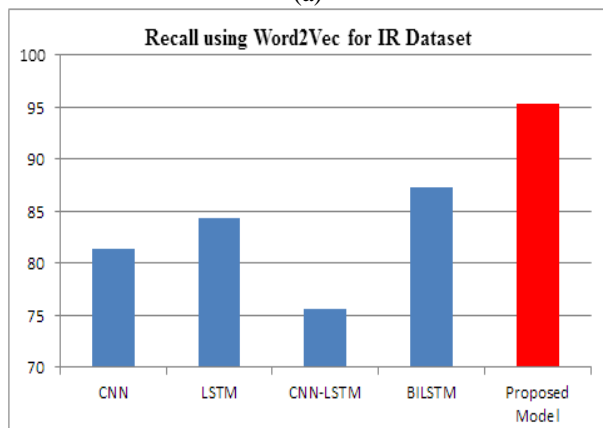
Figure 5: Performance evaluation on faxtext and keras word embedding

For US Airlines dataset the study presented that the proposed model has shown improvement on other techniques. The proposed model with Word2Vec shows an improvement of 7.34% over the BiLSTM model in predicting positive class prediction.

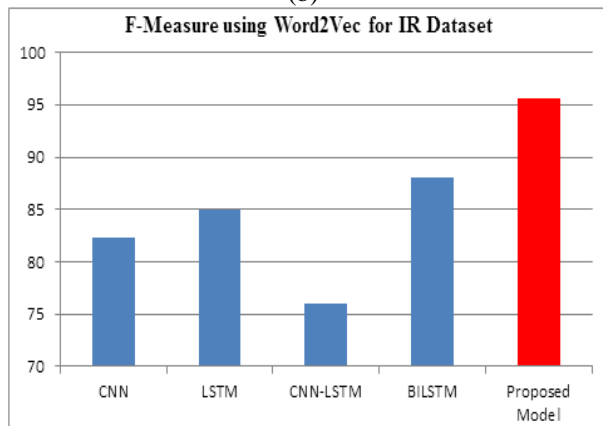
As far as the recall is concerned the improvement is 7.96%. For F-measure, it shows an improvement of 7.65%. The performance of the proposed model over other studied techniques with Keras Embedding is also impressive with an improvement of 9.39% in precision and 9.84% in accuracy.



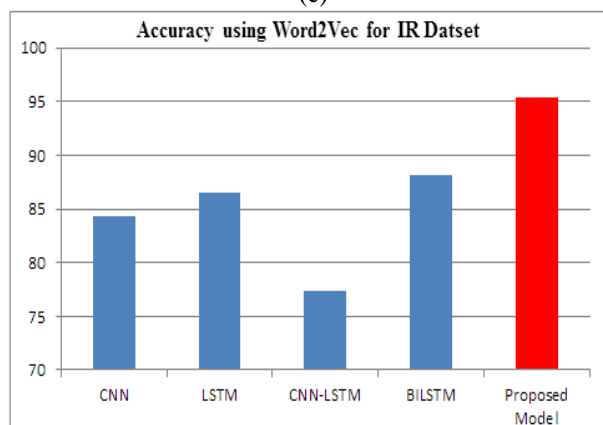
(a)



(b)



(c)

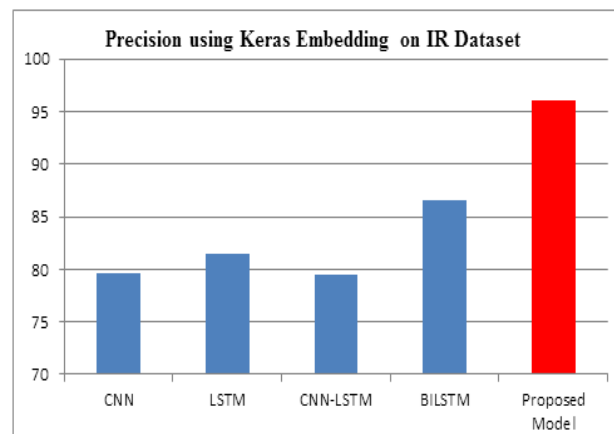


(d)

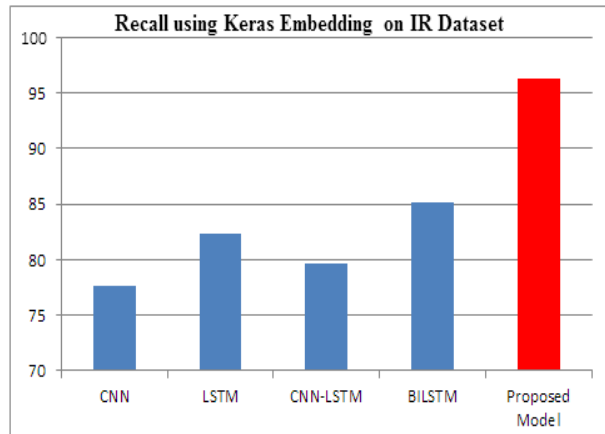
Figure 6: Model performance Word2Vec word embedding for IR dataset

Figure 6 shows the performance of different evaluation parameters using the proposed model on Word2Vec Embedding for the IR dataset. Experiments were performed on another deep learning model on the same dataset. Word2Vec effectively initializes word vectors for the IR datasets, as shown by the higher level of correctness of experimental outcomes. It is clear that the proposed self-attention-based classification model provides better results for all evaluation parameters as compared to other techniques with Word2Vec. The values of precision for CNN were observed at 83.16%, the LSTM model performed precision 85.65%, CNN-LSTM performed 76.32%. The BiLSTM deep learning model performed much better than the other three with the precision of 88.62% and 88.15% classification accuracy. The proposed Model outperformed other deep learning models with a precision of 95.96%. Other performance parameters are also evaluated for different models. Comparison of the proposed model is performed and the proposed self-attention-based model has shown an impressive improvement over other deep learning models with recall 95.32%, F-measure of 95.64%, and accuracy of 95.35%.

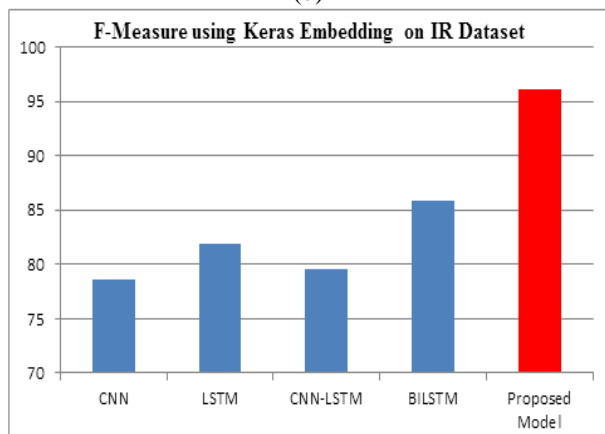
The accuracy of only CNN in the study was merely 82.30%, while the accuracy of experiments on BiLSTM was 87.99% on the IR dataset, indicating that utilizing CNN and BiLSTM separately to conduct sentiment analysis did not yield useful results. Further features of BiLSTM and CNN were combined to optimize the accuracy performance, this approach shows better efficiency than CNN and BiLSTM having accuracy of 91.60% on the IR dataset.



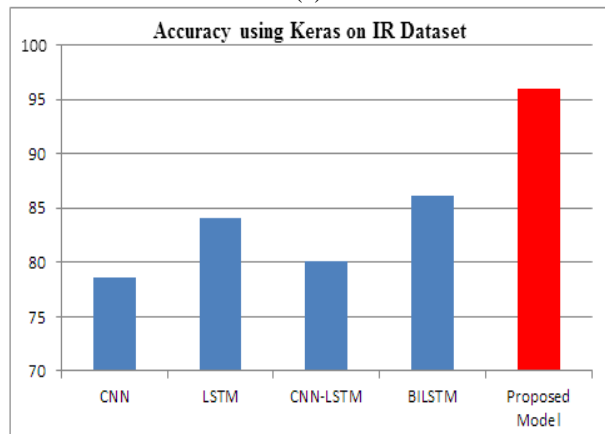
(a)



(b)



(c)

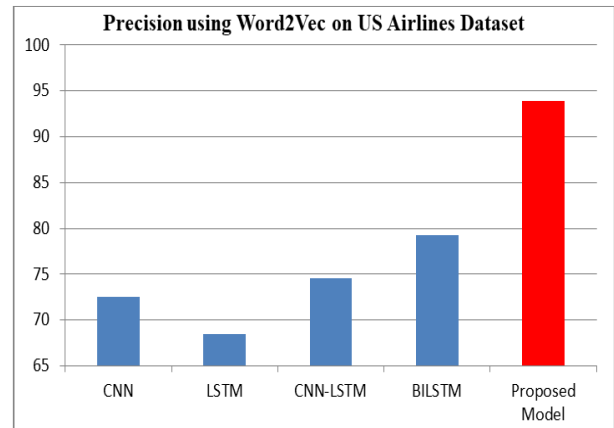


(d)

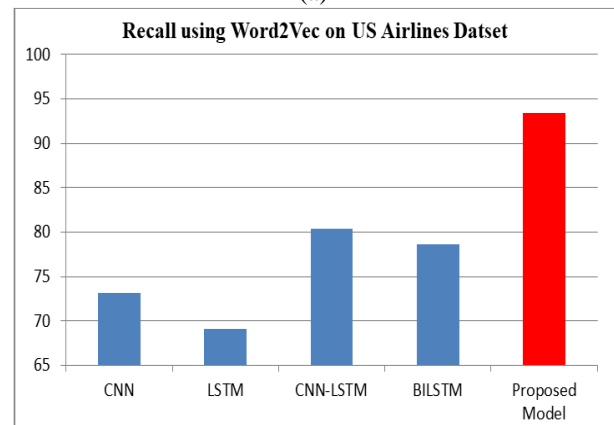
Figure 7: Model performance using Keras word embedding for IR dataset

Different deep learning models were also implemented on same IR dataset using Keras Embedding. Figure 7 displays the model performance on the said embedding. For each of the performance evaluation parameters the proposed model has shown significant improvement over others. The overall accuracy of the proposed optimized model was 9.84% better than BiLSTM method. Precision, Recall, and F-Measure were 9.39%, 11.18%, and 10.29% higher than the BiLSTM model. This indicates clearly that CNN and BiLSTM can't offer great results on their own since CNN can't learn the correlation sequence for long-term dependencies and BiLSTM can't extract local

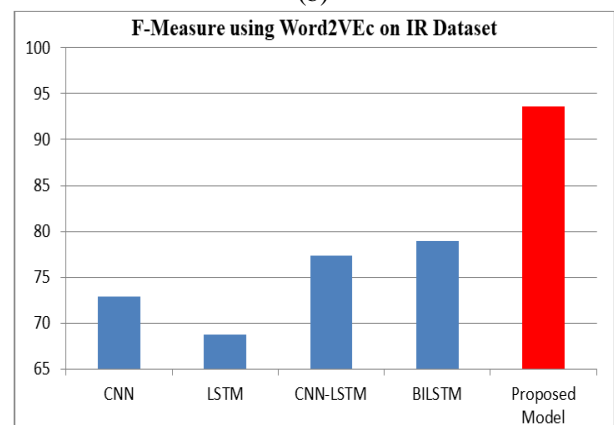
characteristics. The combination of CNN and BiLSTM is merged with self-attention; the model is able to learn each word in tweets more effectively since it contains enough word context information based on previous and future context.



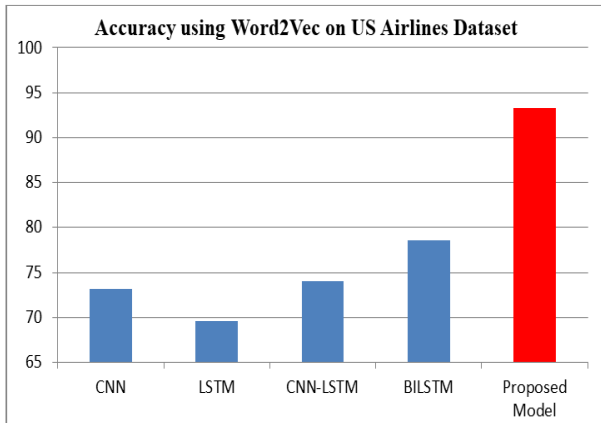
(a)



(b)



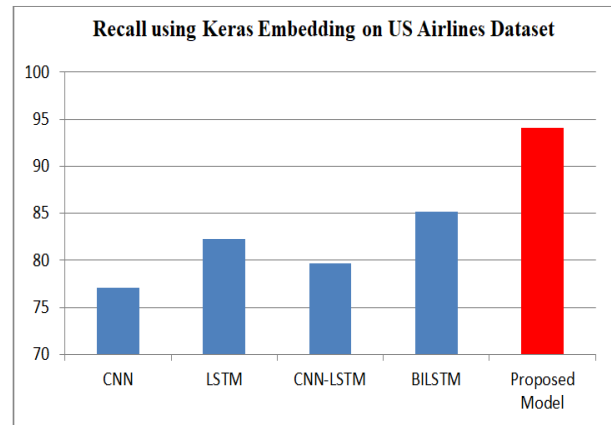
(c)



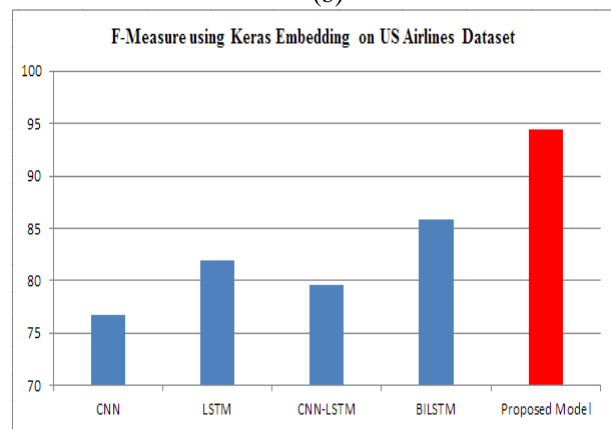
(d)

Figure 8: Model performance using Word2Vec word embedding for US Airlines dataset

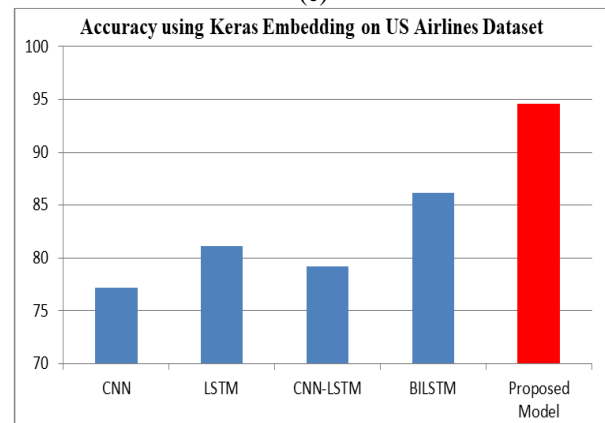
Since the proposed model performed much better on IR dataset using different word embeddings Word2Vec and Keras. Therefore, further experiments were conducted for testing the validity of the performance of the proposed model. The proposed model was then implemented on the US Airlines dataset using both word embeddings that were used for the IR dataset. The overall performance of the proposed model using Word2Vec embedding is shown in Figure 8 and Keras word embedding is represented in Figure 9. On the US Airlines dataset, the performance of the given self-attention-based model is observed to be reduced than the performance of the IR dataset but still, the proposed model outperforms other deep learning techniques in performance with 93.89%, 94.63% precision for Word2Vec and Keras Embeddings respectively. The classification accuracy on US Airlines dataset was also impressive and better than other models with 93.24% for Word2Vec and 94.56% for Keras Embedding.



(b)

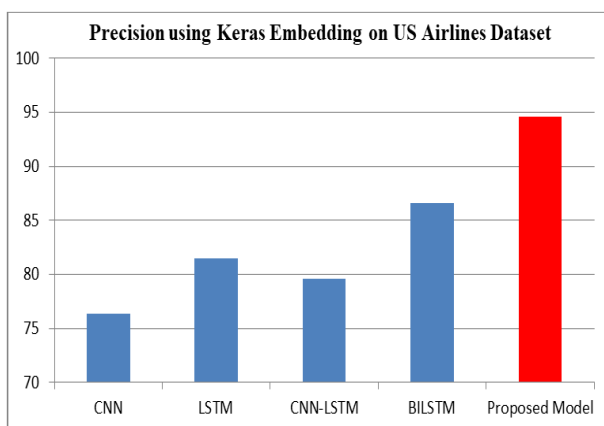


(c)



(d)

Figure 9: Model performance using Keras embedding for US Airlines dataset



(a)

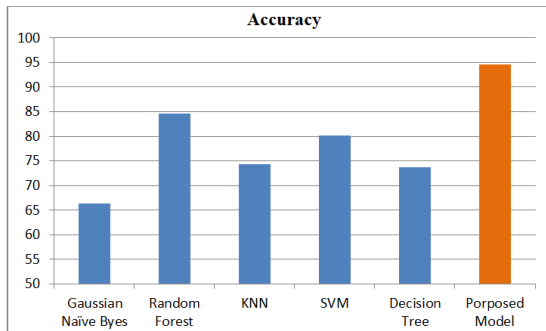


Figure 10: Accuracy of the proposed model over traditional models

Extensive experiments were performed for classic learning techniques on US airlines dataset for classification for validating the effectiveness of the model. Figure 10 demonstrates the accuracy level of the models. The proposed model outperformed the other traditional model also with a very high level of accuracy as compared to other models. The Gaussian Naïve Bayes shows the least performance among all with an accuracy of classification 66.60%. The Decision Tree method performed better than Gaussian Naïve Bayes with an accuracy of 73.5%. On the other hand, KNN, SVM, and Random Forest methods achieved an accuracy of 74.01%, 80%, and 84.5% respectively. The proposed model performed the most optimized accuracy with 94.56%.

5.3 Performance comparison with other state of the art model

The experimental findings were compared to earlier work in text sentiment classification methodologies to ensure that the performance model is verified.

Table 5: Experimental findings for sentiment classification accuracy in %

Models	Accuracy	Reported by
CNN-BiLSTM	90.66	Rhanou et al. [45]
CNN-BiLSTM	94.20	Zi-xian et al. [46]
BiLSTM with Self-Attention	86.24	Jun-Xie et al. [47]
BiLSTM with Muli-Head Attention	92.11	FEI et al. [48]
Conv-LSTM-Conv	89.02	Ghorbani et al. [49]
Text-CNN	91.50	Chuantao et al. [50]
CNN-BiLSTM with Keras	95.98	Proposed Model

Rhanou et al. [45] suggested a model that combines CNN with BiLSTM models using Doc2vec embedding for long text emotion analysis. The composite neural network model presented in [46] is comprised of two parts: a convolutional neural network for extracting local features from text vectors, a BiLSTM that extracts globalized features linked to context of text, and a fusion of the attributes collected by the two complementary models.

The sentences is automatically classified by the trained neural network based hybrid model. The results of experiments reveal that the accuracy rate of text classification is 94.2%, with a total of 10 iterations. For polarity classification of fine-grained sentiment in small sized texts, a BiLSTM model based on Self-Attention-Based using information of aspect-term is presented in Jun-Xie et al. [47]. A layer of word-encoding, a Dual LSTM module, an attention-based module, and a softmax function module are the primary constituents of the model. The vector based on hidden feautres and the vectors of aspects are merged by inserting in the BiLSTM module and the module based on self attention reducing computational complexity imposed by direct vector division. The model [47] achieved 86.24 accuracy which is lesser in comparison to proposed model. The model described in [48] investigates analysis of sentiment for Chinese text on social media by merging Multi-head Attention (MHAT) mechanism with BiLSTM networks for addressing the shortcomings of classic sentiment analysis. The goal of researchers was to add weights of influence to the generated sequence of text due to the MHAT mechanism's ability to learn important information from a distinct representation subspace utilising numerous dispersed computations. The model presented in [48] provided 92.11% accuracy. Gorbani et al. [49] suggested a ConvNet with BiLSTM model that classifies features using CNN, learns context information using BiLSTM, and then reuses the results for CNN to produce an abstract feature before applying to the final dense layer. The model achieved a great accuracy of 89.02 percent. The proposed model, on the other hand, is simpler and requires less complexity analysis, yet it achieves a greater accuracy of 6.96 percent more than [49]. Chuantao et al. [50] presented the BiLSTM deep learning model with two weak-tagging stages. The suggested approach employed weak-tagging for training the proposed model, lowering the detrimental influence of samples of noise in weak-tagging for the categorization of sentiment model's performance of categorization, and increases the accuracy of the sentiment categorization approach, which achieves 91.50 percent accuracy. In comparison to [50], the suggested Keras embeddings model outperformed the [50] model.

Based on the comparison with these previous models it is evident that the proposed model with Keras embeddings outperforms other models and achieves much better accuracy.

6 Conclusion

The research was performed on two different datasets. Word2Vec and Keras word embedding methods were applied for training and evaluation of the model on both datasets. The proposed model integrated the features of CNN with BiLSTM with the self-attention mechanism. ConvNet collects text characteristics and passes text context information to BiLSTM. The attention mechanism improved the classification accuracy as it extracted the context of the sentence more accurately. Hyper-parameters tuning was performed to optimize the model.

Therefore, the proposed model performed classification with improved accuracy and efficiency. The model provided improvement in accuracy by 7.20% using Word2Vec embedding for IR dataset and 9.84% more accuracy in classification using Keras embedding. The model performed effectively on US airlines dataset as well with 8.77% more accuracy with word2vec embedding and 8.03% more accuracy for Keras embedding. The proposed model outperformed other traditional models for the US Airlines dataset.

References

- [1] Liu B, *Sentiment Analysis and Opinion Mining* (Synthesis Lectures on Human Language Technologies), vol. 5, no. 1. San Rafael, CA, USA: Morgan & Claypool, 2012, pp. 1_167. Accessed: Nov. 1, 2020, doi: 10.2200/S00416ED1V01Y201204HLT016.
- [2] Tang HF, Tan SB and Cheng XQ, *Research on sentiment classification of Chinese reviews based on supervised machine learning techniques*. Chin. Inf. Process., vol. 21, no. 6, pp. 88_126, 2007.
- [3] Liu Y, J.W. Bi, and Z. P. Fan. *A method for multi-class sentiment classification based on an improved one- vs-one (OVO) strategy and the support vector machine (SVM) algorithm*. Inf. Sci, vols. 394_395, pp. 38_52, Jul. 2017.
- [4] Zhang J and Zong C, *Deep neural networks in machine translation: An overview*. IEEE Intell. Syst., vol. 30, no. 5, pp. 16_25, Sep./Oct. 2015.
- [5] Yin W, Schütze H, Xiang B, and Zhou B, *ABCNN: Attention-based convolutional neural network for modeling sentence pairs*. Trans. Assoc. Comput. Linguistics, vol. 4, pp. 259_272, Dec. 2016.
- [6] Ansari A, Maknojjia M, and Shaikh A, *Intelligent question answering system based on artificial neural network* in Proc. IEEE ICETECH, Coimbatore, India, Mar. 2016, pp. 758_763.
- [7] Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, and Kuksa P, *Natural language processing from scratch*. J. Mach. Learn. Res., vol. 12 pp. 2493–2537, Aug. 2011
- [8] Levy O, Goldberg Y, and Dagan I, *Improving distributional similarity with lessons learned from word embeddings*. Trans. Assoc. Comput. Linguistics, vol. 3, pp. 211–225, May 2015
- [9] Liu G and Guo J, *Bidirectional LSTM with attention mechanism and convolutional layer for text classification*. Neurocomputing, vol. 337, pp. 325_338, Apr. 2019.
- [10] Bahdanau D, Cho K, and Bengio Y, *Neural machine translation by jointly learning to align and translate*. 2014, arXiv: 1409.0473. [Online]. Available: <https://arxiv.org/abs/1409.0473>
- [11] Rush A M, Chopra S, and Weston J, *A neural attention model for abstractive sentence summarization*. 2015, arXiv: 1509.00685. [Online]. Available: <https://arxiv.org/abs/1509.00685>
- [12] Hermann K M, Kocisky T, Grefenstette E, Espeholt L, Kay W, Suleyman M, and Blunsom P, *Teaching machines to read and comprehend* in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 1684–1692.
- [13] Ranjan R., Daniel A.K. (2021), *Intelligent Sentiments Information Systems Using Fuzzy Logic*, Information and Communication Technology for Intelligent Systems. ICTIS 2020. Smart Innovation, Systems and Technologies, vol 195. Springer, Singapore. https://doi.org/10.1007/978-981-157078-0_55
- [14] Jiang L, Yu M, Zhou M, Liu X, and Zhao T, *Target-dependent Twitter sentiment classification* in Proc. ACL, 2011, pp. 151–160.
- [15] Perez-Rosas V, Banea C, and Mihalcea R, *Learning sentiment lexicons in Spanish*. LREC, vol. 12, p. 73, May 2012
- [16] Zhang M, Zhang Y, and Vo D T, *Gated neural networks for targeted sentiment analysis* in Proc. AAAI Conf. Artif. Intell., 2016, pp. 3087–3093
- [17] Tang D, Qin B, and Liu T, *Aspect level sentiment classification with deep memory network*. Proc. Conf. Empirical Methods Natural Lang. Process. 2016, pp. 214_224.
- [18] Tang D, Qin B, Feng X, and Liu T. *Effective LSTMs for target-dependent sentiment classification* in Proc. 26th Int. Conf. Comput.Linguistics, Tech. Papers (COLING), 2016, pp. 3298_3307.
- [19] Ren Y, Zhang Y, Zhang M, and Ji D, *Context-sensitive Twitter sentiment classification using neural network*. In Proc. AAAI, Phoenix, AZ, USA, Feb. 2016, pp. 215_221.
- [20] Rosenthal S, Farra N, and Nakov P, SemEval-2017 task 4: *Sentiment analysis in Twitter*. Proc. SemEval, Vancouver, BC, Canada, Aug. 2017, pp. 502_518.
- [21] Fan F, Feng Y, and Zhao D, *Multi-grained attention network for aspect-level sentiment classification*. In Proc. Conf. Empirical Methods Natural Lang. Process., 2018, pp. 3422_3433.
- [22] Zhang Q, Lu R, *A multi-attention network for aspect-level sentiment analysis*. Future Internet, vol. 11, no. 7, p. 157, Jul. 2019.
- [23] Xu Q, Zhu L, Dai T, and Yan C, *Aspect-based sentiment classification with multi-attention network*. Neurocomputing, vol. 388, pp. 135_143, May 2020.
- [24] Meng W, Wei Y, Liu P, Zhu Z, and Yin H, *Aspect based sentiment analysis with feature enhanced attention CNN-BiLSTM*. IEEE Access, vol. 7, pp. 167240–167249, 2019.
- [25] Park H J, Song M, and Shin K S, *Deep learning models and datasets for aspect term sentiment classification: Implementing holistic recurrent attention on target-dependent memories*. Knowl.-Based Syst. vol.187, Jan. 2020, Art. No. - 104825.

- [26] Lin Z, Feng M, Santos CN D, Yu M, Xiang B, Zhou B, Bengio Y, *A structured self-attentive sentence embedding*. arXiv preprint 2017, arXiv: 1703.03130.
- [27] Chen H, Sun M, Tu C, Lin Y, Liu Z, *Neural sentiment classification with user and product attention*. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 21 September 2016; pp. 1650–1659.
- [28] Fu, X, Yang, J, Li J, Fang M, Wang H, *Lexicon-enhanced LSTM with attention for general sentiment analysis*. IEEE Access 2018, 6, 71884–71891.
- [29] Dou, Z Y, *Capturing user and product Information for document level sentiment analysis with deep memory network*. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 3 March 2017; pp. 521–526.
- [30] Ranjan R, Daniel A K, *A Deep Learning Model for Extracting Consumer Sentiments using Recurrent Neural Network Techniques*, Int. Jour. Of Com. Sci and Net. Sec., Vol. 21 No. 8 pp. 238-246, 2021
- [31] Yang Z, Yang D, Dyer C, He X, Smola A., and Hovy E., *Hierarchical attention networks for document classification*, in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol., 2016, pp. 1480–1489.
- [32] Chen H, Sun M, Tu C, Lin Y, and Liu Z, *Neural sentiment classification with user and product attention*, in Proc. Conf. Empirical Methods Natural Lang. Process., 2016, pp. 1650–1659.
- [33] Zhang Z, Zou Y, and Gan C, *Textual sentiment analysis via three different attention convolutional neural networks and cross-modality consistent regression*, Neurocomputing, vol. 275, pp. 1407–1415, Jan. 2018.
- [34] Deng D, Jing L, Yu J, and Su S, *Sparse self-attention LSTM for Sentiment lexicon construction*, IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 27, no. 11, pp. 1777–1790, Nov. 2019.
- [35] Ahmad H., Asghar M. U., Asghar M. Khan Z., A. and Mosavi A. H., *A Hybrid Deep Learning Technique for Personality Trait Classification From Text*, in IEEE Access, vol. 9, pp. 146214-146232, 2021, doi: 10.1109/ACCESS.2021.3121791.
- [36] Cheng J., Sadiq M., Kalugina O. A., Nafees S. A. and Umer Q., *Convolutional Neural Network Based Approval Prediction of Enhancement Reports*, in IEEE Access, vol. 9, pp. 122412-122424, 2021, doi: 10.1109/ACCESS.2021.3108624.
- [37] Wang C., Yang X. and Ding L., *Deep Learning Sentiment Classification Based on Weak Tagging Information*, in IEEE Access, vol. 9, pp. 66509-66518, 2021, doi: 10.1109/ACCESS.2021.3077059.
- [38] F. Xu, Z. Pan, and R. Xia, *E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework*, Inf. Process. Manage. vol. 57, no. 5, Sep. 2020, Art. no. 102221.
- [39] Preethi, G.; Krishna, P. V.; Obaidat, M. S.; Saritha, V.; Yenduri, S. ,2017: *Application of deep learning to sentiment analysis for recommender system on cloud*. International Conference on Computer, Information and Telecommunication Systems, pp. 93–97.
- [40] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, *Deep contextualized word representations*, J. Assoc. Comput. Linguistics, vol. 1, pp. 2227_2237, Mar. 2018.
- [41] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, *Gradient-based learning applied to document recognition*, Proc. IEEE, vol. 86, no. 11, pp. 2278_2324, Nov. 1998.
- [42] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, *Sentiment analysis of comment texts based on BiLSTM*, IEEE Access, vol. 7, pp. 51522_51532, 2019.
- [44] F. Y. Zhou, L. P. Jin, and J. Dong, *Review of convolutional neural network*, Chin. J. Comput., vol. 1, pp. 35_38, Jan. 2017.
- [45] M. Rhanou, M. Mikram, S. Yousfi and S. Barzali, *A CNN-BiLSTM Model for Document-Level Sentiment Analysis*, Mach. Learn. Knowl. Extr. 2019, 1, 832–847; doi:10.3390/make1030048
- [46] Zi-xian Liu, De-gan Zhang, Gu-zhao Luo, Ming Lian, Bing Liu., *A new method of emotional analysis based on CNN-BiLSTM hybrid neural network*, Cluster Computing, 2019, https://doi.org/10.1007/s10586-020-03055-9
- [47] JUN XIE, BO CHEN, XINGLONG GU, FENGMEI LIANG, AND XINYING XU, *Self-Attention-Based BiLSTM Model for Short Text Fine-Grained Sentiment Classification 56789,-(.volV)*, IEEE Access, 2019, doi 10.1109/ACCESS.2019.2957510
- [48] FEI LONG, KAI ZHOU, AND WEIHUA OU, *Sentiment Analysis of Text Based on Bidirectional LSTM With Multi-Head Attention*, 2019, IEEE Access, doi 10.1109/ACCESS.2019.2942614
- [49] M. Ghorbani, M. Bahaghighat, Q. Xin, and F. Özen, *ConvLSTMConv network: A deep learning approach for sentiment analysis in cloud computing*, J. Cloud Comput., vol. 9, no. 1, pp. 9_16, Dec. 2020, doi: 10.1186/s13677-020-00162-1.
- [50] Chuantao Wang , Xuexin Yang , and Linkai Ding, *Deep Learning Sentiment Classification Based on Weak Tagging Information*, 2021, IEEE Access, doi 10.1109/ACCESS.2021.3077059

Personality Identification from Social Media Using Ensemble BERT and RoBERTa

Eggi Farkhan Tsani¹, Derwin Suhartono*²

¹ Computer Science Department, Binus Graduate Program – Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia

² Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia

*Corresponding author

E-mail: eggi.tsani@binus.ac.id, dsuhartono@binus.edu

Keywords: social media, big five personality, data augmentation, BERT, RoBERTa

Received: March 30, 2023

Social media growth was fast because many people used it to express their feelings, share information, and interact with others. With the growth of social media, many researchers are interested in using social media data to conduct research about personality identification. The identification result can be used as a parameter to screen candidate attitudes in the company's recruitment process. Some approaches were used for research about personality; one of the most popular is the Big Five Personality Model. In this research, an ensemble model between BERT and RoBERTa was introduced for personality prediction from the Twitter and Youtube datasets. The data augmentation method also introduces to handling the imbalance class for each dataset. Pre-trained model BERT and RoBERTa was used as the feature extraction method and modeling process. To predict each trait in the Big Five Personality, the voting ensemble from BERT and RoBERTa achieved an average f1 score 0,730 for Twitter dataset and 0,741 for Youtube dataset. Using the proposed model, we conclude that data augmentation can increase average performance compared to the model without data augmentation process.

Povzetek: Članek uvaja model združevanja sistemov BERT in RoBERTa za napovedovanje osebnosti iz podatkov Twitterja (X) in Youtube, z izboljšanjem s pomočjo podatkovne augmentacije.

1 Introduction

Based on Leadership IQ's study [1] of 20.000 companies, 46% of new employees resign from their jobs within one and a half years, and 89% of their failures are because of attitudinal reasons. The recruitment process and high turnover because of resignations can be incur high costs for a company. Curriculum vitae screening and face-to-face interviews were not enough to make sure the candidate had a good attitude. One of the approaches for getting a candidate's attitude was to do personality identification. This identification can also use to determine which position of the job was particularly fit for the candidate [2].

Social media has grown so fast around the world. Currently, many people can use social media not only for communication but also to express their thoughts, expectations, and feelings [3]. In January 2021, datareportal survey noted that social media users in Indonesia reached 170 million or 61,8% of the total population [4]. It means more than half of the population in Indonesia uses social media in their daily activities.

Because users use social media to express their feelings, the researchers can use social media data to conduct research about personality prediction. Different approaches were introduced to predicting personalities such as Big Five Personality, MBTI (Myers-Briggs Type Indicator), and DISC (Dominance Influence Steadiness Conscientiousness) [5]. From these three approaches mentioned above, Big Five Personality is the most accepted model to describe personality structure and divide it into personal and group [6]. Table 1 below describes the advantages of the Big Five Personality approach compared to the other two. The Big Five Personality consists of five personality traits that are usually called OCEAN (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) [7].

Table 1: Personality approach comparison

	Big Five	MBTI	DISC
Results	Unique	16 personalities	12 profiles
Predictive	Yes	No	No
Valid	Yes	Yes	No
Reliable	Yes	Yes	No

This research uses two social media datasets to build a prediction model. The first dataset is Twitter data, which consists of 508 users with around 46.000 posts collected manually, and the second dataset is Youtube data, consists of 10.000 clips extracted from 3.000 different videos of people speaking in English to the camera. This dataset is called the First Impression dataset and was downloaded from ChaLearn [8]. Both datasets are based on text and have multi-label cases. The prediction model was built using an ensemble BERT (Bidirectional Encoder Representation from Transformers) and RoBERTa (Robustly Optimized BERT Pretraining Approach) classifier for the personality prediction case.

Compared to other personality prediction research, this research uses a different approach to increase classification performance. Data augmentation using the back translation method was introduced to increase the number of datasets and handle the imbalance class. As a result, classification performance increase around 3-5% compared to classification using the original dataset.

2 Related works

Research about personality prediction has been done previously using various social media data. Facebook, Twitter, and Youtube were three popular social media that were used for research about personality prediction. The dataset available for research either in public or private was collected and labeled to Big Five traits manually.

Research conducted by [9] uses a Facebook dataset called myPersonality consists of 250 users with around 10.000 statuses labeled with the Big Five traits label. Their research used LIWC (Linguistic Inquiry and Word Count), and SPLICE (Structured Programming for Linguistic Cue Extraction) features then used MLP (Multi-Layer Perceptron) classifier to produce 70,78% average accuracy. Another research that uses myPersonality dataset was conducted by [10] using LIWC, SPLICE, and SNA (Social Network Analysis) as feature extraction methods and XGBoost algorithm to achieve the best result with 74,2% average accuracy.

Another personality prediction research that uses social media datasets was also conducted by [11]. The dataset used was Twitter in Bahasa which consists of Twitter posts from 250 users labeled with Big Five traits label. Their research uses SGD (Stochastic Gradient Descent), XGBoost, and super learner to produce a good ROC-AUC (Receiver Operating Characteristic and Area Under Curve) score. Research conducted by [12] also uses Twitter dataset with more data that consists of tweets from 508 users. This research use word-n-gram and

Twitter metadata to process using Random Forest classifier to produce 0,744 f1 scores on average. Research using Twitter dataset was also conducted by [13] using similar data to previous research. This research uses pre-trained models BERT, RoBERTa, and XLNet combined with TF-IGM statistical features. These methods used an averaging model to make better predictions result.

Popular social media dataset also used for research conducted by [14]. Their research uses Youtube vlog dataset which consists of 404 vlogs with audio-video features and transcripts. Decision tree and SVM algorithm were used and produced better results over baseline average performance. Another research using Youtube dataset was also conducted by [15]. This research uses Youtube translations only to create a model using Word2Vec, GloVe, and BERT as feature extraction methods then uses SVM and SVR for classification. This approach produces 0,612 f1 scores as the best prediction result.

Table 2: Previous personality research

Author	Dataset	Classifier	Findings
Tandera et al., 2017	Facebook	MLP	Accuracy 70,78%
Tadesse et al., 2018	Facebook	XGBoost	Accuracy 74,2%
Adi et al., 2018	Twitter	Super Learner	ROC AUC 0,992
Jeremy et al., 2019	Twitter	Random Forest	F1 score 0,744
Christian et al., 2021	Twitter	BERT, RoBERTa, XLNet	F1 score 0,757
Farnadi et al., 2016	Youtube	Decision Tree and SVM	RMSE 0,115
Lopez-Pabon et al., 2022	Youtube	SVM and SVR	F1 score 0,612

Previous research that uses the Big Five Personality approach processes the dataset with imbalance class. For optimizing performance, we modify the data with data augmentation, whereas the original dataset will add modified data so imbalance class can be minimized. With minimizing imbalance class in the data, dataset quality will increase, and modeling process will have a better performance.

The advantage of the ensemble model also reflects from previous research above. For Facebook, Twitter, and Youtube datasets, the best performance resulted from the ensemble model. The best model from Facebook

dataset resulted from ensemble boosting using XGBoost, Twitter dataset resulted from ensemble averaging using BERT, RoBERTa, and XLNet classifiers, and Youtube dataset resulted from SVM and SVR. With this consideration above, this research uses data

augmentation to produce better dataset quality and ensemble model to increase classification performance.

3 Methodology

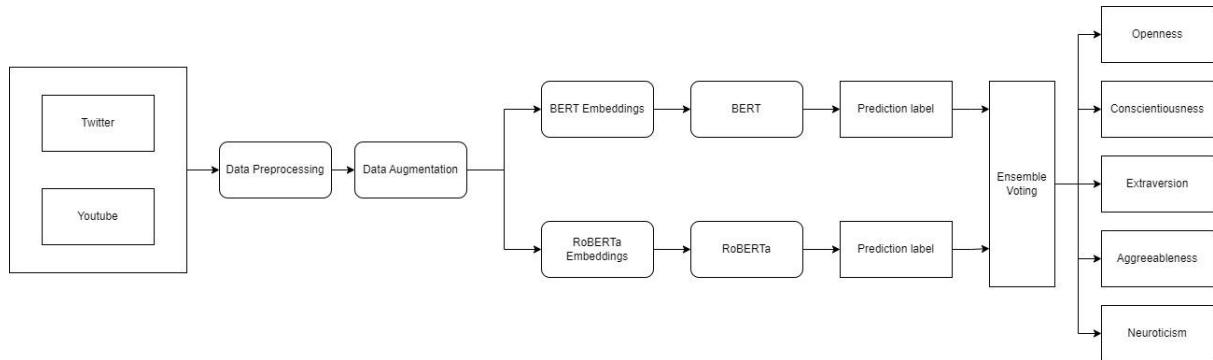


Figure 1: Architecture model

Using Twitter and Youtube as social media data, this research was composed of three phases: data collection, development, and evaluation. The details for each phase can be seen in figure 1. For the initial phase, data collected from previous research [9, 11, 12] has been collected for Twitter dataset. The label for defining the Big Five traits has been annotated by physiological experts. On the other hand, Youtube dataset was an open-source dataset from ChaLearn [7]. Before processing in the development phase, both datasets were preprocessed and augmented to get better-quality datasets.

For the development phase, pre-trained models BERT and RoBERTa were used as embedding and classification methods to produce prediction labels based on the Big Five personality traits. Predicted labels or results from each classifier were ensembled using the voting method to generate final personality labels. The ensemble method was used because ensembles can produce better predictive performance by combining multiple models [16]. After getting the prediction label, the confusion matrix was used as an evaluation metric during the evaluation phase.

3.1 Dataset

This research uses two social media datasets for the experiment. The first dataset is Twitter which was collected manually and consisted of 508 users with around 46.000 Twitter posts in Bahasa. The second dataset, first impression Youtube, consists of 10.000 short videos with text-based english transcripts. This dataset is public and downloaded from ChaLearn [7]. Both datasets are already labeled with Big Five Personality traits so they can be processed with the supervised learning method [17]. Text-

based processing was applied in this research ignoring other types like video and audio. Dataset distribution for Twitter and Youtube were described in table 3 and 4.

Tabel 3: Twitter dataset distribution

values	O	C	E	A	N
High	27.921	14.365	36.187	27.572	22.871
Low	22.191	35.747	13.925	22.540	27.241

Tabel 4: Youtube dataset distribution

values	O	C	E	A	N
High	6.617	5.615	4.407	6.553	5.509
Low	3.281	4.283	5.491	3.345	4.389

3.2 Data preprocessing

Today’s real-world data are highly too noisy, lost, and unsteady [18]. Because of this reason, we need to conduct data preprocessing before modeling the data. Data preprocessing was conducted to remove noise, missing values, and inconsistent data [19]. Data preprocessing consists of several steps such as data cleaning, transformation, and reduction.

Steps to perform preprocessing data in this research are:

1. Remove URL
2. Remove the symbol
3. Translate Bahasa into English Language
4. Converting letters into lowercase

5. Remove stop words
6. Lemmatization

3.3 Data augmentation

From dataset distributions in table 3 and table 4, it can conclude that both datasets have an imbalanced class. In Twitter dataset, the Conscientiousness label has 14.365 high class and 35.747 low class while the Extraversion label has 36.187 high class and 13.925 low class. So, Twitter dataset has an imbalance class in Conscientiousness and Extraversion labels. In Youtube dataset, the Openness label has 6.617 high class and 3.281 low class while the Agreeableness label has 6.553 high class and 3.345 low class. So, Youtube dataset has an imbalance class in Openness and Agreeableness labels.

In this research, the imbalance class can be handled by implementing data augmentation using the back-translation method. This method translates Twitter posts and youtube transcripts from English to Germany and then translates them back to English using T5 (Text to Text Transfer Transformer) model. The T5 model was chosen for language translation because this model can generate good paraphrases from the original language [20]. The translation was used for getting different sentences that are paraphrased from the original data. So, the additional data generated from data augmentation have paraphrased sentences to get better modeling process in the transformer classifier.

3.4 Features

In this research, a pre-trained model for feature extraction was used before modeling with ensemble BERT and RoBERTa. Because embedding was processed using BERT, the feature generated by token embedding, segment embedding, and positional embedding which part of BERT embedding process. This approach was designed to do modeling of two-way representation from left to right and right to left to get the context of the sentence.

Token embedding processed social media status concatenates with special token called classification [CLS] and separator [SEP]. The CLS token was inserted at the beginning of the sentence, and the SEP token inserted at the end of the sentence [21]. The aim of this step is to get input representation for the classification task and separate each sentence. Each word in the sentence was tokenized and mapped to corpus dimension size. Each sentence consists of 12 token representations with 768 fixed dimensions [22]. The second embedding layer used in this feature extraction is segment embedding. This embedding layer was designed to create vector

representation of a sentence. If the input is only one sentence, then the segment embedded only the corresponding vector with index zero. The third embedding layer is positional embedding. This embedding layer was designed as a lookup table representing the number of long sentences. Each row of the table was a position of vector representation of the word.

Similar to BERT, RoBERTa uses token, segment, and positional embedding for extracting features. RoBERTa provides improvement from BERT because RoBERTa uses dynamic masking patterns instead of static masking and separates the segments with separation token $\langle /s \rangle$.

3.5 Model prediction

Deep learning with the transformer model was a popular method for creating the personality prediction system [23]. In transformer, each identical layer in the encoder first computes multi-headed attention between a given token and then run position to the feed-forward network [24]. The latest research was to build a model for personality prediction using transformer and produce a good result. Based on that success, this research used two transformer classifiers BERT and RoBERTa combined with an ensemble method to predict personality traits.

Input resulted from embedding processed using each classifier BERT and RoBERTa. Both classifiers use 16 batch sizes for Twitter dataset and 32 batch sizes for Youtube dataset. We use Adam optimizer with learning rate $1e^{-5}$ because the performance produces better on learning rate $1e^{-5}$. For epochs and loss function, we use 10 epochs with saving the best model to get the best performance from every epoch and binary cross entropy with logit loss which uses sigmoid activation function.

After getting the predicted class from each classifier, we use voting ensemble to produce an average combination to decide the final label for each trait. The final label will be evaluated using a confusion matrix to produce an f1 score as the evaluation result.

3.6 Evaluation metric

Classification system performance describes how good the system classified the data. The confusion matrix is one of the methods that can be used to measure the classification system performance [25]. Basically, the confusion matrix contains information that compares the results of the classification performed by the system with the predicted result.

In this research, the performance measure as a prediction result was f1 score because of imbalanced data

on the dataset. F1 score obtained from the confusion matrix with combining precision and recall formula.

4 Result

In the experiment result, f1 score performance metric was shown for each trait and average. It was shown for each model that we used in this research for Twitter and Youtube datasets.

Tabel 5: Experiment result using Twitter dataset

	M1 ^a	M2 ^b	M3 ^c	M4 ^d	M5 ^e	M6 ^f
O	0,672	0,655	0,671	0,645	0,705	0,701
C	0,500	0,443	0,721	0,704	0,509	0,724
E	0,813	0,827	0,734	0,759	0,849	0,793
A	0,705	0,671	0,809	0,803	0,709	0,826
N	0,610	0,563	0,570	0,526	0,641	0,605
Avg	0,660	0,632	0,701	0,687	0,683	0,730

- ^aM1 represent BERT model
- ^bM2 represent RoBERTa model
- ^cM3 represent BERT model with augmentation
- ^dM4 represent RoBERTa model with augmentation
- ^eM5 represent BERT + RoBERTa model
- ^fM6 represent Proposed model

Table 5 shows all scenario results including classification using one classifier, voting ensemble from two classifiers, and data augmentation for Twitter dataset. The table shows that the proposed model produced better results than BERT or RoBERTa on average. The highest result for each trait produces by ensemble BERT and RoBERTa without data augmentation and proposed models. Openness, Extraversion, and Neuroticism traits produce the best results from ensemble BERT and RoBERTa. For Conscientiousness and Agreeableness traits produce the best results from the proposed model.

From the experiment result, we can conclude that data augmentation on the proposed model produces balanced f1 score for each trait, so it can produce better performance on average results with 0,730 compared to other models. This result is around 5% higher than the model without data augmentation process. Meanwhile, the highest f1 score result was produced by the ensemble BERT and RoBERTa model for the Extraversion trait with 0,849 f1 scores.

Tabel 6: Experiment result using Youtube dataset

	M1 ^a	M2 ^b	M3 ^c	M4 ^d	M5 ^e	M6 ^f
O	0,735	0,800	0,748	0,787	0,790	0,786
C	0,649	0,693	0,599	0,695	0,721	0,713
E	0,493	0,406	0,687	0,700	0,573	0,717
A	0,744	0,791	0,731	0,793	0,788	0,793
N	0,624	0,681	0,601	0,679	0,707	0,697
Avg	0,649	0,674	0,673	0,731	0,716	0,741

- ^aM1 represent BERT model
- ^bM2 represent RoBERTa model
- ^cM3 represent BERT model with augmentation
- ^dM4 represent RoBERTa model with augmentation
- ^eM5 represent BERT + RoBERTa model
- ^fM6 represent Proposed model

Meanwhile, table 6 shows experiment results for Youtube dataset. As shown in the table above, the result for the proposed model can outperform the result from BERT or RoBERTa on average. The highest f1 score result varies for each trait. Extraversion and Agreeableness traits produce the best results from the proposed model with 0,717 and 0,793 f1 scores, respectively. Conscientiousness and Neuroticism traits produce the best results from ensemble BERT and RoBERTa models with 0,721 and 0,707 f1 scores, respectively. For the Openness trait, it produces the best results from RoBERTa classifier with 0,800 f1 scores.

Similar to Twitter, Youtube dataset also concludes that data augmentation on the proposed model produces a balanced f1 score for each trait and produces better average performance with 0,741 compared to other models. This result is around 3% higher than the model without the data augmentation process. Meanwhile, the highest performance result was produced by RoBERTa model for the Openness trait with 0,800 f1 scores.

5 Discussion

This research uses two datasets, which are Twitter and Youtube. For Twitter dataset, this research achieves an average f1 score 0,730. Although this result is still below the previous result [13], this research showed that ensemble using only two classifiers and modified dataset using the data augmentation method can provide a good f1 score. While for Youtube dataset, this research achieves an average f1 score 0,741. Compared to previous results that used Youtube dataset also, this result provided a good result and was better than the research done by [15] that

resulted best f1 score 0,612. One of the reasons for the result is this research modified the dataset to minimize imbalance class with the data augmentation method. Using data augmentation, the result of the f1 score increase about 3% for Youtube dataset and 5% for Twitter dataset.

6 Conclusion

This research shows that personality prediction using text data from social media can produce good results. Although two datasets used in this research have an imbalanced class, they can be fixed with data augmentation using the back translation method. A result from the experiment shows that the proposed model with ensemble BERT and RoBERTa as feature extraction and pre-trained model, back translation as data augmentation method can produce 0,730 average f1 scores for Twitter dataset and 0,741 average f1 scores for Youtube dataset. The back translation method using T5 increase the average f1 score performance on both datasets compared to the processing dataset without data augmentation.

For future development, the dataset used for personality prediction should have a balance class. If the dataset already balances, we don't need to have more processing time to do data augmentation. Besides that, another pre-trained model like ALBERT can be used to reduce the memory and training speed of BERT and RoBERTa [26].

Acknowledgement

We would like to thank Bina Nusantara University for supporting and assisting this research. Also thank to our colleagues who provide insight and contributions for this paper.

References

- [1] Sprockets, "The Importance of Recruiting for Personality: Everything You Need to Know," <https://sprockets.ai/the-importance-of-recruiting-for-personality>.
- [2] R. Shilpa, V. Supriya, P. Sweta, V. R. Vinaya, and S. S. Uday, "Personality prediction using Machine Learning," *Science and Engineering Journal*, vol. 25, no.7, pp. 46-53, 2021. <https://saejournal.com/wp-content/uploads/2021/07/Personality-Prediction-Using-Machine-Learning.pdf>
- [3] R. Valanarasu, "Comparative analysis for personality prediction by digital footprints in social media," *Journal of Information Technology and Digital World*, vol. 3, no. 02, pp. 77-91, 2021. <https://doi.org/10.36548/jitdw.2021.2.002>
- [4] Datareportal, "Digital 2021: Indonesia," <https://datareportal.com/reports/digital-2021-indonesia>.
- [5] N. A. Utami, W. Maharani, and I. Atastina, "Personality classification of Facebook users according to Big Five Personality using SVM (Support Vector Machine) method," *Procedia Computer Science*, vol. 179, pp. 177-184, 2021. <https://doi.org/10.1016/j.procs.2020.12.023>
- [6] N. Abood, "Big five traits: A critical review," *Gajah Mada International Journal of Business*, vol. 21, no. 2, pp. 159-186, 2019. <https://doi.org/10.22146/gamaijb.34931>
- [7] S. Basaran, and O. H. Ejimogu, "A Neural Network Approach for Predicting Personality from Facebook Data," *Sage Journals*, vol. 11, no. 3, 2021. <https://doi.org/10.1177/21582440211032156>
- [8] Computer Vision Center and University of Barcelona, "ChaLearn Looking at People," <https://chalearnlap.cvc.uab.cat/dataset/24/description/>
- [9] T. Tandra, D. Suhartono, R. Wongso, and Y. L. Prasetyo, "Personality Prediction System from Facebook Users," *Procedia Computer Science*, vol. 116, pp. 604-611, 2017. <https://doi.org/10.1016/j.procs.2017.10.016>
- [10] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Personality predictions based on user behavior on the Facebook social media platform," *IEEE Access*, vol. 6, pp. 61959-61969, 2018. <https://doi.org/10.1109/ACCESS.2018.2876502>
- [11] G. Y. Adi, M. H. Tandio, V. Ong, and D. Suhartono, "Optimization for Automatic Personality Recognition on Twitter in Bahasa Indonesia," *Procedia Computer Science*, vol. 135, pp. 473-480, 2018. <https://doi.org/10.1016/j.procs.2018.08.199>
- [12] N. H. Jeremy, C. Prasetyo, and D. Suhartono, "Identifying Personality Traits for Indonesian User from Twitter Dataset," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 4, pp. 283-289, 2019. <https://doi.org/10.5391/IJFIS.2019.19.4.283>
- [13] H. Christian, D. Suhartono, A. Chowanda, and K. Z. Zamli, "Text based personality prediction from

- multiple social media data sources using pre-trained language model and model averaging," *Journal of Big Data*, vol. 8, no. 68, 2021. <https://doi.org/10.1186/s40537-021-00459-1>
- [14] G. Farnadi, G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M. Moens, and M. D. Cock, "Computational personality recognition in social media," *Springer Science and Business Media LLC*, vol. 26, pp. 109-142, 2016. <https://doi.org/10.1007/s11257-016-9171-0>
- [15] F. O. Lopez-Pabon, and J. R. Orozco-Arroyave, "Automatic Personality Evaluation from Transliterations of Youtube Vlogs using Classical and State-of-the-Art Word Embeddings," *Ingeniería e Investigación*, vol. 42, no. 2, 2022. <https://doi.org/10.15446/ing.investig.93803>
- [16] Machine Learning Mastery, "A Gentle Introduction to Ensemble Learning Algorithms," 2021.
- [17] H. Zheng, and C. Wu, "Predicting personality using facebook status based on semi-supervised learning," in *Proceedings of the 2019 11th International Conference on Machine Learning and Computing*, 2019, pp. 59-64. <https://doi.org/10.1145/3318299.3318363>
- [18] A. Souri, S. Hosseinpor, and A. M. Rahmani, "Personality classification based on profiles of social network users and the five-factor model of personality," *Human-centric Computing and Information Sciences*, vol. 8, no. 24, 2018. <https://doi.org/10.1186/s13673-018-0147-4>
- [19] M. U. Maheswari, and J. G. R. Sathiaselan, "Text Mining; Survey on Techniques and Applications," *International Journal of Science and Research*, vol. 6, no. 6, pp. 1660-1664, 2017. https://www.ijsr.net/get_abstract.php?paper_id=ART20174656
- [20] Towards Data Science, "Paraphrase any question with T5 (Text-To-Text Transfer Transformer) - Pretrained model and training script provided," <https://towardsdatascience.com/paraphrase-any-question-with-t5-text-to-text-transfer-transformer-pretrained-model-and-cbb9e35f1555>.
- [21] P. Kaur, G. S. Kohli, and J. Bedi. "Classification of Health-Related Tweets Using Ensemble, Zero-Shot and Fine-Tuned Language Model," in *Proceedings of the 29th International Conference on Computational Linguistics*, 2022, pp. 138-142.
- [22] J. Liu, C. Xia, X. Li, H. Yan, and T. Liu, "A BERT-based Ensemble Model for Chinese News Topic Prediction," *ACM Digital Library*, pp. 18-23, 2020. <https://doi.org/10.1145/3404512.3404524>
- [23] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *31st Conference on Neural Information Processing Systems*, 2017. <https://arxiv.org/abs/1706.03762>
- [24] M. E. Peters, M. Neumann, L. Zettlemoyer, and W. T. Yih, "Dissecting contextual word embeddings: Architecture and representation," in *Proceedings of the 2018 Conference of Empirical Method in Natural Language Processings*, 2020, pp. 1499-1509. <https://doi.org/10.18653/v1/d18-1179>
- [25] A. Kulkarni, D. Chong, and F. A. Batarseh, Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering, pp. 83-106, 2020. <https://doi.org/10.1016/B978-0-12-818366-3.00005-8>
- [26] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "Albert: A lite bert for self-supervised learning of language representations," *arXiv preprint arXiv*, 2019. <https://arxiv.org/abs/1909.11942>

Hybrid Compression Algorithm for Energy Efficient Image Transmission in Wireless Sensor Networks Using SVD-RLE in Voluminous Data Applications

G. Sudha, C. Tharini*

Department of Electronics and Communication Engineering, B.S. Abdur Rahman Crescent Institute of Science & Technology, Vandalur, Chennai, India

E-mail: sudhaganesh74@gmail.com, tharini@crescent.education

* Corresponding author

Keywords: singular value decomposition, dimensionality reduction, threshold, rank matrices, compression ratio

Received: November 21, 2021

WSNs are used in different applications and the enormous volume of data they collect and broadcast across the network overburdens the sensor nodes and this issue can be mitigated by compressing the data before transmitting it over the network. Singular Value Decomposition, a state-of-the-art non-transform-based compression method, primarily for dimensionality reduction in any type of data, is utilized in this study. In this, the difference between the adjacent pixel values of the captured images by WSNs are computed as a preprocessing step, and then compressed, with the compressed data represented by three singular matrices: two orthonormal matrices (X , Y), and one diagonal matrix (Σ), called rank matrix. The resultant data is then applied through a Run Length Encoding step and transmitted. By compressing the image with different thresholds, the rank value of SVD is altered and since the pixel differences which is a relatively small number of bits are only encoded, the outcome is represented with a compression ratio of approximately 12% and also the reconstructed image at the receiver exhibits good Peak Signal to Noise Ratio (PSNR). The use of this strategy in WSNs is also justified by analyzing the amount of energy savings and the nodes' energy usage using standard energy models and the percentage of energy saving varies from 25% to 53% with the decrease in the rank values respectively.

Povzetek: Študija predstavlja hibridni algoritem kompresije SVD-RLE za energetsko učinkovit prenos slik v omrežjih z brezžičnimi senzorji, pri čemer je prihranek energije do 53%.

1 Introduction

Remote monitoring like habitat monitoring, structural health monitoring, traffic surveillance, etc., are the utilization scenarios of Wireless Sensor Networks (WSN). These applications require continuous monitoring and generate huge volume of data. WSN has a number of sensor nodes to perform this operation and they generate the data from the source and transmit towards the sink through a cluster of intermediate nodes as shown in Figure 1. If these voluminous data is transmitted as a raw data, it places burden on the nodes and consume more power which in turn depletes the nodes of its energy. Instead, if the data generated is compressed using an appropriate compression algorithm and the compressed data is then transmitted, the burden on the nodes are reduced, thereby increasing the lifetime of the nodes.

In this approach, a hybrid combination of two state-of-the-art algorithms Singular Value Decomposition (SVD) and Run Length Encoding (RLE) is proposed. SVD represents the entire image data in the form of three matrices: two orthonormal matrices and one rank matrix which are scaling matrices of positive values, for transmission.

The main advantage of SVD is that it can be applied to images of any size instead of equal dimensions in both x and y axes as compared to DCT, DWT, etc.

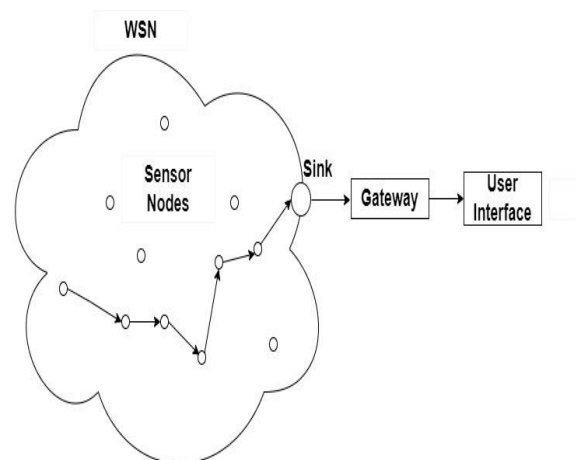


Figure 1: Classical architecture for wireless sensor networks.

Even though SVD significantly reduces the number of bits that must be transferred, RLE is employed to represent the compressed data in terms of lesser number of bits and as RLE is a lossless technique, it improves only the compression ratio and does not affect PSNR. A pre-processing step is introduced before compression and the performance of this algorithm is compared in two stages: before pre-processing and compression (stage I), and after pre-processing and compression (Stage II) with respect to rank values, PSNR, compression ratio along with other state of the art compression algorithms and the stage II performance is found to be more effective in terms of the number of bits sent across the network.

The literature offers a variety of compression methods, for various applications of WSN that involves a huge amount of data collection. These techniques also address the issues of how to improve the efficiency of the nodes coupled with compression, and based on the study, the method of combining both the compression techniques and energy efficiency improvement is the biggest challenging task. In [1], an innovative singular vector sparse reconstruction technique has been developed to improve the conventional Singular Value Decomposition (SVD) based compression technique by focusing on reconstruction based on sparse sampled singular vectors.

As the rank of SVD matrix plays a major role in determining the compression ratio of an image, an improvement was proposed in [2], by building a projection data matrix that spans the subspace of the original data matrix and random sampling of the column space. The proper low rank approximation was then obtained from the projection matrix by employing methods like oversampling and power iterations, and the same was used to compress images.

A technique for the retrieval of quality images were discussed in [3], that involves thresholding based SVD for removing the repetitive data which provides considerable space savings for data storage. A set of compression algorithms that are lossless and adaptive are discussed in [4] with a series of aggregation and routing strategies that shrinks the redundant data before transmission in WSNs.

A high coding efficiency was proposed in [5] that involve frequency tables that depend on adjustment of various parameters like range, step, mutual learning and table initialization. The process of how SVD is used to handle big data sets by identifying the details of pixels that contributes least to the actual image quality and by compressing them and at the same time restoring the actual image quality was discussed in [6]. During data capturing transmission in WSNs, the node topology in which the nodes are organized plays a pivotal role in the energy improvement of nodes and the concept was discussed in [7], with various techniques of arranging the nodes between the source and the sink.

In [8], a truncated SVD for alleviating the errors in outlier detection and to improve the signal quality was elaborated for WSNs within a network. A block partitioning method was employed in [9], by optimally choosing the Eigen values in SVD which can be used in varied applications. In [10], SVD was compared with

Non-negative Matrix Factorization method that showed consistent energy consumption performance with NMF, but a degraded image restoration quality when the block size increases.

A method of data reduction involving SVD technique was elaborated in [11], which proved to be efficient with varying rank values. In [12], [13] a novel code book designing technique was proposed to enhance the effectiveness of image compression using non-transform-based vector quantization and improved differential evolution with a minimal computational time. A comparative analysis was done in [14] involving SVD and Wavelet Difference Reduction (WDR) method for compressing the image and SVD shows better performance at high rank values and WDR shows better performance at high rank values and a trade-off was suggested. In [15], use of data aggregation for redundancy reduction and energy improvement was discussed.

Efficacy of SVD in image compression for compressing the images in wavelets was discussed in [16], by representing images with a very small number of dominant values and analysis of wavelets for various compression techniques was discussed in [17]. Table 1 represents a variety of standard State-of-the-Art lossy and lossless compression algorithms employed in WSNs for the comparison of the proposed method.

Table 1: Summary of related works and contributions.

Reference No.	Approaches	Methodology	Performance / Results
[3]	Singular Value Decomposition (SVD)	Lossy	A PSNR of around 20 dB for rank 50 and around 25 dB for rank 100 is obtained with SSIM of 0.8 and 0.6 respectively.
[18]	Set-Partitioning in Hierarchical Tress (SPIHT)	Lossy	Good reconstruction quality and long computation time as it involves DWT as a preprocessing step. Distributed compression provides energy savings of around 0.2 nJ.
[19]	Discrete Cosine Transform (DCT)	Lossy	A pruned approach is used that gives a PSNR of around 30 dB for standard image data set and an energy consumption of 2.52 μ J for a 8 x 8 block.
[20]	Joint Photographic Experts Group (JPEG)	Lossy	PSNR is 27 dB for standard image data set and the energy

			requirement is 30.67 J for adaptive JPEG.
[21]	Embedded Zerotree Wavelet (EZW)	Lossy	An enhanced EZW is proposed and the PSNR obtained is around 33 dB for the standard test image set compared with around 30 dB for standard EZW.
[22]	Huffman Coding, Run Length Encoding (RLE)	Lossless	As data is exactly retrieved after decompression, compression doesn't save storage space. Achieves a lower compression ratio than lossy techniques.

2 Methodology

The proposed methodology incorporates a hybrid compression technique involving SVD and RLE which is discussed as below:

2.1 SVD based image compression

Singular Value Decomposition (SVD) entails decomposing matrix Z into the form as in Equation (1).

$$Z = XZY^T \tag{1}$$

With the use of this computation, we are able to keep the crucial unique values that the image needs while letting go off the values that are not as crucial to maintaining the image's quality, where X and Y are orthogonal matrices of order $m \times r$ and $r \times n$ respectively, and Z is a diagonal matrix of order $r \times r$, that corresponds to the square roots of the eigenvalues of the matrix $Z^T Z$, that are normally arranged in terms of its magnitude in decreasing order, make up the singular values of a $m \times n$ matrix Z .

The diagonal matrix of SVD represents the rank matrix with singular values of the image on which SVD is applied with the rank values arranged in descending order [1]. A portion of the first few columns (r) of the singular values corresponding to the low frequency content of the image is retained and the remaining with small singular values are discarded for the purpose of compressing an image resulting in dimensionality reduction. Similarly, the X and Y matrices are also trimmed to match with the dimensions of the singular matrix that result in Equation (2).

$$Z_{m \times n} = X_{m \times r} Z_{r \times r} Y_{r \times n}^T \tag{2}$$

The major image content and its contour information are represented by the low frequency data, which has large singular values and also denotes the area where

the grey scale transitions of the image are slow. The high frequency information is represented as smaller singular values that denote a region with rapid variations in gray scale, which represents noise and the image's detailed information. SVD achieves compression by tossing out the singular vectors associated with small singular values that constitutes the image's finer details and hence results in reduced image quality after reconstruction. As the rank value decreases, compression ratio increase, but the image quality decreases. Hence the compression ratio must be limited to achieve significant image quality after reconstruction and this places a limitation in the performance of SVD on image compression and reconstruction.

Rank value in SVD represents the dimension of the non-zero singular matrix. By varying the threshold value, the rank of the matrix is varied. With higher rank value providing very high PSNR and low compression ratio, and lower rank value providing less PSNR and high compression ratio. Compression ratio impacts the number of bits that has to be transmitted through the network which affects the energy savings also. In our proposed method, the rank values of 408, 204 and 51 are considered and as a trade off the rank value is not reduced further so that PSNR, compression ratio and energy savings are maintained effectively. Also, since the preprocessing techniques reduce the magnitude of the pixels, here the trimming of the rank matrix is not needed and hence PSNR is maintained.

The limitation of reduced reconstructed image quality is overcome in our proposed methodology by taking the difference between the adjacent pixel values, then performing the SVD process, which results in the lower magnitude of the pixel values and due to this the compresses values are transmitted without the trimming process as depicted in the following steps. The proposed block diagram is depicted in Figure 2.

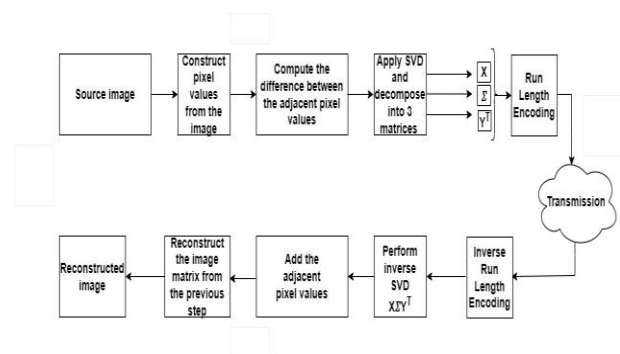


Figure 2: Proposed block diagram.

2.2 Run Length Encoding (RLE)

The output of the SVD process is then applied through RLE in which the continuous runs of zeros and ones are computed and the RLE output is transmitted. For example, if the SVD output is 000011110000000011111111, then instead of transmitting 24 bits, the data is transmitted as 04140818 and in terms of bits it requires only 20 bits and the transmitted data is 01000010011000010001 (as shown in

Figure 3). The reverse process is done at the receiver for reconstruction of bits. If the runs of data are very long, then more space can be saved during the RLE process. In compression of images, the runs of data are very long because of the interpixel redundancy. The pixel values are indicated by bits for wireless transmission and hence space savings is also more and this provides more compression and since RLE is lossless, this provides no information loss.

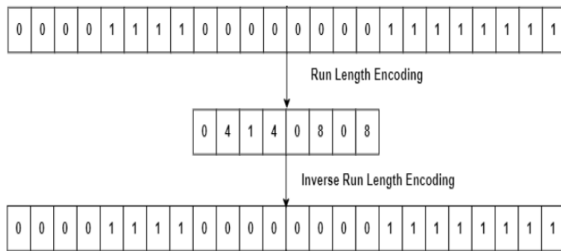


Figure 3: Illustration of RLE process.

The illustration of the proposed hybrid algorithm in the form of a flow chart is represented in Figure 4. With the procedure shown in Figure 4, due to the compression of the difference of the adjacent pixel difference values, there is reduction in the magnitude of the entries in the matrix and hence the need for the trimming of the rank matrices is reduced and hence the PSNR value obtained is considerably higher when compared to the actual SVD and at the same time, the compression ratio also significantly increases.

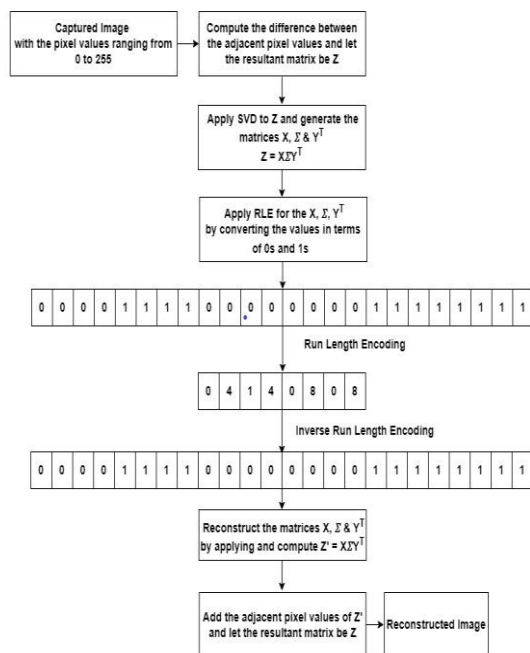


Figure 4: Flowchart of the proposed hybrid SVD – RLE.

3 Results and discussion

The application considered for this work is structural health monitoring of buildings and the below structural images of buildings as shown in Figure 5b are captured using the Raspberry Pi, equipped with a camera module Figure 5a and applied with the proposed SVD algorithm. As Raspberry Pi emulates a sensor node which is similar to its scanty processing ability, it can be chosen to run in Python environment.

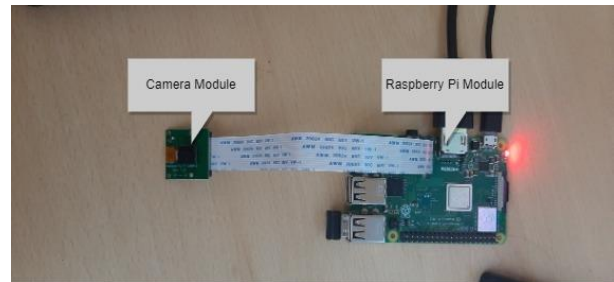


Figure 5a: Raspberry pi setup.

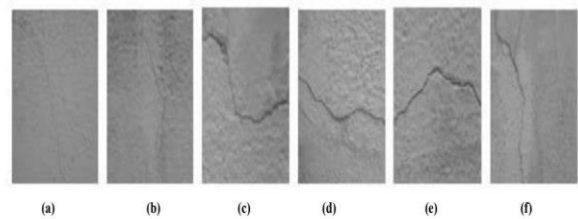


Figure 5b: Image data captured using Raspberry pi.

Consider the image in Figure 5b (b), the actual pixel values of the image and the pixel difference values are as shown in Figure 6 and Figure 7 respectively and the histogram of the pixel values is plotted in Figure 8 and Figure 9 before and after pre-processing respectively. The pre-processing technique of calculating the pixel differences reduces the magnitude of the pixel at the initial stage itself as shown in Figure 6 and 7. For example consider the first few pixels of the image 1 taken for consideration, the pixel values are 141, 139, 135, 128, ...etc before taking the pixel difference and 141, -2, -4, -7... etc after computing the pixel difference which shows that the magnitude of the pixels are drastically reduced before applying SVD. This in turn helps to reduce the compression ratio significantly around 50% with decreasing rank values, instead of applying SVD directly to the image. Also, as SVD is applied here after the pre-processing, there is no need to trim the singular matrices as already the magnitude of the pixels is reduced, which helps to recover the original image after reconstruction by adding the adjacent pixel values.

This is also illustrated by plotting the histogram of the pixel values as in Figure 8 and Figure 9 before and after pre-processing respectively. Most of the pixel values are centered around the value 140 before pre-processing and around 0 after pre-processing. And also, the maximum pixel value of the test image before and

after pre-processing as in Figure 8 and Figure 9 corresponds to 188 (1 occurrence) and 141 (1 occurrence) respectively, so that pixel values are shifted to the left in Figure 9.

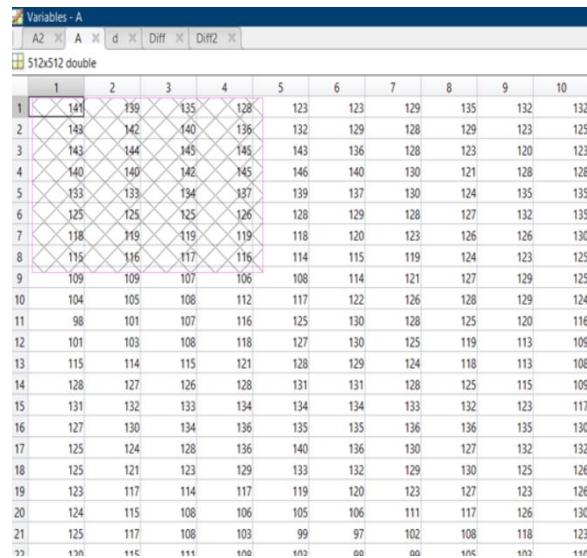


Figure 6: Pixel values of original image.

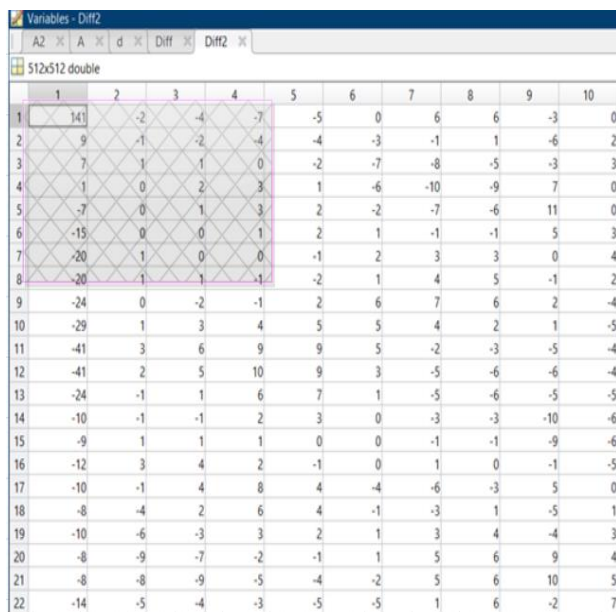


Figure 7: Pixel difference values of original image.

The pixel values obtained before and pre-processing are applied with the SVD process and it is revealed that the suggested algorithm produces small values in the rank matrix, and the other orthogonal matrices as illustrated in Figure 10 and Figure 11 for the actual pixel values and the pixel difference values respectively.

For example, the first entry of the rank matrices is 71053 and 852.6239 in Figure 10 and Figure 11 respectively and the values in Figure 11 reduces along the diagonal elements subsequently when compared to Figure 10. Because of the smaller values in the rank matrix and the other orthogonal matrices, the data is transmitted without trimming, which in turn results in significant PSNR after reconstruction.

The pixel values are applied with the SVD process with different thresholds, which in turn varies the rank values when applied with the SVD process. The high rank value corresponds to more information content and a low rank value provides less information content after the compression process and correspondingly the PSNR value will also decrease. Even if the PSNR value decreases with decrease in the rank value, that is sufficient for the interpretation of the reconstructed image because the rank values are transmitted as such without being trimmed but results in lesser number of bits to be transmitted as the difference values are only compressed.

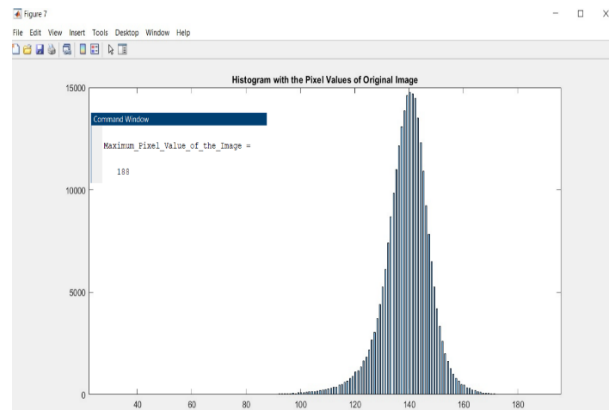


Figure 8: Histogram of pixel values of original image.

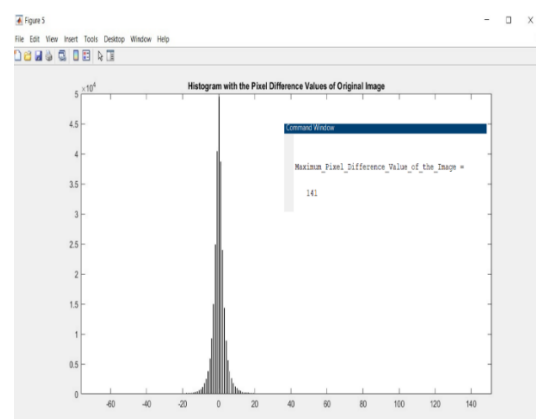


Figure 9: Histogram of pixel difference values.

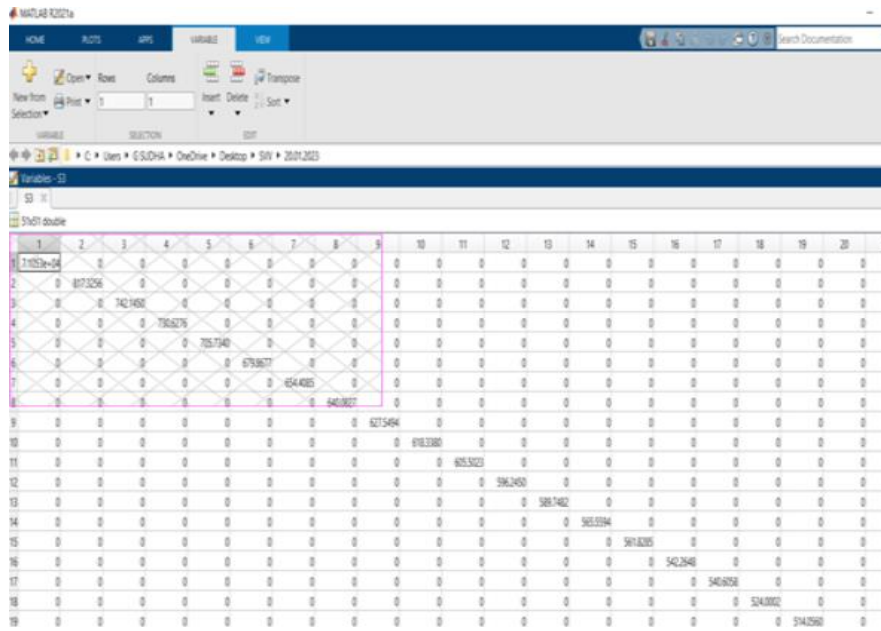


Figure 10: Rank matrix of original image.

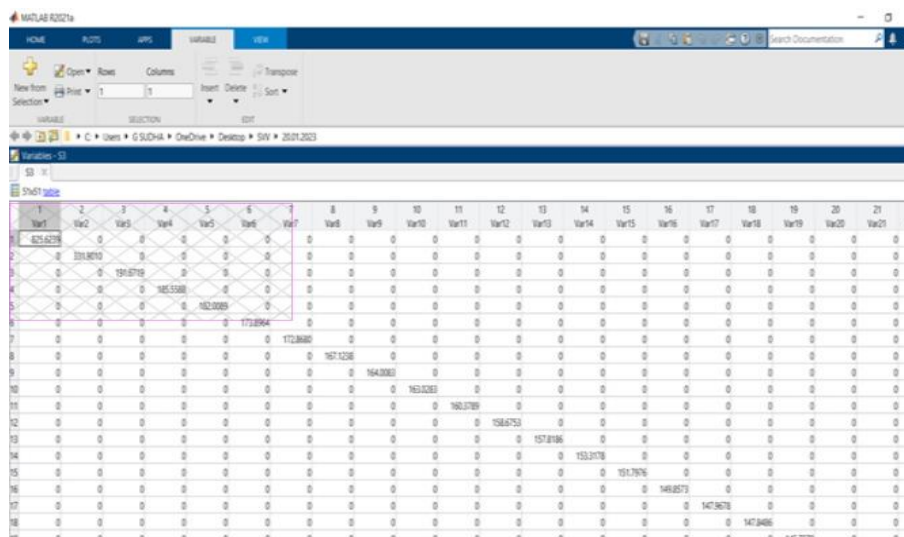


Figure 11: Rank matrix of pixel difference values of original image.

Table 2: Rank and PSNR values for different rank values

Image	Rank	SVD applied to original image	SVD applied to pixel difference values
		PSNR	PSNR
Image 1	408	50.51	60.26
	204	35.54	38.23
	51	33.45	36.54
Image 2	408	47.53	56.62
	204	34.04	37.28
	51	33.11	36.42
Image 3	408	46.70	54.26
	204	39.61	42.51
	51	35.21	38.44
Image 4	408	45.17	53.52
	204	31.48	35.57
	51	29.52	32.54

Image 5	408	46.06	53.95
	204	34.32	37.23
	51	33.15	36.99
Image 6	408	47.15	55.62
	204	27.99	32.78
	51	26.54	30.98

Table 3: Number of output bits that are to be transmitted for different rank values.

Image	Rank	SVD applied to original image	SVD applied to pixel difference values
		RLE Output (No. of Output bits)	
Image 1	408	431666	319466
	204	215894	121798
	51	53960	25426

Image 2	408	431666	309560
	204	215834	138853
	51	52390	24587
Image 3	408	377708	287564
	204	207918	103088
	51	53960	24692
Image 4	408	431666	308546
	204	215834	140357
	51	53960	23892
Image 5	408	377708	287420
	204	161876	128178
	51	53960	24568
Image 6	408	539582	436554
	204	323750	173110
	51	53960	25621

Table 4 gives a comparison of the PSNR metric that aids for effective image reconstruction at the receiver. PSNR for SVD [3] is around 25 dB, SPIHT is around 30 dB, DCT is 30 dB, JPEG is around 27 dB, EZW is 33 dB and the proposed methodology is around 36 dB and the propose techniques can be effectively used for image compression and transmission for all types of images. All the above-mentioned algorithms are tested after applying the LRE process, and since RLE is lossless, it does not affect PSNR.

Table 4: Comparison of PSNR of existing compression methods with the proposed hybrid methodology

Image	Compression Method					
	SVD at rank 51 [3]	SPIHT [18]	DCT [19]	JPEG [20]	EZW [21]	Proposed hybrid compression at Rank 51
	PSNR (dB)					
Image 1	25.42	30.58	32.79	27.52	32.52	36.54
Image 2	25.62	30.91	30.64	28.14	33.12	36.42
Image 3	26.42	32.56	37.81	27.59	33.45	38.44
Image 4	25.32	31.48	30.86	27.56	31.89	32.54
Image 5	26.84	31.98	30.18	27.41	32.75	36.99
Image 6	26.52	31.02	29.29	28.45	29.59	32.98

Table 5 illustrates the number of output bits that are to be transmitted for various compression methods.

The proposed model is also justified in terms of energy consumption requirements by using the standard energy models as in [23]. With an initial energy of 7 Joules for every node in a network of 25 nodes, the energy consumed by the nodes for a hop-by-hop transmission from node 1 to node 25 (for example) is calculated by the formula (3) and (4) where ‘x’ denotes the number of bits, ‘d’ represents the distance between the nodes, E_e denotes the electronics energy.

$$E_t(x, d) = xE_e + kE_f d^2 \quad (3)$$

$$E_r(x) = xE_e \quad (4)$$

Table 5: Comparison of number of output bits that are to be transmitted for existing compression methods with the proposed hybrid methodology

Image	Compression Method					
	SVD at rank 51 [3]	SPIHT [18]	DCT [19]	JPEG [20]	EZW [21]	Proposed hybrid compression
	No. of output bits					
Image 1	53166	86259	77198	82563	85896	25426
Image 2	52239	89564	88930	81475	85786	24587
Image 3	57770	84521	94160	85623	84512	24692
Image 4	52266	89476	69234	85687	84279	23892
Image 5	52266	87493	95680	84568	84352	24568
Image 6	55893	89256	90457	84789	84896	25621

For the output of SVD for different compression ratios, the energy consumed by the source node is plotted in Figure 12, Figure 13 & Figure 14 respectively.

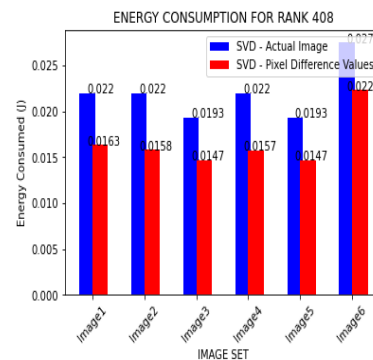


Figure 12: Plot of energy consumption with rank 408 for a node.

The above plots reveal that, based on the values obtained for different rank values, the pre-processed SVD gives less energy consumption when compared to processing the actual image through SVD of 25% for rank 408 to 50% for rank 51 and this energy conservation can be very well utilized in WSNs for voluminous data processing. Also, this process does not involve trimming of SVD matrices as the input pixel value are very less because of the pixel difference values, the PSNR is also maintained for good image reconstruction in the receiver end. The values are also compared with the various compression algorithms and for validation; the energy consumption of a network of nodes for the proposed hybrid compression algorithm is compared with the state-of-the-art algorithms and represented in Table 6.

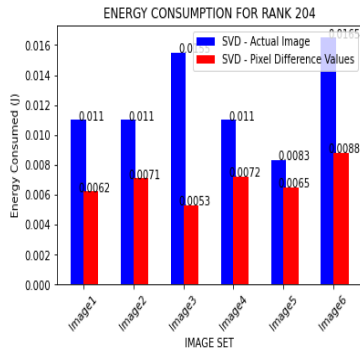


Figure 13: Plot of energy consumption with rank 204 for a node.

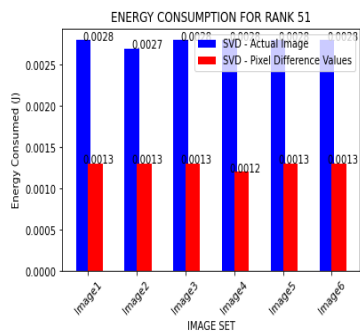


Figure 14: Plot of energy consumption with rank 51 for a node.

Table 6: Comparison of energy consumption of the nodes for existing compression methods with the proposed hybrid methodology.

Image	Compression Method					Proposed hybrid compression
	SVD at rank 51 [3]	SPIHT [18]	DCT [19]	JPEG [20]	EZW [21]	
	Energy consumption (J)					
Image 1	0.0027	0.0044	0.0041	0.0044	0.0044	0.0013
Image 2	0.0027	0.0042	0.0043	0.0044	0.0044	0.0013
Image 3	0.0039	0.0046	0.0042	0.0044	0.0043	0.0013
Image 4	0.0027	0.0044	0.0041	0.0044	0.0043	0.0013
Image 5	0.0027	0.0045	0.0044	0.0044	0.0043	0.0013
Image 6	0.0029	0.0044	0.0044	0.0044	0.0043	0.0013

The proposed method gives a PSNR of around 37 dB with 50% compression. Also, the energy consumed by a network of 25 nodes is 0.0163 J, 0.0062 J, and 0.0013 J respectively for the rank values 408, 204, 51 of the proposed hybrid algorithm, as against 0.022 J, 0.011 J, 0.0028 J for the actual SVD + RLE without preprocessing.

4 Conclusion and future work

SVD is a promising technique for dimensionality reduction for applications involving memory intensive data. In this work a pre-processing step involving difference between the adjacent pixels is taken that takes a smaller number of bits to represent every data compared to actual pixel values and then SVD and RLE is applied.

The PSNR obtained is substantially increased when compared to the conventional SVD, as the process involves no truncation of the matrices related to the rank matrix and the energy consumption for the nodes is also less. The work can be further substantiated by applying the algorithm for data generated from various applications ranging from traffic surveillance, habitat monitoring, industry monitoring, etc. As conventional compression techniques are not feasible to be applied to WSNs due to its impediments like limited resources, limited memory power and the need for prolonging the network lifetime in remote deployments, this contemporary hybrid technique is more promising in terms of PSNR, compression ratio, SSIM and more energy savings over a network of nodes, so that the network lifetime is also enhanced and the reconstructed image is also of a good quality for better interpretation.

The proposed hybrid compression algorithm can further be extended with an additional pre-processing technique like feature extraction, so that only the desired features are compressed and transmitted. Due to this the number of bits that will be transmitted through the network will be substantially reduced and it can be validated by measuring the energy consumption and longevity of the nodes.

References

- [1] Hongran Li et al., “Singular vector sparse reconstruction for image compression”, *Computers and Electrical Engineering*, Vol. 91 (2021). <https://doi.org/10.1016/j.compeleceng.2021.107069>
- [2] Khadeejah James Audu, “Application of singular value decomposition for compressing images”, *Gadua Journal of Pure and Allied Sciences*, 1(2): 82-94 (2022). <https://doi.org/10.54117/gjpas.v1i2.21>
- [3] Ranjeet Kumar et al., “An efficient technique for image compression and quality retrieval using matrix completion”, *Journal of King Saud University – Computer and Information Sciences*, Vol. 34 (2022). <https://doi.org/10.1016/j.jksuci.2019.08.002>
- [4] B. Mtengi et al, “Data compression algorithms for wireless sensor networks: A review and comparison”, *IEEE Access*, Vol. 9 (2021). <https://doi.org/10.1109/ACCESS.2021.3116311>
- [5] Hung-Yi Chen et al., “Improved efficiency on adaptive arithmetic coding for data compression using range-adjusting scheme, increasingly adjusting step, and mutual-learning scheme”, *IEEE Transactions on Circuits and Systems for Video Technology*, (2017). <https://doi.org/DOI:10.1109/TCSVT.2017.2749449>
- [6] Zihan Chen, “Singular value decomposition and its applications in image processing”, *Proceedings of the 1st International Conference on Mathematics and Statistics*, (2018). <https://doi.org/10.1145/3274250.3274261>
- [7] Mario Siller et al., “Wireless sensor networks formation: approaches and techniques”, *Journal of Sensors*, (2016). <https://doi.org/10.1155/2016/2081902>

- [8] Azniza Abd Aziz et al., “Error-control truncated SVD technique for in-network data compression in wireless sensor networks”, *IEEE Access*, (2021). <https://doi.org/10.1109/ACCESS.2021.3051978>
- [9] Helio Pedrini et al., “Adaptive lossy image compression based on singular value decomposition”, *Journal of Signal and Information Processing*, Vol. 10 (2019). <https://doi.org/10.4236/jsip.2019.103005>
- [10] Chong Han et al., “An image compression scheme in wireless multimedia sensor networks based on NMF”, *Information*, MDPI (2017). <https://doi.org/10.3390/info8010026>
- [11] Mohammed K. Al-Obaidi, Anas Fouad Ahmed, “Implementation of image compression based on singular value decomposition”, *Global Journal of Engineering and Technology Advances*, Vol. 11, No. 3, (2022). <https://doi.org/10.30574/gjeta.2022.11.3.0097>
- [12] Awwal Mohammed Rufai et al. “Lossy image compression using singular value decomposition and wavelet difference reduction”, *Digital Signal Processing*, Vol. 24 (2013). <https://doi.org/10.1016/j.dsp.2013.09.008>
- [13] Rahebi J, “Vector quantization using whale optimization algorithm for digital image compression”, *Multimedia Tools and Applications*, Vol. 81 (2022). <https://doi.org/10.1007/s11042-022-11952-x>
- [14] Zermi, N et al., “A lossless DWT-SVD domain watermarking for medical information security”, *Multimedia Tools and Applications*, Vol. 80 (2021). <https://doi.org/10.1007/s11042-021-10712-7>
- [15] Asaad A Alhijaj, et al., “Fuzzy data aggregation approach to enhance energy-efficient routing protocol for HWSNs”, *Informatica*, Vol. 46, No. 7 (2022). <https://doi.org/10.31449/inf.v46i7.4272>
- [16] Fuad Bajaber, et al., “Adaptive decentralised re-clustering protocol for wireless sensor networks”, *Journal of Computer and System Sciences*, Vol. 77, No. 2 (2011). <https://doi.org/10.1016/j.jcss.2010.01.007>
- [17] G. Sudha, C. Tharini, "Analysis of wavelets on discrete wavelet transform for image compression and transmission in wireless sensor networks", *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)* (2022). <https://doi.org/10.1109/IC3IoT53935.2022.9767977>
- [18] He, Fangzhou, “Exploration of Distributed Image Compression and Transmission Algorithms for Wireless Sensor Networks”, *International Journal of Online and Biomedical Engineering (iJOE)*, 15(1):143-155, (2019). <https://doi.org/10.3991/ijoe.v15i01.9782>
- [19] R. J. Cintra, et. al., “Low-complexity 8-point DCT approximations based on integer functions”, *Signal Processing*, 99, 201 – 214, (2014). <https://doi.org/10.48550/arXiv.1612.03461>
- [20] Ma, Tao et al., “A Survey of Energy-Efficient Compression and Communication Techniques for Multimedia in Resource Constrained Systems”, *IEEE Communications Surveys & Tutorials*, (2012). <https://doi.org/10.1109/SURV.2012.060912.00149>
- [21] Boujelbene, R et al, “Enhanced embedded zerotree wavelet algorithm for lossy image coding”, *IET Image Processing*, 13(8), 1364-1374 (2019). <https://doi.org/10.1049/iet-ipr.2018.6052>
- [22] Hanaa ZainEldin, Mostafa A. Elhosseini, Hesham A. Ali, Image compression algorithms in wireless multimedia sensor networks: A survey, *Ain Shams Engineering Journal*, Volume 6, Issue 2 (2015). <https://doi.org/10.1016/j.asej.2014.11.001>
- [23] V. K. Subhashree et al., "Modified LEACH: A qos-aware clustering algorithm for wireless sensor networks," *2014 International Conference on Communication and Network Technologies*, (2014). <https://doi.org/10.1109/CNT.2014.7062737>

Deep Learning Models in Computer Data Mining for Intrusion Detection

Yujun Wang

Department of Science and Technology, Xi'an Siyuan University, Xi'an, Shaanxi 710038, China

E-mail: xawyj@163.com

Keywords: intrusion detection system (IDS), DL, Data mining (DM), and CNN

Received: June 15, 2023

In recent years, the expanded usage of wireless networks for the transfer of enormous amounts of data has caused a multitude of security dangers and privacy issues; accordingly, a variety of preventative and defensive measures, such as intrusion detection (ID) systems, have been developed. ID methods serve a crucial role in safeguarding computer and network systems; yet, performance remains a serious concern for many IDS. The effectiveness of IDS was analyzed by constructing an IDS dataset comprised of network traffic characteristics to identify attack patterns. ID is a classification challenge requiring the use of DL and Data Mining (DM) methods to categorize network data into regular and attack traffic. In addition, the kinds of network assaults have evolved, necessitating an upgrade of the databases used to evaluate IDS. In this study, we present a DL-based IDS that combines an optimization technique called spider monkey swarm with a convolutional neural network (SMSO-CNN). With the use of the well-known NSL-KDD dataset, the SMSO-CNN is assessed and contrasted with the following methods: DNN, k-nearest neighbor, and LSTM. The results show that the SMSO-CNN outperforms compared to other approaches in terms of accuracy.

Povzetek: Predlagana spremenjena strategija algoritma SMO (verzija AI algoritma na osnovi roja) izboljša globalno konvergenco in presega standardni SMO na 20 testnih funkcijah.

1 Introduction

An ID is the process of reviewing the system logs to check for footprints and identify any interaction. ID has been achieved throughout the years using a variety of methods, including statistical, bio-inspired, fuzzy, Markov, etc. [1]. An IDS is important because it may warn companies about potential security issues and provide early detection. It reduces the danger of data breaches, illegal possession of confidential data, and disruption of crucial systems by quickly spotting and reacting to intrusions. In addition, an IDS can help with emergency response efforts by offering important details about the type of an attack and supporting the containment, investigation, and recovery processes [2]. Distributed computer systems' complexity, significance, and informational resources have grown extremely quickly. This reality has led to an increase in computer crimes in recent years, which have increasingly targeted computers and their networks [3]. IDS was any piece of hardware, software, or a hybrid that keeps an eye out for harmful behavior on a system or network of computers. Any IDS's ultimate objective was to apprehend offenders in the act before they seriously harm resources. An IDS guards against compromise, abuse, and assault on a system. Moreover, it examined data integrity, audits network and system settings for vulnerabilities, and monitors network activities. These days, IDS was a crucial part of the security toolkit. Three services were offered by an IDS: monitoring, detection, and alarm generation. IDS are often seen as a firewall feature. Together, firewalls and

IDS improve network security [4]. A foundation for computer and information security. Its fundamental objective was to distinguish between normal system activity and activity that can be considered suspicious or invasive. IDS were required since there are more reported events every year and attack methods are always evolving. The two basic kinds of IDS methods are abuse and anomaly detection. According to the misuse detection technique, an intrusion may be identified by comparing the present activity to a list of invasive patterns. Expert systems, keystroke tracking, and state transition analysis were a few examples of abuse detection. Systems for detecting anomalies assume that an intrusion should cause a departure from the system's typical behavior. Statistical approaches, neural networks, the creation of prediction patterns, association rules, and other methods may all be used to accomplish these strategy [5]. The ability of network managers to identify policy breaches has led to the widespread adoption of Network ID Systems. These policy breaches vary from insiders misusing their access to outside adversaries seeking to get illegal access [6]. For remedial action to be taken, as an alternative, detecting policy violations enables managers to spot weak spots in their defenses [7]. IDS accurately identifies the information and forecasts outcomes that may be used in the future [8]. As they proposed, intelligent network IDS based on SMSO-CNN. SMSO was an extension of the CNN approach, which relies on conditional independence assumption. CNN enhances the attack detection accuracy.

Reference	Objectives	Methods	Strength	Weakness	Results
[9]	The goal of the paper was to discuss the difficulties associated with wireless identification, including signal strength variations, noise, and interference.	DL-based IDS that combines a feature selection approach	Enhanced Model Performance, Improved Efficiency, Transferability to Different Domains.	Need a Complexity and Computational Requirements. Dependency on High-Quality Training Data and Vulnerability to Adversarial Attacks.	The experimental findings demonstrate that, in comparison to previous approaches, the FFDNN-IDS achieves a higher accuracy.
[10]	The purpose of the paper proposing the use of a BD-based (Big Data-based) DL-based system for intrusion detection is to create a practical and scalable method for identifying and containing network intrusions in large-scale and complex environments.	Big data (BD)-based Hierarchical DL System	It improves the identification rate of intrusive assaults relative to single modeling approaches by utilizing numerous DNN models	Privacy and Security Concerns, Adaptability to Evolving Attacks, Resource Requirements for Training	The outcomes shows that how well the system was able to recognise and categorise various forms of intrusions.
[11]	our study aimed to examine an extensive variety of detection of network intrusion models.	BD-related network ID algorithms	Big Data (BD)-related network intrusion detection (ID) algorithms have several strengths that make them effective in detecting and mitigating security threats in large-scale network environments.	Computational Complexity, data storage and management, Privacy and Ethical Considerations.	This survey examined research on big data analytics, focusing on its challenges and various intrusion detection techniques specifically designed for big data analytics.
[12]	It aimed to provide a thorough analysis of contemporary work on the IoT and ML, as well as intelligent approaches and ID structures in computer networks.	Machine learning based ID approaches.	Automated Detection, Enhanced Detection Accuracy, Integration with Other Security Systems.	Dependency on Training Data, Adversarial Attacks, Concept Drift and Evolving Attacks.	Over 95 pertinent works were reviewed for the study, which covered a range of topics relating to security concerns in IoT systems.

[13]	The research aimed to integrate feature selection and preprocessing strategies to improve the performance of Deep Neural Networks (DNN) for intrusion detection.	feature selection and layer design strategies	Enhance the performance of a ML model	Computational Complexity, Generalizability to New Attacks	Results demonstrate that judicious feature selection and layer configuration can effectively shorten the learning period of the model without sacrificing its overall accuracy.
[14]	To identifies the well-known security threats	DL-based ID Systems	Automated Feature Learning, Improved Detection Accuracy, Real-Time Detection and Response.	High Computational Requirements, Dependency on Large Labeled Datasets, Vulnerability to Adversarial Attacks.	The architecture of the Cisco IoT reference model was specifically discussed
[15]	It aimed to investigate and propose effective strategies for data processing and model selection in the context of network intrusion detection using machine learning techniques.	A machine-learning approaches for IDS.	Automatic Detection, Adaptability to Evolving Threats, Improved Detection Accuracy.	Dependency on Labeled training Data, Adversarial Attacks, Privacy and Ethical Considerations.	The study evaluated how nonlinear classifiers beat linear ones and how data normalization and balancing increased most classifiers' overall performance.
[16]	The aim of the study was to improve the IoT security performance.	A deep-learning approach was used to create an algorithm for identifying denial-of-service (DoS) assaults.	Automatic Feature Extraction, Improved Detection Accuracy, Adaptability to Emerging Attack Patterns.	Need lot of data Availability and Representation, Vulnerability to Adversarial Attacks.	The results show that deep-learning models provide more accuracy, enabling more efficient attack prevention on IoT networks.
[17]	The goal of the paper was to compile and assess the body of literature in the area, highlighting the various ML and DL algorithms employed for intrusion detection and the	Taxonomy for ID systems	Clarity and Organization, Decision-Making and Evaluation, Scalability and Flexibility	Security vulnerabilities, Discrimination and exclusion, Lack of standardization and interoperability.	This paper examines and improves the existing challenges and future trends in the field, serving as a valuable reference for researchers conducting

	performance in diverse application domains.				extensive studies.
[18]	Aimed to tackle the problem of interpreting firewall logs.	ML and DL methods	Ability to handle large and complex datasets, Automation and efficiency, Adaptability and generalization, Handling non-linear relationships.	Computing power needed for inference and training. Particularly DL models have a tendency to require high levels of processing, memory, and storage.	Numerous ML and DL algorithms, such as K-Nearest Neighbor (KNN), NB (NB), J48, Random Forest (RF), and Artificial Neural Network (ANN), were evaluated
[19]	The purpose of the research was to investigate and suggest effective model selection strategies and data processing techniques for ML-based network intrusion detection systems.	Numerous ML-based anomaly-based Intrusion Detection Systems (IDS)	Enhanced detection capabilities, reduced false positives, Detection of unknown attacks, Adaptability to changing threats.	Adversarial Attacks, Training Data Limitations.	Our findings show that non-linear classifiers perform better than linear ones overall, and that using data balance and normalization approaches increases the accuracy of most classifiers.
[20]	The purpose of this study was to evaluate several machine learning methods for traffic classification in relation to intrusion detection systems (IDS). The CICIDS2017 dataset, which contains bidirectional traffic flows representing both benign traffic and various modern assaults, was the main focus of the authors' work.	correlation-based feature selection (CFS) technique	Improved Classification Performance, Reduced Overfitting, Faster Training and Inference	Computational Complexity, Sensitivity to Noise, Ignores Non-Correlated Attributes.	The findings demonstrated that decision-tree-based approaches (PART, J48, and random forest) were the most effective, averaging F1 values above 0.999 for the whole dataset
[21]	The aim of this work is to evaluate a	method that combines innovative	Enhanced Feature Representation, Discovering	Complexity, Data Requirements, Interpretability	Performance indicators like precision,

	medical database of diabetes patients using a combination of innovative hierarchical decision attention network, association rules (AR), and multiclass outlier classification with the MapReduce framework.	hierarchical decision attention networks, association rules (AR), and multiclass outlier classification	Associations, Scalability and Efficiency		accuracy, recall, and F-score are used to display the results of the suggested method
[22]	The objective of this paper is to compare various deep learning (DL)-based network intrusion detection systems (IDSsystems). The research also seeks to improve DL models with Generative Adversarial Networks (GANs).	Generative Adversarial Network	Technique is most suited for the NIDS application, Learning Data Representation	Training Instability, Evaluation Metrics, Sensitivity to Hyperparameters	Evaluation metrics are constructed to evaluate each algorithm's performance, and adding synthetic data produced by GAN is meant to increase the NIDS's overall accuracy.
[23]	Emphasized the value of sequential data modeling in cybersecurity, an area where temporal features were key.	Recurrent neural networks (RNNs), a subclass of artificial neural networks (ANNs)	Handling Sequential Data	Vanishing and Exploding Gradient Problems	Compared the various RNN designs to conventional machine learning classifiers, experiments showed a decreased percentage of false positives.
[24]	The purpose of this study is to outline a successful feature engineering approach for Deep Neural Networks (DNNs) in the context of Intrusion	Deep neural network	Capability to Learn Complex Patterns, End-to-End Learning	Need for Large Amounts of Labeled Data, High Computational Requirements.	Using the benchmark ID dataset, a thorough comparison of trials in DNN with various machine-learning methods was conducted

	Detection (ID) in the Internet of Medical Things (IoMT) architecture. The suggested method uses a hybrid strategy that combines Grey Wolf Optimization (GWO) with Principal Component Analysis (PCA), or PCA and PCA, respectively.				
[25]	Aimed to planning and execution of the detection system, it gives thorough information.	Intrusion detection method	Threat Detection, Real-time Monitoring	Dependence on Signature Databases	The test results show that the system satisfies wireless sensor network ID requirements with high accuracy and speed

To overcome the issues of this paper we proposed the SMSO-CNN method which shows that better the outcomes compared to this existing work.

2 Contribution of the study

Thus, this research contributes by demonstrating an implementation of the SMSO-CNN to increase its effectiveness by boosting students' interest, motivation, and the field's relevance to their lives. The following are some of the particular accomplishments of this paper:

- The approach of ID based on Spider Monkey Swarm Optimization is examined.
- To enhance the DL method of CNN which is the ability to handle large amounts of data and assess the effectiveness of the process, an efficient learning component.

3 Proposed methods

In this part, we defined the technique utilized to create the model, outlined the primary processes that were taken to construct the model, and provided an in-depth description of how the steps of the suggested model in Figure 1 were developed. This discussion is broken up into four sections: The first section is devoted to

information gathering. The discretization procedure, feature selection and extraction methods, and other data pre-processing methods are discussed in the second section. The most crucial information is presented in the third section, which details the work done to construct the

suggested model and compile the foundational experiences. The fourth step is to assess the success of each current and new model by comparing their respective parameters.

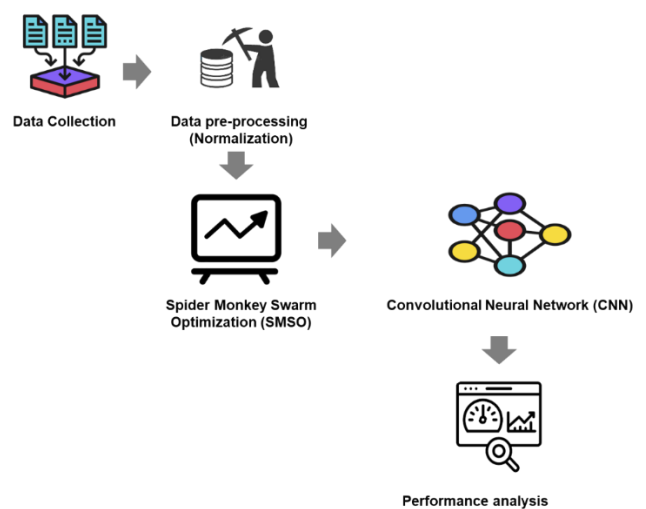


Figure 1: Proposed method framework

A. Data set

The NSL-KDD dataset is a widely used benchmark dataset in the field of network ID and DL research. It was created to addressing some of its limitations and providing a more suitable dataset for evaluating DL models in computer data mining for intrusion detection. The NSL-KDD consists of

41 characteristics, three of which are non-numeric and the other 38 numerical.

The SMSO-CNN model uses CNN to extract features from the sequential nature of SMS messages, capturing patterns and semantics particular to text data. Therefore, using the NSL-KDD dataset, which lacks SMS-specific properties and context, may not offer the essential data representation for efficiently assessing the performance of the SMSO-CNN model. A brand-new data set called NSL-KDD is suggested as a solution to these problems.

B. Data pre-processing

Use a range of preparation procedures to address missing, noisy, and data inconsistency, as well as to clean datasets. Part of the process of cleaning data involves getting rid of things like blank fields, redundant entries, syntax errors, and missing codes. Consistent data-cleaning methods have been applied to the datasets so that the cleaned data can be easily obtained and analyzed. The process of normalization requires the generation of brand-new data vectors. Reducing the likelihood of data redundancy is a major advantage of normalization. The Min-Max normalized approach is crucial for data integration and standardization. The values for each characteristic may be anywhere from 0 to 1, where 0 is the minimum and 1 is the maximum. The normalizing procedure may be expressed as,

$$W_{norm} = \frac{W_j - W_{min}}{W_{max} - W_{min}} \tag{1}$$

When dealing with data in groups, the lowest and highest values are denoted by W_{min} and W_{max} , respectively, where W_j is a data point. The algorithm known as Recursive Feature Elimination (RFE) ranks the most important features and produces a number representing their relative importance. Feature subsets are predicted to shrink when redundancies are eliminated. Sorting the variables from most essential to least interesting may be used to establish the ranking of qualities.

C. Spider monkey swarm optimization coupled with convolutional neural networks (SMSO-CNN)

The CNN's parameters or architecture are optimized using the SMSO algorithm when SMSO is used with a CNN. CNNs are a common deep learning model type for image processing and pattern recognition tasks. Finding the ideal values for various parameters, such as filter sizes, number of layers, learning rates, etc., is typically how CNNs are optimized. Particularly for complex networks or sizable datasets, this procedure can be time-consuming and computationally expensive. The algorithm can be used to automatically look for the ideal values of these parameters by integrating SMSO with CNNs.

SMSO is an algorithm proposed on the premise of swarm intelligence; it employs a cluster of SMs whose behavior is modeled after that of the foraging behavior of honey

SMSs. When there are fewer monkeys in one group, fission occurs and the fusion time is established. The algorithm relies on the social structure of a set of traits from the leader, who decides to either pool resources or move in separate directions in the quest for food. Each subgroup also has a leader; however, they report to the global leader. This is the property of spider monkeys:

- There are 40- 50 spider monkeys in each band.
- The eldest woman in each group acts as a GL and makes almost all of the party's choices.
- Each smaller group is led by a local woman who is in charge of organizing the foraging schedule.

Here are more details about the main parts of Spider Monkey Optimization.

3.1 Setting up the population

Each spider monkey's starting location in the population is represented by its initial parameters, TN_{or} ($o=1, 2, \dots, N$), an N-D vector where N specifies the number of issue variables to be improved. Each SM pinpoints an achievable goal that might fix the issue. It is defined as Eq. for each $TN_{or}(1)$

$$TN_{or} = TN_{minq} + VQ(0,1) \times (TN_{maxq} - TN_{minq}) \tag{2}$$

Where TN_{maxq} and TN_{minq} are minimum and maximum values of TN_{or} in the direction and (0, 1).

3.2 Local leader phase

At this step, the SMO updates its actual role related to the decisions of its local group and local leader (LL), and it also determines the fitness values for the positions of any newly arrived monkeys. This is the stage when Spider monkeys must increase their fitness by replacing their previous positions with new ones. The equation for the o th TN 's position is as follows,

$$TN_{newor} = TN_{or} + VQ(0,1) \times (KK_{kr} - TN_{or}) + VQ(-1,1) \times (TN_{qr} - TN_{or}) \tag{3}$$

In this case, the o th dimensions of the k th LL position corresponds to the r th component of the k th SM. The dimensional TN_{qr} is the r th TN picked at random from the k th group where r is less than or equal to V in the r th dimensions.

3.3 Global leader phase

Members of both the GL and LL groups share their insights to aid in the spider monkeys' stance adjustment. The coordinates may be found by,

$$TN_{newor} = TN_{or} + VQ(0,1) \times (HK_{kr} - TN_{or}) + VQ(-1,1) \times (TN_{qr} - TN_{or}) \tag{4}$$

Where ($r = 1, 2, \dots, N$) is a randomly chosen index and GL_j is the r th dimension of the GL location. At the GLP stage, spider monkeys (TN_{or}) have their positions updated according to the r_i values of the probabilities that are taken into account for calculating their fitness. This manner, the most qualified applicant may best present themselves. The

following equation may be used to determine the probability of r_i :

$$r_i = (\text{fitness } i_x / \text{fitness max}) + 0.1 \quad (5)$$

Where fitness max is the highest possible fitness level for the i th group. In addition, the optimal location is selected by calculating a new fitness algorithm that relies on the created position and comparing it to the previous fitness parameter.

3.4 Global leader learning segment

In the GLL segment, the pessimistic model is used to update and perform the feature extraction. The population is used to choose and create the fitness function value. The optimal value of the place determines the value of the world leader. Instead of updating, the value is increased by one and stored in the Global Limit Count variable.

3.5 Local leader learning phase

According to the fitness values of a community organization, the LLL is changed in the SM location, making it the best possible choice for the local community. It's worth whatever the current regional authority decides it's worth. As it increases by one with each new LLC, no additional updates are supplied.

3.6 Local leader decision phase

If the LLD doesn't update its location using initial randomization or the knowledge of the GL and LL, it does so use the perturbations rate,

$$TN_{newor} = TN_{or} + VQ(0,1) \times (HK_{kr} - TN_{or}) + VQ(0,1) \times (TN_{qr} - KK_{or}) \quad (6)$$

3.7 Global leader decision phase

At this stage, the GL positioning is monitored for a certain amount of time. The GL then divides the population into subsets, always beginning with at least two and going as high as feasible. At the GLD stage, new groups are established and LLL operations are initiated to choose the LL. The GL is unable to change its location. In addition, when the optimum number of distinct groups is reached, it takes its cue from the spider monkey's fusion-splitting social structure and merges all of the smaller groups into a one, super group.

Fitness is calculated by summing the relative relevance of each attribute. Each aspect of the input data is given a score based on the goal variables. When the likelihood of reaching the node drops before it is reached, the relevance of the feature is calculated based on the impurities of the junction with the values. We may get the node's probability

by dividing the ratio of the observed numbers by the total number of specimens. For optimal feature selection, we utilize to determine the fitness function.

$$\text{fitness feature importance} = \frac{\text{Number of specimens that reach the nodes}}{\text{Total number of samples}} \quad (7)$$

Utilizing the low-level co-evolutionary traits, the SMSO hybridized algorithm creates the hybrid mixed capability. There are merge and combine options available as part of the basic hybrid capability. Co-evolutionary is used because variations are employed sequentially, in parallel. The two types are combined, and both contribute to the creation of answers to the challenges. With this adjustment, the hierarchical SMSO generates variations using the strength of SMSO. The velocity is revised using the combined SMSO variations, as suggested,

$$u_j^{l+1} = x * (u_j^l + d_1 q_1 (w_1 - w_j^l) + d_2 q_2 (w_2 - w_j^l) + d_3 q_3 (w_3 - w_j^{l+1})) \quad (8)$$

$$w_j^{l+1} = w_j^l + u_j^{l+1} \quad (9)$$

The most optimal value is chosen by maximizing the fitness value. The proposed method employs the Rosen Brock function, often called the optimization problem. With the in-built localized without a framework to guide and a proper coordinate system, the Rosen Brock product is effectively maximized,

$$e(w) = \sum_{j=1}^{M-1} [100(y_{j+1} - y_j^2)^2 + (1 - y_j)^2] \quad (10)$$

Each variable's goal function is added together to get the best possible outcome. The equation gives the generic form of the optimal solution,

$$\text{Minimize or } Y \text{ or } hbest = \sum_{j=1}^m d_j Y_j \quad (11)$$

Where Y_j the i th control is input and DJ is the optimization problem factor for the i th parameter.

Hence, a function is used to choose the optimal set of characteristics from the subgroup, and data augmentation is calculated if there is any ambiguity among the features.

D. Convolutional neural network

The input data, convolutional, pooling surface, FC overlay, and output vector are the five components that make up a CNN. There are several layer configurations among CNNs. Figure 2 depicts the structure of the CNN that was employed in this investigation.

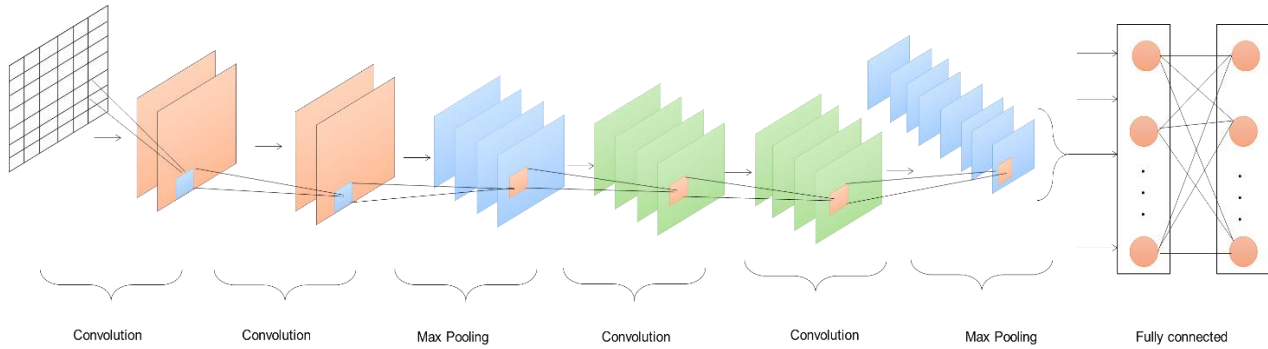


Figure 2: Structure of convolutional neural network

The convolutional gradient task is to identify interesting patterns in the data. Each convolutional kernel in its many layers is associated with a frequency and a divergence coefficient. It is assumed that u_j is the weight parameter, a_j is the divergence amount, and V_{j-1} is the input to convolution layers j while inversion kernel j is active. One such expression for the convolution operation is:

$$V_j = e(u_j \otimes V_{j-1} + a_j) \tag{12}$$

Where, the output result of convolution kernel j , \otimes represents the convolution operation, and $e(x)$ represents the activation function.

The input data is swept repeatedly by the CNN, which then extracts the distinctive information. In addition, the multilayer layer's operational amplifier is changed to *ReLU*. The Linear transfer function is simpler to derive than the exponential, transfer function, and another training algorithm, allowing for faster model training and better protection against gradient disappearing. It is possible to write *ReLU* as:

$$ReLU(V_j) = \begin{cases} V_j & (V_j > 0) \\ 0 & (V_j \leq 0) \end{cases} \tag{13}$$

Downsampling data redundancy is the primary operation of the pooling layer, which also helps to attain invariance and minimize CNN complexity. Pooling layer and maximum pooling are the two most common approaches to completing pooling. If you use averaged pooling, the result is the arithmetic mean of the computation area, whereas if you use max pooling, the result is the largest value of the area.

As max pooling is superior to average pooling at preserving crucial data, it was chosen for this analysis. The mathematical formula for max pooling is:

$$R_i = \max(O_i^0, O_i^1, O_i^2, O_i^3, \dots, O_i^s) \tag{14}$$

Where R_i the return outcome of the pooled region I , Max is the maximum pooling procedure, and O IS is the pooling area i 's element s .

The "classifiers" of a CNN are the FC layers. Its primary purpose is to reconfigure the information from the hidden-layer space that the convolutional as well as pooling layers extracted and weighted.

4 Result and discussion

To apply the recommended methods for detecting intrusion using the Spider Monkey Swarm Optimization - Convolutional Neural Network (SMSO-CNN) approach, DL methodology was used. We employ indicators like accuracy, precision, detective rate, and false alarm rate for analysis.

Confusion matrix

When it comes to assessing IDS, this matrix is among the top options. Each section in this matrix indicates the predicted class, and each row depicts the actual section; the performance of the model is determined by several metrics. The accuracy rate of the classification is determined by comparing the actual number of records categorized with the number of anticipated records. The matrix's contents are summarized by four factors shown in Table 1.

Table 1: Confusion matrix evaluation

Present		Predictive value	
		Positive	Negative
Class	P	TP	FP
	N	FN	TN

Accuracy

Accuracy is a metric used to assess how well a classification model performs in the context of ML and statistics. It calculates the ratio of the model's accurate predictions to all of the predictions made.

$$\text{Accuracy} = \frac{\text{Truepositives} + \text{TrueNegatives}}{\text{Truepositives} + \text{Truenegetives} + \text{Falsepositives} + \text{Falsenegetives}} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{15}$$

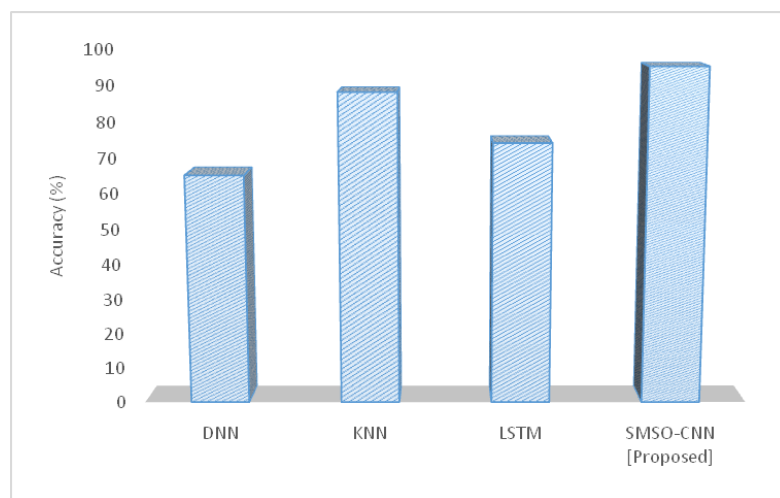


Figure 3: The Accuracy of the proposed and existing system

The accuracy of the suggested strategy is shown in Figure 3. Accuracy percentages are often provided for accuracy levels. Both the existing approach and the one that is being discussed show signs of the potential for inaccurate estimates. This threat is recognized by both systems. The

recommended method, SMSO-CNN, achieves 95% accuracy in contrast to DNN [23], KNN [24], and LSTM [25], 65%, 88%, and 74%, respectively. Thus, the strategy that is suggested has the highest accuracy rate. Table 2 displays the accuracy of the suggested strategy.

Table 2: Comparison of accuracy

Methods	Accuracy (%)
DNN	65
KNN	88
LSTM	74
SMSO-CNN [Proposed]	95

Precision

Precision is a performance parameter that is used to assess the efficacy of a classification model, particularly in binary classification issues. Out of all the occurrences the model predicted as positive, it counts the percentage of accurately predicted positive instances.

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} = \frac{TP}{TP + FP} \quad (16)$$

The precision for the suggested system is shown in Figure 4. Thus, the method that is advised has the best precision. Table 3 displays the recommended approach precision.

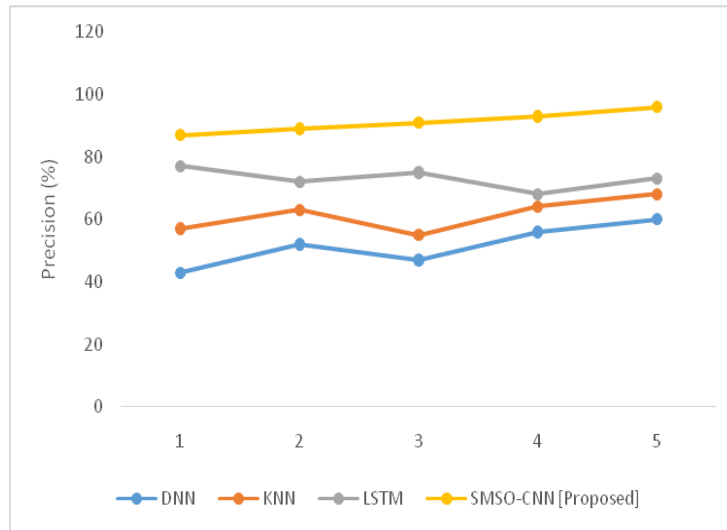


Figure 4: The precision of the proposed and existing method

Table 3: Comparison of precision

Dataset	Precision (%)			
	DNN	KNN	LSTM	SMSO-CNN [Proposed]
1	43	57	77	87
2	52	63	72	89
3	47	55	75	91
4	56	64	68	93
5	60	68	73	96

Detection rate (DR)

TP and FN are the totals for true positives and false negatives, respectively, while DR is the ratio of true positives to all nonself samples detected by the detector set.

$$\text{DetectiveRate} = \frac{TP}{TP+FN} \times 100 \quad (17)$$

There is agreement on the definition of the detection rate, is also called astrue positive rate, and Figure 5 shows the suggested technique.

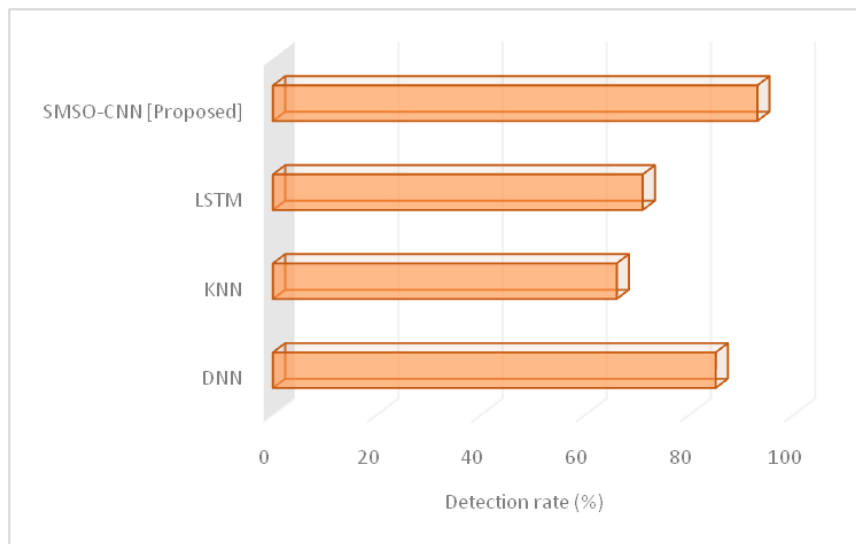


Figure 5: The detection rate of the proposed and existing method

When compared to DNN, KNN, and LSTM, the recommended approach, SMSO-CNN, obtains a 93% detection rate while only having 85%, 66%, and 71%

accuracy, respectively. Thus, the strategy that is suggested has the highest accuracy rate. Table 4 displays the recommended approach detection rate.

Table 4: Detection rate comparison

Methods	Detection rate (%)
DNN	85
KNN	66
LSTM	71
SMSO-CNN [Proposed]	93

False alarm rate

The proportion of benign events that have caused a false alarm is known as the false alarm rate, also called as the false positive rate, whereas the false alarm detection rate (FDR) gauges the percentage of irrelevant notifications.

$$\text{False alarm Rate} = \frac{FP}{FP+TN} \times 100 \tag{18}$$

A high FNR will make the system open to intrusions, while a high FPR will significantly affect how well the IDS performs. Figure6 shows the suggested False Alarm Rate calculation technique.

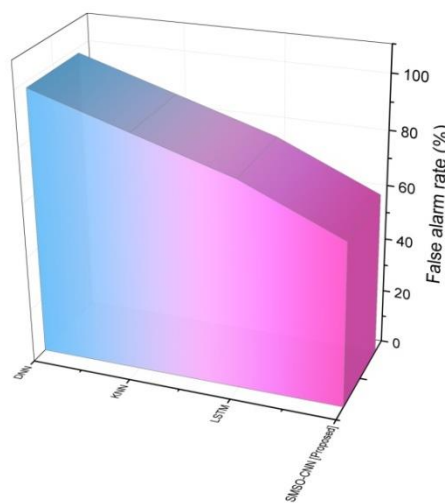


Figure 6: The false alarm rate of the proposed and existing method.

The recommended method, in contrast, obtains a detection rate of 60% whereas DNN, KNN, and LSTM have accuracy rates of 98%, 87%, and 76%, respectively. Thus,

the strategy that is suggested has the highest accuracy rate. Table 4 displays the recommended approach detection rate.

Table 5: Comparison of false alarm rate

Methods	False alarm rate (%)
DNN	98
KNN	87
LSTM	76
SMSO-CNN [Proposed]	60

5 Discussion

Deep Neural Networks (DNNs) are employed in many different fields, but they also have some drawbacks that are large data requirements, Limited performance on small datasets, Vulnerability to adversarial attacks. KNN have some issues in intrusion detection includes the high computational cost, Sensitivity to irrelevant features, Imbalanced dataset challenges and the traditional work of LSTM have some drawbacks in intrusion detection that are training data imbalance, and computationally expensive. In order to overcome these issues SMSO-CNN models were used in this paper. The outcomes show that the SMSO-CNN method performs better than the others in terms of accuracy. According to the results, the suggested SMSO-CNN method performs better than previous methods and is useful for recognizing network assaults. The IDS improves its accuracy in separating ordinary traffic from attack traffic by fusing DL approaches with spider monkey swarm optimization.

6 Conclusion

In this study, we detailed the planning, development, and testing of an SMSO-based DL IDsystem. After researching several DL methods, researchers concluded that there is no one best way to do intrusion detection. Our research led us to the conclusion that the NSL-KDD collection is one of the most reliable control sets for use in simulating IDSs. By studying and assessment of the NSL-KDD dataset, the suggested hybrid model based on DL technology as the last step increases the detection performance increases accuracy, and decreases false alarms. Preparing and analyzing the dataset before the pre-processing step is crucial to building an effective model. This includes picking the right features, shrinking the dimensions, and using the discretization approach to boost IDaccuracy. Research into IDtechnologies has gained traction in response to the growing importance of network security threats. In-depth study on data mining IDmethods led to the development of an outlier-based algorithm for intrusion detection. People have long worried about keeping their networks safe and secure. There are still a lot of studies to be done and difficult problems to be addressed quickly as the network continues to evolve and cyber assaults become more varied.

References

- [1] Zamani, M. and Movahedi, M., 2013. ML methods for intrusion detection. arXiv preprint arXiv:1312.2177.
- [2] Tsai, C.F., Hsu, Y.F., Lin, C.Y. and Lin, W.Y., 2009. IDby ML: A review. expert systems with applications, 36(10), pp.11994-12000.
- [3] Tabash, M., Abd Allah, M. and Tawfik, B., 2020. IDmodel using NB and DL technique. Int. Arab J. Inf. Technol., 17(2), pp.215-224.
- [4] Vinchurkar, D.P. and Reshamwala, A., 2012. A review of an IDsystem using neural network and ML. J. Eng. Sci. Innov. Technol, 1, pp.54-63.
- [5] Florez, G., Bridges, S.A. and Vaughn, R.B., 2002, June. An improved algorithm for fuzzy data mining for intrusion detection. In 2002 Annual Meeting of the North American Fuzzy Information Processing Society Proceedings. NAFIPS-FLINT 2002 (Cat. No. 02TH8622) (pp. 457-462). IEEE.
- [6] Brugger, S.T., 2004. Data mining methods for network intrusion detection. University of California at Davis.
- [7] Sahani, R., Rout, C., ChandrakantaBadajena, J., Jena, A.K. and Das, H., 2018. Classification of IDusing data mining methods. In Progress in Computing, Analytics, and Networking: Proceedings of ICCAN 2017 (pp. 753-764). Springer Singapore.
- [8] Noel, S., Wijesekera, D. and Youman, C., 2002. Modern intrusion detection, data mining, and degrees of attack guilt. Applications of data mining in computer security, pp.1-31.
- [9] Kasongo, S.M. and Sun, Y., 2019. A DL method with filter-based feature engineering for wireless IDsystem. *IEEE Access*, 7, pp.38597-38607.
- [10] Zhong, W., Yu, N. and Ai, C., 2020. Applying BD-based DL system to intrusion detection. *BD Mining and Analytics*, 3(3), pp.181-195.
- [11] Sudar, K.M., Nagaraj, P., Deepalakshmi, P. and Chinnasamy, P., 2021, January. Analysis of intruder detection in BD analytics. In *2021 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-5). IEEE.
- [12] Da Costa, K.A., Papa, J.P., Lisboa, C.O., Munoz, R. and de Albuquerque, V.H.C., 2019. IoT: A survey on

- ML-based ID approaches. *Computer Networks*, 151, pp.147-157.
- [13] Woo, J.H., Song, J.Y. and Choi, Y.J., 2019, February. Performance enhancement of DNN using feature selection and preprocessing for intrusion detection. In *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)* (pp. 415-417). IEEE.
- [14] Idrissi, I., Azizi, M. and Moussaoui, O., 2020, October. IoT security with DL-based ID Systems: A systematic literature review. In *2020 Fourth international conference on intelligent computing in data sciences (ICDS)* (pp. 1-10). IEEE.
- [15] Sahu, A., Mao, Z., Davis, K. and Goulart, A.E., 2020, May. Data processing and model selection for ML-based network intrusion detection. In *2020 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)* (pp. 1-6). IEEE.
- [16] Susilo, B. and Sari, R.F., 2020. ID in IoT networks using a DL algorithm. *Information*, 11(5), p.279.
- [17] Liu, H. and Lang, B., 2019. ML and DL methods for ID systems: A survey. *Applied Sciences*, 9(20), p.4396.
- [18] Aljabri, M., Alahmadi, A.A., Mohammad, R.M.A., Abounour, M., Alomari, D.M. and Almotiri, S.H., 2022. Classification of firewall log data using multiclass ML models. *Electronics*, 11(12), p.1851.
- [19] Sahu, A., Mao, Z., Davis, K. and Goulart, A.E., 2020, May. Data processing and model selection for ML-based network intrusion detection. In *2020 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)* (pp. 1-6). IEEE.
- [20] Rodríguez, M., Alesanco, Á., Mehavilla, L. and García, J., 2022. Evaluation of ML Methods for Traffic Flow-Based Intrusion Detection. *Sensors*, 22(23), p.9326.
- [21] Jayasri, N.P. and Aruna, R., 2022. BD analytics in health care by data mining and classification methods. *ICT Express*, 8(2), pp.250-257.
- [22] Sekhar, C., Kumar, P.H., Venkata Rao, K. and Krishna Prasad, M.H.M., 2022. A Comparative Study on Network ID System Using DL Algorithms and Enhancement of DL Models Using Generative Adversarial Network (GAN). In *High-Performance Computing and Networking: Select Proceedings of CHSN 2021* (pp. 143-155). Singapore: Springer Singapore.
- [23] Vinayakumar, R., Soman, K.P. and Poornachandran, P., 2017. Evaluation of recurrent neural network and its variants for ID system (IDS). *International Journal of Information System Modeling and Design (IJISMD)*, 8(3), pp.43-63.
- [24] RM, S.P., Maddikunta, P.K.R., Parimala, M., Koppu, S., Gadekallu, T.R., Chowdhary, C.L. and Alazab, M., 2020. An effective feature engineering for DNN using hybrid PCA-GWO for ID in IoT architecture. *Computer Communications*, 160, pp.139-149.
- [25] Li, W., Yi, P., Wu, Y., Pan, L. and Li, J., 2014. A new ID system based on KNN classification algorithm in a wireless sensor network. *Journal of Electrical and Computer Engineering*, 2014.

A Modified Spider Monkey Optimization Algorithm Based on Good-Point Set and Enhancing Position Update

Sabreen Fawzi Raheem¹, Maytham Alabbas^{2*}

¹ Basra Technical Institute, Southern Technical University, Basrah, Iraq

² Department of Computer Science, College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq

E-mail: sabreen.fawzi@stu.edu.iq¹, ma@uobasrah.edu.iq^{2*}

* Corresponding author

Keywords: good-point-set method, swarm intelligence, enhancing position update, spider monkey optimization

Received: November 23, 2022

Spider monkey optimization (SMO), developed based on the behavior of spider monkeys, is a recently added swarm intelligence algorithm. SMO is a stochastic population-based metaheuristic algorithm. Spider monkeys have a fission-fusion social structure. In addition to being an excellent tool for solving optimization problems, SMO provides good exploration and exploitation capabilities. In this work, we present a modified strategy for improving the performance of the basic SMO in two ways: (a) we use the good-point-set method instead of random initial population generation; and (b) by changing both the local leader and global leader phases, we modify the SMO position update approach to increase global convergence while avoiding local optima. We evaluate the proposed modified SMO algorithm on 20 popular benchmark functions. The current findings prove that the proposed approach outperforms the standard SMO regarding result quality and convergence rate.

Povzetek: Predstavljeni SMO, stohastičen algoritem, ki temelji na socialni strukturi pajkovih opic, izboljšuje osnovno različico z dvema pristopoma: metodo dobre točke in modificiranim posodabljanjem položaja za večjo globalno konvergenco.

1 Introduction

Nature has inspired many researchers and is therefore considered a rich source of inspiration. Nowadays, most algorithms are inspired by nature. These algorithms depend primarily on one of the successful properties of a biological system. Numerous natural phenomena have inspired academics to create population-based optimization (PBO) methods. Because these PBO methods evaluate fitness, the population of possible solutions is expected to gravitate toward the areas of the potential search spaces with the best fitness. In PBO algorithms, natural selection offers near-optimal solutions to complex optimization problems.

In recent years, researchers have gained interest in swarm intelligence (SI). Studies have proven that algorithms based on SI have enormous potential for numerical optimization. The last few years have seen the development of many related algorithms. In this regard, there are several algorithms available, but not limited to particle swarm optimization (PSO) [1], ant colony optimization (ACO) [2], cuckoo search [3], bacteria foraging optimization (BFO) [4], bat algorithm [5], and artificial bee colony (ABC) [6].

Among the latest techniques to be developed is spider monkey optimization (SMO) [7], a newcomer to the class of SI algorithms. Spider monkey foraging behavior based on fission-fusion social structure (FFSS), which they use to find great food sources and mate, served as the model

for this SMO algorithm. In the same way as any other PBO strategy, it includes the intrinsic population solution that denotes the spider monkey food source.

While searching for the best solution, the SMO algorithm tries to find a suitable balance between exploration and exploitation. It ensures that the local optimal solution is traversed correctly during exploitation, and it explores the global search space to prevent the problem of entrapment in the local optimum during exploration. It has been discovered that SMO effectively explores local search [8].

As SMO is a relatively new algorithm, there is little literature available on it. The authors in [9] developed a modified position update in the SMO method, which includes two revisions to the basic SMO algorithm. The golden section search (GSS) technique was used to modify both the local leader phase (LLP) and the global leader phase (GLP). In [10], Kumar et al. proposed individual fitness as a new technique for updating the position of spider monkeys in the LLP, GLP, and local leader decision (LLD) phases. In [11], Kavita and Kusum proposed a modified version of SMO (TS-SMO), which uses tournament selection to increase the exploration capabilities of SMO and prevent premature convergence. Urvinder and Rohit [12] introduced a modified SMO (MSMO) method based on a dual-search technique for linear antenna array synthesis (LAA). Singh et al. [13] suggested a modified version of SMO (SMONM), which uses Nelder-Mead (NM) transformations to improve the ability of the LLP. The proposed SMONM approach has

the same number of phases as the conventional SMO, except for the LLP, which is modified using the NM reflection, expansion, and contraction transformations. If there is no improvement in the fitness function value after updating the solution with the original LLP, NM modifications are applied.

The rest of this paper is organized as follows: Section 2 briefly explains the fundamental SMO algorithm. Section 3 describes the primary concept of the proposed approach. Section 4 discusses the test functions, simulation results, and comparison outcomes. Section 5 provides a comprehensive explanation.

2 Spider monkey optimization (SMO)

Bansal et al. [7] introduced the SMO algorithm, a nature-inspired evolutionary algorithm. The fission-fusion social structure (FFSS) of spider monkeys (SMs) is modeled in SMO. SMs live in groups of forty to fifty monkeys and separate into subgroups to look for food to reduce competition. The group is led by the most senior female, who is responsible for finding food sources. If she cannot find enough food for the group, she divides it into small subgroups of three to eight individuals. The local leader (LL) leads the subgroups and plans their foraging paths each day. The members of this group are responsible for finding food sources and adjusting their location according to the distance to the food source. It is important for these group members to interact with each other to maintain social bonds, especially if they encounter a stalemate.

2.1 The SMO algorithm's phases

SMO consists of seven phases. Below is a detailed description of each phase of SMO. The steps of the SMO algorithm's phases are given in Algorithm 1.

2.1.1 Initialization phase

Initially, SMO produces N of SMs with uniform distribution. Each $SM_i = 1, \dots, N$ is a D -dimensional vector. The number D indicates how many variables are involved in the optimization problem, and SM_i is the i^{th} individual in the population. Each individual is generated using Eq. 1.

$$SM_i = Lb_j + rand(0,1) * (Ub_j - Lb_j), \forall j \in \{1, \dots, D\}, \forall i = \{1, \dots, N\}, \quad (1)$$

where $rand(0,1)$ is a random number between 0 and 1, the lower bound of the solution location is given by Lb_j , and the upper bound of the solution location is given by Ub_j .

2.1.2 Local leader phase (LLP)

During LLP, SMs change their present positions depending on information from the LL's experience as well as the experience of local group members. The

fitness value of the newly acquired location is computed. If the new position's fitness value exceeds the old position's, the SM replaces its position with the new one. In this phase, the position update equation for the i^{th} SM (member of the k^{th} local group) is calculated as follows:

$$SM_{new_{ij}} = SM_{ij} + rand(0,1) * (LL_{kj} - SM_{ij}) + rand(-1,1) * (SM_{rj} - SM_{ij}), \quad (2)$$

where LL_{kj} stands for the j^{th} dimension of the k^{th} local group leader position, SM_{ij} stands for the j^{th} dimension of the i^{th} SM, and SM_{rj} is the j^{th} dimension of the k^{th} SM that is selected at random within the k^{th} group, such that ($r \neq i$).

2.1.3 Global leader phase (GLP)

The GLP phase begins once the LLP is completed. During this phase, each SM updates its position using Eq. 3, considering the experience of the global leader (GL) and local group members.

$$SM_{new_{ij}} = SM_{ij} + rand(0,1) * (GL_j - SM_{ij}) + rand(-1,1) * (SM_{rj} - SM_{ij}), \quad (3)$$

where GL_j is the j^{th} dimension of the GL position and $j \in \{1, 2, \dots, D\}$ is a randomly selected index.

SM_i 's location is updated throughout this phase based on the probability $prob_i$ that can be calculated using Eq. 4.

$$Prob_i = \frac{fitness_i}{\sum_{i=1}^N fitness_i}. \quad (4)$$

2.1.4 Global leader learning (GLL)

A greedy selection is used to update the GL's position. Consequently, the GL's position is updated based on the SM's position that has the best fitness. In addition, the *GlobalLimitCount* ($GLLimit$) is increased by one if the GL's position has not been updated.

2.1.5 Local leader learning (LLL)

A greedy selection method is used to update the position of the LL in that group, i.e., to reflect the position of the new LL. Selection is made based on the position of the SM in the group with the best fitness. A comparison is then conducted between the new position of the LL and the earlier one, and *LocalLimitCount* ($LLLlimit$) is increased by one if no changes are made.

2.1.6 Local leader decision (LLD)

As long as no LL position has changed above a *LocalLeaderLimit* threshold, whether it is random or using Eq. 5, all group members are updated. This is determined by the perturbation rate (pr), which specifies how much perturbation is in the current position.

If $rand(0,1) \geq pr$ **then**
Using Eq. 1.

Else

$$SM_{new_{ij}} = SM_{ij} + rand(0,1) * (GL_j - SM_{ij}) + rand(0,1) * (SM_{ij} - LL_{kj}). \tag{5}$$

2.1.7 Global leader decision (GLD)

After a certain number of iterations, the global leader's (GL's) position is observed. If it is not changed, the GL divides the population into subgroups. Initially, two groups are separated from the population, and the number of groups is increased until it reaches the maximum number of groups (MG).

The pseudocode for the SMO algorithm is presented in Algorithm 1 [7].

Algorithm 1: SMO algorithm
<ol style="list-style-type: none"> 1: Initialize the population, the Local Leader Limit (LLlimit), the Global Leader Limit (GLlimit), and the perturbation rate (pr). 2: Calculate fitness 3: Choose leaders (both global and local) through greedy selection. 4: While (termination criteria is not) do 5: Create new locations for all group members based on your own experience, local leader experience, and group member experience. Using Eq. (2) 6: Use the greedy selection process to choose between an existing location and a newly generated location based on fitness, and then choose the better one. 7: Using Eq. (4), compute the probability p_i for each group member. 8: Create new locations for all group members based on p_i's selection, using self-experience, global leader experience, and group member experiences. Using Eq. (3). 9: By applying the greedy selection process to all groups, you can update the position of local and global leaders. 10: If a Local group leader does not update her position after a set number of times (LLLimit), the algorithm will re-direct all members of that group to foraging. 11: If the Global Leader does not update her position for a set number of times (GLLimit), she divides the group into smaller groups using the steps below. 12: End while

3 The current work

In the current work, two modifications have been made to make the basic SMO algorithm work better. The first modification is that we used the good-point-set method to generate a suitable initial population for the SMO algorithm. Then, we modified both LLP and GLP of the basic SMO algorithm.

3.1 Modification of the initial population

Initializing the population is one of the most crucial steps in metaheuristic optimization. A successful initial population can generate reasonable solutions faster and accelerate convergence. The basic SMO algorithm

randomly generates the initial population using the stochastic method. However, the initial solutions do not cover the entire search space in quality and may be good or bad. To address these challenges, we used the good-point-set (GPS) method [14] to generate the initial population of the SMO algorithm. GPS is generally used to create several individuals with uniform distributions to preserve the population's diversity. The GPS pseudocode is presented in Algorithm 2 [15, 16,17,18].

Algorithm 2: Pseudocode of the GPS method

*P is the lower prime number with $P \geq 2 * D + 3$.*

- 1: **For** $i=0$ to N **do**
- 2: **For** $j=0$ to D **do**
- 3: $x = 2 * n + 3$,
- 4: **While** $P \neq x$ **do**
- 5: Set individual counter $k = 2$
- 6: **For** $k = 2$ to $x - 1$ **do**
- 7: **If** $mod(x, i) == 0$ **then**
- 8: $x = x + 1$
- 9: **Else**
- 10: $P = x$
- 11: **End if**
- 12: **End for**
- 13: **End while**
- 14: $GPS_{i,j} = i * 2 * \cos\left(2 * \pi * \frac{j}{P}\right)$
- 15: $X_{ij} = Lb_j + GPS_{i,j} * (Ub_j - Lb_j)$
- 16: **End for**
- 17: **End for**

3.2 Modification of the position update in SMO

Maintaining diversity in the LLP and GLP of SMO can balance exploring the entire search space and exploiting the best solutions found nearby. To achieve this balance, the proposed algorithm modifies both LLP and GLP of the basic SMO algorithm using the modified Eqs. 6 and 7, respectively.

$$SM_{new_{ij}} = SM_{ij} + rand(0,1) * (LL_{kj} - SM_{ij}) + rand(-1,1) * (SM_{r1j} - SM_{r2j}). \tag{6}$$

$$SM_{new_{ij}} = SM_{ij} + rand(0,1) * (GL_{kj} - SM_{ij}) + rand(-1,1) * (SM_{r1j} - SM_{r2j}). \tag{7}$$

Where $r1 \in \{1,2, \dots, N\}$ is a different index from i that was chosen at random from the population, and $r2 \in \{1,2, \dots, N\}$ is a different index from i , where $r1 \neq r2$.

4 Experiments

4.1 Test functions

The improved SMO algorithm is evaluated using 20 well-known benchmark optimization functions, f_1 - f_{20} , as

shown in Table 1. These continuous optimization problems cover a range of difficulties, search spaces, and multimodality.

Table 1: The current 20 benchmark optimization functions for testing

NO	Function	type	Search Range	optimal value	Formulation
f_1	Rosenbrock	UN	[-50,50]	0	$f_1(x) = \sum_{i=1}^{D-1} 100 (x_{i+1} - x_i^2)^2 + (x_i - 1)^2$
f_2	sphere	US	[-100,100]	0	$f_2(x) = \sum_{i=1}^D x_i^2$
f_3	Elliptic	UN	[-100,100]	0	$f_3(x) = \sum_{i=1}^D (10^6)^{(i-1)/(D-1)} x_i^2$
f_4	Sum Squares	US	[-10,10]	0	$f_4(x) = \sum_{i=1}^D i x_i^2$
f_5	Quartic	US	[-1.28,1.28]	0	$f_5(x) = \sum_{i=1}^D i x_i^4$
f_6	kowalik	MS	[-5,5]	0.0003075	$f_6(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1(b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4} \right]^2$
f_7	Scaffer's F6	MN	[-100,100]	0	$f_7(x) = 0.5 + \frac{\sin^2 \left(\sqrt{\sum_{i=1}^D x_i^2} \right) - 0.5}{(1 + 0.001 (\sqrt{\sum_{i=1}^D x_i^2}))^2}$
f_8	Rastrigin	MS	[-5.12,5.12]	0	$f_8(x) = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10)$
f_9	Griewank	MN	[-600,600]	0	$f_9(x) = \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$
f_{10}	Ackley	MN	[-32.768,32.768]	0	$f_{10}(x) = 20 + e - 20 \exp \left[-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} \right] - \exp \left[\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i) \right]$

f_{11}	Himmelblau	MS	[-5,5]	78.3323	$f_{11}(x) = \frac{1}{D} \sum_{i=1}^D (x_i^4 - 16 x_i^2 + 5 x_i)$
f_{12}	Step	US	[-100,100]	0	$f_{12}(x) = \sum_{i=1}^D (x_i + 0.5 _i)^2$
f_{13}	Beale	UN	[- 4.5,4.5]	0	$f_{13}(x) = [1.5 - (1 - x_2)]^2 + [2.25 - x_1(1 - x_2^2)]^2 + [2.625 - x_1(1 - x_2^3)]^2$
f_{14}	Easom	UN	[-100,100]	-1	$f_{14}(x) = -\cos x_1 \cos x_2 e^{((-x_1-\pi)^2 - (x_2-\pi)^2)}$
f_{15}	Schwefel	MS	[-500,500]	0	$f_{15}(x) = 418.9829 * D - \sum_{i=1}^D x_i \sin(\sqrt{ x_i })$
f_{16}	Levy N.13	MN	[-10,10]	0	$f_{16}(x) = \sin^2(3\pi x_1) + (x_1 - 1)^2(1 + \sin^2(3\pi x_2)) + (x_2 - 1)^2(1 + \sin^2(2\pi x_2))$

f_{17}	Goldstein-Price	MN	[-2,2]	3	$f_{17}(x) = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 13x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] * [30 + (2x_1 - 3x_2)^2 (18 - 32x_1 + 12x_1^2 - 48x_2 - 36x_1x_2 + 27x_2^2)]$
f_{18}	Colville	MN	[-10,10]	0	$f_{18}(x) = [100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1)]$
f_{19}	Booth	UN	[-10,10]	2	$f_{19}(x) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$
f_{20}	Dekkers and Aarts	MN	[-20,20]	2	$f_{20}(x) = 10^5 + x_1^2 + x_2^2 - (x_1^2 + x_2^2)^2 + 10^{-5}(x_1^2 + x_2^2)^4$

4.2 Parameter setting

The present work and the basic SMO share the following primary control parameters:

- The Swarm size (N) = 80,
- The Minimum Group (MG) = 5,
- LocalLeaderLimit = N
- GlobalLeaderLimit = D * N,
- Pr grows linearly over iterations by Eq. 8, where initial pr \in [0.1, 0.4],

$$pr = pr + (0.4/iter). \quad (8)$$

4.3 Experimental results

To evaluate the proposed approach, we compared it to the basic SMO algorithm on 20 popular benchmark functions. Table 2 shows the mean and standard deviation (SD) of the solutions obtained by each algorithm based on 15 runs. The best outcomes highlighted in bold typeface.

As shown in Table 2, the proposed approach outperforms the basic SMO algorithm in most cases in terms of convergence rate, convergence speed, and global optimization capability.

5 Conclusions

This work proposes two modifications to address the shortcomings of the basic SMO algorithm. The first modification improves the distribution of the initial population using the GPS method during the initialization phase. The second modification modifies the LLP and GLP phases to account for the intensified and diversified breathing space for local searches. We evaluated the proposed algorithm on 20 standard benchmark functions. The simulation results show that the proposed algorithm outperforms the basic SMO algorithm and provides accurate and robust solutions. We aim to improve the performance of the modified SMO algorithm in a number of ways using powerful soft computing tools. By using soft computing tools, we can control the parameters of the SMO algorithm, guide the

search process, and develop hybrid algorithms to create algorithms that are more efficient and effective than the basic SMO algorithm.

Table 2: Comparison of test problem results

Function	D	Basic SMO		Modified SMO	
		Mean	SD	Mean	SD
Rosenbrock	10	967.95828	2311.63421	7.93559	0.34621
	30	28.66994	1.93780	27.33593	0.36470
sphere	30	2.24285e-13	8.39023e-13	2.89926e-17	8.26299e-18
	60	7.85583e-17	7.66269e-18	2.63342e-17	7.62556e-18
Elliptic	30	5.75254e-17	9.29459e-17	2.78168e-17	7.55754e-18
	60	7.63723e-17	7.43771e-18	3.06328e-17	4.73009e-18
Sum Squares	30	7.62003e-10	2.85115e-09	3.04396e-17	7.97876e-18
	60	7.02444e-17	1.20772e-17	3.74735e-17	6.21793e-18
Quartic	30	1.16326e-23	4.03573e-23	1.49493e-23	4.31037e-23
	60	2.04903e-23	4.45292e-23	1.15639e-22	1.99527e-22
Kowalik	4	0.00105094	0.0006879576	0.00032137	6.26241e-05
Schaffer's F6	10	0.03602	0.01491	0.00906	0.00242
	30	0.03327	0.00898	0.00972	1.03182e-09
Rastrigin	30	114.81071	19.08043	0	0
	60	103.08621	67.12331	0	0
Griewank	30	1.39276e-06	4.74962e-06	0	0
	60	9.75133e-11	2.09656e-10	0	0
Ackley	30	9.37203e-12	3.50519e-11	3.28626e-15	1.42108e-15
	60	4.23365e-15	8.86203e-16	3.99681e-15	0
Himmelblau	100	- 78.3323	9.07×10^{-12}	- 78.3323	3.48×10^{-14}
Step	30	0	0	0	0
Beale	2	1.33×10^{-13}	4.00×10^{-16}	4.91×10^{-8}	2.36×10^{-8}
Easom	2	-1	0	-1	9.05×10^{-13}
Schwefel	30	3.82×10^{-4}	7.28×10^{-13}	3.82×10^{-4}	6.78×10^{-13}
Levy N.13	2	9.03×10^{-20}	7.79×10^{-20}	1.97×10^{-20}	1.72×10^{-20}
Goldstein-Price	2	3	4.81×10^{-8}	3	1.1×10^{-15}
Colville	4	5.18×10^{-2}	5.59×10^{-2}	8.24×10^{-3}	1.83×10^{-2}
Booth	2	1.75×10^{-18}	1.92×10^{-12}	1.33×10^{-18}	1.15×10^{-18}
Dekkers and Aarts	2	-24776.52	7.27×10^{-12}	-24776.52	7.28×10^{-12}

References

- [1] M. Clerc, Particle swarm optimization. John Wiley & Sons, 2010.
- [2] M. Dorigo and T. Stützle, "Ant colony optimization: overview and recent advances," Handbook of metaheuristics, pp. 311-351, 2019.
- [3] M. Shehab, A. T. Khader, and M. A. Al-Betar, "A survey on applications and variants of the cuckoo search algorithm," Applied Soft Computing, vol. 61, pp. 1041-1059, 2017. <https://doi.org/10.1016/j.asoc.2017.02.034>
- [4] C. Guo, H. Tang, B. Niu, and C. B. P. Lee, "A survey of bacterial foraging optimization," Neurocomputing, vol. 452, pp. 728-746, 2021. <https://doi.org/10.1016/j.neucom.2020.06.142>
- [5] X.-S. Yang and X. He, "Bat algorithm: literature review and applications," International Journal of Bio-inspired computation, vol. 5, no. 3, pp. 141-149, 2013. <https://doi.org/10.1504/IJBIC.2013.055093>
- [6] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," TECHNICAL REPORT-TR06 2005.
- [7] J. C. Bansal, H. Sharma, S. S. Jadon, and M. Clerc, "Spider monkey optimization algorithm for numerical optimization," Memetic computing, vol. 6, no. 1, pp. 31-47, 2014.
- [8] K. Gupta, K. Deep, and J. C. J. C. I. Bansal, "Improving the local search ability of spider monkey optimization algorithm using quadratic approximation for unconstrained optimization," vol. 33, no. 2, pp. 210-240, 2017. <https://doi.org/10.1111/coin.12081>

- [9] S. Kumar and R. Kumari, "Modified position update in spider monkey optimization algorithm," in *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)*, 2014: Citeseer.
- [10] S. Kumar, R. Kumari, and V. K. Sharma, "Fitness based position update in spider monkey optimization algorithm," *Procedia Computer Science*, vol. 62, pp. 442-449, 2015. <https://doi.org/10.1016/j.procs.2015.08.504>
- [11] K. Gupta and K. Deep, "Tournament selection based probability scheme in spider monkey optimization algorithm," in *Harmony Search Algorithm*: Springer, 2016, pp. 239-250. DOI: 10.1007/978-3-662-47926-1_23
- [12] U. Singh and R. Salgotra, "Optimal synthesis of linear antenna arrays using modified spider monkey optimization," *Arabian Journal for Science and Engineering*, vol. 41, no. 8, pp. 2957-2973, 2016.
- [13] P. R. Singh, M. Abd Elaziz, and S. Xiong, "Modified Spider Monkey Optimization based on Nelder–Mead method for global optimization," *Expert Systems with Applications*, vol. 110, pp. 264-289, 2018. <https://doi.org/10.1016/j.eswa.2018.05.040>
- [14] H. Liu, Z. Cai, and Y. Wang, "A new constrained optimization evolutionary algorithm by using good point set," in *2007 IEEE Congress on Evolutionary Computation*, 2007, pp. 1247-1254: IEEE. DOI: 10.1109/CEC.2007.4424613
- [15] S. F. Raheem and M. Alabbas, "Dynamic Artificial Bee Colony Algorithm with Hybrid Initialization Method," *Informatica*, vol. 45, no. 6, 2021. <https://doi.org/10.31449/inf.v45i6.3652>
- [16] M. Tang, W. Long, H. Wu, K. Zhang, and Y. A. W. Shardt, "Self-Adaptive Artificial Bee Colony for Function Optimization," *Journal of Control Science and Engineering*, vol. 2017, p. 13 pages, 2017. <https://doi.org/10.1155/2017/4851493>
- [17] Athraa Qays Obaid, Maytham Alabbas, "Hybrid Variable-Length Spider Monkey Optimization with GoodPoint Set Initialization for Data Clustering", *Informatica* 47 (2023) 67–78, <https://doi.org/10.31449/inf.v47i8.4872>.
- [18] S. F. Raheem and M. Alabbas, "Optimal k-means clustering using artificial bee colony algorithm with variable food sources length," *International Journal of Electrical Computer Engineering*, vol. 12, no. 5, 2022. DOI: 10.11591/ijece.v12i5.pp5435-5443.

Acne Vulgaris Detection and Classification: A Dual Integrated Deep CNN Model

Md Baharul Islam^{1,2,*}, Masum Shah Junayed^{2,3}, Arezoo Sadeghzadeh², Nipa Anjum⁴, Afsana Ahsan Jeny^{2,3}, A. F. M. Shahen Shah⁵

¹Department of CSE, Florida Gulf Coast University, Fort Myers, FL, USA

²Department of Computer Engineering, Bahcesehir University, Istanbul, Turkey

³Department of CSE, University of Connecticut, Storrs, CT, USA

⁴Department of CSE, Khulna University of Engineering & Technology, Khulna, Bangladesh

⁵Department of Electronics and Communication Engineering, Yildiz Technical University, Turkey

E-mail: bislam.eng@gmail.com

*Corresponding Author

Keywords: acne classification, pattern recognition, deep CNN model, skin disease, medical image analysis

Received: June 24, 2022

Recognizing acne disease and evaluating its type is vital for the efficacy of the medical treatment. This report collects a dataset of 420 images and then labels them into seven different classes by a well-experienced dermatologist. After a pre-processing step, including local and global contrast enhancement and noise removal by a smoothing filter, the dataset size is enhanced using augmentation. The images of the dataset and the augmented ones are all fed into a novel integrated dual deep convolutional neural network (CNN) model to recognize acne disease and its type by classifying it into seven groups. First, two CNN-based units are designed to extract deep feature maps, later combined in a feature aggregation module. The aggregated features provide rich input information and classify the acne by a softmax. The proposed architecture's optimizer, loss function, and activation functions are all tuned so that both CNN units are trained with minimum kernel size and fewer training parameters. Thus, the computational cost is minimized. Compared with three machine learning-based classifiers and five pre-trained models, our model achieves competitive state-of-the-art performance with an accuracy of 97.53% on the developed dataset.

Povzetek: Razviti sistem uporablja CNN model za prepoznavanje in klasifikacijo aken (acne vulgaris) z natančnostjo 97.53% na naboru 420 slik.

1 Introduction

Acne is an unwanted skin disease occurring in the pilosebaceous unit. It is most commonly seen on the face, forehead, chest, upper back, and shoulders. Acne is usually related to hormonal fluctuations during adolescence, although some adults continue to experience acne into their 40-50s, too (caused by oil and dead skin cells) [1]. According to the surveys conducted in [2, 3], around 80% of adolescents and young adults are affected by acne, and approximately 40-50 million Americans have acne problems. Acne causes pain, redness, bleeding, and many other physical problems. Its psychological and emotional effects on patients can be far worse than physical issues. The changes in the beauty of the skin's appearance result in several psychological problems such as anger, depression, anxiety, fear, shame, embarrassment, low self-esteem, poor self-image, etc. Acne also has negative impacts on the social life of the patients, e.g., lack of confidence, limited employment chances, social withdrawal, and suicidal tendencies at worst [7].

Dermatologists diagnose acne by a simple visual inspection based on comedones, pustules, nodules, cysts, etc. It is

a subjective diagnosis depending on the experts' experience and ability. There is no particular test for acne, and in only special and critical cases, the X-ray, CT scan, or MRI tests are suggested by dermatologists [40]. Some professionals occasionally employ dermoscopic images for clinical diagnosis [9]. However, these images are acquired by a non-invasive method which is time-consuming. Additionally, there are several skin analysis systems, e.g., VISIA from Canfield and ANTERA 3D from Miravex, which are also expensive and cannot always detect acne accurately. These types of equipment also require to be operated and analyzed by well-experienced experts. Due to the lack of dermatologists, especially in under-developing countries, people do not receive timely treatments for acne. Even in developed countries, it is estimated that for an appointment with a dermatologist, patients are asked to wait for an average of 32 days¹, which delays the treatment procedure. On the other hand, if some types of acne, such as cysts (caused by severe infection), are not cured on time, they are likely to turn into permanent scars, or they need to be surgically removed by

¹<https://www.firstderm.com/appointment-wait-time-see-dermatologist/>

dermatologists.

Hence, it is vital to provide an automated system to recognize acne and identify its type. In this case, patients can receive timely treatment without expensive equipment or expert help at the initial stage. Generally, providing such a system is challenging due to several main reasons: (a) high diversity of acne lesions and human skin tones, (b) significant variations in the size, shape, and position of acne, and (c) dependency on the age, gender, and skin types.

Skin disease detection, especially acne detection based on deep neural networks and machine learning techniques, has attracted much attention among researchers. Several approaches have been reported in the literature in the last decade [12, 16, 14, 34, 29]. Even some of them are for acne severity grading [27, 35, 39, 10]. However, limited research has been carried out in the acne classification [5, 32, 21] while it is a vital issue in getting appropriate treatment in the early stages. The main challenge of acne recognition systems is their inability to classify different types of acne vulgaris. On the other hand, acne classification approaches are also required to improve their performance.

Another main challenge in acne classification is the lack of an appropriate and publicly available dataset. One of the earliest datasets was used in [5] with 35 images in 5 different types (i.e., nail, comedones, papules, pustule, and nodule). However, these small numbers of images are not enough, especially for deep learning-based systems. Later, another dataset was developed in [32] with 3000 skin images in 7 classes (including normal skin, papule, cyst, blackhead, pustule, whitehead, and nodule).

Although this dataset is rich in the number of images and classes and sufficient for deep learning models, it is not publicly available. We have even tried to contact the corresponding author through email to access their dataset, but they are not reachable. Additionally, the dataset does not include two more acne types, i.e., excoriated and keloidalis acne. Another dataset was proposed in [21] (collected from <http://dermnet.com>) with 300 images in 5 classes (i.e. closed comedo (whitehead), cystic, keloidalis, open comedo (blackhead), pustular). Although it includes keloidalis acne and has a reasonable number of images, it still suffers from a lack of excoriated acne and has limited classes.

In [29], another dataset was utilized with total 871 images (in 4 classes). This dataset also has a limited number of classes and is not publicly available. Recently, a publicly available dataset has been developed in [35] called "ACNE04," which is suitable only for grading acne in 4 classes (i.e., mild, moderate, severe, and very severe), not for acne type classification.

A new acne dataset with seven classes has been developed to address the abovementioned challenges. Then, an acne recognition and classification system are proposed based on an integrated dual deep CNN model. Overall, the main contributions of this work can be summarized as follows:

- Due to the unavailability of the public acne classifica-

tion dataset, an acne dataset is created, including 420 acne images in 7 classes. Contrary to the datasets proposed in [32] and [21], our developed dataset contains two different types of acne, i.e., Keloids and Excoriated. So, suppose the dataset of [32] is publicly available. In that case, its combination with ours results in a dataset with nine classes which can be significantly valuable for further research and analysis on acne classification problems.

- Providing an acne disease-free skin (DF) class (including normal skin and the skin with other diseases such as Eczema, Skin Cancer, Fungal Infection, etc.) enables our method not only to classify acne into six different types but also to distinguish it from the other skin diseases.
- The images of the dataset are captured using a smartphone camera. As smartphones are extensively available and used by almost everyone, the proposed method can be considered a remote screening system that ordinary people can efficiently utilize without needing expensive equipment and expert help.
- An integrated dual CNN-based automatic acne recognition and classification system is proposed. In this model, the extracted feature maps from two CNN models (without fully connected layers) are concatenated and aggregated, resulting in comprehensive rich information from input images and so high acne classification accuracy. The kernel size, the training parameters, and the number of layers is minimized by adjusting the optimizer, loss function, and activation functions, which reduces computational cost while the accuracy is still high.
- The performance of the proposed method is compared with the results of three machine learning-based classifiers, i.e., KNN, SVM, and MLP, and five pre-trained deep learning-based models, i.e., GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet. Our model receives competitive performance in the recognition and classification of acne.

The rest of the paper is organized as follows: Section 2 reviews the related works reported in the literature for acne detection, classification, and grading. The developed dataset, its setup, pre-processing, and augmentation steps are all explained in Section 3. Our proposed model is presented in detail in Section 4. Experimental results and the related discussions and comparisons are demonstrated in Section 5. Finally, conclusions are drawn in Section 6 as well as the future directions.

2 Related works

Many approaches have been reported in the literature for acne detection, classification, and grading in the last

decade. Generally, these approaches can be categorized into two main groups: conventional computer vision- and machine learning-based and deep learning-based methods. This section summarizes some of the primary studies in these two groups.

2.1 Conventional approaches

Conventional automatic acne detection/recognition methods are primarily based on feature extraction for classification. Malik et al. [27] proposed an acne grading system based on feature extraction and support vector machine (SVM). They classified the acne severity into four classes, i.e., mild, moderate, severe, and very severe. Khongsuwan et al. [24] proposed a method for counting the number of points for acne vulgaris. Their system achieved a prediction accuracy of 83.75% on the cropped part of the skin. The Ultra-Violet (UV) fluorescence light was applied to capture images before converting them to Gray-Scale and RGB. In this method, the quality of the images was improved using adaptive histogram equalization, and the number of points (acne) was counted based on the extended maxima transform. This image processing technique can quickly analyze acne images, but it would become tricky when the number of images is high.

Later, Alamdari et al [5] developed a mobile application for acne detection, classification and segmentation. They collected 35 images in five classes from various dermatology resources. They used Fuzzy C-Min (FCM) to classify images with no acne (normal skin) from the skin images with acne disease (with an accuracy of 100%). Moreover, they performed another classification task based on SVM (with the linear kernel) and fuzzy c-means techniques to distinguish acne scarring from inflammatory acne. This classification task achieved an average accuracy of 80% and 66.6% for FCM and linear SVM methods, respectively. As the images of their dataset were captured by a cell-phone, they suffered from difficulty visualizing the small lesion and ascertaining the depth of involvement. Additionally, they used a limited number of images for segmentation and classification, i.e., 35 images, while more images from more subjects are required to provide an accurate evaluation.

Kittigul et al. [25] detected acne based on robust applied features and then classified it using five designed features. They achieved an average accuracy of 68%, which is insufficient for clinical purposes. Hameed et al. [13] presented a hybrid technique using Naive Bayes Classifier (NBC) and image processing to detect and classify acne into three different types. Using 40 images in each of the three classes, they achieved an accuracy of 93.42%. Acne patterns were segmented using adaptively regularized fuzzy c-means (ARFCM) clustering technique and Morphological opening, creating the mask for all training images. Fourteen Haralick features were extracted from all patterns of the masks, which were later fed into NBC to perform the classification.

Although numerous approaches have been proposed based on traditional image processing techniques, they still suffer from noise and low accuracy due to the variations in characteristics of acne vulgaris, such as color variations and color complexity.

2.2 Deep learning-based methods

Deep neural networks (DNNs) such as CNNs have been extensively used for image classification. According to their high recognition and classification ability, sometimes they can perform even better than human beings in specific tasks, such as traffic sign recognition, face recognition, and handwriting digit recognition [30, 15]. Contrary to conventional computer vision methods, CNN-based models extract more and deeper features which enhance the classification accuracy and enable the system to deal with more classification types [11]. Hence, they are the focus of interest among the researchers not only for acne detection and classification but also in every field of medical image analysis [23, 8, 26, 22, 38, 28, 17, 33, 18, 19, 20].

Shen et al. [32] proposed a new automatic CNN-based diagnosis method for facial acne vulgaris to classify different types of acne vulgaris. This method extracted image features based on CNNs, classified by a classifier. The skin area was detected by applying a binary classifier for skin and non-skin classes. Then, the type of acne was determined using a seven-class classifier. Zhao et al. [39] proposed a grading system to assess the severity of facial acne vulgaris using 4,700 selfie images in 5 groups from "clear" to "severe." Based on the transfer learning approach, the features of the images were extracted using a pre-trained model (ResNet 152). Then, the target severity level was learned from the labeled images by adding and training a fully connected layer. The irrelevant background was minimized using OpenCV models to find facial landmarks. Key skin patches were extracted from the selfie images based on these landmarks. They trained their model after rolling each skin patch to improve testing results. The Root Mean Squared Error (RMSE) was 0.482 when applying the skin patch rolling data augmentation.

Junayed et al. [21] utilized Deep Residual Neural Network to build a model called "AcneNet" in which 1800 acne images (original images plus the augmented ones) in five classes were used. Training, validation, and testing accuracy were 86.28%, 86.11%, and 95.89%, respectively. Two pre-trained models, Inception-V3 and MobileNet, were also implemented on the same dataset and compared with their proposed method. Although this method was slightly underfitting, its performance was competitive.

Alom et al. [6] worked on skin cancer segmentation and classification using dermoscopic images. They proposed NABLA-N Net based on the R2U-Net model, which was composed of three different architectures: NABLA-N Net (A), NABLA-N Net (B), and NABLA-N Net (AB). An Inception Recurrent Residual Convolution Network (RRCNN) was used for recognizing skin cancer from der-

moscopic images. They used the transfer learning technique with NABLA-2 Net (AB) and got a testing accuracy of 96.36% without applying any augmentations. This accuracy was increased to 96.03% by employing augmentation along with transfer learning. They classified images into seven classes and got a testing accuracy of 81.12% without using data augmentation. This accuracy was also increased to 87.09% after applying data augmentation. Although this model was not proposed to detect acne, it performed well in skin cancer segmentation and classification.

Another automatic diagnosis system for skin disease was proposed by Shanthi et al. in [31] based on AlexNet architecture. They used the DermNet dataset in which 105 images were used for training (its 10% was taken as validation) and 69 images for testing. The dataset had four classes: Acne, Urticaria, Eczema Herpeticum, and Keratosis. They obtained 96.32% training accuracy and 62.1% validation accuracy. Testing accuracy for each type of Acne, Keratosis, Eczema Herpeticum, and Urticaria was achieved as 85.7%, 92.3%, 93.3%, and 92.8%, respectively.

Rashataprucksa et al. [29] tried to overcome the weak performance of traditional image processing techniques in acne detection and classification. They compared the performance of the Faster Region-based Convolution Neural Network (Faster R-CNN) and Region-based Fully Convolutional Network (R-FCN) on a dataset with 871 images (four classes of acne). Achieving a mean average precision of 28.3% for R-FCN, they proved that R-FCN performed comparatively better than Faster R-CNN. Although this method was more accurate and faster than traditional image-processing methods, its accuracy was still low for real-life clinical applications.

3 Dataset development

Identifying the type of acne is a crucial factor for having a successful treatment. One of the main challenges in automatic acne classification is providing a proper dataset. It is essential to have a dataset with a sufficient number of images and classes, especially for deep learning-based systems. A dataset of 420 pictures in 7 different types is proposed to overcome this challenge. These seven classes include Acne of Closed Comedo (ACC with 68 photos), Acne of Cystic (AC with 50 images), Acne of Excoriated (AE with 56 images), Acne of Keloidalis (AK with 71 images), Acne of Open Comedo (AOC with 53 photos), Acne of Pustular (AP with 62 images), and acne disease-free skin (DF with 60 images) (including the images of normal skin and the skin images with other diseases rather than acne).

The authors have captured seventy-seven images of this dataset from the subjects who visited the Department of Dermatology at Bangabandhu Sheikh Mujib Medical University (BSMMU) and Dhaka Medical College (DMC). The informed consent was obtained from all subjects before capturing the acne images by the 13-MP smartphone cam-

era. Some samples of these 77 images are illustrated in Fig. 2. An example for each of these acne types with the related descriptions is illustrated in Fig. 1. The rest of the images have been collected from public platforms of Baumann Cosmetic Dermatology (<http://www.derm.net/>) and New Zealand Dermatologists (<https://dermnetnz.org/>). All of the images in our dataset have been labeled by a well-experienced dermatologist in 7 classes. Due to different image sizes, all images are resized to 224×224 images.

4 Proposed CNN model

The main flowchart of the proposed method is illustrated in Fig. 3. This figure shows that the input images are fed into two CNN-based models after passing the pre-processing step and augmented. The extracted features from these dual CNN models are concatenated, aggregated using fully connected layers, and passed to the softmax classifier for classification. The details related to each step are presented in the following subsections.

4.1 Preprocessing and augmentation

Contrast is essential in medical imaging to better represent the images, especially for acne recognition and classification. Unlike the reported approaches in the literature in which only one contrast enhancement technique, i.e., local contrast or global contrast, was applied, a novel preprocessing technique is proposed in this paper through which local and global contrasts are incorporated. The primary goal of combining the local and global contrasts is to create an informational image that clearly shows the acne's location while simultaneously improving the image quality. A novel statistical function is generated to enhance the local contrast of the acne images in the region of $Q(m, n)$ by using the local mean of (E) and local standard deviation of (σ). The following equations (1 and 2) calculate the E and σ , respectively:

$$E(Q(m, n)) = \frac{1}{(2u + 1)^2} \sum_{m=0}^h \sum_{n=0}^w (Q(m, n)) \quad (1)$$

$$\sigma = \sqrt{\frac{1}{(2u + 1)^2} \sum_{m=0}^h \sum_{n=0}^w [Q(m, n) - E(Q(m, n))]^2} \quad (2)$$

where $(2u + 1)^2$ and $E(Q(m, n))$ represent the local contrast and the mean of the original input image, respectively. Here, $Q(m, n)$ denotes a region with the height of m and the width of n in 3 channels of RGB, i.e. $(m \times n \times 3)$, in which $(m, n) \in \mathbb{R}$. A statistical function that utilizes these parameters is described as:

$$Q_L(m, n) = E(Q(m, n)) + \phi[Q(m, n) - E(Q(m, n))] \quad (3)$$







Type/Photo	Description	Type/Photo	Description
ACC 	It seems as though tiny pimples have risen to the surface of the skin. Unlike common pimples, these lesions are not inflammatory and do not pain. When a clogged comedo forms, it is because a block of skin cells and oil got stuck within the hair follicle.	AK 	Acne Keloidalis is an inflammatory disease that lasts a long time. It can cause blisters and lesions on the back of the neck, which can progress to scars.
AC 	Acne Cystic is the most severe kind of acne. These are formed when cysts grow deep beneath the skin. Bacteria, oil, and dry skin cells may gather in pores, creating an environment that encourages acne.	AOC 	Whitehead is essentially the same as folliculitis, except that material within the follicle gets oxidized, and the follicles become blocked. Its sebaceous glands get blocked by excessive sebum. Pores filled with blackheads are known as open comedones.
AE 	A kind of acne caused by an atypical circumstance is excoriated acne. To peel off the skin is the meaning of the term excoriate. Blotchy, picky, and touchy acne is a characteristic of overwhelming urges to pinch or squeeze every pore on the skin, even minor blemishes.	AP 	Papules are almost identical to pustules. However, pustules have a white or yellow substance on the surface that looks like sebum.

Figure 1: Six different types of acne, including Acne of Closed Comedo (ACC), Acne of Cystic (AC), Acne of Excoriated (AE), Acne of Keloidalis (AK), Acne of Open Comedo (AOC), and Acne of Pustular (AP), and their characteristics.

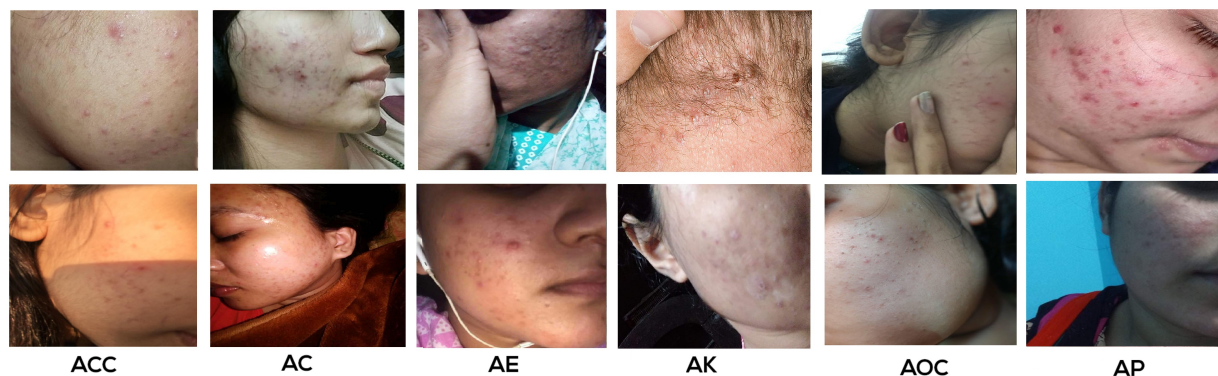


Figure 2: A sample of collected acne dataset

Here, the local contrast-enhanced image and the contrast gains (range greater than 1) are represented by $Q_L(m, n)$ and ϕ , respectively. Afterward, a top-hat maximization technique is used to enhance the global contrast. The top-hat filter operation is accomplished and stated as follows:

$$Q_{top}(m, n) = Q(m, n) - Q(m, n) \circ s_e \quad (4)$$

In this step, \circ and s_e denote the opening operation and the structural element, respectively. The opening operation is employed here to boost the global contrast, and the procedure is simple enough to be completed in a minimal amount of time. The output of $Q_{top}(m, n)$ is fed into the maximizing function, which creates the improved picture. The final enhanced image is created by combining the outputs of $Q_g(m, n)$ and $Q_{con}(m, n)$ calculated by equations 5 and 6, respectively:

$$Q_g(m, n) = \max_{z \in \Delta(m)} \left(\max_{\alpha \in ((m, n))} (Q_{top}(m, n)) \right) \quad (5)$$

$$Q_{con}(m, n) = \sum (Q_g(m, n), Q_L(m, n)) - Q(m, n) \quad (6)$$

It is worth mentioning that, in addition to the contrast enhancement, all images pass a smoothing filter through which their probable noises are removed.

Providing a considerable amount of training data is critical in Deep Learning (DL)-based models. If the dataset number is enormous, the overall model has a very flexible function with many tunable parameters for training. Additionally, increasing the number of training data in CNN reduces the probability of overfitting, generalizes the model to different input patterns, and so makes it robust [37]. Thus, to take advantage of the essential training data, seven other augmentation techniques, i.e., scaling, flipping horizontally, rotating 30° randomly to the right or left, shading, padding, affine transformation, and translation, are employed. These augmentations enhance the number of images in the dataset by producing additional images equal to seven times the original set.

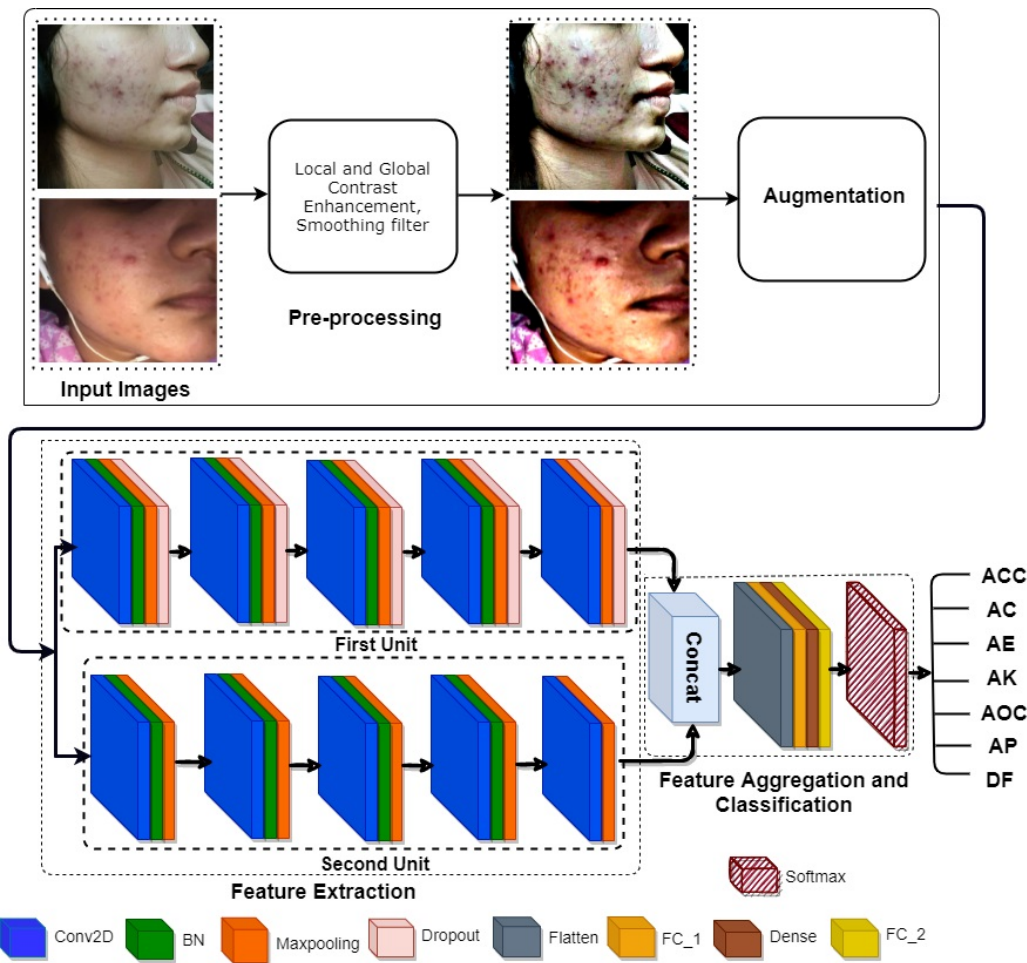


Figure 3: An overview of the proposed deep CNN-based acne classification system for identifying and categorizing acne vulgaris. In this system, the input images are fed into an integrated dual-CNN model after applying pre-processing and augmentation.

4.2 Dual CNN-based feature extractor

With the emergence of high GPUs and the considerable number of data, deep CNN-based models have been the focus of interest and extensively applied for detecting and classifying disease images in the last decade[36]. Deep neural networks with many layers obtain high accuracy for feature extraction and classification. However, increasing the number of layers raises the requirement for many training images, parameters, and high computational time. An integrated deep CNN-based feature extractor is proposed for acne recognition and classification to solve these limitations and get highly informative feature maps with fewer layers. In our proposed method, two separated CNN models (with no fully connected and classification layers), i.e., first and second units in Fig. 3, are designed and trained parallel for feature extraction.

The architecture details of these two CNN models, including the number of layers and the filter, kernel, and output sizes, are all summarized in Table 1. This table shows that both CNN units take the input images of size $224 \times 224 \times 3$ (RGB images with three channels). The

first unit comprises five convolution blocks, containing a convolutional layer, a Max Pooling (MP) layer (for reducing the space size for data representation), and two regularization layers, i.e., a batch normalization layer (BN) and a dropout layer. Regularization is a strategy to improve the model by changing the learning algorithm. It also improves the model’s performance on invisible information, reduces overfitting, and enhances generalization with improved convergence. In this unit, the filters in 5 blocks are 16, 32, 64, 96, and 128, respectively. The kernel size (KS) in all convolutional and MP layers are 3×3 and 2×2 , respectively. The padding for the first three convolutional layers is applied as "same" and for the rest two convolutional layers as "valid." ReLu activation function is used in all of the convolutional layers as follows:

$$RELU(x) = MAX(0, X) \tag{7}$$

The negative values of the matrix are considered 0, and the positive values are kept unchanged. Five dropout layers in 5 blocks are set as 0.25, 0.25, 0.4, 0.4, and 0.25, which means 25%, 25%, 40%, 40%, 40%, and 25% of neurons in

Table 1: The summary of the proposed two CNN feature extractors, including layers, their configurations, and output shape.

First Unit					Second Unit				
Layers	Filter	Configuration	Stride	Output Shape	Layers	Filter	Configuration	Stride	Output Shape
Conv2D	16	KS: 3 × 3; padding: same; ReLU	2	224 × 224 × 16	Conv2D	32	KS: 7 × 7; padding: same; ReLU	1	224 × 224 × 32
BN	-	-	-	224 × 224 × 16	BN	-	-	-	224 × 224 × 32
MP	-	KS: 2 × 2	2	112 × 112 × 16	MP	-	KS: 2 × 2	2	112 × 112 × 32
Dropout	-	0.25	-	112 × 112 × 16	Conv2D	32	KS: 5 × 5; padding: same; ReLU	1	112 × 112 × 32
Conv2D	32	KS: 3 × 3; padding: same; ReLU	2	112 × 112 × 32	BN	-	-	-	112 × 112 × 32
BN	-	-	-	112 × 112 × 32	MP	-	KS: 2 × 2	2	56 × 56 × 32
MP	-	KS: 2 × 2	2	56 × 56 × 32	Conv2D	48	KS: 3 × 3; padding: same; ReLU	1	56 × 56 × 48
Dropout	-	0.25	-	56 × 56 × 32	BN	-	-	-	56 × 56 × 48
Conv2D	64	KS: 3 × 3; padding: same; ReLU	1	56 × 56 × 64	MP	-	KS: 2 × 2	2	28 × 28 × 48
BN	-	-	-	56 × 56 × 64	Conv2D	64	KS: 3 × 3; padding: same; ReLU	2	28 × 28 × 64
MP	-	KS: 2 × 2	2	28 × 28 × 64	BN	-	-	-	28 × 28 × 64
Dropout	-	0.4	-	28 × 28 × 64	MP	-	KS: 2 × 2	2	14 × 14 × 64
Conv2D	96	KS: 3 × 3; padding: valid; AF: ReLU	1	28 × 28 × 96	Conv2D	128	KS: 3 × 3; padding: same; ReLU	2	14 × 14 × 128
BN	-	-	-	28 × 28 × 96	MP	-	KS: 2 × 2	2	7 × 7 × 128
MP	-	KS: 2 × 2	2	14 × 14 × 96					
Dropout	-	0.4	-	14 × 14 × 96					
Conv2D	128	KS: 3 × 3; padding: valid; ReLU	1	14 × 14 × 128					
MP	-	-	-	7 × 7 × 128					
Dropout	-	0.25	-	7 × 7 × 128					

hidden layers. These are set to 0 at each training phase update to prevent the model from overfitting while improving the accuracy. The output of the first CNN unit is a set of feature maps produced from the last dropout layer and has a high level of detail on acne disorders useful for acne type classification.

Similarly, the second unit contains five convolution blocks, each composed of a convolutional layer, a BN layer, and an MP layer, but no dropout layers. Another difference with the first unit is its filters 32, 32, 48, 64, and 128 in five blocks, respectively. The kernel sizes in the first and second convolutional layers are 7×7 and 5×5 , respectively. In the rest three convolutional layers, the kernel size is 3×3 . Padding in all convolutional layers is applied as "same." Other characteristics of the second unit, such as the kernel size in the MP layer and the activation function, are the same as the first unit. The second unit provides different feature maps from the first one.

4.3 Feature aggregation and classification

The acquired sets of feature maps from two CNN-based feature extractors have specific information about the input images. Consequently, their combination forms a comprehensive feature map, resulting in rich, robust, deep information from the inputs and high classification accuracy. As illustrated in Fig. 3, these two sets of feature maps are first assembled in the concatenation layer to obtain powerful feature aggregation and high-dimension feature representation with fewer semantic correlations. More discriminative shape information is provided by aggregating all features using a flattened layer, two fully connected layers (FC_1 and FC_2), and a dense layer. The number of neurons employed in FC_1 , dense, and FC_2 layers are 1048, 128, and 512, respectively. The final aggregated features are fed into a softmax classifier to recognize and classify the acne disease.

To better predict and classify, categorical cross-entropy is employed as the loss function, and the model is trained using the ADAM optimizer. The multi-class cross-entropy

loss function is defined as follows:

$$Loss = - \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (8)$$

where $y_i = \begin{cases} 1, & \text{if the element is in class } i \\ 0, & \text{otherwise} \end{cases}$ and \hat{y}_i is the probability that the element is in class i . The minus sign shows that the loss value gets smaller as the distributions become closer. Adam Optimizer helps the CNN model minimize errors, making it more reliable and efficient. In our proposed model, we use the automatic learning rate reduction technique. The initial learning rate is 0.0003.

5 Experimental results and discussion

In this section, the experimental setup and results are presented. Additionally, the performance of our proposed method is compared with three conventional machine learning classifiers and five pre-trained models, i.e., GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet, on our developed dataset.

5.1 Experimental setup

All experiments presented in this paper are carried out using an intel core i9 PC with a 3.60 GHz CPU, 64 GB RAM, and Nvidia Geforce Rtx 2080 super GPU with 8 GB video RAM. All training and testing are conducted in an Anaconda python environment with a visual code editor using Keras and TensorFlow frameworks. The number of epochs is defined as 60 with a batch size of 32.

Evaluation metrics

The performance of the proposed method is evaluated in terms of accuracy ($\frac{TP+TN}{(TP+TN+FP+FN)}$), Precision ($\frac{TP}{(TP+FP)}$), Recall/Sensitivity ($\frac{TP}{(FN+TP)}$), F1-

score $\left(\frac{2*(Precision*Recall)}{(Precision+Recall)}\right)$, Specificity $\left(\frac{TN}{(FP+TN)}\right)$, and Matthews Correlation Coefficient $(MCC = \frac{(TP*TN)-(FP*FN)}{\sqrt{((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))}})$ Score. In these evaluation metrics, TP, TN, FP, and FN stand for True Positive, True Negative, False Positive, and False Negative, respectively. True Positive refers to the correct prediction done by the classifier when the actual class of the data and the predicted class are both 1 (True). On the other hand, when the actual class is 0 (False) and the predicted class is also 0 (False), it is considered True Negative. In False Positive, the actual class of the data is 0 (False), while the classifier predicts it as 1 (True). It is named False because the model mispredicted the class and Positive due to the predicted class being 1 (True). Conversely, a False Negative happens when the actual class is 1 (True) and the predicted one is 0 (False). Similarly, False shows the misclassification, and Negative refers to the predicted class as 0 (False).

Accuracy, as the classification rate, is defined as the number of correct predictions divided by the total number of predictions. Recall (Sensitivity) is the true positive rate that observes the actual positive values correctly identified. The precision determines the number of positive class predictions which belong to the positive class, while the F1-score is the consonant mean of Precision and Recall, which measures testing accuracy. MCC measures the accuracy of the classifier by comparing observed and expected results.

Table 2: Performance comparison of our method with and without applying data augmentation in terms of sensitivity, specificity, and accuracy.

Dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)
Without Augmentation	70.11	82.76	82.18
With Augmentation	91.42	98.56	97.53

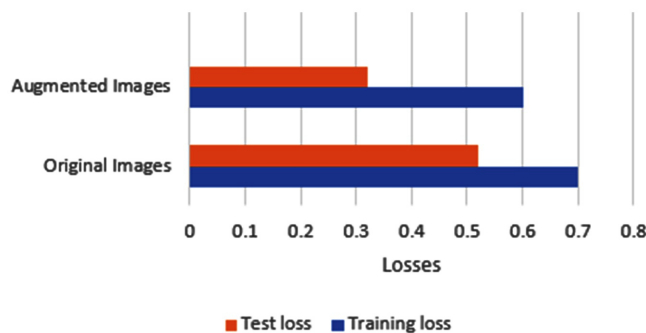


Figure 4: Performance comparison of our method with and without applying data augmentation in training and test loss.

5.2 Performance assessment

The performance of the proposed method for the automatic acne classification is evaluated on the developed dataset in

terms of sensitivity, specificity, and accuracy. The generated dataset’s images are resized to 224×224 RGB images as input. Then, a 10-fold cross-validation strategy is employed for two dataset scenarios: without and with data augmentation. Hence, the whole dataset without augmentation (420 images) is randomly divided into ten equal-size subsamples. Among them, a single subsample is selected as a testing set (i.e., 10% of the whole dataset, which is 42 images), and the remaining nine subsamples (i.e., 90% of the dataset) are used as the training set. This process is repeated ten times, while each of the ten subsamples is used exactly once as the testing set during the whole validation process. Similarly, a 10-fold cross-validation strategy is also applied for the dataset with augmentation. Still, this time the number of images is increased to 3360 (i.e., 420 original images plus 2940 augmented images based on seven augmentation techniques).

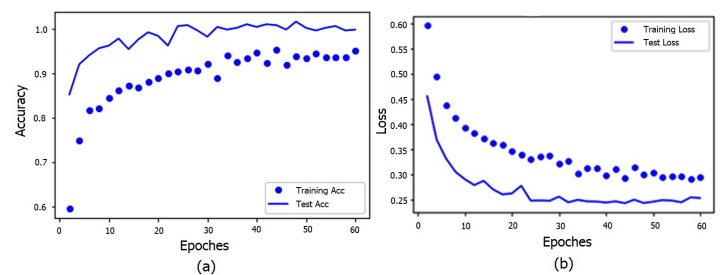


Figure 5: Performance of the proposed system on our developed dataset (with augmentation) in terms of (a) training and validation (test) accuracy, and (b) training and validation (test) loss.

The average results of 10-fold cross-validation for both scenarios are summarized in Table 2. As presented in the table, sensitivity, specificity, and accuracy are increased by 21.31% (from 70.11 to 91.42%), 15.8% (from 82.76% to 98.56%), and 15.35% (from 82.18% to 97.53%), respectively, for augmented images. Additionally, the impact of the augmentation on the performance of the proposed method is investigated in terms of training and test loss in Fig. 4. As illustrated in this figure, after augmentation, the training loss and test loss are decreased by 0.1 (from 0.7 to 0.6) and 0.2 (from 0.52 to 0.32), respectively. These results conclude that the augmentation prevents the model from overfitting and improves the system’s performance. Hence, in all experiments, the system’s performance is evaluated on the augmented dataset.

Moreover, the effects of the convolution block numbers in both CNN units are further investigated to select the optimum number. Twenty-five different combinations of convolution blocks are implemented with two optimizers (i.e., stochastic gradient descent (SGD) and ADAM) as presented in Table 3. According to the feature maps of each block (which are not shown here for brevity), the first convolution blocks usually detect and extract the edges of the images. As the number of blocks increases and the network becomes more profound, the feature maps look more like an

Table 3: The number of convolution blocks and the type of optimizers applied on the model.

No of Conv. Blocks in the First Unit	Filters of the First Unit	NO of Conv. Blocks in the Second Unit	Filters of the Second Unit	Optimizer	Accuracy
3	32, 64, 96	3	32, 48, 64	SGD	82.31
		4	32, 32, 48, 64	ADAM	83.63
		5	32, 32, 48, 64, 128	SGD	79.78
		6	16, 32, 32, 48, 64, 128	ADAM	76.92
		7	16, 32, 32, 48, 64, 128, 256	SGD	81.66
4	16, 32, 64, 128	3	32, 48, 64	ADAM	82.01
		4	32, 32, 48, 64	SGD	83.33
		5	32, 32, 48, 64, 128	ADAM	83.98
		6	16, 32, 32, 48, 64, 128	SGD	81.50
		7	16, 32, 32, 48, 64, 128, 256	ADAM	82.23
5	16, 32, 64, 96, 128	3	32, 48, 64	SGD	78.81
		4	32, 32, 48, 64	ADAM	83.19
		5	32, 32, 48, 64, 128	SGD	82.47
		6	16, 32, 32, 48, 64, 128	ADAM	81.62
		7	16, 32, 32, 48, 64, 128, 256	SGD	87.02
6	16, 32, 64, 64, 96, 128	3	32, 48, 64	ADAM	85.28
		4	32, 32, 48, 64	SGD	83.26
		5	32, 32, 48, 64, 128	ADAM	84.55
		6	16, 32, 32, 48, 64, 128	SGD	87.29
		7	16, 32, 32, 48, 64, 128, 256	ADAM	89.61
7	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	SGD	91.56
		4	32, 32, 48, 64	ADAM	93.83
		5	32, 32, 48, 64, 128	SGD	94.34
		6	16, 32, 32, 48, 64, 128	ADAM	95.17
		7	16, 32, 32, 48, 64, 128, 256	SGD	96.71
8	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	ADAM	97.53
		4	32, 32, 48, 64	SGD	95.09
		5	32, 32, 48, 64, 128	ADAM	94.18
		6	16, 32, 32, 48, 64, 128	SGD	93.41
		7	16, 32, 32, 48, 64, 128, 256	ADAM	95.62
9	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	SGD	95.01
		4	32, 32, 48, 64	ADAM	94.98
		5	32, 32, 48, 64, 128	SGD	95.27
		6	16, 32, 32, 48, 64, 128	ADAM	95.71
		7	16, 32, 32, 48, 64, 128, 256	SGD	95.55
10	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	ADAM	96.48
		4	32, 32, 48, 64	SGD	94.98
		5	32, 32, 48, 64, 128	ADAM	95.24
		6	16, 32, 32, 48, 64, 128	SGD	95.38
		7	16, 32, 32, 48, 64, 128, 256	ADAM	95.77
11	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	SGD	95.35
		4	32, 32, 48, 64	ADAM	96.50
		5	32, 32, 48, 64, 128	SGD	95.11
		6	16, 32, 32, 48, 64, 128	ADAM	94.96
		7	16, 32, 32, 48, 64, 128, 256	SGD	95.75
12	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	ADAM	96.10
		4	32, 32, 48, 64	SGD	95.83
		5	32, 32, 48, 64, 128	ADAM	95.75
		6	16, 32, 32, 48, 64, 128	SGD	96.00
		7	16, 32, 32, 48, 64, 128, 256	ADAM	95.89

abstract representation than the original image. The simple patterns, such as edges and shapes, are detected based on lower-level feature maps, while the high-level concepts are encoded using deeper feature maps. In our integrated dual-CNN feature extractor, the required features of acne and its types are extracted with five convolution blocks for both CNN units. As summarized in Table 3, using less than five convolution blocks provides fewer features insufficient for getting high accuracy. Although more features are extracted by increasing the number of convolution blocks, it does not necessarily always increase the accuracy. And instead, it leads to overfitting and false positives. Going deeper results in sparser feature maps (the filters detect fewer features). Consequently, the deeper feature maps provide more information about the class of the image than the image itself, which is helpful but less visually interpretable. In the first convolution blocks, simple shapes (available in every image) are detected, while the deeper networks seek more complex features that don't appear in every image. This phenomenon happens in our system when the number of convolution blocks in both CNN units is more than 5. Hence, the highest accuracy (96.71%

with SGD and 97.53% with ADAM) is achieved using five convolution blocks in both CNN-based feature extractors. Comparing the performance of our model based on two optimizers of SGD and ADAM, it is observed that ADAM achieves better results as it is an extension of SGD.

Hence, the performance of the final proposed system is presented in Figs. 5 (a) and (b) in terms of accuracy and loss, respectively, on the augmented dataset. The dotted line is related to the training set in both graphs, and the solid line is related to validation/test data. The X and Y axes in this figure demonstrate the number of epochs and accuracy/loss, respectively. It is observed that the model's performance is almost stable after 30 epochs and the training and test accuracy reach 94.81% (with a loss of 0.38) and 97.53% (with a loss of 0.29), respectively, after 60 epochs. Overall, the test set is well performed than the training set.

5.3 Comparison with classifiers

In this study, acne classification is carried out with conventional machine learning classifiers to demonstrate the capability of our proposed dual CNN-based acne classifi-

Table 4: Comparison of the proposed model performance with different machine learning-based classifiers.

Classifiers	Precision	F1-Score	Sensitivity	Specificity	Accuracy
KNN	72.35%	74.41%	75.89%	90.25%	90.10%
MLP	75.79%	76.52%	76.92%	93.45%	92.86%
SVM	78.14%	82.48%	80.82%	94.45%	94.06%
Proposed Softmax	91.37%	91.36%	91.42%	98.56%	97.53%

recognition system. The softmax classifier is replaced with machine learning-based classifiers. The extracted integrated feature maps from the dual-CNN feature extractor are fed into three classifiers: SVM, MLP, and KNN. Their tuned hyperparameters, i.e., initial learning rate, minimum batch size, learning algorithm, maximum epochs, and learning factor of f_c , are 0.0003, 32, ADAM, 60, and 10, respectively.

The performance of these classifiers is evaluated in terms of precision, F1-score, sensitivity, specificity, and accuracy and compared with our proposed model based on the softmax classifier in Table 4. As presented in this table, the accuracy of KNN, MLP, and SVM classifiers and our model is 90.10%, 92.86%, 94.06%, and 97.53%, respectively. Comparing the results, our proposed CNN model based on a softmax classifier achieves at least 3.5% more accuracy than the other classifiers. Additionally, its processing time is less than the different conventional classifiers.

5.4 Comparison with pre-trained models

Our proposed method can determine the type of acne and distinguish between acne and other skin diseases such as eczema and cancer. To further highlight our proposed acne classification capability, it is compared with five pre-trained models: GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet. Table 5 displays the performance of these models as well as ours in terms of accuracy, precision, F1-score, sensitivity, specificity, and MCC for each of the acne types (i.e., ACC, AC, AE, AK, AOC, AP, DF) and their average results. As presented in this table, the average accuracy of GoogleNet, MobileNet, VGG19, ResNet50, AlexNet, and our proposed model is 94.13%, 94.90%, 95.58%, 96.24%, 95.89%, and 97.53%, respectively, among which ours is the highest. Not only in terms of accuracy but also in terms of other evaluation metrics, i.e., precision (91.37%), F1-score (91.36%), sensitivity (91.42%), specificity (98.56%), and MCC (89.94%), our proposed method outperforms the others. Additionally, the accuracy of each class in our proposed method is higher than that of other methods. The highest accuracy belongs to the DF class, which refers to the acne-free skin images. As our developed dataset has the benefit of having a type containing normal skin and other skin disease images (except acne), our proposed method has successfully trained for recognizing acne disease. If the probe image does not contain acne, it is classified as DF with high accuracy of 99.40%.

Fig. 6 presents the confusion matrices of our proposed model and the other five pre-trained deep learning-based models. As demonstrated in this figure, the number of true positives in all seven classes is higher in our proposed model, which proves its competitive performance compared to GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet.

As another evaluation tool, our proposed method's Receiver Operating Characteristic (ROC) Curve and five pre-trained models are illustrated in Fig. 7. It presents the performance of the models at different thresholds, in which the x-axis is the false positive rate and the y-axis is the true positive rate. In this probability curve, the Area Under the ROC Curve (AUC) indicates the classification capability of the corresponding model. As illustrated in this figure, the AUC scores of GoogleNet, MobileNet, VGG-19, ResNet-50, AlexNet, and our proposed CNN are 91.32, 92.91, 91.76, 91.88, 93.24, and 94.67, respectively. These models are all implemented with the ADAM optimizer. As the higher AUC shows a better performance, our proposed method obtains the best performance having the highest AUC of 94.67.

5.5 Comparison with the state-of-the-arts

To have a fair comparison between our proposed method and the state-of-the-art approaches, we must implement them on the proposed dataset. No available source codes are found for the related works to implement them. Consequently, our model is only compared with one state-of-the-art acne classification approach proposed in [21] by implementing it on the same dataset accessible in <http://dermnet.com>. The results are presented in Table 6. As it is noted in this table, our proposed dual CNN-based acne classification system achieves higher performance (96.74%) than the state-of-the-art approach of [21] (with a reported accuracy of 95.89%) on the same dataset of 1800 acne images in 5 different classes.

Additionally, as the proposed model has competitive performance in acne classification, it also inspired us to evaluate it for acne grading. Hence, it is implemented on an acne grading dataset called "ACNE04" [35], which is publicly available. The grading system is implemented on the same dataset of 1457 acne images in 4 classes, i.e., mild, moderate, severe, and very severe. The results are also summarized in the same table (Table 6). Comparing the results, our proposed acne classification system can be successfully performed for acne grading by achieving higher accuracy of 86.36% which is 2.25% higher than that of the state-of-the-art approach in [35].

5.6 Failure cases

Although our proposed acne classification system can successfully recognize and classify acne into six different types, there are a few misclassification cases, as shown in Fig. 8 with its actual class and the predicted label. How-

Table 5: Comparison between the performance of the proposed model and the other five pre-trained models.

Methods	Disease types	Accuracy (%)	Precision (%)	F1-Score (%)	Sensitivity (%)	Specificity (%)	MCC (%)
GoogleNet	ACC	93.45	75.00	76.60	78.26	95.86	72.81
	AC	94.05	77.08	78.72	80.43	96.21	75.29
	AE	94.94	81.25	82.11	82.98	96.89	79.17
	AK	93.75	77.08	77.89	78.72	96.19	74.26
	AOC	94.35	79.17	80.00	80.85	96.54	76.71
	AP	90.48	72.92	68.63	64.81	95.39	63.19
	DF	97.92	93.75	92.78	91.84	98.95	91.57
Total Average		94.13	79.46	79.53	79.70	96.57	76.14
MobileNet	ACC	94.64	81.25	81.25	81.25	96.88	78.13
	AC	95.54	79.17	83.52	88.37	96.59	81.11
	AE	93.75	81.25	78.79	76.47	96.84	75.17
	AK	94.94	83.33	82.47	81.63	97.21	79.52
	AOC	94.94	79.17	81.72	84.44	96.56	78.84
	AP	92.26	77.08	74.00	71.15	96.13	69.54
	DF	98.21	93.75	93.75	93.75	98.96	92.71
Total Average		94.90	82.14	82.21	82.44	97.02	79.29
VGG-19	ACC	95.24	83.33	83.33	83.33	97.22	80.56
	AC	95.54	85.42	84.54	83.67	97.56	81.93
	AE	94.94	83.33	82.47	81.63	97.21	79.52
	AK	96.43	87.50	87.50	87.50	97.92	85.42
	AOC	94.35	77.08	79.57	82.22	96.22	76.35
	AP	93.75	79.17	78.35	77.55	96.52	74.70
	DF	98.81	95.83	95.83	95.83	99.31	95.14
Total Average		95.58	84.52	84.51	84.53	97.42	81.95
ResNet-50	ACC	96.43	87.50	87.50	87.50	97.92	85.42
	AC	94.64	79.17	80.85	82.61	96.55	77.76
	AE	95.24	85.42	83.67	82.00	97.55	80.91
	AK	96.54	81.25	83.87	86.67	96.91	81.34
	AOC	94.94	79.17	81.72	84.44	96.56	78.84
	AP	93.75	83.33	79.21	75.47	97.17	75.67
	DF	99.11	97.92	96.91	95.92	99.65	96.39
Total Average		96.24	84.81	84.82	84.94	97.47	82.33
AlexNet	ACC	95.54	81.25	83.67	86.67	96.91	81.34
	AC	95.24	83.33	83.33	83.33	97.22	80.56
	AE	96.73	89.58	88.66	87.76	98.26	86.75
	AK	97.02	93.75	90.00	86.54	98.94	88.35
	AOC	96.43	85.42	87.23	89.13	97.59	85.19
	AP	93.75	77.08	77.89	78.72	96.19	74.26
	DF	96.54	81.36	88.89	97.96	96.31	87.39
Total Average		95.89	84.54	85.67	87.16	97.35	80.41
Proposed model	ACC	97.92	89.58	92.47	95.56	98.28	91.33
	AC	96.13	87.50	86.60	85.71	97.91	84.34
	AE	96.73	89.58	88.66	87.76	98.26	86.75
	AK	97.92	93.75	92.78	91.84	98.95	91.57
	AOC	98.21	95.83	93.88	92.00	99.30	92.86
	AP	96.43	85.42	87.23	89.13	97.59	85.19
	DF	99.40	97.92	97.92	97.92	99.65	97.57
Total Average		97.53	91.37	91.36	91.42	98.56	89.94

Table 6: The performance comparison of the AcneNet and ACNE04 datasets in the model.

Datasets	No. of classes	Dataset Sizes	Accuracy
AcneNet [21] (Classification)	5	1800	95.89 [21]
			96.74 (Our)
ACNE04 [35] (Grading)	4	1457	84.11 [35]
			86.36 (Our)

ever, among these eight instances, two misclassifications (Figs. 8 (a) and (b)) has the low confidence score close to 0.5. Deeply analyzing the acne images of these misclassification cases, we draw the inferences that the proposed model mostly has difficulties classifying the tiny acne, as illustrated in Figs. 8 (d) and (h). Misclassifications with high confidence scores have occurred in Figs. 8 (c), (e), (f), and (g). Several reasons can be included, such as low image quality, a unique acne case, etc.

6 Conclusions and future works

This paper introduced a deep learning-based lightweight system to recognize and classify different acne vulgaris using a novel acne dataset. Firstly, an acne dataset with 420

images in 7 classes was developed. Then, these images were modified by applying a pre-processing system. The number of shots was increased to 3360 by using seven different augmentation methods. An integrated dual CNN-based model was proposed to recognize acne and classify it into seven groups. The whole feature extractor was composed of two CNN models with different conv2d, BN, and max-pooling layers. The extracted feature maps from these two models were first concatenated and then aggregated using fully connected layers. The final comprehensive feature maps were fed into a softmax layer for classification. The performance was investigated for convolution blocks in two feature extractor units and two different SGD and ADAM optimizers.

In addition, its performance was evaluated for both original images without augmentation and the extended dataset with augmented images to analyze the influence of augmentation. It was compared with three conventional machine learning-based classifiers and five pre-trained deep learning-based models and received competitive performance. The proposed method's feasibility has been confirmed by conducting experiments and achieving a state-of-the-art accuracy of 97.53% in acne classification. It was

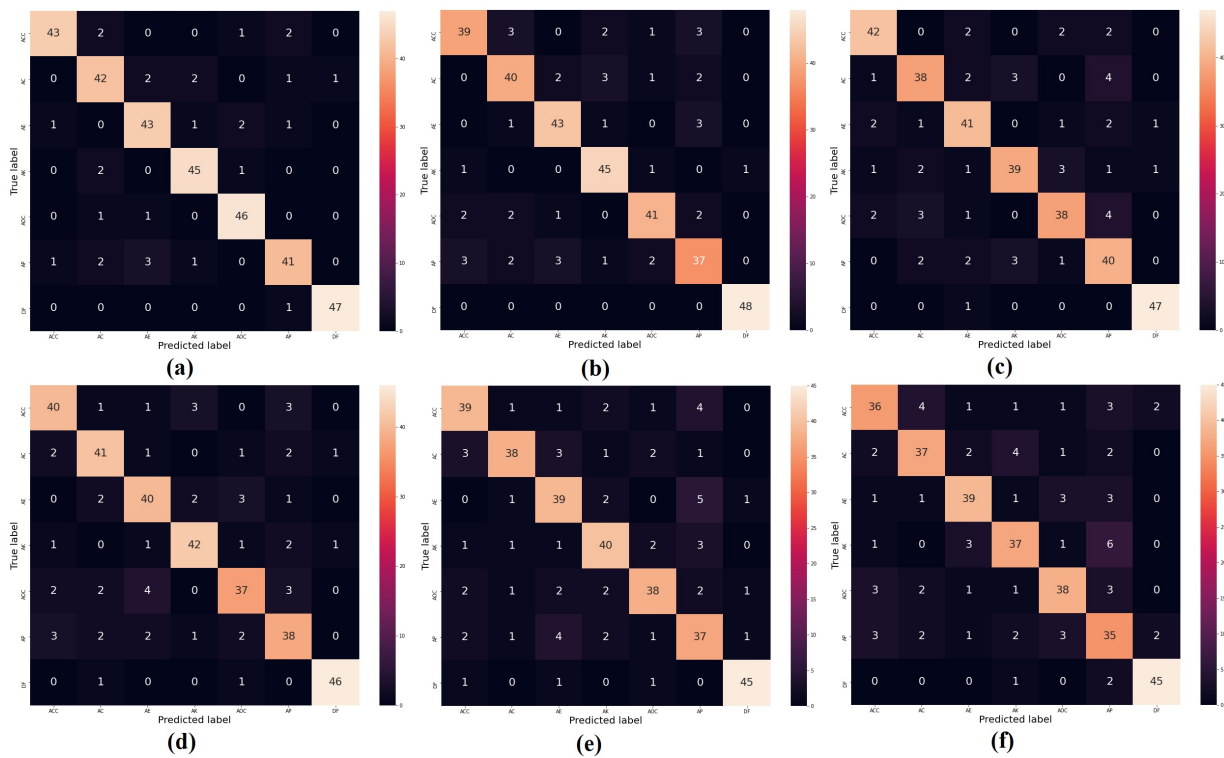


Figure 6: Confusion matrix for acne classification. Here, (a), (b), (c), (d), (e), and (f) represent the confusion matrix of the proposed model, AlexNet, ResNet-50, VGG-19, MobileNet and GoogleNet

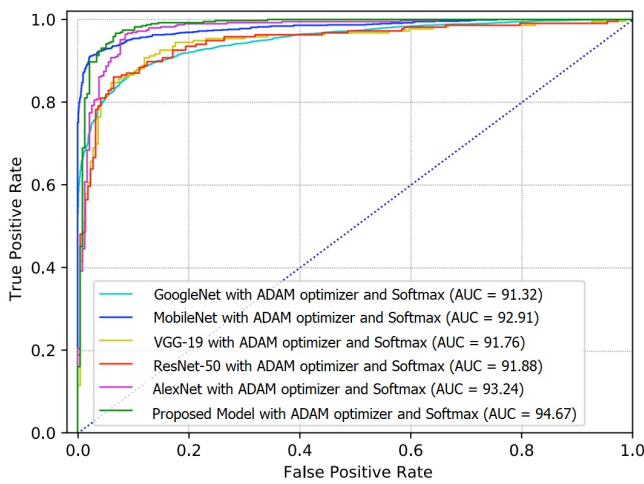


Figure 7: Comparison between the AUC of the ROC curves belonging to our proposed deep dual-CNN model and five pre-trained models.

also implemented on an acne grading dataset and achieved good performance with an accuracy of 86.36%. In the future, we want to expand our study with an updated and extended form of our proposed acne dataset with additional images and acne classes. As our proposed model is highly accurate for acne classification and computationally efficient, it will also be adjusted and applied to other similar dermatological disease classification and identification

tasks in our future research direction.

Experimental code availability

For further investigation, comparison, and analysis of this study by the research community, the experimental code and model are accessible upon request to the corresponding author through email.

Compliance with ethical standards

This article does not contain any studies with human participants and/or animals performed by any authors.

Conflict of interest

We (authors) certify that this article has no actual or potential conflict of interest.

Authors’ contributions

All authors contributed to this paper. **Afsana Ahsan Jeny, Masum Shah Junayed, Nipa Anjum:** Methodology, Experiment, Writing- Original draft preparation. **Md Baharul Islam:** Investigation, Conceptualization, Supervision, Writing- Reviewing and Editing, **Arezoo Sadeghzadeh, A. F. M. Shahen Shah:** Writing- Reviewing and Editing.



Figure 8: Misclassification cases of the proposed model.

References

- [1] Goulden V, Clark SM, Cunliffe WJ. Post-adolescent acne: a review of clinical features. *British journal of dermatology*. 1997 Jan;136(1):66-70. <https://doi.org/10.1046/j.1365-2133.1997.d01-1144.x>
- [2] Dharshana S, Singh A, Sharma S, Mohan S, Joshi A. Depression, mood change and self-esteem among adolescents aged 12-25 years with acne vulgaris in India. *Annals of Tropical Medicine and Public Health*. 2016;9(1). <https://doi.org/10.4103/1755-6783.168712>
- [3] Picardo M, Eichenfield LF, Tan J. Acne and rosacea. *Dermatology and therapy*. 2017 Jan;7(1):43-52. <https://doi.org/10.1007/s13555-016-0168-8>
- [4] Zaenglein AL. Acne vulgaris. *New England Journal of Medicine*. 2018 Oct 4;379(14):1343-52. <https://doi.org/10.1056/NEJMc1702493>
- [5] Alamdari N, Tavakolian K, Alhashim M, Fazel-Rezai R. Detection and classification of acne lesions in acne patients: A mobile application. In *IEEE international conference on electro information technology (EIT) 2016 May 19 (pp. 0739-0743)*. IEEE. <https://doi.org/10.1109/eit.2016.7535331>
- [6] Alom MZ, Aspiras T, Taha TM, Asari VK. Skin cancer segmentation and classification with NABLA-N and inception recurrent residual convolutional networks. *arXiv preprint arXiv:1904.11126*. 2019 Apr 25. <https://doi.org/10.48550/arXiv.1904.11126>
- [7] Ayer J, Burrows N. Acne: more than skin deep. *Postgraduate medical journal*. 2006 Aug 1;82(970):500-6. <https://doi.org/10.1136/pgmj.2006.045377>
- [8] Abayomi-Alli OO, Damasevicius R, Misra S, Maske-liunas R, Abayomi-Alli A. Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold. *Turkish Journal of Electrical Engineering and Computer Sciences*. 2021;29(8):2600-14. <https://doi.org/10.3906/elk-2101-133>
- [9] Barata C, Celebi ME, Marques JS, Rozeira J. Clinically inspired analysis of dermoscopy images using a generative model. *Computer Vision and Image Understanding*. 2016 Oct 1;151:124-37. <https://doi.org/10.1016/j.cviu.2015.09.011>
- [10] Chen L, Li Y, Han L, Yuan L, Sun Y, Tang X. Classification and Treatment System for Facial Acne Vulgaris Based on Image Recognition. In *Elderly Health Services and Remote Health Monitoring 2020 (pp. 65-71)*. Springer, Singapore. https://doi.org/10.1007/978-981-15-7154-1_6
- [11] Cireşan DC, Meier U, Masci J, Gambardella LM, Schmidhuber J. High-performance neu-

- ral networks for visual object classification. arXiv preprint arXiv:1102.0183. 2011 Feb 1. <https://doi.org/10.48550/arXiv.1102.0183>
- [12] Gu Y, Ge Z, Bonnington CP, Zhou J. Progressive transfer learning and adversarial domain adaptation for cross-domain skin disease classification. *IEEE journal of biomedical and health informatics*. 2019 Sep 23;24(5):1379-93. <https://doi.org/10.1109/jbhi.2019.2942429>
- [13] Hameed AZ, Awad WK, Irsan NA, Abdulbaqi AS. Hybrid technique for skin pimples image detection and classification. *International Journal of Horticultural Science and Technology*. 2020;29:4102-9. <http://sersc.org/journals/index.php/IJAST/article/view/5164>
- [14] Hameed N, Shabut AM, Ghosh MK, Hossain MA. Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. *Expert Systems with Applications*. 2020 Mar 1;141:112961. <https://doi.org/10.1016/j.eswa.2019.112961>
- [15] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision 2015* (pp. 1026-1034). <https://doi.org/10.1109/iccv.2015.123>
- [16] Jeny AA, Sakib AN, Junayed MS, Lima KA, Ahmed I, Islam MB. SkNet: a convolutional neural networks-based classification approach for skin cancer classes. In *23rd International Conference on Computer and Information Technology (ICCIT) 2020 Dec 19* (pp. 1-6). IEEE. <https://doi.org/10.1109/iccit51783.2020.9392716>
- [17] Jiang S, Li H, Jin Z. A visually interpretable deep learning framework for histopathological image-based skin cancer diagnosis. *IEEE Journal of Biomedical and Health Informatics*. 2021 Jan 15;25(5):1483-94. <https://doi.org/10.1109/jbhi.2021.3052044>
- [18] Junayed MS, Islam MB, Sadeghzadeh A, Rahman S. CataractNet: An automated cataract detection system using deep learning for fundus images. *IEEE Access*. 2021 Sep 15;9:128799-808. <https://doi.org/10.1109/access.2021.3112938>
- [19] Almoosawi NM, Khudeyer RS. ResNet-34/DR: a residual convolutional neural network for the diagnosis of diabetic retinopathy. *Informatica*. 2021 Dec 17;45(7). <https://doi.org/10.31449/inf.v45i7.3774>
- [20] Ullah Z, Odeh A, Khattak I, Hasan MA. Enhancement of Pre-Trained Deep Learning Models to Improve Brain Tumor Classification. *Informatica*. 2023. 47; 165–172. <https://doi.org/10.31449/inf.v47i6.4645>
- [21] Junayed MS, Jeny AA, Atik ST, Neehal N, Karim A, Azam S, Shanmugam B. AcneNet-a deep CNN-based classification approach for acne classes. In *12th International Conference on Information & Communication Technology and System (ICTS) 2019 Jul 18* (pp. 203-208). IEEE. <https://doi.org/10.1109/icts.2019.8850935>
- [22] Junayed MS, Sakib AN, Anjum N, Islam MB, Jeny AA. Eczemanet: A deep cnn-based eczema diseases classification. In *2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS) 2020 Dec 9* (pp. 174-179). IEEE. <https://doi.org/10.1109/ipas50080.2020.9334929>
- [23] Kaymak R, Kaymak C, Ucar A. Skin lesion segmentation using fully convolutional networks: A comparative experimental study. *Expert Systems with Applications*. 2020 Dec 15;161:113742. <https://doi.org/10.1016/j.eswa.2020.113742>
- [24] Khongsuwan M, Kiattisin S, Wongseree W, Leelasantitham A. Counting number of points for acne vulgaris using UV fluorescence and image processing. *Biomedical Engineering International Conference 2012 Jan 29* (pp. 142-146). IEEE. <https://doi.org/10.1109/bmeicon.2012.6172038>
- [25] Kittigul N, Uyyanonvara B. Acne detection using speeded up robust features and quantification using K-Nearest neighbors algorithm. In *Proceedings of the 6th International Conference on Bioinformatics and Biomedical Science 2017 Jun 22* (pp. 168-171). <https://doi.org/10.1145/3121138.3121168>
- [26] Loram I, Siddique A, Sánchez MB, Harding P, Silverdale M, Kobylecki C, Cunningham R. Objective analysis of neck muscle boundaries for cervical dystonia using ultrasound imaging and deep learning. *IEEE journal of biomedical and health informatics*. 2020 Jan 9;24(4):1016-27. <https://doi.org/10.1109/jbhi.2020.2964098>
- [27] Malik AS, Ramli R, Hani AF, Salih Y, Yap FB, Nisar H. Digital assessment of facial acne vulgaris. In *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings 2014 May 12* (pp. 546-550). IEEE. <https://doi.org/10.1109/i2mtc.2014.6860804>
- [28] Pour MP, Seker H. Transform domain representation-driven convolutional neural networks for skin lesion segmentation. *Expert Systems with Applications*. 2020 Apr 15;144:113129. <https://doi.org/10.1016/j.eswa.2019.113129>

- [29] Rashataprucksa K, Chuangchaichatchavarn C, Triukose S, Nitinawarat S, Pongpruthipan M, Piromsopa K. Acne detection with deep neural networks. In 2020 2nd International Conference on Image Processing and Machine Vision 2020 Aug 5 (pp. 53-56). <https://doi.org/10.1145/3421558.3421566>
- [30] Sermanet P, Chintala S, LeCun Y. Convolutional neural networks applied to house numbers digit classification. In Proceedings of the 21st international conference on pattern recognition (ICPR2012) 2012 Nov 11 (pp. 3288-3291). IEEE. <https://doi.org/10.48550/arXiv.1204.3968>
- [31] Shanthi T, Sabeenian RS, Anand R. Automatic diagnosis of skin diseases using convolution neural network. *Microprocessors and Microsystems*. 2020 Jul 1;76:103074. <https://doi.org/10.1016/j.micpro.2020.103074>
- [32] Shen X, Zhang J, Yan C, Zhou H. An automatic diagnosis method of facial acne vulgaris based on convolutional neural network. *Scientific reports*. 2018 Apr 11;8(1):1-0. <https://doi.org/10.1038/s41598-018-24204-6>
- [33] Son HM, Jeon W, Kim J, Heo CY, Yoon HJ, Park JU, Chung TM. AI-based localization and classification of skin disease with erythema. *Scientific Reports*. 2021 Mar 5;11(1):1-4. <https://doi.org/10.1038/s41598-021-84593-z>
- [34] Tang P, Liang Q, Yan X, Xiang S, Zhang D. GP-CNN-DTEL: Global-part CNN model with data-transformed ensemble learning for skin lesion classification. *IEEE journal of biomedical and health informatics*. 2020 Feb 28;24(10):2870-82. <https://doi.org/10.1109/jbhi.2020.2977013>
- [35] Wu X, Wen N, Liang J, Lai YK, She D, Cheng MM, Yang J. Joint acne image grading and counting via label distribution learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision 2019 (pp. 10642-10651). <https://doi.org/10.1109/iccv.2019.01074>
- [36] Yadav SS, Jadhav SM. Deep convolutional neural network-based medical image classification for disease diagnosis. *Journal of Big Data*. 2019 Dec;6(1):1-8. <https://doi.org/10.1186/s40537-019-0276-2>
- [37] Zhang C, Tavanapong W, Wong J, Groen PC, Oh J. Real data augmentation for medical image classification. In *Intravascular imaging and computer-assisted stenting, and large-scale annotation of biomedical data and expert label synthesis* 2017 Sep 14 (pp. 67-76). Springer, Cham. https://doi.org/10.1007/978-3-319-67534-3_8
- [38] Zhang L, Wang X, Yang D, Sanford T, Harmon S, Turkbey B, Wood BJ, Roth H, Myronenko A, Xu D, Xu Z. Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. *IEEE transactions on medical imaging*. 2020 Feb 12;39(7):2531-40. <https://doi.org/10.1109/tmi.2020.2973595>
- [39] Zhao T, Zhang H, Spoelstra J. A computer vision application for assessing facial acne severity from selfie images. *arXiv preprint arXiv:1907.07901*. 2019 Jul 18. <https://doi.org/10.48550/arXiv.1907.07901>
- [40] Zhao X, Hu X, Liao Y, He T, Zhang T, Zou X, Tian J. Accurate MR image super-resolution via lightweight lateral inhibition network. *Computer Vision and Image Understanding*. 2020 Dec 1;201:103075. <https://doi.org/10.1016/j.cviu.2020.103075>

Low-cost GNSS Receivers for Geodetic Monitoring Purposes

Veton Hamza

Faculty of Civil and Geodetic Engineering, Univeristy of Ljubljana

E-mail: veton.hamza@fgg.uni-lj.si

Thesis Summary

Keywords: GNSS, low-cost GNSS receivers, displacements, monitoring.

Received: November 15, 2023

This article is an extended abstract of the doctoral dissertation entitled “Cost-effective GNSS receivers for geodetic monitoring” [1]. For several years, geodetic Global Navigation Satellite System (GNSS) receivers have been argued as proper sensors to monitor natural hazards and engineering structures. Nevertheless, their use is questionable in areas where there is a risk of instrument damage considering their high costs. Therefore, low-cost GNSS receivers are considered as an alternative for such applications. Through various evaluation tests designed and conducted in the doctoral dissertation, substantial analyses for data and positioning quality, size of detected displacements, and application of low-cost GNSS receivers in geodetic monitoring were obtained. The developed Low-Cost GNSS Monitoring System (LGMS) represents a suitable solution for near real-time continuous geodetic monitoring with high accuracy and reduced costs. Furthermore, it can be applied in the geodetic monitoring of natural hazards like subsidences, sinking phenomena, and landslides, as well as, other engineering structures such as viaducts, bridges, dams, mining areas, and chimneys, provided it meets the accuracy standards.

Povzetek: Predstavljena je doktorska disertacija z naslovom »Uporaba cenovno ugodnih sprejemnikov GNSS za geodetski monitoring«.

1 Introduction

Low-cost Global Navigation Satellite System (GNSS) receivers are considered an alternative to geodetic counterparts, particularly for projects and applications constrained by budget limitations. Low-cost GNSS receivers are highly desirable for geodetic monitoring of natural hazards and engineering structures, as they enable cost-effective monitoring at numerous points of interest while offering several advantages. These advantages include affordability, compact size, low energy consumption, ease of replacement in case of damage, and flexibility for configuration and data processing using open-source software and applications.

The main objective of the doctoral dissertation was to investigate the capabilities of using low-cost GNSS receivers for developing a Low-cost GNSS Monitoring System (LGMS) with high accuracy and decreased costs that can continuously monitor displacements in near-real-time.

2 Methods

a. Quality analysis of GNSS observation

A series of tests were conducted under various conditions, including open-sky and adverse scenarios, to assess the quality of GNSS observations acquired with low-cost GNSS receivers [2,3]. As a reference for

comparison high-ended geodetic GNSS receivers were used. The results indicated that, in both open-sky and adverse scenarios, low-cost GNSS receivers provided lower data quality (worse results for carrier-to-noise ratio, multipath, phase noise, cycle slips, and others) compared to geodetic GNSS receivers [2,3]. However, the difference in cycle slips and phase noise was not very significant and it ensured first information for sufficient positioning quality even though the results from all tests were in favor of geodetic GNSS receivers [2].

b. Positioning performance evaluation

To perform a more comprehensive assessment of positioning quality for low-cost GNSS receivers, additional experimental tests were designed and performed. These tests encompassed a zero baseline test, a short baseline test, a comparison of coordinate precision, and 3D geodetic adjustment by using GNSS and Terrestrial Positioning System (TPS) observations [1,4].

The results from the zero baseline test confirmed that the low-cost GNSS receiver (ZED-F9P) has minor receiver noise (sub-millimeter error of estimated coordinates). In a short baseline test, the superiority of GNSS receivers was confirmed, while the precision of obtained horizontal and vertical positions was almost equal for geodetic and low-cost GNSS receivers in cases when long sessions (50h) were considered. However, geodetic GNSS receivers obtained higher positioning accuracy. The precision of obtained coordinates in short sessions (3h) was again in favor of geodetic GNSS

receivers. The 3D geodetic adjustment of GNSS and TPS observations was shown to improve the positioning accuracy of low-cost GNSS receivers. Nevertheless, it was not an optimal solution for continuous geodetic GNSS monitoring applications with a high risk of instrument damage.

c. Displacement detection

To identify the size of detected displacements by low-cost GNSS receivers three evaluation tests were conducted [2,5,6]. The first was focused on analyzing the impact of using geodetic GNSS receivers as a base station [6]. In the second one, the impact of employing low-cost GNSS antennas with known calibrated parameters (Survey calibrated) was analyzed in static relative and Precise Point Positioning (PPP) methods [2]. In the third, the size of detected displacements in the Real-Time Kinematic (RTK) method was analyzed [5].

The findings revealed that 3D displacements of 10 mm are detectable in the static relative method by low-cost GNSS receivers (ZED-F9P receiver and ANN-MB antenna) with a high level of reliability [6]. The use of low-cost calibrated antennas (Survey calibrated) decreased the size of detected 3D displacements to 5 mm in 30-minute sessions [2]. However, the size of detected displacements increased to 20 mm in the PPP (8h sessions) [2] and RTK (15s sessions) [5].

d. Application of LGMS

The LGMS (Figure 1) was developed only from low-cost GNSS sensors (ZED-F9P receiver and Survey calibrated antenna), costing around 500 EUR per unit. Notably, the LGMS was capable of functioning in areas lacking an electricity network [7]. It underwent testing during a six-month monitoring period of the Laze landslide, where four LGMS were deployed. The system demonstrated consistent operation, continuously collecting GNSS observations that were post-processed to estimate displacements [5].

In three of the monitoring stations, no horizontal displacements were detected, but slow vertical movements were observed in those stations. By this, it was confirmed that LGMS can detect slow movements with sub-centimeter accuracy continuously and with reduced costs.



Figure 1. LGMS in Laze landslide.

3 Conclusion

The findings of the doctoral dissertation indicate that low-cost GNSS receivers currently do not provide observations of the same quality as geodetic GNSS receivers, but rather slightly inferior. Nonetheless, these GNSS sensors can still deliver positioning solutions with quality that is adequate for numerous surveying projects. Furthermore, low-cost GNSS receivers were shown as well-suited sensors for developing LGMSs. The system was successfully tested in the Laze landslide and represents a suitable solution for near-real-time continuous monitoring with high accuracy and decreased costs.

It is noteworthy that the LGMS can find application in monitoring natural hazards like subsidence, sinking phenomena, and landslides, as well as, various other engineering structures such as viaducts, bridges, dams, mining areas, chimneys, provided it meets the necessary accuracy standards.

References

- [1] Hamza, V. Cost Effective GNSS Receivers for Geodetic Monitoring, University of Ljubljana, 2023.
- [2] Hamza, V.; Stopar, B.; Ambrožič, T.; Sterle, O. Performance Evaluation of Low-Cost Multi-Frequency GNSS Receivers and Antennas for Displacement Detection. *Appl. Sci.* 2021, *11*, 22, doi:10.3390/app11146666.
- [3] Hamza, V.; Stopar, B.; Sterle, O.; Pavlovčič-Prešeren, P. Low-Cost Dual-Frequency GNSS Receivers and Antennas for Surveying in Urban Areas. *Sensors* 2023, *23*, 19, doi:10.3390/s23052861.
- [4] Hamza, V.; Stopar, B.; Sterle, O. Testing the Performance of Multi-Frequency Low-Cost GNSS Receivers and Antennas. *Sensors* 2021, *21*, 16, doi:10.3390/s21062029.
- [5] Hamza, V.; Stopar, B.; Sterle, O.; Pavlovčič-Prešeren, P. A Cost-Effective GNSS Solution for Continuous Monitoring of Landslides. *Remote Sens.* 2023, *15*, 22, doi:10.3390/rs15092287.
- [6] Hamza, V.; Stopar, B.; Ambrožič, T.; Turk, G.; Sterle, O. Testing Multi-Frequency Low-Cost GNSS Receivers for Geodetic Monitoring Purposes. *Sensors* 2020, *20*, 16, doi:10.3390/s20164375.
- [7] Hamza, V.; Stopar, B.; Prešeren, P.P.; Sterle, O. Uporabnost Cenovno Ugodnih Instrumentov GNSS v Nalogah Geodetskega Monitoringa.; Kuhar, M., Ed.; Slovensko združenje za geodezijo in geofiziko: Ljubljana, 2023; pp. 105–120.

CONTENTS OF Informatica Volume 47 (2023) pp. 1-599**Papers**

- AHMAD, S. & Z. KHAN, M. ALI, M. ASJAD. 2023. A Novel Framework Based on Integration of Simulation Modelling and MCDM Methods for Solving FMS Scheduling Problems. *Informatica 47*: 501-514.
- AL-SHAREEDA, M.A. & S. MANICKAM, M.A. SAARE. 2023. Enhancement of NTSA Secure Communication with One-Time Pad (OTP) in IoT. *Informatica 47*: 1-10.
- ALSAEED, D. 2023. LOCUS: A Mobile Tourism Application and Recommender System for Personalized Places and Activities. *Informatica 47*: 201-212.
- ALIJA, S. & E. BEQIRI, A.S. GAAFAR, A.K. HAMOUD. 2023. Predicting Students Performance Using Supervised Machine Learning Based on Imbalanced Dataset and Wrapper Feature Selection. *Informatica 47*: 11-20.
- BINH, H. 2023. Guest Editorial Preface: Information and Communication Technology. *Informatica 47*: 301-302.
- CHEGIREDDY, R.P.R. & A.S. NAGESH. 2023. A Novel Method for Human Mri Based Pancreatic Cancer Prediction Using Integration of Harris Hawks Variants & Vgg16: A Deep Learning Approach. *Informatica 47*: 115-130.
- CHEHAL, D. & P. GUPTA, P. GULATI, T. GUPTA. 2023. Comparative Study of Missing Value Imputation Techniques on E-Commerce Product Ratings. *Informatica 47*: 373-382.
- CHEHAL, D. & P. GUPTA, P. GULATI. 2023. Predicting the Usefulness of ECommerce Products' Reviews Using Machine Learning Techniques. *Informatica 47*: 275284.
- CUI, L. 2023. Application of Adaptive Artificial Bee Colony Algorithm in Reservoir Information Optimal Operation. *Informatica 47*: 193-200.
- FAN, W. & C. CHANG, N. YAO, L. XU, H. JU. 2023. AHP Algorithm for Indoor Air Pollution Detection and Evaluation System Design. *Informatica 47*: 431-440.
- HAMZA, V. 2023. Low-cost GNSS receivers for geodetic monitoring purposes. *Informatica 47*: 603-604.
- HE, Z. 2023. Improved Genetic Algorithm in Multi-objective Cargo Logistics Loading and Distribution. *Informatica 47*: 253-260.
- HOMBALIMATH, A. & N. MANGLA, A. BALODI. 2023. Designing A Permissioned Blockchain Network for The Insurance Claim Process Using Hyperledger Fabric and Composer. *Informatica 47*: 393-416.
- HUČ, A. 2023. Detecting Temporal and Spatial Anomalies in Users' Activities for Security Provisioning in Computer Networks. *Informatica 47*: 295-296.
- ISLAM, M. & M.S. JUNAYED, A. SADEGHZADEH, N. ANJUM, A. JENY, A.F.M.S. SHAH. 2023. Acne Vulgaris Detection and Classification: A Dual Integrated Deep CNN Model. *Informatica 47*: 587-602.
- ISLAM, M.A. & M.H.A. BISWAS. 2023. Potential Impact of Climate Change on Groundwater Level Declination in Bangladesh: A Mathematical Modeling Computation. *Informatica 47*: 261-274.
- ITO, K. & T. MURAYAMA, S. YADA, S. WAKAMIYA, E. ARAMAKI. 2023. Complaints with Target Scope Identification on Social Media. *Informatica 47*: 335-348.
- KABASSI, K. 2023. A Framework for Evaluating Distance Learning of Environmental Science in Higher Education Using Multi-Criteria Group Decision Making. *Informatica 47*: 487-500.
- KAUR, H. & T. KAUR, R. GARG. 2023. A Prediction Model for Student Academic Performance Using Machine Learning. *Informatica 47*: 97-108.
- KIMURA, Y. & T. KOMAMIZU, K. HATANO. 2023. An Automatic Labeling Method for Subword-Phrase Recognition in Effective Text Classification. *Informatica 47*: 315-326.
- KOCUVAN, P. & S. ERŽEN, I. TRUCCOLO, F. RIZZOLIO, M. GAMS, E. DOVGAN. 2023. An Efficient Meta-Platform for Providing Expert Medical Help to Italian and Slovenian Users. *Informatica 47*: 469-476.
- LE, T. & N. HUYNH-DUC, C.T. NGUYEN, M. TRAN. 2023. Dynamic Cost Estimation of Reconstruction Project Based on Particle Swarm Optimization Algorithm. *Informatica 47*: 327-334.

- LI, C. & Z. MU. 2023. Analysis Platform of Rail Transit Vehicle Signal System Based on Data Mining. *Informatica* 47: 441-450.
- LI, L. 2023. Dynamic Cost Estimation of Reconstruction Project Based on Particle Swarm Optimization Algorithm. *Informatica* 47: 173-182.
- LUÉVANO, A.N. & A.H. BARRIENTOS, N.K. VALVERDE. 2023. A Hybrid Modelling Framework for E-Commerce Supply Chain Simulation: Complex Adaptive Systems Perspective. *Informatica* 47: 159-172.
- MADATOV, K. & S. BEKCHANOV, J. VIČIČ. 2023. Automatic Detection of Stop Words for Texts in the Uzbek Language. *Informatica* 47: 143-150.
- MAHLOUS, A.R. 2023. Threat Model and Risk Management for a Smart Home IoT System. *Informatica* 47: 51-64.
- NEMER, Z.N. & W.N. JASIM, E.J. HARFASH. 2023. Implementation of Multiple CNN Architectures to Classify the Sea Coral Images. *Informatica* 47: 43-50.
- NGUYEN, T. & D. DANG. 2023. On Integrating Multiple Restriction Domains to Automatically Generate Test Cases of Model Transformations. *Informatica* 47: 21-42.
- PIPALWA, R. & A. PAUL, T. MUKHERJEE. 2023. Prediction of Heart Disease Using Modified Hybrid Classifier. *Informatica* 47: 65-72.
- QUOC, T.N. & H.L. THANH, H.P. VAN. 2023. Khmer-Vietnamese Neural Machine Translation Improvement Using Data Augmentation Strategies. *Informatica* 47: 349360.
- RAHEEM, S. & M. ALABBAS. 2023. A Modified Spider Monkey Optimization Algorithm Based on Good-Point Set and Enhancing Position Update. *Informatica* 47: 579-586.
- RANJAN, R. & A. DANIEL. 2023. CoBiAt: A Sentiment Classification Model Using Hybrid ConvNet-Dual-LSTM with Attention Techniques. *Informatica* 47: 523-536.
- RATHI, M. & A. SINHA, S. TULSYAN, A. AGARWAL, A. SRIVASTAVA. 2023. Assessing Mental Health Crisis in Pandemic Situation with Computational Intelligence. *Informatica* 47: 131-140.
- SALMAN, I. & J. VOMLEL. 2023. Learning the Structure of Bayesian Networks from Incomplete Data Using a Mixture Model. *Informatica* 47: 83-96.
- SARASWAT, S.K. & V.K. DEOLIA, A. SHUKLA. 2023. Computational Analysis of Uplink NOMA and OMA for 5G Applications: An Optimized Network. *Informatica* 47: 383392.
- SCIUS-BERTRAND, A. & M. BUI, A. FISCHER. 2023. A Hybrid Deep Learning Approach to Keyword Spotting in Vietnamese Stele Images. *Informatica* 47: 361-372.
- SGHAIER, A. & A. MEDDEB. 2023. Real Time QoS in WSN based Network Coding and Reinforcement Learning. *Informatica* 47: 477486.
- SHARMA, R. & R. PIPPAL. 2023. Blockchain based Efficient and Secure Peer-to-Peer Distributed IoT Network for NonTrusting Device-to-Device Communication. *Informatica* 47: 515-522.
- SHULAJKOVSKA, M. & G. NOVESKI, M. SMERKOL, J. GRABNAR, E. DOVGAN, M. GAMS. 2023. EU Smart Cities: Towards a New Framework of Urban Digital Transformation. *Informatica* 47: 151-158.
- SINGH, S. & V.K. SEHGAL. 2023. Deep Learning-Based CNN Multi-Modal Camera Model Identification for Video Source Identification. *Informatica* 47: 417-430.
- SUDHA, G. & C. THARINI. 2023. Hybrid Compression Algorithm for Energy Efficient Image Transmission in Wireless Sensor Networks Using SVD-RLE in Voluminous Data Applications. *Informatica* 47: 555-564.
- TANDON, V. & R. MEHRA. 2023. An Integrated Approach for Analysing Sentiments on Social Media. *Informatica* 47: 213-220.
- THAKUR, N. & N.U. KHAN, S.D. SHARMA. 2023. A Two Phase Ultrasound Image Despeckling Framework by Nonlocal Means on Anisotropic Diffused Image Data. *Informatica* 47: 221-234.
- TOŠIĆ, A. 2023. Tradeoffs In Using Blockchain Technology for Security, Privacy, And Decentralization: theoretical And Empirical Perspectives. *Informatica* 47: 297298.
- TRUEMAN, T.E. & A.K. JAYARAMAN, J. S, G.A. NARAYANASAMY. 2023. A Multichannel Convolutional Neural Network for Multilabel Sentiment

Classification Using Abilify Oral User Reviews. Informatica 47: 109-114.

TSANI, E. & D. SUHARTONO. 2023. Personality Identification from Social Media Using Ensemble BERT and RoBERTa. Informatica 47: 537-554.

TÜZEMEN, S. & Ö. BARIŞ-TÜZEMEN, A.K. ÇELİK. 2023. Sentiment Analysis and Machine Learning Classification of COVID-19 Vaccine Tweets: Vaccination in the shadow of fear-trust dilemma. Informatica 47: 73-82.

VO, A. & T.N. PHAM, V.B. NGUYEN, N.H. LUONG. 2023. Lightweight Multi-Objective and Many-Objective Problem Formulations for Evolutionary Neural Architecture Search with the Training-Free Performance Metric Synaptic Flow. Informatica 47: 303-314.

WANG, X. 2023. Personalized Recommendation System of E-learning Resources Based on Bayesian Classification Algorithm. Informatica 47: 451-458.

WANG, Y. 2023. Deep Learning Models in Computer Data Mining for Intrusion Detection. Informatica 47: 565-578.

YANG, F. 2023. Optimization of Personalized Recommendation Strategy for E-commerce Platform Based on Artificial Intelligence. Informatica 47: 235-242.

YANG, W. & M. ZHAN, Z. HUANG, W. SHAO. 2023. Design and Development of Mobile Terminal Application Based on Android. Informatica 47: 285-294.

YAO, Y. 2023. Design of Ecological Land Remediation Planning and Remediation Mode Based on Spatial Clustering Algorithm. Informatica 47: 183-192.

ZHAO, Z. & J.Z.K. YANG, S.WANG, M. ZENG. 2023. Data Processing of Municipal Wastewater Recycling Based on Genetic Algorithm. Informatica 47: 459-466.

ZHAO, H. & A. SHARMA. 2023. Logistics Distribution Route Optimization Based on Improved Particle Swarm Optimization. Informatica 47: 243-252.

JOŽEF STEFAN INSTITUTE

Jožef Stefan (1835-1893) was one of the most prominent physicists of the 19th century. Born to Slovene parents, he obtained his Ph.D. at Vienna University, where he was later Director of the Physics Institute, Vice-President of the Vienna Academy of Sciences and a member of several scientific institutions in Europe. Stefan explored many areas in hydrodynamics, optics, acoustics, electricity, magnetism and the kinetic theory of gases. Among other things, he originated the law that the total radiation from a black body is proportional to the 4th power of its absolute temperature, known as the Stefan–Boltzmann law.

The Jožef Stefan Institute (JSI) is the leading independent scientific research institution in Slovenia, covering a broad spectrum of fundamental and applied research in the fields of physics, chemistry and biochemistry, electronics and information science, nuclear science technology, energy research and environmental science.

The Jožef Stefan Institute (JSI) is a research organisation for pure and applied research in the natural sciences and technology. Both are closely interconnected in research departments composed of different task teams. Emphasis in basic research is given to the development and education of young scientists, while applied research and development serve for the transfer of advanced knowledge, contributing to the development of the national economy and society in general.

At present the Institute, with a total of about 900 staff, has 700 researchers, about 250 of whom are postgraduates, around 500 of whom have doctorates (Ph.D.), and around 200 of whom have permanent professorships or temporary teaching assignments at the Universities.

In view of its activities and status, the JSI plays the role of a national institute, complementing the role of the universities and bridging the gap between basic science and applications.

Research at the JSI includes the following major fields: physics; chemistry; electronics, informatics and computer sciences; biochemistry; ecology; reactor technology; applied mathematics. Most of the activities are more or less closely connected to information sciences, in particular computer sciences, artificial intelligence, language and speech technologies, computer-aided design, computer architectures, biocybernetics and robotics, computer automation and control, professional electronics, digital communications and networks, and applied mathematics.

The Institute is located in Ljubljana, the capital of the independent state of Slovenia (or S^onia). The capital

today is considered a crossroad bet between East, West and Mediterranean Europe, offering excellent productive capabilities and solid business opportunities, with strong international connections. Ljubljana is connected to important centers such as Prague, Budapest, Vienna, Zagreb, Milan, Rome, Monaco, Nice, Bern and Munich, all within a radius of 600 km.

From the Jožef Stefan Institute, the Technology park “Ljubljana” has been proposed as part of the national strategy for technological development to foster synergies between research and industry, to promote joint ventures between university bodies, research institutes and innovative industry, to act as an incubator for high-tech initiatives and to accelerate the development cycle of innovative products.

Part of the Institute was reorganized into several high-tech units supported by and connected within the Technology park at the Jožef Stefan Institute, established as the beginning of a regional Technology park “Ljubljana”. The project was developed at a particularly historical moment, characterized by the process of state reorganisation, privatisation and private initiative. The national Technology Park is a shareholding company hosting an independent venture-capital institution.

The promoters and operational entities of the project are the Republic of Slovenia, Ministry of Higher Education, Science and Technology and the Jožef Stefan Institute. The framework of the operation also includes the University of Ljubljana, the National Institute of Chemistry, the Institute for Electronics and Vacuum Technology and the Institute for Materials and Construction Research among others. In addition, the project is supported by the Ministry of the Economy, the National Chamber of Economy and the City of Ljubljana.

Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
Tel.: +386 1 4773 900, Fax.: +386 1 251 93 85
WWW: <http://www.ijs.si>
E-mail: matjaz.gams@ijs.si
Public relations: Polona Strnad

INFORMATICA
AN INTERNATIONAL JOURNAL OF COMPUTING AND INFORMATICS
INVITATION, COOPERATION

Submissions and Refereeing

Please register as an author and submit a manuscript at: <http://www.informatica.si>. At least two referees outside the author's country will examine it, and they are invited to make as many remarks as possible from typing errors to global philosophical disagreements. The chosen editor will send the author the obtained reviews. If the paper is accepted, the editor will also send an email to the managing editor. The executive board will inform the author that the paper has been accepted, and the author will send the paper to the managing editor. The paper will be published within one year of receipt of email with the text in Informatica MS Word format or Informatica L^AT_EX format and figures in .eps format. Style and examples of papers can be obtained from <http://www.informatica.si>. Opinions, news, calls for conferences, calls for papers, etc. should be sent directly to the managing editor.

SUBSCRIPTION

Please, complete the order form and send it to Dr. Drago Torkar, Informatica, Institut Jožef Stefan, Jamova 39, 1000 Ljubljana, Slovenia. E-mail: drago.torkar@ijs.si

Since 1977, Informatica has been a major Slovenian scientific journal of computing and informatics, including telecommunications, automation and other related areas. In its 16th year (more than twentyeight years ago) it became truly international, although it still remains connected to Central Europe. The basic aim of Informatica is to impose intellectual values (science, engineering) in a distributed organisation.

Informatica is a journal primarily covering intelligent systems in the European computer science, informatics and cognitive community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the Refereeing Board.

Informatica web edition is free of charge and accessible at <http://www.informatica.si>.

Informatica print edition is free of charge for major scientific, educational and governmental institutions. Others should subscribe.

Informatica

An International Journal of Computing and Informatics

Web edition of Informatica may be accessed at: <http://www.informatica.si>.

Subscription Information Informatica (ISSN 0350-5596) is published four times a year in Spring, Summer, Autumn, and Winter (4 issues per year) by the Slovene Society Informatika, Litostrojska cesta 54, 1000 Ljubljana, Slovenia.

The subscription rate for 2022 (Volume 46) is

- 60 EUR for institutions,
- 30 EUR for individuals, and
- 15 EUR for students

Claims for missing issues will be honored free of charge within six months after the publication date of the issue.

Typesetting: Blaž Mahnič, Gašper Slapničar; gasper.slapnicar@ijs.si

Printing: ABO grafika d.o.o., Ob železnici 16, 1000 Ljubljana.

Orders may be placed by email (drago.torkar@ijs.si), telephone (+386 1 477 3900) or fax (+386 1 251 93 85). The payment should be made to our bank account no.: 02083-0013014662 at NLB d.d., 1520 Ljubljana, Trg republike 2, Slovenija, IBAN no.: SI56020830013014662, SWIFT Code: LJBASI2X.

Informatica is published by Slovene Society Informatika (president Niko Schlamberger) in cooperation with the following societies (and contact persons):

Slovene Society for Pattern Recognition (Vitomir Štruc)

Slovenian Artificial Intelligence Society (Sašo Džeroski)

Cognitive Science Society (Olga Markič)

Slovenian Society of Mathematicians, Physicists and Astronomers (Dragan Mihailović)

Automatic Control Society of Slovenia (Giovanni Godena)

Slovenian Association of Technical and Natural Sciences / Engineering Academy of Slovenia (Mark Pleško)

ACM Slovenia (Ljupčo Todorovski)

Informatica is financially supported by the Slovenian research agency from the Call for co-financing of scientific periodical publications.

Informatica is surveyed by: ACM Digital Library, Citeseer, COBISS, Compendex, Computer & Information Systems Abstracts, Computer Database, Computer Science Index, Current Mathematical Publications, DBLP Computer Science Bibliography, Directory of Open Access Journals, InfoTrac OneFile, Inspec, Linguistic and Language Behaviour Abstracts, Mathematical Reviews, MatSciNet, MatSci on SilverPlatter, Scopus, Zentralblatt Math

Informatica

An International Journal of Computing and Informatics

An Efficient Meta-Platform for Providing Expert Medical Help to Italian and Slovenian Users	P. Kocuvan, S. Eržen, I. Truccolo, F. Rizzolio, M. Gams, E. Dovgan	469
Real Time QoS in WSN based Network Coding and Reinforcement Learning	A. Sghaier and A. Meddeb	477
A Framework for Evaluating Distance Learning of Environmental Science in Higher Education using Multi-Criteria Group Decision Making	K. Kabassi	487
A Novel Framework Based on Integration of Simulation Modelling and MCDM Methods for Solving FMS Scheduling Problems	S. Ahmad, Z.A. Khan, M. Ali, M. Asjad	501
Blockchain based Efficient and Secure Peer-to-Peer Distributed IoT Network for Non-Trustng Device-to-Device Communication	R.K. Sharma and R.S. Pippal	515
CoBiAt: A Sentiment Classification Model Using Hybrid ConvNet-Dual-LSTM with Attention Techniques	R. Ranjan and A.K Daniel	523
Personality Identification from Social Media Using Ensemble BERT and RoBERTa	E.F. Tsani, D. Suhartono	537
Hybrid Compression Algorithm for Energy Efficient Image Transmission in Wireless Sensor Networks using SVD-RLE in Voluminous Data Applications	G. Sudha, C. Tharini	545
Deep Learning Models in Computer Data Mining for Intrusion Detection	Yujun Wang	555
A Modified Spider Monkey Optimization Algorithm Based on Good-Point Set and Enhancing Position Update	S.F. Raheem, M. Alabbas	569
Acne Vulgaris Detection and Classification: A Dual Integrated Deep CNN Model	M.B. Islam, M. S. Junayed, A. Sadeghzadeh, N. Anjum, A.A. Jeny, A. F. M. Shahan Shah	577
Low-cost GNSS Receivers for Geodetic Monitoring Purposes	V. Hamza	593

