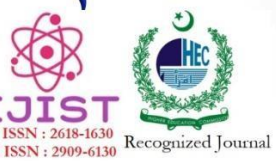
50sea Journals

Special Issue
National University of Modern Languages



Special Efforts
Prof. Dr. Basit Shehzad
Post Doc University of Cambridge UK

Prof. Dr Ali Iqtatad Mirza
Chief Editor
International Journal of Innovations in Science & Technology

Journal.50sea.com

Special Thanks

Prof Dr. Basit Shahzad

Postdoc (Trustworthy Electronic Voting framework using Blockchain)2018
University of Cambridge, United Kingdom.

PhD. (Computer Science) University Technology Petronas (UTP), Malaysia, 2014
MS. (Software Engineering),College of E&ME,National University of Science & Technology, Pakistan, 2005
MSc. (Computer Science),Hamdard University, Karachi, Pakistan, 2000

We are grateful for the special efforts of Prof. Dr. Basit Shahzad for providing us intime support to publish the special issue. The articles published in this issue proved state of the art and a true contribution in research and development.

We are pay humble thanks to Prof. Dr Basit and we are looking forward for the same collaboration in future.

Abstracting and Indexing

Instructions for Authors

The editorial board encourages and welcome true researches, laboratory experiments and real time field observations to get published in IJIST. The authors are advised to prepare their manuscript according to the template of IJIST.

Please see the checklist before submitting your manuscript to IJIST.

- The manuscript is prepared according to the template of IJIST.
- Symbols and names are used according to international standards.
- Page no and Line no are adjusted on the manuscript.
- Figure and Table are clearly cited.
- Author names and their affiliation are typed clearly.
- There is no any limit to the length of manuscript.
- Abstract is comprised of 250 words.
- Author's contribution and the statement narrating no of conflict of interest is mentioned in the end.
- Each Figure and Table is numbered and cited in the text.
- Spelling and English grammar is checked.
- It is "Open Access" journal that publish articles on payment of publishing fee by authors or by their institutions.
- All the articles are published under Creative Common License CC-BY therefore, authors mush agree with same license.

Aims and Scopes

The authors are advised to submit their manuscript in accordance with disciplines as below:

- Administrative Science
- Agriculture/Forestry
- Climatology
- Criminology
- Development Study
- Environment
- GIS
- Geography
- Meteorology
- Physics
- Remote Sensing
- Social Science
- Urban Planning
- Economics
- Chemistry
- Bio-Chemistry
- Computer Science

Peer Review Process

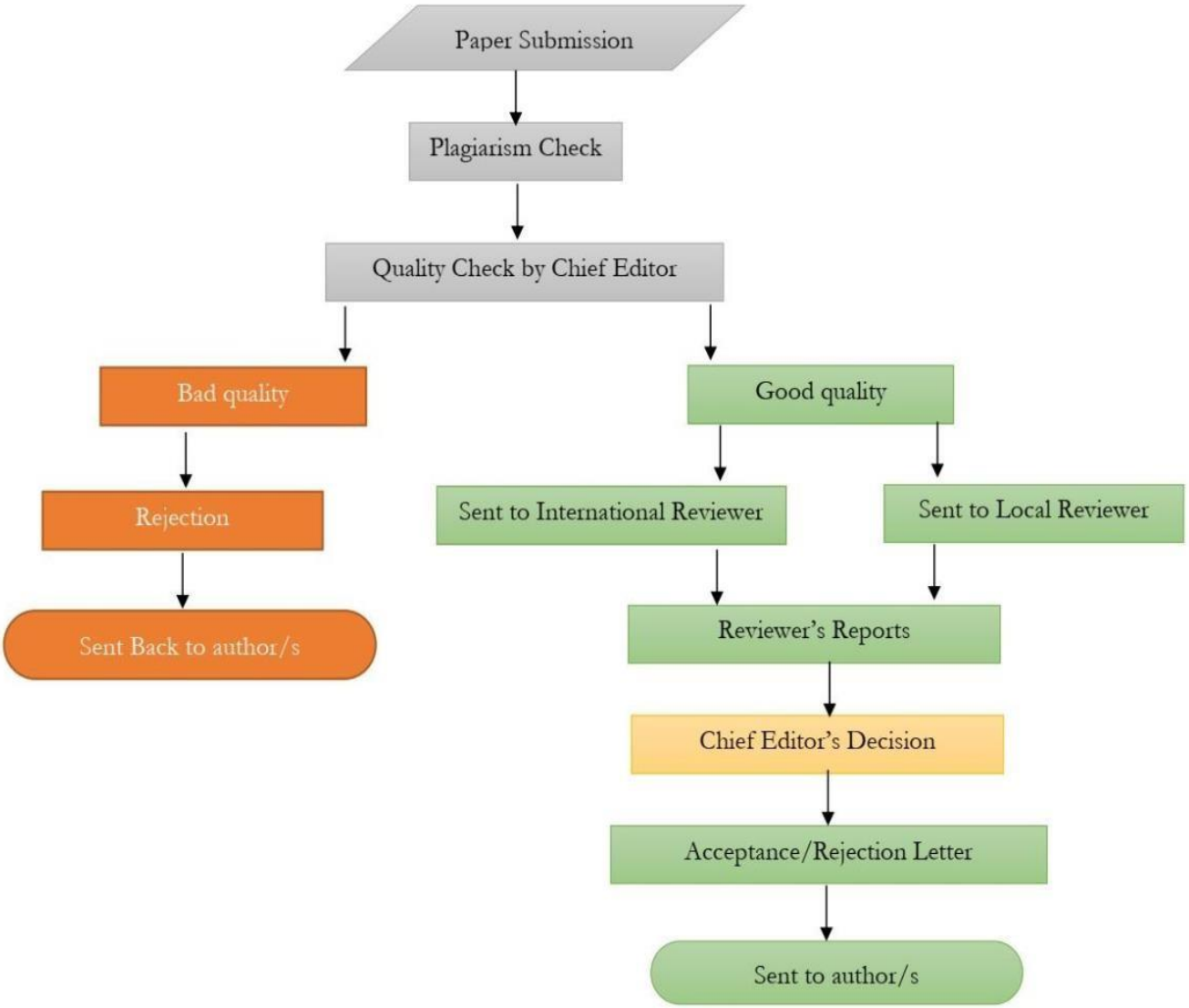


Table of Contents

International Journal of Innovations in Science & Technology (IJIST)

ISSN 2618-1630

Special Issue | Dec 21- Feb 22

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

LSM: A Lightweight Security Mechanism for IoT Based Smart City Management Systems using Blockchain

Hafiz Humza Saeed[1], Abdullah Bin Masood[2], Hassaan Khaliq Qureshi[1]
[1]National University of Science and Technology (NUST), Islamabad, Pakistan.
[2]University of Cyprus (UCY), Nicosia, Cyprus.
[2]CYENS - Centre of Excellence, Nicosia, Cyprus.
*Email ID:[1](hsaeed.msee18seecs, hassaan.khaliq)@seecs.edu.pk,[2]amasoo01@ucy.ac.cy
Citation | Saeed. H. H, Masood. A. B and Qureshi. H. K, "LSM: A Lightweight Security Mechanism for IoT Based Smart City Management Systems using Blockchain". International Journal of Innovations in Science and Technology" Vol 3, Special Issue, pp: 1-14, 2021.
Received Dec 15, 2021; Revised | Dec 19, 2021, Accepted | Dec 19, 2021; Published | Dec 20, 2021.

Smart cities utilize digital technologies for the improvement of its services' quality and performance by reducing resources' cost and consumption, with a commitment of action and efficiency to its citizens. The increased urban migration has led to many problems in cities, such as traffic congestion, waste management, noise pollution, energy consumption, air pollution, etc., as nowadays COVID-19 pandemic has seized the whole world. So, it is necessary to carry out its standard operating procedures (SOPs), including less human interaction. Thus, technology plays a vital role via Internet-of-Things (IoT) based systems. In this paper, a lightweight security mechanism (LSM) is proposed to enrich the IoT based systems. Blockchain technology is integrated, and its completely decentralized peer-to-peer (P2P) technology enables the users' authentication and authorizes legitimate procedures. The IoT based management system is developed to monitor some of the aforementioned problems and solve solid waste, air, and noise monitoring systems. The Ethereum blockchain is used to implement a smart contract based framework for the system's security and access control. The evaluation of performance of the LSM demonstrates that it is an efficient and lightweight tool in terms of cost, resources, and computation and superior over related security studies.

Keywords: Smart City, Internet-of-Things, Lightweight Security Mechanism, Blockchain, Smart Contract.

1. INTRODUCTION

According to the 2017 census, Pakistan's urban population was 32% in 1998, which has increased to 40% and is predicted to reach 50% by 2025 [1]. With the rise of population, the burden on city administrations to provide essential services to all citizens has also

increased. Moreover, the COVID-19 pandemic has held onto the entire world. Thus, it is vital to complete its standard operating procedures (SOPs), which incorporate less human interaction utilizing technology for several purposes [2], [3]. There has been a significant development in the intelligence of digital devices, such as smart

machines, smartphones, and smart sensors, leading to the high-quality pursuits of the Internet-of-Things (IoT) to meet administrative requirements [4]. IoT uses the internet to connect different devices in different areas to collect and analyze information without human-to-human interaction [5].

IoT works on mechanism of transferring the data from sensor to cloud through gateway to store data from where different integrated devices share their information to communicate or to exchange data with each other. This working mechanism reduced the human interaction with computer as sensors automatically exchange their information with each other. Human have just to monitor the data on Graphical-User Interface (GUI) developed for different types of data. IoT includes three layers, the perception layer, the network layer and the application layer [6]. The perception layer includes a group of devices authorized for the Internet which are able to sense the objects, and for exchange of data with other devices by Internet communication systems. Radio Frequency Identification Devices (RFID), cameras, sensors, Global Positioning Systems (GPS) are examples of layer of perception of the devices. Information transmits from perception layer to application layer through the network layer. IoT systems use a combination of short-range systems of communication technologies such as Bluetooth and ZigBee to carry the data of the devices of the perception to a gateway near based on the functionality of the parties' appellants. The technologies of the Internet such as Wi-Fi, 2G, 3G and 4G carry information on long distances based on the implementation [7].

IoT devices have to be built-in with various devices and units, enabling them to interact and engage seamlessly with each other in an impervious way to reduce human resources [8]. Thus, this large volume of data can also pose many problems as it is centralized and monitored from time to time by a single provider. The cloud is a processing and storage technology that cannot guard its consumers' security and privacy [9]. The work was being started, and different techniques were proposed, but the integration of IoT with blockchain technology gained the limelight of researchers and developers [10], [11]. Blockchain made its space in the market due to its decentralized, distributed, and tamperproof ledger properties. Blockchain technology has proven to be sufficient for economic purposes like Bitcoin and can be of incredible value [12], [13]. Integrating it with the IoT based management system enables an extra layer of security and data integrity by authorizing only authentic users. It maintains transaction archives throughout countless nodes that are Peer-to-Peer (P2P) coupled, making it tamperproof [14], [15].

By smart contract, it allows more functionality to play with the IoT and blockchain's integration with each other. A blockchain based scheme by issuing tokens for user access to fog-enabled IoT devices using a smart contract is developed (J. K. Mudhar et al., 2020) [16]. The tokens are issued to the user by the admin in an off-chain procedure by which the question of the token's confidentiality and integrity is raised. However, the feasibility of the system in a real environment is not tested likewise (J. Oh et al., 2021) [17]. In (A. Ouaddah, 2019) [18], a Fair Access and PPDAC is introduced as a lightweight and privacy-preserving access control-based on blockchain, mainly the open access and public type.

On the other hand, (P. Velmurugadass et al., 2021) [19] constructed a blockchain based architecture that is used for data integrity and privacy in the IaaS cloud. However, Proof-of-Work (PoW) is not suitable for IoT systems as they are resource constraints. Apart from blockchain, (M. Masud et al., 2021) [20], a one-way cryptographic hash, bitwise XOR, and nounce (number used only once) are used to provide a lightweight and secure communication. In (G. Sharma et al., 2019) [21], (M. Wazid et al., 2019) [22], proposed a lightweight authentication scheme that proved to be as insecure against privileged insider attacks.

Considering and overcoming the issues raised in the aforementioned studies, LSM, a lightweight security mechanism is proposed. LSM has a strong authentication with accurate verification and reduced the computational overhead. Its performance evaluation makes it a lightweight mechanism for security, resources, optimization, and time. The application chosen to demonstrate the LSM feasibility and potential in a real environment, IoT based smart city management system, is developed to monitor and provide a solution to solid waste, air, and noise monitoring management systems. These smart applications aimed to lead automation to reduce human-to-human or human-to-computer interaction due to the COVID-19 pandemic.

Waste management is a primary expenditure in many modern cities since both the cost for the service and the storage of waste in landfills are relatively high. In current scenario, collection and management of waste is quite difficult without the use of modern technology [23], [24]. To overcome these waste management problems, an IoT based system can be deployed to allow the terminals, namely "Smart Bins," to monitor the available data to manage and call the garbage truck when necessary. IoT based system also offers statistics on air quality in saturated areas, parks, and health tracks. In this way, humans can locate the healthiest route outdoors. This provision requires that the air pollution sensors be deployed in the metropolis and share the statistics freely with all authorized residents [25], [26].

The noise is also a form of pollution as the carbon dioxide (CO_2) in the air. In this case, the metropolis experts have already issued particular legal guidelines to decrease the quantity of noise in the metropolis [27], [28]. However, despite being written on boards (Quiet zone), people keep making noise in the hospital's regions. IoT based framework will observe noise levels for the authorities to take necessary actions. This service can improve the decorum of hospital areas and the silence at night.

Benefiting from IoT characteristics and the distributed nature of blockchain, proposed LSM: a lightweight security mechanism for IoT based smart city management systems. The main contributions of this paper are given below:

- A computationally efficient smart contract based lightweight security mechanism (LSM) for IoT based smart city management system is proposed.
- LSM is secure against various attacks like a spoof, Sybil, and replay.
- LSM only permits the registered and verified users to access the IoT data through the smart contract they authorized for IoT devices.

Section I presents the introduction, literature review, objective and contributions. System architecture and testbed implementation for LSM is provided in section II. Section III evaluates the performance of the LSM in terms of security, time overhead, and benchmark studies. Further, the concluding remarks are offered in section IV.

2. MATERIAL AND METHODS

LSM's architecture is shown in Figure 1. The flowchart of the IoT based smart city management is shown in Figure 2. The architecture of the developed system is consisting of two components whose functionalities are discussed below:

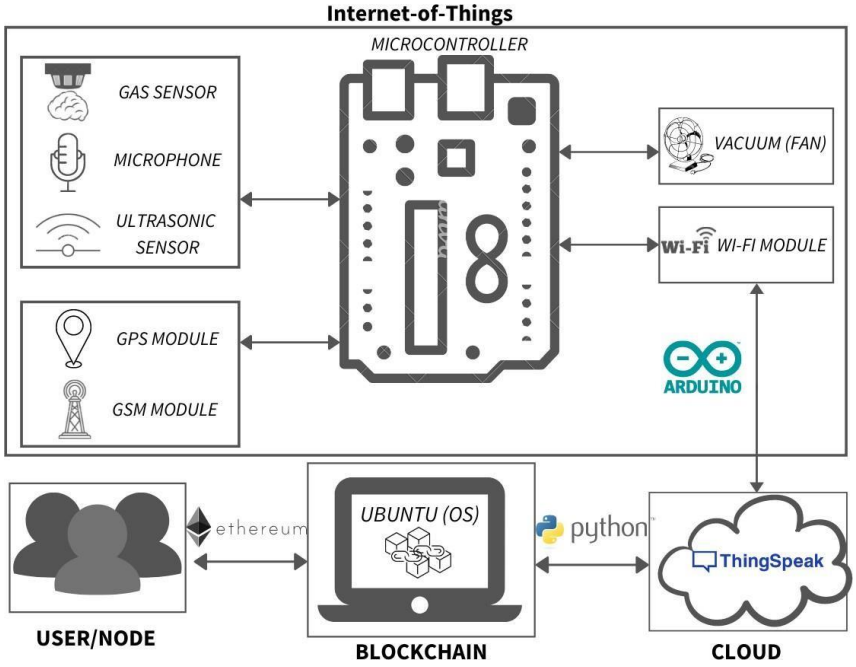
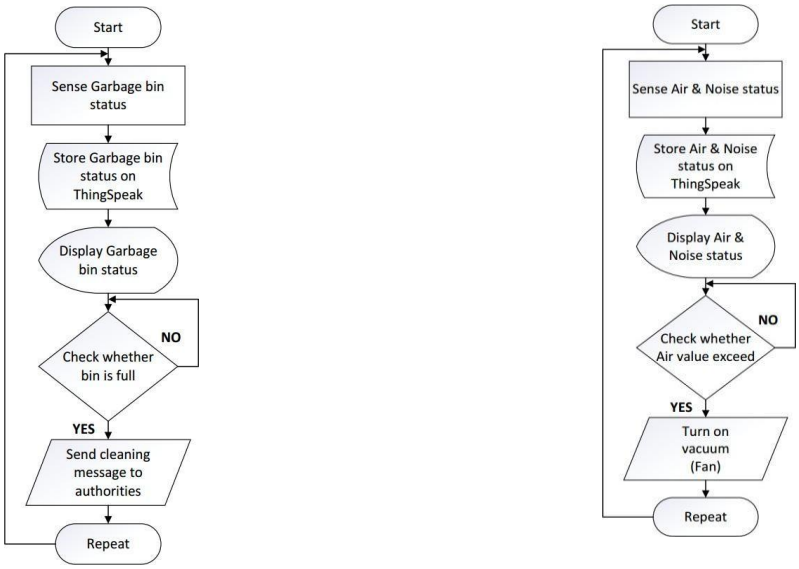


Figure 1. System Architecture of LSM



(a) IoT based Garbage System **(b)** IoT based Air & Noise System

Figure 2. Flowchart of IoT based Smart City Management System

Hardware Components: Different sensors and modules are incorporated in this system architecture to represent IoT system. *microphone*, *MQ-6*, and *ultrasonic sensors* are used for noise, air, and garbage monitoring, respectively. *Mini-fan* is used as a vacuum for demonstration purposes. *Wi-Fi*, *GSM*, and *GPS* modules transmit IoT data, messages, and locations, respectively. All sensors and modules are interfaced with an *Arduino-Uno* board, which is an 8-bit microcontroller integrated circuit.

Software Components: *Arduino-Genuino* software is used for code compilation and configuration of the modules. To store data on the cloud, the *Thing Speak* server is used. *Ethereum* is a popular platform that can process any complex algorithm code through *Ethereum Virtual Machine (EVM)*. So, *Ubuntu operating system (OS)* is used in this system architecture for *Ethereum* blockchain development. Communication between *Ethereum* and cloud is done through a *python script* that includes the *JSON-RPC* protocol and the *Web3py* library[1], which are lightweight and efficient for a resource constraint environment. An *Ethereum* node can call or deploy a smart contract using *Go-Ethereum (Geth)* client. A smart contract is a bunch of rules or provisions of an agreement that executes on a blockchain to audit and authorize these concurred terms without the association of an outsider. *Solidity*, a high-level language, is used to write a smart contract. *Remix IDE* and *Truffle suite* are used to develop and deploy the smart contract.

The IoT based smart city management system is implemented to demonstrate the potential to carry out the LSM as a Proof-of-Concept (PoC). The garbage monitoring system will update their data after every five minutes. The air monitoring system will update after every fifteen minutes, and the noise monitoring system will continuously update its data. IoT devices are connected with the cloud to upload their data and communicate with each other. Users are connected to the blockchain. The advantage of this method is that users get IoT data only when they request it. Resource optimization is done via this technique. User authentication is done via a smart contract in the blockchain and brings confidentiality, integrity, authenticity, and various security attacks like a spoof, sybil, and replay. This study aims to get a lightweight security mechanism in IoT as they operate in a resource constraint environment. Table 1 shows the notations used in this paper.

IoT based Garbage System

In this system, *ultrasonic sensor* and *GPS* module are attached to the garbage bin through which data and location information is fetched. By using the distance formula *"$s = v \times t$",* the*ultrasonic sensor* measures the bin status. Through the *Wi-Fi* module on it, statistics are uploaded on the cloud, enabling the users to monitor it from anywhere. Whenever the garbage bin is full, the message is sent through the *GSM* module to authorities to take necessary actions.

Table 1. Table of Notations

Notations	Description
Data*device*	IoT sensor's data
EA*user*	Ethereum's address
EA*reg.*	Registered users
Hash*reg.*	Registered hash
ID*device*	IoT sensor's number
Sign*user*	User credentials

[1]https://web3py.readthedocs.io/en/stable

Algorithm 1: Smart Contract

Input: $Sign_{user}$, ID_{device}

Output: $Data_{device}$ **Data:**

$EA_{reg.}$, $Hash_{reg.}$

// Checking if the user is authorized to the system

if $EA_{user} \neq EA_{reg.}$ **then**
    return false
    // Checking if the user is authorized to the device
    else if $keccak256(Sign_{user}, ID_{device}) == Hash_{reg.}$ **then**
        return true, $Data_{device}$
        else
            return false

IoT based Air & Noise Monitoring System

Different sensors such as *MQ-6* and microphones are used to fetch the value of noise and air pollution from the surroundings in this system. In addition, *GPS* module is used to access the location. Through the *Wi-Fi* module on it, statistics are uploaded on the cloud, enabling the users to monitor it from anywhere. A relay is used to interface the fan with the air sensor. Whenever a gas value passes a specific value, the signal is given to the relay, and the fan is operated.

Blockchain Integration

Ethereum blockchain and its nodes are developed using the *Geth* implementation on *Ubuntu OS*. The genesis file is created using puppeth to trigger the *Ethereum* blockchain. Clique, *Proof-of-Authority (PoA)*, consensus protocol is opted [29]. *Keccak256* algorithm is used to create *Ethereum* addresses [30, 31, 32]. Elliptic Curve Digital Signature Algorithm (ECDSA) generates private and public keys. The smart contract is developed utilizing the *Remix IDE* platform. The functionality of the smart contract is presented in Algorithm 1. The *Truffle suite* is utilized for the deployment of the smart contract. Smart contract

transactions cannot be changed and are permanently stored in a transparent framework. The deployed code of the smart contract cannot be changed and is only triggered by the sender's transaction message.

3. RESULT AND DISCUSSION

The developed IoT based smart city management system is illustrated in Figure 3. The specifications of the devices on which the developed system is evaluated are shown in Table 2. The results are system-dependent. They can vary from system to system as their specifications change. The smart city application's results are taken by deploying the IoT system in Lahore city[2]31.5204° N, 74.3587° E.

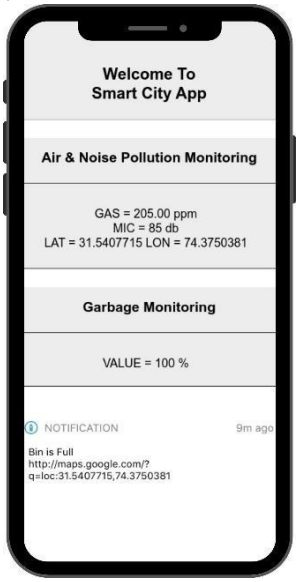


Figure 3. IoT based Smart City Management System

Table 2 Specification of Devices

Device	Model	Processing Speed
Laptop	HP-450-Notebook	Intel Core i5, 3rd Gen.
Internet Router	PTCL DSL-G2452D	8 Mbps

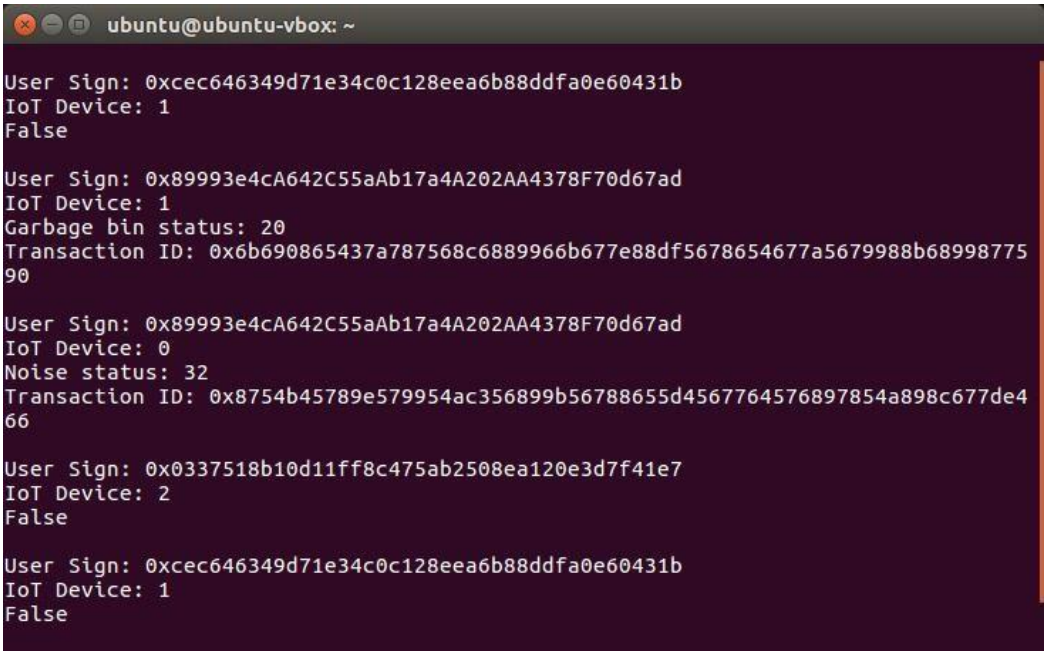


Figure 4. User Authentication via LSM

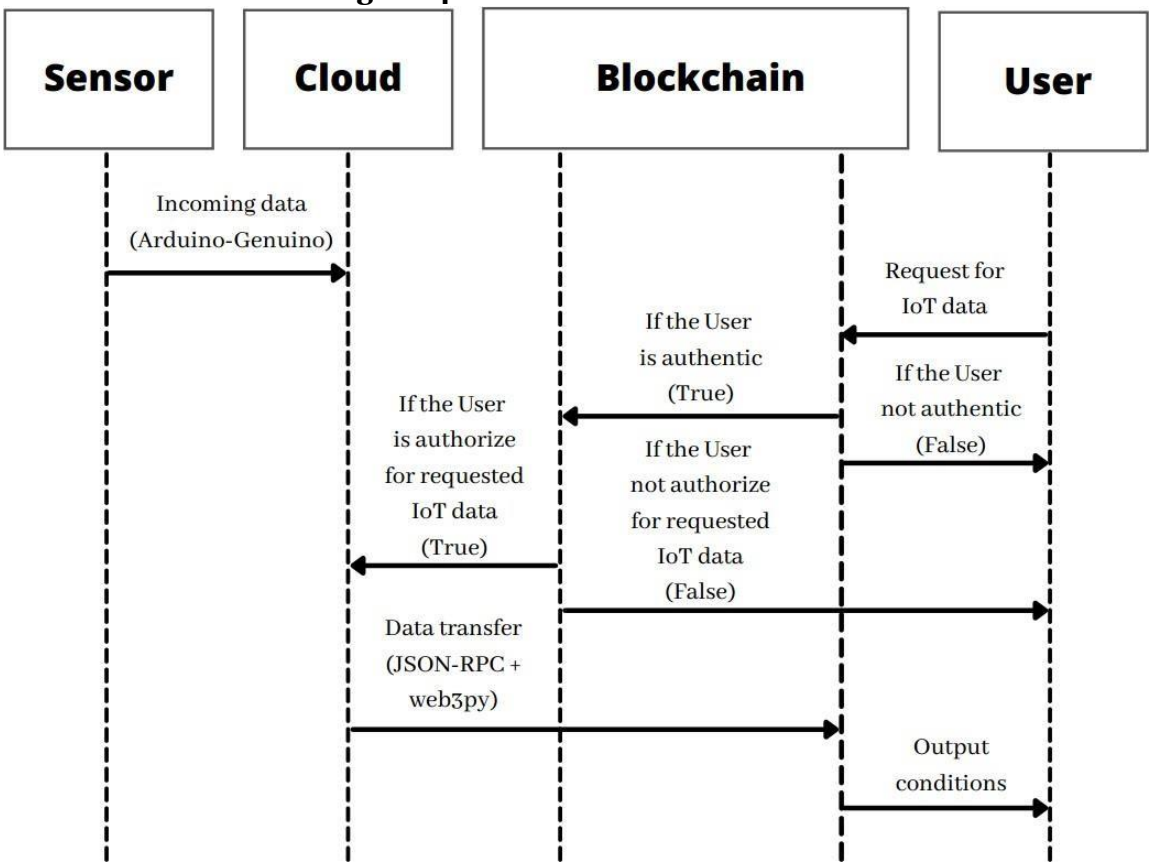


Figure 5. Sequence Diagram of LSM

The incoming sensor's data is transmitted via a *python script* using *JSON-RPC* and *Web3py* library. The *python script* loads data from the cloud's URL[3] and directs it to the blockchain. When the user is required to screen any of the IoT's data, it will enter itself's credentials ($Sign_{user}$) and the required sensor number (ID_{device}) from its *Ethereum* address (EA_{user}). Then, the smart contract will first check the authenticity of the user by comparing EA_{user} with the registered users ($EA_{reg.}$). If the user is authentic, then it will check the combine hash of $Sign_{user}$ and ID_{device}using *Keccak256* algorithm with registered hashes ($Hash_{reg.}$). Different numbers are assigned to various sensors. The number "0" is given to the *microphone* and "1" to the *ultrasonic sensor*, and "2" to the *gas sensor*. If the hash matches and the user is authorized to access the requested IoT's sensor data, then the respective values are then sent to the requested user; otherwise, it will return "false", as illustrated in Figure 4. The complete sequence diagram of LSM is illustrated in Figure 5.

Security Analysis

- **Integrity:** For the integrity of the data in the system, data is signed before sending data to the recipient, using the ECDSA algorithm supported by Ethereum. The recipient confirms this against the smart contract's address.
- **Identification:** $Sign_{user}$ and ID_{device} is required to access the IoT system. Each device and userregistered with the system has a separate ID and sign.

[3]https://thingspeak.com/

- **Non-repudiation:** All transactions are signed with their respective $Sign_{user}$. Therefore, the sender cannot repudiate having performed a transaction.
- **Authentication:** The user must first be registered with the IoT system. If the user is already registered, the smart contract has the associated credentials. As soon as the smart contract verifies the existence and validity of the details provided by the user, it can interact with the IoT system.
- **Spoof attack:** To successfully launch a spoof attack, attacker need a $Sign_{user}$,$ID_{device}$and EA_{user}.If the attacker somehow gets the ID_{device} and EA_{user}, still needs the $Sign_{user}$.
- **Sybil attack:** In a Sybil attack, the attacker needs to create a fake identity to enter into the system. In LSM, users and devices are not allowed to have more than one ID. The message is signed with the private key. Therefore, creating a fake identity in the system has been reduced and is almost infeasible.
- **Replay attack:** In LSM, all messages generated in the system are assigned to a unique transaction ID and timestamp. Therefore, a replay message with a previously accepted transaction ID will be rejected. So, protection against replay attacks is coped.

Time Overhead

If users directly access the cloud for the IoT data, the delay is less as compared to the blockchain. Because users are interacting with the cloud direct now, but there is no significant security in the cloud. On the other hand, in blockchain, users interact with the cloud via blockchain due to which processing and propagation delay increases. So, this is our trade-off between delay and security. But, as the number of users increases in the cloud, the total delay increases because of the rise in queuing delay. But in blockchain, despite the increase in the number of users, total delay remains almost constant, as illustrated in Figure 6. The total delay is calculated using the total delay equation, illustrated in Figure 7.

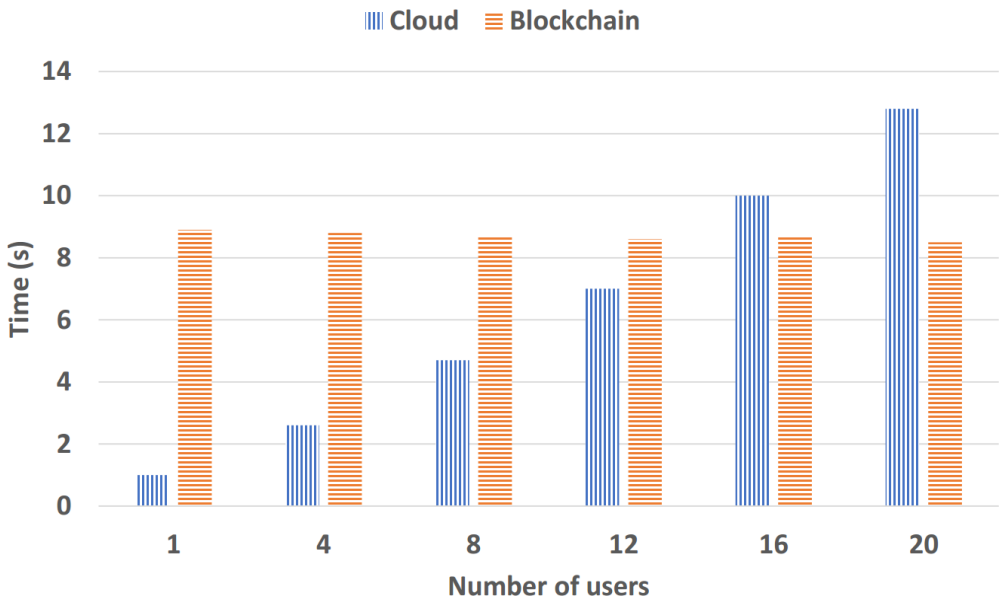


Figure 6. Total Delay Graph over Number of Users Access IoT Data via Cloud vsBlockchain

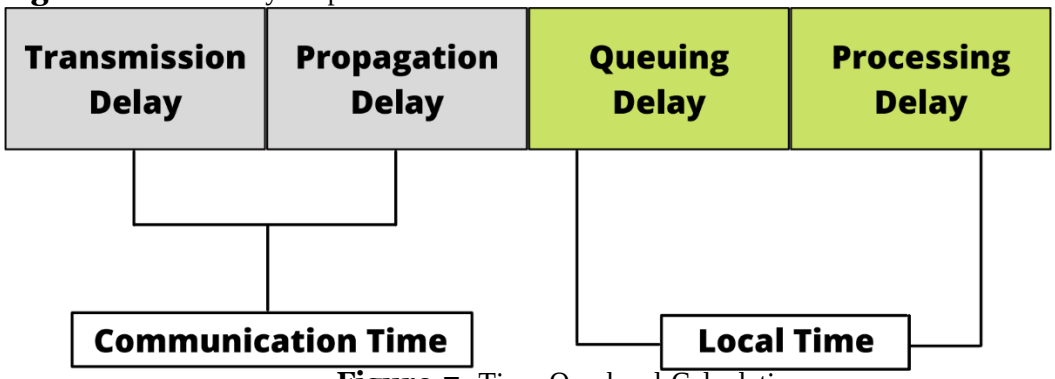


Figure 7. Time Overhead Calculations

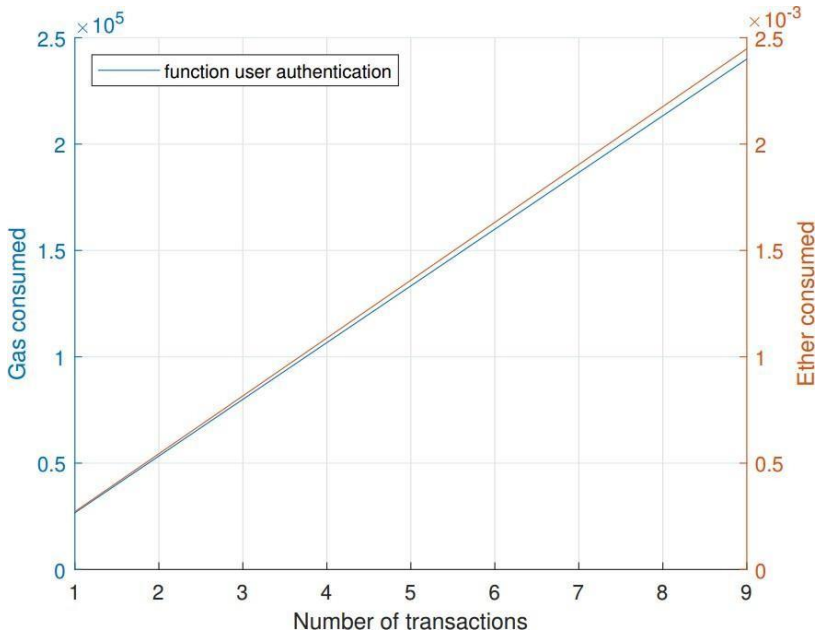


Figure 8. Computational Graph of LMS in terms of Gas and Ether consumed

Comparison with Relevant Work

Gas is described as a resource that is paid for transaction verification. By increasing the gas limit, the average block size increases, affecting the increase in cost. A large block size means more space to store the Ethereum blockchain. LSM consumed 26664 gas, illustrated in Figure 8. While (J. K. Mudhar et al., 2020)'s request access smart contract consumed 51402 gas. (P. Velmurugadass et al., 2021) used PoW operations which include more CPU power consumption as compared to PoA. So does the energy consumption also increase in PoW which has a negative effect on the system's delay. PoA is a lightweight consensus protocol, and its equipment is also cost effective as compared to PoW.

LSM can scale well regarding the number of devices without affecting the system as the cloud manages it. The computational effort is independent of the number of devices. LSM is tested in a real environment and has a permission access control compared to (J. Oh et al., 2021), and (A. Ouaddah, 2019).

4. CONCLUSION

This paper demonstrated the lightweight security mechanism LSM for an IoT based smart city application management system by integrating blockchain technology within the network, enhancing the user authentication and access control using the smart contract. The performance evaluation illustrates LSM as a lightweight in terms of cost, resources, and computation. While secure in the spoof, sybil, and replay attacks. In the near future, the plan is to extend the system with more applications and add machine learning/deep learning to make smart cities more efficient and autonomous to cope with the recent and zero-day attacks.

DECLARATIONS

- **Acknowledgment:** This manuscript has not been published or submitted to any other journals.

- **Conflict of interest:** The authors declare that they have no conflict of interest.
- **Informed consent:** Informed consent was obtained from all the individuals included in the study.
- **Authors' contributions:** This work is conceptualized by H. K. Q. Associated application concepts and experimentation design is done by H. H. S. A. B. M. provided technical input to improve the quality and presentation of work. All the authors read and approved the final manuscript.
- **Project details:** This project has received funding from the European Union's Horizon 2020 Re- search and Innovation Programme under Grant Agreement No 739578 and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

REFRENCES

[1] N. V. M. Lopes and S. Farooq, "*Smart city governance model for Pakistan*," "Smart Governance for Cities: Perspectives and Experiences" 2020, pp. 17–28.

[2] D. Skegg, P. Gluckman, G. Boulton, H. Hackmann, S. S. A. Karim, P. Piot, and C. Woopen, "*Future scenarios for the covid-19 pandemic*," "The Lancet" 2021, vol. 397, pp. 777–778.

[3] M. Salman, Z. U. Mustafa, N. Asif, N. Shehzadi, T. M. Khan, T. H. Mallhi, Y. H. Khan, F. Saleem et al., "*Awareness of covid-19 among illiterate population in Pakistan: A cross-sectional analysis*," "Disaster Medicine and Public Health Preparedness" 2021, pp. 1–17.

[4] S. Balne and G. Sindhu, "*Network protocol challenges of internet of things (iot) features-review*," "International Journal of Innovative Research in Science, Engineering and Technology" 2021, vol. 10, pp. 2305–2309.

[5] M. A. Khan and K. Salah, "*Iot security: Review, blockchain solutions, and open challenges*," "Future Generation Computer Systems" 2018, vol. 82, pp. 395–411.

[6] Z. Lv, L. Qiao, A. Kumar Singh, and Q. Wang, "*Ai-empowered iot security for smart cities*," "ACM Transactions on Internet Technology" 2021, vol. 21, pp. 1–21.

[7] M. H. Kashani, M. Madanipour, M. Nikravan, P. Asghari, and E. Mahdipour, "*A systematic review of iot in healthcare: Applications, techniques, and trends*," "Journal of Network and Computer Applications" 2021, pp. 103164

[8] A. Sudha et al., "*An analytical review on privacy-preserving and public auditing in cloud storage*," "IEEE Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)" 2021, pp. 74–80.

[9] M. M. Khandekar and R. V. Dhumal, "*Review paper on cloud computing*," "Case Studies for Research in Computer Science and Engineering" 2021, pp. 17.

[10] N. Rifi, E. Rachkidi, N. Agoulmine, and N. C. Taher, "*Towards using blockchain technology for iot data access protection*," "IEEE 17th international conference on ubiquitous wireless broadband (ICUWB)" 2017, pp. 1–5.

[11] P. Chinnasamy, C. Vinothini, S. Arunkumar, S. A. Sundarraj, and S. Annlin, "*Blockchaintechnology in smart-cities*," "Blockchain Technology: Applications and Challenges" 2021, pp. 179.

[12] H.-N. Dai, Z. Zheng, and Y. Zhang, "*Blockchain for internet of things: A survey*," "IEEE Internet of Things Journal" 2019, vol. 6, pp. 8076–8094.

[13] A. F. Aysan, H. B. Demirtas¸, and M. Sarac¸, "*The ascent of bitcoin: Bibliometric analysis of bitcoin research*," "Journal of Risk and Financial Management" 2021, vol. 14, pp. 427.
[14] M. Shurman, A. A.-R. Obeidat, and S. A.-D. Al-Shurman, "*Blockchain and smart contract for iot*," "IEEE 11th International Conference on Information and Communication Systems (ICICS)" 2020, pp. 361–366.
[15] S. Hakak, W. Z. Khan, G. A. Gilkar, M. Imran, and N. Guizani, "*Securing smart cities through blockchain technology: Architecture, requirements, and challenges*," "IEEE Network" 2020, vol. 34, pp. 8–14.
[16] J. K. Mudhar, S. Kalra, and J. Malhotra, "*An efficient blockchain based authentication scheme to secure fog enabled iot devices*," "IEEE Indo–Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)" 2020, pp. 75–80.
[17] J. Oh, S. Yu, J. Lee, S. Son, M. Kim, and Y. Park, "*A secure and lightweight authentication protocol for iot-based smart homes*," "Sensors" 2021, vol. 21 pp. 1488.
[18] A. Ouaddah, "*A blockchain based access control framework for the security and privacy of iot with strong anonymity unlinkability and intractability guarantees*," "Advances in Computers" 2019, vol. 115, pp. 211–258.
[19] P. Velmurugadass, S. Dhanasekaran, S. S. Anand, and V. Vasudevan, "*Enhancing blockchain security in cloud computing with iot environment using ecies and cryptography hash algorithm*," "Materials Today: Proceedings" 2021, vol. 37, pp. 2653–2659.
[20] M. Masud, G. S. Gaba, K. Choudhary, M. S. Hossain, M. F. Alhamid, and G. Muhammad, "*Lightweight and anonymity- preserving user authentication scheme for iot-based healthcare*," "IEEE Internet of Things Journal" 2021, pp. 1.
[21] G. Sharma and S. Kalra, "*A lightweight user authentication scheme for cloud-iot based healthcare services*," "Iranian Journal of Science and Technology, Transactions of Electrical Engineering" 2019, vol. 43, pp. 619–636.
[22] M. Wazid, A. K. Das, S. Shetty, J. JPC Rodrigues, and Y. Park, "*Ldakm-eiot: Lightweight device authentication and key management mechanism for edge-based iot deployment*," "Sensors" 2019, vol. 19, pp. 5539.
[23] K. R. Vanapalli, H. B. Sharma, V. P. Ranjan, B. Samal, J. Bhattacharya, B. K. Dubey, and S. Goel, "*Challenges and strategies for effective plastic waste management during and post covid-19 pandemic*," "Science of The Total Environment" 2021, vol. 750, pp. 141514.
[24] S. Nanda and F. Berruti, "*Municipal solid waste management and landfilling technologies: a review*," "Environmental Chemistry Letters" 2021, vol. 19, pp. 1433–1456.
[25] S. W. Hunt, D. A. Winner, K. Wesson, and J. T. Kelly, "*Furthering a partnership: Air quality modeling and improving public health*," "Journal of the Air & Waste Management Association" 2021, vol. 71, pp. 682-688.
[26] S. Malleswari and T. K. Mohana, "*Air pollution monitoring system using iot devices*," "Materials Today: Proceedings" 2021.
[27] R. Chandrappa and D. B. Das, "*Noise pollution*," "Environmental Health-Theory and Practice" 2021, pp. 141–148.
[28] D. Dobrilovic´, V. Brtka, G. Jotanovic´, Z˘. Stojanov, G. Jaus˘evac, and M. Malic´, "*Architecture of iot system for smart monitoring and management of traffic noise*," "5th EAI International Conference on Management of Manufacturing Systems" 2022, pp. 251–266.

[29] S. Joshi, "*Feasibility of proof of authority as a consensus protocol model*," "arXiv preprint arXiv:2109.02480" 2021.
[30] T. A. Alghamdi, I. Ali, N. Javaid, and M. Shafiq, "*Secure service provisioning scheme for lightweight iot devices with a fair payment system and an incentive mechanism based on blockchain*," "IEEE Access" 2019, vol. 8, pp. 1048–1061.
[31] M. Hussain, M. Beg et al., "*Fog computing for internet of things (iot)-aided smart grid architectures*," "Big Data and cognitive computing" 2019, vol. 3, pp. 8.
[32] A. Ahmad, M. Saad, J. Kim, D. Nyang, and D. Mohaisen, "*Performance evaluation of consensus protocols in blockchain- based audit systems*," "IEEE International Conference on Information Networking (ICOIN)" 2021, pp. 654–656.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Requirements decision-making as a process of Argumentation: A Google Maps Case Study with Goal Model

Javed Ali Khan[1]*, Affan Yasin[2], Engr. Muhammad Assam[3], Wasiat Khan[1], Syed Yaqub Shah[1], Rehan Ali Khan[4]
[1]Department of Software Engineering, University of Science and Technology Bannu, Pakistan
[2]School of Software, Tsinghua University, Beijing, China.
[3]College of computer science and technology, Zheijiang University, Hanghzou, China
[4]College of electrical engineering, Zheijiang University, Hanghzou, China
*Correspondence: Javed Ali Khan, engr_javed501@yahoo.com.
Citation | Khan. J. A, Yasin. A, Assam. E. M, Khan. W, Shah. S. Y and Khan. A. R, "Requirements decision-making as a process of Argumentation: A Google Maps Case Study with Goal Model", International Journal of Innovations in Science and Technology, Vol 3, Special Issue, pp: 15-33, 2021.
Received |Dec 14, 2021; Revised | Dec 19, 2021 Accepted | Dec 19, 2021; Published | Dec 20, 2021.

In social media platforms, crowd-users extensively interact and contribute information related to software applications. Usually, crowd-users discuss software features or hot issues and record their opinions about the software applications under discussion either in textual form or via end-user votes. Such requirements-related information is considered a pivotal alternative source for requirements engineers to the already existing in-house stakeholders in order to illustrate decision-making. Also, requirements decision-making for Crowd requirements engineering is a difficult task, as it is always based on incomplete knowledge and requires trade-offs from multi-perspectives. However, existing requirements models and associated tools are still lacking, which enable requirements engineers to make informed decision-making and capture conflicting requirements knowledge. This paper elaborates the interaction among the crowd-users about the Google Map mobile application in the Reddit forum to recover conflicting requirements-related information using the goal modeling approach. For this purpose, we extracted critical arguments from a crowd-users conversation in user forums regarding a given design; built a graphical argumentation model based on the extracted information; aligned types of arguments with goal-oriented modeling constructs in the non-functional requirements framework; conducted exiting goal-model analysis to the requirements model to reach consensus based on argumentation and reasoning, such as supporting, attacking, undefined, and conflicting. The proposal is described with illustrative example models and the associated evaluation processes of design decision-making situation for Google Map interface design.
Keywords: CrowdRE, Argumentation, Goal Modeling, Requirements, user forum.

INTRODUCTION

With the persistent use of online user forums, app stores, and other prominent social media platforms, i-e Twitter, Facebook, etc., it becomes pivotal to listen to the end-users and understand their requirements, needs, and suggestions for software design and evaluation [1, 2]. Recently, research studies have shown that such a large amount of freely available crowd-users data contributes too many different aspects of crowd requirements engineering and helps in requirements decision-making [3]. In recent studies, researchers have shown greater interest in computational argumentation [4,5] in requirement engineering to identify conflicting crowd requirements information [6,7] and help illustrating decision-making [8].

Kanchev et al. [9] proposed the Canary framework, which captures and identifies large copra of requirement-related knowledge from the end-user forum, then using a high-level query language, pivotal and useful requirements-related information is captured for requirements analysts and engineers to make future requirements-related decisions. Further, Kanchev and Chopra [9] found that user forums contain rich information about software requirements having supporting and rebuttal arguments. At the same time, Kurtanovic and Maalej [10, 11] applied supervised machine learning techniques on amazon store to MINE rationale concepts from end-user reviews. Reviews are classified into different rationale categories: issues, alternatives, criteria, decision, and justification, and also found that how users argue and support their decision regarding up-gradation and application switching.

Furthermore, a typical scenario in user forums is that an end-user starts a conversation topic by sharing their experience of using a software application in the form of a comment, upon which other users comment either in favor or against that initial comment. Other users might give suggestions for a new feature or provide an alternative solution in response to that comment. Also, end users can express like and dislike emotions on other user's comments. From such discussions threads and their nested structure in a user forum, we can quickly identify key requirements and arguments on a given topic and organize them in a structured way using the existing goal-modeling approach. Such arguments from users can be a good source of information for requirement decision-making [12, 13]. It inspires us to perform an exploratory study on the crowd-users comments in the Reddit user forum to recover distant aspects of social media related to requirements engineering and how they can be restructured and modeled to leverage requirements analysts. Mainly, we are interested in different systematic approaches to model, structure, and analyze requirements knowledge recovered from the user forum in a more structured way.

In this paper, we contributed in the following aspects: (1) **we extracted key arguments from a crowd-users conversation in user forums regarding a given design decision topic; (2) built a graphical argumentation model based on the extracted information; (3) align types of arguments with goal-oriented modeling constructed in the non-functional requirements framework; (4) conducted exiting goal-model analysis to the requirements model to reach consensus based on argumentation reasoning.**

We answered the following research questions with the proposed research approach: **(1) How the typical user forum dataset looks like? (2) How to model key requirements arguments identified from the raw user data and build an argumentation model? (3) How to apply existing goal-model semantic on the argumentation model to recover conflicting requirements information?** In other

words, how user forums data can be appropriately used during requirements decision-making if we treat them as an argumentation process.

The main structure of research paper is: in literature review, which elaborates on the existing literature work; for instance, it highlights the dataset taken for the proposed approach; in Argumentation for Requirements decision using goal model approach section, we introduce our proposed approach to construct a goal model from the end-comments; in Argumentation integration with model evaluation section, we elaborate on the analysis process and reports the results; In the discussion section, we discuss the research findings, authenticity, applicability, and marketing the proposed approach; In the conclusion section, we conclude the paper and discusses future work.

Literature review

In this section, related work on requirement engineering (RE) with argumentation is elaborated in detail. The Dung's abstract argumentation framework [14] and Toulmin arguments [15] are the two most frequently used frameworks in RE research problems, i-e, security requirements [16], risk assessment [17], privacy requirements [18], requirements decision-making [19], consistent goal-based requirements [20], useful information [21] and legal requirements [22,23] to identify conflicting RE knowledge.

Argumentation theory [24] was first introduced in security requirements by Haley et al. [25, 26]. They proposed that system security requirements can be satisfied using formal and informal arguments, called satisfaction arguments. Later, Franqueira et al. [27] extended Haley's research work by proposing a risk assessment method (RISA) to identify refutations and justifications for the security requirements satisfaction and overcome incomplete and uncertain information. At the same time, Kovacs et al. [28] validated the RISA framework with a complex BitMessage chat application (case study) to analyze security requirements. During analysis, they identified new security requirements to improve the BitMessage client app.

Jureta et al. proposed Acceptability evaluation framework (ACE) [29], which formally caters the discussions between stakeholders and engineers on the validity of RE artifacts in the form of a graph, for which an acceptability condition is defined, if it holds, that means that validity is achieved for the given artifact. Yehia et al. [30] proposed the CaRE framework, transforming informal, ambiguous, conflicting, and incomplete stakeholder requirements into complete, consistent requirements using abstract argumentation. Zee et al. [31, 32] developed a framework to trace the elements of Goal-oriented requirement language (GRL) to the evidence and arguments of system stakeholders. The framework is based on the ASPIC+ framework, extended with practical and evidence-based argumentation. Also, using a UML metamodel, they integrated the ASPIC+ argumentation model with GRL elements. Van et al. [33] extended the Zee work and developed the RationalGRL framework, which helps to map arguments diagrams to the goal models. For this, first, the arguments are identified as end-users goals, then these arguments are converted into the abstract framework (AF), having nodes and attacking arguments. Later, the AF is then translated into anaimed model. Finally, Ghanavati et al. [34] evaluate and analyze the previously developed RationalGRL framework with an example.

Bagheri et al. [35] proposed an argumentation-based approach to capture the inconsistencies in the requirement specification using abstract argumentation semantics. Also, the preference function is utilized to recover the most inconsistent requirement pairs. Elrakaiby et al. [36] proposed an argumentation-based framework to capture ambiguities arrived during requirements elicitation by interview. Murukannaiah et al. [37] proposed an Arg-ACH approach to resolve stakeholder's goals conflicts by identifying and capturing inconsistencies between the goals and beliefs. Kurtanovic and Maalej [38, 39] applied supervised machine learning techniques on amazon store to mine rationale concepts from user reviews. Reviews are classified into different rationale categories: issues, alternatives, criteria, decision, and justification, and also found that how users argue and support their decision regarding up-gradation and application switching. Recently, Khan et al. [40,41,42] proposed the CrowdRE-Arg approach, which analyzes end-users conversations in the online user forum and identifies conflict-free requirements-related information using abstract argumentation semantics. The CrowdRE-Arg approach is supported by machine learning classifiers and other developed algorithms to automate the proposed framework.

In summary, much research efforts have been made in adopting argumentation to address RE related decision-making processes in a situation when : (i) we need to deal with conflicting information, e.g., security requirements decisions are inherently inconsistent with conflicting viewpoints (attackers and defenders); (ii) it involves weighing, comparing, or evaluating arguments, e.g., goal-oriented analysis are precisely the process of weighing, comparing and evaluating the alternatives supported by different arguments from multiple stakeholders; (iii) when we need to make decisions, e.g., design decisions, compliance decisions, priorities of requirements decisions, etc.

EXAMPLE CASE

The example we have selected for our case study is a discussion post on Reddit[1] forum about Google Maps, encircled with red color, as shown in Figure. 1. Reddit is a social web forum where registered users, referred to as *Redditors,* can express their agreement or disagreement with up-votes and down-votes. Every day, millions of people worldwide have different demographics and skills posts, likes, and comments. The Reddit community is broken down into smaller communities referred to as *subreddits,* where each subreddit represents a different general topic, e.g., "Games," and is managed by the moderators. Redditors can subscribe to various subreddits so that their contents are displayed on the home page of Redditors. Redditors can submit a post in the form of a story, link, or video while selecting one of the subreddits they subscribed to the Reddit user forum. Posts with a high value of up-votes appear at the top in subreddits. Redditors earn "karma" for submitting either a post or make comments on any post.

[1] https://www.reddit.com/ Access on 25-5-2021

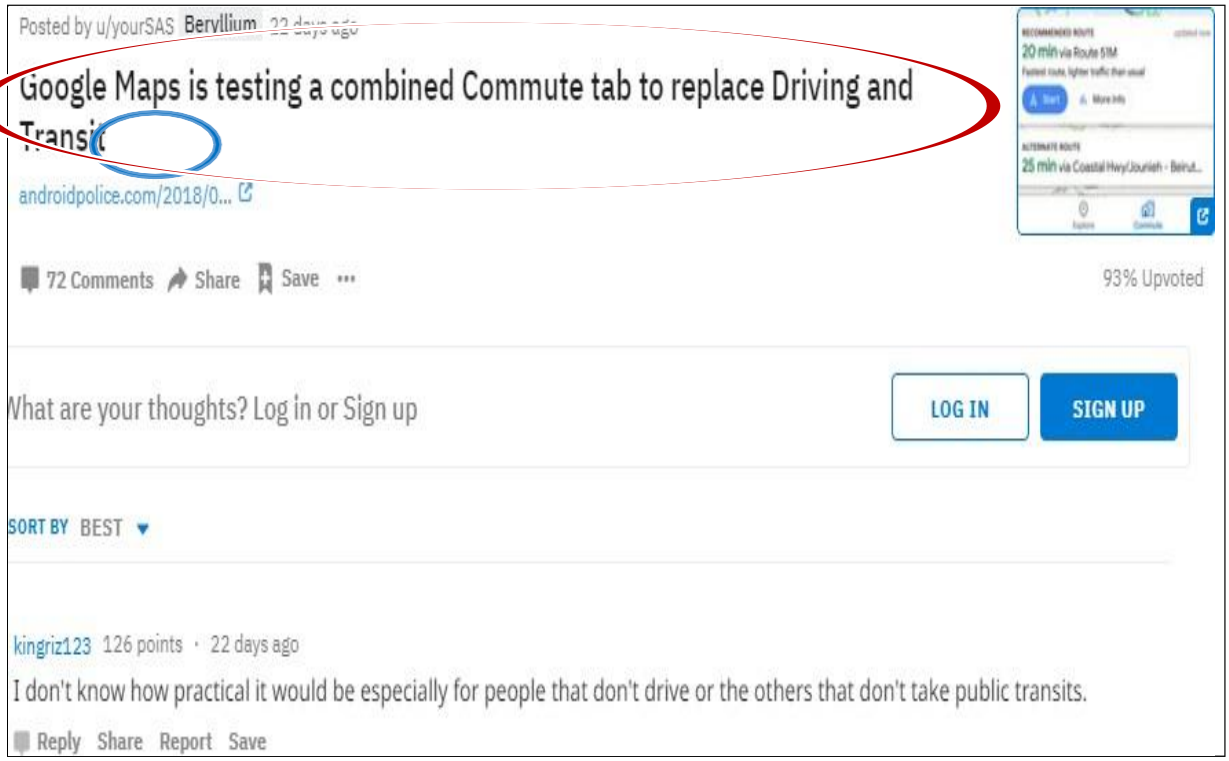


Figure 1. Example under consideration of Google map mobile app.

Also, Reddit users may get "Reddit gold" if other users like or value his post or comment. Posts on the Reddit forum are displayed in hierarchical order. The most common features of the Google Map mobile application are, explore locations, street view, turn-by-turn view, public transits, finding parking lots. In the example, different crowd-users (contributors) discuss the change made to the Google Map mobile application in its layout by replacing the driving and transit button with the commute button in the bottom tab of the application.

In total, 53 Redditors joined the crowd-user discussion in the user forum. They contributed 72 posts in the end-user forum against the main discussion topic, encircled with the blue color, as shown in Figure 1. For the crowd-user discussion topic, we consider only 54 end-user comments as relevant to the requirements engineering domain during the manual analysis using the content analysis approach [43], while the remaining 18 comments were either classified as irrelevant or deleted by the administrator or the crowd-users itself. Therefore, we need to ignore its corresponding replies to preserve the natural flow of user comments.

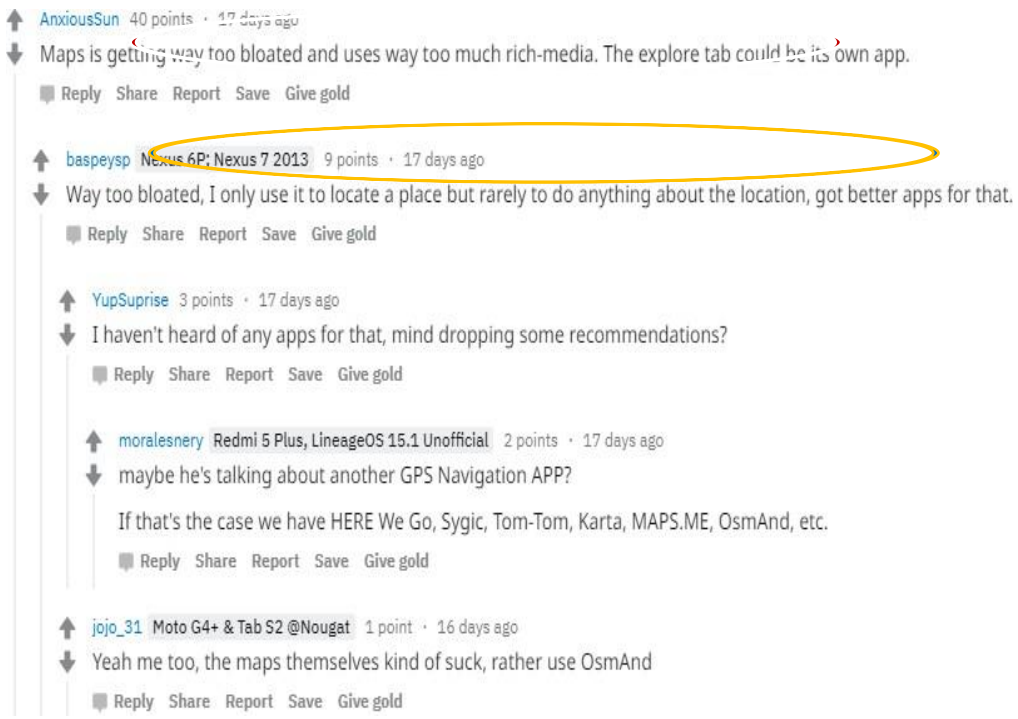


Figure 2. Typical comment structure on online user forum.

Figure. 2 shows an end-user conversation except between different crowd-users that discusses a Google Map design issue. A crowd-user named "AnxiousSun" submits an attacking argument in response to the main discussion topic, encircled with red color. Similarly, Baspeysp registered a supporting claim in response to the "AnxiousSun" attacking claim, encircled with a yellow color, followed by other crowd-users arguments, classified as alternative options to the Google Map mobile application. We use it as a running example to illustrate the analysis process we propose.

ARGUMENTATION FOR REQUIREMENTS DECISIONS USING GOAL MODELING APPROACH

In this paper, we proposed an approach that extracts requirements-related arguments from user forum discussions and use them to support requirements decisions by using the semantics of existing goal-modeling tools. The proposed research approach includes the following steps: (1) We use a goal-oriented modeling technique to organize the discussion process between the crowd-users. (2) Next, we manually identify the key requirements decision arguments by using the content analysis approach [44], which is modeled as the root node of the goal model to identify whether it is conflict-free or not (3) Major requirements-related arguments of crowd-users are categorized according to its nature: alternative operationalization, facts/belief/justification, emotional label, issue/challenge, each of them is elaborated in Table. 1 (4) According to the different nature of the arguments, we suggest different continuing argumentation processes as requirements elicitation or evaluation processes to follow. For example, the potential of argumentation as a tool for convincing and persuading others, understanding a topic collaboratively, and finding an agreed solution that can be adopted for the purpose of requirements. In particular, computational argumentation

theory [45] supports decision-making based on the rationale to reach a goal by resolving the conflicts, by identifying attacking and supporting arguments [45]. Furthermore, the crowd-users comments in the user's forums are lengthy as compared to the app store [46] and Twitter [46], which often contain rationale information that can be worthy for making requirements decisions based on argumentation [47]. We identified the original user discussion in the user forum and extracted information relevant to requirements. We parse them manually using the content analysis approach [48] and go through the steps mentioned above to explore the feasibility of the proposed method.

Identifying the decision topic as a Root Node

When there is a software design or requirements decision to make, one often initiates a new discussion topic in the forum discussion. In our case, it is "Adding a commute Tab in Google Map interface," for which we would like to collect the arguments of crowd-users from the user forum, which are registered by them in response to the main discussion topic, as depicted in Figure 2. The decision point is the root goal, as shown in Figure 3, for which end-users propose multiple design alternatives, issues, and supporting and attacking arguments to reach that root goal. The root node is encircled with red color in Figure 3.

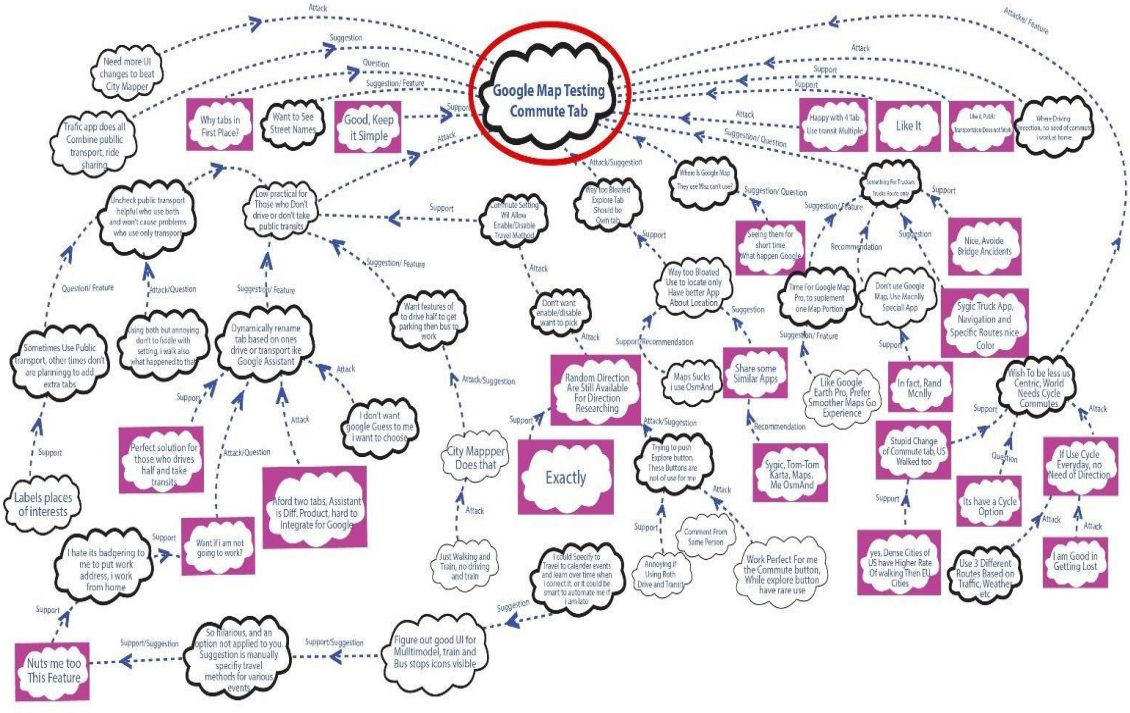


Figure 3. Identifying Discussion Topic as the Root Goal constructed in Goal Mode

Type of Arguments as Goal

We represented requirements in our model with the NFR framework[2] graphical elements. Some examples from the proposed approach and icon used from the NFR goal model are shown in Table 1. While conducting a detailed analysis about the topic under discussion, we follow the flow of the crowd-users discussion in the user forum. In the Reddit forum, user comments are displayed in a hierarchical structure, as shown in Figure 2. Where each comment constitutes a relation toward another comment. Based on which we can derive a goal model shown in Figure 3. We can see that crowd-users' discussion contains information relevant to requirements decisions and can be modeled as an argumentation structure. Identified requirements-related artifacts from the crowd-users conversation in the user forum are discussed below.

In general, arguments are made of a claim and with its associated evidence. In social media forums, end users' claims could be categorized into different types: a software feature, a crowd-user preference on the software interface design, or crowd-user feedback about their experience with the system, depending on the topic under discussion. Relation type is either attacking, supporting, issue, and suggestion or alternative. Relation type helps in requirements decision-making and identifying conflict-free requirements/features.

Arguments that suggests alternatives or Features: Users could suggest multiple alternatives concerning the topic under discussion. Each of the proposed alternatives is considered an operationalization of a problem represented by a soft goal of the NFR framework, as depicted in Table 1. For example, a user desires that "*a commute button should be changed automatically based on whether you selected public transport to travel or you are driving by yourself, like Google have their other app Google Assistant*" is an alternative of "Adding a commute Tab in Google map interface", which are both serving a higher level softgoal of "Better User Experience for Google Map". While the symbol ~~Suggests~~> is used to represent the alternative or feature relationship between the requirements arguments, as shown in Table. 1.

Arguments that providing facts or evidence: Crowd-users could provide evidence that is either supporting or attacking the goal or sub-goal. Such information is an essential reference for requirements-related decision-making. By identifying the key arguments from the crowd-user conversations in the user forum, we can find users' concerns and preferences [49]. We represented such arguments in our model by using the modeling concept of "Claim" in the NFR framework and the "supporting" and "attacking" relation between arguments are represented by "Make" and "Break" contribution links that are used to either support or weaken a relationship between two arguments in the NFR framework. For example, a crowd-user attack on the former argument is depicted in "Row 2" of Table. 1, as "google can only afford two tabs also, google assistant is a different team app." This argument is further supported by other arguments developed during the ongoing crowd-user discussion in the user forum. To represent the attacking and supporting relationships in the NFR goal model constructed from the crowd-user comments in the user forum, we use ~~Supports~~>, and ~~Attacks~~> symbols.

[2]http://istar.rwth-aachen.de/tiki-index.php

Table 1. Types of arguments identified from user forum

Argumentation types	Examples from Reddit forum	Graphical Icon	Associated Link
Alternative as an NFR Operationalisation	'Can't they just rename the tab dynamically based on whether you selected drive or public transport in the "how you get to work?" question they asked you? This is how Google Assistant handles things.'		Suggests>, goal refinements with contribution, or correlation links
Fact, Belief, Justification as an NFR Claim	"Google app can afford two tabs and google assistant is a different team product, it's impossible for Google to have a good integration across all products."		Supports>, Attacks> goal refinements with contribution links
Emotional Label as an NFR Label	"Like it, Happy with 4 Tabs, Perfect solution."		✓, ✗, W^+, W^- goal labels
Issue, Challenge as an NFR Softgoal	"Using both (drive and transit) but annoying, don't want to fiddle with setting. I walk also, what happen to that? Why tabs in first place? It has a cycle option?"		Questions> goal refinements with contribution, or correlation links

Arguments that expresses emotional attitude: End-users express their emotional attitude in social forums towards software functionalities/features under discussion. Their votes are either positive or negative. Votes against software features should be considered important indicators for whether end-users are satisfied with a given/intended functionality or whether they provide alternative solutions as suggestions to relieve their grudges. Examples of emotional attitude along with desired modeling symbol are shown in Row no. 3 of Table 1. Emotional attitudes are modeled as a label to a given NFR node; positive ones are labeled with "Satisfied" (✓), while negative ones are labeled with "Denied" (✗). When both positive and negative comments exist, we could evaluate them qualitatively as "weakly satisfied" (W^+) or "weakly denied" (W^-). These are built-in NFR semantics used to identify whether the root node (requirement) is conflict-free or not based on their supporting and attacking arguments or emotions.

Arguments that imposing questions or challenges: While doing a detailed and comprehensive analysis of user comments in the user forum using the content analysis approach, it was observed that users may ask questions or pose challenges in response to certain software features or the root node. In response to their question, other crowd-users

may enquire clarification questions or challenge the validity of certain claims. Such arguments are either pending for decisions or lead to a new round of requirements elicitation. Such question helps requirement engineers in clarifying the requirements. A few examples are shown in Table 1, in which a crowd-user asks a question against the root goal "*Why tabs in the first place?*" and a question "*It has a cycle option?*". Furthermore, in response to these questions, other crowd-users submit suggestions or alternatives, which are further discussed by other end-users in the user forum. For example, in Figure 6, following up the question being asked, requirement engineers can further refine the goal model and get useful information about the feature or issue, its alternatives, advantages, and limitations to refine the requirements and help software engineers to implement in smoothly.

ARGUMENTATION INTEGRATED WITH MODEL EVALUATION

In requirements engineering, especially in evaluating goal models, both qualitative and quantitative evaluation approaches are proposed, e.g., the label propagation algorithm in the NFR framework and the formal reasoning approaches goal models. To reach the final requirements decision, we need to handle various decision-making situations as depicted in the following meta-patterns.

(a) Goal satisfaction as Winning Arguments

When a goal receives supporting evidence, sentiments, or comments, it is a winning argument, as shown in Figure 4. Three end-users support the root goal by expressing their positive feelings, i-e, "like it," "good, keep it simple," and "I like it with one tab less." Therefore, such evidence results in the satisfaction of the root goal, a winning argument. There are other cases in the user forum where the support of an argument is the justification/explanation for the claims rather than a simple expression of "like" or "dislike." But even with "like" and "dislike," we can already capture user preferences. By applying the existing NFR semantics, it is identified that the root node is identified as conflict-free and is therefore referred to as a winning argument or requirements.

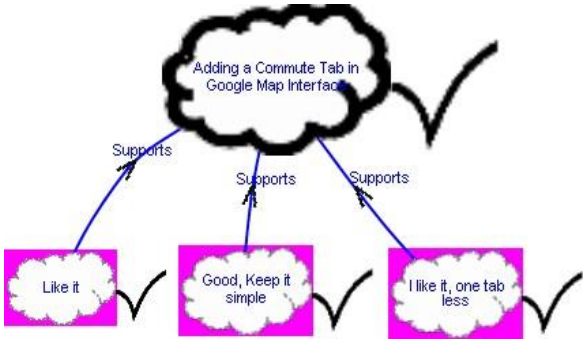


Figure 4. Goal satisfaction as winning arguments

(b) Goal Denial by Counter Arguments

When a goal or sub-goal is attacked by counteracting evidence, sentiments, or crowd-user comments, it is a losing argument or requirement. In Figure 5, two examples of the root

node being attacked by sub-arguments in the ongoing crowd-users discussion in the user forum. In the first example, a user commented, "*where there are driving directions, there is no need for commute if work at home*" attacks the root node, which results in the denial of the root node by the definition of abstract argumentation and NFR goal model semantics. In the second example, the end-user claims that "*Maps are too bloated, explore tab should be another app*.", which is supported by another claim, "*Way too bloated, use to locate only, have a better app about location*," which again results in the denial of the root node by the definition of abstract argumentation and NFR goal model semantics. With the supports and attacks relations, we can adopt the formalisms from the abstract argumentation, where we identify the set of requirements that are conflict-free, complete, grounded, or stable.

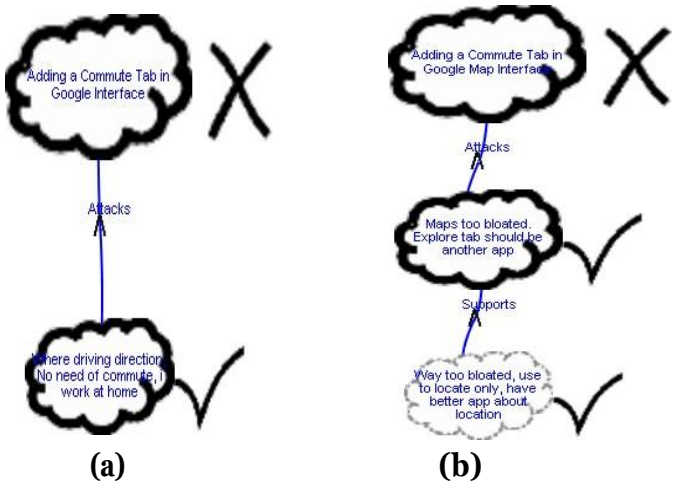


Figure 5. Goal denial by counter arguments

In this case, the arguments "*where there are driving directions, there is no need for commute if work at home*", "*Maps are too bloated, explore tab should be another app*", and "*Way to bloated, use to locate only, have better app about location*" are identified as conflict-free requirements or arguments. There is also the extension mechanisms where rebut denotes a disagreement with the claim, whereas an undercut denotes adisagreement with the support, i.e., an explanation or justification.

(c) Goal satisfaction as Winning Arguments

In the end-user forum, there are cases where crowd-users raise certain questions or challenges in response to the main discussion topic, which needs further elaboration or clarification, against which other crowd-users register comments in the ongoing discussion. One such example is explained in Figure 6, where a crowd-User1 raised a question, "*How practical for those who don't drive or don't take public transits*?" against the main discussion topic "*Adding a commute tab in the google map interface*." Against which another crowd-User2 registered a comment "*commute setting will allow you to enable/disable different travel methods…*" which supports the aforementioned argument. Next, in response to that comment, a Crowd-User3 submits an argument, "*I don't want to do enable/disable all the time, I want to pick….*", which shows a possible new feature for the requirements engineer and an attacking argument at the

same time. Next, Crowd-User4 objects to the above comment that "*Random direction for reaching any place are still available for searching*", a claim based on his experience using the Google map mobile app.

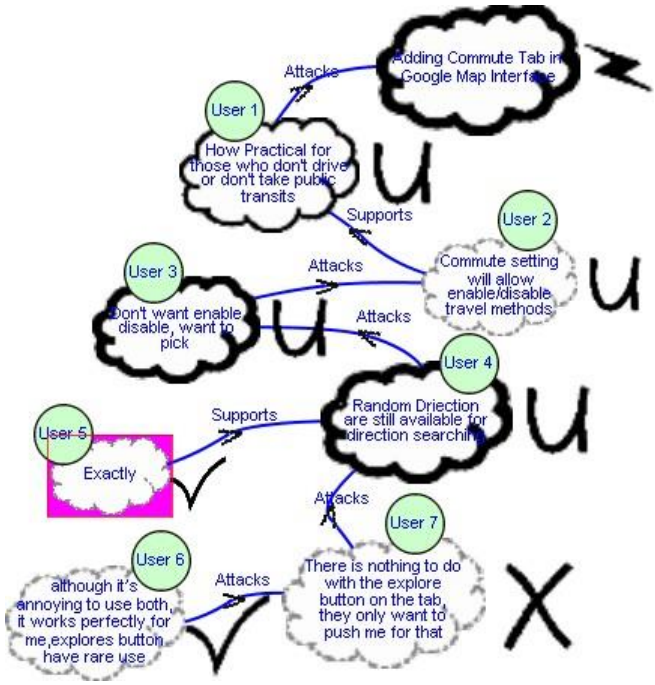


Figure 6. Goal refinement by arguments imposing questions and challenges (parital)

Similarly, a Crowd-User5 supports the previous comment by expressing his positive emotion, while a Crowd-User7 attacks it by claiming that "*there is nothing to do with the explore button on the tab, they only want to push me for that*." Finally, a Crowd-User6 submitted a comment in response that "*although it's annoying to use both, it works perfectly for me*." The symbol ↯ represents conflicting output for a certain requirements-related argument, while the symbol ∪ represents undecided results for a crowd-user comment in the discussion topic on the user forum. It can be concluded from Figure 6 that the root node "*Adding a commute tab in google map interface*" is classified as a conflicting design requirement for the Google Map mobile application, which is due to the certain crowd-users arguments that are labeled as "undecided" during the evaluation of the root node using existing NFR goal model semantics. However, requirements engineers can model such dialogical argumentation and trace the arguments in requirements decision-making by weighing their corresponding supporting and attacking arguments, which leads to a final decision about the given design question. In this case, based on their supporting and attacking arguments, the main discussion topic is identified as conflicting and needs to be ignored.

(d) Goal evaluation by clustering and aggregation of crowd opinions

The overall NFR goal diagram of the ongoing discussion between the crowd-users in the user forum about the Google Map mobile application is depicted in Figure 7. In total, 63/72 crowd-comments were manually collected, among which 54 end-users comments

were recognized as relevant, nine were irrelevant and ignored using the content analysis approach [9], while the remaining nine end-users comments were either deleted by the forum administrator due to the violation of term and conditions or deleted by the end-user because of that we need to delete their corresponding sub-arguments to preserve the original structure of the discussion topic in the forum. Furthermore, 48 unique crowd-users took part in the discussion topic; one user commented five times, four users commented three times, and four users commented twice about the Google Maps mobile app. After the detailed review and analysis of the flow of discussion using the content analysis approach [41], 19 end-users claims were identified that are classified as either supporting or attacking, 24 operationalization/suggestions/ features/alternatives were identified, and seven questions or challenges were raised out of which the end-users further discussed three questions while four questions were left unanswered. Six emotional attitudes are classified again into supporting or attacking opinions.

Finally, we aggregate the annotated crowd-user comments from the online user discussion in the user forum to construct the argumentation-based goal model and reach a conclusion based on their supporting and attacking arguments by applying the existing semantics of the NFR goal model, as shown in Figure 7. the crowd-users register 16 sub-arguments in response to the root node that is represented with the identified relationship (suggestion, supporting, attacking, or question), out of which ten user comments are concluded and referred as winning arguments with (✓) symbol, as shown in Fig. 8. The conclusion of the root node is derived by evaluating the aggregated evidence of their supporting and attacking arguments in the ongoing user discussion in the Reddit user forum. Using the existing semantics of the NFR goal model, the root node "*Adding a commute tab in Google Map interface*" is identified as conflicting, having the symbol ↯ based on their supporting and attacking arguments. Therefore, it is concluded that current goal modeling approaches cannot identify the conflicting viewpoints in the user conversations over the social media platform. It is necessary to develop an argumentation-based goal modeling approach to identify conflict-free arguments using extended abstract argumentation semantics. Further, as shown in Figure. 7, three major crowd-users arguments deny the proposed change in the Google map interface. Also, certain crowd-users provide suggestions in response to the main discussion topic that is considered alternative solutions and leaves it pending further decision. Questions are being raised in the user conversations, which are yet to be answered. These are undecided routes in the goal model which need further elicitation; hence these are not considered in a decisive situation. Some end-users expressed their likes to the root node.

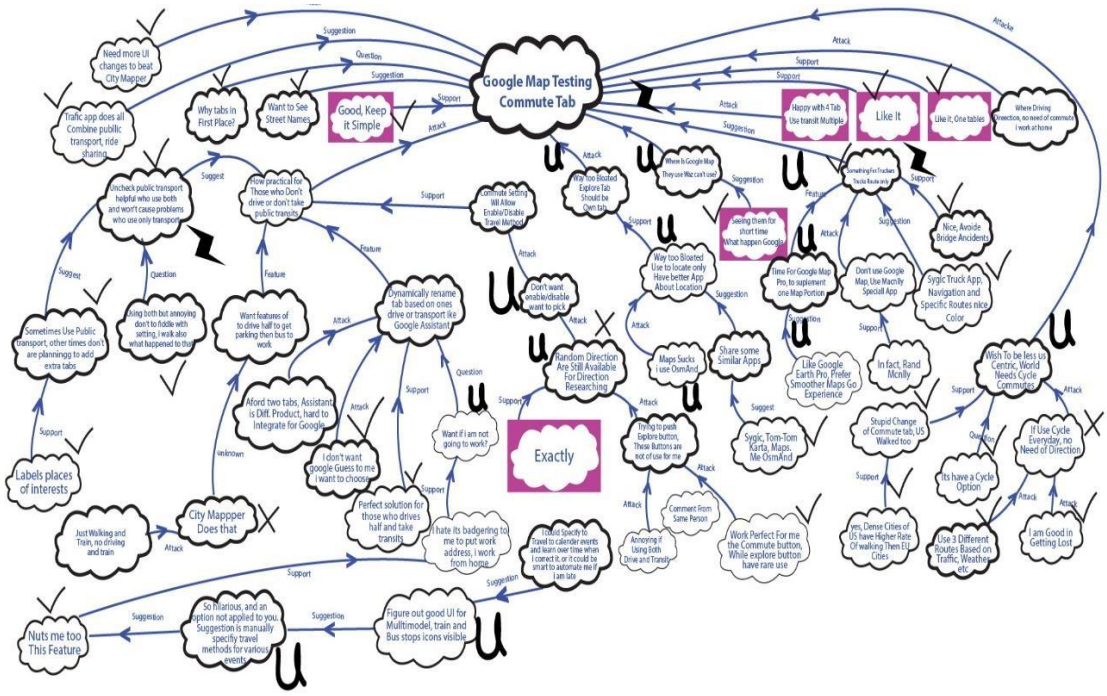


Figure 7. Requirements decision by aggregation of crowd opinions

To conclude, in total, we identified eight attacking arguments that are registered against the main discussion topic in the goal model, getting four supporting arguments or sentiments from the crowd-users in the forum. In contrast, three crowd-users comments were undecided, and one route was neutral. This information provides important insights for requirements engineers in the forthcoming requirements decision-making. It explicitly shows crowd-user arguments, resolves the ambiguities, and identifies potential important requirements/features. However, we will require argumentation-based modeling tools to efficiently identify the conflicting arguments or requirements in the social media platform that would help requirements engineers decide emerging new features or exiting hot issues in the ongoing user conversations.

DISCUSSION

This paper reports the initial results of adopting arguments to crowd-generated requirements information from the end-user forum. The proposed approach is feasible for such application scenarios, where eliciting and modeled arguments align with goal-oriented requirements models, such as the NFR framework. Requirements-related decision-making is supported by the underlying formalisms offered by the computational argumentation framework, where arguments and counter-arguments are recognized and traded-off. However, we will require argumentation-based modeling tools to efficiently identify the conflicting arguments or requirements in the social media platform that would help requirements engineers decide emerging new features or exiting hot issues in the ongoing user conversations as current goal modeling tools lack in identifying conflict-free arguments using the existing semantics. In our proposed approach, Where argumentation provides a

natural and important tool to formalize requirements engineering process and artifacts, especially for handling inconsistent and incomplete requirements information and capturing dialogical requirements elicitation. When such argumentative abilities are explicitly modeled, requirements decisions are made traceable, rationalized, and verifiable.

Authenticity of the study: Crowd Requirements engineering (CrowdRE) is an emerging field used to elicit, validate, model, prioritize, and negotiate software requirements for market-driven software applications. The main source of collecting requirements-related information for market-driven applications is social media platforms. Whereas, in the online software applications distribution platforms such as google play store, Apple's play store , Amazon software application store, or other mobile applications stores, end-users easily search, install, and give feedback's or comments on software applications. According to a research study, nearly over 3 million software applications are available in the Google Play and App store, as of June 2020, with over 75 billion software application downloads per month. This software applications distribution is not limited to mobile applications anymore. Other software types such as desktop applications, software plugins, and other open-source software are now available via app stores. For example, the Eclipse Marketplace for the Eclipse Development Environment has approximately 25 million active Bundles, plugins, and different products available. Also, specialized crowd-users feedback platforms based on the principle of crowdsourcing, where crowd-users can suggest, comment, and vote on the possible idea proposed by the other crowd-users, are getting immense popularity; one such platform is UserVoice. Considering such a large amount of pivotal information for requirements engineers and software developers, we take a step to the model end-user conversation in the social media platform with the existing goal of modeling software to recover conflicting requirements-related information. The novelty of the proposed work is utilizing the existing NFR goal modeling tool and its semantics to recover conflicting requirements information, according to our knowledge, which has been demonstrated with the Google Maps mobile application case study.

Applicability of the proposed Approach: The proposed approach is an experimental study on utilizing the existing NFR goal modeling tool and its semantics to identify conflicting requirements-related information. The proposed approach can be utilized as the first step in modeling crowd requirements to identify conflicting, incomplete, unclear, and ambiguous requirements information. Also, the proposed approach encourages to development of an automated requirements validation and modeling approach to efficiently model crowd requirements. For this purpose, we can adopt natural language processing, machine learning classifiers, and argumentation theory to automate the process. Additionally, the modified proposed approach can automatically identify conflict-free new features, design alternatives, issues, and their winning arguments by using abstract argumentation theory from the user conversation in the user forums underneath rationale. It will help requirements engineers decide conflict-free emergent requirements or hot issues faced by end-users underneath rationale. The process will reveal the end-users tactic knowledge and make the decision-making process transparent by capturing the rationale and documenting the requirements.

Furthermore, we elaborated earlier that the proposed approach concluded that existing NFR goal modeling tools are limited in modeling and identifying conflict-free requirements-related information in the social media platforms based on experimental study. Hence, we can grab this opportunity to develop and market the CrowdRE modeling tool to identify conflict-free emergent requirements or hot issues faced by end-users underneath rationale in the social media platforms.

CONCLUSION AND FUTURE WORK

This paper analyzed and evaluated the interaction between the crowd-users about the Google Map mobile application in the Reddit forum to recover conflicting requirements-related information using the goal modeling approach. For this purpose, we first extracted critical arguments from a crowd-users conversation in user forums regarding a given design decision, a new feature proposed, or an issue identified. Secondly, we build a graphical argumentation model using the extracted arguments from the end-user discussion. Next, we aligned the extracted arguments with goal-oriented modeling constructs in the non-functional requirements framework. Finally, we utilized the exiting goal-model analysis and semantics to the requirements argumentation model to reach a consensus-based on argumentation theory and reasoning. Additionally, we demonstrate the proposed approach with a crowd-users conversation on Google Map interface design from the Reddit user forum. The current argumentation modeling case study is based on historical data, which is after the fact. In the future, if we can embed the argumentation analysis during an ongoing discussion and use it to direct the requirements process on the run, the interests of requirements decision-makers can be better served, as more valuable and timely recommendations can be provided. In the future, the scalability of the proposed method needs to be studied, and the prospective automated tool chain is to be implemented to address the needs of significant complex engineering problems.

REFRENCES

[1] E. Yehia., F. Alessio, and J. Mylopoulos: "*CaRE: A refinement calculus for requirements engineering based on argumentation semantics*". In. Preceding of 26 IEEE international conference on Requirement Engineering (RE@Next! 18), pp 364-369 (2018)

[2] R.Z. Moghaddam, B.P. Bailey, and C. Poon: "*Ideatracker: an interactive visualization supporting collaboration and consensus building in online interface design discussion*" In: Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction- Volume Part I. INTERACT'11, Berlin, Heidelberg, Springer-Verlag (2011) 259-276.

[3] P. M. Dung, "*On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games*," Artif. Intell., vol. 77, no. 2, pp. 321–357, 1995.

[4] J. A. Khan, L. Liu, L. Wen, and A. Raian: *"Crowd Intelligence inRequirements Engineering: Current Status and Future Directions*". Proc. Int. Conf. Requirements Engineering: Foundation for Softwarequality, Essen, Germany, March 2019, pp 245–261

[5] J. Mylopoulos, L. Chung, and B. Nixon: "*Representing and using non-functional requirements: A process-oriented approach. Software Engineering*", IEE transaction on, 18(6): 483-497, 1992.

[6] E. Guzman, R. Alkadhi, and N. Seyff: "*A Needle in a Haystack: What Do Twitter Users Say about Software*"?. In. Preceding of RE 2016 International conference: pp. 96-105 (2016).

[7] R. Sebastiani, P. Giorgini, J. Mylopoulos: "*Simple and Minimum-Cost Satisfiability for Goal Models*". CAiSE 2004: 20-35
[8] P. Giorgini, J. Mylopoulos, E. Nicchiarelli, and R. Sebastiani: "*Reasoning with Goal Models*". ER 2002: 167-181
[9] V. N. L. Franqueira., Thein T. Tun., Y. Yu, R. Wieringa, B. Nuseibeh: "*Risk and argument: A risk-based argumentation method for practical security"* In. 19th IEEE International Requirement Engineering Conference, RE 2011, pp. 239–248 (2011).
[10] I, J. Jureta, J. Mylopoulos, and S. Faulkner: "*Analysis of multi-party agreement in requirements validation*". In. 16th IEEE International Requirement Engineering Conference, RE 2009, pp. 57–66 (2009).
[11] G. M. Kanchev and A. K Chopra: "*Social media through the requirements lens: A case study of Google maps*". CrowdRE@RE 2015: 7-12
[12] D. Ionita, J. Bullee, R. J. Wieringa: "*Argumentation-based security requirements elicitation: The next round*". In. 1st IEEE international workshop on evolving security and privacy requirements engineering (ESPRE), pp 7-12 (2014).
[13] Eclipse Foundation. Eclipse marketplace[EB/OL]. https://marketplace.eclipse.org/, Last accessed on December 2019.
[14] User Voice. Roadmap prioritization from product feedback[EB/OL]. https://www.uservoice.com/, Last accessed on December 2019

[15] E. Toulmin, "*The Uses of Argumentation"*, Cambridge Univ. Press, vol. 37, no. 2, pp. 168–182, aug 1959.
[16] M. V. Zee, D. Marosin, F. Bex, S. Ghanavati.: "*RationalGRL: A framework for rationalizing goal models using argument diagrams*". In. Proc. of the 35 Int. Conf. on Conceptual Modeling (ER'2016), Short paper, pp 553-560 (2016).

[17] T. T. Tun, A. K. Bandara, C. Haley, I. Omoronyia, and B. Nuseibeh.: "*Privacy Arguments: Analysing selective disclosure requirement for mobile application*". In. 20th IEEE international requirement engineering conference, pp 131-140 (2012)
[18] E. Yehia, F. Alessio, P. Spoletini, G. Stefania, and B. Nuseibeh,: "*Using argumentation to explain ambiguity in requirement elicitation interviews*". In. 25th IEEE international conference on requirement engineering, (2017).
[19] I. Mirbel and S. Villata: "*Enhancing Goal-based Requirements Consistency: an Argumentation based Approach*". In. Int. Work. Compute. Log. Multi-Agent Syst., pp. 110–127 (2012).
[20] M. R. Itzel and A. Perni: "*Argumentation-based discussion for user forum: A Research Preview*". In. Preceding of the 20th International working conference on requirements engineering: Foundation for software quality (REFSQ 2014), pp 232-238 (2014).
[21] G. M. Kanchev, P. K. Murukannaiah, A. K. Chopra, P. Sawyer. "*Canary: Extracting requirements-related information from online discussions*". In. Preceding of 25th International Requirement Engineering Conference (RE17), pp 31-40 (2017).
[22] S. Ingolfo, A. Siena, J. Mylopoulos, A. Perini. "*Arguing regulatory compliance of software requirements*". Data Knowledge Engineering (DKE), vol. 87, pp. 279– 296, (2013).
[23] R. Muthuri, G. Boella, J. Hulstijn, L. Humphreys. "*Argumentation-based legal requirements engineering- The role of legal interpretation in requirement acquisition*". In. 24th IEEE international requirement engineering conference workshops (RELAW), pp 249-258 (2016).

[24] J. A. Khan, L. Lin, Y. Jia, and L. Wen: "*Linguistic Analysis of Crowd Requirements: An experimental study*". In. preceding of 27th IEEE international conference on requirement engineering workshop (Empri 18), pp (24-31) (2018).
[25] SC.B. Haley, J.D. Moffett, R. Laney, and B. Nuseibeh, "*Arguing Security: Validating Security Requirements Using Structured Argumentation*," Proc. Third Symp. Requirements Eng. for Informa- tion Security with the 13th Int'l Requirements Eng. Conf., 2005.
[26] C. B. Haley, R. Laney, J. D. Moffett, and Bashar, N.: "*Security requirements engineering: A framework for representation and analysis*", TSE, vol. 34, no. 1, pp. 133–153, 2008.
[27] Y Yu, V. N. L. Franqueira, T. T. Tun, R. J. Wieringa, B. Nuseibeh: "*Automated analysis of security requirements through risk-based argumentation*". JSS, 106, 102-116 (2015).
[28] A. Kovacs, I. Karakatsanis, and D. Svetinovic.: "*Argumentation-based security requirements analysis: BitMessage case study*". In. IEEE international conference on Internet of Things (iThings), and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom), pp 408-414 (2014).
[29] I. J. Jureta, F. Stephane, S. Pierre-Yves: "*Clear justification of modeling decision for goal-oriented requirement engineering*". Requirement Engineering Journal (2008), 87-115 (2008).
[30] E. Yu, and J. Mylopoulos: "*Why goal-oriented requirement engineering*". In. proceeding of the 4th International workshop on requirement engineering: Foundation of software quality. Vol 15, page 15-22, 1998.
[31] M. V. Zee, S. Ghanavati: "*Capturing Evidence and Rationales with Requirements Engineering and Argumentation-Based Techniques"*. In Proc. of the 26th Benelux Conf. on Artificial Intelligence (BNAIC2014), (2014).
[32] M. V. Zee, F. Bex, and S. Ghanavati: "*Rationalization of Goal Models in GRL using Formal Argumentation*". In. Proc. of RE: Next! Track at RE'15, August 2015.
[33] Van et al.: "*Bug Report, Feature Request, or Simply Praise? On Automatically Classifying App Reviews*". In. Proc. 23rd IEEE Int'l Requirements Eng. Conf. (RE 15), pp. 116–125(2015).
[34] S. Ghanavati, M. V. Zee, and F. Bex: "*Argumentation-based methodology for Goal-oriented Requirement Language (GRL)*". In. Proceedings of the 29h international conference on Advance information system engineering (CAiSE 2017), pp 97-102 (2017).
[35] E. Bagheri and F. Ensan: "*Consolidating Multiple requirement specifications through argumentation*". In. Proceeding of the ACM symposium on applied computing (SAC), pp 659-666 (2011).
[36] Elrakaiby, P. Baroni, M. Giacomin, A. Hunter, H. Prakken, C. Reed, G. Simari, M. Thimm, and S. Villata (2017*)"Towards Artificial Argumentation"*, AI Magazine, 38(3):25-36.
[37] P. K. Murukannaiah, A. Kalia, P.R. Telang, and P. S. Munidar: "*Resolving goal conflicts via argumentation-based analysis of competing hypotheses*". In. Preceding of 23th IEEE International conference on Requirement engineering (RE 2015), pp 156-165 (2015).
[38] Z. Kurtanovic, W. Maalej. "*Mining user rationale from software reviews*". In. Preceding of 25th international conference on requirement engineering (RE 17). (2017).
[39] Z. Kurtanovic, W. Maalej. "*On user rationale in software engineering*". Requirement engineering journal, pp 357-379 (2018).
[40] J. A. Khan, Y. Xie, L. Liu, L. Wen. "*Analysis of requirements-related arguments in user forums*". Proceedings of the IEEEInternational Conference on Requirements Engineering 2019; 2019-September: 63–74

[41] J. A. Khan, L. Liu, L. Wen L,A. Rain. "*Conceptualising, extracting and analysing requirements arguments in users*' forums: TheCrowdRE-Arg framework. Journal of Software: Evolution and Process: e2309
[42] J. A. Khan, L. Liu, and L. Webn. "*Requirements knowledge acquisition from online user forums*". IET Software 2020; 14(3): 242–253
[43] K.A.Neuendorf.: "*The content analysis Guidebook*" .Sage Publications,2002.
[44] Google. Play store[EB/OL]. https://play.google.com/store/, Last accessed on 2020-09-04.
[45] Apple. App store[EB/OL]. https://www.apple.com/ios/app-store/, Last accessed on 2020-09-04.
[46] Amazon. Amazon software[EB/OL]. https://www.amazon.com/design-downloadbusiness-education-software/b?node=229534, Last accessed on 2020-09-04.
[47] Statista. Number of available applications in the google play store from december 2009 to june2020[EB/OL]. https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/, Last accessed on July 2020.
[48] Statista. Number of available apps in the apple app store from 2008 to 2020[EB/OL]. . https://www.statista.com/statistics/268251/number-of-apps-in-the-itunes-app-store-since-2008/, Last accessed on July 2020.
[49] Statista. Cumulative number of apps downloaded from the apple app store from july 2008 tojune 2017 (in billions)[EB/OL]. https://www.statista.com/statistics/263794/number-of-downloads-from-the-appleapp-store, Last accessed on July 2020.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Finger-vein Image Enhancement and 2D CNN Recognition

Noroz Khan Baloch[1], Saleem Ahmed[2], Ramesh Kumar[2],
[1] Dept. of Electronics Engg. Dawood University of Engineering & Technology Karachi, Pakistan.
[2] Dept. of Computer System Engg. Dawood University of Engineering & Technology Karachi, Pakistan.
Noroz Khan Baloch, Email ID: b.noroz820@yahoo.com.
Citation | Baloch. N. K, Ahmad. S and Kumar. R, "Finger-vein Image Enhancement and 2D CNN Recognition". International Journal of Innovations in Science and Technology. Vol 3, Special Issue, pp: 33-44, 2021.
Received | Dec 15, 2021; Revised | Dec 25, 2021 Accepted | Dec 26, 2021; Published | Dec 30, 2021.

Finger vein recognition technology is a novel biometric technology with multiple features such as live capture, stability, difficulty in stealing and imitating, and more in the field of information security that has been utilized in a wide range of applications. In this proposed method, the finger region is separated from the background using a Sobel Edge detector and a Poly ROI which helps shape the finger. The background separation enhancement of low contrast using dual contrast limited adaptive histogram equalization which works on the visual characteristics of the finger-vein image dataset. When dual CLAHE is applied, the finger-vein histogram intensity is separated all across the image. Following the implementation of DCLAHE, an enhanced 2D-CNN model is utilized to recognize objects with the updated dataset. By maximizing the values of a preprocessed dataset, the 2D CNN model learns features. This model has a 94.88% accuracy rate.
Keywords: biometric; contrast limited adaptive histogram equalization; Sobel edge detector; poly region of interest and two dimensional convolution neural network.

INTRODUCTION

Personal identification is currently achieved using biometric technology used in a variety of security applications such as mobile phones, computer access, border crossing, banks, and ATMs. Fingerprint, iris, face, voice, and finger-vein recognition are some of the methods used (FV) [1, 2]. Because the vein patterns beneath the skin differ, the finger-vein recognition system has gotten a lot more attention in recent years. As a result, finger-vein recognition suggests a safe and practical technique for human biometric classification [3]. Traditional methods for recognizing finger veins rely on extracting lines from the input image or amplification and feature extraction [4]. Near-infrared (NIR) light is used to capture vein

images on the finger, and a charged coupled device (CCD) camera is used to record vein images beneath the finger [5]. Because the dataset needed to be normalized for the preprocessing step, the bitmap image was turned into a grayscale dataset with a single dimension for preprocessing and 2D CNN implementation [15]. When NIR penetrates the dorsal side of the finger and a CCD camera penetrates the palmar side, the region of the finger vein appears darker [6].

Finger vein datasets can be found in a variety of formats. The proposed approach for finger vein recognition was evaluated using the SDUMLA-HMT database [7]. Other finger vein datasets include HKPU-FV [8], UTFV [9], MMCBNU 6000 [10], THU-FV [11], and others. Local binary pattern (LBP) [12], Gabor filters (GF) [13], local derivative pattern (LDP) [12], and other techniques have been utilized to recover finger-vein images from the backdrop. Because these techniques have a weakness in similar settings such as translation and rotation during the capture of finger vein images, the deep learning technique known as convolutional neural network (CNN) is used for the classification of finger vein images. CNN/ConvNet is a deep learning algorithm that is faster than standard techniques [14].

Our suggested technique starts with a preprocessed Dual CLAHE [7] image dataset and feeds these preprocessed finger-vein images into a two-dimensional CNN model. Before classifying finger veins and applying enhancement techniques, our model uses a Sobel edge detector and a polygonal region of interest (ROI) to extract the background of the finger region [15]. As a result, we employ our image for histogram equalization follow extraction, which improves the image dataset. A histogram is a graphic representation of the distribution of greyscale values of pixels in a greyscale image [16]. The histogram's equalization distributes grey values uniformly throughout the image. Dual contrast limited adaptive histogram equalization (CLAHE) is used twice in the suggested technique to boost grey values in a finger vein image. Based on the folder names, the processed dual CLAHE finger vein images are classified. These pictures are then organized and utilized in the suggested 2D CNN model. In contrast to prior methods for extracting finger veins from picture datasets, 2D CNN algorithms can detect finger veins on a large scale of the dataset by learning features and recognizing them without removing background characteristics. According to the findings, 2D CNN is capable of locating finger veins with a high degree of accuracy.

The paper is organized as follows. The related study on finger-vein detecting technology is described in Section 2. Section 3 explains the finger-vein background extraction, Dual CLAHE, and 2D CNN architecture. The model's result and explanation are shown in section 4, and the conclusion is presented in section 5.

RELATED WORK

According to prior studies [1, 4], preprocessing and feature deduction approaches are commonly used in the detection of finger veins. The Gabor Filters have variable forms and directions to locate the finger-vein pattern [17]. The researcher utilized the local binary pattern (LBP) [18] and local derivative pattern (LDP) techniques to extract finger-vein patterns in [12]. The Gabor filter is used to match the SIFT features of vein patterns in the finger-vein classification [19, 20]. The principal component analysis (PCA) technique is used to create different traits [21]. Finger-vein feature extraction using sparse representation, the maximum curvature approach, HOG, and SVM are some of the documented strategies [23].

The algorithms for recognizing finger veins described above are standard. To improve the performance of finger-vein recognition, the researcher uses deep learning CNN-based algorithms instead of standard approaches [1]. In [24], the author developed a CNN technique for finger recognition that combines a reduced-complexity four-layer with fused convolutional-subsampling. In [24] suggested a CNN approach that consists of seven layers, five of which are convolutional layers and two of which are fully connected layers. The challenge of irregular size for features is acceptable resolved with a PCA-based dimensionality reduction strategy in the [26] paper, which employs a method that uses a multimodal finger fusion based on CNN. These procedures, on the other hand, are either more difficult or incapable of delivering the required results. On the other extreme, these treatments are either more complex or incapable of producing the desired effects.

THE PROPOSED METHOD

Overview. In this study, finger-vein recognition is broken down into six steps, as shown in the flow diagram in Figure 1. The SDUMLA-HMT picture database [7], which is a free collection of finger-vein images, is used to read the first photos. As observed in Figure 2(a) and Figure 2(b), the finger-vein images in the dataset are standardized, which is normalized in the second phase, which changes the image dataset into greyscale images of size, as seen in Figure 2(a) and Figure 2(b). The third step is to extract our backdrop from the finger-vein pictures dataset (see Figure 2(c)). The Sobel edge detector and polygonal ROI are used to obtain the finger region. The finger-vein picture dataset is first run through the dual CLAHE, which produces dark vein ridges, before being fed into the 2D CNN deep learning model. The architecture of the 2D CNN with twenty layers, as shown in Figure 9, is shown below.

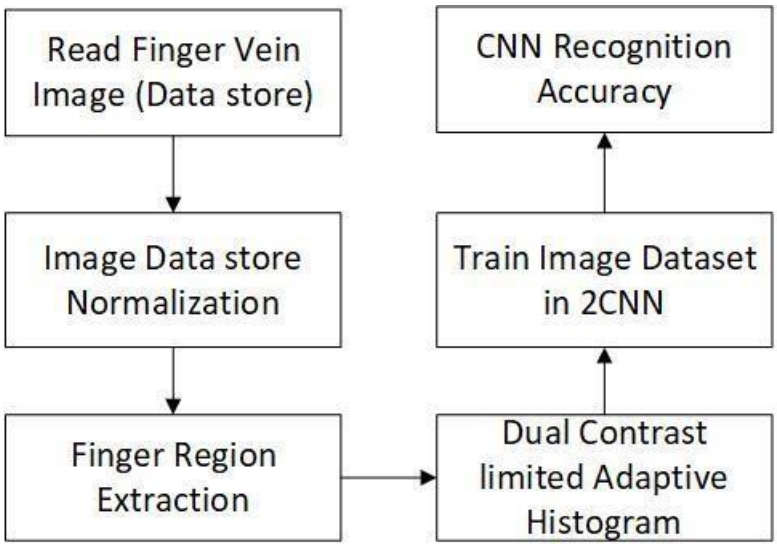


Figure 1. The proposed method's flow diagram.

Sobel Edge Detector and Poly ROI.

At the points where the gradient is the highest, it appears ahead of the edges. By performing a 2-D spatial gradient quantity on a picture, the Sobel method accentuates regions of high spatial frequency that correlate to edges. It only takes the finger boundary of the finger-

vein image [31]. So after the Sobel edge detector detects some areas of the finger and detects some of the backgrounds at the corner of the image, it cannot discriminate the finger region, so for this, it uses polygonal ROI to shape the finger part. Polygonal ROI creates a binary mask for the finger region to subtract the finger from the background portion and take the finger to detect finger-vein edges [32]. Figure 2(c) illustrates the result of separating the finger region from the background part.

Dual CLAHE. Considering that the contrast of the dataset's finger-vein images is too low, the finger-vein cannot be distinguished from the background or skin region, we utilized a dual CLAHE enhancement technique [7], in which it is applied twice to the finger-vein image. The finger-vein image is cropped and reduced using this technique from the black region, which is the background portion of the original image that has been removed. We used dual CLAHE to enhance the vein region twice after cropping the finger-vein image, as shown in Figure 3. CLAHE is a modified form of Adaptive Histogram Equalization (AHE) that works on small parts of the contextual region [28] in the finger-vein picture known as tiles 22 on both CLAHE and AHE. CLAHE with the same exponential distribution [7], [27], with clipping limits of 0.03 on the first and 0.04 on the second.

$$N_{avg} = \frac{N_r X \times N_r Y}{N_{gray}} \quad (1)$$

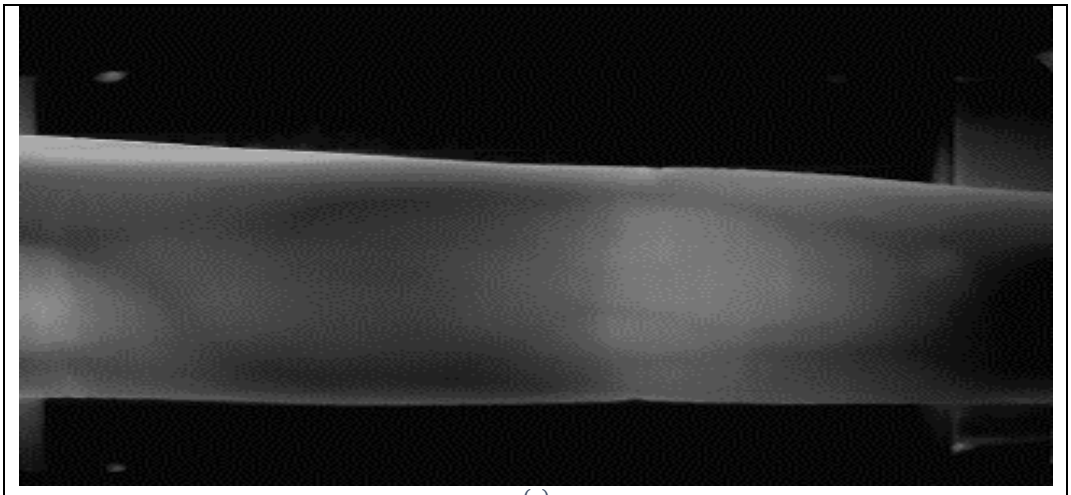
The average number of pixels is represented by N_{avg}, the number of gray-scale pixels is represented by N_{gray}, and pixel amounts in the X and Y directions are represented by N_rX and N_rY, respectively, as indicated in Eq. (1).

The CL (clipped limit) is a value that is limited to a specific threshold where grayscale is darker and is considered finger-vein as follows:

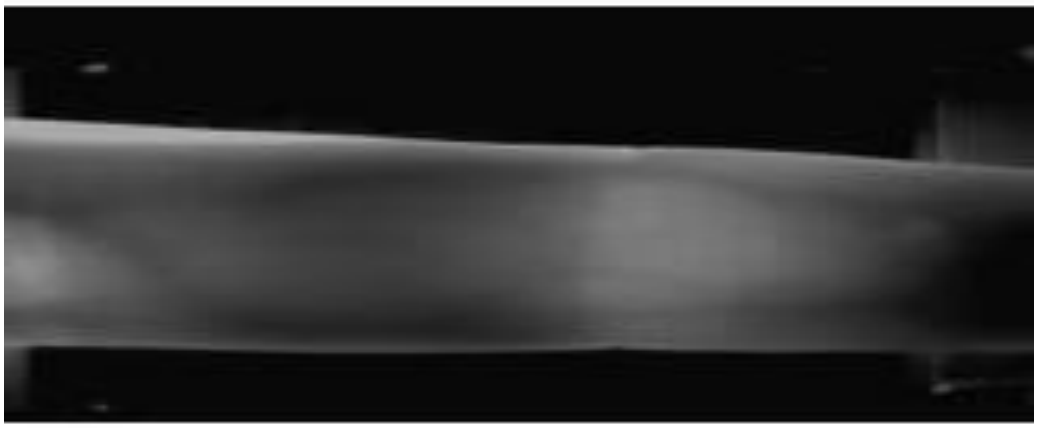
$$N_{CL} = N_{clip} \times N_{avg} \quad (2)$$

$$\mathrm{N_{avggray}} = \frac{N_{\Sigma clip}}{N_{gray}} \quad (3)$$

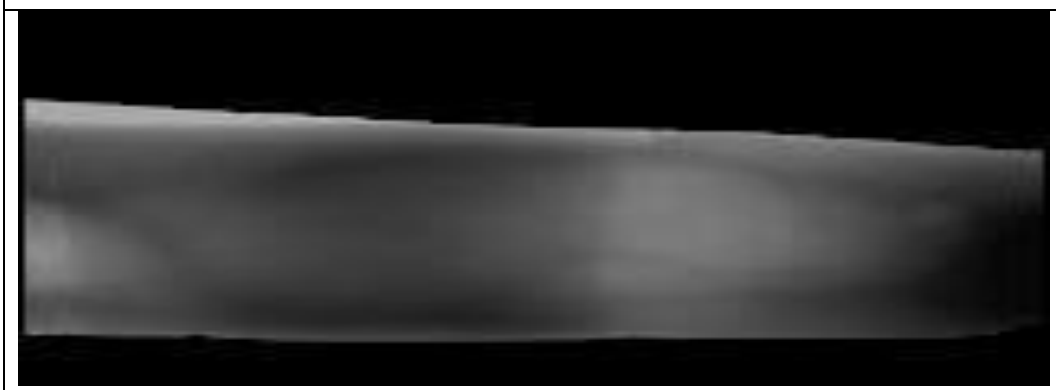
Within the range of the [0 1] array, the actual CL is N_{clip}, where the real CL is N_{CL}. Only when it exceeds the pixel number N_{CL} is it clipped. $N_{\Sigma clip}$ is the total number of pixels clipped, and the remaining average grey-level pixels are distributed as in Eq. (3) [7], [27].



(a)



(b)



(c)

Figure 2. (a) Original Finger Vein 240×320×3, (b) Normalized Image 120×160×1, (c) Finger Region Extraction using Sobel edge detector and polygonal ROI.

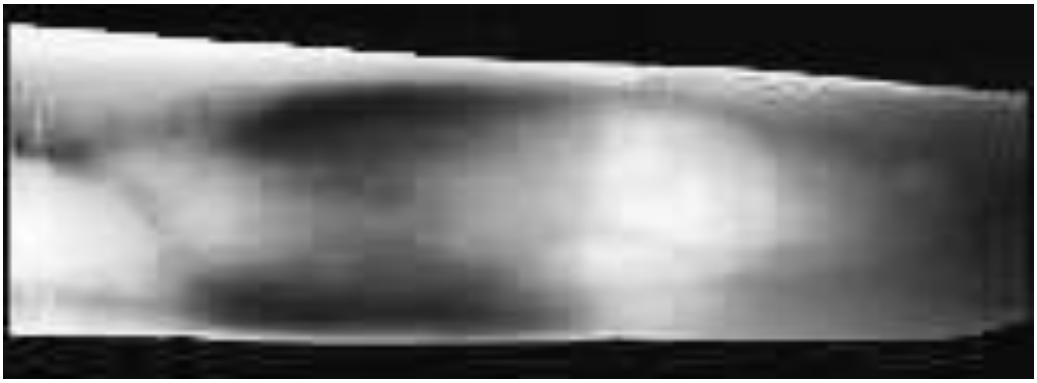


Figure 3. Dual CLAHE implementation.

2D CNN Network Architecture. The image input layer has a size of pixels, and there are four 2D convolutional layers, four batch normalization layers, and four rectified linear units (ReLU) with three max-pooling layers, two fully connected layers, one softmax layer, and a classification output layer. Learning characteristics from finger veins is 94.38% accurate.
The convolutional layer concatenates different sizes of filters on the finger-vein image dataset to learn features of the finger-vein image for the image recognition system, as demonstrated in Figure 4 which shows the initial convolutional layers. The first convolutional layer employs five filters and [2 2 2 2] padding to maintain the image size by adding two zeros to the top, bottom, left, and right of the finger-vein image matrix. Ten, twenty, and thirty filters of various sizes are used in the convolutional layers, as shown in Figure 9.

$$y = \sum_{c=1}^{N} \omega_c * x_c + b \quad (4)$$

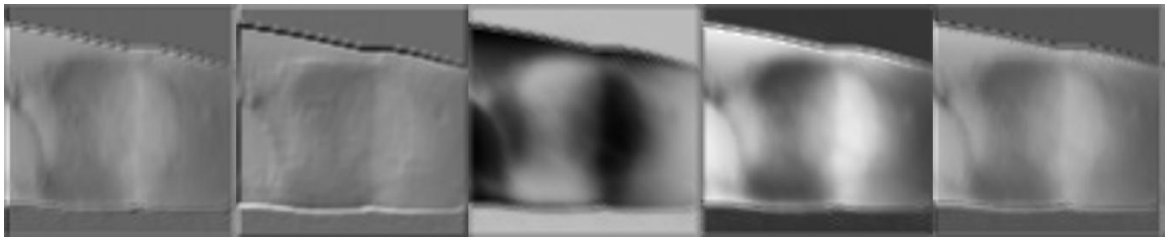


Figure 4. First Convolution Output Layer.

ω_c is the weight array of a 2D filter in the c^{th} input channel, x_c is the 2D input of the c^{th} channel filter, and b is the bias of a filter in the above Eq. (4). As seen in Figure 5, batch normalization substitutes this bias with a β–shift factor. The sizes of channels may fluctuate with the bias factor after convolving the finger-vein input images with different filters, hence batch normalization normalizes the channels for ReLU activation [28].

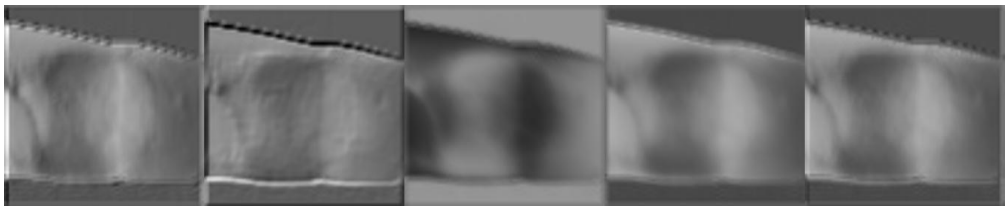


Figure 5. Batch Normalization Layer Output Channels.

The ReLU layer, as shown in Figure 6 [5], is an activation function that increases nonlinearity after batch normalization. ReLU employs the maximum pixel value in this model. Pixels less than zero are referred to as 0, whereas pixels larger than zero are referred to as y. Non-linear functions such as tanh and sigmoid were previously employed. Eq. (5) [1], [5] shows how the researcher determines the ReLU:

$$f(y) = \max(0, y) \tag{5}$$

Figure 6. ReLU Layer Output Channels.

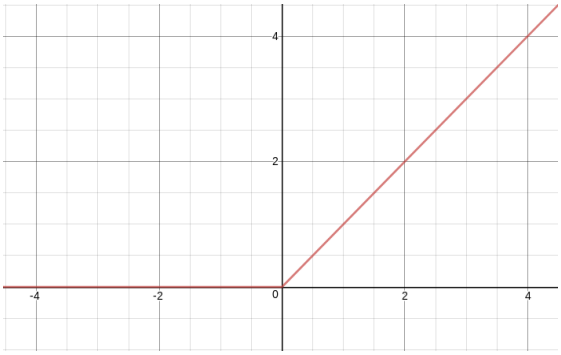


Figure 7. Graph of ReLU Layer.

The max-pooling layer with a filter matrix will take the maximum values, taking only the maximum number of channel values. The stride of the max-pooling layer is [2 2] with padding. The maximum pooling output is shown in Figure 7.

We apply this function of the fully connected layer two times $t = f_c(x, \omega, b)$, where x represents the input of this layer, w represents the weight matrix, b represents the bias, and t represents the output [29], which is as described in Eq. (6):

$$t_{i^F} = \sum_i w_{ii^F} x_i + b_i \tag{6}$$

For each image location used to classify pixels into one of the initial labels where the linked object classes belong, the second final layer of 2D CNN generates a probability

distribution. For multi-elegance category issues, the category output layer employs the pass-entropy lack of reciprocally absolute instruction [30].

Figure 8. Layer Output Channels of Max-pooling.

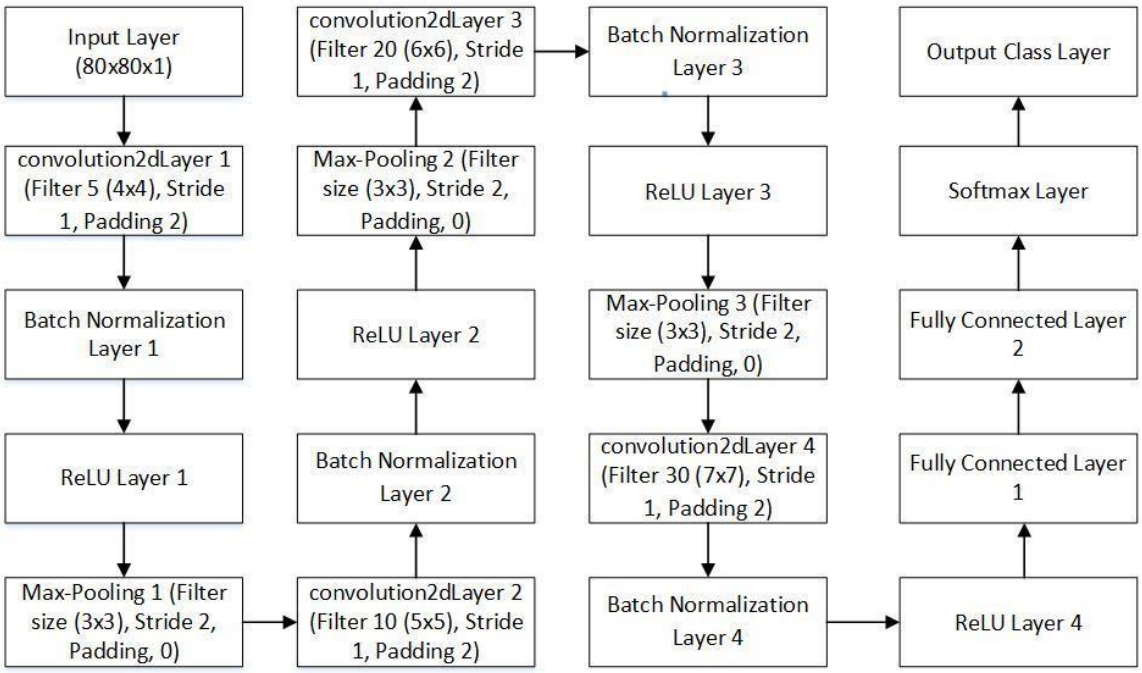


Figure 9. Basic architecture of 2D CNN.

RESULT

The preprocessed dataset of the dual CLAHE finger-vein picture dataset is randomly divided into 80 percent training, 20 percent validation, and 20 percent test images, with four index, middle, and ring finger classes in each folder. These divided data are fed into an Adam optimizer-based train network, which optimizes the learned values from our CNN model and detects features in each image. In this version, we extract four preprocessed finger-vein pictures from each folder of the left and proper hands, labeling them with the names of the index, center, and ring palms. The training option employs an Adam optimizer with a maximum epoch of 80, with one epoch equaling one complete cycle and twelve iterations. The training progress report in Figure 10 shows that after 960 iterations, the validation accuracy for learning the dual CLAHE preprocessed image dataset with various classes is 94.88 %.

The Adam optimizer is used to train the 2D CNN model and update network weights. The epoch values, iteration, iteration per epoch, and validation frequency are shown in table (1) below. Figure 11 depicts a model loss in which the weight may be replenished and the loss reduced on subsequent review. The blue line in Figure 10 depicts the training (smoothed), which is less noisy. The light blue line represents training, which classifies individual mini-batches, while dotted or dashed lines indicate validation, which classifies the complete collection. In Figure 10, the blue line illustrates the training (smoothed), which is less noisy. The light blue line is the training that classifies individual mini-batches, and validation is represented by dotted or dashed lines that classify the entire set. The training loss on each mini-batch is shown in Figure 11 in light orange, the smoothed training loss in orange, and the validation loss in dotted or dashed lines.

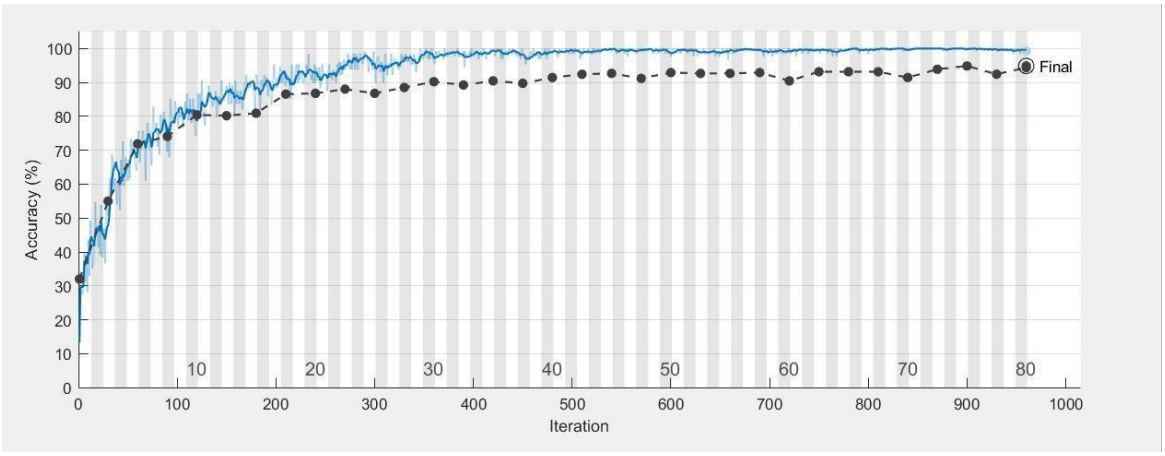


Figure 10. 2D CNN accuracy graph.

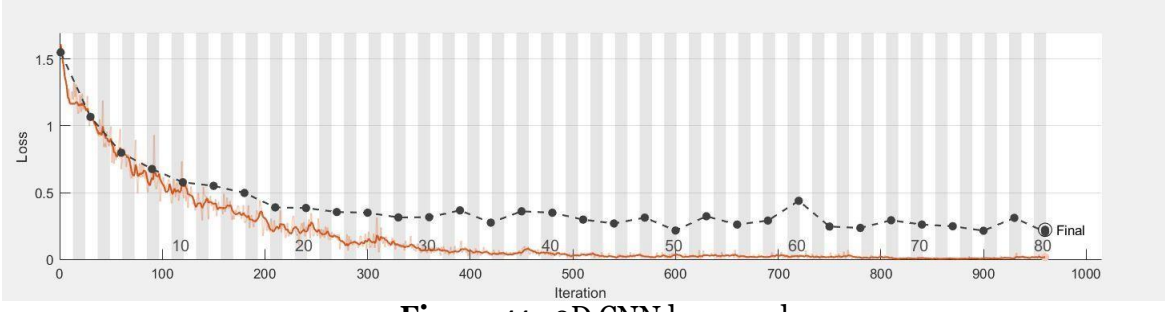


Figure 11. 2D CNN loss graph.

Table 1. Table describing sample of IJIST.....

#	Parameter	Values
1	Validation accuracy	94.87%
2	Epoch	80
3	Iteration	960
4	Iteration per epoch	12

5	Validation Frequency	30 iteration

CONCLUSION

In this research work, we used 2544 images to separate the finger region from the background using a Sobel edge detector and poly ROI, which shaped the acquired edges of the finger and created a binary mask to separate out the finger. The finger-vein dataset is enhanced using dual CLAHE and after enhancing the finger-vein images, these datasets are further utilized in 2D CNN where the model learns the features of every segmented finger-vein. The finger-vein images are divided into distinct classes of folders such as index, middle, and ring folders, with 2D CNN taking four images from each folder. Hence, our version became capable of examining the traits of every photo as well as distinguishing every photograph inside folders. The suggested model's accuracy is 94.88%.

Acknowledgement. The author expresses gratitude to the Department of Electronic Engineering, and Computer Science at Dawood University of Engineering and Technology in Karachi, Pakistan, where the research has been carried out.

Author's Contribution. Noroz Khan is the author and conducted this research under the supervision of Saleem Ahmed. Ramesh Kumar, the co-supervisor.

Conflict of interest. There is no conflict of interest declared by the authors.

REFRENCES

[1] Liu, Wenjie, Weijun Li, Linjun Sun, Liping Zhang, and Peng Chen, "*Finger vein recognition based on deep learning*," In 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), p. 205-210, IEEE, 2017.

[2] Pham, Tuyen Danh, Young Ho Park, Dat Tien Nguyen, Seung Yong Kwon, and Kang Ryoung Park, "*Nonintrusive finger-vein recognition system using NIR image sensor and accuracy analyses according to various factors*," Sensors 15, no. 7 (2015): 16866-16894.

[3] Mulyono, David, and Horng Shi Jinn, "*A study of finger vein biometric for personal identification*," In 2008 International Symposium on Biometrics and Security Technologies, pp. 1-8. IEEE, 2008.

[4] Hong, Hyung Gil, Min Beom Lee, and Kang Ryoung Park, "*Convolutional neural network-based finger-vein recognition using NIR image sensors*," Sensors 17, no. 6 (2017): 1297.

[5] Das, Rig, EmanuelaPiciucco, Emanuele Maiorana, and Patrizio Campisi, "*Convolutional neural network for finger-vein-based biometric identification*," IEEE Transactions on Information Forensics and Security 14, no. 2 (2018): 360-373

[6] Shin, Kwang Yong, Young Ho Park, DAT Tien Nguyen, and Kang Ryoung Park, "*Finger-vein image enhancement using a fuzzy- based fusion method with gabor and retinex filtering*," Sensors 14, no. 2 (2014): 3095-3129.

[7] Baloch, Noroz Khan, Zuhaibuddin Bhutto, Abdul Sattar Chan, Mudasar Latif Memon, Kashif Saleem, Murtaza Hussain Shaikh, and Saleem Ahmed, "*Finger-vein Image Dual Contrast Enhancement and Edge Detection*," International Journal of Computer Science and Network Security 19, no. 11 (2019): 184-192.

[8] X. Qiu, W. Kang, S. Tian, W. Jia, and Z. Huang, "*Finger Vein Presentation Attack Detection Using Total Variation Decomposition*," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 2, pp. 465-477, Feb. 2018.
[9] B. T. Ton and R. N. J. Veldhuis, "*A high-quality finger vascular pattern dataset collected using a custom-designed capturing device*," International Conference on Biometrics (ICB), pp. 1-5, Madrid, 2013.
[10] Y. Lu, S. J. Xie, S. Yoon, Z. Wang and D. S. Park, "*An available database for the research of finger vein recognition*," 6th International Congress on Image and Signal Processing (CISP), pp. 410-415, Hangzhou, 2013.
[11] W. Yang, X. Huang, F. Zhou, Q. Liao, "*Comparative competitive coding for personal identification by using finger vein and finger dorsal texture fusion*", Inf. Sci., vol. 268, pp. 20-32, Jun. 2014.
[12] Lee, Eui Chul, Hyunwoo Jung, and Daeyeoul Kim. "*New finger biometric method using near infrared imaging.*" Sensors 11, no. 3 (2011): 2319-2333.
[13] Kejun, Wang, Liu Jingyu, P. Popoola Oluwatoyin, and Feng Weixing, "*Finger vein identification based on 2-D gabor filter,*" In the 2nd International Conference on Industrial Mechatronics and Automation, vol. 2, pp. 10-13. IEEE, 2010.
[14] Ezhilmaran, D., and P. Rose Bindu Joseph, "*Fuzzy inference system for finger vein biometric images*." In 2017 International Conference on Inventive Systems and Control (ICISC), pp. 1-4. IEEE, 2017.
[15] Lin, Cheng-Jian, Shiou-Yun Jeng, and Mei-Kuei Chen, "*Using 2D CNN with Taguchi parametric optimization for lung cancer recognition from CT images,*" Applied Sciences 10, no. 7 (2020): 2591.
[16] Garg, Priyanka, and Trisha Jain, "*A comparative study on histogram equalization and cumulative histogram equalization*," International Journal of New Technology and Research 3, no. 9 (2017): 263242.
[17] Shin, Kwang Yong, Young Ho Park, Dat Tien Nguyen, and Kang Ryoung Park, "*Finger-vein image enhancement using a fuzzy-based fusion method with gabor and retinex filtering,*" Sensors 14, no. 2(2014): 3095-3129.
[18] Pham, Tuyen Danh, Young Ho Park, Dat Tien Nguyen, Seung Yong Kwon, and Kang Ryoung Park, "*Nonintrusive finger-vein recognition system using NIR image sensor and accuracy analyses according to various factors*," Sensors 15, no. 7 (2015): 16866-16894.
[19] Peng, Jialiang, Ning Wang, Ahmed A. Abd El-Latif, Qiong Li, and Xiamu Niu, "*Finger-vein verification using Gabor filter and SIFT feature matching,*" In 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 45-48. IEEE, 2012.
[20] Pang, Shaohua, Yilong Yin, Gongping Yang, and Yanan Li, "*Rotation invariant finger vein recognition,*" In Chinese Conference on Biometric Recognition, pp. 151-156. Springer, Berlin, Heidelberg, 2012.
[21] Wu, Jian-Da, and Chiung-Tsiung Liu, "*Finger-vein pattern identification using principal component analysis and the neural network technique,*" Expert Systems with Applications 38, no. 5 (2011): 5423-5427.
[22] Xin, Yang, Zhi Liu, Haixia Zhang, and Hong Zhang, "Finger vein verification system based on sparse representation," Applied optics 51, no. 25 (2012): 6252-6258.

[23] Syarif, Munalih Ahmad, Thian Song Ong, Andrew BJ Teoh, and Connie Tee, "*Enhanced maximum curvature descriptors for finger vein verification*," Multimedia Tools and Applications 76, no. 5 (2017): 6859-6887.
[24] Radzi, Syafeeza Ahmad, Mohamed Khalil Hani, and Rabia Bakhteri, "*Finger-vein biometric identification using convolutional neural network*," Turkish Journal of Electrical Engineering & Computer Sciences 24, no. 3 (2016): 1863-1878.
[25] Gopinath, P., and R. Shivakumar. "*Exploration of finger vein recognition systems." (2021).*
[26] *Wang, Li, Haigang Zhang, and Jingfeng Yang,* "Finger Multimodal Features Fusion and Recognition Based on CNN," In 2019 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 3183-3188. IEEE, 2019.
[27] Ganesan, Thenmozhi, Anandha Jothi Rajendran, and Palanisamy Vellaiyan, "*An Efficient Finger Vein Image Enhancement and Pattern Extraction Using CLAHE and Repeated Line Tracking Algorithm,*" In International Conference on Intelligent Computing, Information and Control Systems, pp. 690-700. Springer, Cham, 2019.
[28] Sledevic, Tomyslav, "*Adaptation of convolution and batch normalization layer for CNN implementation on FPGA*," In 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream), pp. 1-4. IEEE, 2019.
[29] Liu, Kui, Guixia Kang, Ningbo Zhang, and Beibei Hou, "*Breast cancer classification based on fully-connected layer first convolutional neural networks,*" IEEE Access 6 (2018): 23722-23732.
[30] Borra, Surya Prasada Rao, N. V. S. S. Pradeep, N. Raju, S. Vineel, and V. Karteek, "*Face recognition based on convolutional neural network,*" International Journal of Engineering and Advanced Technology 9, no. 4 (2020): 156-162.
[31] Muthukrishnan, Ranjan, and Miyilsamy Radha. "*Edge detection techniques for image segmentation.*" International Journal of Computer Science & Information Technology 3, no. 6 (2011): 259.
[32] Li, Yongxiao, and Woei Ming Lee. "*PScan 1.0: flexible software framework for polygon based multiphoton microscopy.*" In SPIE BioPhotonics Australasia, vol. 10013, p. 1001333. International Society for Optics and Photonics, 2016.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Simulink Analysis and Mathematical Modeling of Parameters Variation for Thyristor based Speed Controller of Single Phase Induction Motor

Muhammad Shahzad Bajwa[1], Muhammad Usman Keerio[2], Noor Hussain Mugheri[3], Rameez Akbar Talani[4] , Rizwan Aziz Siddiqui[5], Riaz Hussain Memon[6]
Quaid-e-Awam University of Engineering, Science and Technology Nawabshah
* Correspondence: Muhammad Shahzad Bajwa , shahzadbajwa@quest.edu.pk
Citation | Bajwa. M. S, Keerio. M. U, Mugheri. N. H, Talani. R. A, Siddique. R. A and Memo. R. H, "*Simulink Analysis and Mathematical Modeling of Parameters Variation for Thyristor based Speed Controller of Single Phase Induction Motor*". International Journal of Innovations in Science and Technology, Vol 3, Special Issue, pp: 45-58, 2021.
Received | Dec 14, 2021; Revised | Dec 17, 2021 Accepted | Dec 19, 2021; Published | Dec 31, 2021.

The thyristor is a power electronics device that is widely used in various electrical appliances due to its lower on-conduction losses, easyavailability, lower switching loss, greater efficiency and cost-benefit. Mostly a thyristor is used in rectifiers and variable-speed drives. Almost 70% of loads used in the world consists of induction motors in various types and forms. In this work, a thyristor-based controller is used to control the speed of a single phase induction motor by adjusting the firing angle for the gate terminal of the thyristor. Depending upon the firing angle, the output voltage, output current, speed, power factor and the total harmonic distortion are varied which is analyzed through MATLAB/Simulink. Further curve fitting technique is used to formulate the mathematical relationships between varying parameters concerning thyristor's firing angle. The findings of this work are helpful to achieve the best curve fit model for varying parameters concerningthe thyristor firing angle.

Keywords: AC-AC Voltage controller, Firing angle control, Capacitor start induction motor, Speed, Power factor, Total Harmonic Distortion.

Introduction

Induction motors are used globally because of their simplicity, ruggedness and low cost, etc. Its applications varies from domestic scale i.e., refrigerators, washing machines, pumps, etc to the industrial scalei.e., electric vehicles, industrial robots, elevators and so on.

To improve the efficiency and performance of induction motors, speed control of the induction motor is required purposely to satisfy the load demand rather than to run with a constant speed [1].

The construction of a single-phase induction motor can't self-start due to the absence of revolving magnetic flux. Therefore, an additional winding also known as

the auxiliary winding is attached with the main winding to create an induction motor self-start. These two windings are displaced electrically at an angle of 90° and are attached in parallel arrangements across the main AC supply. The common types of single-phase induction motors are split-phase motors, Capacitor-start, Capacitor-start run motors, shaded pole motors, etc [2].

To manage the speed of the induction motor, a variety of methods are adopted such as changing the pole, adding rheostat in the stator, voltage control, voltage/frequency control, etc. Therefore, thespeed control of the induction motor is carried out by changing either the stator parameters such as the varying number of poles, supply voltage, supply frequency, etc, or rotor parameters such as varying external resistance attached with the rotor circuit, cascaded control, etc [3].

The conventional or traditional speed control methods experience various disadvantages including frequent maintenance, mechanical wear and tear, bulk and low efficiency. Consequently, the use of power electronic devices and controllers play a very important role in enhancing the mentioned disadvantages and also facilitate soft starting [4-7].

Depending upon the methodology, a range of techniques are applied for speed control. Recently a technique is used in the voltage control method for controlling the speed by connecting back to back anti-parallel thyristor pairs called AC voltage controllers with stator part of the induction motor. Different values of the thyristor's firing angles generate different sequences of pulses that are applied to the gate of thyristors as shown in Figure1.

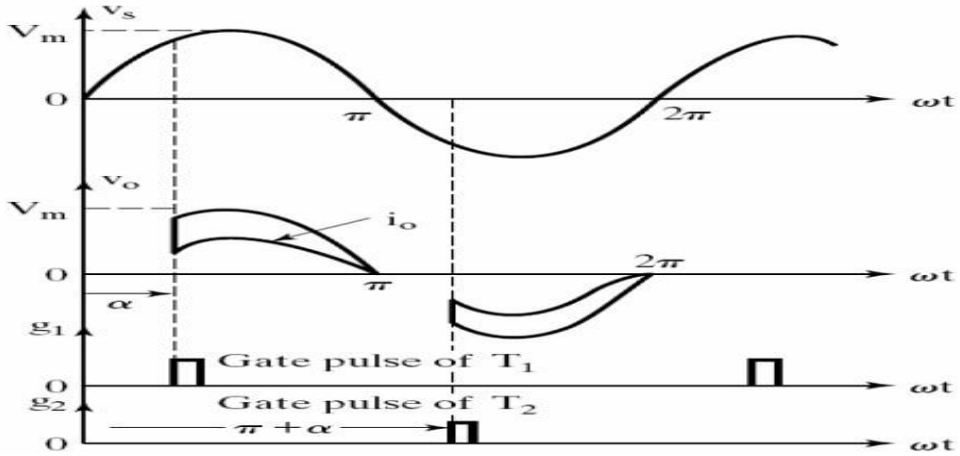


Figure 1. Input voltage, Output voltage, Thyristors pulses waveforms

By adjusting the value of the thyristor's firing angle, the output voltage changes and that's the reason behind thechange that occurs in speed since speed and voltage have a proportional relationship [8-9].

The benefits of using a thyristor-based controller include the smooth variation in the output voltage but the shape of the input sine wave is changed due to triggering of thyristor T_1 and T_2 and change of current due to inductive nature of load hence harmonics will be generated [10].

Due to switching in thyristors, harmonics will be generated. Harmonics are integral multiple frequencies containing a waveform which are added further with the

fundamental frequency. If the fundamental frequency of a waveform is f, then its integral multiples will be 2f, 3f, 4f, 5f and so on.These integral multiples are said to be the harmonic frequencies or harmonic orders.

The study of the harmonics is necessary because they cause electromagnetic interference, poor power factors, heating and temperature rise of machines, decreasing lifetime duration and reducing the performance of the machine. Thus it is very important to identify dominant harmonic orders as well as total harmonic distortion (THD) in the system and to employ appropriate methods to mitigate the harmonic distortion [11-13].

Due to thyristor-based speed controller of single phase induction motor,this work not only enumerates the variation in other parameters but also establish the mathematical relationshipamong various parameters and thyristor's firing angles.

Material and Methods

Simulation Modeling

To show the effect of thyristor-based speed controller on various parameters of induction motor, a single phase capacitor start induction motor connected with the thyristor-based controller is modeled in MATLAB/Simulink as shown in Figure 2
The parameters used in Simulink modeling as shown in Table 1.

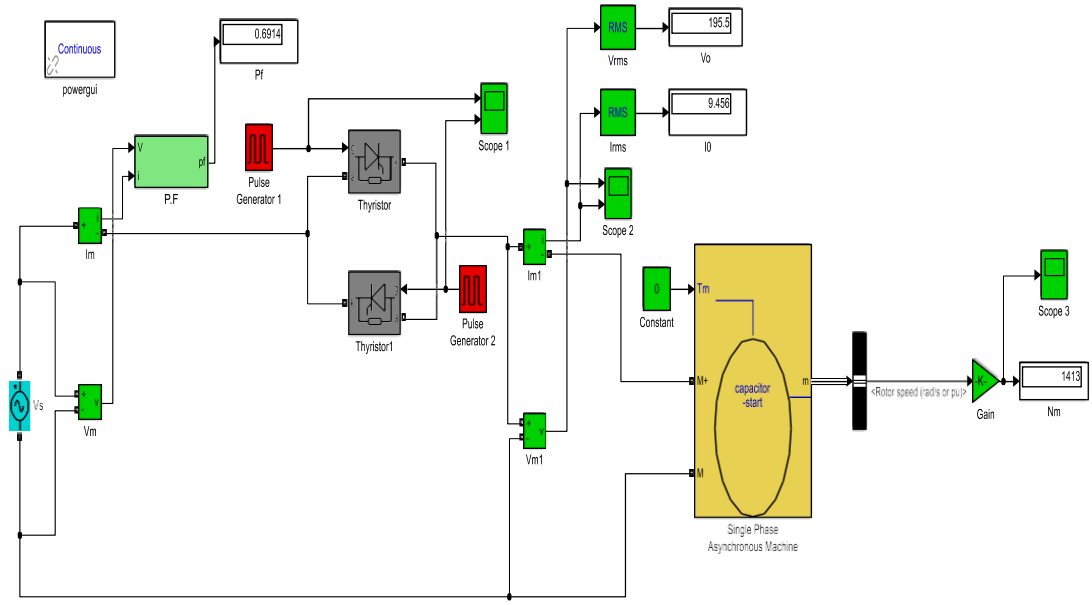


Figure 2. Simulink model of thyristor-based controller with capacitor-start induction motor

Table 1. Parameters used in simulink Model

Parametrs	Values
Input supply voltage(rms)	220V
Supply frequency	50Hz
Power Rating	0.25*746VA
Thyristor's firing angle	0-140^0

Result and discussion.

The Thyristor's firing angle manages the voltage and hence the speed. The firing angle or the gate pulse for the gate terminal of anti-parallel thyristors were provided through pulse generator block in simulink which was like a simple pulse width modulation technique. Through triggering the thyristor T1 and thyristor T2, the output rms voltage was controlled. The pulsating sequence waveforms were given for 0^0 and 70^0 thyristor's firing angles as shown in Figure 3. and Figure 4.

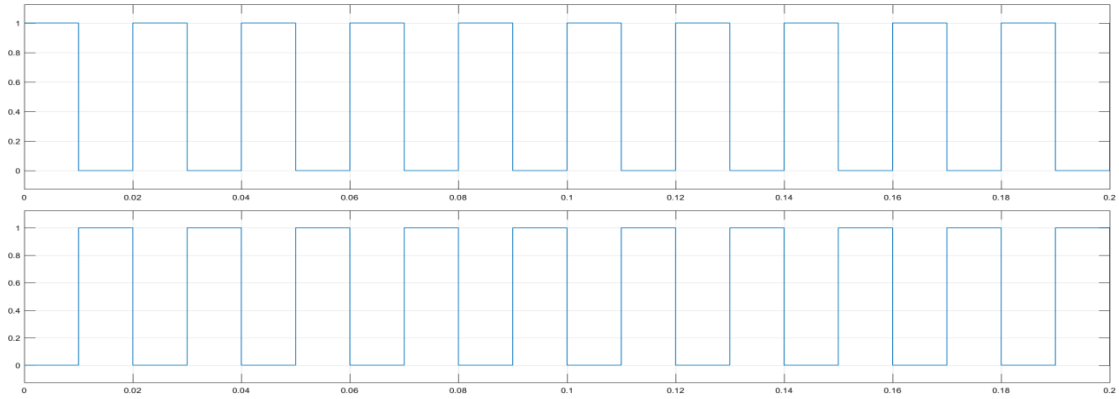


Figure 3. Gate pulses for T_1& T_2 at firing angle of 0^0

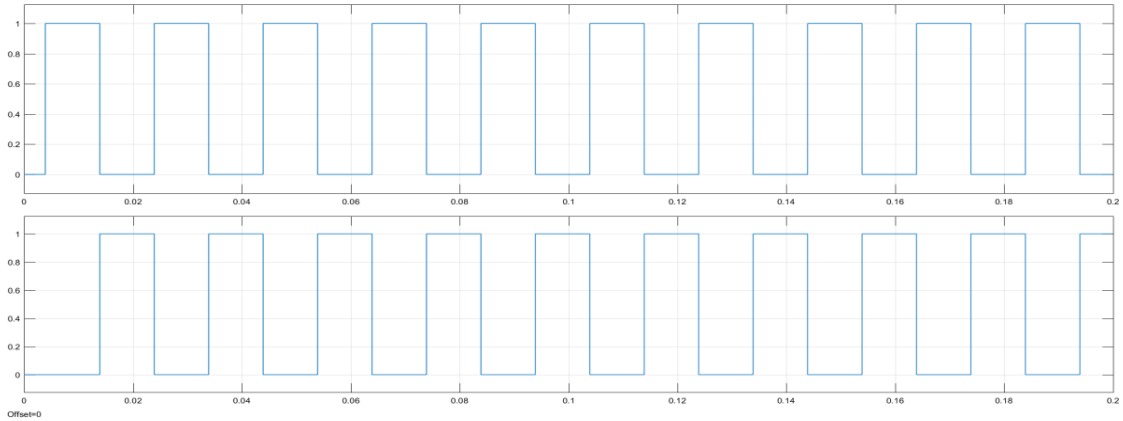


Figure 4. Gate pulses for T_1& T_2 at firing angle of 70^0

Similarly, gate pulses were generated for various thyristor's firing angles. It is indicated in Figure 3. and Figure 4. that the sequence of gate pulses was moving in the right direction with increasing thyristor's firing angles. These different pulses will chop or cut the sinusoidal supply voltages and this chopped form will be supplied to the statorcircuit. Hence by changing the thyristor's firing angles increasingly give a falling speed.The results are shown below; Figure 5. and Figure 6. show the rotor speed at thyristor's firing angle for 0^0 and 70^0.

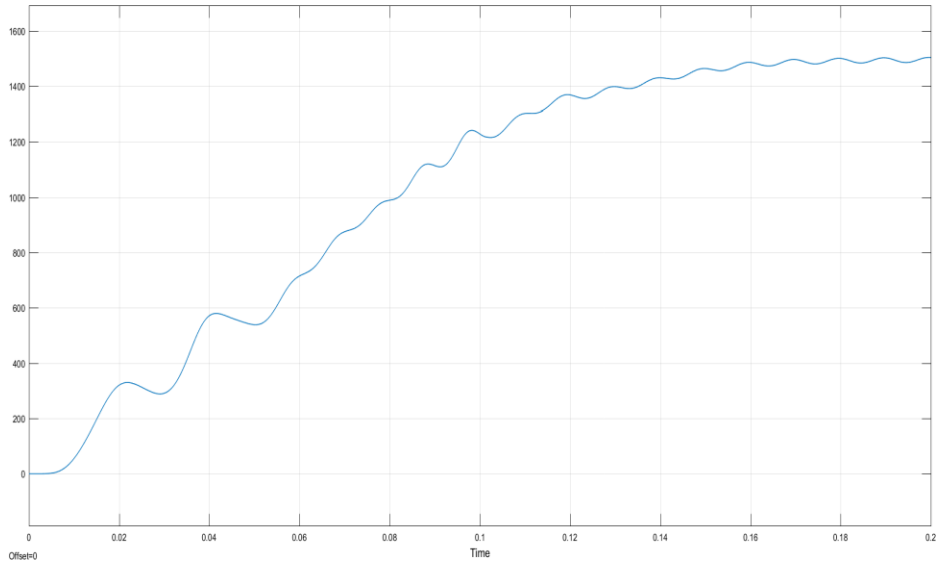


Figure 5. Rotor speed at a firing angle of 0^0

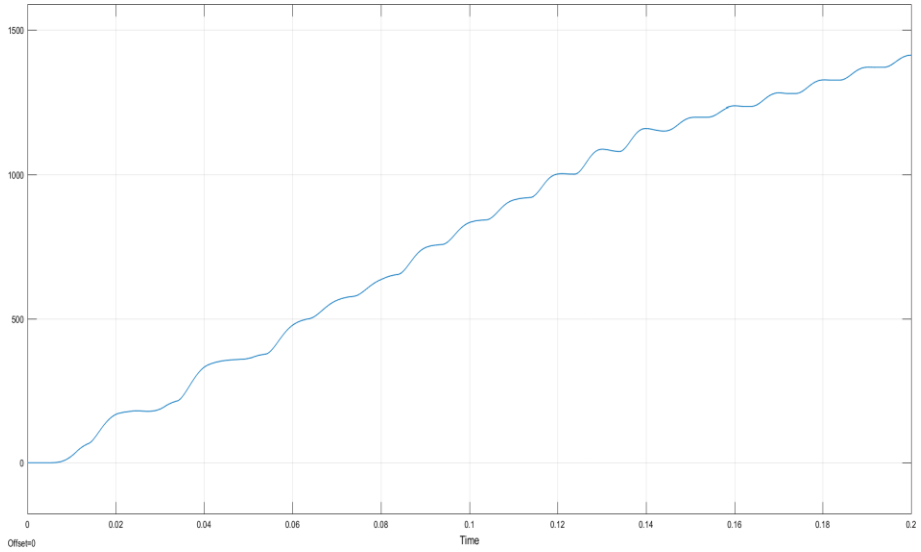


Figure 6. Rotor speed at a firing angle of 70^0

It can be seen that the speed of the induction motor comes to its steady-state level when time (t) is 0.18 seconds as shown in Figure 5.It was observed that by increasing the thyristor's firing angle,the time required to reach steady-state speed was also increased as shown in Figure 6. for the firing angle of 70^0. Figure 7. and Figure 8. show the output voltage and current waveforms at thyristor's firing angle for 0^0 and 70^0.

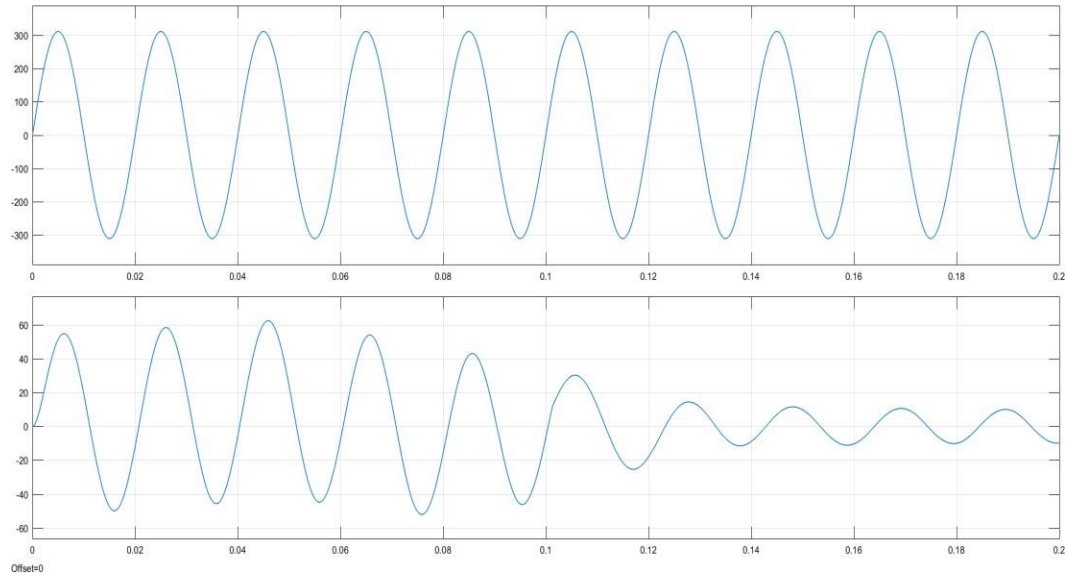


Figure 7. Output waveforms for voltage and current at a firing angle of 0^{o}

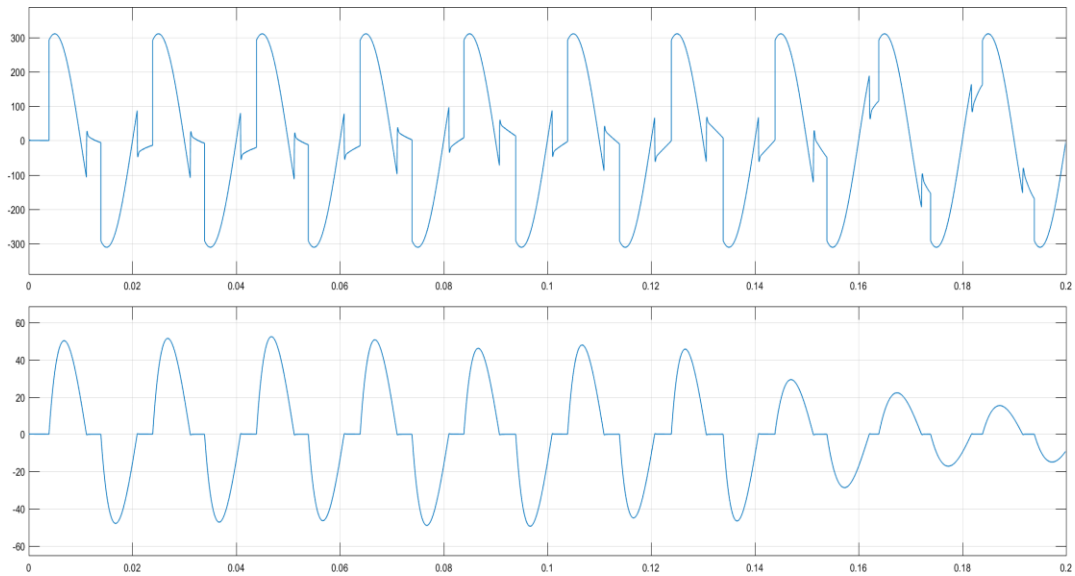


Figure 8. Output waveforms for voltage and current at a firing angle of 70^{o}

From the output current waveform, it is apparent from Figure 7 that at time t = 0.1 seconds the auxiliary winding gets disconnected and the induction motor only runs with the main winding current. Figure 8. Shows that as the thyristor's firing angle increases the waveform distortion in voltage and current waveform distortion also increases. When the thyristor's firing angle increases the %THD increases while the rms output voltage decreases.

The Fast Fourier Transform (FFT) of the output voltage and the output current are shown in Figure 9. and Figure 10. at a firing angle of 70^{o}.From the FFT analysis, the %THD in output voltage waveform was 19.45% and the %THD in output current waveform was 22.66%, while the 3rd, 5th and 7th harmonic orders were leading in nature. Hence the design of an appropriate filter was required to mitigate the leading

harmonic orders and to reduce the %THD within the international standards limit.Table.2 provides the information regarding variation in parameters such as rms output voltage, current, motor speed, power factor, %THD of output voltage and current for different thyristor's firing angles.

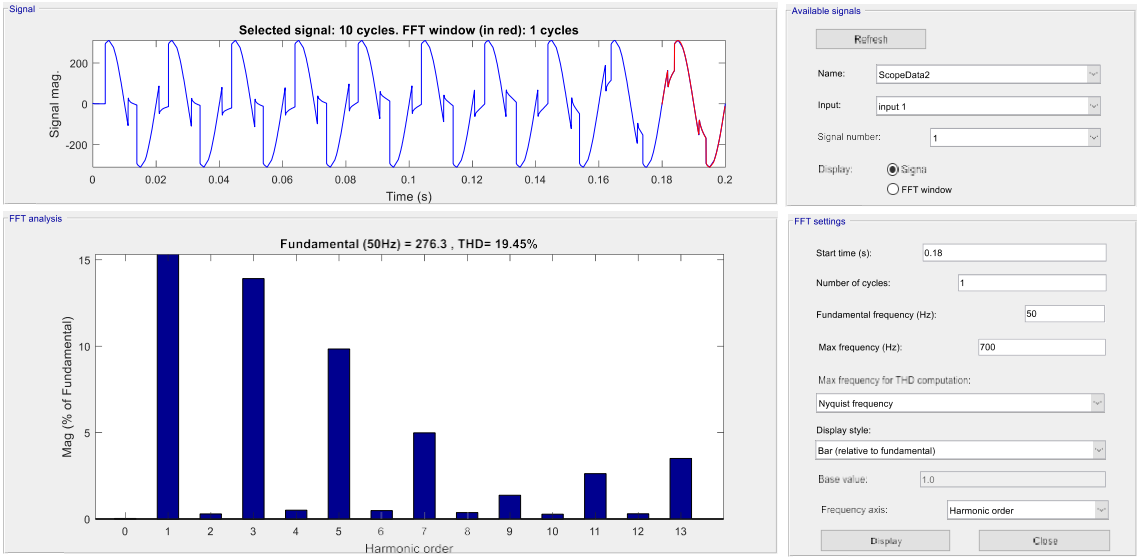


Figure 9. FFT analysis of output voltage at a firing angle of 70°

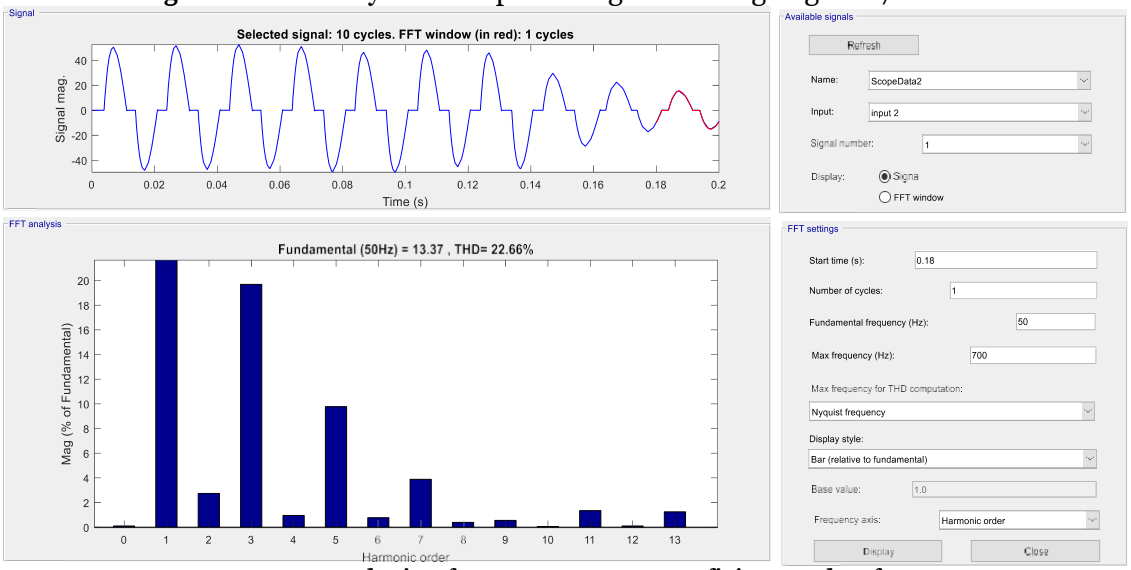


Figure 10. FFT analysis of output current at a firing angle of 70°

Table 2. Result of variations in various parameters by changing thyristor's firing angle.

S.No	Firing angle α (Degree)	Output voltage (V)	Output current (A)	Motor speed Nm (RPM)	Power factor Cosϕ	THD_v Output (%)	THD_i Output (%)
1	0^0	219.9	7.102	1503	0.15	0.28	0.32
2	20^0	219.9	7.093	1503	0.16	0.25	0.25
3	40^0	219.9	7.169	1501	0.18	0.29	0.29
4	60^0	218.4	8.245	1492	0.46	3.22	4.69
5	80^0	160	14.87	1236	0.63	47.17	47.17
6	100^0	105.4	17.79	706.5	0.58	84.92	84.92
7	120^0	63.01	10.98	246.9	0.40	122.06	72.20
8	140^0	27.36	4.79	47.11	0.22	189.95	100.01

Mathematical Modeling

It was observed from Table 2.that parameters such as output voltage, output current, speed, power factor, harmonic distortion vary by changing the thyristor's firing angle. In this section, mathematical expressions of the lines that relate to the thyristor's firing angle and different system parameters were derived using the curve fitting technique.

The goodness of the fitted line of the equation was determined by a parameter called residual or error. Residual is the difference between fitted function and the given data which is collectively known as Sum of Square due to Error (SSE). The value residual square (R-Square) should be 0.9 or 1 while the value of SSE should be zero or less than zero for the best curve fit [14].

Output Voltage Variation

The obtained output voltage value variations were graphically drawn for thyristor's firing angles from 0^0 to 140^0 as shown in Figure.11. Goodness of fit: SSE: 7.176e-021 and R-square: 1

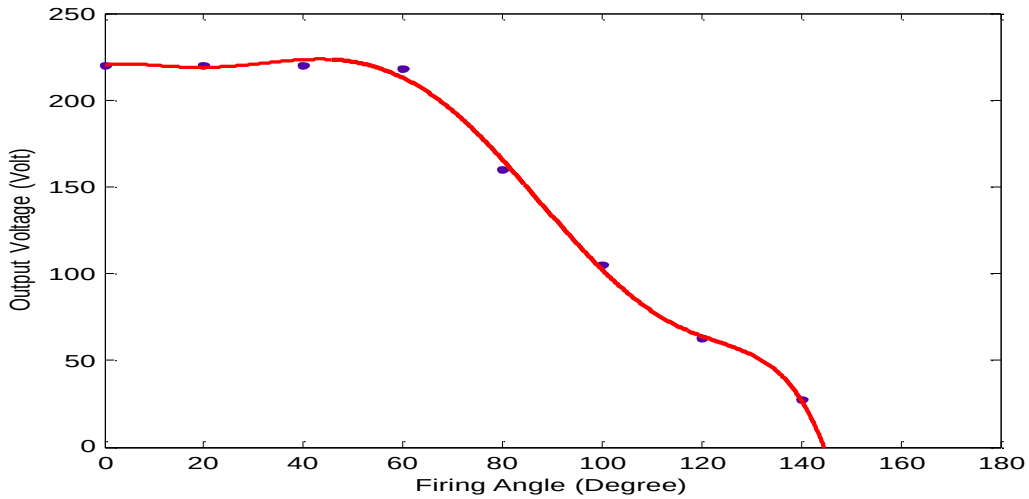


Figure 11. Firing angle Versus Output voltage

Output Voltage

$$= 0.00000000008451\ \alpha^7 - 0.00000004284\alpha^6 + 0.000008561\alpha^5 - 0.0008492\alpha^4 + 0.04303\alpha^3 - 1.036\alpha^2 + 9.063\alpha + 219.9 \quad (1)$$

Goodness of fit: SSE: 7.176e-021 and R-square: 1

The significance of this method was that the output voltage can be calculated at any other thyristor's firing angle. For example, if $\alpha = 90^0$ is taken, then by using equation 1, the predicted value of output voltage becomes 123.83v.

From the plotted curve, in Figure 11 it was observed that for 0°-60° range of firing angles, there was no drop in voltage, while voltage decreases rapidly after 70° of firingangle. After 140° firing angle, there was a very low value of output voltage for which the speed is not feasible.

Output Current Variation

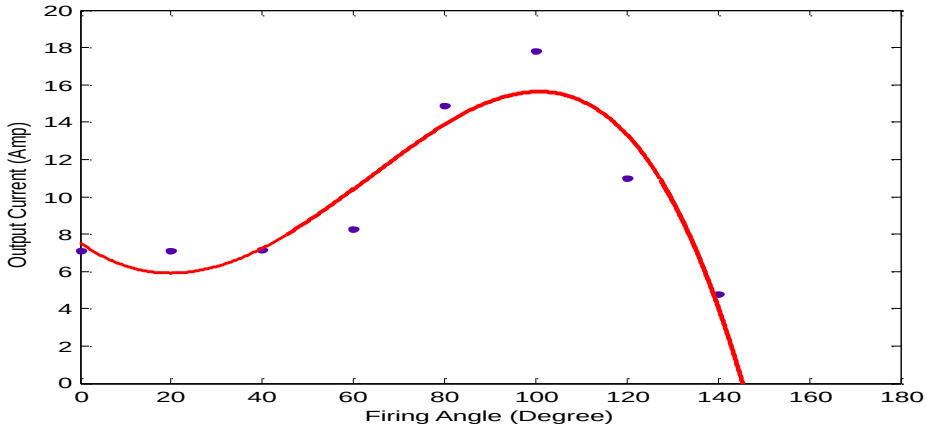


Figure 12. Firing angle Versus Output current

Output Current

$$= 0.0000000003314\alpha^6 - 0.0000001217\alpha^5 + 0.0000157\alpha^4 - 0.0008587\alpha^3 + 0.02011\alpha^2 - 0.162\alpha + 7.089 \quad (2)$$

Goodness of fit: SSE: 0.4288 and R-square: 0.9969

In Figure 12. the output current value is constant like an output voltage for 0°-60° range of firing angle, while its value was increased for 70°-100° range of firing angle. After 100° of firing angle the output current decreases.

Speed Variation

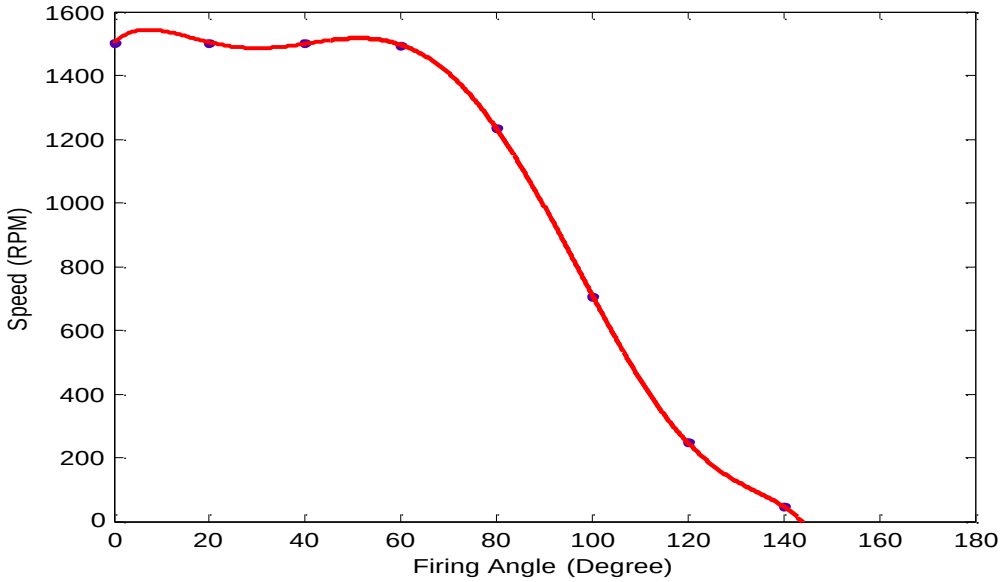


Figure 13. Firing angle Versus Speed

$$\textbf{Speed} = -0.0000000000601\ \alpha^7 + 0.0000000189\alpha^6 - 0.000001137\alpha^5 - 0.000157\alpha^4 + 0.01939\alpha^3 - 0.6714\alpha^2 + 7.054\alpha + 1503 \qquad (3)$$

Goodness of fit: SSE: 2.959e-021 , R-square: 1

The graph of speed versus thyristor's firing angle was shown in Figure 13. This graph is similar to output voltage versus thyristor's firing angle graph because the value of speed fully depends upon the controlled value of output voltage. It was observed that the speed decreases in the same manner as output voltage variations.

Power factor Variation

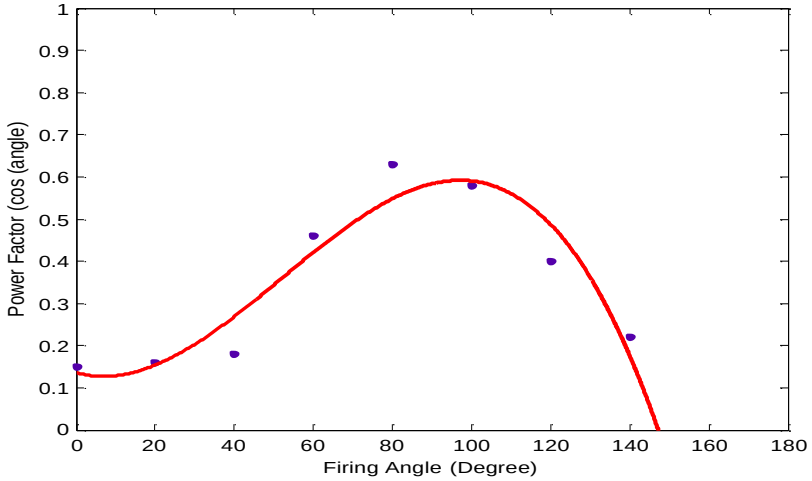


Figure 14. Firing angle versus power factor

Power Factor

$$= -0.000001269\alpha^3 + 0.0001983\alpha^2 - 0.002595\alpha + 0.1358 \quad (4)$$

Goodness of fit: SSE: 0.02594, R-square: 0.9031

In Figure 14 the relationship between the power factor and thyristor's firing angle was graphically drawn. The value of the power factor increased for 0°-90° range of firing angle. While its value was again decreased after 90° of firing angle due to the power loss.

THD_V Variation

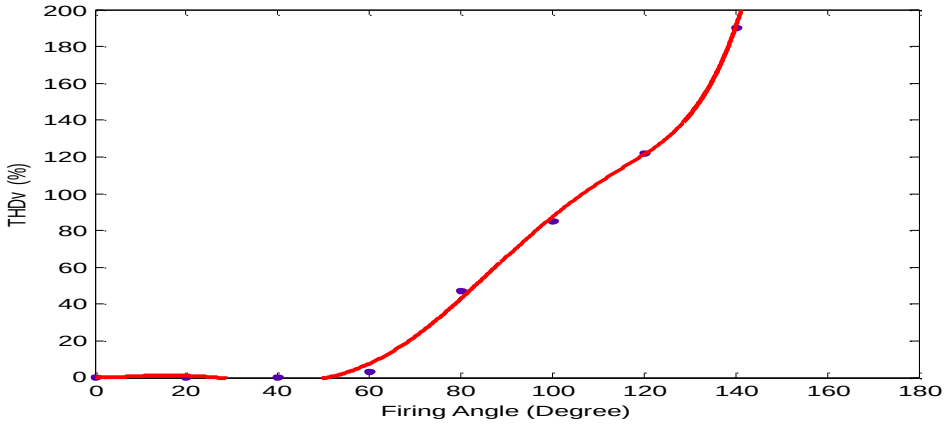


Figure 15. Firing angle Versus THD_v

$$\mathbf{THDv} = -0.00000000006573\alpha^7 + 0.00000003322\alpha^6 - 0.0000066\alpha^5 + 0.0006495\alpha^4 - 0.03263\alpha^3 + 0.78\alpha^2 - 6.79 + 0.28 \quad (5)$$

THD_i Variation

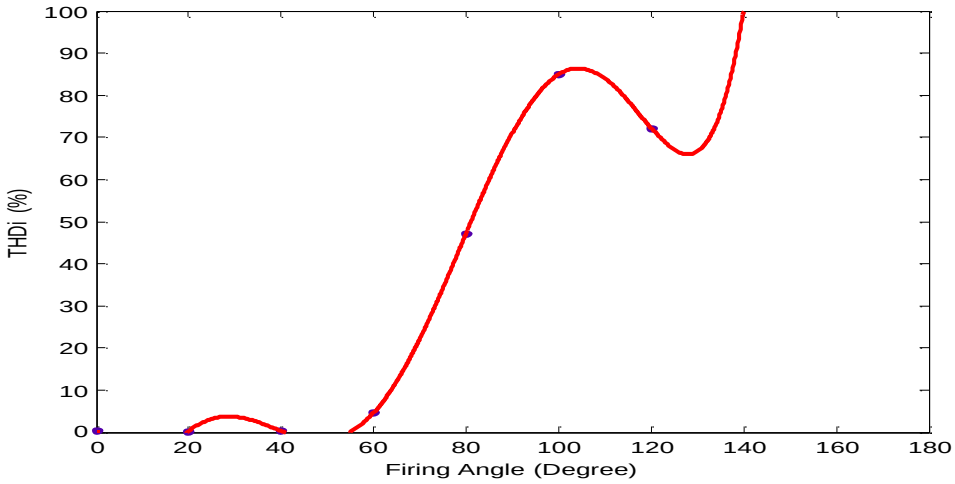


Figure 16. Firing angle Versus THD_i

$$\mathbf{THDi} = -0.0000000000176\alpha^7 + 0.00000001129\alpha^6 - 0.000002677\alpha^5 + 0.0002989\alpha^4 - 0.01629\alpha^3 + 0.4092\alpha^2 - 3.671\alpha + 0.32 \quad (6)$$

Goodness of fit: SSE: 2.323e-022 , R-square: 1

The $\%THD_v$and $\%THD_i$ variations concerning the thyristor's firing angles were graphically shown in Figure 15 and Figure 16. It became clear that both $\%THD_V$ and $\%THD_i$ were increasing with thyristor's firing angle. The mathematical equations obtained from 1 to 6 using the curve fitting techniques can be used to calculate the mentioned parameter variations at any other thyristor's firing angle.

Conclusion.

The paper analyzed the effect on system parameters due to thyristor-based speed controller of a single phase induction motor by varying the thyrister's firing angle. It is observed that the parameters such as output voltage, output current, speed, power factor, harmonic distortion varies by changing the thyristor's firing angle. Also mathematical expressions of the lines that relate to the thyristor's firing angle and different system parameters were derived using the curve fitting technique. The scope of this work is that prior to use this system, it is necessary to identify precise and accurate phenomena of different varying parameters concerning to thyristor's firing angle.

Acknowledgement. The first author would like to thank the Quaid-e-Awam University of Engineering, Science and Technology Nawabshah, Sindh for this research work and this work has not been published or submitted to other journals and all authors are agreed with the content of manuscript and this work will be published in this journal.

Author's Contribution. Conceptualization, M.S.B and M.U.K.; Methodology, M.S.B..; Investigation, M.S.B.; writing—original draft preparation, M.S.B; writing—Review and editing, M.S.B, M.U.K, N.H.M, R.A.T, R.H.M ; Supervision, M.U.K.

Conflict of interest. The Authors declare no conflict of interest for publishing this manuscript in IJIST.

Project detail. Nil

REFRENCES

1. K.Sundareswaran,"An improved energy-saving scheme for capacitor-run induction motor," IEEE Transactions on Industrial Electronics, vol. 48, no. 1, pp. 238-240, Feb. 2001.
2. M. Dey, S. Nehal, and A. Imran,"Monitoring the Variation of Speed for Single Phase Induction Motor Using Antiparallel Back to Back Connected Silicon Controlled Rectifier," In Proceeding 3rd International Conference on Mechanical Engineeering and Renewable Energy , pp. PI-279,2015.
3. S. Rahman, and A.A. Abidin, "A Review on Induction Motor Speed Control Methods", International Journal Of Core Engineering & Management vol.3, no.5, 2016.
4. M. Jannati, S. A. Anbaran, S. H. Asgari, W. Y. Goh, A. Monadi, M. J. A. Aziz, and N. R. N. Idris,"A review on Variable Speed Control techniques for efficient control of Single-Phase Induction Motors: Evolution, classification, comparison," Renewable and Sustainable Energy Reviews,Vol 75, pp.1306-1319,2017.
5. M.A. Hannan, J. A. Ali, P.J. Ker, A. Mohamed, M.S. Lipu, and A. Hussain,"Switching techniques and intelligent controllers for induction motor drive: Issues and recommendations, ". IEEE access, V.6, pp.47489-47510, 2018.
6. H.C. Chuang and C.T. Lee, "The efficiency improvement of AC induction motor with constant frequency technology", Energy, Vol.174, pp.805-813, 2019.
7. M.A. Niazi, Q. Hayat, B. Khan, and M. Afaq, "Speed Control of Three Phase Induction Motor using Variable Frequency Derive Control System". International Journal of Current Engineering and Technology, E-ISSN, pp.2277-4106, 2020.
8. R. P. Akabari, and H. buch, "Modeling of split phase induction motor with single phase cycloconverter," Journal of Information, Knowledge and Research in Electrical Engineering," Vol.02, no.2, 2013.
9. R. Gupta, R. Lamba, and S. Padhee, "Thyristor Based Speed Control Techniques of DC Motor: A Comparative Analysis", International Journal of Scientific and Research Publications, Vol. 2, Issue 6, June 2012.
10. M.A Mahar, M.A. Uqaili, and A.S Larik," Harmonic analysis of AC-DC topologies and their impacts on power system," Mehran University of Engineering and technology, Vol.30, Jan, 2011.
11. A. Altintas,"A Comparative Study on ac Voltage Controllers in Terms of Harmonic Effectiveness", Erciyes University Journal of Science and Technology, Vol. 21, No.1- 2, pp. 79-86, 2005.
12. M. U. keerio, M. S. Bajwa , A. S, Saand, M. A. Memon, "Harmonic measurement in computer laboratory and design of passive harmonic filter using MATLAB ," International Journal of Advanced Computer Science and Applications, Vol. 8, No.12, pp.1-9, 2017
13. M. S. Bajwa, M. U. keerio, N.H.Mugheri, M.A. Koondhar, I.A. Channa, "An Experimental Analysis of Harmonic Distortion Indices In Home Based Electrical Appliances and Improvement Using Passive Filter " Quaid-E-Awam University

Research Journal of Engineering, Science & Technology, Nawabshah , Vol.19,

14. M. S. Bajwa, A. P. Memon, J. A. Ansari, and M. T. Bhatti,"An Experimental Investigation Based on Mathematical And Software Modeling of Total Harmonic Distortion in Personal Computer " , Bahria university journel of information and communication Technologies Vol.9 ,Issue 1 , pp.62-73, 2016.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Evaluating Artificial Intelligence and Statistical Methods for Electric Load Forecasting

Usman Dialwar[1], Abdul Khaliq[2], Nadeem Kureshi[3]
[1,2,3] Department of Electrical & Computer Engineering, Sir Syed CASE Institute of Technology, Islamabad, Pakistan
* Correspondence: Usman Dialwar , mud.sscaseit@gmail.com.
Citation | Dilawer. U, Khaliq. A and Kureshi. N, "Evaluating Artificial Intelligence and Statistical Methods for Electric Load Forecasting". International Journal of Innovations in Science and Technology, Vol 3, Special Issue, pp: 59-83, 2021.
Received | Dec 14, 2021; Revised | Dec 19, 2021 Accepted | Dec 27, 2021; Published | Jan 1, 2022.

Electric Load Forecasting (ELF) is one of the challenges being faced by the Power System industry. With the ever-growing consumer demand, power generating companies struggle to manage and provide an uninterrupted power supply to the users. Over the past few decades, the introduction of smart grids and power deregulation has changed load forecasting dynamics. Most of the current research focuses on short-term load forecasting (STLF), involving an hour to a week's time forecasting. Various techniques are being used for accurately predicting the electric load. However, gold standards are yet to be defined mainly because of the subject's variety, non-linearity, and un-predictive form. In this study critical review of 25 publications has been carried out to find the most efficient method for ELF. The novelty of this study is that comparative and scientific analyses are carried out to find the most proficient techniques for load forecasting. Also, various parameters are combined for comparison in this study after analyzing published reviews on the subject. Artificial Neural Networks (ANN) and Auto-Regressive Moving Average (ARMA) models outperform other methods basing upon statistical analysis, i.e., Mean Absolute Percentage Error (MAPE) and comparative acceptance, in the research community.
Keywords: Electric load forecasting, Power load, Modelling electricity loads, Long term/ Short term forecasting, Performance management.

Introduction

Electric Load Forecasting (ELF) has been a prime area of concern since the advent of electricity. Predicting future load helps power utility companies to plan and meet the power generation with consumer's demands. ELF is also one of the significant factors for regulatory bodies, industries, trading and insurance companies [1]. With technological advancement, the integration of smart devices in various technical fields has become a norm, and the power industry is not an exception. Additionally, due to global warming issues, inclination towards renewable energies resulted in the introduction of smart equipment in power generation and grid systems. The same has resulted in the availability of digitized data, which on the other hand, became helpful for analysis and future prediction [2]. The consumer's electricity demand

is increasing day by day, as the world has moved towards an automated version of almost everything. Traditional power generating companies face challenges to meet user demands, and their return on investments are declining. A strong change in the power sector is observed during the 1990s with deregulation and market competition [3]. On the other hand development of smart electronics devices has gained popularity in generating power more efficiently. The volatility of electricity is adamant with the fact that it has to be provided promptly. A huge amount of electricity cannot be stored; hence equating generation with the user demands is a tough task. ELF has thus emerged as a vibrant field for the scientific community. An accurate load prediction enables decision-making by the power operators. The power industry thus invested a lot in this field to compete in the market and avoid burning extra fuel or running machinery to generate abundant electricity.

ELF is generally categorized in long, medium, and short-term forecasting on a temporal basis. Though no standard categorization has been laid so far, all of them are interconnected in the broader perspective. Long Term Load Forecasting (LTLF) – spand the load prediction for three years or more. For less than three year time period, it is termed medium forecasting (MTLF). Finally, forecasting is carried out in the short term (STLF) from an hour/ half-hour to a week's time [1]. With the growing renewable power generation systems, the introduction of smart grid systems, and privatizations, short-term and very short-term forecasting have gained popularity. This study is aimed to review published literature to look for the best technique for electric load forecasting. Rest of this paper is divided into four major positions. In the first, literature review is carried out following the explanation of research methodology applied in this study. Findings with comparative analysis and results are explained in the next part. Finally, the discussion is carried out before concluding the study.

Literature Review

Calculation of load is one of the significant factors for power companies. All the operations and planning of power generation, transmission, maintenance, etc., are based on future load value. The forecasting helps in decision making as well as reducing the risk of non-availability of power. Several conventional methods of forecasting are already in practice. Over the period, various techniques have been researched to improve load forecasting.

The qualitative methods forecast are based upon the opinions and discussion with domain experts. These methods are employed when historical data is not available for a forthcoming event. Estimates are generally vague and can lead to a blackout. The Quantitative techniques involve Time Series Analysis and Econometric Analysis. A variable of interest is defined in time series such that its value is estimated relying on the relevant historical data. Baseband model, Trend model, Linear Regression models are few examples. The econometric analysis considers the drivers such as business index, weather index, etc. such that they further leads to estimate demand requirements. Recently, Artificial Intelligence (AI) has out-performed the conventional methods in the fields where non-linear and complex data is involved. The non-linear demands, transmission losses, climate factors, etc., and their relationships have made load forecasting a potential field for application of AI techniques. Artificial Neural Networks (ANN), Support Vector Machines (SVM), Genetic Algorithm, Fuzzy Logic, Self-Organizing Maps, Extreme Learning Machines are few AI techniques that can be employed in load forecasting. Various reviews and analyses done on the subject are consulted to develop a comprehensive meta-analysis approach for this study. Table 1 depicts the methods followed by previously published studies on the subject.

Table 1. Previous Reviews on ELF

Reference Review	Method
[1]	Commonly used by Expert Community
[4]	
[5]	MAPE Percentages
[6]	
[7]	RMSE Percentages
[8]	Data and Error measured (MAPE, RMSE)

The main contributions of this paper are:

- Results are generated based on a comparative and statistical analysis of different studies already published in the subject field. Since both analyses have different implications. Comparative analysis shows the acceptance of different techniques in the research community. While on the other hand, the statistical analysis compares results in mathematical form.
- Previous reviews on ELF are first analyzed to select the parameters to compare various studies on the subject further.
- A systematic review is carried out for considering the studies published in various journals. The aim is to cover the subject domain in a wholesome and diverse manner.
- The research community can benefit by realizing the theoretical and statistical performance of various methods from this study.

Research Approach

Research is an ongoing process where methods and theories are developed and supported by logic and proof. Its main objective is to combine published methods and theories of one category and compare them with that of another in a systematic manner to reach some conclusion [9]. Meta-analysis is a common field of almost all research disciplines. It consists of five basic steps involving finding relevant studies on the subject, developing consistent criteria for comparison, recording relevant information from the study as per the criteria, analyzing information to compile them in broad contours, and finally drawing conclusions basing upon these findings [10]. This study aims to review the academic literature to explore the most efficient methods being used for STLF. Critical analysis is carried out to analyze the dynamics and performance of various methods and techniques employed.

The general framework of this study comprises two phases. In the first phase, research papers and articles are searched in most databases on the internet centered upon specific keywords. In the next phase, developed methods and their results are analyzed statistically. The general framework of this research is shown in Figure 1.

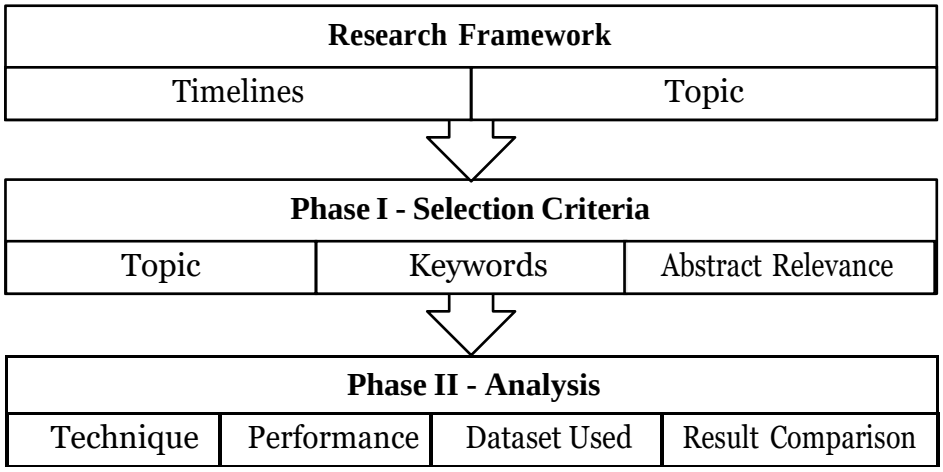
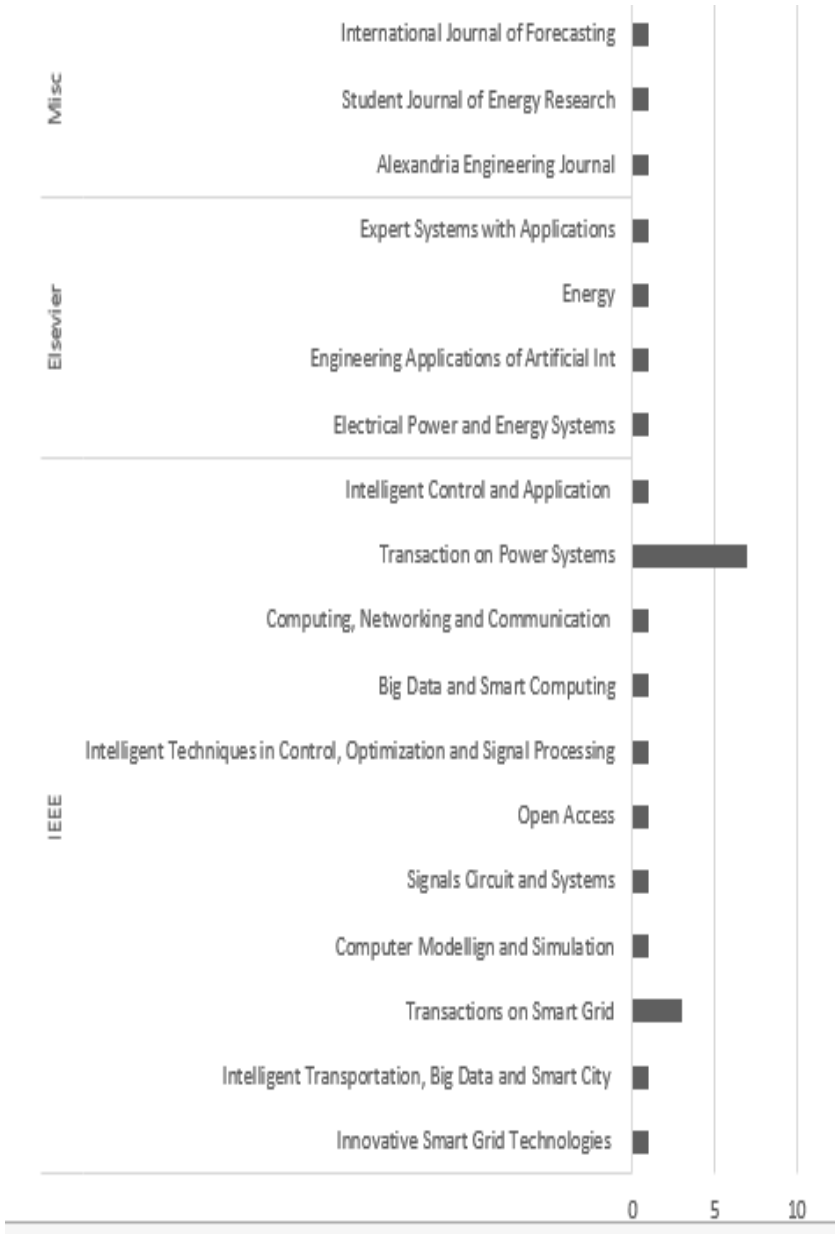


Figure 1. General Framework.

Research Catalogue

In phase I, to systematically review the literature, search is mainly carried out from 2000-2020 by using specific keywords and search engines. Scopus and IEEE Xplore are the most reliable databases in the scientific community. Both these databases are used with the keywords "electric load forecasting," "power load," "modeling electricity loads," and "long term/ Short term forecasting." Eight thousand five hundred five papers and articles came out due to search initially, including papers from areas of computing, power market, and wind energy. The advanced research tool is used to narrow down the search to a specific area of power engineering, resulting in 5,009 papers. The search is further refined based upon the title of the papers to locate 1865 papers relevant to electric load forecasting and STLF. Keeping in view the time constraint, the scope of the project, and resources available, 25 papers on STLF are selected for review and meta-analysis purposes. The topmost journals that contributed towards the selected topics are found to be IEEE Transactions on Power Systems, IEEE Transactions on Power Grids, International Journal of Forecasters, and International Journal of Electrical Power and Energy System, as shown in Graph 1.



Graph 1. Journal Wise Publications.

Analysis Approach.

Finally, in phase II, each publication is studied in detail for comparative analysis after selecting publications during the initial phase. Owing to the variability of consumer's load demands due to various meteorological conditions, socio-economic conditions, a two-pronged approach is applied in this study. Firstly, specific criteria are developed to analyze and compare the studies in detail. Since each step of the research contributes to the studies' final results, criteria are developed in such a sense that it covers complete research methodology. Secondly, the proposed methods are compared for statistical analysis as per their Mean Absolute Percentage Error (MAPE) results. This study assumes that all the results published

in the studies are correct, methods used by the majority of the expert community are best, and finally, MAPE percentages of the studies are compared.

Criteria of Analysis.

Different performance measures and results are used in different papers as per their requirements. However, meaningful meta-analysis can only be done based on some criteria. This criterion needs to be selected very thoughtfully. If it misses the relevant parameters of the respective research theme, then the chances are high that meta-analysis may not make correct assessments. Henceforth, various systematic literature reviews and studies on the subject are consulted before defining comparison criteria for this study. Criteria given in Table 2 are used to compare the papers in this study.

Table 2. Criteria for Analysis of Various Studies

Category	Description
Proposed Method	Essence of this study, as we want to check which methods used for STLF are more reliable and efficient.
Dataset Used	Number of samples or data used as input plays critical role in estimation.
Overview of the Methodology	What methodology is used by the author Long/ Medium/ Short/ Very Short Term Forecasting
Performance Measure	How results are compared with other methods and what are proposed method's strengths and weaknesses
Prediction Term	Time duration for which prediction is made.

Research Findings.

During this research, researchers found that several techniques are used by researchers while estimating load forecasts. Since no standardized model exists for the types employed, the rise of multidisciplinary collaborations in the scientific community has made the types of techniques more ambiguous to categorize. However, it is found that most of the expert community has classified techniques in two main areas; statistical and artificial intelligence-based, as shown below.

Statistical Methods

These econometrics-based mathematical models are generally based on relationships between two or more variables. The relationship is multiplicative or additive. These techniques mostly use the historical load series to forecast the future load [11].

Autoregressive (AR) and Moving Average (MA)

ARMA model is the integration of AR and MA models [12]. These are two basic models used for studying the statistical properties of a non-stationary process. Most researchers use their different combinations for forecasting purposes. In the AR model, the present value of a load series can be expressed in combination with past loads [13]. This model can predict the load value based upon past values of the load having some correlation. The equation of the AR model can be written as follows:

$$y_k - \sum_{i=1}^{p} \alpha_i y_{k-i} = e_k$$

Where α_i's are the unknown coefficients of the AR Model,e_k is the random noise and p the order of the AR model that tells us the no of past values involved in the process.

The MA model is used where load value is forecasted from past values of the input random noise values. The equation can be written as:

$$y_k = e_k + \sum_{i=1}^{q} \beta_i e_{k-1}$$

Where β_is are the unknown coefficients of the Model, e_kis the random noise and q the order of the MA model. Written in the notation of ARMA (p, q), these models combine the strengths of the AR and MA model to forecast the load value. Present values of load can be expressed in the form of past values of load and current and past value of noise, as shown in the equation below:

$$y_k - \sum_{i=1}^{p} \alpha_i y_{k-1} = e_k + \sum_{i=1}^{q} \beta_i e_{k-i}$$

[13] Developed an ARMA model by adding Gaussian noise to incorporate nonlinearity and then selecting a suitable model, with an order, to predict load. Parameters are estimated using gradient-based methods, and finally, the model is validated for its adequacy with the real data. The model performed well compared to simple ARMA and ANN. [12] developed a basic ARMA model and compared it with Projection Pursuit Regression (PPR) to be better performing.

Another variant is Auto-Regressive Integrated Moving Average (ARIMA), which considers the non-linearity involved in a time series. The AR, MA, and ARMA models are applicable for stationary processes only. However, when non-stationary data is involved, data has to be transformed to a stationary form. The equation of the ARIMA model is:

$$\alpha(B).\bigvee d \ . \ y_k = \beta(B). \ e_k$$

Where α, β are the unknown coefficients. e_k defines the noise. [14] an employed modified version of the ARIMA model by incorporating temperature and operator's knowledge into the model. The proposed model performed better than the ARMA model for predicting next year's hourly data. The ARMA and ARIMA are used successfully by [11] to forecast the load for the Kuwaiti electric network. Their approach mainly uses segmentation and decomposition of time series into similar regions and contours to make the forecast.

Kalman Filtering Algorithm

A certain level of uncertainty generally terms long-term forecasting. To cope up with this, Kalman Filters were introduced in 1960 to minimize the mean of the squared model's error. The algorithm comprises a set of equations that gives efficient recursive means to estimate the state of an observed sequence [15]. This technique has few powerful characteristics where it can control the highly noisy systems and cater to small unknown variables of the system. This algorithm can address unknown variables like weather, abrupt load demands, and customer requirements in load forecasting. The mechanism works in two stages. In the predictor stage, the algorithm predicts the load's current state based upon its previous states. Its covariance and the corrector stage information from the metering device are collected to an estimated state vector by employing the weighted average. The Kalman Filter method generally does not take into account the non-linear issues of load forecasting. Hence its modified versions are employed as done by [15]. The proposed modified Kalman Filter versions Extended Kalman Filter (EKF) and Unscented Kalman Filters (UKF) to estimate the non-linear behavior better using Jacobian Matrices.

Regression Models

Regression models are widely used statistical methods in forecasting. The main gist is learning more about the relationship between dependent and independent variables of the process. Multiple regression is based on minimizing the sum of squares of the difference between observed and predicted values. [16] used regression technique to develop a semi-parametric additive model for 24-hour demand forecast. They developed 48 models on a half-hourly basis, using selected historical load and temperature data. Forecast residuals and forecast errors are calculated using the modified bootstrap method, and finally, empirical distributions are constructed around the forecast errors for load prediction.

Non-Linear Predictors

Non-linear dynamics of the power industry are explored using non-linear chaotic dynamic and evolutionary strategy by many studies. [17] used non-linear chaotic dynamic based predictor PREDICT2 for analysis of non-linear load during training stage with emphasis on optimizing the objective function. A new Evolutionary strategy is proposed to solve the optimization problem with a candidate solution vector, having a random value with a standard deviation. [18] applied Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks to remove gradient problems in the past load data for ELF.

Exponential Smoothing (ES)

This forecasting technique works on the weighted average of the past observations. The highest weight is given to the present value of the load, then the next lower weight to the preceding value of the present value and even lower to the observation before. Due to the simplicity and accuracy, the ES technique is used quite frequently for load forecasting. The ES techniques have been divided into three further divisions. Single exponential smoothing (Brown's Method) is used when there is no pattern in the given data, Double exponential smoothing (Holt's Method) when the trend is observed in the data, and finally, Triple Exponential Smoothing (Holt-Winters Method) when data reveals significant seasonal configurations. [19] developed five ES weighted models, including a Singular Value Decomposition SVD based model to reduce the data to lower dimensions with uncorrelated variables. In [20], they proved that their proposed Seasonal Holt-Winters Exponential Smoothing method outperformed ARMA and PCA models. They used the models to forecast the seasonal demands of European data. They added an index and smoothing equation for forecasting the load. Also, ARMA and PCA models are developed to compare the performance. [21] calculated load forecast for Irish market using Double Seasonal Holt Winter's Exponential Smoothing with Error Correction. Seasonal parameters are initialized from the historical load data, and a model is proposed using the exponential smoothing algorithm. Finally, the GRG nonlinear error of predicted value and actual data is calculated.

Artificial Intelligence (AI) Models

AI systems have been developed for forecasting and estimating with the advent of advanced technology and high computational powers.

Support Vector Machines (SVM)

Presented by Vapnik in 1995, the SVM is classification and regression techniques. SVM mainly extracts the decision rules having satisfactory generalization ability from the training data called support vectors [22]. Input space is mapped nonlinearly into a higher space dimension constructing an optimal hyper plane. In the training phase of the SVM, linearly constrained quadratic programming is carried out, which is unique but time-consuming. [23] used Self Organizing Mapping (SOM) technique to organize the input data into clusters. SVMs

are then applied to each data subset to forecast the load for the next day. This hybrid method proved helpful in addressing the non-stationary load time series. [22] and [24] applied VMD is applied to decompose input data into subseries based on the certain center frequency and bandwidth. The nonlinear mapping function is used to map data in a high dimension, where the SVR function is used to relate forecast values with input.

Artificial Neural Network (ANN)

Developed in 1990 by Warren McCulloch and Walter Pitts, the ANN has been applied in several areas, including forecasting and classifications [25]. ANN is a non-linear circuit that can perform non-linear curve fitting. It processes information in line with the human biological systems. Inspired by the working of the human brain, the NN can process a certain piece of information using its basic unit called a neuron. Information received at the input node of the neuron is accumulated, processed, and then further forwarded to the next neuron through the output node. The ANN system is trained on the relevant historical data to identify the similarities and patterns of the input data. Then based upon this prior knowledge about the data and system, the network gives generalized output. In its most simplistic form, the network consists of an input layer, a hidden layer, and an output layer. The input I am sent to the hidden layer and associated weights performs a certain function f(x) to give an output. Based upon its topology, the ANN is generally categorized into Feed Forward (FF-NN) and Feedback or Recurrent NN.

Feed Forward Neural Networks (FF-NN)

Usually preferred for forecasting and consists of various combinations of input, hidden, and an output layer. In its simplistic form Single Layer Perceptron, no hidden layer exists. The forecasts are obtained using a linear combination of inputs and weight vectors, which are obtained using a learning algorithm that minimizes some cost function e-g MSE. With the addition of an intermediate layer, the NN takes Non-Linear Multi-Layer Perceptron (MLP). Neurons are arranged in layers and connected through weight vectors with the next layer. Neuron b takes the input from its predecessor neuron, if it exists, computes the weighted sum w, eliminates the bias, and gives the output after applying the activation function g. The equation is given by:

$$\sum_{j=1}^{n} \beta_j \, g(w_j \, . \, x_i + b_j) = \; y_i$$

Where, x_i is the input, w_j is the weight, b is the neuron of hidden layers.

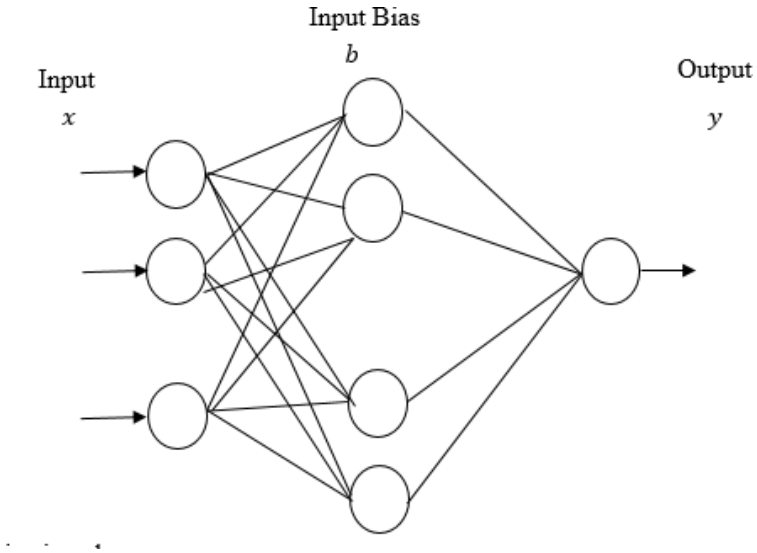


Figure 2. Basic Structure of a NN

Feedback Neural Networks

Unlike the FF-NN, the feedback NN is dynamic. Whenever a new input pattern is given the output of neurons is computed. Their output depends on the state of the system. Feedback of the neurons is modified due to the feedback system, and hence the NN enters into a new state. To overcome the vanishing gradient problem of the NN, Nonlinear Autoregressive Models with Exogenous Inputs (NARX) have been developed. This three-layer FF-NN with good learning capabilities has a sigmoid activation function in its hidden layer, linear activation function in the output layer, and delay lines for storing previously predicted values.

NN-based STLF has been enhanced using Multi-resolution analysis (MRA) by [26]. Four models are developed with different input variables among load, temperature, differenced load from the first, and MRA with the differenced load. The final models comprise sub-models of the first three models to decompose the load series using individual fitting. The proposed Model with load, temperature and first-order differenced load as input predicted the load most accurately. [25] also proposed Wavelet-based NN (WNN), using previously used algorithms for generation, selection, and generalization, to compare its prediction performance. However, they concluded that results of WNN are comparable with naïve methods and MLP NN on GEFCom dataset. [27] improved the BP NN with the introduction of GA. They used PSO to improve the convergence speed and PCA to reduce the matrix dimensionality.

[28] introduced Artificial Immune System with ANN. The aim is to check the benefits of the robust AIS like computational strengths as its distributed, diverse, anomaly detection, and self-organizing learning abilities. The performance of AIS-based FF-NN has comparable results on the MAPE scale with that of BP NN. However, further studies may reveal the true potential of AIS in the field. Different types and scales of NN have been used for the last two decades by researchers for load forecasting. Having received considerable success, the NN is also criticized for having too many input parameters, leading to data overfitting. [29] conducted a detailed review of various models of NN with traditional statistical methods. They compared large NN with linear models, including Naïve forecasting, methods with one and

more smoothing filters, smoothing filters with linear regression combination of smoothing filers, and NN. The conclusion is that large NN can perform well because they consider more historical data and can interpolate high dimensional functions, which improves the profile load forecasting. [30] also worked to investigate the non-linear characteristics of the power load series is identified using MLP. An attractor is then developed in a phase plane to train ANN. [31] proposed a set of probabilistic models as constrained quantile regression models to average and predict the future data.

While using NN [32] employed a wavelet-based ensemble scheme. Selection of mother wavelet and decomposition level is a tricky affair. Here ensemble of wavelets is used, and their output is aggregated to get the best features as output. Wavelet-based ensemble networks, algorithm incorporating Levenberg–Marquardt (LM) for improved learning, Conditional Mutual Information Feature Selection (CMIFS) method is employed for feature selection, and Partial Least Square Regression (PLSR) is used for forecasting purposes. In another scheme, [33] exploited that NN learns load dynamics without memorizing the data for a long time with accurate results. Challenges are faced while extrapolating the relationships different from those extracted from training data. Five models of three-layered FF NN are used for forecasting. Redundant hidden neurons are also eliminated by observing duplication in co-linearity with the output.

Hybrid Models

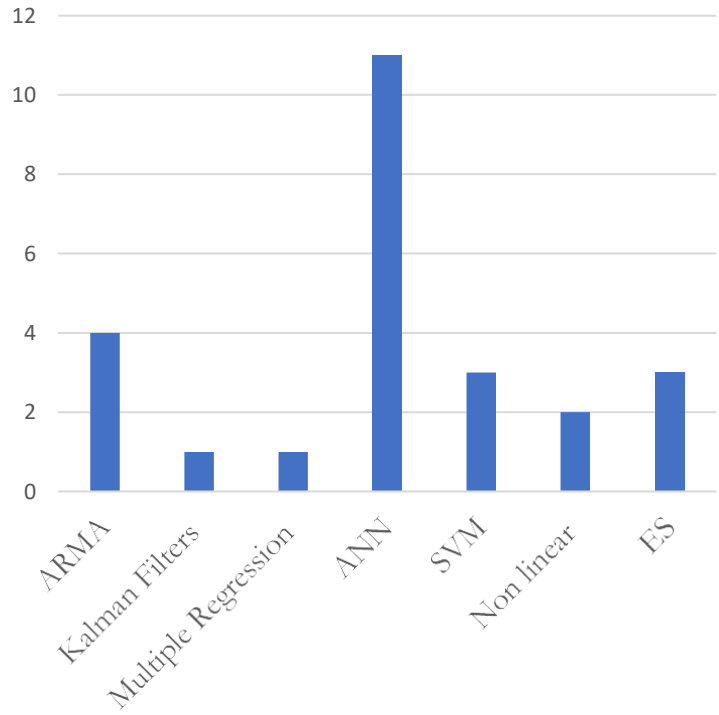
Many in the past have published applications of combining the strengths of different models into hybrid models. Also, for load forecasting, various methods have been combined to produce efficient methods. The probabilistic nature of power systems makes it a potential field for employing various methods to estimate the forecast for time.

Results

Since there is no gold standard yet for forecasting and methods for prediction, reviewers consider various assumptions while comparing the publications. This study assumes that all the results published in the studies are correct, methods used by the majority of the expert community are best. However, for statistical analysis, MAPE percentages of all studies are also compared.

Comparative Analysis

After selecting publications for this study during the initial phase, each publication is studied in detail for comparative analysis as elaborated in Table III. The dynamics of the topic and unpredictability of load influence the researchers to use different variables and performance standards to measure their proposed methods. In addition, heterogeneity of data due to socio-economic conditions and consumer's profile, non-linear environmental conditions, including weather, humidity for different countries, add complexity to the comparison of studies. At times a simple or a particular method favors a particular situation, even the sophisticated techniques. This study focuses on the comparison benchmarks mentioned in Table II, including Proposed Method, Data set used, Overview, and Prediction Term. ANN stands out to be the most used method for ELF after going through all the publications during this study, as shown in Graph II.



Graph 2. No of Publications Studied as per Method

Eleven out of 25 publications used ANN in one form or the other. ARMA-based models are the next most frequently used method for ELF by the researchers.

Table 3. Comparison of Studies

Ref	Technique	Dataset/ Training & Testing	Overview
[12]	Hybrid using Auto-regression Integrated moving average (ARIMA) & projection pursuit regression (PPR)	5 min interval time series of Sichuan Electric Power Company, China with 864 observations for randomly selected data from 26-28 Sep 2016. 576 for training and 288 for testing out of the total of 864	• ARIMA is modeled, estimating its order and parameters using Bayesian Information Criteria (BIC) and correlation function. • Next, the PPR model is developed. Integration of both models is carried out to address the linear and non-linear dynamics of the load forecast.
[11]	Hybrid model using Autoregressive Moving Average ARMA and ARIMA.	Daily load data of Kuwaiti Electric network from 2006 – 2008.	• Data is segmented to locate the identical patterns and calculate their probability. Afterward, the series is decomposed using MA for load pattern segmentation. • For load forecasting, curve fitting is conducted to identify region

			similarity, contours, and related points.
[13]	ARMA including Gaussian and Non-Gaussian Processes	Hourly data of 3 months between 1998-1999 of Taipower Company, Taiwan	• Historical data is processed for the Gaussian test using the Bi-spectrum process. • If data is Gaussian, second-order statistics are calculated; otherwise, the MA model is applied. • The correct model identified in the previous step is used to estimate the parameter representation of the model.
[14]	ARIMA Model integrated with operators knowledge	Hourly load and peak load data from Iran's national grid from 1996-1998. Data from 1996-1997 is used for training, while that of 1998 for testing purposes.	• ARIMA model is proposed considering the historical data, estimating the parameters. • 16 Modified ARIMA models are used for forecasting along with the temperature and operators knowledge
[18]	Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks	Three-year record data of various feeders from West Canada with 1997 records 1,597 records for training and 400 out of total 1997 (80%/20% split ratio)	• Features from past data are collected based on socio-economic and weather conditions. • Principal component analysis (PCA) and Normalization of selected features are performed. • LSTM and GRU networks for Many to Many and One to Many configurations are developed. • These networks are better to vanish and explode gradient problems in the data.
[17]	A hybrid approach based on non-linear chaotic dynamic predictor	Hourly electricity load of the year 2002 from New England, Albert, and Spain. Random selection of 4 weeks, one each in 4 months of a year, for testing. Rest of the data used for training	• The time-series data, set as input to the model, is divided into two segments. One segment is used to predict data in the second segment. • The population is initialized using candidate random variables. Next, the population parameters are recombined to produce off-springs and then mutate.

			• This ES is used to tune the prediction parameters.
[16]	Semiparametric additive models using Modified Bootstrap method	Half hourly demand and temperature data of Melbourne from 1997-2009 from Australian National Electricity Market. Data from 2004-2008 was used for training, and 2009 data for testing.	• A Semiparametric model is developed to forecast demand and temperature values using their historical data. Cross-validation is done to select the variables for use in models. • Forecast residuals are calculated by sequentially substituting into random forecasted model values. • The modified bootstrap method is used to obtain forecast errors.
[15]	Modified Non-linear Kalman Filter, Extended Kalman Filter (EKF), and Unscented Kalman Filter (UKF). Weather and Wind Speed data is accumulated from the website.	Reference Energy Disaggregation Dataset (REDD) anonymously collected from Boston, US	• Standard KF is modeled using past data, temperature, and wind speed data. • The model predicts the value based on past data along with its covariance. • The output is recursively updated using the law of minimizing mean square error. • The non-linear Modified filters, EKF and UKF, are applied to calculate the prediction.
[12]	ES with Holt-Winters, ARIMA, and PCA	30 Weeks hourly/ half hourly data of 6/ 4 European countries from Apr-Oct 2005. The first 20 weeks of each data is used for training and the last ten weeks for testing.	• Seasonal Holt-Winters Exponential Smoothing is applied to forecast two seasonal demands. • Additional seasonal index and extra smoothing equations are added for the double seasonal method. • The initial level and seasonal values are estimated by averaging the observations and minimizing the squared sum of errors.

			• ARMA and PCA models are also developed to compare the performance
[21]	Double Seasonal Holt-Winters Exponential Smoothing with error correction	Half-hourly data of 15 months from an Irish supply company from Jan 2013- March 2014	The seasonal parameters defined in Days, Weeks, and Seasons are initialized from the historical load data. The model is proposed using the exponential smoothing algorithm. Finally, the GRG nonlinear error of predicted value and actual data is calculated.
[19]	Five exponentially weighted methods incl new Singular Value Decomposition SVD based ES	Half hourly observation from 2007-2009, first two years used for training and last year for testing	• SVD based approach is used to reduce the data to lower dimensions with uncorrelated variables. • Modified Holt Winter ES (HWT) • Discounted weight regression (DWR)
[23]	A hybrid approach to combine SOM with SVM	Hourly data of one year from 2003-2004 of New York City, US	• In the first stage, SOM is used to group the training data with similar properties. • The SVM network of 24 machines is then applied with regression and risk minimization principles to forecast the next day's load.
[24]	Simulated Annealing with SVM	Taiwanese load data from 1045-2003, 40 years training set from 1945-1984, 10 yrs validation 1985-1994, 9 yrs testing 1995-2003	• The past data is normalized using the simulated annealing algorithms. • Then SVMs are applied for load forecasting. • The proposed model is compared with ARIMA and Regression NN.
[22]	Hybrid model using Variational Mode Decomposition Self Recurrent	Half-hour load data from National Electricity Market, Queensland, Australia, and hourly load data from New York Independent System,	• VMD is applied to decompose input data into subseries based on certain center frequencies and bandwidth. • The nonlinear mapping function is used to map data in a high

	Support Vector Regression Cuckoo Bird Cuckoo Search (VMD-SR-SVRCBCS)	USA. Both datasets were distributed into 3 x parts for the training, validation, and testing phase	dimension, where the SVR function is used to relate forecast values with input.
[30]	ANN-based on Multilayer Perceptron	Daily peak demand of one year for 1995, 9 months data is used for training and two months for testing	• The time series is extended to confirm its chaotic character using correlation dimension and Lyapunov Spectrum. • The state-space of a differential equation is created for the time series taking into account all its variables. • Then model based on correlation dimension and state space of the data is developed.
[27]	Neural Network optimized using Particle Swarm Optimization (PSO) and Principal component analysis (PCA)	1-year data of a Power Grid Corporation Previous one-year data for training	• PSO is used to initiate the model from initial weights and thresholds. • PCA is used to reduce the input dimension as per the set threshold with GA optimization. • Load is forecasted for the next 24 hours.
[28]	Feed Forward-Neural Network (FF-NN) trained by the Artificial Immune System (AIS)	Day, time, temperature, and 720 samples each from historical load data of Kuala Lumpur, Malaysia, and the other from North Carolina, US 65% of the data is used for training and 35% for testing	• The AIS-based algorithm is developed with initial weights selected randomly between 0 and 1. • FF-NN on MLP architecture is proposed where input parameters are multiplied with weights. • Regression is performed to correlate the predicted values with the past load series.
[34]	Convolutional Neural Network (CNN) with K	1.4 million records of electricity data from 2012-2014 containing hourly load data from the power	• Raw data is pre-processed, converted into two subsets, training and testing, based upon

	Means clustering is employed.	industry. 1,003,716 samples from 2012 - 2013 are used for training and 469300 samples for testing.	selected feature analysis using K Means Clustering. • CNN is trained on one subset and then validated on the testing subset.
[35]	Feed Forward Deep Neural Network (FF-DNN) and Recurrent Deep Neural Network (R-DNN)	Hourly data of NEW England, the USA from 2007-2012 comprising 52600 records. 43824 samples are used for training, while the rest are used for the testing phase.	• The data is analyzed in the time and frequency domain to model it comprehensively. • In the next stage, Rectifier Activation Function (ReLU) is used to model FF-DNN and R-DNN. • Separate results are computed considering only Time Domain and Time & Frequency Domain features.
[36]	Modified Deep Residual Network adopting ensemble strategy	North American utility data set with hourly data from 1985-1992. 2-year data from 1991-1992 is used as test data, rest of the data is used for training. To check the generalization, ISO-NE data is used.	• A two-level basic structure is formed for forecasting 24 hours data with the Scaled Exponential Linear Units (SELU) activation function. • Output is fed into Deep Residual Network (ResNet) constructed from a stack of three residual blocks. • Modifications are made, ResNetPlus, by employing several residual side blocks and averaging the output of each main residual block with these side blocks to improve error backpropagation of the network. • The next ensemble strategy is used to improve the generalization capability of the network.
[26]	NN with Wavelet decomposition	Hourly load data of North America from 1988-1992	• Features are extracted into Low and High-Frequency components using Multi-Resolution Analysis. Input variables are selected by applying correlation functions. • Four models of NN have been developed based upon MLP.

			• The model with inputs of load, temperature, and first-order differenced performed the best among other NN.
[37]	Wavelet Neural Networks	Hourly load data of NEW England from 2003-2005 as training and year 2006 data used for testing.	• Wavelets are used to decompose the load into Low and High-Frequency components • MLP based NN is then applied for load forecasting.
[29]	Large NN and regression methods	Hourly data from 1996-1997 of a city of Brazil. Data is split for the training, testing, and validating phase.	• Various models are developed based upon Naïve forecasting, methods with one and more smoothing filters, smoothing filters with linear regression combination of smoothing filers and NN and large NN
[33]	NN	Load data of the previous 1 hour is used to predict the next 20 minutes load for a power company in the US.	• They used relative load curves of past data instead of load increments to improve the forecasting accuracy as is done in traditional NN models. • Input variables are selected based upon their string statistical correlation with outputs. • Supervised training is carried out for the proposed NN using the previous load data and minimizing the error function.
[38]	Hybrid NN based on Wavelets	Hourly load data from ISO England for the year 2009-2010.	• ELM-LM algorithm is developed by randomly initializing the weights and biases to estimate the output weights. • Wavelet transform is used to employ frequency components along with temporal dimensions of the past load series. • PLSR is used to combine the forecasts of different wavelets. • Hourly load data is fed into 24 FF-NN with the detailed extraction of frequency components using wavelet transforms.

Statistical Analysis

For performance measurement, researchers have used various forecasting standards like including Mean Absolute Percentage Error (MAPE), Root Mean Square Error, etc. However, MAPE is used more frequently in statistical studies. The difference value is calculated by taking the absolute difference between the proposed method and other methods with which it is compared. The MAPE difference of the proposed and other methods are calculated. Then the smallest value is located to identify the best method as depicted in Table IV. It is revealed that ARMA models and ANN gave the MAPE values. A stable difference criteria has also been defined by setting the value of alpha from 0.01 to 5. This means that difference values less than 0.01 and greater than 5 are ignored in this study. The mean MAPE for ANN and ARMA Models is 0.799 and 0.8446, respectively.

Table 4. Comparison of MAPE and Standard Deviation

Cat	Ref	Proposed Method	Benchmark Method	% MAPE	Mean-MAPE
ARMA / ARIMA	[12]	Hybrid using ARIMA & Projection Pursuit Regression (PPR)	ARIMA PPR	0.634 0.403	**0.8446**
	[11]	Hybrid model using ARMA and ARIMA.	Real data	0.5	
	[13]	ARMA including Gaussian and Non Gaussian Processes	ARMA ANN	0.05 0.58	
	[14]	ARIMA Model integrated with operators knowledge	ARIMA ANN Operators	1.24 1.27 2.08	
Non-Linear Regression	[18]	Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks	FNN Modified FNN	5.1 2.79	5.05
	[17]	A hybrid approach based on non-linear chaotic dynamic predictor	ANN ARIMA	5 4.5	
	[16]	Semi parametric additive models using Modified Bootstrap method - Regression	ANN Hybrid	0.85 0.4	0.85
Kalman	[15]	Modified Non-linear Kalman Filter - Kalman	EKF	0	0
Exponential Smoothing (ES)	[20]	ES with Holt Winters.	ARMA PCA AR	0.059 0.05 0.086	0.69
	[21]	Double Seasonal Holt-Winters Exponential Smoothing with error correction	Naïve w/o EC	1.66 5.3	

	[19]	Five exponentially weighted methods incl new Singular Value Decomposition SVD based ES	ANN HWT with SM NEW SVD Weather-based	0.02 0.01 0.01 0.016	
Vector Machines (SVM)	[23]	Hybrid approach to combine SOM with SVM	ISO SVM	1.15 0.65	2.42
	[24]	Simulated Annealing with SVM	ARIMA GRNN	8.55 3.42	
	[22]	Hybrid model using VMD-SR-SVRCBCS	ARIMA GRNN BPNN SVR	7 5.3 5.1 3.8	
	[30]	ANN based on Multilayer Perceptron	Others	0.4%	**0.799**
	[27]	NN optimized using PSO and PC Analysis	No PC Reduction	1.1	
	[28]	FF-NNtrained by theArtificial Immune System (AIS)	Data 1 - AIS Data 2 – AIS	0.473 1.347	
Neural Networks (NN)	[34]	CNN with K Means clustering is employed.	LR SVR SVR &K Means NN NN & K-Means	25 8.97 0.89 0.163 0.115	
	[35]	FF-DNN and Recurrent Deep Neural Network (R-DNN)	Time Frequency	12 0.01	
	[36]	Modified Deep Residual Network adopting Ensemble Strategy	Temperature -1 Temperature -2 Temperature -3	0.02 0.05 0.11	
	[26]	4 models based on NN with Wavelet decomposition with different inputs	Model-1 Model-2 Model-3	0.17 0.42 0.94	

[37]	Wavelet NN	NN w/o weather NN & weather Similar day	0.188 0.07 0.22
[29]	Large NN and regression methods	Smoothing with small NN Large NN	0.1 1
[33]	NN	Forecaster-1 Forecaster-2 Forecaster-3	0.59 0.39 0.23
[38]	Hybrid NN based on Wavelets	Abductive MLR RBFNN Random forest	0.79 0.83 0.56 0.92

Discussion and Future Research.

Load forecasting has become a topic of significance in the past few decades. Researchers have used various techniques to identify the best-performing methods. However, the non-linear dynamics of the topic imply that no one method can be classified as the best. Availability of historical load data is the prime factor in forecasting. However, heterogeneity in this data itself challenges the analysis. The data is dispersed in different patterns with different power companies. It is calculated on an hourly basis, whereas at the other places, it is recorded on a seasonal basis.

Most of the statistical methods employ past load series and weather information for prediction. These past load data are used as input to Regression techniques and the weather and its functional relationship. The same is then solved regressively to reduce the square error of the prediction. Exponential smoothing models are developed by a linear combination of time series and other variables. Kalman filtering use filtering techniques to reduce the noise in data to predict future load. When combined with Wavelet decomposition forecasting is improved further as it employ frequency component of data series as well. The ANN techniques have performed quite well for ELF. However, their main concern is data fitment. The NN employs layers of neurons and a large number of parameters that raise the concern over parameterization in performing the task. Large NN performs better in forecasting results, but the theory behind this remains a black box.

Meteorological conditions also risk load forecasting. Although in today's digital world, previous data and future weather forecasts are also available. Still, the unpredictability of the weather, humidity conditions plays a significant role in load forecasting. Then the socio-economic conditions of the consumers dictate the variability of load demands. One cannot consider the functions, gatherings, or other related activities at a specific place. Another important concern is about the transmission network dynamics. Equipment failures and

accidents make the power unavailable in one region, thus causing demand at another generating region.

Technological advancements, especially in the form of renewable energies, have modified the dynamics of power sectors. The load forecasting will be an area of concern to fulfill consumer's power requirements. Based on this study, the following areas are elaborated for future research:

- Implement the techniques found during this study on the real-world load data to verify their performance. Challenges found in this study, like availability of data, weather constraints, diverse consumer power demand, etc., will be considered.
- Increasing use of electric appliances and wide adoption of electrical transportation systems significantly impact electricity requirements. Load forecasting in this regard will enable power utilities to meet user's load requirements.
- Load forecasting is evolving day by day with the latest technological developments. The growing acceptance of renewable power generation systems, especially solar systems, makes users' load demand unpredictable. Research in renewable power generation and forecasting is also an area of interest for the future.
- Study the feasibility of integrating renewable power generation systems into the main power grid.

Conclusion.

Meta-Analysis is carried out by studying 25 publications on ELF modeling proposed by researchers and compared with various other forecasting methods. The criterion for comparison is selected, including technique employed, data set used, overall methodology, performance, and MAPE measurement. The comparative results show that various non-linear factors play a significant role in ELF, importantly weather conditions. Few methods are preferred because of their fast computation power and linear relationship among variables. ANN and ARMA are found to be the best performing methods. ANN is mostly used when changes occur at a faster pace like frequent changes in weather or environmental conditions. However, with larger NN, the issues of data over fitment need to be taken into consideration. The ARMA models are attractive due to ease in their practical interpretation. They are usually criticized for their limitation to deal with non-linearity behavior of processes.

Author's Contribution. All the authors contributed equally.
Conflict of interest. We declare no conflict of interest for publishing this manuscript in IJIST.
Project details. Nil

REFERENCES

1. Hammad, M. A., Jereb, B., Rosi, B., & Dragan, D. Methods and Models for Electric Load Forecasting: *A Comprehensive Review. Logistics & Sustainable Transport*, 11(1), 51-76, 2020.
2. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4), 1030-1081, 2014.
3. Leung, T. C., Ping, K. P., & Tsui, K. K. What can deregulators deregulate? The case of electricity. *Journal of Regulatory Economics*, 56(1), 1-32, 2019.
4. Kuster, C., Rezgui, Y., & Mourshed, M. Electrical load forecasting models: A critical systematic review. *Sustainable cities and society*, 35, 257-270, 2017.
5. Zhao, H., & Tang, Z.. The review of demand side management and load forecasting in smart grid. *In 2016 12th World Congress on Intelligent Control and Automation (WCICA)* (pp. 625-629). IEEE, 2016.
6. Zor, K., Timur, O., & Teke, A. (2017, June). A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting. *In 2017 6th International Youth Conference on Energy (IYCE)* (pp. 1-7). IEEE, 2017.
7. Almalaq, A., & Edwards, G. A review of deep learning methods applied on load forecasting. *In 2017 16th IEEE international conference on machine learning and applications (ICMLA)* (pp. 511-516). IEEE, 2017.
8. Upadhaya, D., Thakur, R., & Singh, N. K. A systematic review on the methods of short term load forecasting. *In 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC)* (pp. 6-11). IEEE, 2019.
9. Mohammed, Z. C., Humberto, A. R. J., Kiplel., M. C. & Chepkoech, M. A Meta-Analysis on Sustainable Supply Chain Management: An Analytical approach. European Journal of Logistics, *Purchasing and Supply Chain Management.* 2017.
10. Neuman, L. W. (2007). Social research methods, 7/E. *Pearson Education India.*
11. Almeshaiei, E., &Soltan, H. A methodology for electric power load forecasting. *Alexandria Engineering Journal*, 50(2), 137-144, 2011.
12. Yang, L., & Yang, H. A Combined ARIMA-PPR Model for Short-Term Load Forecasting. *In 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)* (pp. 3363-3367), 2019. IEEE.
13. Huang, S. J., & Shih, K. R. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. *IEEE Transactions on power systems*, 18(2), 673-679, 2003.
14. Amjady, N. (2001). Short-term hourly load forecasti using time-series modeling with peak load estimation capability. IEEE Transactions on power systems, 16(3), 498 505.
15. Gaur, M., &Majumdar, A. (2017). *One-Day-Ahead Load Forecasting using nonlinear Kalman filtering algorithms*.
16. Fan, S., & Hyndman, R. J. Short-term load forecasting based on a semi-parametric additive model. *IEEE Transactions on Power Systems*, 27(1), 134-141, 2011.
17. Unsihuay-Vila, C., De Souza, A. Z., Marangon-Lima, J. W., & Balestrassi, P. P. Electricity demand and spot price forecasting using evolutionary computation combined with chaotic nonlinear dynamic model. *International journal of electrical power & energy systems*, 32(2), 108-116, 2010.

18. Dong, M., & Grumbach, L. A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction. *IEEE Transactions on Smart Grid*, 11(1), 470-482, 2019.
19. Taylor, J. W. Short-term load forecasting with exponentially weighted methods. *IEEE Transactions on Power Systems*, 27(1), 458-464, 2011.
20. Taylor, J. W., & McSharry, P. E. Short-term load forecasting methods: An evaluation based on european data. *IEEE Transactions on Power Systems*, 22(4), 2213-2219, 2007.
21. Kavanagh, K. Short Term Demand Forecasting for the Integrated Electricity Market. *Student Journal of Energy Research*, 2(1), 1, 2017.
22. Zhang, Z., Hong, W. C., & Li, J. Electric load forecasting by hybrid self-recurrent support vector regression model with variational mode decomposition and improved cuckoo search algorithm. *IEEE Access*, 8, 14642-14658, 2020.
23. Fan, S., & Chen, L. Short-term load forecasting based on an adaptive hybrid method. *IEEE Transactions on Power Systems*, 21(1), 392-401, 2006.
24. Pai, P. F., & Hong, W. C. Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Conversion and Management*, 46(17), 2669-2688, 2005.
25. Ribeiro, G. T., Mariani, V. C., & dos Santos Coelho, L. Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting. *Engineering Applications of Artificial Intelligence*, 82, 272-281, 2019.
26. Reis, A. R., & Da Silva, A. A. Feature extraction via multiresolution analysis for short term load forecasting. *IEEE Transactions on power systems*, 20(1), 189-198, 2005.
27. Lv, Y. C., Xu, X., Xu, R. L., &Ren, H. (2019, January). Research on Short-Term Load Forecasting Approach for Smart Grid. *In 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)* (pp. 602-605). IEEE.
28. Hamid, M. A., & Rahman, T. A. Short term load forecasting using an artificial neural network trained by artificial immune system learning algorithm. *In 2010 12th International Conference on Computer Modelling and Simulation* (pp. 408-413). IEEE, 2010.
29. Hippert, H. S., Bunn, D. W., & Souza, R. C. Large neural networks for electricity load forecasting: Are they overfitted?. *International Journal of forecasting*, 21(3), 425-434, 2005.
30. Michanos, S. P., Tsakoumis, A. C., Fessas, P., Vladov, S. S., &Mladenov, V. M. (2003, July). Short-term load forecasting using a chaotic time series. In Signals, Circuits and Systems, 2003. SCS 2003. *International Symposium on* (Vol. 2, pp. 437-440). IEEE.
31. Wang, Y., Zhang, N., Tan, Y., Hong, T., Kirschen, D. S., & Kang, C. Combining probabilistic load forecasts. *IEEE Transactions on Smart Grid*, 10(4), 3664-3674, 2018.
32. Li, S., Wang, P., & Goel, L. A novel wavelet-based ensemble method for short-term load forecasting with hybrid neural networks and feature selection. *IEEE Transactions on power systems*, 31(3), 1788-1798, 2015.
33. Charytoniuk, W., & Chen, M. S. Very short-term load forecasting using artificial neural networks. *IEEE transactions on Power Systems*, 15(1), 263-268, 2000.
34. Dong, X., Qian, L., & Huang, L. Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach. In 2017 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 119-125). IEEE, 2017.
35. Din, G. M. U., &Marnerides, A. K. Short term power load forecasting using deep neural networks. *In 2017 International Conference on Computing, Networking and Communications (ICNC)* (pp. 594-598). IEEE, 2017.

36. Chen, K., Chen, K., Wang, Q., He, Z., Hu, J., & He, J. Short-term load forecasting with deep residual networks. *IEEE Transactions on Smart Grid*, 10(4), 3943-3952, 2018.
37. Chen, Y., Luh, P. B., & Rourke, S. J. (2008, June). Short-term load forecasting: Similar day based wavelet neural networks. *In 2008 7th World Congress on Intelligent Control and Automation* (pp. 3353-3358). IEEE.
38. Li, S., Wang, P., & Goel, L. A novel wavelet-based ensemble method for short-term load forecasting with hybrid neural networks and feature selection. *IEEE Transactions on power systems*, 31(3), 1788-1798, 2015.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Numerical Analysis of Impact of Relative Humidity on Crossflow Heat Exchangers with Staggered Configuration at Maximum Operating Temperature

Arshan Ahmed[1], Atta ul Mannan Hashmi[1], Fahad Rafi Butt[2], Shahbaz Ghani[1] and Imran Akhtar[1]

[1] Department of Mechanical Engineering, NUST College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Islamabad, Pakistan.

[2] Research Associate, Digital Pakistan Lab, NUST College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Islamabad, Pakistan.

* Correspondence: Arshan Ahmed; arshan721@yahoo.com.

Citation | Ahmad. A, Hashmi. A. M, Butt. F. R, Ghani. S and Akhtar. I. "Numerical Analysis of Impact of Relative Humidity on Crossflow Heat Exchangers With Staggered Configuration at Maximum Operating Temperature", International Journal of Innovations In Science and Technology, Special Issue, pp: 84-95, 2021.

Received | Dec 14, 2021; Revised | Jan 08, 2022 Accepted | Jan 12, 2022; Published | Jan 14, 2022

Heat exchangers are employed in numerous applications of industry, automotive and air conditioning systems. The efficacy of heat exchangers depends upon various factors e.g., Reynolds number (Re) of the fluids, geometry of heat exchanging surfaces, and the Prandtl number of the cooling air. In this paper, the working of a crossflow heat exchanger with elliptical tubes is simulated numerically for $5000 \leq Re \leq 20000$ at its maximum operating temperature of 323K. The tubes were arranged in a staggered way. The radical investigations were done at one-of-a-kind relative humidity ranges within the cooling air ranging from 0% to 80%. The relative humidity was modeled in the shape of mass fractions of water vapors in the air. The thermos-physical properties of dry and moist air were employed for the analysis. The impact of this changing of relative humidity on forced convection heat transfer of heat exchangers is examined in the form of percentage change in Nusselt number. With the increase in moisture content in the air, the Nusselt number was observed increased up to 4.5%. The paper provides a tool to analyze the Nusselt number of the elliptical-shaped heat exchanger while operating in moist atmospheric conditions.

Keywords: Heat transfer; relative humidity; Nusselt number; Prandtl number and Reynolds number.

INTRODUCTION

Heat exchangers are used in numerous industrial applications to transfer heat from one fluid to another. The efficacy of the amount of heat transfer depends upon various factors including Reynolds number (Re), area of heat transfer surfaces, and Prandtl number (Pr) of air. The case in which heat transfer from the surface requires external means to move the fluid

over it, is termed Forced Convection Heat Transfer (FCHT). For higher velocities, natural convection heat transfer through buoyant forces is considered negligible as compared to FCHT. Nusselt number (Nu) provides the means to access the temperature gradient in the non-dimensional form and in this paper, the Nusselt number presented depends only on forced convection. Heat exchanger used for the analysis is of crossflow type with a staggered arrangement of tubes having an elliptical shape.

Different types of performance analysis with changing parameters have already been done on heat exchangers other than humid air, especially in forced convection mode. Suyi and Shizhou [1] devised correlations for Nu depending upon Reynolds and Prandtl numbers of air up to the longitudinal and transverse pitch (SL/ST) ratio of 4. Iskra and Simonson [2] developed a relationship for convective mass transfer as a function of Reynolds number and relative humidity in a rectangular duct. Reynoso-jardón et al. [3] numerically conducted an analysis of crossflow heat exchangers for increased heat transfer by varying the transverse and longitudinal pitches. Sayed Ahmed et al. [4] concluded that for a staggered wing-shaped bundle of tubes Nusselt number increases with Reynolds number and it reaches its maximum value with the angle of attack of 45 degrees. Kanthimathi et al. [5] numerically deduced that increasing the array of tubes has an increasing effect on heat transfer. Gharbi et al. [6] came up with an expression for Nusselt number based on Reynolds number for the range $5000 \leq Re \leq 25000$ under specified conditions.

Still et al. [7] carried out experiments to calculate the effect of relative humidity on convective heat transfer of small cylinders in controlled conditions. They concluded that the relationship of heat transfer rates in dry and humid air is only a function of the molar fraction of water vapors in the air and does not depend on the Reynolds number and temperature of the air. Uguru-okorie and Ikpotokin [8] compared staggered and aligned tubes configurations for heat transfer rates experimentally and found an increase of 4.28% in a staggered configuration. Zhang et al. [9] numerically forecasted the change occurring in Nusselt number of natural convection heat transfer at different temperatures with relative humidity. Zukauskas & Ulinskas [10] presented the correlations for heat transfer and pressure drop for aligned and staggered crossflow heat exchangers in the range of $1000 < Remax < 2 \times 106$ with $0.7 < Pr < 500$ and $ST/SL < 0.7$.

From the literature assessment, it could be deduced that elliptical-fashioned tubes provide better warmth transfer costs in comparison to other shapes at exceptional Reynolds numbers. Apart from other design and overall performance analyses, research for the impact of various air properties in terms of relative humidity on the elliptical-formed tubes heat exchanger is essential to gauge its performance in a damp ecosystem.

This novel work gauges the efficiency in the form of exchange in Nusselt quantity of warmth exchangers due to trade in relative humidity of the cooling air on the most working temperature which the heat exchanger comes throughout at its operation. The trade that is predicted to come back relies upon the thermos-physical houses of air. The alternate has been investigated at the variety of maximum Reynolds numbers which the warmth exchanger comes to throughout some point of its operation i.e. $5000 < Re < 20000$.

PROBLEM SETUP

Model Characteristics. The heat exchanger available in the facility has been used for the analysis. It is a crossflow heat exchanger with elliptical-shaped tubes with a staggered

arrangement. The major and minor axis ratio of the tubes of the heat exchanger is 5.76. Heat exchanger has a total of 1140 × water tubes divided equally into two flow-pass configurations. The tubes are divided into 7 × staggered rows with transverse and longitudinal pitches mentioned in Figure 1. The air flows over the bank parallel to the longitudinal axis of the tube. The inlet air is uniform with constant dynamic viscosity and thermal conductivity as per the relative humidity [11].

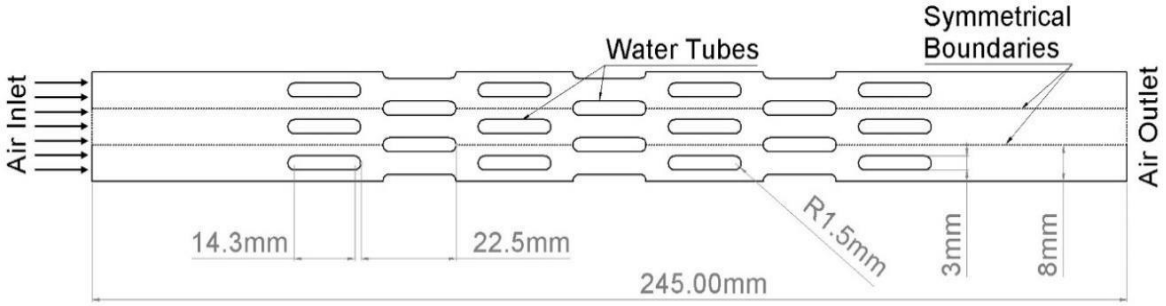


Figure 1. Geometry details and boundary conditions.

Boundary Conditions. The air outlet was ready as a pressure outlet. Symmetric boundary situations were carried out at the lateral obstacles of the area. The investigations were completed on the maximum air temperature of 323K which the warmth exchanger faces during its operation inside the facility. 323K was also set as the working temperature for the simulation. Thermo-bodily houses of air were also considered at an identical temperature [11]. The range of Reynolds number, already described, for the analysis encompasses the ambit of air velocities which heat exchanger faces in its operation. The relative humidity was introduced in the air in the form of mass fractions of water vapors using the species transport model in Ansys Fluent®. The variation of mass fraction of water vapors concerning relative humidity [12] at 323K is shown in Figure 2.

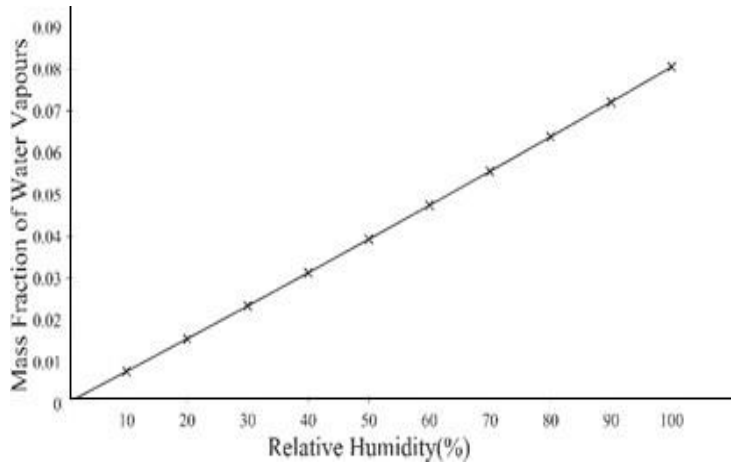


Figure 2. Mass fraction of water vapors with respect to relative humidity in air at 323K.

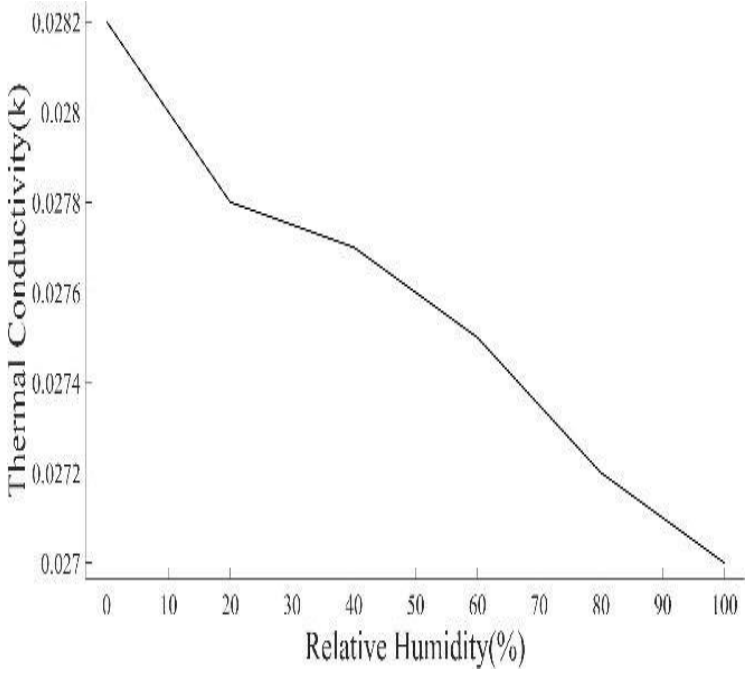


Figure 3. Thermal conductivity of air at 323K

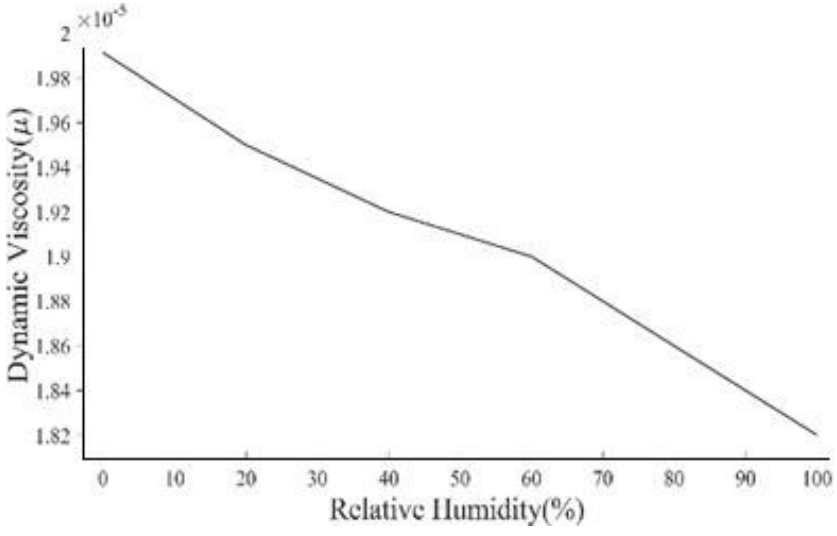


Figure 4. Dynamic viscosity of air at 323K.

To minimize the effect of turbulence on heat transfer, turbulence intensity was set to 5% for all the cases of Reynolds numbers. At the measured water flow rate of 1.95 x 10-5m³/sec per tube and assuming the waterside temperature difference as 10K, the tube walls were kept at a constant heat flux of 11500 W/m². Tubes walls were also applied with no-slip boundary conditions. Along with symmetric boundary conditions, lateral boundaries have zero diffusive flux. In the analysis, the air was treated as an ideal gas and its specific heat was calculated by applying the mixing law [13]. Thermal conductivity and dynamic viscosity of

humid air with different relative humidity levels [11] at 323K were shown in Figure 3 and figure 4 respectively. The fluid properties were considered constant at the specified humidity levels. Gravity was assumed to be acting in the direction of airflow while the buoyancy and radiation effects were neglected being minimal.

Ansys Fluent® was used for solving the governing equations of the grid with quadrilateral volumes for this model to the minimal scaled residual level of 10^{-6}. The pressure speed coupling approach with the okay-ε turbulence model was used to solve the computational grid because for fixing the elliptical bank of tubes, the prediction of the RNG model was far better than popular-ε turbulence models [14], so the former was used inside the analysis. The coupled scheme was used to clear up the discretized shape of equations. Second-order upwind method was followed for momentum and turbulent kinetic strength equations.

The geometry was modeled using the two-dimensional based grid with $y+ \approx 1$. To make sure that the solution was impartial of mesh size and density, a mesh independence test [15] was achieved for the calculation of the average Nusselt variety. In the mesh independence study, the solution becomes independent of the grid when the number of cells reaches 386869. The zoomed grid was depicted in Figure 5.

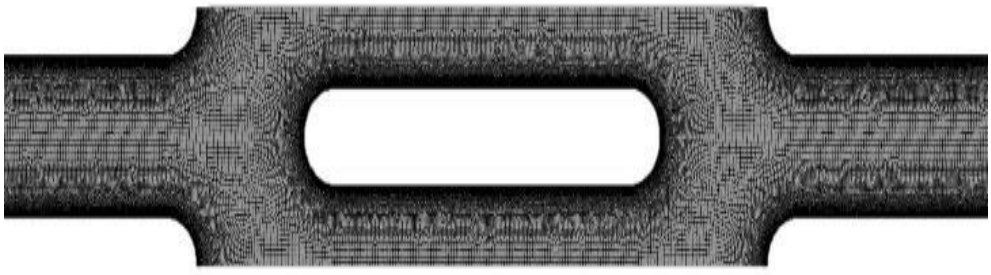


Figure 5. Blow-up of a single tube in computational domain grid.

The dimensionless temperature gradient on the surface due to convective warmness within the form of the Nusselt number is dependent on the Reynolds and Prandtl numbers [16]. A variety of literature on the effect of Reynolds numbers on the convective heat transfer of elliptical-shaped tubes is already available, so in this paper, the concentration is kept on the analysis of the influence of relative humidity on the Nusselt number of elliptical-shaped tubes heat exchanger.

Before performing the analysis, the CFD model is validated with the one presented by Gharbi et al. [6] within the range $5000 < Re < 20000$. During validation, the mass fraction of water vapors in the inlet is set to zero so that the heat transfer process involves dry air only. The comparison of the results is presented in Figure 6. Both the models show the same trend of rising in Nusselt number with the increasing Reynolds number with the maximum results' deviation of 7.5%. Contours of static temperature at 323K for the range of Reynolds number as displayed in Figure 8. depict the change in the thermal boundary layer with the change in Reynolds number, attributing to greater Nusselt numbers at higher velocities.

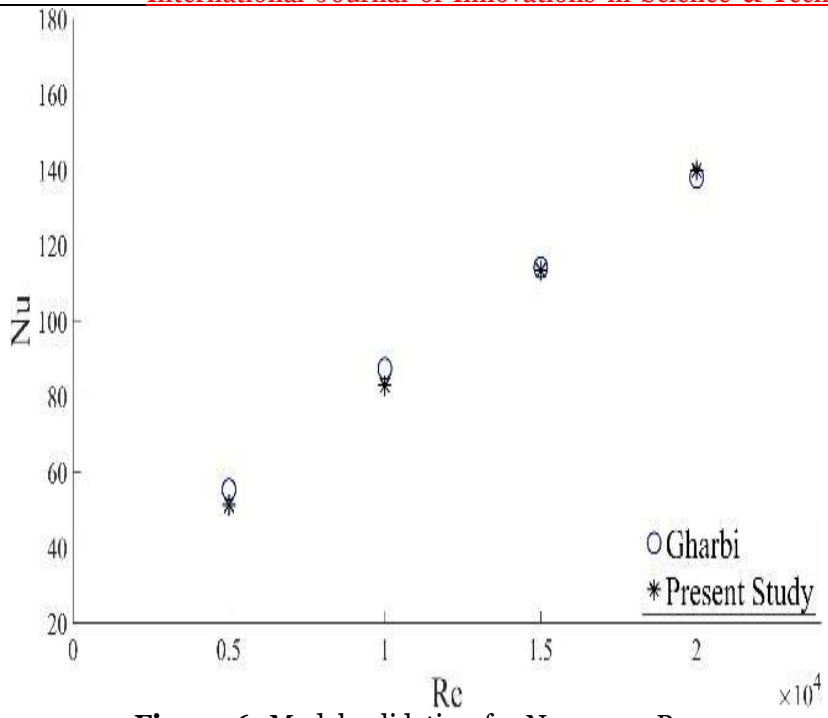


Figure 6. Model validation for Nu versus Re.

METHODOLOGY FOR HEAT TRANSFER CALCULATIONS

The amount of heat required to be dissipated from the water was calculated using the equation:

$$Q = \dot{m}_w c_{pw} \Delta T_w \quad (1)$$

where $\dot{m}_w$ is the mass flow rate of water in a single tube in one pass, c_{pw} is the specific heat of water and ΔT_w is the temperature difference in the inlet and outlet of a water tube in one pass.

The air side heat transfer coefficient $\bar{h}$ is calculated using the equation:

$$Q = \bar{h} A \Delta T_{lm} \quad (2)$$

Where Q/A is the heat flux which is emitted by the water tube and ΔT_{lm} is the log mean temperature difference which is calculated by using the equation:

$$\Delta T_{lm} = \frac{T_{a_{in}} - T_{a_{out}}}{ln(\frac{T_{a_{in}} - T_{wall}}{T_{a_{out}} - T_{wall}})} \quad (3)$$

Average Nusselt number is calculated for the bank of tubes to analyze the temperature gradient and amount of heat transfer from the equation:

$$\overline{Nu} = \frac{\bar{h} L_c}{k_a} \quad (4)$$

Where L_c is the characteristic length of the tube and k_a is the thermal conductivity of air at that specified humidity level.

Prandtl number is defined as the ratio of momentum diffusivity to thermal diffusivity and is given by the expression:

$$Pr = \frac{c_p \mu}{k} \quad (5)$$

This percentage exchange of Nusselt number was computed as:

$$\% \, Change \, in \, Air \, Nu = \frac{Nu_{humid} - Nu_{dry}}{Nu_{dry}} \times 100 \qquad (6)$$

Results and Discussion

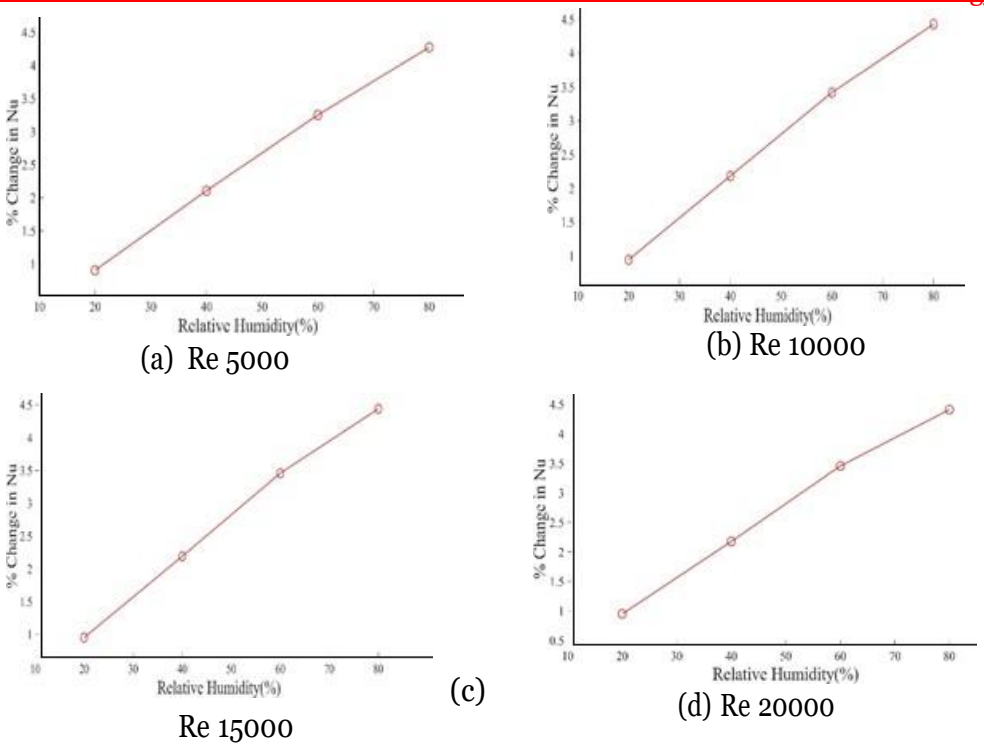
Influence of Relative Humidity (Mass Fraction of Water Vapours) on Nusselt Number.

In the analysis, as already stated, the relative humidity is simulated in the form of mass fractions of water vapors in the air. The greater the relative humidity, the greater will be the mass fraction. Mass fraction also depends upon the temperature of the air. Water content at higher temperatures will be more in the air as compared to the lower temperatures at the same humidity levels. However, for the case in this paper, the temperature of the inlet air is kept only at 323K. With the change in relative humidity, the thermophysical properties of the air were observed changed and thus the Prandtl number.

Since the Nusselt number is depending on the Prandtl range, so it also adjustments with the trade-in relative humidity. In this paper, this alteration has been calculated numerically the usage of the CFD code within the shape of percentage exchange in Nusselt quantity for relative humidity from 0% to 80%.

The share of relative humidity is numerous from 0% to eighty% with an increment of 20% at the range of the Reynolds range already mentioned. The results show that the Nusselt range are adjusted with the trade-in relative humidity. This transformation is attributed to the trade-in Prandtl wide variety of the air with the variant within the relative humidity. At 323K, the functionality of air to preserve water content material increases significantly with the boom in the percentage of relative humidity. For that reason, growing the warmth switch coefficient of the air has increased the Nusselt variety. The same can also be defined from the fact that the molecular weight of the air with 0% humidity content material is greater than that of water vapors and with the aid of increasing the humidity ratio in the air, its molecular weight decreases and it will increase the Nusselt quantity.

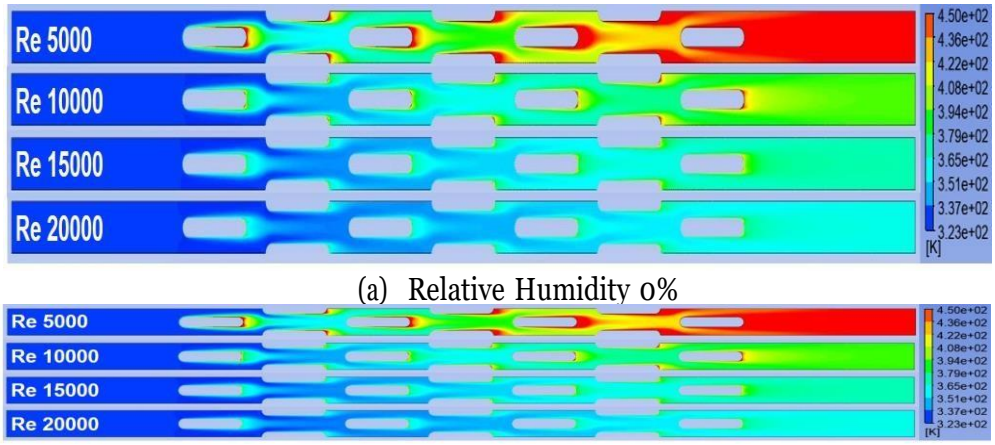
The change in molecular weight or density becomes more pronounced when the water content in the air increases. Thus, the change in Nusselt number becomes more pronounced at higher relative humidity levels. The percentage change in the Nusselt number concerning relative humidity at 323K is presented graphically in Figure 7.



(a) Re 5000 (b) Re 10000

(c) Re 15000 (d) Re 20000

Figure 7. Change in Nu with relative humidity at 323K for different Re.



(a) Relative Humidity 0%

(b) Relative Humidity 80%

Figure 8. Temperature contours for range of Re at minimum and maximum relative humidity levels at inlet temperature 323K.

The heat transfer coefficient was observed increased with the increase in Reynolds number and relative humidity. For the range of Reynolds number selected for the study, the change produced in the Nusselt number due to the change in humidity content of air that was observed independent of Reynolds number.

Conclusion.

Relative humidity affects the thermos-physical properties in the air, which influence the airside Nusselt number of heat exchanger. This effect has been calculated and analyzed using the CFD code at 323K on a crossflow heat exchanger with elliptical-shaped tubes arranged in a staggered fashion. Prandtl number changes with an increase in relative humidity of the air. From the analysis, it can be concluded that the change in Nusselt number due to the change in relative humidity in air does not depend on the Reynolds number. The Nusselt number changes up to 4.5% by increasing the relative humidity from 0% to 80% at 323K. The share exchange in Nusselt, wide variety with the increase in relative humidity will grow if the operating temperature is elevated because the ability of air to maintain water vapors increases with the temperature upward thrust. In similar manner, the proportion will lower upon lowering the running temperature. In the same way, the percentage will decrease upon lowering the operating temperature.

Acknowledgement. This manuscript has not been published or submitted to other journals previously. This research is supported by NUST College of Electrical & Mechanical Engineering, National University of Sciences and Technology, Pakistan.

Author's Contribution. Arshan Ahmed and Atta ul Mannan Hashmi developed the theory and performed the numerical computations. Fahad Rafi Butt verified the analytical results. Shahbaz Ghani and Imran Akhtar conceived the idea to investigate the effect of relative humidity on Nusselt number and supervised the findings of this work. All authors discussed the results in detail and contributed to the final manuscript.

Conflict of interest. The authors declare no conflict of interest for publishing this manuscript in IJIST.

REFRENCES

[1] H. Suyi and P. Shizhou, "*Convection and heat transfer of elliptical tubes*" "Heat Mass Transfer" 1995, vol. 30, no. 6, pp. 411–415.

[2] C. R. Iskra and C. J. Simonson, "*Effect of humidity on the convective mass transfer coefficient in a rectangular duct*" "Saudi Medical Journal" 2012, vol. 33, pp. 3–8.

[3] E. Reynoso-jardón, G. Urquiza, O. Tenango-pirin, and Y. Mariaca-beltrán, "*Numerical study of a cross flow heat exchanger varying the transverse and longitudinal space to increase heat transfer*" "International Journal of Combinatorial Optimization Problems and Informatics" 2020, vol. 11, pp. 106–114.

[4] S. A. E. Sayed Ahmed, E. Z. Ibrahiem, O. M. Mesalhy, and M. A. Abdelatief, "*Heat transfer characteristics of staggered wing-shaped tubes bundle at different angles of attack*," "Heat Mass Transfer" 2014, vol. 50, no. 8, pp. 1091–1102.

[5] T. Kanthimathi, P. Bhramara, and G. N. S. Ram, "*Numerical analysis on the effect of heat Transfer rate by varying the tube arrays in a cross-flow heat exchanger*" "International Journal of Scientific & Engineering Research" 2016, vol. 7, no. 6, pp. 13–16.

[6] N. El Gharbi, A. Kheiri, M. El Ganaoui, and R. Blanchard, "*Numerical optimization of heat exchangers with circular and non-circular shapes*" "Case Studies in Thermal Engineering" 2015, vol. 6, pp. 194–203.

[7] M. Still, H. Venzke, F. Durst, and A. Melling, "*Influence of humidity on the convective heat transfer from small cylinders*" "Experiments in Fluids" 1998, vol. 24, no. 2, pp. 141–150.

[8] D. Uguru-okorie and I. Ikpotokin, "*Experimental comparison of staggered and in-line tube-bank thermal performance"* "International Journal of Mechanical Engineering and Technology" 2018, vol. 9, pp. 1061–1070.

[9] J. Zhang, A. Gupta, and J. Baker, "*Effect of relative humidity on the prediction of natural convection heat transfer coefficients*" "Heat Transfer Engineering" 2007, vol. 28, no. 4, pp. 335–342.

[10] B. Prasad, A. Tawfek, and V. Rao, "*Heat transfer from aerofoils in cross-flow*" "International Communications in Heat and Mass Transfer" 1992, vol. 19, no. 6, pp. 879–890.

[11] P. T. Tsilingiris, "*Thermo-physical and transport properties of humid air at temperature range between 0 and 100°C*" "Energy Conservation and Management" 2008, vol. 49, no. 5, pp. 1098–1110.

[12] Water moisture content of humid air calculator, http://www.quadco.engineering.com/en/know-how/cfd-calculate-water-fraction-humid-air.html.

[13] ANSYS Inc, "*ANSYS Fluent® User's Guide*" 2013, Release 15.

[14] L. M. Smith and W. C. Reynolds, "*On the Yakhot-Orszag renormalization group method for deriving turbulence statistics and models*," "Physics of Fluids A: Fluid Dynamics" 1992, vol. 4, no. 2, pp. 364–390.

[15] Importance of mesh independence study & why you should care, https://www.maxshear.com/training/importance-of-mesh-independence-study-amp-why-you-should-do-it.

[16] F. P. Incropera, D. P. Dewitt, T. L. Bergman, and A. S. Lavine, "*Fundamentals of heat and mass transfer*" "John Willey & Sons" 2006, ed. 6th, p. 370.

[17] A. Zukauskas, "*Heat Transfer from tubes in crossflow*" "Advances in Heat Transfer" 1972, vol. 8, pp. 93–160.

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Estimated Zones of Saint-Venant Equations for Flood Routing with Over Bank Unsteady Flow in Open Channel

Talat Nazir, Sohail Ahmad Awan
Department of Mathematics, COMSATS University Islamabad, Abbottabad Campus, Abbottabad 22060, Pakistan.
* Correspondence: Talat Nazir talat@cuiatd.edu.pk
Citation | Nazir. T, and Awan. S. A, "Estimated Zones of Saint-Venant Equations for Flood Routing with Over Bank Unsteady Flow in Open Channel". International Journal of Innovations in Science and Technology" Vol 3, Special Issue, pp: 96-109, 2021.
Received | Dec 9, 2021; Revised | Dec 20, 2021 Accepted | Dec 21, 2021; Published | Dec 25, 2021.

Abstract.

In this paper, we learn about the control of open channel water glide under the flood routing conditions. Generally, for flood routing in rivers, the Saint-Venant equations will be used which can be solved by finite distinction method. Saint-Venant equations will be converted into nonlinear equations and will be solved using the Preissmann scheme in the finite difference method. Using the Newton Raphson method, the set of equations will be changed into linear equations and will be solved by the space method. Our aims are to the estimated zones of Saint-Venant equations for flood routing by using the finite difference method with over bank unsteady flow in an open channel. The effectiveness of this method to optimize the choice of finite difference method is more accurate than other methods having adequate space and time steps.

Keywords: Flood routing; Saint-Venant equations; sprace method; over bank flow; hydraulic radius.

1. INTRODUCTION

Over the last two decades, the study of open channel water flow has become an active research area due to its numerous advantages. It has been extensively investigated by many researchers. The hydraulic and hydrological problems involved in the computation of flood waves are based on Saint Venant (SV) equations, the pair of continuity and momentum equations. This study is important in an open channel under non-uniform or unsteady flow. It is worth mentioning that in this case SV equations cannot be solved analytically. Stoker [1] obtained the approximate solution of SV equations using the explicit finite difference method described in [2]. Ooi and Weyer [3] described the control design for an irrigation channel derived from physical data. Many researchers have obtained approximate solution of SV equations in a particular case of open channels ([4], [5], [7], [8], [9], [10], [11], [12], [13] and [14]). In these models, river waves can be categorize as gravity, diffusion, or kinematic waves, that relates to exceptional types of momentum equations.

The flood routing is fully based on the unsteady flow (long wave or surges) and the water storage equations. At the point of the channel, a flood hydrograph will find out from the well-

known hydrograph at various points upstream or downstream via the known channel characters with the characteristics of side inflow or outflow between these two points. Nguyen and Kawano [15] obtained simultaneous solutions for flood routing in the open channel network by using the Preissmann method for their proposed model. Kazezylmaz-Alhan et al. [16] discussed the reliability of the finite difference method for solving proposed diffusion and kinematic wave equations that describe the overland flow. Kohne et al. [17] presented the diffusion and kinematic wave method for runoff and surplus computation. Das [18] presented the Muskingum model to find the flood path and obtained its coefficients by employing the optimization method. Sulistyono and Wiryanto [19] investigated the flood routing by dynamic wave model in the trapezoidal channels.

The main objective of this paper is to develop a quantitative method for identifying river wave types in the case of flood routing with the overbank flow. Different theoretical cases relate to the ratios between the central channel and flooded location concerning breadth and glide will be analyzed. Moreover, we define the estimated zones of SV equations for flood routing by using the Preissmann method with over bank unsteady flow in an open channel.

2. MATERIALS AND METHODS

For flood routing issues like river waves and dynamic modeling of one-dimensional, the numerical answer of SV equations will be used. In the case of the flooded region in the river, let B_1 and B_2 be the breath of the central channel and flooded region of the channel, respectively as given in Figure 1. Also let A_1 and A_2 be the cross-sectional area of the central channel and flooded region of the channel, respectively.

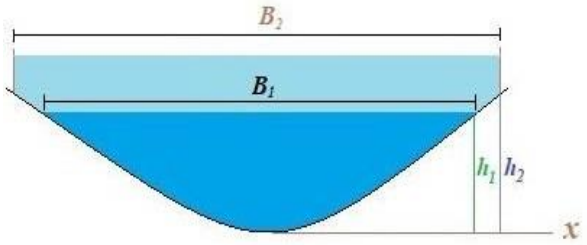


Figure 1. Open Channel having two breaths and cross-sectional areas for flow.

SV equations consist of two equations; continuity equation and momentum equation. The continuity (mass) equation is given as

$$\frac{\partial A_2}{\partial t} + \frac{\partial Q}{\partial x} = q,$$

which further implies that

$$B_2 \frac{\partial h}{\partial t} + \frac{\partial Q}{\partial x} = q. \tag{1}$$

Whereas the momentum equation is given as follows:

$$\frac{\partial Q}{\partial t} + \frac{\partial (\frac{Q2}{A2})}{\partial x} + gA_2(\frac{\partial h}{\partial x} + S_f) - gA_2S_0 + qV = 0,$$

$$\frac{\partial VA}{\partial t} + \frac{\partial (V_2A_2)}{\partial x} + gA2\,(\frac{\partial h}{\partial x} + S_f) - gA_2S_0 + qV = 0,$$

$$A_2\frac{\partial V}{\partial t} + A_2\frac{\partial (V_2)}{\partial x} + gA2\,(\frac{\partial h}{\partial x} + S_f) - gA_2S_0 + qV = 0.$$

Thus, the momentum equation becomes

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + g\frac{\partial h}{\partial x} + g(S_f - S_0) + \frac{qV}{A_2} = 0, \tag{2}$$

where is the flow high (m), is the velocity of flow (m/s), is the acceleration due to gravity (m/s²), is energy line slope, is the slope of the riverbed, is discharge (m³/s), is the horizontal distance (m), is time (s) and the q is a lateral inflow.

To derive this system, a basic assumption will be made: that there is one-dimensional flow in both regions of the channel (central region and flooded region), and there is lateral inflow or outflow.

Generally, the side channel will be rougher than the central channel. The velocity V in the central region is greater than as compared to the flooded region. In this case, the Manning formula can be applied separately to both regions in determining the velocity of both channels. In this case, the manning formula might be applied one by one to every section in determining the velocity of the section. After that, the discharge in the section will be computed. The total discharge will be equal to the sum of these discharges. Since velocity V in the central channel will be greater than the velocity in the flooded region so the component of discharge in the flooded area is small as compared to the discharge in the central channel; hence discharge Q can be approximated as follows.

$$Q \approx A_1 V, \tag{3}$$

where the cross-sectional area A_1 depends on x and t. As

$$A_1 = B_1 h. \tag{4}$$

Differentiating equation (3), we have

$$\begin{aligned}\frac{\partial Q}{\partial x} &= \frac{\partial (A_1 V)}{\partial x} = V\frac{\partial A1}{\partial x} + A_1 \frac{\partial V}{\partial x} \\ &= V\frac{\partial (B_1 h)}{\partial x} + (B_1 h)\frac{\partial V}{\partial x} \\ &= B_1 h\frac{\partial V}{\partial x} + (B_1 h)\frac{\partial V}{\partial x}.\end{aligned} \tag{5}$$

Substituting equation (5) in equation (1), we get

$$B_2 \frac{\partial h}{\partial t} + (B_1 V)\frac{\partial h}{\partial x} + (B_1 h)\frac{\partial V}{\partial x} = q,$$

$$B_2 \frac{\partial h}{\partial t} + B_1 (V\frac{\partial h}{\partial x} + h\frac{\partial V}{\partial x}) = q$$

and so,

$$\frac{B_2}{B_1}\frac{\partial h}{\partial t} + V\frac{\partial h}{\partial x} + h\frac{\partial V}{\partial x} = \frac{q}{B_1}. \tag{6}$$

If η is the ratio of flooded region breath B_2 and central channel breath B_2, that is, $\eta = \frac{B_2}{B_1}$, then we obtain

$$\eta\frac{\partial h}{\partial t} + V\frac{\partial h}{\partial x} + \frac{A1}{B1}\frac{\partial V}{\partial x} = \frac{q}{B1}. \tag{7}$$

Generally, in the momentum equation, the term will be calculated by the Manning formula. Since the velocity in the central region is greater than the velocity in the flooded region, thus the term of the Manning formula in the flooded region is smaller as compared to the central region of the channel. As

$$Q = \frac{A_1 R^{\frac{2}{3}} S_f^{\frac{1}{2}}}{n},$$

which implies that

$$VA_1 = \frac{A_1 R^{\frac{2}{3}} S_f^{\frac{1}{2}}}{n},$$

and hence

$$S_f = (Vn)^2 R^{\frac{-4}{3}} = (Vn)^2 R^{-m},$$

where R is the hydraulic radius (m), n is the coefficient of roughness and m is constant ($m = \frac{4}{3}$) and ($h << B_1$) is for large rivers. For central channel, the hydraulic radius is given by

$$R = \frac{B_1 h}{B_1 + 2h} = \frac{h}{1 + \frac{2h}{B_1}} \approx h. \quad (9)$$

Substituting equation (9) in equation (8), we get

$$S_f \approx n^2 V^2 h^{-m}. \quad (10)$$

As Froude number is a dimensionless number which is used in hydrodynamics to specify that how a particular model works in relation to a real system, so

$$F = \frac{V}{\sqrt{gh}},$$

that is,

$$gh\,F^2 = V^2. \quad (11)$$

Substituting equation (11) in (10) implies

$$S_f = n^2 g F^2 h^{1-m}.$$

As

$$\frac{\partial S_f}{\partial x} = n^2 g F^2 (1-m) h^{-m} \frac{\partial h}{\partial x},$$

so, we have

$$\frac{\partial h}{\partial x} = \frac{h^m}{n^2 g F^2 (1-m)} \frac{\partial S_f}{\partial x}. \quad (12)$$

Using (12) in (2), the momentum equation becomes

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{h^m}{n^2 g F^2 (1-m)} \frac{\partial S_f}{\partial x} + g(S_f - S_0) + \frac{qV}{A_2} = 0,$$

that is,

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{h^m}{\beta}\frac{\partial Sf}{\partial x} + g(S_f - S_0) + \frac{qV}{A_2} = 0, \quad (13)$$

where $\beta = n^2 F^2 (1-m)$. Continuity equation (7) and the momentum equation (13) provide the general form of SV equations with flooded region and the side channels rougher than the central channel. In such case, SV equations depends on parameter that appears in the continuity equation. In the particular case $\eta = 1$, equivalent to flood routing with no overbank flow, was considered widely by Moussa and Bocquillon [12]. These two equations can be solved by using finite difference method. According to this method, Preissmann method is used to solve SV equations.

3. RESULT AND DISCUSSION

The Pressman model has a vast application in flood routing hydrograph in the open channel as given in Figure 2.

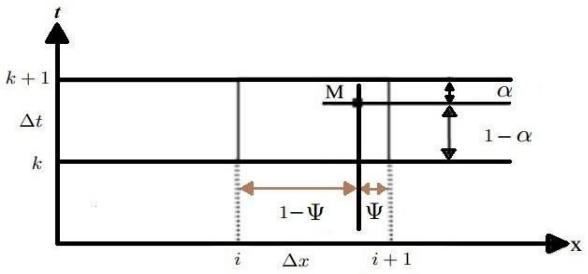


Figure 2. Preissmann method

By using above model, the time derivative is as follows

$$\frac{\partial f}{\partial t}\Big|_i^{k+1} = \psi\left[\frac{f_{i+1}^{k+1}-f_{i+1}^{k}}{\Delta t}\right] + (1-\psi)\left[\frac{f_i^{k+1}-f_i^{k}}{\Delta t}\right]. \quad (14)$$

Similarly, the space derivative is given as follow

$$\frac{\partial f}{\partial x}\Big|_i^{k+1} = \alpha\left[\frac{f_{i+1}^{k+1}-f_i^{k+1}}{\Delta x}\right] + (1-\alpha)\left[\frac{f_{i+1}^{k}-f_i^{k}}{\Delta x}\right]. \quad (15)$$

And the others terms are given as

$$f = \alpha\left[\frac{f_{i+1}^{k+1}+f_i^{k+1}}{2}\right] + (1-\alpha)\left[\frac{f_{i+1}^{k}+f_i^{k}}{2}\right]. \quad (16)$$

In the SV equations; for the terms $\frac{\partial h}{\partial t}$ and $\frac{\partial V}{\partial t}$ we will use equation (14), and for terms $\frac{\partial h}{\partial x}$, $\frac{\partial Sf}{\partial x}$ and $\frac{\partial V}{\partial x}$ we will use equation (15), and for terms $\frac{A1}{B_1}$, V, $\frac{q}{B_1}$, $\frac{V}{A_2}$ and S_f, we will use equation (16). The continuity equation (7), for $\psi = \frac{1}{2}$ becomes

$$\eta\left[\frac{h_{i+1}^{k+1}-h_{i+1}^{k}}{2\Delta t}+\frac{h_i^{k+1}-h_i^{k}}{2\Delta t}\right] + [(\alpha)\frac{V_{i+1}^{k+1}+V_i^{k+1}}{2}+(1-\alpha)\frac{V_{i+1}^{k}+V_i^{k}}{2}]\times$$
$$[(\alpha)\frac{h_{i+1}^{k+1}-h_i^{k+1}}{\Delta x}+(1-\alpha)\frac{h_{i+1}^{k}-h_i^{k}}{\Delta x}]+\left[(\alpha)\frac{\left(\frac{A_1}{B_1}\right)_{i+1}^{k+1}+\left(\frac{A_1}{B_1}\right)_i^{k+1}}{2}\right.$$
$$\left.+(1-\alpha)\frac{\left(\frac{A_1}{B_1}\right)_{i+1}^{k}+\left(\frac{A_1}{B_1}\right)_i^{k}}{2}\right]\times[(\alpha)\frac{V_{i+1}^{k+1}+V_i^{k+1}}{\Delta x}+(1-\alpha)\frac{V_{i+1}^{k}+V_i^{k}}{\Delta x}]$$
$$-q\left[(\alpha)\frac{\left(\frac{1}{B_1}\right)_{i+1}^{k+1}+\left(\frac{1}{B_1}\right)_i^{k+1}}{2}+(1-\alpha)\frac{\left(\frac{1}{B_1}\right)_{i+1}^{k}+\left(\frac{1}{B_1}\right)_i^{k}}{2}\right]=0,$$

which implies that

$$\eta[h_{i+1}^{k+1}+h_i^{k+1}]-\eta[h_{i+1}^{k}+h_i^{k}]+\frac{\Delta t}{\Delta x}\alpha^2(V_{i+1}^{k+1}+V_i^{k+1})(h_{i+1}^{k+1}-h_i^{k+1})$$
$$+\frac{\Delta t}{\Delta x}\alpha(1-\alpha)(V_{i+1}^{k+1}+V_i^{k+1})(h_{i+1}^{k}-h_i^{k})$$
$$+\frac{\Delta t}{\Delta x}\alpha(1-\alpha)(V_{i+1}^{k}+V_i^{k})(h_{i+1}^{k+1}-h_i^{k+1})$$
$$+\frac{\Delta t}{\Delta x}(1-\alpha)^2(V_{i+1}^{k}+V_i^{k})(h_{i+1}^{k}-h_i^{k})$$
$$+\frac{\Delta t}{\Delta x}\alpha^2\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k+1}+\left(\frac{A_1}{B_1}\right)_i^{k+1}\right)(V_{i+1}^{k+1}-V_i^{k+1})$$
$$+\frac{\Delta t}{\Delta x}\alpha(1-\alpha)\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k+1}+\left(\frac{A_1}{B_1}\right)_i^{k+1}\right)(V_{i+1}^{k}-V_i^{k})$$
$$+\frac{\Delta t}{\Delta x}\alpha(1-\alpha)\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k}+\left(\frac{A_1}{B_1}\right)_i^{k}\right)(V_{i+1}^{k+1}-V_i^{k+1})$$
$$+\frac{\Delta t}{\Delta x}(1-\alpha)^2\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k}+\left(\frac{A_1}{B_1}\right)_i^{k}\right)(V_{i+1}^{k}-V_i^{k})-q\Delta t\alpha\left(\left(\frac{1}{B_1}\right)_{i+1}^{k+1}+\left(\frac{1}{B_1}\right)_i^{k+1}\right)$$
$$-q\Delta t(1-\alpha)\left(\left(\frac{1}{B_1}\right)_{i+1}^{k}+\left(\frac{1}{B_1}\right)_i^{k}\right)=0. \quad (17)$$

Finally, a function T_i for interval i will be obtained as follows.

$$T_i(h_i^{k+1}, V_i^{k+1}, h_{i+1}^{k+1}, V_{i+1}^{k+1}) = \eta[h_{i+1}^{k+1} + h_i^{k+1}] + \eta D_1 + \frac{\Delta t}{\Delta x}\alpha^2\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k+1} + \left(\frac{A_1}{B_1}\right)_i^{k+1}\right)(V_{i+1}^{k+1} - V_i^{k+1}) + \frac{\Delta t}{\Delta x}\alpha D_2(V_{i+1}^{k+1} - V_i^{k+1}) + \frac{\Delta t}{\Delta x}\alpha D_3\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k+1} + \left(\frac{A_1}{B_1}\right)_i^{k+1}\right) + D_4 + \frac{\Delta t}{\Delta x}\alpha^2(V_{i+1}^{k+1} + V_i^{k+1})(h_{i+1}^{k+1} - h_i^{k+1}) + \frac{\Delta t}{\Delta x}\alpha D_5(h_{i+1}^{k+1} - h_i^{k+1}) + \frac{\Delta t}{\Delta x}\alpha D_6(V_{i+1}^{k+1} + V_i^{k+1}) + D_7 - D_8 - q\Delta t\alpha\left(\left(\frac{1}{B_1}\right)_{i+1}^{k+1} + \left(\frac{1}{B_1}\right)_i^{k+1}\right) = 0. \quad (18)$$

In equation (18), D_1 to D_8 are coefficients which are given below:

$$D_1 = -h_{i+1}^{k} - h_i^{k},$$

$$D_2 = (1-\alpha)\left(\left(\frac{A_1}{B_1}\right)_{i+1}^{k} + \left(\frac{A_1}{B_1}\right)_i^{k}\right),$$

$$D_3 = (1-\alpha)(V_{i+1}^{k} - V_i^{k}),$$

$$D_4 = \frac{\Delta t}{\Delta x} D_2 D_3,$$

$$D_5 = \alpha(1-\alpha)(V_{i+1}^{k} + V_i^{k}),$$

$$D_6 = (1-\alpha)(h_{i+1}^{k} - h_i^{k}),$$

$$D_7 = \frac{\Delta t}{\Delta x} D_5 D_6,$$

$$D_8 = q\Delta t(1-\alpha)\left(\left(\frac{1}{B_1}\right)_{i+1}^{k} + \left(\frac{1}{B_1}\right)_i^{k}\right).$$

Similarly, the momentum equation (13), for $\psi = \frac{1}{2}$ becomes

$$\left[\frac{V_{i+1}^{k+1} - V_{i+1}^{k}}{2\Delta t} + \frac{V_i^{k+1} - V_i^{k}}{2\Delta t}\right] + \left[(\alpha)\frac{V_{i+1}^{k+1} + V_i^{k+1}}{2} + (1-\alpha)\frac{V_{i+1}^{k} + V_i^{k}}{2}\right] \times \left[(\alpha)\left(\frac{V_{i+1}^{k+1} - V_i^{k+1}}{2}\right) + (1-\alpha)\left(\frac{V_{i+1}^{k} - V_i^{k}}{2}\right)\right] + \frac{1}{\beta}\left[(\alpha)\left(\frac{(h^m)_{i+1}^{k+1} + (h^m)_i^{k+1}}{2}\right) + (1-\alpha)\left(\frac{(h^m)_{i+1}^{k} - (h^m)_i^{k}}{2}\right)\right] \times \left[(\alpha)\left(\frac{(S_f)_{i+1}^{k+1} - (S_f)_i^{k+1}}{\Delta x}\right) + (1-\alpha)\left(\frac{(S_f)_{i+1}^{k} - (S_f)_i^{k}}{\Delta x}\right)\right] - gS_0 + g\left[(\alpha)\left(\frac{(S_f)_{i+1}^{k+1} + (S_f)_i^{k+1}}{2}\right) + (1-\alpha)\left(\frac{(S_f)_{i+1}^{k} + (S_f)_i^{k}}{2}\right)\right] + q\left[\alpha\left(\frac{\left(\frac{V}{A_2}\right)_{i+1}^{k+1} + \left(\frac{V}{A_2}\right)_i^{k+1}}{2}\right) + (1-\alpha)\left(\frac{\left(\frac{V}{A_2}\right)_{i+1}^{k} + \left(\frac{V}{A_2}\right)_i^{k}}{2}\right)\right] = 0,$$

which implies that

$$\frac{\Delta x}{g\Delta t}[V_{i+1}^{k+1}+V_i^{k+1}]-\frac{\Delta x}{g\Delta t}[V_{i+1}^{k}+V_i^{k}]+\frac{1}{g}\alpha^2(V_{i+1}^{k+1}+V_i^{k+1})(V_{i+1}^{k+1}-V_i^{k+1})$$
$$+\frac{1}{g}\alpha(1-\alpha)(V_{i+1}^{k}+V_i^{k})(V_{i+1}^{k+1}-V_i^{k+1})+\frac{1}{g}(1-\alpha)^2(V_{i+1}^{k}+V_i^{k})(V_{i+1}^{k}-V_i^{k})$$
$$+\frac{1}{g}\alpha(1-\alpha)(V_{i+1}^{k}+V_i^{k})(V_{i+1}^{k+1}-V_i^{k+1})+\frac{1}{g\beta}\alpha^2((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})$$
$$\times((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})+\frac{1}{g\beta}\alpha(1-\alpha)((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})((S_f)_{i+1}^{k}-(S_f)_i^{k})$$
$$+\frac{1}{g\beta}\alpha(1-\alpha)((h^m)_{i+1}^{k}+(h^m)_i^{k})((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})+\frac{1}{g\beta}(1-\alpha)^2$$
$$\times((h^m)_{i+1}^{k}+(h^m)_i^{k})((S_f)_{i+1}^{k}-(S_f)_i^{k})-\frac{2\Delta x}{g}gS_0+\Delta x\,\alpha((S_f)_{i+1}^{k+1}+(S_f)_i^{k+1})$$
$$+\Delta x(1-\alpha)((S_f)_{i+1}^{k}+(S_f)_i^{k})+\frac{q\Delta x}{g}\alpha\left(\left(\frac{V}{A_2}\right)_{i+1}^{k+1}+\left(\frac{V}{A_2}\right)_i^{k+1}\right)$$
$$+\frac{q\Delta x}{g}(1-\alpha)\left(\left(\frac{V}{A_2}\right)_{i+1}^{k}+\left(\frac{V}{A_2}\right)_i^{k}\right)=0. \quad (19)$$

Consequently, a function U_i for interval i will be obtained as follows.

$$U_i(h_i^{k+1},V_i^{k+1},h_{i+1}^{k+1},V_{i+1}^{k+1})=\frac{\Delta x}{g\Delta t}[V_{i+1}^{k+1}+V_i^{k+1}]+D_9$$
$$+\frac{\alpha^2}{g}((V_{i+1}^{k+1})^2+(V_i^{k+1})^2)+D_{10}(V_{i+1}^{k+1}-V_i^{k+1})+D_{11}(V_{i+1}^{k+1}+V_i^{k+1})+D_{12}$$
$$+\frac{1}{g\beta}\alpha^2((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})+\alpha D_{13}((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})$$
$$+\frac{1}{g\beta}\alpha D_{14}((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})+D_{15}+D_{16}+\Delta x\,\alpha((S_f)_{i+1}^{k+1}+(S_f)_i^{k+1})+D_{17}$$
$$+\frac{q\Delta x}{g}\alpha\left(\left(\frac{V}{A_2}\right)_{i+1}^{k+1}+\left(\frac{V}{A_2}\right)_i^{k+1}\right)+D_{18}=0, \quad (20)$$

where D_9 to D_{18} are given as:

$$D_9=-\frac{\Delta x}{g\Delta t}[V_{i+1}^{k}+V_i^{k}],$$
$$D_{10}=\frac{1}{g}\alpha(1-\alpha)(V_{i+1}^{k}+V_i^{k}),$$
$$D_{11}=\frac{1}{g}\alpha(1-\alpha)(V_{i+1}^{k}-V_i^{k}),$$
$$D_{12}=\frac{1}{g}\alpha^2[(V_{i+1}^{k})^2-(V_i^{k})^2]$$
$$D_{13}=\frac{1}{g\beta}(1-\alpha)((S_f)_{i+1}^{k}-(S_f)_i^{k}),$$
$$D_{14}=(1-\alpha)((h^m)_{i+1}^{k}+(h^m)_i^{k}),$$
$$D_{15}=D_{14}D_{13},$$
$$D_{16}=-2\Delta xS_0,$$
$$D_{17}=\Delta x(1-\alpha)((S_f)_{i+1}^{k}+(S_f)_i^{k}),$$

$$D_{18} = \frac{q\Delta x}{g}(1-\alpha)\left(\left(\frac{V}{A_2}\right)^{\mathbf{k}}_{\mathbf{i+1}} + \left(\frac{V}{A_2}\right)^{k}_{i}\right).$$

As in favor of each interval, the two vector functions and are the functions of four variables, namely. In the favor of each interval, we define two equations, and similarly, for intervals equation will arise. For each interval, we have a node. And for each node, there will be two unknowns (flow rate and depth), thus if there are unknowns, then two equations are formed. Thus, we will obtain two more equations from up and downstream conditions [20].

Finally, a system of nonlinear equations will be generated; to change the nonlinear equations to linear, we will use the Newton-Raphson method. According to this method, the derivative of functions is given as follows

$$\frac{\partial T_i}{\partial h^{\mathbf{k+1}}_{\mathbf{i}}} = \eta + \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right]\left[\frac{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2 - (A_1)^{\mathbf{k+1}}_{\mathbf{i}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2}\right]$$
$$+\frac{\Delta t}{\Delta x}\alpha D_3\left[\frac{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2 - (A_1)^{\mathbf{k+1}}_{\mathbf{i}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2}\right] - \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right]$$
$$-\frac{\Delta t}{\Delta x}\alpha D_5 + q\Delta x\alpha\left[\frac{\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2}\right],$$

that is,

$$\frac{\partial T_i}{\partial h^{\mathbf{k+1}}_{\mathbf{i}}} = \eta + \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right]\left[1 - \frac{(A_1)^{\mathbf{k+1}}_{\mathbf{i}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2}\right] + \frac{\Delta t}{\Delta x}\alpha D_3$$
$$\left[1 - \frac{(A_1)^{\mathbf{k+1}}_{\mathbf{i}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2}\right] - \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right] - \frac{\Delta}{\Delta x}\alpha D_5 + \left[\frac{q\Delta t\alpha}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2}\right]\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}.$$

(21)

Also

$$\frac{\partial T_i}{\partial h^{\mathbf{k+1}}_{\mathbf{i+1}}} = \eta + \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right]\left[\frac{(B_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\left(\frac{dA_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}} - (A_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i+1}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\right]^2}\right]$$
$$+\frac{\Delta t}{\Delta x}\alpha D_3\left[\frac{(B_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\left(\frac{dA_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}} - (A_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i+1}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\right]^2}\right] + \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right]$$
$$+\frac{\Delta t}{\Delta x}\alpha D_5 + q\Delta x\alpha\frac{\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i}}\right]^2},$$

that is,

$$\frac{\partial T_i}{\partial h^{\mathbf{k+1}}_{\mathbf{i+1}}} = \eta + \frac{\Delta t}{\Delta x}\alpha^2\left[V^{\mathbf{k+1}}_{\mathbf{i+1}} - V^{k+1}_{i}\right]\left[1 - \frac{(A_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\left(\frac{dB_1}{dh}\right)^{\mathbf{k+1}}_{\mathbf{i+1}}}{\left[(B_1)^{\mathbf{k+1}}_{\mathbf{i+1}}\right]^2}\right]$$

$$+\frac{\Delta t}{\Delta x}\,\alpha D_3\,[1-\frac{(A_1)_{i+1}^{k+1}\,(\frac{dB_1}{dh})_{i+1}^{k+1}}{[(B_1)_{i+1}^{k+1}]^2}]+\frac{\Delta t}{\Delta x}\alpha^2[V_{i+1}^{k+1}-V_i^{k+1}]+\frac{\Delta t}{\Delta x}\,\alpha D_5$$

$$\frac{\Delta t}{\Delta x}[\alpha\frac{}{[(B_1)_{i+1}^{k+1}]^2}]\,(\frac{dB_1}{dh})_{i+1}^{k+1}. \qquad (22)$$

Similarly, we have

$$\frac{\partial T_i}{\partial h_i^{k+1}}=-\frac{\Delta t}{\Delta x}\,\alpha^2\,((\frac{A_1}{B_1})_{i+1}^{k+1}+(\frac{A_1}{B_1})_i^{k+1})-\frac{\Delta t}{\Delta x}\,\alpha D_2+\frac{\Delta t}{\Delta x}\alpha^2[h_{i+1}^{k+1}-h_i^{k+1}]$$
$$+\frac{\Delta t}{\Delta x}\,\alpha D_6, \qquad (23)$$

$$\frac{\partial T_i}{\partial h_{i+1}^{k+1}}=-\frac{\Delta t}{\Delta x}\,\alpha^2\,((\frac{A_1}{B_1})_{i+1}^{k+1}+(\frac{A_1}{B_1})_{i+1}^{k+1})-\frac{\Delta t}{\Delta x}\,\alpha D_2+\frac{\Delta t}{\Delta x}\,\alpha^2\,[h_{i+1}^{k+1}-h_i^{k+1}]+\frac{\Delta t}{\Delta x}\,\alpha D_6, \qquad (24)$$

$$\frac{\partial U_i}{\partial h_i^{k+1}}=\frac{\alpha^2}{g\beta}((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})\,m(h^{m-1})_i^{k+1}-\frac{\alpha^2}{g\beta}((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})\,(\frac{\partial (S_f)_i^{k+1}}{\partial h_i^{k+1}})$$
$$+\alpha D_{13}(m(h^{m-1})_i^{k+1})+\frac{\alpha D_{14}}{g\beta}\frac{\partial (S_f)_i^{k+1}}{\partial h_i^{k+1}}+\Delta x\,\alpha\frac{\partial (S_f)_i^{k+1}}{\partial h_i^{k+1}}-\frac{V_{i+1}^{k+1}(\frac{dA_2}{dh})_{i+1}^{k+1}}{[(A_2)_{i+1}^{k+1}]^2}. \qquad (25)$$

As $S_f=n^2gF^2h^{1-m}$, we have

$$\frac{\partial (S_f)_i^{k+1}}{\partial h_i^{k+1}}=n^2gF^2(1-m)(h^{-m})_i^{k+1},$$

that is,

$$\frac{\partial (S_f)_i^{k+1}}{\partial h_i^{k+1}}=g\beta(h^{-m})_i^{k+1},$$

where $\beta=n^2F^2(1-m)$, and the above equation (25) becomes

$$\frac{\partial U_i}{\partial h_{i+1}^{k+1}}=\frac{\alpha^2}{g\beta}((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})\,m(h^{m-1})_i^{k+1}-\alpha^2((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})(h^{-m})_i^{k+1}$$
$$+\alpha D_{13}(m(h^{m-1})_i^{k+1})+\alpha D_{14}(h^{-m})_i^{k+1}+\Delta x\,\alpha g\beta(h^{-m})_i^{k+1}-\frac{dA_2}{[(A_2)_{i+1}^{k+1}]^2}(\frac{dA_2}{dh})_{i+1}^{k+1}. \qquad (26)$$

Also

$$\frac{\partial U_i}{\partial h_{i+1}^{k+1}}=\frac{\alpha^2}{g\beta}((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})\,m(h^{m-1})_{i+1}^{k+1}+\frac{\alpha^2}{g\beta}((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})\,(\frac{\partial (S_f)_i^{k+1}}{\partial h_{i+1}^{k+1}})$$
$$+\alpha D_{13}(m(h^{m-1})_{i+1}^{k+1})+\frac{\alpha D_{14}}{g\beta}\frac{\partial (S_f)_i^{k+1}}{\partial h_{i+1}^{k+1}}+\Delta x\,\alpha\frac{\partial (S_f)_i^{k+1}}{\partial h_{i+1}^{k+1}}-\frac{V_{i+1}^{k+1}(\frac{dA_2}{dh})_{i+1}^{k+1}}{[(A_2)_{i+1}^{k+1}]^2},$$

that is,

$$\frac{\partial U_i}{\partial h_{i+1}^{k+1}}=\frac{\alpha^2}{g\beta}((S_f)_{i+1}^{k+1}-(S_f)_i^{k+1})\,m(h^{m-1})_{i+1}^{k+1}+\alpha^2((h^m)_{i+1}^{k+1}+(h^m)_i^{k+1})(h^{-m})_{i+1}^{k+1}$$
$$+\alpha D_{13}(m(h^{m-1})_{i+1}^{k+1})+\alpha D_{14}(h^{-m})_{i+1}^{k+1}+\Delta x\,\alpha g\beta(h^{-m})_{i+1}^{k+1}-\frac{dA_2}{[(A_2)_{i+1}^{k+1}]^2}(\frac{dA_2}{dh})_{i+1}^{k+1}. \qquad (27)$$

In the similar way, the partial derivatives of remaining equations are

$$\frac{\partial U_i}{\partial V_i^{k+1}} = \frac{\Delta x}{g\Delta t} - \frac{2\alpha^2}{g} V_i^{k+1} - D_{10} + D_{11} + \frac{q\Delta x\alpha}{g(A_2)_i^{k+1}} \quad (28)$$

and

$$\frac{\partial U_i}{\partial V_{i+1}^{k+1}} = \frac{\Delta x}{g\Delta t} + \frac{2\alpha^2}{g} V_i^{k+1} + D_{10} + D_{11} + \frac{q\Delta x\alpha}{g(A_2)_i^{k+1}}. \quad (29)$$

As a result, in support of interval, equations can be produced. Hence, it is ample to get the other two equations from the upstream and downstream boundary conditions, so for all nodes, two unknowns and values will be obtained. The upstream borderline (node No. 1) may additionally be the inflow hydrograph, where is the discharge of influx hydrology, and the downstream boundary circumstance (node No.) may be the glide phase, the place is the overflow height, is the overflow coefficient factor, and is the weight.

Thus, the downstream and upstream conditions are given as

(h_K, V_K) ——— (h_{K+1}, V_{K+1})

(h_1, V_1) ——— (h_2, V_2)

As

$$T_0(h_1, V_1) = Q_1 - A_1 V_1, \quad (30)$$

$$T_{k+1}(h_{k+1}, V_{k+1}) = A_{k+1} V_{k+1} - CW(h_{k+1} - H)^{1.5}. \quad (31)$$

The initial condition is the steady flow before the flood, so the partial derivative of equation (30) and (31) are

$$\frac{\partial T_0(h_1, V_1)}{\partial h_1} = \frac{\partial Q_1}{\partial h_1} - \frac{\partial A_1 V_1}{\partial h_1} = -V_1 \frac{\partial A_1}{\partial h_1}, \quad (32)$$

$$\frac{\partial T_0(h_1, V_1)}{\partial V_1} = \frac{\partial Q_1}{\partial V_1} - \frac{\partial A_1 V_1}{\partial V_1} = -V_1 \frac{\partial V_1}{\partial V_1} = -V_1, \quad (33)$$

$$\frac{\partial T_{k+1}(h_{k+1}, V_{k+1})}{\partial h_1} = \frac{\partial A_{k+1} V_{k+1}}{\partial h_1} - \frac{\partial (C.W\,(h_{k+1} - H)^{1.5})}{\partial h_1} = V_{k+1} \frac{\partial A_{k+1}}{\partial h_1} - (1.5)CW\,(h_{k+1} - H)^{0.5}, \quad (34)$$

and

$$\frac{\partial T_{k+1}(h_{k+1}, V_{k+1})}{\partial V_{k+1}} = \frac{\partial A_{k+1} V_{k+1}}{\partial V_{k+1}} - \frac{\partial (C.W\,(h_{k+1} - H)^{1.5})}{\partial V_{k+1}} = A_{k+1} \quad (35)$$

To solve non-linear equations from (32) to (35), we will use the Newton-Raphson algorithm.

The Newton-Raphson method is one of the most familiar iterative schemes used to solve non-linear equations. First, we will write given equations in vector form:

$$T_i(x_1, x_2, \ldots, x_{2N}) = 0,$$

where $i = 1, 2, 3, \ldots, 2N$, $X = (x_1, x_2, \ldots, x_{2N})$denotes the vector of unknown variables. By Taylor series expansion, we have

$$T_i(x + \delta x) = T_i(x) + \sum_{i=1}^{2N} \frac{\partial T_i}{\partial x_j} \delta x_j + O(\delta x)^2$$

for $i = 1, 2, 3, \ldots, 2N$. In the above equation, first partial derivatives form a Jacobian matrix. If we take

$$J = \sum_{i=1}^{2N} \frac{\partial T_i}{\partial x_j}$$

then

$$T_i(x + \delta x) = T_i(x) + \mathrm{J}\, \delta x_j + O(\delta x)^2.$$

If we neglect higher order terms and set the left hand-side equal to zero, then we obtain a set of linear equations given by

$$J\, \delta x_j = -T_i(x).$$

The above system of linear equations in matrix form can be further solved by using Gaussian elimination method or LU decomposition method for the unknown value of x, and hence the approximate solution is obtained by

$$x_{new} = x_{old} + \delta x.$$

The iteration procedure will be carried out until a reset convergence stage is achieved. Now in our case, we write our above-mentioned equations in the same way as above

$$T_0(h_1^{n+1}, V_1^{n+1}) = 0.$$

By Taylor series expansion, we can write as follows

$$T_0(h_1^{n+1}, V_1^{n+1}, \Delta h_1^{n+1}, \Delta V_1^{n+1}) = T_0(h_1, V_1) + \frac{\partial T_0}{\partial h_1^{n+1}} \Delta h_1^{n+1} + \frac{\partial T_0}{\partial V_1^{n+1}} \Delta V_1^{n+1}.$$

Put $T_0(h_1^{n+1}, V_1^{n+1}, \Delta h_1^{n+1}, \Delta V_1^{n+1}) = 0$ and let $T_0(h_1, V_1) = t_0$, so the above equation will become

$$t_0 + \frac{\partial T_0}{\partial h_1^{n+1}} \Delta h_1^{n+1} + \frac{\partial T_0}{\partial V_1^{n+1}} \Delta V_1^{n+1} = 0,$$

that is,

$$\frac{\partial T_0}{\partial h_1^{n+1}} \Delta h_1^{n+1} + \frac{\partial T_0}{\partial V_1^{n+1}} \Delta V_1^{n+1} = -t_0. \tag{36}$$

Also,

$$T_i(h_i^{n+1}, V_i^{n+1}) = 0,$$

By the Taylor series expansion, we have

$$T_i(h_i^{n+1}, V_i^{n+1}, \Delta h_i^{n+1}, \Delta V_i^{n+1}) = T_i(h_i^{n+1}, V_i^{n+1}) + \frac{\partial T_i}{\partial h_i^{n+1}} \Delta h_i^{n+1} + \frac{\partial T_i}{\partial V_i^{n+1}} \Delta V_i^{n+1} + \frac{\partial T_i}{\partial h_{i+1}^{n+1}} \Delta h_i^{n+1} + \frac{\partial T_i}{\partial V_{i+1}^{n+1}} \Delta V_i^{n+1}.$$

Put $T_i(h_i^{n+1}, V_i^{n+1}, \Delta h_i^{n+1}, \Delta V_i^{n+1}) = 0$ and $T_i(h_i^{n+1}, V_i^{n+1}) = t_i$, we get that

$$\frac{\partial T_i}{\partial h_i^{n+1}} \Delta h_i^{n+1} + \frac{\partial T_i}{\partial V_i^{n+1}} \Delta V_i^{n+1} + \frac{\partial T_i}{\partial h_{i+1}^{n+1}} \Delta h_i^{n+1} + \frac{\partial T_i}{\partial V_{i+1}^{n+1}} \Delta V_i^{n+1} = -t_i. \tag{37}$$

Similarly, we obtain the following equations

$$\frac{\partial U_i}{\partial h_i^{n+1}} \Delta h_i^{n+1} + \frac{\partial U_i}{\partial V_i^{n+1}} \Delta V_i^{n+1} + \frac{\partial U_i}{\partial h_{i+1}^{n+1}} \Delta h_i^{n+1} + \frac{\partial U_i}{\partial V_{i+1}^{n+1}} \Delta V_i^{n+1} = -u_i. \tag{38}$$

$$\frac{\partial T_{N+1}}{\partial h_{N+1}^{n+1}} \Delta h_1^{n+1} + \frac{\partial T_{N+1}}{\partial V_{N+1}^{n+1}} \Delta V_{N+1}^{n+1} = -t_{N+1}. \tag{39}$$

In the above equation (36) to (39), t_0, t_i, u_i and t_{N+1} be the values of T_0, T_i, U_i and T_{N+1}, respectively. Finally, we get a system of $2N + 2$ equations and $2N + 2$ unknowns, where unknowns are h and V. The system of equations are transformed into the matrix form is as follows

$$\begin{bmatrix} \frac{\partial T_1}{\partial h_1} & \frac{\partial T_1}{\partial V_1} & 0 & . & . & . & . & 0 \\ \frac{\partial T_1}{\partial h_1} & \frac{\partial T_1}{\partial V_1} & \frac{\partial T_1}{\partial h_2} & \frac{\partial T_1}{\partial V_2} & . & . & . & . \\ \frac{\partial U_1}{\partial h_1} & \frac{\partial U_1}{\partial V_1} & \frac{\partial U_1}{\partial h_2} & \frac{\partial U_1}{\partial V_2} & . & . & . & . \\ 0 & 0 & \frac{\partial T_2}{\partial h_2} & \frac{\partial T_2}{\partial V_2} & \frac{\partial T_2}{\partial h_3} & \frac{\partial T_2}{\partial V_3} & . & . \\ 0 & 0 & \frac{\partial U_2}{\partial h_2} & \frac{\partial U_2}{\partial V_2} & \frac{\partial U_2}{\partial h_3} & \frac{\partial U_2}{\partial V_3} & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & \frac{\partial T_{N+1}}{\partial h_{N+1}} & \frac{\partial T_{N+1}}{\partial V_{N+1}} \end{bmatrix} \begin{bmatrix} \Delta h_1 \\ \Delta V_1 \\ \Delta h_2 \\ \Delta V_2 \\ \Delta h_3 \\ \Delta V_3 \\ . \\ . \\ . \\ \Delta h_{N+1} \\ \Delta V_{N+1} \end{bmatrix} = \begin{bmatrix} -t_0 \\ -t_1 \\ -u_1 \\ -t_2 \\ -u_2 \\ . \\ . \\ . \\ . \\ -u_N \\ -t_{N+1} \end{bmatrix}$$

The matrix elimination programs, such as Gaussian elimination or LU decomposition method will be used to solve the above given system. However, Fread [9] used the Sparse characteristics of the Jacobian coefficient matrix with maximum value are considered for consecutive elements and developed an effective algorithm to solve such kind of Ribbon matrix problem. No matter which solution is used to solve the matrix, it is ultimately the case. The step is to calculate the correction value of the unknown of the next iteration. Thus, we obtain the following

$$Q_i^{k+1} = Q_i^k + \delta Q_i,$$
$$h_j^{k+1} = h_j^k + \delta h_j.$$

Numerical Experiment. The above method requires main and boundary conditions. The main condition is a stable flow before the flow change, where the upstream condition can be flooded hydraulic pressure, and the downstream boundary condition can be a displacement relationship. We consider a rectangular open channel with a length of 3 km, a breath of 3 m, having bed slope 0.0005 and a Manning roughness coefficient at the beginning of the channel is 0.25. When the flood entered with the specified water level (in flow water level) with a spillway at its end having height of 1 m, width of 15 m, and coefficient factor 1.6. Then, the water level curve and flow water level on each point will be drawn at any time by the above method. For calculations, MATLAB software has been used.

In Figures 3 and 4, the flow of water graph at 500 m and 1500 m in intervals from the flood beginning point has been drawn with Preissmann scheme. According to these figures, we observed that through the increase of distance from the starting point of flood, difference between results increases.

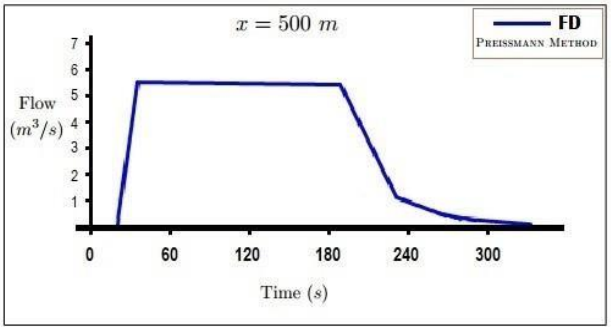


Figure 3. Hydrographs computed at the 500 m distance from the flood start

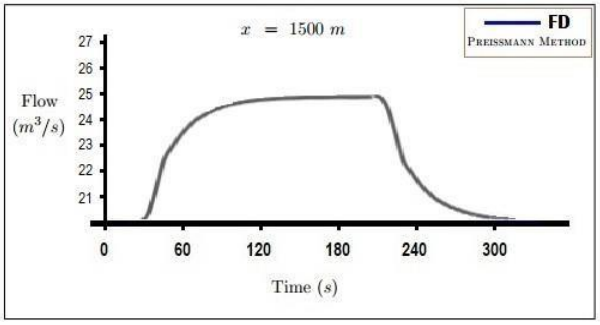


Figure 4. Hydrographs computed at the 1500 m distance from the flood start

4. Conclusion.

We study the control of open channel water flow under flood routing conditions by employing the Saint Venant equations. The Saint Venant equations converted the structure of flow in the given channel into nonlinear partial differential equations which can be solve numerically by using numerical methods such as finite difference method and Newton Raphson method. By using this method, we develop iteration scheme to estimate the flow and height of water in the given channel. We also study the estimated zone of Saint-Venant equations for flood routing with over bank unsteady flow in the open channel. Some numerical experiments are also presented.

Acknowledgement. The authors are thankful to the Higher Education Commission for providing research grant to develop this work.

Author's Contribution. All authors contributed equally to this research work of manuscript.

Project. This research work is developed under the Higher Education Commission Project no. 5398 titled "Control of open channel water flow and detecting leaks and dumps".

Conflict of interest. No conflict of interest for publishing this manuscript in IJIST.

REFRENCES

[1] J.J. Stoker, "*Water waves the mathematical theory with applications*" Wiley (Interscience), New York, 1957.

[2] K.O.A. Askari, M. Shayannejad, S. Eslamian, and B. Navabpour, "*Comparison of solutions of saint-venat equations by characteristics and finite difference methods for unsteady flow analysis in open channel*" International Journal of Hydrology Science and Technology, 2018, Vol. 8, No. 3, 229-243.

[3] J.A. Baker, "*Finite-Element computational fluid dynamics*" McGraw-Hill, New York, NY 1983.

[4] S.K. Ooi and E. Weyer, "*Control design for an irrigation channel from physical data*" Control Engineering Practice, 2008, 16 (9), 1132-1150.

[5] M.H. Chaudhry, "*Open-channel flow*" Springer Verlag, 2008.

[6] X.Y. Chen, K.W. Chau, and A.O. Busari, "*A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model*" Engineering Applications of Artificial Intelligence, 2015, Vol. 46, No. A, 258-268.

[7] J.H.D. Vieira, "*Conditions governing the use of approximations for the saint-venant equations for shallow water flow*" Journal of Hydrology, 1983, vol 60, pp: 43-58.

[8] J.C.I. Dooge and B. M. Harley, "*Linear routing in uniform channel*" Colorado state university, 1967.

[9] M.G. Ferrick, "*Analysis of river wave types*" Water Resource. Res., 1985, 21, 209-212.

[10] D.L. Fread, "*Channel routing. In: Hydrological Forecasting*", M. G. Anderson and T. P. Burt (Ed.), Wiley, New York, 1985.

[11] J.W. Thiong'o, "*Investigations of fluid flows in open rectangular and triangular channels*" Doctoral dissertation, 2013.

[12] R. Moussa and C. Bocquillon, "*Criteria for the choice of food-routing methods in natural channels*" Journal of Hydrology, 1996, 186, 1-30.

[13] C.L. Wu, K.W. Chau, and Y.S. Li, "*Methods to improve neural network performance in daily flows prediction*" Journal of Hydrology, 2009, Vol. 372, No. 1-4, 80-93.

[14] T. Nazir, "*The control of various channels water flows*" IEEE Xplore, 2018 International Conference on Frontiers of Information Technology, DOI: 10.1109/FIT.2018.00066, 2018, 339-344.

[15] Q.K. Nguyen and H. Kawano, "*Simultaneous solution for flood routing in channel networks*" Journal of Hydrology Eng. ASCE., 1995, Vol. 121, No. 10, 744-750.
[16] C.M. Kazezylmaz-Alhan, M. A. Miguel, and P. Rao, "*On numerical modeling of overland flow*" Applied Mathematics and Computation, 2005, Vol. 166, No. 3, 724-740.
[17] J. M. Kohne, T. Wohling, V. Pot, P. Benoit, S. Leguedois, Y. Bissonnais, and J. Simunek, "*Coupled simulation of surface runoff and soil water flow using multi-objectives parameter estimation*" Journal of Hydrology, 2011, Vol. 403, No. 1-2, 141-156.
[18] A. Das, "*Reverse stream flow routing by using Muskingum models*" Sadhana, 2009, Vol. 34, No. 3, 483-499.
[19] B.A. Sulistyono and L.H. Wiryanto, "*Investigation of flood routing by a dynamic wave model in trapezoidal channels*" In AIP Conference Proceedings, 2017, Vol. 1867, No. 1.
[20] R. Taormina and K.W. Chau, "*Data-driven input variable selection for rainfall-runoff modeling using binary coded particles swarm optimization and extreme learning machines*" Journal of Hydrology, 2015, Vol. 529, No. 3, 1617-1632.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Towards Skin Cancer Classification Using Machine Learning and Deep Learning Algorithms: A Comparison

Iqra Kiran[1], Muhammad Zohaib Siddique [2], Ateeq Ur Rehman Butt[1], Dr. Azhar Imran Mudassir[3], Muhammad Azeem Qadir[1], Sundus Munir[4]
[1] Department of Computer Science (National Textile University).
[2] Department of Computer Science (Riphah International University).
[3] Department of Creative Technologies (Air University Islamabad).
[4] Department of Computer Science (Lahore Garrison University).
* Correspondence: Name and Email ID of corresponding author.
Citation | Kiran. I, Siddique. Z, Butt. A. R, Mudassir. A. I, Qadir. M. A and Munir.S, "*Towards Skin Cancer Classification Using Machine Learning and Deep Learning Algorithms: A Comparison*". International Journal of Innovation in Science and Technology, Vol 3, Special Issue, pp: 110-118, 2021
Received | Dec 13, 2021; Revised | Dec 20, 2021 Accepted | Dec 21, 2021; Published | Dec 25, 2021.

Skin cancer is an uncontrolled development of abnormal skin cells potentially due to excessive exposure to sun, history of sunburns, less melanin, Precancerous skin lesions, moles, etc. This occur when unrepaired DNA damages the cells of the skin. It is one of the diseases that are viewed on its quick evolution and the most common type of cancer that endangers life. Researchers have implemented several machine learning and deep learning techniques for classification of skin cancer. In this research paper, different cancer categories are classified using significant attributes. We have used International Skin Imaging Collaboration (ISIC) dataset for classification purposes. This dermoscopic attributes dataset includes 1000 images and 10016 instances, seven categories, 5 features and 2 Meta attributes. We implemented K-Nearest Neighbor, Logistic Regression, Convolutional Neural Network, Naïve Bayes, and Decision Tree for classification and compared their performance. In order to implement classification algorithm, we used Orange which is an open-source machine learning, data mining, and data visualization toolkit. The models are evaluated based on matrices that include Accuracy, C. Automation, F1 score, Precision, Recall, and AUC. Furthermore, frequency of features is visualized using graphical method and the ROC analysis is also performed for the classifiers. It is observed that CNN technique provided the highest accuracy of 89% and the mentioned results are the highest results of classification with the state of the art techniques. For future, the improved and recent dataset and ensemble modelling techniques based on deep learning can used to enhance classification results. The research can also be extended for other cancer types using CNN.

Keywords: skin cancer; classification; ISIC; CNN; LR; NB; DT and KNN.

INTRODUCTION

Skin is the principal and greatest organ of the human body. Skin covers the bones, muscles, and all body parts. Even a small problem can lead to major damage. Various types of skin cancers may occur due to some infections. Usually, people visit the consultant when cancer reaches its critical stage and the patient may face trouble in recovering [1]. Skin cancer is an aggressive form of cancer amongst various types of cancers. The number of patients affected by skin cancer has risen to 53% over the last decade. In the United States, 01 in 52 women, and 01 in 32 men were infected with melanoma, and approximately 10 million people have died from melanoma. Latest studies revealed that 98% of the patients who survive early identified with melanoma and only 17% survive when melanoma was left incurable at its initial level [2][3]. Approximately 178,560 new melanoma cases have been reported in the US since 2018, including non-invasive 87,290 and invasive 91,270 cases. Furthermore, melanoma-related deaths have reached up to 9,320, which involves 3330 women and 5990 men [4], [5]. Figure 1. shows the different types of skin lesions.

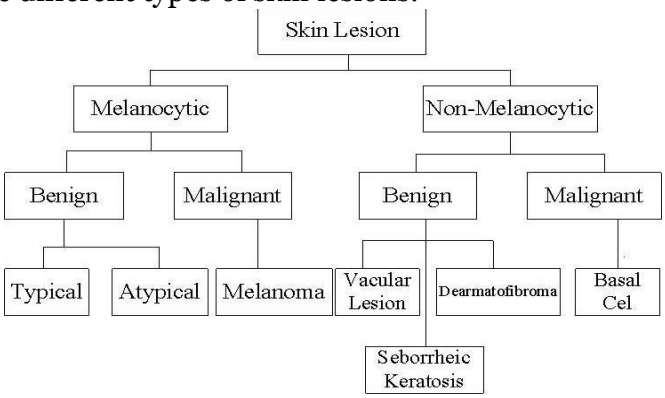


Figure 1. Skin lesion classification tree

Skin cancer is categorized into seven forms. Actinic keratoses (akiec) Is a rough, scaly patch, which is commonly found at face, lips, and back. Basal Cell Carcinoma (BCC) is a category of a non-melanocytic malignant lesion. It is a public type, but the least dangerous form of a tumor. It grows slowly and is most common in areas of skin that is exposed to the sun more often like face. Mostly on the neck, skull, and upper torso. Squamous Cell Carcinoma (SCC) is mostly on dark skin color. It is mostly on the legs and feet. Benign keratosis (bkl) natural, harmless, non-cancerous growth on the skin is seborrheic keratosis. It occurs generally as a red, black, or brown growth on the back, arms, chest, or face, dermatofibroma (df) is a common type of benign skin tumor seen most often on the legs, small, slow-growing, typically firm, red-to-brown bump, Pigmented nevi (moles) are skin lesions that are generally black, brown or skin-colored, melanoma (melanoma), and vascular skin (vascular), is the most commonly diagnosed lesions in skin cancer. The physicians developed various strategies for skin lesion assessment comprising CASH, 7-point checklist, and ABCD rule [6].

ISIC database is the world's biggest publicly available repository for dermoscopic images of skin lesions. In 2018 ISIC conducted an image analysis competition in which participants were to practice a dermoscopic image using the HAM 10,000 dataset to recognize one of seven classes: melanoma, melanocytic NV, bcc, actinic keratosis, bkl, df, and vascular

lesion [7]. This initiative is a follow-up to a similar approach last year, which was held in collaboration with the 2017 International Biomedical Imaging Symposium [8].

Dubal et al. [9] used a Neural Network and ABCD rule has been used to classify the image to a high degree of accuracy. Their techniques were applied as an expert software program, where users provide input in terms of cancer images and identified skin lesions are beginning or malignant. Farooq et al. [10] applied different algorithms on the different datasets and provided comparative analysis. They used SVM and NN classifiers to classify the segmented moles. R. Ashraf et al. presented the use of deep learning in their research work for the classification of skin cancer images in an effective way [11]. Mhaske et al. [12] compared different ML algorithms by using the ph2 dataset. They classified melanoma skin cancer on supervised and unsupervised Machine learning. Murthi et al. suggested a computer-aided melanoma skin cancer identification using ANN. They calculate accuracy on MATLAB and achieved the highest accuracy of 96%. Ramlakhan et al. [13] used the classical machine learning techniques to design a technique to classify benign and malignant lesions. When tests were conducted on 83 images, with an accuracy of 66.7 percent. Some scientists also focus on skin diseases other than skin cancer. AUR Butt et al. presented a computer-aided diagnosis for segmentation and classification of burnt human skin by using different machine learning algorithms by incorporating the comparison between all the used algorithms [14].

Islam et al. [15] concentrated on the development of a portable classification system for pigmented skin lesions. The classification system proposed for skin lesions uses image processing and artificial intelligence to evaluate the texture-based characteristics derived from the image of the disease. The arsenic detection accuracy and recall rates were 88% and 84%. A mobile-based optimization approach for the classification of benign and malignant lesions were implemented by Aleem et al. [16]. Trained and tested on the dataset, the smartphone app contained only 84 images. Research and testing on small dataset, results in 80% sensitivity and 75% specificity. Hekler et al. [17] integrated human intelligence and artificial intelligence to identify skin cancer. Using 11,444 dermoscopic images a specific CNN was trained to classify skin lesion images into five groups. These images have also been classified by dermatologists, and it has been discovered that human and artificial intelligence combines to accomplish superior results. Their proposed methodology achieved 82.95% accuracy. Mobile phones are also utilized in the field to classify skin lesions.

F. A. Khan et al. presented the use of deep convolutional neural networks in their research work for the segmentation and classification of burnt skin images of human beings [18]. One such attempt Ahmed et al. [19] did where 48,373 dermoscopic images were trained for binary classification of skin lesions using the Convolutional Neural Network model MobileNetV2. Using the qualified model, a skin lesion image with an accuracy rate of 91.33% was graded as benign or malignant. After the classification model was trained, an iOS-based mobile app was designed to assess its efficiency on unseen images. Abbas et al. [20] used the standard approach to machine-learning to identify benign and malignant skin lesions. 900 images were included in their proposed approach. Second, segmentation of the image using an edge detection technique to remove the ROI. Texture-based characteristics were extracted and SVM was applied to them to achieve the overall classification mark and 99.02% accuracy was obtained. F. A. Khan et al. presented the use of DCNN for segmentation and depth classification of burnt human skin by claiming that their obtained results are the best and

highest results among the previous results of the state-of-the-art techniques of related works [21].

Deep learning methods are also being used for the classification of skin lesions by using pre-trained learning models. Mahbod et al. [22] have suggested a fully automated classification scheme. Features were produced using AlexNet [23], VGG16 [24], and ResNet-18 [25] in the implemented classification scheme, and passed for final prediction to the SVM classifier. The experimental classification scheme was tested on 150 images and the melanoma and seborrheic keratosis yielded a region under the curve of 83.83 % and 97.55% respectively. In this article, the flow is the programmer's design section in which the system's mathematical model that defines the console's input and the output state is clarified. The classification results obtained by these algorithms and their precision are also discussed. Finally, the results are discussed in terms of accuracy for various kinds of classifiers.

In this work, we experimented with the ISIC 2018 dataset. The classification is performed using Convolutional Neural Network, Logistic Regression, K-Nearest Neighbor, Naïve Bayes, and Decision Tree.

MATERIAL AND METHODS

Dataset site

HAM 1000 [8] (Human against a machine with 10000 images) released by International Skin Imaging Collaboration (ISIC) including 10016 instances. The dataset is a multi-classification with seven different labels. The main objective of using a dataset is to classify the different categories of skin cancer. It is publicly available for academics used to perform machine learning processes. The dataset consists of dermoscopic attributes. 40% dataset are considered as training and 60% are considered as testing. The dataset consists of 5 features and 2 Meta attributes. The description of data is defined in Table 2.

Table 2. Dataset description.

Dx	**Akies, bcc, bkl, df, mel, nv, vasc**
dx-type	Confocal, consensus, follow-up, Histo
Sex	Female, male, unknown
Age	25-80
Localization	Abdomen, acral, back, chest, ear, face, foot, genital, hand, lower extremity, neck, scalp, trunk, unknown, upper extremity

Decision Tree

Decision Tree (DT) is supervised learning among which one of the dissimilar methods is built for classification. It uses inductive reasoning to generate a tree structure in which each node indicates an attribute while the node's each descending branch represents one of the possible outcomes for that attribute. Each node of the tree is the distinguishing equation when classifying all data. It is a common method that provides both classification and predictive function at the same time [15]. We applied DT for the classification of skin cancer by setting criterion gain ratio, highest depth of the tree measured as 100. The minimal instance in leaf

size = 2, and did not split the minimal size that is less than 5. DT showed an overall accuracy of 86%, precision 67%, and recall 70%.

Naive Bayes

Naive Bayes (NB) is essentially developed on the basis of base theorem that is used to test the theory of probability. It utilizes the theory of probability for data classification. The algorithm provides accuracy, precision, and recall rate were 87%, 66%, and 69% respectively.

Logistic Regression

Logistic Regression (LR) is a classification algorithm for Machine Learning employed to predict the probability of categorical dependent variables. We apply LR for skin cancer classification. We used ridge regularization (L2) and set the probability as 1. We achieved 88% overall accuracy by using logistic regression and 62% precision and 70% recall.

Convolutional Neural Network

A convolutional neural network is a kind of machine learning which consists of many neural network layers. Two different types of convolutional and pooling layer. The last stage is typically made by a fully connected layer. In our proposed method ReLU is an activation function and the learning rate is set 0.001. The maximum number of iterations was 200 and the number of hidden layer neurons was 100. CNN achieved the highest accuracy which is 89% and is best among all.

K-Nearest Neighbor

K-Nearest Neighbor is also one of the supervised learning methods being used for classification problems. KNN selects data on the base of the k value of the nearest neighbor then decides the relevance with the given points. We apply K-Nearest Neighbor with k value 5 and 40% data split for training and 60% for testing then made a 2-fold of cross-validation. By using KNN we achieved a precision of 67%, and recall 66%, whereas we achieved an accuracy of 81% by using the KNN algorithm.

RESULT

The proposed work has been evaluated using core i3 with 4 GB of RAM and developed using an open-source machine learning, data visualization, and data mining toolkit known as Orange (Version: 3-3.26.0). ISIC dataset was used for classification which contains 10016 instances. Seven categories of skin cancer are used for classification. We have used multiple algorithms in our work for determining the cancer type classification, which is represented in Table 3.

Table 3. An average accuracy of all model

Model	Accuracy	C. Automation	F1	Precision	Recall
KNN	81%	0.68	0.67	0.67	0.68
Decision Tree	86%	0.70	0.68	0.67	0.70
Neural Network	89%	0.70	0.67	0.65	0.70
Naïve Bayes	87%	0.69	0.67	0.66	0.69
Logistic Regression	88%	0.70	0.65	0.62	0.70

The graphical representation of performance of algorithms is presented below in Figure 2.

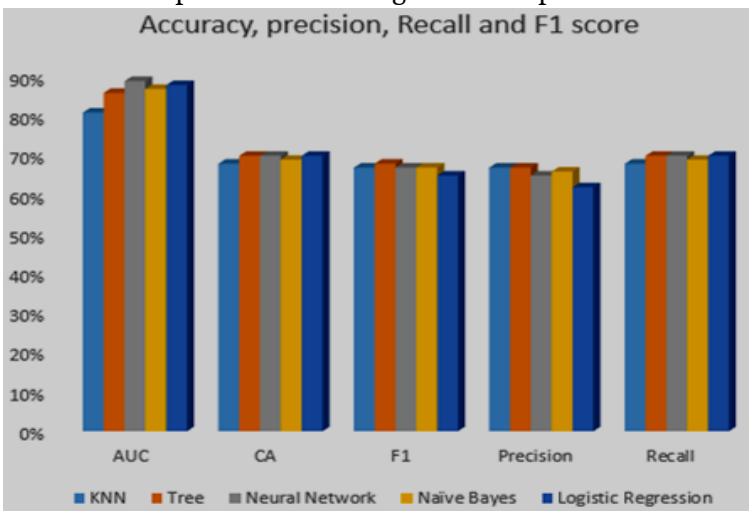


Figure 2. An average accuracy of all models

Figure 3. shows ROC analysis for the algorithm implemented in this paper.

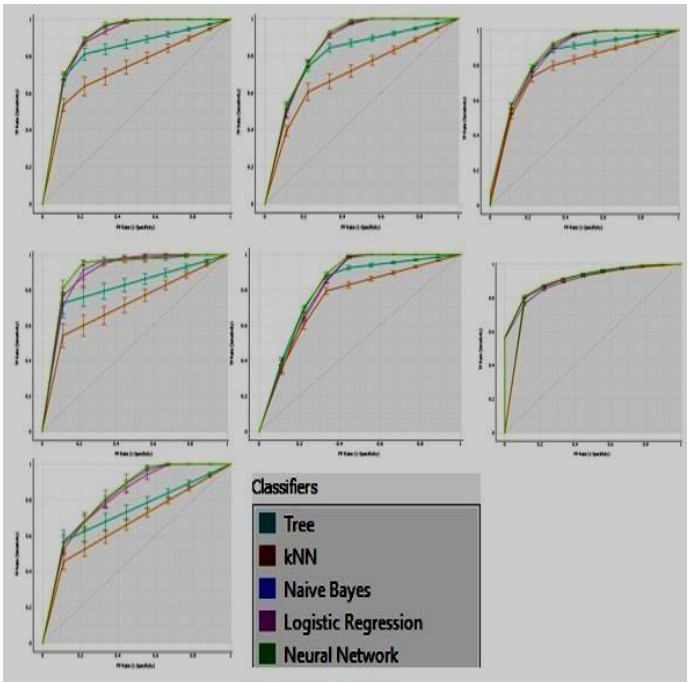


Figure 3. ROC Analysis

Figure 4 shows the visualizing frequencies of sex vs dx (a) and age vs dx (b).

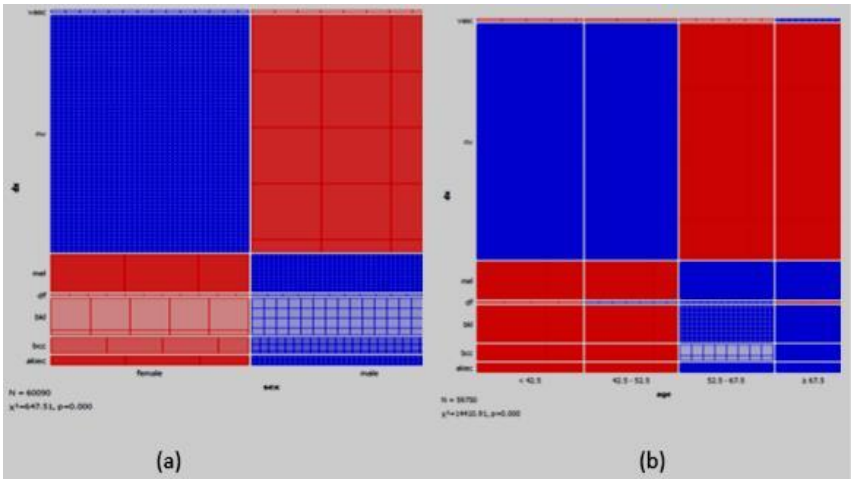


Figure 4. Graphical method for visualizing frequencies.

DISCUSSION

We have observed that CNN offers the highest accuracy of 89% and has recall rate of 0.70 which is similar decision tree and logistic regression. Meanwhile, all the other algorithms offer comparable accuracy. Logistic regression has the second highest accuracy of 88% while KNN provides the lowest accuracy for the classification, 81%. Moreover, CNN offers precession of 0.65 which is as average value when compared to the precision value of other algorithms. Decision tree and KNN provides the highest precision of 0.67. Furthermore, Decision tree, neural network and logistic regression has highest recall of 0.70. The F1 score is highest for decision tree which is perhaps due to the highest values attained by this classifier for recall and precision among other classifiers.

The specificity and sensitivity of all classifier based on seven categories of skin cancer can be noted from the graphs in Fig 3.2. The X-axis shows that FP (specificity) and Y-axis shows TP (sensitivity). The visualizing frequencies of sex vs dx (a) and age vs dx (b) can be observed from Figure 3.3. In (b) age contain < 42 to > = 67. Nv affects a patient of male and female as compared to other categories.

CONCLUSION

This research article discusses the classification technique for skin cancer. The investigation was conducted on the Intel Core i3 CPU having 4GB of RAM. Orange v3-3.26.0 has been used to examine & train the classification model. Few existing classification methods for the medical diagnosis of cancer patients have been discussed on basis of accuracy. Five machine learning technique was applied to the ISIC dataset. The results showed that CNN outperforms other models. We achieved 89% accuracy to classify 7 categories of skin cancer. In future, recent and improved dataset can be used for achieving even better accuracy. Furthermore, ensemble models of deep leaning algorithms can also be used to enhance the performance of classification model. In addition to this, the proven classification algorithms can be used for detection of less common skin cancers like Kaposi sarcoma, Merkel cell carcinoma, Sebaceous gland carcinoma.

Acknowledgement. Acknowledgements are considered necessary.

Author's Contribution. Corresponding author should explain the contribution of each co-author completely.

Conflict of interest. Authors are advised to explain that there exists no conflict of interest for publishing this manuscript in IJIST.

Project details. If this research was conducted as a result of a project, please give details like project number, project cost and completion date etc....

REFRENCES

[1] U. Jamil, A. Sajid, M. Hussain, O. Aldabbas, A. Alam, and M. U. Shafiq, "*Melanoma segmentation using bio-medical image analysis for smarter mobile healthcare,*" J. Ambient Intell. Humaniz. Comput., 2019, vol. 10, no. 10, pp. 4099–4120.

[2] J. Burdick, O. Marques, J. Weinthal, and B. Furht, "*Rethinking Skin Lesion Segmentation in a Convolutional Classifier*," J. Digit. Imaging, vol. 31, 2018, no. 4, pp. 435–440.

[3] K. D. Miller et al., "*Cancer treatment and survivorship statistics, 2016*," CA. Cancer J. Clin.,2016, vol. 66, no. 4, pp. 271–289.

[4] T. Akram, M. A. Khan, M. Sharif, and M. Yasmin, "*Skin lesion segmentation and recognition using multichannel saliency estimation and M-SVM on selected serially fused features*," J. Ambient Intell. Humaniz. Comput., 2018.

[5] E. Nasr-Esfahani et al., "*Melanoma detection by analysis of clinical images using convolutional neural network*," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, 2016, vol. 2016-Octob, pp. 1373–1376.

[6] M. A. Khan, M. Y. Javed, M. Sharif, T. Saba, and A. Rehman, "*Multi-model deep neural network-based features extraction and optimal selection approach for skin lesion classification,*" 2019 Int. Conf. Comput. Inf. Sci. ICCIS 2019, 2019, pp. 1–7.

[7] M. Properties and H. O. Mucosa, "*Technical Report*", 2015, pp. 2–3.

[8] P. Tschandl, C. Rosendahl, and H. Kittler, "*Data Descriptor: The HAM 10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions*," Nat. Publ. Gr., 2018, vol. 5, pp. 1–9.

[9] P. Dubai, S. Bhatt, C. Joglekar, and S. Patii, "*Skin cancer detection and classification*," Proc. 2017 6th Int. Conf. Electr. Eng. Informatics Sustain. Soc. Through Digit. Innov. ICEEI 2017, 2018, vol. 2017-Novem, pp. 1–6.

[10]M. A. Farooq, M. A. M. Azhar, and R. H. Raza, "*Automatic Lesion Detection System (ALDS) for Skin Cancer Classification Using SVM and Neural Classifiers,"* Proc. - 2016 IEEE 16th Int. Conf. Bioinforma. Bioeng. BIBE 2016, 2016, pp. 301–308.

[11]R. Ashraf, I. Kiran, T. Mahmood, A. Ur Rehman Butt, N. Razzaq, and Z. Farooq, "*An efficient technique for skin cancer classification using deep learning*," Proc. - 2020 23rd IEEE Int. Multi-Topic Conf. INMIC 2020, 2020.

[12]H. R. Mhaske and D. A. Phalke, "*Melanoma skin cancer detection and classification based on supervised and unsupervised learning*," 2013 Int. Conf. Circuits, Control. Commun. CCUBE 2013, pp. 1–5, 2013.

[13]K. Ramlakhan and Y. Shang, "*A mobile automated skin lesion classification system*," Proc. - Int. Conf. Tools with Artif. Intell. ICTAI, 2011, pp. 138–141.

[14]A. U. Rehman Butt, W. Ahmad, R. Ashraf, M. Asif, and S. A. Cheema, "*Computer Aided Diagnosis (CAD) for Segmentation and Classification of Burnt Human skin*," 1st Int. Conf. Electr. Commun. Comput. Eng. ICECCE 2019, 2019, no. July, pp. 24–25.

[15]C. L. Chang and C. H. Chen, *"Applying decision tree and neural network to increase quality of dermatologic diagnosis*," Expert Syst. Appl., 2009, vol. 36, no. 2 PART 2, pp. 4035–4041.

[16]R. Z. B and E. Conchon, "*eHealth 360°,"* , 2017, vol. 181, pp. 407–418.

[17]A. Hekler et al., *"Superior skin cancer classification by the combination of human and artificial intelligence,"* Eur. J. Cancer, 2019, vol. 120, pp. 114–121.
[18]F. A. Khan et al., *"Computer-aided diagnosis for burnt skin images using deep convolutional neural network,"* Multimed. Tools Appl., 2020, vol. 79, no. 45–46, pp. 34545–34568.
[19]A. Ech-Cherif, M. Misbhauddin, and M. Ech-Cherif, *"Deep Neural Network Based Mobile Dermoscopy Application for Triaging Skin Cancer Detection*," 2nd Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2019, 2019, pp. 1–6.
[20]Z. Abbas, M. U. Rehman, S. Najam, and S. M. Danish Rizvi, *"An Efficient Gray-Level Co-Occurrence Matrix (GLCM) based Approach Towards Classification of Skin Lesion,"* Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019, 2019, pp. 317–320.
[21]F. A. Khan, A. U. Rehman Butt, M. Asif, H. Aljuaid, A. Adnan, S. Shaheen, *"Burnt Human Skin Segmentation and Depth Classification Using Deep Convolutional Neural Network (DCNN),"* J. of Medical Imaging and Health Informatics, 2020, vol. 10, no. 10, pp. 2421-2429.
[22] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinger, *"Skin Lesion Classification Using Hybrid Deep Neural Networks",* Institute for Pathophysiology and Allergy Research, Medical University of Vienna, Austria Department of Research and Development, TissueGnostics GmbH , Austria Department of Computer Science , Loughborough University , U . K . Department of Biomedical, 2019, pp. 1229–1233.
[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "*ImageNet classification with deep convolutional neural networks,"* Commun. ACM, 2017, vol. 60, no. 6, pp. 84–90.
[24] F. O. R. L. Arge and C. I. Mage, *"V d c n l -s i r,"* , 2015, pp. 1–14.
[25] V. Sangeetha and K. J. R. Prasad, *"Syntheses of novel derivatives of 2-acetylfuro[2,3-a]carbazoles, benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-acetyloxycarbazole-2- carbaldehydes*," Indian J. Chem. - Sect. B Org. Med. Chem., 2006, 1954, vol. 45, no. 8, pp. 1951.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Multirate Adaptive Equalization

Muhammad Yasir Siddique Anjum [1], Muhammad Ali Raza Anjum [2], Usman Riaz [3]
[1] National University of Modern Languages, Pakistan.
[2]Army Public College of Management and Sciences, Pakistan.
[3] Centre for Advanced Studies in Engineering, Pakistan.
* Correspondence: Muhammad Yasir Siddique Anjum <yasir.siddique@numl.edu.pk>.
Citation | Anjum. M.Y.S, Anjum. M.A.R, Riaz. U, "Multirate adaptive Equalization", International Journal of Innovations in Science and Technology, Vol 3, Special Issue, pp: 119-125, 2021.
Received | Dec 13, 2021; Revised | Dec 18, 2021 Accepted | Dec 19, 2021; Published |Jan 1, 2022.

Finite Impulse Response (FIR) filter model emulates the Inter Symbol Interference (ISI) in a wireless communication channel. An equalizer, typically an Infinite Impulse Response (IIR) filter, behaves as an inverse filter to the FIR filter to remove the effects of the ISI. IIR filters are generally avoided due to tractability issues, and an FIR filter, with an adaptive signal processing algorithm to minimize the error due to the ISI, is deployed at the receiver. However, the filter is observed to quickly reach a steady state where further iterations do not yield a reduction in the error. This can be attributed to relatively slow variations in the steady state error which prevent further reduction of the errors. This work focuses on converting the low frequency error variations to high frequency variations by the use of multirate signal processing. As such, the steady state error can be damped as well, providing further reduction in the error and an enhanced adaptive filter performance.
Keywords: adaptive; filter; ISI; multirate; equalization.

INTRODUCTION

It has been well established that in order to remove the effects of ISI, an equalizer has to be employed at the wireless communication receiver [1]. Two choices are available in this regard: an FIR filter and an IIR filter. FIR filters are preferred over IIR filters due to their simplicity and ease of implementation. However, FIR filters are unable to sufficiently minimize the error due to ISI, and a certain allowance has to be made for the magnitude of error in order to deploy and reap the benefits of FIR filters at the wireless communication receiver.

A well-known criterion for making such an allowance is the Minimum Mean Square Error (MMSE) criteria [2]. Many adaptive signal processing algorithms are available that try to reduce the error based on the MMSE criteria. Least Mean Square Algorithm (LMS), Normalized LMS algorithm (NLMS), and Recursive Least Squares algorithm (RLS) are popular in this regard [3,11]. These algorithms are distinguished by their implementation

complexity and convergence speed, but an exclusive focus on the error analysis is not always to be found.

In iterative solvers, the error is observed to settle into a steady state after rapid initial convergence, where further iterations do not yield significant reduction in MSE [12]. This can be attributed to the ill-conditioning of the input covariance matrix, arising due to the disparity in the magnitude of its eigenvalues. Error associated with small magnitude eigenvalues dampens quickly whereas the error associated with large magnitude eigenvalues tends to linger on. Further observed is the frequency of the eigenvectors associated with large magnitude eigenvalues, which is relatively lower compared to the frequency of eigenvectors associated with small magnitude eigenvalues. These two factors, i.e., relatively large magnitude of an eigenvalue and the low frequency of its associated eigenvector, causes the MSE to settle into the steady state early, where in further reduction in MSE with increased number of iterations is not possible.

In this work, we proposed that the magnitude of larger eigenvalues can be reduced and the low frequencies of their associated eigenvectors can be converted to high frequencies by down sampling the error vector, wgicg will cause the error associated with large magnitude eigenvalues to dampen quickly as well. In this way, MSE can be reduced further and the filter convergence can be enhanced. This is the premise behind the presented work.

EQUALIZATION IN WIRELESS COMMUNICATION SYSTEMS

The equalization problem. Basic model of the equalization problem in a wireless communication system is depicted in Figure 1 [2].

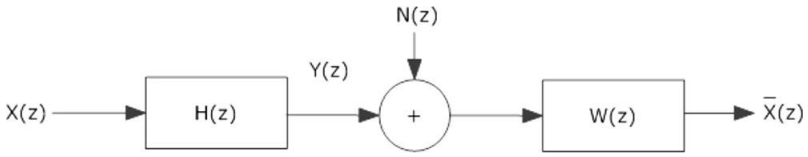


Figure 1. Model of equalization problem.

Output $Y\ (z)$ of a multipath channel $H(z)$ in the presence of Additive White Gaussian Noise (AWGN) $N(z)$ of zero mean and σ^2 variance can be expressed as:

$$Y\ (z) = H(z)X(z) + N(z) \tag{1}$$

With the corresponding MMSE equalizer $W\ (z)$ output being:

$$W(z) = \frac{\phi_{xx}(z)H(z^{-1})}{\phi_{xx}(z)\phi_{hh}(z) + \phi_{nn}(z)} \tag{2}$$

Such that $\phi_{xx}(z)$, $\phi_{hh}(z)$, and $\phi_{nn}(z)$ represent the power spectral densities of the input, channel, and noise respectively. MMSE equalizer endeavors retrieve $X(z)$ by minimizing the error between $X(z)$ and the equalizer output $\bar{X}(z)$.

Adaptive solution to equalization problem. Adaptive solution to the equalization problem is depicted in Figure 2.

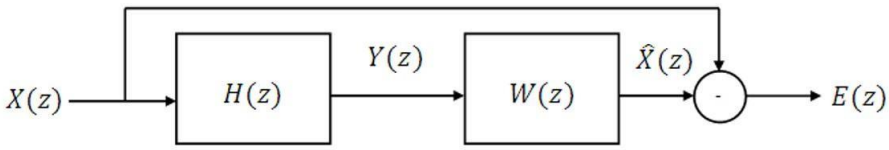


Figure 2. Modified equalization model.

An adaptive Wiener filter provides the ratio of the cross power spectral density $\phi_{xy}(z)$ to the auto power spectral density $\phi_{yy}(z)$ [2]:

$$W(z) = \frac{\phi_{xy}(z)}{\phi_{yy}(z)} \tag{3}$$

Which, in time domain, can be written in the following matrix form:

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{p} \tag{4}$$

(4) Is known as the Wiener-Hopf equation [2]. $\mathbf{w}$ Is the so-called Wiener filter and represents the desired response of the adaptive equalizer. $\mathbf{R} = E\{\mathbf{y}\mathbf{y}^{\mathbf{H}}\}$ Is the filer input autocorrelation matrix, and $\mathbf{p} = E\{\mathbf{x}\mathbf{y}^{\mathbf{H}}\}$ the input-output cross correlation vector.

ANALYSIS OF ERROR

Error controlling matrix. (4) Can be rewritten as:

$$\mathbf{M}\mathbf{w} = \mathbf{p} - \mathbf{R}\mathbf{w} + \mathbf{M}\mathbf{w} \tag{5}$$

Where $\mathbf{M}$ represents the error controlling matrix. (5) can be rearranged as:

$$\mathbf{w} = (\mathbf{I} - \mathbf{M}^{-1}\mathbf{R})\mathbf{w} + \mathbf{M}^{-1}\mathbf{p} \tag{6}$$

$\mathbf{I}$ Represents the identity matrix. (6) can be iteratively solved as [2]:

$$\mathbf{w}[\mathbf{n}+\mathbf{1}] = (\mathbf{I} - \mathbf{M}^{-1}\mathbf{R})\mathbf{w}[\mathbf{n}] + \mathbf{M}^{-1}\mathbf{p} \tag{7}$$

Or:

$$\mathbf{e}[\mathbf{n}+\mathbf{1}] = (\mathbf{I} - \mathbf{M}^{-1}\mathbf{R})\mathbf{e}[\mathbf{n}] \tag{8}$$

Such that $\mathbf{e}[\mathbf{n}] = \mathbf{w}[\mathbf{n}] - \mathbf{w}$, with $\mathbf{w}$ representing the exact solution to (4). (8) Shows that the convergence of error depends on the eigenvalues of the error controlling matrix $(\mathbf{I} - \mathbf{M}^{-1}\mathbf{R})$.

Computation of eigenvalues. Given that the equalizer $W(z)$ employed in Figure 2 is a two-tap filter, with the correlation factorα, then $\mathbf{R}$ in (4) can be represented as [2]:

$$\begin{bmatrix} 1+\alpha^2 & -\alpha & 0 & 0 \\ -\alpha & 1+\alpha^2 & -\alpha & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -\alpha & 1+\alpha^2 \end{bmatrix} \tag{9}$$

With$\alpha = 1$, $\mathbf{R}$ can be viewed a second difference matrix with Dirichlet boundary conditions [12]. It has eigenvalues of the form:

$$\lambda^R_j = 2 - 2\cos\left(\frac{j\pi}{N+1}\right) \tag{10}$$

Where$j = 1, \ldots, N$, such that N represents the size and λ^R_j the j–the eigenvalue of the $\mathbf{R}$ matrix. Setting $\mathbf{M} = \mathbf{I}$ in the error controlling matrix $\mathbf{D} = \mathbf{I} - \mathbf{M}^{-1}\mathbf{R}$ leads to the following expression for the eigenvalues:

$$\lambda^D_j = 2\cos\left(\frac{j\pi}{N+1}\right) - 1 \tag{11}$$

Maximum value of $\lambda^D{}_j$ will be one when the cosine is zero, which is not satisfactory. In order for the error to converge, all the eigenvalues must be less than one. Selecting $\mathbf{M}^{-1} = \mathbf{I}/2$ causes all the eigenvalues of **D** to be less than one.

$$\lambda^D{}_j = \cos\left(\frac{j\pi}{N+1}\right) - 1 \tag{12}$$

First four eigenvalues of **D** computed from (12) are: $\cos(\pi/5) = 0.8090, \cos(2\pi/5) = 0.3090, \cos(3\pi/5) = -\cos(2\pi/5) = 0.3090, \cos(4\pi/5) = -\cos(\pi/5) = 0.8090$. Note that the lower frequency $\cos(\pi/5)$ has larger eigenvalue magnitude (0.8090) compared to the eigenvalue magnitude (0.3090) associated with the higher frequency$cos(2\pi/5)$. This will cause the higher frequency to dampen faster in iteration process. The lower frequency will dampen slowly and tend to linger on. If lower frequency is converted to a higher frequency, by the process of down sampling, it can be made to dampen faster like the higher frequency, and the convergence and the accuracy can be improved. The following general formula was derived for the eigenvalues of the error controlling matrix with an arbitrary value α:

$$\lambda^D{}_j = \alpha\cos\left(\frac{j\pi}{N+1}\right) - \frac{(\alpha^2 - 1)}{2} \tag{13}$$

An eigen distribution of **D** computed from (15) for $N = 10$ versus α is displayed in Figure 3 for illustration.

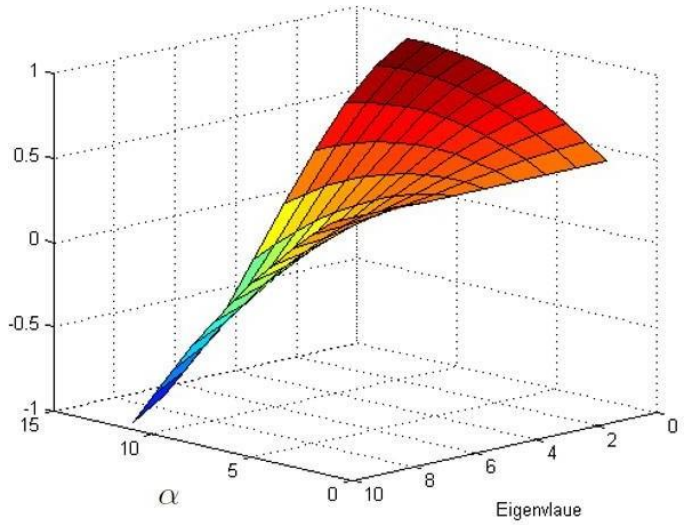


Figure 3. Eigen distribution of the error controlling matrix **D** for$N = 10$.

Computation of eigenvectors. The following general expression was derived for the eigenvectors of error controlling matrix **D** with an arbitrary valueα:

$$v_j[n] = \sin\frac{j\pi}{N+1}n \tag{14}$$

Eigenvector plot of **D** $(N = 10)$ versus different values of α is displayed in Figure 4. Note that the large eigenvalues are associated with low frequency eigenvectors, and vice versa.

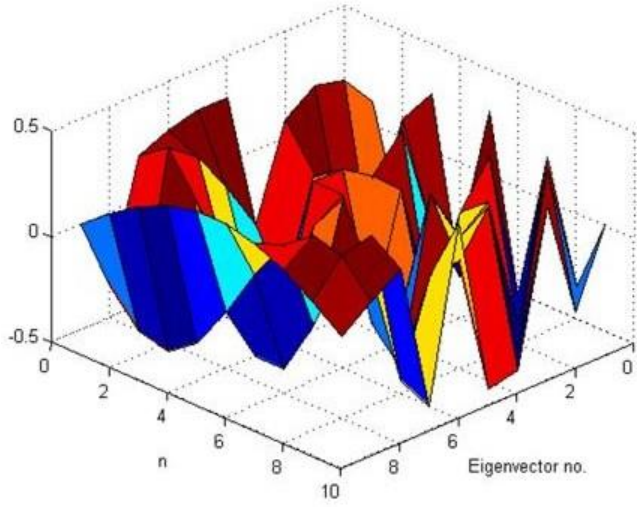


Figure 4. Eigenvectors of error controlling matrix **D** with $N = 10$ versus different values ofα.

DOWNSAMPLING AND ANALYSIS OF DOWNSAMPLED ERROR

Downsampling of eigenvalues. Now the error in (8) is downsampled by a factor of two, and the downsampled eigenvalues are compared with those of the original system($N = 10, \alpha = 0.5$). Downsampling is performed according to the following expression:

$$\lambda^{D}_{j} = \alpha\cos\left(\frac{j\pi}{N/2+1}\right) - \frac{(\alpha^2 - 1)}{2} \qquad (15)$$

Downsampling will reduce the eigenvalues to five in contrast to the original ten. A graphical comparison of the original and the down sampled eigenvalues is displayed in Figure 5, which shows that the magnitude of larger eigenvalues is considerably reduced after down sampling.

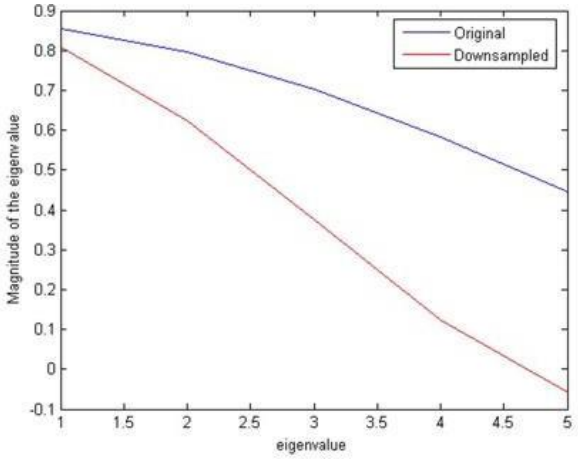


Figure 5. Graphical comparison of the downsampled and original eigenvalues of the error controlling matrix for$N = 10, \alpha = 0.5$, and downsampling factor of2.

Analysis of downsampled error. Re-writing (8) as:

$$\boldsymbol{e[n+1] = De[n]} \tag{16}$$

With $\mathbf{D = I - M^{-1}R}$. Alternatively for (16):

$$\boldsymbol{e[n] = D^k e[0]} \tag{17}$$

For $k = 0$:

$$\boldsymbol{e[0] = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n} \tag{18}$$

Multiplying (18) by $\mathbf{M}$:

$$\mathbf{Me[0] = c_1 Mv_1 + c_2 Mv_2 + \cdots + c_n Mv_n} \tag{19}$$

Leads to:

$$\mathbf{Me[0] = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_n \lambda_n v_n} \tag{20}$$

Finally:

$$\mathbf{M^k e[0] = c_1 \lambda^k{}_1 v_1 + c_2 \lambda^k{}_2 v_2 + \cdots + c_n \lambda^k{}_n v_n} \tag{21}$$

Equation (21) shows that, for the error to dampen quickly, magnitude of the eigenvalues should be as small as possible; ideally, they must be zero. However, due to the poor conditioning of the input covariance matrix, a disparity can be found in the magnitude of eigenvalues. Large magnitude eigenvalues cause the error to enter into a steady state and prevent further reduction of MSE. The steady state error can be reduced by downsampling the error vector, which makes the associated eigenvalues to dampen quickly as well when their higher powers are taken in to account according to (21). A plot of the MSE achieved from the original and the downsampled error sequence is displayed in Figure 6. The plot confirms that the downsampled error reduces much faster than the original one.

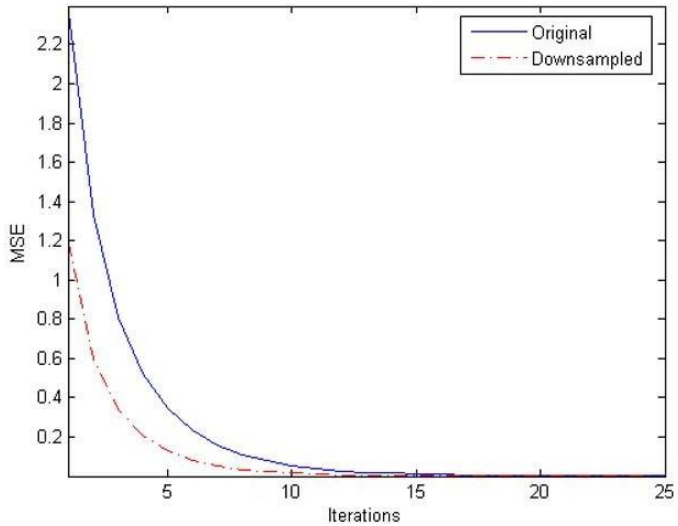


Figure 6. Comparison of the MSE achieved from the original and the downsampled error sequence.

CONCLUSION

An improvement in the error performance of adaptive wireless channel equalizer was found by downsampling the error vector by a factor of two. The improvement was demonstrated analytically and numerically. Further possibilities for improvement in error

performance of adaptive equalizer can be explored by increasing the downsampling factor beyond two.

Acknowledgement. No funding to declare.

Author's Contribution. Muhammad Yasir Siddique Anjum wrote the first draft of the manuscript, and conducted the data analysis. Muhammad Ali Raza Anjum provided technical expertise of multirate signal processing, and helped edit the manuscript. Usman Riaz supervised the study, provided factual review, and helped edit the manuscript.

Conflict of interest. There exists no conflict of interest for publishing this manuscript.

Project details. The research was not conducted as a part or result of a project.

REFRENCES

1. T. S. Rappaport et al., "*Wireless communications: principles and practice*". Prentice hall PTR New Jersey, 1996, vol. 2.
2. B. Farhang-Boroujeny, "*Adaptive filters: theory and applications*" John Wiley & Sons, 2013.
3. M. Umer, "*Adaptive lms based channel equalization*," International Journal of Technology and Research, vol. 2, no. 4, p. 111, 2014.
4. R. Martinek and J. Zˇ´ıdek, "*The real implementation of nlms channel equalizer into the system of software defined radio*," 2012.
5. S. K. Sahoo and M. N. Mohanty, "*Effect of ber performance in rls adaptive equalizer," International Journal of Advanced Computer Research (IJACR)"*, vol. 2, no. 4, 2012.
6. R. A. Fayadh, F. Malek, H. A. Fadhil, N. A. Al-Shareefi, H. Saad et al., "*Mmse equalized rake-receiver using adaptive filter with a family of partial update algorithms in wireless communication systems*," International Journal of Engineering and Technology (IJET), vol. 5, no. 6, pp. 5169–5177, 2014.
7. R. Candido, M. Eisencraft, and M. T. Silva, "*Channel equalization for synchronization of chaotic maps," Digital Signal Processing*, vol. 33, pp. 42–49, 2014.
8. R. N. Devi, T. Saikumar, and K. K. Rao, "*Nlms algorithm based cma channel equalization through an adaptive mmse equalizer*," in Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (INDIA 2012) held in Visakhapatnam, India, January 2012. Springer, 2012, pp. 679–688.
9. Research progress of low-altitude wireless channel modeling and equalization techniques.
10. B. Alshehry, A. Odeh, and E. Abdelfattah, "*Adaptive equalization: Lms, rls and cma,*" 2014.
11. J. Ai, G. Yue, X. Cheng, and S. Li, "*Low complexity rls channel estimation for sc-fde in 60 ghz communications,*" in 2012 IEEE 14th International Conference on Communication Technology. IEEE, 2012, pp. 186–191.
12. S. Gilbert, "*Computational science and engineering,*" Massachusetts Institute of Technology, Wellesley-Cambridge Press, Massachusetts, USA, 2007.

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Asphalt Pavement Potholes Localization and Segmentation using Deep RetinaNet and Conditional Random Fields

Rana Ghazanfar Ali [1], Syed M. Adnan[2],Nudrat Nida [3],Wakeel Ahmad [4], Farooq Bilal [5]
[1,2, 4, 5]Department of Computer Science, University of Engineering & Technology Taxila, Pakistan
[3]Air University, Islamabad, Aerospace & Aviation CampusKamra, Pakistan.
* Correspondence: Rana Ghazanfar Ali, ghazanfar.ali@uettexila.edu.pk

Citation | Ali. R. G, Adnan. S. M, Nida. N, Ahmad. W and Bilal. F, "*Asphalt Pavement Potholes Localization and Segmentation Using Deep RetinaNet and Conditional Random Fields*", International Journal of Innovations in Science and Technology, Vol 3, Special Issue, pp; 126-139, 2021.
Received | Dec 15, 2021; Revised | Dec 22, 2021 Accepted | Dec 23, 2021; Published |Jan 01, 2022.

The main aspect of maintaining the roads and highways' durability and long life is to detect potholes and restore them. A huge number of accidents occur on the roads and highways due to the pothole. It also causes financial loss to vehicle owners by damaging the wheel and flat tire. For the strategies of the road management system and ITS (Intelligent Transportation System) service, it is one of the major tasks to quickly and precisely detect the potholes. To solve this problem, we have proposed a deep learning methodology to automatically detect and segment the pothole region within the asphalt pavement images. The detection of the pothole is a challenging task because of the arbitrary shape and complex structure of the pothole. In our proposed methodology, to accurately detect the pothole region, we used RetinaNet that creates the bounding box around the multiple regions. For the segmentation we used Conditional Random Field that segments the detected pothole regions obtained from RetinaNet. There are three steps in our methodology, image preprocessing, Pothole region localization, and Pothole segmentation. Our proposed methodology results show that potholes in the images were correctly localized with the best accuracy of 93.04%. Conditional Random Fields (CRF) also show good results.
Keywords: RetinaNet, Pothole Segmentation, Conditional Random Fields (CRF), CAD tool, Region proposal.

INTRODUCTION

Road Maintenance is essential, especially in developed countries. The highway system of a developed country contains thousands of centerline kilometers of pavement. Such systems consist of asphalt, concrete, or composite pavements ranging in condition, age, and performance. Pavement distress is defined as a disorder of pavement structure that reduces serviceability or leads to a reduction in serviceability. Pavement distresses contain symptoms indicating problems of deterioration like Rutting, Fatigue Cracking, Longitudinal Cracking,

Transverse Cracking, Block Cracking, Patches, and Potholes. Among all pavements, potholes are the most commonly damaging road conditions [1].

A pothole is a shallow or a deep hole in the pavement surface resulting from the loss of pavement surfacing material as shown in the diagram Figure. 1. it may be a Bowl-shaped hole of various sizes on the pavement surface.

Potholes are a nuisance, especially in the developing world, and can often result in vehicle damage or physical harm to the vehicle occupants. Therefore, an automated process is needed for vehicles that would allow them to avoid accidents.

Potholes in road surfaces are mostly caused by water as cracks that allow the water to seep through and mix with asphalt and it slowly creates the cavity underneath the crack.

In a road maintenance management system, the assessment of road surface distresses is one of the important tasks for constructing repair and maintenance strategies. So, there is a dire need of a system for automatic detection.

Figure 1. Road image(s) with Pothole [1]

Over the past years, many road maintenance network programs have been established to check the ongoing performance of the road network, to determine further pavement conditions and check long-term needs, to support investment plans, make the decision, and to point out maintenance and restoration. Like many others, the transportation department of the United States established Long-Term Pavement Performance (LTPP) for data collection, storage, analysis, and product development of road networks [20]. LTPP is an ongoing and active program. In this program assessment of pavement surface is an important aspect that requires reliable quality measurements like cracks, potholes, etc.

Broadly speaking, the assessment process of pavement condition is categorized into three major parts, the first part is data collection. Most of the data collection is done by vehicles that are equipped with video cameras for surface imaging. The second part is to identify and classify the type of distress like potholes, cracks, etc. The third and last part is to assess the distress. Also, Pothole detection techniques are further divided into three main categories [21] as follows:

- **3D Reconstruction:** Detection methods that are based on a 3D reconstruction of the pavement surface rely on 3D point clouds provided by either laser scanners or stereovision algorithms using a pair of video cameras. The 3D laser scanning systems can be further distinguished into time-of-flight scanners that employ reflected laser pulses to directly create 3D point coordinates and hybrid systems that use digital cameras to capture consecutive images of lines projected by infrared lasers.
- **Vibration-based approaches:** In vibration-based accelerometers are used to access pavement conditions. These techniques lack accuracy and reliability and cannot be used at bridge expansion joints.
- **Vision-Based Approaches:** These include the techniques that are based on image processing analysis by extracting and comparing the features. Various vision-based approaches have been studied for automating the detection of potholes. In [1], a supervised

approach for automated potholes is presented. This technique is based on pothole texture extraction and comparison where texture inside the pothole candidate is described and compared with the texture of the surrounding region. This implies the existence of several pothole texture samples through which the system is trained. The pothole is identified based on these training results. The dataset comprised of 120 images including both testing and training and reasonable accuracy was achieved.

Another study [2] aimed to detect a variety of road related objects such as lanes, road signs, and potholes. This paper states that pothole is characterized by a distinctive black color on road and that can be predominant characteristics to detect pothole, but they didn't achieve desirable results and more effective filtering was needed to improve the accuracy of the approach.

Apart from that certain image processing technique was employed [3]. In which the images are collected by pining a camera on the front windscreen of a vehicle. The images were then converted into greyscale. Gaussian filter was applied as preprocessing step followed by canny edge detector to have a binary image. The binary image was dilated several times to remove the unwanted edges closed to the outer boundary. After that contour detection was applied to detect potholes which were then filtered out to discard those, that don't meet the size constraint of potholes. This technique achieved a precision of 81.8% and recall of 74.4%.

In 2013 [4], an unsupervised method was presented which was based on spectral clustering, segmenting image, shape extraction followed by identification on the extraction of potholes based on seed points. Dataset was comprised of 50 images with single and multiple potholes and the technique was able to achieve the desired accuracy.

Gabor filter can also be used [5] for road pavements detection in the time-frequency domain in which filter was convolved with input image being processed and the binary image was obtained by thresholding the real component of the input image. Different orientation of the Gabor filter was applied to obtain binary images and all of them were combined resulting in an output image that contained pavement detected area.

Kanza Azhar et al [6] has presented Histogram of Oriented Gradients (HOG) features for the visual analysis of pavement images. Using HOG features pavement images and labeled as pothole and non-pothole images. And Normalized cut segmentation is used for localizing the pothole region in the pothole labeled images. HOG features and Normalized cut segmentation proved to be robust for the identification of pothole images. Dataset was comprised of 120 images and reasonable accuracy, precision, and recall was obtained.

A pothole detection system has been developed in [7] in which the mean shift algorithm has been used with Gaussian kernels to remove speckle noise. The pothole damage detection technique is based on assuming that any dark strong edge in the image is considered as a pothole edge if it has certain size constraints.

A novel framework has been proposed for segmenting road images in a hierarchical manner that can deal with several road defects. The experiments have shown that the approach has achieved satisfactory results on various road images [9].

For detecting and quantifying the pavement defects a depth sensor is used by Jahanshahi et al. [10]. In which pixels are classified as deep or flat using thresholding. Then, defective regions are described by using the maximum depth of the region.

The rate of creation and deployment of computer vision methods for civil engineering applications has been exponentially increasing. Koch et. al [11] presented a comprehensive survey of the state of the art in computer vision-based defect detection in asphalt civil infrastructure.

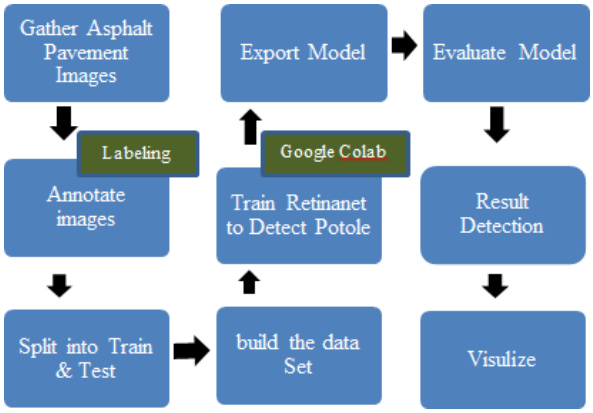


Figure 2. The architecture of Proposed Methodology for pothole detection

In 2019 [22], uses an open hardware device and a prototype to build an IoT-enabled device for potholes detection. In this paper, they have used a Raspberry pi board to build an interactive and context-aware system that can sense the surface in the physical environment. The system for pothole detection using Kinect sensor achieved 84% accuracy.

In 2020 [23], uses a new 1500 image dataset of Indian roads for pothole detection. In this paper firstly the dataset was annotated and tested using YOLO (You Only Look Once), the dataset is also trained and tested on YOLOv3, YOLOv2, YOLOv3-tiny, and the result are compared. The model is tested on different pothole images and it detects with reasonable accuracy. The accuracy achieved by these networks are Yolov3 58.79%, Yolov2 64.05%, Yolov2 68.57%, Tiny Yolov3 72.12.

In 2021[24] Surekha Arjapure presented a deep learning technique Mask Region-Based Convolutional Neural Network to detect and segment potholes and to calculate their area. They used a dataset of 291 images. Dataset has been collected manually on local roads of Mumbai city and nearby highways. An experimental result in this research gives an overall accuracy of 90% for pothole detection.

Retina Net detector when trained and tested on the Koch et al [1] and (Zamir, 2014) dataset using the pre-trained models classify potholes with high precision and recall rate than all existing state of the art detectors. We have used the Conditional Random Field (CRF) semantic method for segmentation which is trained on the deeplabv3 model using the backbone pre-trained ResNet model. This method extracts the affected patches of asphalt surface more precisely with variable boundaries to recognize the pothole. In the previous works in this area of research, the potholecharacteristic retrieval is fully automated but natural weather conditions and environmental limitations are the difficulties of developing a fully automated pothole analysis system. The study specifies the research gapof robust techniques in real-time pothole detection.

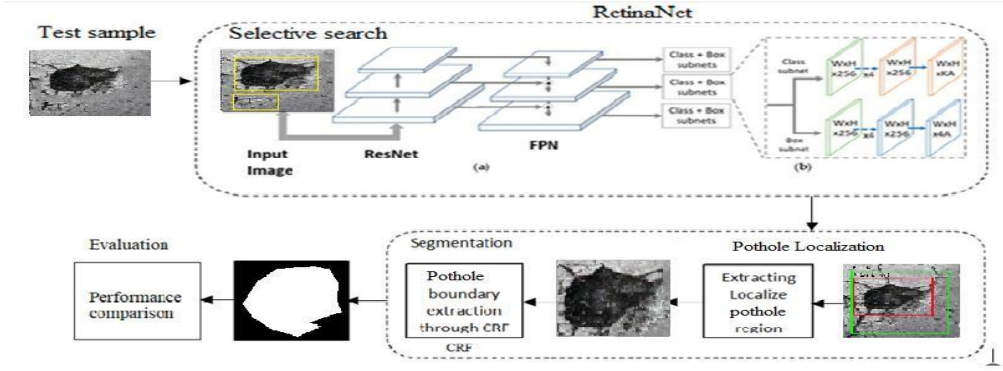


Figure 3.Proposed methodology for localization and Segmentation of Pothole region

In our proposed methodology firstly, we preprocess the images to refine the image quality. In data preprocessing, we have applied data augmentation techniques to increase the data set for better model learning at the training stage. After that, we trained the RetinaNet for potholes detection. Feature pyramid network used by retina net on top of the pre-trained model in a fully convolutional manner. This FPN layer enables the network to take an arbitrary size image and gives the result at multiple levels in the FPN layer in the form of proportionally sized feature maps. The top level of the feature map consists of grid cells that cover the large region of the image and suitable for the detection of large objects. This FPN is attached with two sub-networks, classification subnet and regression subnet, which helps the RetinaNet model to learn to detect the pothole region. When this trained RetinaNet model is applied to the test images, it localizes the pothole region with high precision of 0.93. After the detection of the pothole, the localized region is processed through CRF to statistically segment the pothole, as shown in Figure 3.

Material and Methods.

Image Processing

In the image preprocessing phase firstly, we have rescaled all the images. After rescale we have applied different data augmentation techniques to make out the best from the dataset.

Random cropping center: We have cropped our images to make the pothole object in the center of the image to get better information about the pothole.

Rotating the image: Changing the position of an object at a pivot point to some angle means the rotation of an image. The image with or without a pothole will look the same upon rotation and looks like that image is taken from a different angle.

Mirroring: Mirroring is used to create a mirror image of the object. We have flipped the image in the horizontal position to enlarge the dataset.

Salt and pepper noise: By changing the value of some of the pixels of an image from 0 to q we can add the salt and pepper noise. It will represent that image was taken on a dusty day.

RetinaNet Based Pothole detection

Different techniques have been explored regarding potholes detection. Keeping in view those techniques, we propose RetinaNet for Pothole detection as they are one stage detection model. It uses the Feature Pyramid Network (FPN)[12] to extract multiscale information and performs regression and classification box in one stage. It also used the focal loss method to improve the model's performance by addressing the issue of class

imbalance. In our proposed methodology we have used different backbone models like Restnet101, Densenet169, Efficientnet82, and VGG19 for training and testing of the Retinanet model on our dataset. Retinanet model uses the FPN as the top backbone layer above these pre-trained backbone neural networks.

Feature Extraction using Feature Pyramid Network (FPN)

Many object detection techniques use sliding window operation to get the region of interest (ROI). Later these older techniques were replaced in deep learning like CNN with better accuracy and precision, but it is also extensive. Therefore, we used FPN as the backbone which extracts features from the image. It takes an arbitrary image as an argument and gives results in equally sized feature maps by combining the semantic strong features with low resolution and weak features with high resolution. As Potholes have variations in size and shape, FPN uses deep features for road damage detection and model these variations. FPN uses the Pothole's color, intensity, and size at multiple levels for accurate and precise modeling and each layer as pothole local representation. FPN is attached with two sub-networks, classification subnet, and regression subnet.

Classification subnet: The classification subnet uses the output of FPN to predict the object (Pothole class) at each spatial position for each anchor and K object classes. This subnet contains the four 3*3 convolution layers with D channel filters from pyramid level and each channel followed by ReLU activation. In the end, the Pothole area is classified and localized by using sigmoid activators.

Box Regression Subnet: Regression is the subnetwork of the RetinaNet which is used for bounding box regression. Like classification subnet, it also uses the FPN's output and is attached with each feature pyramid level. It works parallel with the classification subnet for localization, pothole area. It localizes the pothole area concerning the anchor box if the pothole exists in the image.

Focal Loss

Focal loss is designed to address the class imbalance that occurs in single object detection methods when the dataset has a large number of background classes and few foreground classes. Due to class imbalance, the training of the dataset becomes inefficient and reduces the model performance as most locations are shown negative which gave no useful signal. Focal loss is based on entropy loss and it is minimized by adjusting the gamma values while training the Pothole dataset [25].

$$CEtrp\ (pt) = -\log\ (pt) \quad (1)$$

$$FLs(pt) = -(1 - pt)\ \gamma \log(pt) \quad (2)$$

Here CEtrp in (1) is the entropy loss and FLs in (2) is the focal loss. The gamma in (2) is the focusing parameter, its values range from 0 to 5.

RetinaNet Training Parameters

We have firstly trained our dataset on the retina net for accurate classification and localization of the pothole region in the image to get high accuracy. Secondly, we have used the RetinaNet to classify deep learning [7, 23]. We have used labeled ground truth in the training process of the deep learning network. RetinaNet model was pre-trained on the COCO dataset. To localize the pothole precisely, the RetinaNet was refined by using the parameters such as the learning rate which was initially set as 5, then 0.0001, and then change to 1 during training, 75 epochs, and 20,000 steps are used for training in RetinaNet. We

choose small anchor size 16 and skipped the biggest anchor of size 512 in our dataset to localize the pothole.

Detection of Pothole at Test time

To identify the pothole region in the images sigmoid cross entropy was used. The sigmoid layer is a nonlinear activation function, it changes the values which range between 0 and 1. To obtain the threshold confidence score, the obtained deep feature vector form image is passed to the sigmoid layer. At each feature pyramid network level, we have decoded box predictions. From these predictions, the top-level predictions are combined with non-maximum suppression along with a 0.5 threshold to get the final results of detections.

Segmentation of Pothole using Conditional Random Fields (CRF)

RetinaNet localizes the pothole region in asphalt surface images. We have used CRF to segment the pothole region so that we can detect the affected area of the asphalt surface. CRF is a type of discriminative undirected probabilistic graphical model that can represent the relationship between different variables [13]. CRF model helps us to use the observed one to estimate the unobserved ones. As this model creates the decision boundary between different classes, we have only two classes one is background class and the other is pothole class. To predict the pothole boundary, we have used the CRF model which gives better results than other state of the art methods.

For using the CRF, we have used the output of the RetinaNet which was the localized region of the pothole. We cropped that localized region from the image and give it to the CRF for the image segmentation. CRF arranged the image data points as a graph. We suppose that the image is an undirected graph C= (V, E) where V is nodes and each node corresponds to variable Xi, and E is the edge set. Let n=|V| shows the number of nodes in the graph and I be the input variable which was the localized area of the pothole in our case and Y=(y1,y2,...,yN) be the random output variable which can be the segmented pothole region or can be the segmented non-pothole area where V=I ∪Y and every Yv(v € V) gets a value from possible discrete labels range, which is in our case is pothole or non-pothole. In CRF we assume that each output variable Yv when conditioned on the input variable 'I' will obey the Markov property [14].

$$P(Yv \mid I, Yw, w \neq v) = p(Yv \mid I, bw, w \in S(v)) \quad (3)$$

Where S(v) is a set of adjacent nodes and the (Y, I) is the input and output variable of CRF. In Pothole segmentation we used i € I represent the image that has to be segmented and y € Y represents the segmented results of the image for each 'i' pixel in the CRF graph. The basic objective is to solve energy function.

$$E(y, i) = \sum i\, Ei(yi, ii) + \sum ij\, Eij(yi, yj, ii, ii) \quad (4)$$

In this equation E(y,i) is the Energy function [15]. The CRF requires a mash of each original image to perform the coarse boundary refinement for each predicted mask to get the pothole segmentation final result.

RESULTS AND EXPERIMENTS

Evaluation Protocols

The experiment was performed using a laptop with Intel(R) Core(TM) m3-7Y30 CPU @ 1.00GHz and 1.61GHz, 8GB memory, and a 64-bit operating system. This algorithm was implemented on Google Colab using windows 10. More sophisticated, we used python APIs, on the Collaboratory using GPU: 1xTesla K80, compute 3.7, having 2496 CUDA

cores, 12GB GDDR5 VRAM, CPU: 1xsingle core hyperthreaded Xeon Processors @2.3Ghz i.e(1 core, 2 threads), 12.6 GB memory.

Dataset

The dataset we used to perform our experiments is provided by Koch [1] and the Go (Zamir, 2014) google street view dataset.

Evaluation metrics

For validating each phase of the proposed methodology following evolution metrics were used.

Pothole detection

Pothole region detection was performed by using the Retina Net interaction over union (IOU) as shown in eq (5). IOU is predicted and ground truth box overlapping criteria. For Pothole region detection IOU's acceptable range is between 0.5 and 1, where from 0 to 0.3 is the background region. In the (5) True Positive (TP) are the correctly predicted boxes and others are False Positive (FP) and FN is False Negative.

$$IoU = 2 * TP / (TP + FN + FP) \quad (5)$$

$$mAP = \text{mean}\ (Tp/Tp+Fp) \quad (6)$$

To evaluate the classification phase, the mean average precession (map) is used to compute the average precision of pothole detection/ classification.

Pothole Segmentation

For the evaluation of the segmentation phase, we have used the following parameters Accuracy (AC), Specificity (SP), Dice Score (DI), Sensitivity (SE), and Jaccard coefficient (JC).

$$Ac = (tp + tn) / tp + fp + fn + tn \quad (7)$$

$$SP = tp / (tp + fp) \quad (8)$$

$$SE = tp / (tp + fn) \quad (9)$$

$$Di = 2 * (2 * tp / (fn + (2*tp) + fp)) \quad (10)$$

$$Jc = 2 * (tp / (tp + fn + fp)) \quad (11)$$

In these equations tp, tn, fp, fn represent true positive, true negative, false positive, and false negative pixels. In our case, the porthole pixel region is called true positive, so if the pothole is detected it's the region of true positive pixels, other than it is the region of non-pothole pixels or background is true negative pixels.

RESULTS AND DISCUSSION

The obtained results from the proposed methodology for pothole detection and segmentation will be discussed in this section.

Detection of pothole using RetinaNet

In our proposed model, the probability to the pothole area and remaining background is assigned by the convolution layers extracted fields, so that our region of interest which is the pothole in our model can be detected. Feature pyramid network (FPN) generated the convolution features from the proposal region, the RetinaNet extracted those features and maps with the trained model so that it can classify the proposed region.

In our proposed method, the pothole area was considered as positive and the rest of the image was considered as background or negative. For the pothole area detection, the IOU threshold is set to 0.5. The area which was below this threshold is considered as background or the negative class. If the threshold value was more than the threshold value it means a pothole was detected in the image other than that the region has not pothole if the IOU

threshold value is less than 0.3. The RetinaNet localized the pothole at the regression layer with a good map result shown in Table 2. The results in the form of images are shown in Figure 3. The detected pothole area is more precisely localized with the help of RetinaNet as shown in Figure 3. RetinaNet model was tested and trained using different pre-trained neural network models one by one and their map was calculated which is shown in Table 2. As the feature maps of Efficientnet82 and Mobilenet160 were different block by block, they didn't perform well and showed a low map as shown in Fig 8 and Table 2. Densenet169 and Vgg19 have shown better results in comparison with Mobilenet160 and Efficientnet82 but resnet101 has shown way better results than all of these models. The Resnet101 model has an activation function that connects it to FPN in RetinaNet[16]. Also, the Resnet101 model solves used the block of the residual network to solve the vanishing gradient problem so that feature maps can reach and jump over the layers it shows better results than other models. By seeing the results and the performance which shows that Resnet101 has high precision in localizing the Pothole region as shown in Table 2.

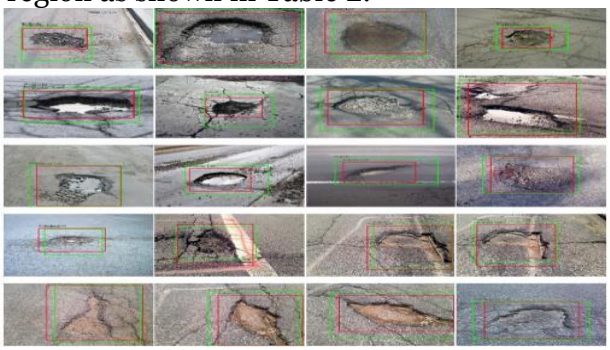


Figure4. high scoring test samples of Pothole detection

Table 1. Pothole localization mAP score using RetinaNet model with Resnet101 as the backbone

Image	Image1	Image2	Image3	Image4	Image5	Image6	Image7	mAP
Precision	0.93	0.99	0.91	0.99	0.89	0.91	0.92	0.93

Pothole Segmentation

For the evaluation of the segmentation phase, an accurate pothole region was required. To obtain the pothole region from the images we have used Condition Random Field (CRF). The resultant image of CRF segmentation was visually similar to the pothole ground truth as shown in figure 4. The performance of the segmentation results was measured by computing SE, SP, AC, F1 score on all the testing images. The proposed methodology got average values of AC as 0.932, F1 score as 0.8624, SP as 0.962, SE as 0.912. We have also calculated the DI and JC at the pixel level. The value of DI is 0.912 and the value of JC is 0.896. We got better results of segmentation because the Retinanet localized the Pothole more accurately.

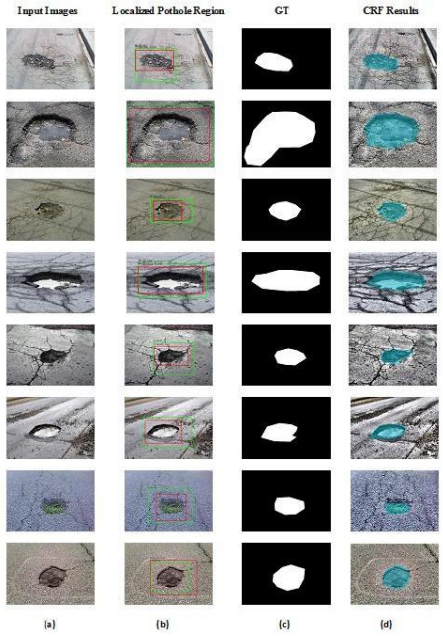


Figure 5. (a) input images, (b) are localized pothole region images, (c) are original ground truth masks and (d) are resultant segmented images

Table 2 Different RetinaNet models mAP results

Backbones	mAP
Resnet101	0.934
Vgg19	0.893
Densenet169	0.856
Efficientnet82	0.814
Mobilenet160	0.764

Comparative Analysis

Comparison with other state of the art methods

Our Proposed methodology has given better results with other state of the art techniques as shown in Table 3. The techniques used in this research have a lower average accuracy than our proposed methodology. Our methodology resulted in improved pothole localization by showing a pothole through its outline. Also, our proposed methodology

shows better results of segmentation than other state of the art techniques due to the good pothole region localization by using the RetinaNet model.

Table 3 Comparison of Performance

Method	Accuracy	Precision	Recall
Retinanet	**0.93**	**0.89**	**0.85**
Koch et al. [1]	0.85	0.81	0.86
Ryu et al. [17]	0.73	0.80	0.73
Wang et al. [18]	0.86	0.83	0.87
HOG Based Approach [19]	0.90	0.86	0.94
Reddy, D. R et al. [22]	0.84	0.78	0.88
Dharneeshkar et al. [23]	72.16	0.76	0.4
Surekha Arjapure et al.[24]	0.90	0.88	0.83

5. CONCLUSION

In this paper, we have proposed a novel technique based on RetinaNet for accurate pothole localization and Conditional Radom Field for efficient and precise pothole segmentation. Our technique consists of three steps, image preprocessing in which data augmentation was done, Pothole localization, and Pothole segmentation. The localized region of the pothole was obtained by RetinaNet shows a good result as compared to other state of art methods by getting deep features, using that localized pothole region the CRF segmentation show improved performance. Our method detects the different types of potholes of asphalt surfaces by making the mechanism of training data efficient. The experiment was performed on the dataset provided by Koch and the google street view dataset.

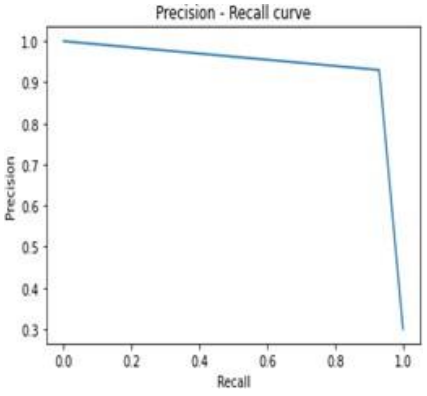


Figure 6. The area under the Precision-Recall curve for Pothole region separation

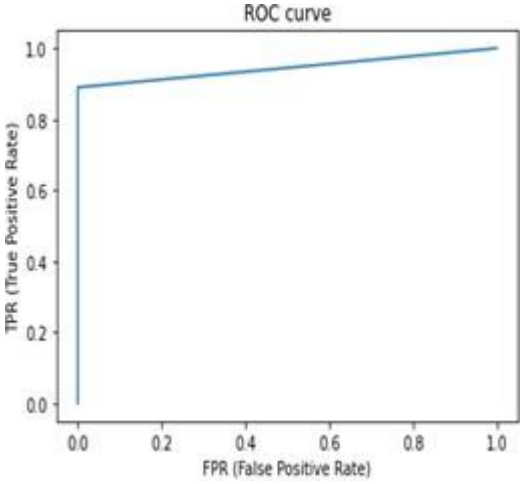


Figure7. The area under the ROC curve for the Pothole Region Separation

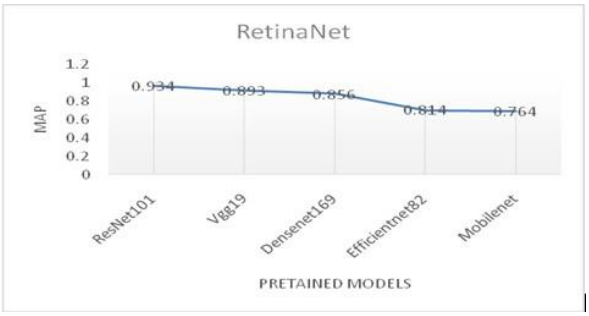


Figure 8. mAP score of different Pretrained models on RetinaNet

ACKNOWLEDGMENT

The authors would like to thank Department of Computer Science, University of Engineering and Technology Taxila for R&D support and Christian Koch [1] for providing the dataset.

Author's Contribution. All the authors contributed equally.

Project details. NIL

Conflict of interest. We declare no conflict of interest for publishing this manuscript in IJIST.

REFERENCES

[1] C. Koch and I. Brilakis, "*Pothole detection in asphalt pavement images*", Advanced Engineering Informatics, vol. 25, pp. 507–515, 2011.
[2] A. Danti, J. Kulkarni, and P. Hiremath, "*An Image Processing Approach to Detect Lanes*", Pot Holes and Recognize Road Signs. International Journal of Modeling and Optimization, vol. 6, pp. 658-662, 2012.
[3] S. Nienaber, M. Booysen and R. Kroon, "*Detecting Potholes Using Simple Image Processing Techniques and Real-world Footage*", 34th Southern African Transport Conference, 2015.
[4] E. Buza, S. Omanovic and A. Huseinovic, "*Pothole Detection with Image Processing and Spectral Clustering*", Recent Advances in Computer Science and Networking, vol. 810, pp. 4853, 2013.
[5] M. Salman, S. Mathavan, K. Kamal and M. Rahman, "*Pavement Crack Detection Using the Gabor Filter*", Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems, The Hague, The Netherlands, October 6-9, 2013.
[6] K. Azhar, F. Murtaza, M. H. Yousaf, H. A. Habib, "*Computer Vision Based Detection and Localization of Potholes in Asphalt Pavement Images*", IEEE Canadian Conference on Electrical and Computer Engineering, 2016.
[7] A. B. Makone, and A. G. Rathod., "*Pothole Dimensions Measurement using Mean Shift-based Speckle Filtering*" International Journal of Science Techonology and Engineering, vol. 2, 2016.
[8] L. Huidrom, L. K. Das, S. K. Sud,"*Method for automated assessment of potholes, cracks and patches from road surface video clips*", Procedia - Social and Behavioral Sciences, vol. 104, pp. 312 – 321, 2013.
[9] Y. Jia, Z. He and H. Zhu, "*A Hierarchical Segmentation Approach towards Roads and Slopes for collapse Recognition*", International Journal of Signal Processing and Pattern Recognition, vol. 6, pp.153-164, 2013.
[10] M. Jahanshahi, F. Jazizadeh, S. Masri and B. Gerber, "*Unsupervised approach for autonomous pavement-defect detection and quantification using an inexpensive depth sensor*", Journal of Computer and Civil Engineering, vol. 27, pp. 743-754,2013.
[11] C. Koch, K. Doycheva, V. Kasireddy, B. Akinci, P. Fieguth, "*A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure*", Advanced Engineering Informatics, vol. 30, pp. 208-210, 2016.
[12] T. Lin, P. Doll´ar, R. B. Girshick, K. He, B. Hariharan, and S. J. Belongie, "*Feature pyramid networks for object detection*," Computer Research Repository, vol. abs/1612.03144, 2016.
[13] Conditional random field [Online]. Available: https://en.wikipedia.org/wiki/Conditional_random_field [Accessed].
[14] T. L. XH, 1 and J. Ma, "*Conditional Random Fields for Image Labeling*", Mathematical Problems in Engineering, 2016.
[15] P. Krähenbühl, V. Koltun, "*Efficient inference in fully connected crfs with gaussian edge potentials*," Advances in Neural Information Processing Systems, 2011.
[16] L. Ale, N. Zhang, L. Li, "*Road damage detection using RetinaNet.*," IEEE International Conference on Big Data (Big Data); 2018.
[17] S. Ryu, T. Kim, Y. Kim, "*Image-based pothole detection system for ITS service and road management system,*" Mathematical Problems in Engineering. 2015.
[18] P. Wang, "*Asphalt pavement pothole detection and segmentation based on wavelet energy field*", Mathematical Problems in Engineering, 2017.

[19] K. Christian, K. Gorgieva, V. Kasireddy, B. Akinci, P. Fieguth, "*A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure*," Advanced Engineering Informatics, vol. 29, pp. 196–210, 2015.
[20] "*Fedral Highway Administration*" , LTPP- Beyond FY 2009: What needs to be done? Technical Report, Publication No. FHWA-HRT-09-052, 2009
http://www.fha.dot.gove/pavement/ltpp/pubs/09025/09052.pdf
[21] Taehyeong Kim, Seung-Ki Ryu, "*Review and Analysis of Pothole Detection Methods*" Journal of Emerging Trends in Computiing and Information Sciences , Vol. 5, No. 2014, ISSN 2079-8409.
[22] Reddy, D. R., Goud, G. P. C., & Naidu, C. D. "*Internet of Things Based Pothole Detection System using Kinect Sensor*" 2019, Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2019.
[23] J, D., V, S. D., S A, A., R, K., & Parameswaran, L. "*Deep Learning based Detection of potholes in Indian roads using YOLO*". 2020 International Conference on Inventive Computation Technologies (ICICT). 2020.
[24] Surekha Arjapure, D.R.Kalbande, "*Deep Learning Model for Pothole Detection and Area Computation*",2021 International Conference on Communication information and Computing Technology (ICCICT), 2021.
[25] LIN, T.-Y., GOYAL, P., GIRSHICK, R., HE, K. & DOLLÁR, P. "*Focal Proceedings of the IEEE international conference on computer vision*", pp: 2980-2988, 2017.

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Compact Frequency Selective Surface (FSS) for X-Band Shielding

Taiba Khalil[1], Muhammad Ali Riaz[1], Humayun Shahid[1], Muhammad Jamil Khan[1], Yasar Amin[1]
[1] Department of Telecommunication Engineering (University of Engineering and Technology Taxila).
* Correspondence: Taiba Khalil, Email: taiba.khalil14@gmail.com
Citation | Khalil. T, Riaz. M. A, Khan. M. J and Amin. Y, "Compact Frequency Selective Surface (FSS) for X-Band Shielding". International Journal of Innovations in Science and Technology, Special Issue, pp: 140-152, 2021.
Received | Dec 01, 2021; Revised | Dec 22, 2021 Accepted | Dec 26, 2021; Published | Jan 06, 2022.

With the increase in the usage of electromagnetic devices, electromagnetic interference increased many folds. Frequency Selective Surface (FSS) provide effective shielding from unwanted frequency ranges. A thin, conformal band-stop FSS is presented in this research that provides effective electromagnetic shielding properties in X-band. The FSS acts as a band stop filter at 10 GHz. The proposed FSS has 54.7% fractional bandwidth. The design is of the dimensions 6.79 x 6.79 x 0.127 mm^3, employing Rogers RT 5880 substrate with 0.0009 dielectric constant. It has an attenuation of at least -57.97 dB. The proposed FSS shows oblique incidence angle independence for both TE and TM modes, up to 60^o scan angle. The incidence angle independence makes the FSS response stable for both normal and varying angles of the incident waves. The design has a copper cladding of 0.018 mm, making the overall FSS thickness of 0.145 mm. The thin substrate makes the design flexible and easily bendable for curved surfaces. Its thin structure makes it easily applicable on buildings, vehicles and military aircrafts for electromagnetic shielding purposes. The conformability and shielding properties make the design suitable for various other applications.
Keywords: Frequency Selective Surface; X-Band; Electromagnetic Shielding; Flexible; compact.

INTRODUCTION

Rapid technological enhancements and the invention of smart devices have connected several devices with each other. This device-to-device connectivity and various other technologies like the Internet of Things (IoT), space communication and radars have increased the electromagnetic interference. 5G also aims to provide a huge data rate with less latency, [1,2] increasing the electromagnetic interference. X-Band (8-12 GHz) is one of the major contributors to increasing electromagnetic interference [3]. X-Band frequency ranges are immune to atmospheric attenuation which makes them ideal for weather monitoring and space communication. It has a shorter wavelength and higher image resolution for radar applications,

maritime traffic control and vehicle detection [4]. Electromagnetic interference causes serious threats to sensitive electronic systems, healthcare and military devices. The electromagnetic interference can be minimized by using effective shielding, for example metallic shields can be used for isolation from electromagnetic signals. But they block the desired electromagnetic radiations as well [5].

FSS provides effective shielding properties that can be used for shielding against the electromagnetic radiations owing to their selective filtering properties. FSS are periodic surfaces having identical elements in an infinite array arranged in one or two dimensions [6]. FSS has the capability of passing and blocking frequencies of desired range [7]. An array of periodic metallic patches or apertures on a dielectric substrate forms a frequency selective surface. The metallic apertures act as band pass filters, and the metallic patches act as bandstop filters at resonance frequency [8].

FSS can be used as polarizers, sub reflectors, filters, beam multiplexers in optical and microwave applications. Dual-band FSS for GSM shielding based on the paper substrate is presented to ensure effective shielding, having a low profile at affordable cost. A dual layer absorptive/transmissive FSS is proposed absorbing WLAN signals and transmitting GSM signals with reduced multipath fading [9]. FSS can be made electronically switchable by making it switch from transparent and reflection state. This makes it a spatial filter for reconfiguring building architecture [10]. FSS is known as spatial filters as they are different from classical filters. FSS deals with signals having different polarizations and angles of incidence [6]. Its capability to show a stable response for various incident angles for TM and TE modes makes it an excellent tool for use in Electromagnetic Compatible (EMC) applications, where polarization modes and angles are usually unknown [11]. A spiral slot frequency selective surface filter for shielding from undesired electromagnetic noise produced by wireless charging batteries is proposed in reference [12]. Spiral slot with seven to ten windings is cut on the metallic ground plane. The helical antenna is composed of five wire windings and the FSS is placed in between two helical antennas. FSS and Substrate Integrated Waveguide (SIW) technology is used for airborne radome applications [13,14].

The Electromagnetic Band Gap (EBG) structures can be used to suppress waves in the desired stop band. A hybrid of FSS and EBG is used to suppress surface waves [15]. Previously, multi-layer FSS was presented for shielding purposes in applications. Two FSS were cascaded using Koch fractal elements for shielding, and providing 20 dB attenuation with a wide stopband [16]. Multi-layer FSS structures needed to be carefully aligned and more FSS layers might cause processing errors [17]. Extremely High Frequency (EHF) ranges are used in transmitters of satellites, which increases the electromagnetic interferences and it can damage sensitive external transceivers. Due to a large number of transmitters and receivers outside the satellite, satellites can get interference from nearby transmitters. 2D cross-dipole FSS structure with specific stop band property can be used for shielding against out-of-band signals [18]. Miniaturized FSS for shielding in the millimeter wave region (40 – 70 GHz) is presented in reference [19]. The structure is single layered having Jerusalem cross and Fan shape etched on both sides of the substrate. It has a shielding effectiveness of more than 16 dB and can be used for mutual coupling reduction. FSS is analyzed in terms of its S-

Parameters, polarization independence and shielding effectiveness. Various methods have been used for calculating the shielding effectiveness of an enclosure. The transmission line approach for calculating the shielding effectiveness of a double layer frequency selective surface enclosure, with different angles of incidence is much faster than other numerical methods [20].

FSS can be made flexible by using flexible substrate; it can provide conformal shielding by using it on curved surfaces. Stretchable FSS embedded with silicone elastomers are able to be wrapped on doubly curved surfaces [21]. FSS shielding for X band by using miniature ring shaped elements was proposed in [22]. Effective shielding in X and Ka bands were achieved by a single layer Ultra-Wide Band (UWB) polarization stable FSS [23]. Shielding in X-band is of immense importance as it is used for communication and radar engineering. Gas plasma encapsulated in FSS with switchable attenuation from 24dB to 44dB is used for electromagnetic shielding in X-band. This low loss device has three orders of structure and a large area that makes it useful for beam steering and RCS control as well [24]. Compact FSS structure consisting of eight segment polygon having an attenuation of 35dB is used and presented for effective shielding in X-band. The structure shows polarization independence in TE and TM modes [25]. A convoluted ring loop FSS structure having a low profile is developed for X-band shielding. For miniaturization, four stubs are added at a 90^{o} angle to each other. The structure has an attenuation of 37 dB for a oblique angle of incidence up to 45^{o} [26]. A sub-wavelength polarization independent FSS for X-band shielding is proposed in reference [27]. It shows attenuation of greater than 25 dB for the shielding band. A FSS for shielding X-band SATCOMS is proposed in reference [28]. It shows bandstop characteristics with attenuation of up to 32 dB in the shielding band. An FSS consisting of a convoluted circular loop for electromagnetic shielding applications is proposed in reference [29]. Its effective shielding is centered at 9.7 GHz with an attenuation of atleast 47 dB.

The proposed design of this study was comprised of a thin FSS layer printed on Rogers's 5880 substrate having a thickness of 0.127 mm. The thin dielectric substrate makes the structure flexible, making it easier to employ on curved surfaces. The design stops the frequency at 10 GHz (X-band) and transmits the rest of the frequency ranges. It has an attenuation of at least -57.97 dB. The proposed design shows oblique incidence angle independence in terms of both TE and TM modes for up to scan angle of 60^{o}. The design consists of a compact unit cell to provide stability for various incidence angles. It employs an extremely thin substrate that is highly flexible and easily applicable on curved surfaces. It can be used in antennas for enhancing the gain and directivity of antenna.

Material and methods.

The proposed design is simulated on Rogers 5880 substrate in Figure. 1 and 2. The substrate has a dielectric loss tangent of 0.0009 and relative permittivity of 2.2. The thickness of substrate is 0.127 mm that makes the FSS flexible. The overall unit cell has dimensions of 6.79 x 6.79 mm^2. The design consists of an X-shaped metallic loop placed on the substrate. The proposed design is printed on one side of the dielectric substrate. The thickness of copper cladding is 18 μ m.

The parameter 'P' represents the length and width of the substrate and 'L' represents the length and width of the metallic X-shaped square loop. The design consists of V-shaped inward stubs on all sides of the square loop. The design parameters are given in Table 1. The positioning and size of these stubs provide the incidence angle independent in terms of TE and TM modes.

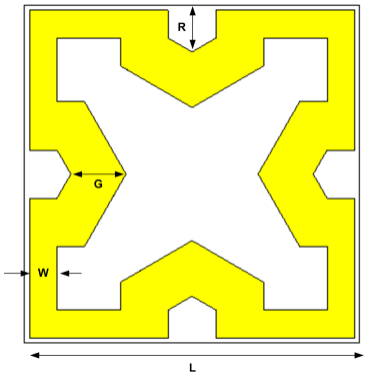


Figure 1. Proposed FSS Unit Cell Front View

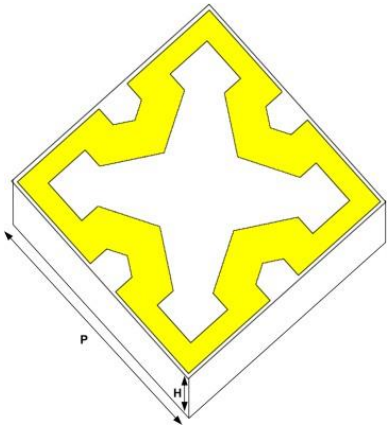


Figure 2. Proposed FSS Unit Cell Perspective View

Table 1. Parameter List

Heading	Heading
P	6.79
L	6.6
G	1.12
R	0.935
W	0.56
H	0.127

Result and discussion.

Ansoft HFSS was used for simulation and optimization of the proposed design. The proposed FSS is designed in order to shield against frequencies in the X-band (8-12 GHz). S_{11} and S_{21} parameters were calculated for the normal angle of incidence in Figure 3. Shielding

effectiveness (SE) of the proposed design was calculated for checking shielding against X-band shown in Figure 4. The design shows good shielding properties for X-band with its resonance frequency at 10 GHz. It has an attenuation of at least -57.97 dB.

Oblique Incidence Angle Independence - The design showed independence for oblique angles of incidence for both TE and TM modes. Figure 5 shows SE of TE and TM modes, for the angle of incidence up to 60°. The polarization independence was provided with the help of V-shaped stubs pointed inwards in the square loop. The widths of these stubs affect the response of FSS, for different incident angles. The fractional bandwidth increased slightly with the increase in the angle of incidence for TE modes in Figure 5. The fractional bandwidth decreased with the increase in the angle of incidence for TM modes in Figure 6.
The change in the resonant frequency was not significant for the change in angle of incidence for both TE and TM modes. The resonant frequency remains in the allowable limits for both TE and TM modes, by changing the angles of incidence up to 60°. A few abnormalities were observed from 15.1 GHz to 18 GHz in case of TE modes because of the varying angle of the incident wave. The angles greater than 30° interferes with the higher frequency range of X-band. It showed minor shielding effects for attenuation up to 27 dB for scan angle of 30° and up to 45 dB for the scan angle of 60°.

Effects of Varying Length i.e a A frequency shift was observed in the resonance frequency by varying the length of the inner copper loop 'L' in Figure 7. By increasing 'L' from 6.6 mm to 6.7 mm, the resonant frequency shifts to the lower frequency range, from 10 GHz to 8.36 GHz. By decreasing 'L' from 6.6 mm to 6.4 mm, the resonant frequency shifts to a higher frequency range of 10 GHz to 11.52 GHz. In this case the frequency shift was indirectly proportional to varying 'L'. The frequency shift can also be observed by varying the thickness of the hexagonal stubs 'G' in Figure 8.

By increasing 'G' from 1.08 mm to 1.12 mm, the resonant frequency shifted to a lower frequency range, from 10 GHz to 9.57 GHz. In this case, the frequency shift was directly proportional to 'G'. However, the frequency shift was slightly smaller for varying 'G' than 'L'.

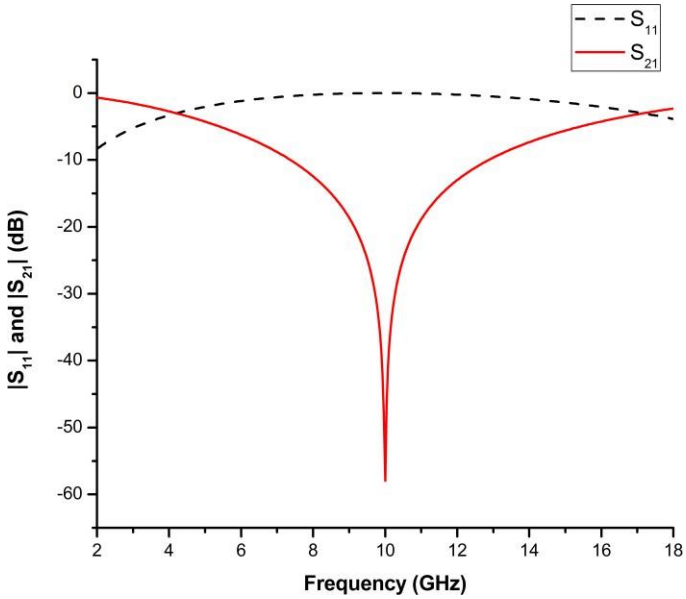


Figure 3. S_{11} and S_{21} Parameter

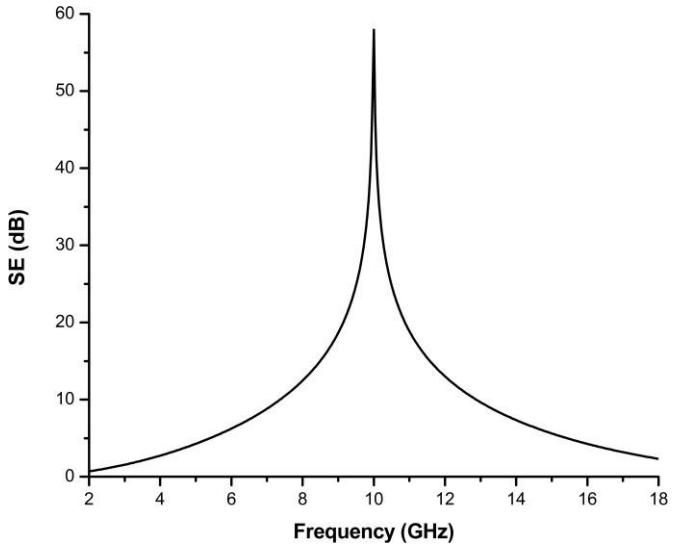


Figure 4. Shielding Effectiveness

Loss Tangent – The loss tangent of the employed substrate affected the SE of the proposed FSS in Figure 9 by varying loss tangent 'δ', due to which the SE value also varies. When δ=0 the SE reaches to the maximum value of -57.97 db. By increasing δ from 0 to 1 the SE starts decreasing until it reaches to the minimum value of δ =1. For the proposed design δ=0.0009 for Rogers 5880 substrate.

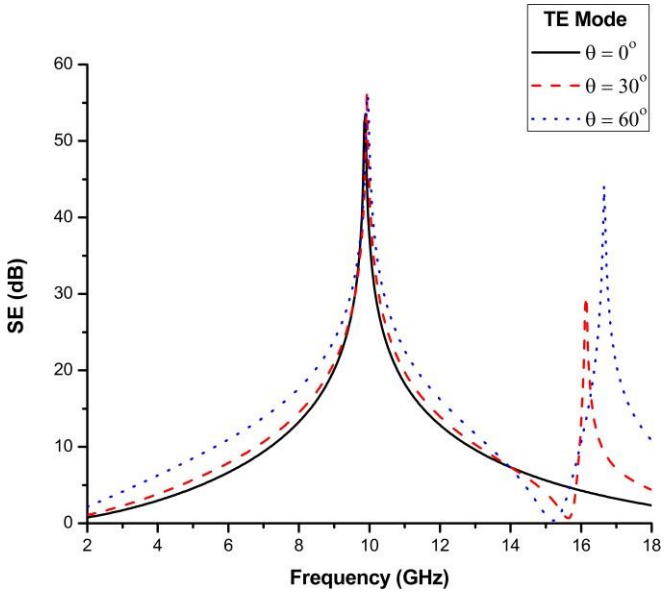


Figure 5. Effects of Varying Angles of Incidence on Proposed FSS for TE Modes.

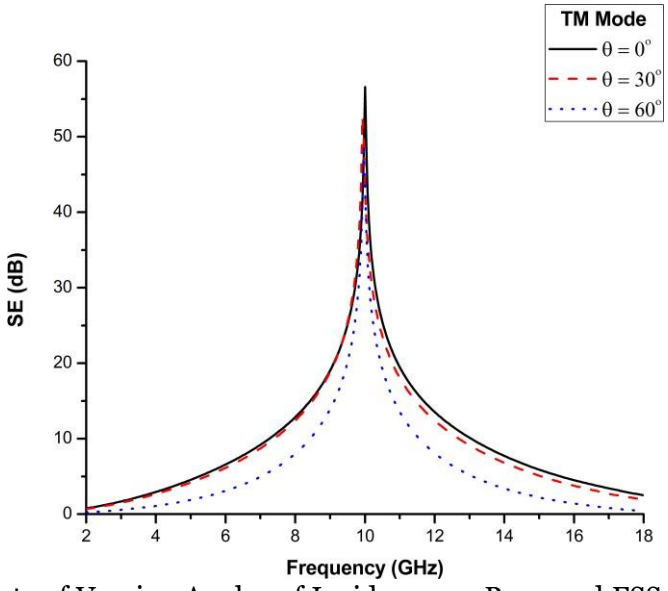


Figure 6. Effects of Varying Angles of Incidence on Proposed FSS for TM Modes

Relative Permittivity – The relative permittivity used for the proposed design is 2.2. By increasing the relative permittivity, the resonant frequency was shifted to the lower frequency range as shown in Figure 10. For the increase in the relative permittivity, from 2.2 to 6.2, the resonant frequency shifts from 10 GHz to 7.5 GHz. For the decrease in the relative permittivity from 2.2 to 1.2, the frequency shifted to the higher frequency range at 10.78 GHz. The frequency shift was indirectly proportional to the relative permittivity.

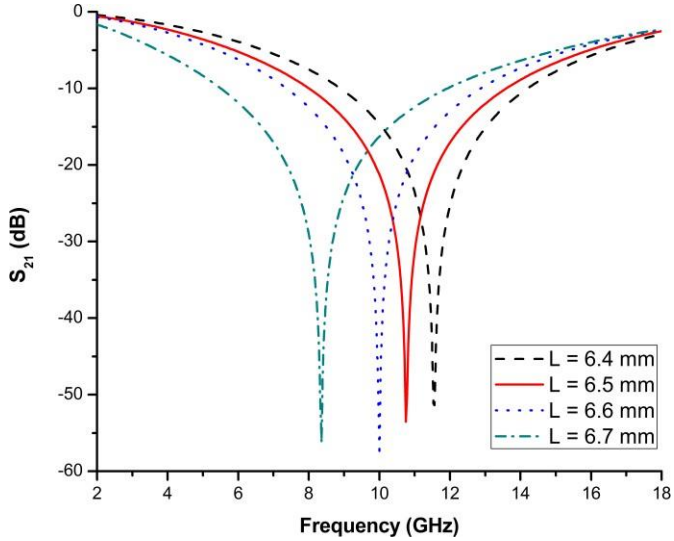


Figure 7. Varying length of square loop 'L'

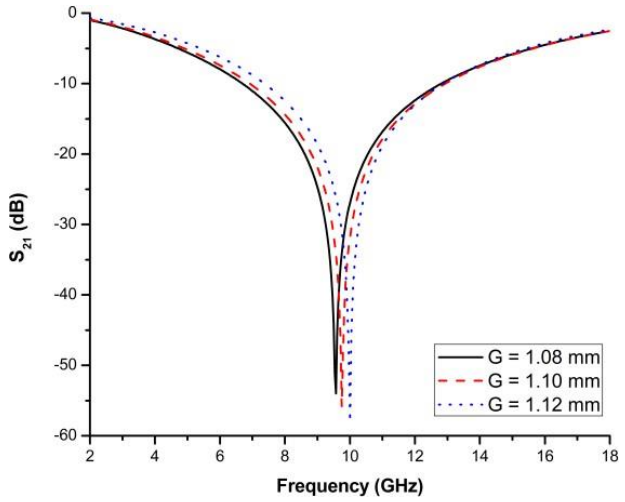


Figure 8. Varying Length of hexagonal stub 'G'

Fss Prototype Fabrication - The proposed FSS design was fabricated on a Rogers RT 5880 substrate sheet of 9 x 12 inches. The copper patch on the substrate was 18 μm in thickness. The FSS prototype was fabricated by etching the metallic patch on the substrate. The FSS unit cells were of the order 33 x 44 across the sheet as shown in Figure 11.

The proposed design was compared with other studies in the relevant field in Table. II. The results shows that the proposed FSS is thinnest among others, making it extremely flexible and applicable on curved surfaces.

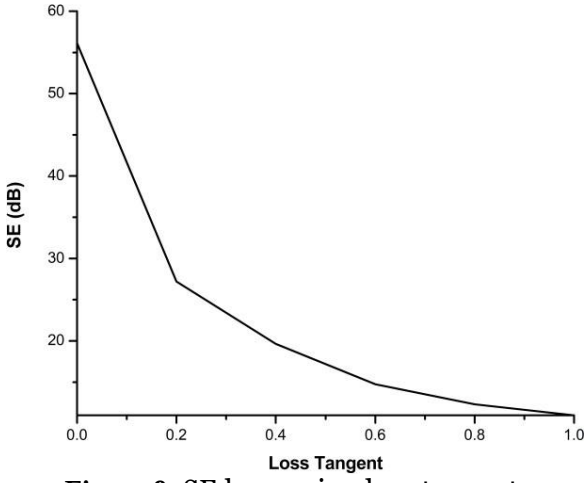


Figure 9. SE by varying loss tangent

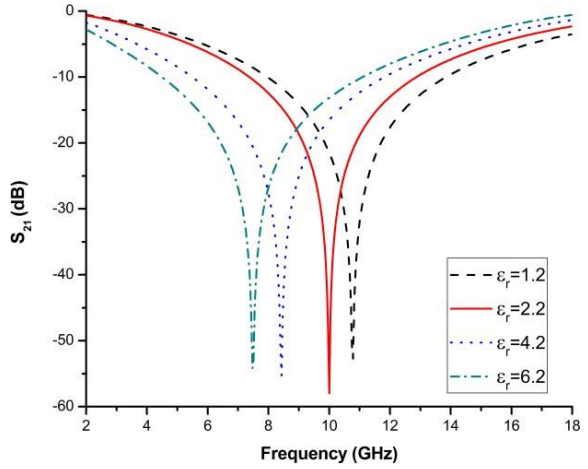


Figure 10. Frequency shift by varying permittivity

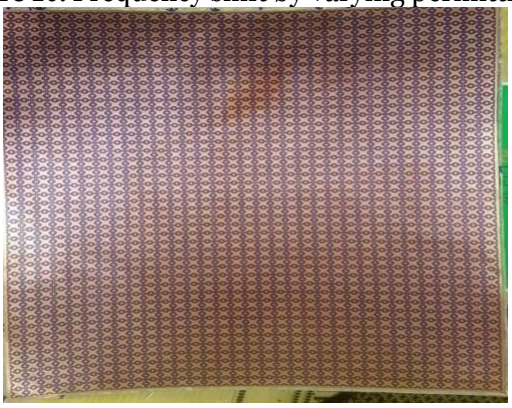


Figure 11. Fabricated FSS Prototype

Table 2. Comparison

Reference	Substrate	Size (mm^3)	FSS Sheet Thickness (mm)	Oblique Incidence Angle Independence	Flexible
30	FR4	12x12x3.2	3.235	Up to 45^0	No
31	Arlon Di880	8.8x8.8x0.762	0.762	Up to 60^0	No
32	FR4	10x10x0.7	0.797	No	No
33	Rogers 5880	6.8x6.8x0.127	0.162	Up to 60^0	Yes
25	FR4	5.3x5.3x1.6	1.635	Up to 60^0	No
22	FR4	7.1x7.1x1.6	1.635	Up to 45^0	No
Proposed Work	Rogers 5880	6.79x6.79x0.127	0.145	Up to 60^0	Yes

Conclusion.

An FSS for X-band shielding is proposed. The design is of 6.79 x 6.79 x 0.127 mm^3 dimensions. The proposed FSS show attenuation of at least -57.97 dB with a fractional bandwidth of 54.7 %. The FSS absorbs the frequency at 10 GHz and transmits the rest of the frequency ranges. The design is oblique incident angle independent in terms of TE and TM modes up to a scan angle of 60^0. The proposed design is extremely thin and flexible which makes it easily applicable on curved surfaces. It can be used in antennas to gain enhancement due to its reflective properties. It can be used in radomes, aircrafts and buildings for shielding purposes.

Acknowledgement.

All admires and thankfulness to Almighty Allah who gave me wisdom and knowledge to explore the world. I am thankful to my Research Project Supervisor Engr. Dr. Muhammad Ali Riaz for his aspiring guidance and kind advice during the project work. I am sincerely grateful to Engr. Dr. Humayun Shahid for his precious time that he gave me during the completion of my project and friendly advice whenever I needed his help. I am thankful to my family for their endless love and support during my work.

Author's Contribution.

The concept and idea of research was suggested by Dr. Muhammad Ali Riaz. The whole research and paper writing was performed under his supervision and reviewed by him. Taiba Khalil designed the proposed FSS and performed simulations to get the required results. She also wrote the research paper under the guidance of Dr. Humayun Shahid. The softwares used for the simulations are provided by Dr. Yasar Amin and Dr. Muhammad Jamil Khan.

Project details. The research was done for completion of Masters in Engineering Research Project. Total amount spent on project in 8000 PKR, for dielectric substrate and fabrication of the FSS Sheet. The project was completed in February 2019.

Conflict of interest. There is no conflict of interest for submitting this paper to IJIST.

REFRENCES

[1] Sani Yahya, M., and Rahim, SKA, "*15 GHz grid array antenna for 5G mobile communications system,*" Microwave and OpticalTechnology Letters, 2016, vol. 58, no. 12, pp. 2977–2980.

[2] Hakimi, S., Rahim, S. K. A, *"Millimeter-wave microstrip Bent line Grid Array antenna for 5G mobile communication networks,"* In Proceedingsof the IEEE Asia-Pacific Microwave Conference (APMC),Sendai, Japan, 2014, pp. 622–624.

[3] Bruder, J. A., Carlo, J., Gurney, J., Gorman, J " *IEEE standard for letter designations for radar-frequency bands,*" IEEEAerospace & Electronic Systems Society, 2003, vol. 6, no. pp. 1– 3.

[4] Damini, A., McDonald, M., Haslam, G. E, "*X-band wideband experimental airborne radar for SAR, GMTI and maritime surveillance,*" IEE Proceedings-Radar, Sonar and Navigation, 2003, vol. 150, no. 4, pp. 305–312.

[5] Seman, F. C., Khalid, N. K, "*Investigations on fractal square loop FSS at oblique incidence for GSM applications,*" In IEEE ElectricalPower, Electronics, Communicatons, Control and Informatics Seminar(EECCIS), Malang, Indonesia, 2014, vol. 58, no. , pp. 62–66.

[6] Munk, B. A, "*Frequency selective surfaces: theory and design,*" New York (USA), John Wiley and Sons, 2005.

[7] Anwar, Rana, Mao, L., Ning, H, "*Frequency selective surfaces: A review of Applied Sciences,"* Microwave and Optical TechnologyLetters, 2018, vol. 8, no. 9, pp. 1689.

[8] Unal, E., Gokcen, A., Kutlu, Y, " *Effective electromagnetic shielding,*" IEEE Microwave magazine, 2006, vol. 7, no. 4, pp. 48–54.

[9] Kiani, G. I., Weily, A. R., Esselle, K. P, "*A novel absorb/transmit FSS for secure indoor wireless networks with reduced multipath fading,*" IEEE Microwave andWireless Components Letters, 2006,vol. 16, no. 6, pp. 378–380.

[10] Kiani, G. I., Ford, K. L., Olsson, L. G., Esselle, K. P., Panagamuwa, C. J, "*Switchable frequency selective surface for reconfigurable electromagnetic architecture of buildings,*" IEEE Transactions on Antennas and Propagation, 2009, vol. 58, no. 2, pp. 581–584.

[11] Huang, F. C., Chiu, C. N., Wu, T. L., Chiou, Y. P, "*A circularring miniaturized-element metasurface with many good features for frequency selective shielding applications,"* IEEE Transactions onElectromagnetic Compatibility, 2015, vol. 57, no. 3, pp. 365–374.

[12] Firdaus, K., Sakakibara, K., Amano, Y., Hirayama, H., Kikuma, N., Tabata, T., Kojima, S. H, "*Design of spiral-slot frequency selective surfaces for shielding from noises of wireless power transfer,*" In Proceedings of the IEEE International Workshopon Antenna Technology (iWAT), South Korea., 2015, pp. 345–347.

[13] Kanth, V. K., Raghavan, S, "*EM Design and Analysis of Frequency Selective Surface Based on Substrate-Integrated Waveguide Technology for Airborne Radome Application.*," IEEE Transactionson Microwave Theory and Techniques, 2019, vol. 67, no. 5, pp. 1727– 1739.

[14] Varikuntla, K. K., Velu, R. S, "*Design and development of angularly stable and polarisation rotating FSS radome based on substrate-integrated waveguide technology,"* IET Microwaves, Antennas and Propagation, 2019, vol. 13, no. 4, pp. 478–484.

[15] Varikuntla, K. K., Velu, R. S, "*Design and development of angularly stable and polarisation rotating FSS radome based on substrate-integrated waveguide technology,"* IET Microwaves, Antennas and Propagation, 2019, vol. 13, no. 4, pp. 478–484.

[16] Manicuba, R. H., D'Assuncao, A. G., Campos, A. L, "*Wide stop-band cascaded frequency selective surfaces with Koch fractal elements,*" In Digests of the 14th Biennial IEEE Conference on ElectromagneticField Computation, Chicago, IL, USA., 2010, vol. 16, pp. 1–1.

[17] Ma, X., Wan, G., Zhang, W., Mu, Y., Tang, X, "*Synthesis of second-order wide-passband frequency selective surface using double- periodic structures,"* IET Microwaves, Antennas and Propagation, 2006, vol. 13, no. 3, pp. 373–379.

[18] Turner, I, "*Use of frequency selective surfaces to reduce coupling between antennas on satellites,*" In Proceedingsof the IEEE International Symposium on Electromagnetic Compatibility (EMC), Dresden, Germany., 2015, pp. 340–343.

[19] Kesavan, A., Karimian, R., Denidini, T. A, "*A novel wideband frequency selective surface for millimeter-wave applications,*" IEEEAntennas and Wireless Propagation Letters, 2016, vol. 15, pp. 1711– 1714.

[20] Sureshkumar, T. R., Venkatesh, C., Salil, P., Subbarao, B, "*Transmission line approach to calculate the shielding effectiveness of an enclosure with double-layer frequency selective surface,*" IEEE Transactions on Electromagnetic Compatibility, 2015, vol. 57, no. 6, pp. 1736–1739.

[21] Gurrala, P., Oren, S., Liu, P., Song, J., Dong, L, "*Fully conformal square-patch frequency-selective surface toward wearable electromagnetic shielding,*" IEEE Antennas andWireless Propagation Letters, 2017, vol. 16, pp. 2602–2605.

[22] Yong, W. Y., Rahim, S. K. A., Seman, F. C., Ramili, M. R., Remili, N. A, "*Miniaturisation of ring shape element frequency selective surface for X-band shielding,*" In Proceedings of the IEEE AsiaPacific Microwave Conference (APMC), Kuala Lumpar, Malaysia., 2017, pp. 877–880.

[23] Sohail, I., Ranga, Y., Matekovits, L., Esselle, K. P., Hayt, S. G, "*A low-profile single-layer UWB polarization stable FSS for electromagnetic shielding applications*," In Proceedings ofthe IEEE International Workshop on Antenna Technology: SmallAntennas, Novel EM Structures and Materials, and Applications(iWAT), Sydney, Australia., 2014, pp. 220–223.

[24] Cross, L. W, "*Study of X-band plasma devices for shielding applications,*" In Proceedings of the IEEE MTT-S International MicrowaveSymposium (IMS2014), Tampa, Florida, USA., 2014, pp. 1–4.

[25] Bilal, M., Saleem, R., Khan, F. A., Quddus, A., Shafique, M. F, "*Frequency selective surface for X-band shielding applications,*" In Proceedings of the IEEE 16th Mediterranean MicrowaveSymposium (MMS), Abu Dhabi, United Arab Emirates., 2016, pp. 1–3.

[26] X. Wu, G. Jiang, X. Wang, P. Xie and X. Li, "*A Multi-Level-Denoising Autoencoder Approach for Wind Turbine Fault Detection,*" in *IEEE Access,* vol. 7, pp. 59376-59387, 2019

[27] M. Karamirad, C. Ghobadi, J. Nourinia, S. Abbasi and B. Mohammadi, "*Sub-Wavelength Polarization-Independent Frequency Selective Surface for X-band Shielding,*" *2020 28th Iranian Conference on Electrical Engineering (ICEE),* 2020, pp. 1-3.

[28] M. Bilal, R. A. Wagan, R. Saleem, A. Q. Satti, T. Shabbir and S. M. Abbas, "*Fractal X-Shaped FSS Employed Electromagnetic Shield for X-band SATCOMs," 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting,* 2020, pp. 993-994.

[29] M. Idrees, S. Buzdar, S. Khalid and M. A. Khalid, "A Novel Miniaturized Frequency Selective Surface for EMI Shielding Applications," *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST),* 2021, pp. 1003-1006.

[30] I. S. Syed, Y. Ranga, L. Matekovits, K. P. Esselle and S. Hay, "*A Single-Layer Frequency-Selective Surface for Ultrawideband Electromagnetic Shielding,"* in *IEEE Transactions on Electromagnetic Compatibility*, vol. 56, no. 6, pp. 1404-1411, Dec. 2014

[31] S. Ünaldı, S. Çimen, G. Çakır and U. E. Ayten, "*A Novel Dual-Band Ultrathin FSS With Closely Settled Frequency Response,*" in *IEEE Antennas and Wireless Propagation Letters,* vol. 16, pp. 1381-1384, 2017

[32] Sarika, R. Kumar, M. R. Tripathy and D. Ronnow, "*Fractal frequency selective surface based band stop filters for X-band and Ku-band applications*," *2017 3rd International Conference on Advances in Computing,Communication & Automation (ICACCA) (Fall)*, 2017, pp. 1-4

[33] M. Nauman, R. Saleem, A. K. Rashid and M. F. Shafique, "*A Miniaturized Flexible Frequency Selective Surface for X-Band Applications*," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 2, pp. 419-428, April 2016

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Voice Spoofing Countermeasure Based on Spectral Features to Detect Synthetic Attacks Through LSTM

Gulam Qadir [1], Saima Zareen[1], Farman Hassan[1], Auliya Ur Rahman[1]
[1]Affiliation of authors University of Engineering and Technology Taxila, Punjab Pakistan
* Correspondence: Farman Hassan, Email ID: farmanhassan555@gmail.com.
Citation | Qadir.G, Zareen.S, Hassan.F, Rahman.A "Voice Spoofing Countermeasure Based on Spectral Features to Detect Synthetic Attacks Through LSTM". International Journal of Innovations in Science and Technology, Special Issue, pp: 153-165, 2022.
Received | Dec 15, 2021; Revised | Dec 29, 2021 Accepted | Jan 06, 2022; Published | Jan 07, 2022.

With the growing number of voice-controlled devices, it is necessary to address the potential vulnerabilities of Automatic Speaker Verification (ASV) against voice spoofing attacks such as Physical Access (PA) and Logical Access (LA) attacks. To improve the reliability of ASV systems, researchers have developed various voice spoofing countermeasures. However, it is hard for the voice anti-spoofing systems to effectively detect the synthetic speech attacks that are generated through powerful spoofing algorithms and have quite different statistical distributions. More importantly, the speedy improvement of voice spoofing structures is producing the most effective attacks that make ASV structures greater vulnerable to stumble on those voice spoofing assaults. In this paper, we proposed a unique voice spoofing countermeasure which is successful to hit upon the LA attacks (i.e., artificial speech and transformed speech) and classify the spoofing structures by the usage of Long Short-Term Reminiscence (LSTM). The novel set of spectral features i.e., Mel-Frequency Cepstral Coefficients (MFCC), Gammatone Cepstral Coefficients (GTCC), and spectral centroid are capable to seize maximum alterations present in the cloned audio. The proposed system achieved remarkable accuracy of 98.93%, precision of 100%, recall of 92.32%, F1-score of 96.01%, and an Equal Error Rate (EER) of 1.30%. Our method achieved 8.5% and 7.02% smaller EER than the baseline methods such as Constant-Q Cepstral Coefficients (CQCC) using Gaussian Mixture Model (GMM) and Linear Frequency Cepstral Coefficients (LFCC) using GMM, respectively. We evaluated the performance of the proposed system on the standard dataset i.e., ASVspoof2019 LA. Experimental results and comparative analysis with other existing state-of-the-art methods illustrate that our method is reliable and effective to be used for the detection of voice spoofing attacks.

Keywords: ASVspoof2019 LA dataset; Deep Learning; Spoofing countermeasure; Synthetic Speech; Voice anti-spoofing.

INTRODUCTION

ASV verifies the identity of users based on the voice presented to the ASV systems. Within the previous couple of years, we've witnessed a rapid evolution in voice biometrics-primarily based on user authentication. ASV structures are embedded in several gadgets inclusive of smart speakers (Amazon Alexa, Google home), and smartphones for the authentication in diverse software domains i.e., e-trade, banking structures, home automation, and special utility logins [1]. Google domestic gets speech commands from the users and performs several features which include putting reminders, text or name, remaining, and starting doorways, and unlocking cell telephones [2]. These applications are based on ASV systems [3]. Banking systems are also using voice-based authentication systems to verify customers such as BBVA's and Barclays Wealth have been using voice biometrics for verification of telephone callers. The Grant bank has developed a voice-based application, which allow customers to perform transactions simply by using voice commands [4].

We have witnessed an exponential and rapid growth in voice-driven authentication systems due to the Covid-19 pandemic. Social distancing and lockdown have limited the capability of facial, fingerprint, or iris recognition. This pandemic has urged the world to shift the verification measures based on a human to machine and human to human interactions to voice-based authentication systems [5]. Consequently, voice-primarily based authentication has emerged as a most possible and simple answer than every other biometrics approach which includes the iris, facial, and fingerprint. Moreover, voice-based authentication is considered to be the most economical and efficient than other biometrics methods [6]. ASV is playing a significant role in the biometric verification process. ASV uses the acoustic features of a person to authenticate the users [7]. However, intruders can mislead ASV systems by voice spoofing attacks i.e., voice replay, mimics or twins, voice conversion (VC), and syntenic speech [8,9]. Among these spoofing attacks, synthetic spoofing attacks (text-to-speech (TTS) and VC are threats to ASV systems that occur due to the rapid development of synthetic methods [10,11].

Stand-alone voice anti-spoofing techniques are developed to enhance the security and reliability of ASV systems. ASVspoof challenge series has been providing datasets [9,12,13] and provided their standard metrics for voice anti-spoofing speaker verification. In this work, we cognizance of voice anti-spoofing la attacks which include detecting true and spoofed speech produced with the aid of VC and TTS spoofing structures, detecting unseen attacks, and type of acknowledged assaults. Traditional techniques focus on feature engineering and hand-crafted features i.e., Cochlear filter cepstral coefficients (CFCCIF) [14], Linear cepstral coefficients (LFCC) [15], and Constant-Q cepstral coefficients (CQCC) [16] have shown better results against spoofing attacks. GMM is used in traditional methods [14-18] as a backend classifier.

The research community has also explored various deep learning methods [19-26] to detect LA attacks. In [19], deep learning models were investigated for anti-spoofing. Combining Convolutional Neural Network (CNN) and Recurrent Neural Network (RNNs) showed robustness against spoofing attacks. In [18], a deep residual community (ResNet) turned into used with temporal pooling. In [21], ResNet becomes employed with the most

margin of cosine loss and frequency masks augmentation. In [22], mild convolutional gate RNN become adopted to enhance the lengthy-time period dependency for the detection of voice spoofing assaults. In [23], a technique of characteristic genericization based on a mild CNN machine changed into a proposed that performed properly towards l. a. attacks. In [23], a technique of feature genericization based on a light CNN system was proposed that performed well against LA attacks. In [24], the transfer learning approach was explored with the ResNet network. The research community introduced model fusion based on sub-band modeling [25] and various features [26-28] to enhance the performance of voice anti-spoofing systems. In [29], three features such as MFCC, CQCC, and short-term Fourier transform (STFT) were integrated to detect voice spoofing attacks detection. ResNet was employed for classification purposes. It was observed that these three different variants of ResNet with MFCC, CQCC, and STFT produce better results than the baselines (CQCC-GMM [9] and LFCC-GMM [9]). In [30], two features such as CQCC and log power magnitude were used to design spoofing countermeasure. A deep neural network (DNN) was employed to discriminate the spoof and authentic audio. The DNN was based on a squeeze-excitation network and residual networks. This [30] framework yielded better results than the existing state-of-the-art methods, but the fusion of squeeze-excitation network and residual network substantially increased the training time. Few works [31-33] have used machine-learned features. In [31], two different deep learning models such as light convolutional neural network (LCNN) and gated recurrent neural network (GRNN) were used as features extractors from bonafide and spoof samples. Extracted functions had been used to train three gadget mastering algorithms which include SVM, linear discriminant analysis, and its probabilistic version (PLDA) for discriminating spoof and legitimate audio. In [32], DNN becomes hired to generate body-level posteriors and bottleneck capabilities to distinguish spoof and authentic audio. In [33], RNN based sequence level and DNN based frame-level features were used to design spoofing countermeasures for LA attacks. Different machine learning algorithms such as LDA, SVM, and gaussian density function were employed for classification purposes. Three model structures such as stacked autoencoder, multi-task joint learned DNN, and spoofing discriminant DNN was used for DNN based frame-level features. LSTM-RNN and BiLSTM-RNN were used for RNN based sequence-level features. These techniques achieved better classification results, but the computation cost was maximum.

In this work, we proposed a novel set of integrated spectral features for a voice anti-spoofing framework that improves the detection of unseen attacks, synthetic speech and classifies the cloning algorithms. The proposed system is robust to capture the alterations produced by the spoofing systems in speech signals. Experimental results show that our method outperformed all other existing state-of-the-art methods on the ASVspoof2019 LA dataset. The main contributions of our work are as under:

We proposed a novel set of integrated features, which better capture the maximum distortions created by the voice spoofing algorithms and traits of the speaker-induced variations in the genuine audio. The proposed system is capable of successfully classifying the spoofing systems and detecting unseen LA attacks. Our method is capable to improve the

security of automatic speaker verification systems against speech synthetic and voice converted attacks. We performed rigorous experiments on the ASVspoof2019 LA dataset to show the significance of the proposed system for the detection of LA attacks.

PROPOSED METHODOLOGY

This section presents a detailed description of the proposed voice spoofing countermeasure. The main objective of the proposed framework is to discriminate authentic and cloned audio, classify spoofing systems, and detect unseen LA attacks. The proposed system comprises two stages such as features extraction and classification. In the initial stage, the proposed spoofing countermeasure system takes audio as input and extracts three features i.e., 14-dim MFCC, 14-dim GTCC, and 1-dim spectral centroid. In the second stage, we employed LSTM for classification purposes. The detailed working mechanism of the proposed system is shown in Figure 1.

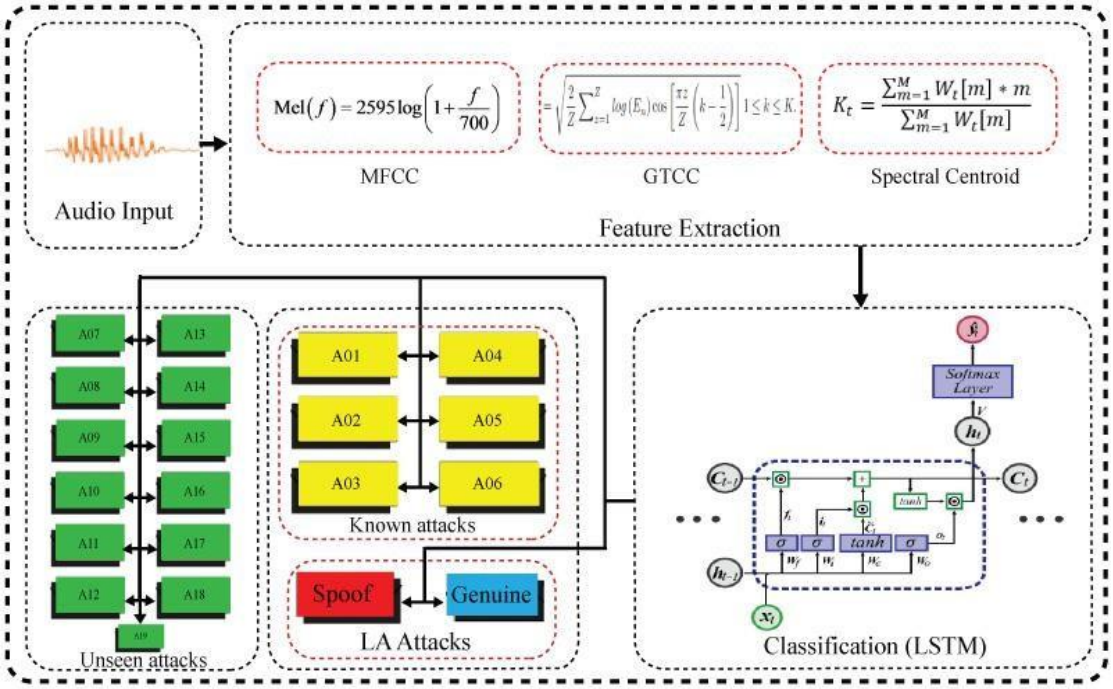


Figure 1. Proposed System.

Feature Extraction

To develop a robust voice spoofing countermeasure, we need to propose a robust audio feature descriptor that is capable of capturing the algorithmic artifacts from the spoof speech signals. For this purpose, we proposed a novel set of integrated spectral features that can extract highly discriminative information from the audio signals to accurately detect LA attacks, classify the spoofing systems, and detect the unseen voice spoofing attacks. The detailed feature extraction process is discussed below:

MFCC

MFCC takes voice as an input, and it calculates cepstral coefficients from it. The stepwise computation of MFCC from audio is explained in the below Figure 2. We extracted 14-dim MFCC features from audio by pre-emphasis. Pre-emphasis is used to compensate for the high-frequency part of the speech signal followed by the frame blocking in which speech signals are segmented into frames of 15-20ms to investigate the speech over a brief period. Next, each frame of audio is multiplied by the hamming window to keep the continuity between the frames, enhance harmonics, and minimize the edge effects. In the next step, we employed the Fast Fourier Transform (FTF) to get the magnitude spectrum of each frame. The power spectrum obtained after employing FFT is then mapped to mel-scale. Next, we multiplied the magnitude frequency by 40 triangular bandpass filters to get the log energy.

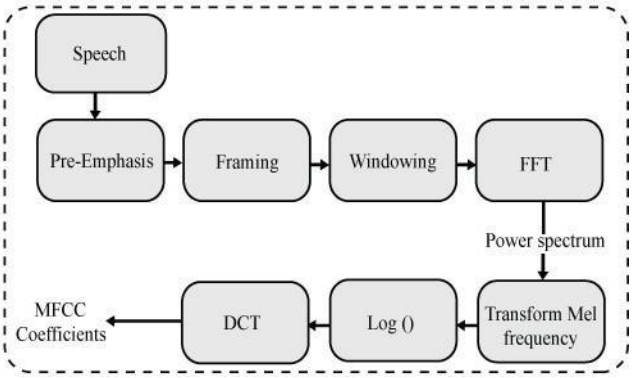


Figure 2. MFCC feature extraction process.

Later, we computed Discrete Fourier Transform (DFT) of the list of Mel log powers. Finally, the amplitude of the spectrums is selected as MFCCs.

GTCC

The computational process of the GTCC features is the same as the MFCC extraction scheme. We extracted 14-dim GTCC features from the audio. First, the speech signal is windowed into noticeably short frames of duration 10-50ms to examine over a short period. Making speech signals for a short duration has two purposes i.e., non-stationary signals are assumed to be stationary for such a limited interval of time, and features are extracted efficiently. Subsequently, we employed 48 GT filters to the signal of FFT to compute the energy of the sub-band. Finally, the log of each sub-band is computed followed by applying DCT. The GTCC features are computed as below:

$$GTCC_m = \sqrt{\frac{2}{N}} \sum_{n=1}^{N} \log(X_n) \cos\left[\frac{\pi n}{N}(m - \frac{1}{2})\right] 1 \leq m \leq M(5) \quad (1)$$

Where Xn is the energy of the speech signal in the nth spectral band, N is the number of GT filters and M is the GTCC.

Spectral Centroid

Spectral centroid is a measure that is used to characterize the spectrum in digital signal processing. It represents the mid-factor of the mass of the entire energy spectrum, in addition to the power distribution, across the excessive and occasional-frequency bands. The spectral centroid is computed as below:

$$W_e = \frac{\sum_{l=1}^{N} K_e[l] * l}{\sum_{l=1}^{N} K_e[l]} \quad (2)$$

The spectral centroid is computed as below:

Where the Kel is the magnitude of FT at current window e and frequency bin l.

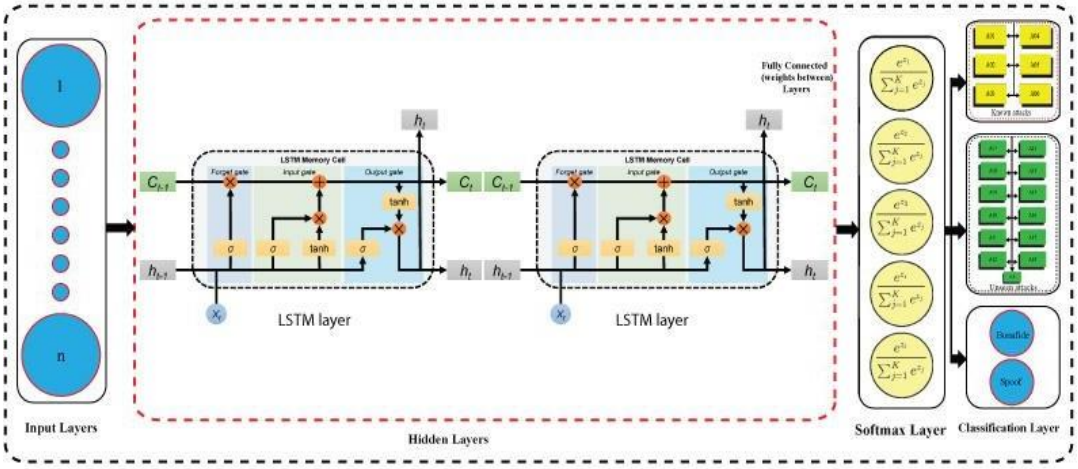


Figure 3. LSTM architecture.

Classification

Audio is a time-series data and LSTM is well suited to be used for the classification of time series data. Therefore, in this work, we also employed LSTM for classification purposes. We used different input parameters such as 2, 3, 4, 5, etc., layers, different hidden units such as 100, 200, 300, etc., and different optimizers such as adam, sgd, etc., to get better performance results. After using various configurations, we achieved better classification results on the following parameters: using an adam optimizer, two LSTM layers, 500 hidden units, mini-batch size of 64, and several epochs 25. Figure 3. shows the LSTM architecture being used for all the experimentation purposes.

RESULTS AND DISCUSSION

Dataset

We used the ASVspoof2019 LA dataset for experimentation purposes. This dataset consists of two different subset datasets i.e., ASVspoof2019 PA and ASVspoof2019 LA. PA contains samples of voice replay attacks, and the LA dataset contains synthetic and converted speech. Each dataset is further subdivided into three subsets i.e., training, development, and evaluation sets. The details of the ASVspoof2019 LA dataset are given in below Table 1.

Evaluation Metrics

To evaluate the performance of the proposed system, we used an accuracy, Equal error rate (EER), F1-score, precision, and recall. Countermeasures having lower EER values indicate better classification performance of the systems to detect spoofing attacks. We compared the performance of our method with baseline methods and other existing systems based on an EER value.

Table 1. Details of the ASVspoof2019 LA dataset.

Subset	Bonafide	Spoof	
	#Utterence	#Utterence	Attacks
Training	2,580	22,800	A01-A06
Development	2,548	22,296	A01-A06
Evaluation	7,355	63,882	A07-A19

Performance evaluation on known attacks

The objective of this experiment is to classify the cloning algorithms. Six different TTS and VC cloning algorithms i.e., A01, A02, A03, A04, A05, and A06 are used to generate spoof samples of the ASVspoof2019 LA dataset. These attacks are also called known attacks. There are 22,800 spoof samples of the training set and 22,296 development set that is generated by using these 6 cloning algorithms. Each algorithm-generated 3,800 samples of training and 3,716 of the development set. A01, A02, A03, and A04 are TTS while A05 and A06 are VC algorithms. We used the training samples for training the model and the development set for the testing purpose. From the results reported in Table 2, we can observe that the proposed system successfully classified all the spoofing systems. Moreover, our method performed well on A06 and achieved an accuracy of 99.70%, EER of 0.10%, a precision of 99.70%, recall of 100%, and an F1-score of 99.85%. The system performed second-best on A04 and achieved an accuracy of 99.40%, EER of 0.40%, the precision of 100%, recall of 98.80%, and F1-score of 99.39% while the proposed system performed worst on A03 and achieved an accuracy of 91.30%, EER of 5.93%, the precision of 92.63%, recall of 89.30%, and F1-score of 90.93%. The detailed results of the spoofing systems in terms of accuracy, EER, F1-score, precision, and recall are reported in Table 2. Overall, our system performed well and successfully detected all the cloning algorithms. From the results, we can conclude that the proposed system is robust to capture the variations in signals of spoof audios generated by cloning algorithms.

Table 2. Performance evaluation on the Cloning algorithms/spoofing systems.

Spoofing System	Accuracy %	EER %	F1-score%	Precision %	Recall%
A01	99.10	0.30	99.55	99.10	100
A02	95.60	3.23	95.42	91.24	100

A03	91.30	5.93	90.93	92.63	89.30
A04	99.40	0.40	99.39	100	98.80
A05	91.50	2.73	97.77	92.90	98.83
A06	99.70	0.10	99.85	99.70	100

Performance evaluation on unseen attacks

The objective of this experiment was to evaluate the performance of the proposed system to detect unseen LA attacks i.e., A07, A08, A09, A10, A11, A12, A13, A14, A15, A16, A17, A18, and A19. The evaluation set of the ASVspoof2019 LA dataset contains spoof samples of unseen attack types while the training and development sets contain spoofed samples of known attack types as discussed in section 3.3. We used 25,380 samples of the training set for training the model and 6,000 samples of unseen attacks for testing purposes. From the results reported in Figure4, it is observed that the proposed system performs well on A08 attacks and achieved an accuracy of 96.89%. Our method performed worst on the A17 unseen attack and achieved an accuracy of 80.01%. Overall, the system has reliable performance results on unseen attack types. The detailed results of our method to detect unseen attacks in terms of accuracy are given in Fig 4. From the effects pronounced in determine 4, we will take a look at that the proposed machine has correctly detected all unseen l. a. attacks. Our method is likewise dependable for use for the detection of unseen la spoofing assaults generated through the effective VC, TTS, VC-TTS spoofing algorithms. The proposed system is robust and capable to capture the algorithmic artifacts produced by VC, TTS, VC-TTS algorithms in bonafide audio with better accuracy.

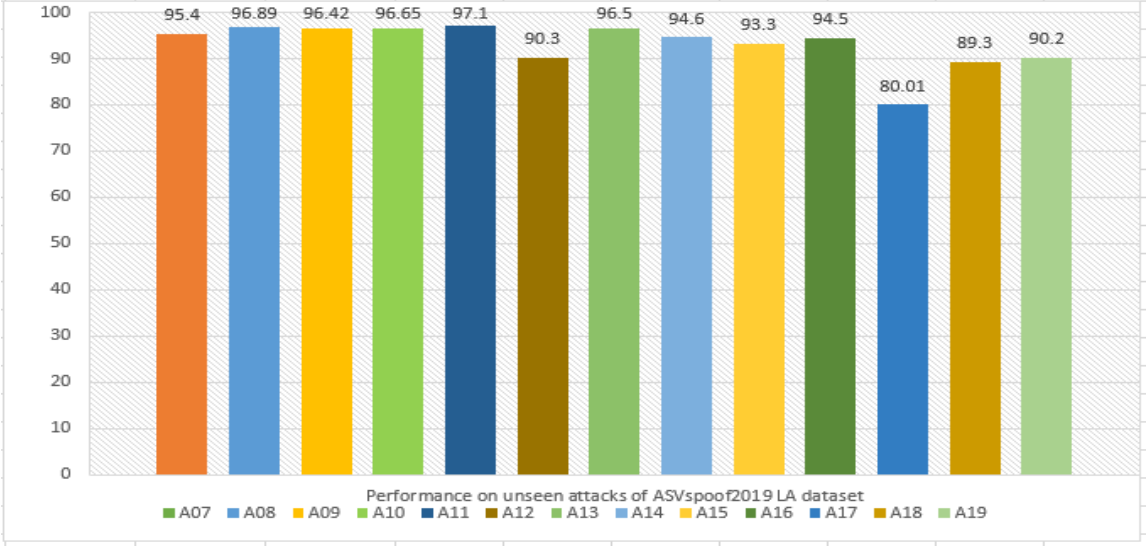


Figure 4. Performance evaluation on unseen attacks.

Performance evaluation on LA attacks

We designed this experiment to evaluate the performance of the proposed system to detect LA attacks. For this purpose, we extracted 29-dim spectral features comprising of (14-dim MFCC, 14-dim GTCC, and 1-dim SC) from the ASVSpoof2019 LA dataset. We used 20,005 samples of the training set for training the LSTM model and 18,483 samples of an evaluation set for testing the trained model. Experimental results are shown in Figure 5. The proposed method achieved an accuracy of 98.93%, EER of 1.07%, the precision of 100%, recall of 98.77%, and F1-score of 99.38%. The baseline method (CQCC-GMM [9] and LFCC-GMM [9]) achieved an EER of 9.57% and 8.09%, whereas, the proposed method (MFCC-GTCC-Spectral centroid-LSTM) achieved an 8.5% smaller EER than the baseline methods. From the results reported in Figure5, we can conclude that our method performs well in terms of accuracy and EER. The proposed system is capable of capturing the most discriminatory characteristics from audio signals of spoofed samples generated by TTS and VC algorithms. Experimental results signify the effectiveness of our method that can be implemented in ASV systems for the reliable detection of LA attacks.

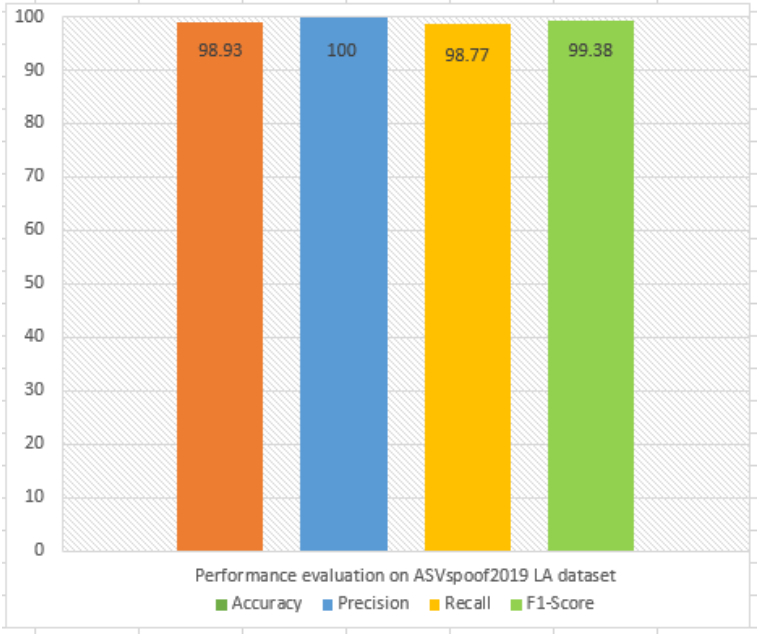


Figure 5. Performance evaluation on ASVspoof2019 LA dataset.

Comparison with other methods

We designed this experiment to check the effectiveness of the proposed system against other existing state-of-the-art methods [24, 16,18, 20, 23, 19, 21, 17]. As shown in Table 3, we compared the results of our method with the baseline [9], and other existing state-of-the-art methods to demonstrate the superiority of the proposed system. The EER and min-tDCF values of the existing state-of-the-art methods and our system are reported in Table 3. It can be observed that our method performed well against all the existing methods and baselines methods [9]. Chen et al. [21], perform the second-best by achieving an EER of 3.49% and

min-tDCF of 0.092, whereas, the baseline methods (CQCC-GMM and LCFF-GMM) [9] performed worst by achieving an EER of 9.57% and 8.09%. From the results reported in Table 3, we observe an accuracy gain of 8.5%, 7.02%, 6.59%, 5.31%, 5.21%, and 4.52% than the state-of-the-art methods [9,28,19,22,24]. These results signify the effectiveness and superiority of the proposed spoofing countermeasure. Experimental results and comparative analysis show that our voice anti-spoofing technique outperforms all the existing techniques in terms of EER and min-tDCF values. From these consequences, we can finish that the proposed machine can reliably be used for the detection of la attacks.

Table 3. Performance comparison with other methods.

System	EER%	min-tDCF
Baseline (CQCC-GMM) [9]	9.57	0.237
Baseline (LFCC-GMM) [9]	8.09	0.212
Chettri et al. [28]	7.66	0.179
Monterio et al. [19]	6.38	0.142
Gomez-Alanis et al. [22]	6.28	--
Aravind et al. [24]	5.32	0.151
Lavrentyeva et al. [27]	4.53	0.103
Wu et al. [23]	4.07	0.102
Tak et al. [25]	3.50	0.090
Chen et al. [21]	3.49	0.092
Proposed (MFCC-GTCC-Spectral Centroid-LSTM)	1.07	0.0343

CONCLUSION

This paper has presented a novel voice spoofing countermeasure to locate l. a. assaults. Additionally, the proposed gadget is capable to come across the spoofing systems (TTS and VC) that have been used to generate the spoofed samples of the ASVspoof2019 l. a. dataset. We proposed a novel set of integrated features, which captures maximum alterations and algorithmic artifacts present in speech signals. We employed LSTM for classification purposes to discriminate authentic and fake audio, classify the known attack types, and detect all the unseen LA attacks. Experimental results show that the proposed system outperformed the baseline and the existing state-of-the-art methods. The proposed system gives better classification results by achieving an 8.5% smaller EER than the baseline [9]. In the future, we aim to apply the proposed system on PA and deepfake datasets.

Acknowledgment. We are grateful to the ASVspoof organizers for providing the dataset.

Author's Contribution. The corresponding author equally contributed to this work.

Conflict of interest. Authors claim that there exists no conflict of interest for publishing this manuscript in IJIST.

REFERENCES

1. Voiceprint: The New WeChat Password. 2015 [cited 2021 04 April]; Available from: https://blog.wechat.com/2015/05/21/voiceprint-the-new-wechat-password/.
2. Millward, S. "Open Sesame: Baidu Helps Lenovo Use Voice Recognition to Unlock Android Phones". 2012 [cited 2020 12 March]; Available from: https://www.techinasia.com/baidu-lenovo-voice-recognition-android-unlock.
3. Vigderman, A. What Is Home Automation and How Does It Work? Mar 15, 2021 [cited 2021 April 16]; Available from: https://www.security.org/home-automation/.
4. Fernández, L. Efma recognizes Garanti Bank's mobile voice assistant. 2017 [cited 2020 3 March]; Available from: https://www.bbva.com/en/efma-recognizes-garanti-banks-mobile-voice-assistant/.
5. Access on 9.10.2021, Available online at: http://cobramoto.fi/phlms/growth-of-voice-assistants
6. Access on 9.10.2021, Available online at: lhttps://venturebeat.com/2021/07/14/how-voice-biometrics-is-saving-financial-services-companies-millions-and-eliminating-fraud/.
7. K. Delac and M. Grgic, "A survey of biometric recognition methods," in Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine. IEEE, 2004, pp. 184–193.
8. Z. Wu, N. Evans, T. Kinnunen, J. Yamagishi, F. Alegre, and H. Li, "Spoofing and countermeasures for speaker verification: A survey," speech communication, 2015. vol. 66, pp. 130–153,
9. M. Todisco, X. Wang, V. Vestman, M. Sahidullah, H. Delgado, A. Nautsch, J. Yamagishi, N. Evans, T. H. Kinnunen, and K. A. Lee, "ASVspoof 2019: Future horizons in spoofed and fake audio detection," Proc. Interspeech, pp. 1008–1012, 2019.
10. M. R. Kamble, H. B. Sailor, H. A. Patil, and H. Li, "Advances in anti-spoofing: from the perspective of ASVspoof challenges," APSIPA Transactions on Signal and Information Processing, 2020. vol. 9,
11. R. K. Das, T. Kinnunen, W.-C. Huang, Z.-H. Ling, J. Yamagishi, Z. Yi, X. Tian, and T. Toda, "Predictions of subjective ratings and spoofing assessments of voice conversion challenge 2020 submissions," in Proc. Joint Workshop for the Blizzard Challenge and Voice Conversion Challenge 2020, 2020, pp. 99–120.
12. Z. Wu, T. Kinnunen, N. Evans, J. Yamagishi, C. Hanilc¸i, M. Sahidullah, and A. Sizov, "ASVspoof 2015: the first automatic speaker verification spoofing and countermeasures challenge," in Sixteenth Annual Conference of the International Speech Communication Association, 2015, Conference Proceedings
13. T. Kinnunen, M. Sahidullah, H. Delgado, M. Todisco, N. Evans, J. Yamagishi, and K. A. Lee, "The ASVspoof 2017 challenge: Assessing the limits of replay spoofing attack detection," in Proc. Interspeech 2017, 2017, pp. 2–6. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2017-1111
14. T. B. Patel and H. A. Patil, "Combining evidence from Mel cepstral, cochlear filter cepstral and instantaneous frequency features for detection of natural vs. spoofed speech," in Sixteenth Annual Conference of the International Speech Communication Association, Conference Proceedings

15. M. Sahidullah, T. Kinnunen, and C. Hanilc¸i, "A comparison of features for synthetic speech detection," in Sixteenth Annual Conference of the International Speech Communication Association, 2015.
16. M. Todisco, H. Delgado, and N. Evans, "A new feature for automatic speaker verification anti-spoofing: Constant Q cepstral coefficients," in Proc. Odyssey, 2016, vol. 45, pp. 283–290. [Online]. Available: http://dx.doi.org/10.21437/Odyssey.2016-41
17. L. Wang, Y. Yoshida, Y. Kawakami, and S. Nakagawa, "Relative phase information for detecting human speech and spoofed speech," in Sixteenth Annual Conference of the International Speech Communication Association, Conference Proceedings.
18. J. Sanchez, I. Saratxaga, I. Hernaez, E. Navas, D. Erro, and T. Raitio, "Toward a universal synthetic speech spoofing detection using phase information," IEEE Transactions on Information Forensics and Security, 2015.vol. 10, no. 4, pp. 810–820,
19. Zhang, C. Yu, and J. H. Hansen, "An investigation of deep-learning frameworks for speaker verification antispoofing," IEEE Journal of Selected Topics in Signal Processing, 2017. vol. 11, no. 4, pp. 684–694,
20. J. Monteiro, J. Alam, and T. H. Falk, "Generalized end-to-end detection of spoofing attacks to automatic speaker recognizers," Computer Speech & Language, 2020.p. 101096
21. T. Chen, A. Kumar, P. Nagarsheth, G. Sivaraman, and E. Khoury, "Generalization of audio deepfake detection," in Proc. Odyssey the Speaker and Language Recognition Workshop, Conference Proceedings, 2020, pp. 132–137.
22. Gomez-Alanis, A. M. Peinado, J. A. Gonzalez, and A. M. Gomez, "A light convolutional GRU-RNN deep feature extractor for ASV spoofing detection," Proc. Interspeech, 2019. pp. 1068–1072,
23. Z. Wu, R. K. Das, J. Yang, and H. Li, "Light convolutional neural network with feature genuinization for detection of synthetic speech attacks," Proc. Interspeech, , 2020.pp. 1101–1105
24. P. Aravind, U. Nechiyil, N. Paramparambath, et al., "Audio spoofing verification using deep convolutional neural networks by transfer learning," arXiv preprint arXiv:2008.03464, 2020.
25. H. Tak, J. Patino, A. Nautsch, N. Evans, and M. Todisco, "Spoofing attack detection using the non-linear fusion of sub-band classifiers," Proc. Interspeech, 2020. pp. 1106–1110
26. X. Tian, Z. Wu, X. Xiao, E. S. Chng, and H. Li, "Spoofing detection from a feature representation perspective," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 2119–2123.
27. Lavrentyeva, S. Novoselov, A. Tseren, M. Volkova, A. Gorlanov, and A. Kozlov, "STC antispoofing systems for the ASVspoof2019 challenge," Proc. Interspeech, , 2019.pp. 1033–1037
28. B. Chettri, D. Stoller, V. Morfi, M. A. M. Ram´ırez, E. Benetos, and B. L. Sturm, "Ensemble models for spoofing detection in automatic speaker verification," in Proc. Interspeech, 2019, pp. 1018–1022. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2019-2505
29. Alzantot, M., Z. Wang, and M.B. Srivastava, "Deep residual neural networks for audio spoofing detection". arXiv preprint arXiv:1907.00501, 2019.

30. Lai, C.-I., et al., "ASSERT: Anti-spoofing with squeeze-excitation and residual networks". arXiv preprint arXiv:1904.01120, 2019.
31. Gomez-Alanis, A., et al., "A Light Convolutional GRU-RNN Deep Feature Extractor for ASV Spoofing Detection". Proc. Interspeech 2019, 2019: p. 1068-1072.
32. Alam, M.J., et al. "Spoofing Detection on the ASVspoof 2015 Challenge Corpus Employing Deep Neural Networks". in Odyssey. 2016.
33. Qian, Y., N. Chen, and K. Yu, "Deep features for automatic spoofing detection. Speech Communication", 2016. 85: p. 43-52.

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Automatic Vehicle Number Plate Recognition Approach Using Color Detection Technique

Muhammad Ayaz[1], Dr. Said Khalid Shah[1], Dr. Muhammad Javed[1], Muhammad Assam[2], Wasiat Khan[3], Fahad Najeeb[1]
[1] Department of Computer Science University of Science and Technology Bannu, Pakistan.
[2] College of Computer Science and Technology Hangzhou 310027, China.
[3]Department of Software Engineering, University of Science and Technology, Bannu, KP Pakistan
*Muhammad Ayaz Email: ayazkhanttt@gmail.com
Citation|Ayaz.M, Shah.S.K, Javed.M, Assam.M, Khan.W, Najeeb.F "*Automatic Vehicle Number Plate Recognition Approach Using Color Detection Technique*". International Journal of Innovations in Science and Technology, Special Issue, pp: 166-176, 2022.
Received | Dec 07, 2021; Revised | Dec 28, 2021 Accepted | Dec 31, 2021; Published |8 Jan, 2022.

An Automatic Vehicle Number Plate Recognition System (AVNPR) is a key research area in image processing. Various techniques are developed and tested by researchers to improve the detection and recognition rate of AVNPR system but faced problems due to issues such as variation in format, lighting conditions, scales, and colors of number plates in different countries or states or even provinces of a country. Douglas Peucker Algorithm for shape approximation has been used in this research to detect the rectangular contours and the most prominent rectangular contour is extracted as a number plate (NP) and the connected component analysis is used to segment the characters followed by optical character recognition (OCR) to recognize the number plate characters. A custom dataset of 210 vehicle images with different colors at various distances and lighting conditions was used for the proposed method captured on my smart phone Galaxy J7 Model SM-j700F at roads and parking. The dataset contains various types of vehicles (i.e. Trucks, motorcars, mini-buses, tractors, pick-ups etc). The proposed method shows an average result of 95.5%. The novelty used in this method is that it works for different colors simultaneously because in Pakistan, several colors are used for vehicle NPs.

Keywords: Number plate recognition, color detection, HSV color, Douglas Peucker Algorithm, OCR

INTRODUCTION

The number plate recognition system is a crucial factor of an intelligent transportation system. In this system, vehicles are recognized based on their number plates. It has a variety of applications, including a border management system, automatic car parking system, automatic toll collection on highways and enforcement of traffic laws [1, 2, 3]. It is also used to detect stolen vehicles, fake NPs holders, and non-standard number plates. An intelligent gate management system is also based on this criterion to restrict the entry of unauthorized

vehicles into an organization [4, 5]. It has wide applications to recognize high-valued vehicles (i.e., government officials, police vans, school buses, ambulances, etc.) to provide a safe way to them in bulk of vehicles and can be used to manage one-way roads and restrict heavy trucks to enter cities and towns [2].

In general, the number plate recognition system consists of four steps [6, 7]. In the first step, the input image is preprocessed for noise removal and foreground enhancement. In the second step, the region of interest (ROI) is detected and extracted. In the third step, the characters are segmented, and in the fourth step, recognition is done. Researchers adopt different approaches and each approach has its pros and cons. The majority of systems are limited to some specific formats and territories. There are two main approaches used to recognize the number plates; traditional image processing approach and the deep learning approach. In the traditional image processing approach, simple mathematical functions are used to locate and extract the number plate and character segmentation is used to save the characters as separate images. Template matching is used to recognize the characters and convert them into text form.

In the deep learning approach, a supervised learning strategy is used. An object detection model is developed and trained using a relevant dataset to recognize the desired object in a video or image. Nowadays deep learning-based models are used widely for the number plate detection and recognition with high accuracy. Hundreds and thousands of image based datasets are used with proper annotation to train the model and it is not an easy task that need high power systems with GPUs to work smoothly. A model trained for one format may not be compatible for other formats of NPs. On the other hand, in traditional image processing approach, the localization/detection ratio of NPs does not work well in natural scene images and videos having noise and bad illumination and it is the main step of AVNPR system and the remaining steps are highly dependent on this step [8][9].

This section contains different methods obtained from literature survey and are used for number plate detection and recognition. *Pavani and Mohan* [1] used simple image processing techniques and connected component analysis (CCA) was used for plate extraction and again CCA was used to segment the characters. *Sharma et al.* [2] used a Sobel edge detector to detect the foreground edges and thresholding technique was used for image binarization to locate the number plates. Morphological operations were used to enhance the plate area for extraction and minimization of the non-plate locations. Template matching was used to recognize the number plate characters. This system was developed for the specific format of Nepali number plates. *Yogheedha et al.* [8] used Otsu's thresholding and background subtraction to locate the number plate and Built-in OCR was used to recognize the characters. Out of 14 images, the system recognized 13 images of the car. It was developed as a training project for Perlis University of Malaysia's intelligent gate management system to restrict the entry of private vehicles. *Kulkarni et al*. [9] developed a model for helmet detection of motorcycle riders and also number plate recognition system was used to find those riders having no helmet. They used thresholding and morphological operations for plate detection and OCR was used to recognize the plate. *Wang et al.* [10] used the color detection method to detect the red lights and then the middle of the lights was extracted as the region of interest. It worked for the rear NP of a vehicle having red color backlights. *Babbar et al.* [11] used preprocessing techniques

and binarization to filter out the strong objects. CCA was used to locate NP and aspect ratio to extract the NP successfully and OCR was used for NP recognition.

Rajathilagam et al. [12] used traditional image processing techniques to develop the system. The input image was converted to grayscale and dynamic adaptive thresholding was used for NP localization. Morphological operation dilation was used to highlight the region of interest and erosion was used to eliminate the false NP areas. OCR was used to recognize the plate characters. That method worked only for very clear images which are taken at a very short distance. *Indira et al*. [13] used Euclidian distance to locate the possible NP region with the predefined threshold value. The stroke width was used with maximally stable extreme regions (MSER) to detect and extract the NP and template matching was used to recognize the plates. *Haider and Khuram* [14] used simple morphological operations to detect the number plate. After morphological operations and image filtering, CCA was used to extract the number plate and feature matching was used for character recognition. *Hendry and Chen [15]* used the sliding window technique with a deep learning-based object detection model "You Only Look Once (YOLO)" to detect and recognize NPs of Taiwan motor car. They used AOLP dataset for training the model having 2049 images of Taiwan LPs.*Babu. et al.* [16] used a custom data set of 6500 car NPs images with annotation and trained Darknet framework YOLO v3 for NPs detection [17, 18]. The same model was trained again for character recognition of plates with 91% accuracy. *Pustokhina et al.* [19] used the improved Bernsen method to enhance the image illumination. Connected Component Analysis was used to detect the number plates. Optimal k-means clustering was used to segment the characters and Convolutional Neural Network was used to recognize the characters. *Sahoo et al.* [20] used preprocessing techniques to enhance the image quality and the prewitt edge detector was used to detect the vertical edges. Thresholding technique was used to segment the images to remove the weak blobs. Morphological operations dilation and erosion were used to locate the vehicle number plate. Aspect ratio was used to extract the rectangular blob as expected number plate region. The extracted image (NP) was enhanced using histogram equalization to improve the shapes of characters and CCA was used to segment the characters. K-nearest neighbor and distance tree classifier were used to recognize the characters. *Tenzin et al*. [21] used preprocessing for image enhancement and Robert's edge vertical detection with a fixed value of thresholding (0.0375) to filter the image for NP and the rectangular shape was extracted as NP. The bounding box method was used to segment the characters and template matching was used to recognize the plate characters. *Tourani et al. [16]* used histogram equalization for image enhancement and trained Darkent framework YOLOv3 for NP detection and again the same model was trained for character recognition. A custom dataset of Iranian vehicles with proper annotation was used for training the model. *Pustokhina et al.,* [19] developed a model for mobile application using Tinny YOLO for NP detection and again convolutional neural network (CNN) was used for character recognition. *Saif et al.* [20] developed and trained a model for Bangladeshi standard NPs format using YOLOv3 for NP detection using custom dataset and again Darknet framework YOLOv3 was trained with a dataset of characters to recognize the plate characters. *Henry et al.* [21] used Tiny YOLOv3 for NP detection and direct recognition without segmentation was done using YOLOv3-SPP (Spatial Pyramid Pooling) model. *Qadri and Asif* [22] used a color-based technique to extract the NPs having a yellow color. After preprocessing, the input image was searched for yellow pixels and set to 1 and other than

yellow pixels were set 0 to extract the NP region easily. OCR was used to recognize the characters. *Khan et al.* [23] used LAB color space transformation and L-channel was selected and contours were drawn. Otsu's thresholding was used after preprocessing to segment the ROI. Morphological operation erosion and dilation were used and eroded image was subtracted from dilated image to highlight NP's characters and 2d convolution was used to extract the NP region. Support Vector Machine was used to recognize the characters.

M Rizwan Asif et al. [24] used traditional image processing techniques to detect and recognize vehicle number plates using blue color properties. Image was converted from RGB to YDbDr color space to detect the blue color and intensity adjustment was used to eliminate low intensity objects. Thresholding was used for image segmentation and CCA was used to detect the candidate regions and aspect ratio was used to detect the expected plate among the list of ROIs. [25] Converted RGB image to HSI format for color segmentation. Morphological operations were used to highlight the ROI. Connected component analysis was used with aspect ratio to filter the rectangular shape (expected LP) among the existing blobs. *F. S. Khan et al*. [26] combined shape and color properties based on late fusion to detect objects. [27] Used two different approaches and comparatively studied the results. Initially, they used color-based approach to detect the LP of vehicles and then used the shape-based LP detection method.

In this research, a color-based NP detection technique with shape approximation is used. The acquired image is converted to HSV color format to filter the image for NP localization and extraction. In the previous literature, color base techniques were utilized for single color [28]. The novelty used in this method is that it works for different colors simultaneously because in Pakistan several colors are used for vehicle NPs. After detecting the NP, we proceed with the recognition step.

PROPOSED METHOD

In this study, the proposed method was discussed in detail **Figure 1**. showed the flow diagram of the proposed system.

.

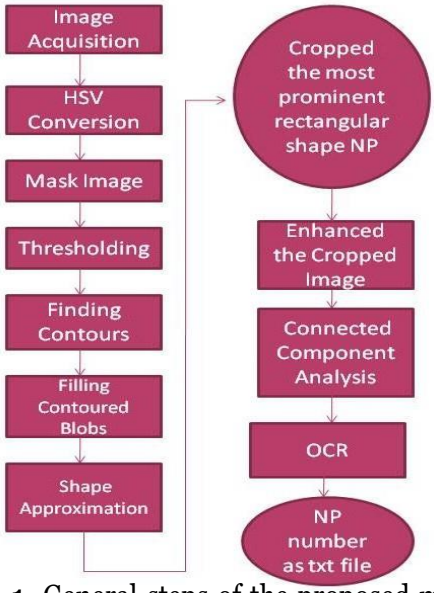


Figure 1. General steps of the proposed method

In this research, traditional image processing techniques were used to detect and extract the number plate region. Connected component analysis was used to segment the characters and built-in OCR was used to recognize the characters. The input image was converted to HSV color format and the mask was generated. Binarization and morphological operations were used to locate the strong blobs and Ramer Dauglous Peacuker shape approximation algorithm was used with aspect ratio to detect the rectangular shapes and the most prominent rectangular shape was obtained as NP area. The extracted NP image was enhanced using pre-processing operations and connected component analysis was used to segment the plate characters and built-in OCR was used to recognize the plate. Three primary steps i.e. extraction, segmentation and recognition were elaborated in detail as below.

Number Plate Detection and Extraction

The input image was resized and pre-processing techniques were used to enhance the quality of the input image. It was then converted from RGB color format to HSV color format to easily filter the desired color intensity using a NumPy array of lower and upper bound values. Separate lower and upper bound values were used for each NP's color detection because each color had a separate range of values. The mask image was generated as shown in Figure. 2.

Figure 2. (a) Input image with yellow plate, (b) Mask image, (c) Input Image with green Plate and (d) Mask image

The image was converted to grayscale and thresholding was used to eliminate the non-plate regions because in many cases, there is a probability of existence of some non-plate regions. As the thresholding operation was used to remove the non-plate regions but it also erases some portion of ROI called outlines of number plate and it was difficult to take it as a single contour region. Morphological operation gradient was used to strengthen the outlines of the region of interest (ROI) that was used to highlight the object to contour the ROI easily. Edge detection method was used to find the edges and contour function was used to find all the contours.

Shape Approximation: In most of the cases the largest contour contains vehicle number plate but the problem arises when the input image contains a non-plate blob/area which is

larger than the NP area with color intensities similar to the region of interest [29, 30]. To eliminate this problem Ramer Dauglous Peacuker shape approximation algorithm was used. Before shape approximation method, the internal region of each contour is filled completely to approximate the shape accurately. Without filling the blob, shape approximation function approximates each NP character as separate contour.

Ramer Dauglous Peacuker algorithm developed in 1973 by David H. Douglas and Thomas K. Peucker was widely used for geometrical shape approximation [31, 32]. This algorithm converts a curve/contour composed of a linear line segment to a similar curve by reducing number of points. The points on the converted curve are nothing but subset of the points that defines the original curve. The algorithm treats each curve/contour C as an order set of points and the distance dimension is $\varepsilon > 0$.

$$C= (p_1, p_2,.....p_n) \quad (1)$$

First of all the algorithm automatically marks the first and the last points to be kept. The line segment obtained from p_1 to p_n by joining the kept points is the first approximation of the curve C. The algorithm then searches the point in the inner points that has largest distance from the line segment p_1 to p_n. If the distance of the largest point is smaller than ε , then simplification is been done and all the inner points can be put away without the simplification curve being worse than ε . If the distance of the largest distance point is greater than ε then the largest distance point must be kept. The algorithm repeats itself with the first point and largest distance point and with the largest distance point and end point. When the repetition is complete, the output curve represents the set of all kept point called shape approximation [33].

Character Segmentation

After successful detection and extraction of the number plate region, the next step is character segmentation. This step is mainly used to prepare the extracted plate image as an input to the recognition step to easily recognize the characters. In this step, the input image was resized and converted to the grayscale image. Thresholding was used for image binarization. Connected Component Analysis was used to declare each area composed of similar and connected pixels as an individual object. Width, height and area properties were used for each connected component to filter out the objects that are too small or too large.

Character Recognition

This step was used to read the plate image and convert the characters and digits into readable form called character recognition. There are different approaches used for this purpose. In this system a built-in OCR based on deep learning called Pytesseract was used for character recognition [34, 35].

Results and Discussion

The proposed method is based on color intensities with shape approximation. Due to the prominent colors used in our country for NPs it becomes easy to detect the ROI easily using color based segmentation and in more than 80% cases, the plates using their color intensities are easily distinguished shown in **Figure 3**.

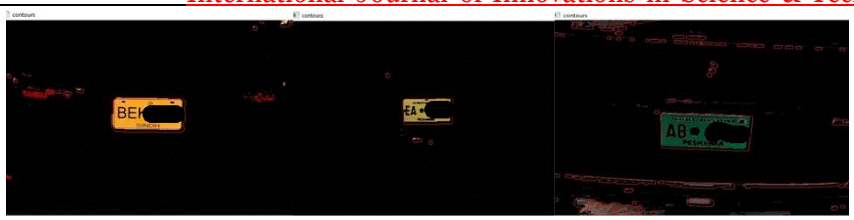


Figure 3. Shows some sample images after finding contours

To increase the accuracy rate we used shape approximation in the second phase to overcome the problems that shown in **Figure 4.**

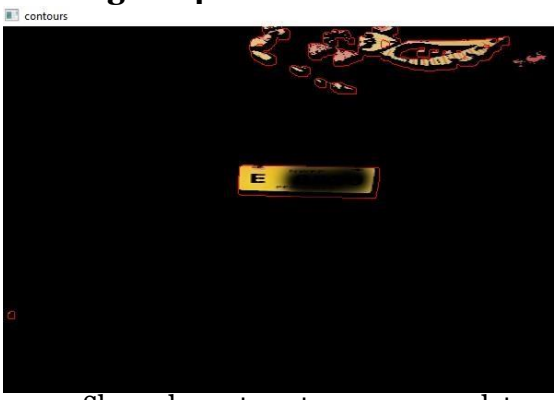


Figure 4. Shows largest contour as non –plate area

There are three main colors of NPs in Pakistan used in different territories/provinces for the last decade. The green color is used for government vehicles. For public vehicles, the yellow color is used in Khyber Pakhtunkhwa and Sindh provinces; green-yellow is used in Punjab and Baluchistan provinces. Sample images of Number Plates used in our country are shown in Figure. 5.

Figure 5. Sample Images of NP

A custom dataset of 210 vehicle images with different colors at various distances and lighting conditions was used for the proposed method captured on my smart phone Galaxy J7 Model

SM-j700F at roads and parking. The dataset contains various types of vehicles (i.e. Trucks, motorcars, mini-buses, tractors, pick-ups etc). The proposed method shows an average result of 95.5%. Table 1 shows the result of extraction and recognition of all the four types of NPs.

Table 1. Shows the results of the proposed system

Type	Items	Extraction	Recognition
Yellow	90	88	83
Green	50	46	43
Green_Yellow	70	67	65
Total	210	201	191

There is no public dataset of Pakistani vehicle NPs as per the knowledge obtained by the survey and efforts were also made for comparative study. The dataset used for comparisons with some state-of-the-art methods are shown in Table 2.

Table 2. Comparisons

S#	Methods	Year	Accuracy
1	Qadri and Asif [28]	2009	74 %
2	Haider and Khurshid [10]	2017	85 %
3	K. Deb et al. [31]	2017	91 %
4	Babar et al. [7]	2018	87 %
5	Pustkkhina et al. [14]	2020	93 %
6	**Proposed**	**2021**	**95.5 %**

CONCLUSION AND FUTURE WORK

There is no public dataset of Pakistani vehicle NPs as per the knowledge obtained by the survey and efforts were also made for comparative study. The number plate recognition system is not used in Pakistan currently due to various factors. Still this research article proposed an efficient method for the development on an automatic number plate recognition system in Pakistan. Despite being trained on a small dataset, it can show better and accurate results comparable to other number plate's recognition systems. The proposed method may provide a base for the implementation of an automatic number plate recognition system in Pakistan. In the proposed research article, simple image processing techniques are used to first extract the NP, and then OCR is used to recognize the plate. The extraction rate of the plate may be increased by imposing rules on the drivers to keep the number plates safe from dust and scratches.

Acknowledgement

The Author says thanks to Mr. Waheed Ullah Post Graduate Student at department of Mathematics University of Science and Technology Bannu for his valuable suggestions and support regarding the implementation of Ramer Dauglous Peacuker Algorithm.

Conflict of Interest

It is declared by the authors that this manuscript has no conflict of interest.

REFERENCES

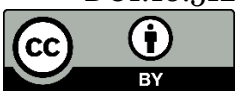
1. T. Pavani and DVR Mohan, "*Number Plate Recognition by using open CV- Python''*, International Research Journal of Engineering and Technology (IRJET) Mar 2019, Volume: 06 Issue: 03.
2. Sharma G "*Performance Analysis of Vehicle Number Plate Recognition System Using Template Matching*", 2018. T. Pavani and DVR Mohan, "*Number Plate Recognition by using open CV-Python''*, International Research Journal of Engineering and Technology (IRJET) Mar 2019, Volume: 06 Issue: 03.
3. Sharma G "*Performance Analysis of Vehicle Number Plate Recognition System Using Template Matching*", 2018.
4. W. Weihong and T. Jiaoyang, "*Research on License Plate Recognition Algorithms Based on Deep Learning in Complex Environment*," *In IEEE Access*, vol. 8, pp. 91661-91675, 2020, Article (CrossRef Link).
5. R. Naren Babu, V. Sowmya and K. P. Soman, "*Indian Car Number Plate Recognition using Deep Learning,"* In Proceedings of the 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies *(ICICICT),* Kannur Kerala, India, pp. 1269 1272, 2019. Article (CrossRef Link).
6. Izidio, D.M.F., Ferreira, A.P.A., Medeiros, H.R. et al. "*An embedded automatic license plate recognition system using deep learning*". Des Autom Embed Syst 24, 23–43 (2020). https://doi.org/10.1007/s10617-019-09230-5
7. Techniques. J Inform Tech Softw Eng 8: 232. Article (CrossRef Link)
8. K. Yogheedha, A. S. A. Nasir, H. Jaafa and S. M. Mamduh, "*Automatic Vehicle License Plate Recognition System Based on Image Processing and Template Matching Approach*", International Conference on Computational Approach in Smart Systems Design and Applications *(ICASSDA),* ,pp. 1-8, 2018. Article (CrossRef Link)
9. Y. Kulkarni, S. Bodkhe, A. Kamthe and A. Patil, "*Automatic number plate recognition for motorcyclists riding without helmet*", 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, 2018, pp. 1-6, Article (CrossRef Link)
10. Wang, J., Bacic, B. & Yan, W.Q. "*An effective method for plate number recognition*". Multimed Tools Appl 77, 1679 1692 (2018).Article (CrossRef Link)
11. S. Babbar, S. Kesarwani, N. Dewan, K. Shangle and S. Patel, *"A New Approach for Vehicle Number Plate Detection*," 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, 2018, pp. 1-6, Article (CrossRef Link)
12. R. Rajathilagam, K. Sivamani, R. Seetharaman and D. Nedumaran, "*Neural Network based Vehicle Number Plate*". *"Recognition System*," 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC), Chennai, India, 2019, pp. 102-104, Article (CrossRef Link)
13. Indira K., Mohan K.V., Nikhilashwary T. Bhattacharyya S., Mukherjee A., Bhaumik H., Das S.,Yoshida K (2019) "*Automatic License Plate Recognition*". In: (eds) Recent Trends in Signal and Image Processing. Advances in Intelligent Systems and Computing, vol 727. Springer, Singapore. Article (CrossRef Link)
14. S. A. Haider and K. Khurshid, "An implementable system for detection and recognition of license plates in Pakistan," *In Proceedings of the International Conference on*

Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, pp. 1-5, 2017. Article (CrossRef Link)

15. S. Tenzin, P. Dorji, B. Subba and T. Tobgay, "*Smart Check-in Check-out System for Vehicles using Automatic Number Plate Recognition,"* In Proceedings of the 11th International Conference on Computing, Communication and Networking Technologies *(ICCCNT)*, Kharagpur, India, pp. 1-6, 2020, Article (CrossRef Link).
16. Hendry, Rung-Ching Chen,Automatic License Plate Recognition via sliding-window darknet YOLO deep learning,Image and Vision Computing,Volume 87,2019,Pages 47-56, ISSN02628856. Article (CrossRef Link).
17. A. Tourani, A. Shahbahrami, S. Soroori, S. Khazaee and C. Y. Suen, "*A Robust Deep Learning Approach for Automatic Iranian Vehicle License Plate Detection and Recognition for Surveillance Systems*," *IEEE Access*, vol. 8, pp. 201317-201330, 2020, Article (CrossRef Link).
18. S. Zhang, G. Tang, Y. Liu and H. Mao, "*Robust License Plate Recognition With Shared Adversarial Training Network*," *in IEEE Access*, vol. 8, pp. 697-705, 2020, Article (CrossRef Link)
19. I. V. Pustokhina et al., "*Automatic Vehicle License Plate Recognition Using Optimal K-Means With Convolutional Neural Network for Intelligent Transportation Systems*," *IEEE Access*, vol. 8, pp. 92907-92917, 2020. Article (CrossRef Link)
20. Ashok Kumar Sahoo, "*Automatic recognition of Indian vehicles license plates using machine learning approaches*", Materials Today: Proceedings, 2020, ISSN 2214-7853. Article (CrossRef Link)
21. A. Kumar and D. Verma, "*Number Plate Reorganization using Image Processing and machine learning Approaches: A Review*," International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, pp. 1-4, 2020. Article (CrossRef Link).
22. https://opencvpythontutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_colorspaces/py_colorspaces.html#convertingcolorspas
23. W. Wang, J. Yang, M. Chen and P. Wang, "*A Light CNN for End-to-End Car License Plates Detection and Recognition*,"*in IEEE Access*, vol. 7, pp. 173875-173883, 2019. Article (CrossRef Link)
24. Al-Mahbashi L.T.A., Abu Talip Yusof N., Shaharum S., Abdul Karim M., Mohd Faudzi A. , "*Development of Automated Gate Using Automatic License Plate Recognition System*," In Proceedings of the 10th National Technical Seminar on Underwater System Technology, Singapore, 2018.. Article (CrossRef Link)
25. N. Saif et al., "*Automatic License Plate Recognition System for Bangla License Plates using Convolutional Neural Network*," In Proceedings of the TENCON Conference (TENCON), Kochi, India, , pp. 925-930, 2019. Article (CrossRef Link)
26. C. Henry, S. Y. Ahn and S. -W. Lee, "*Multinational License Plate Recognition Using Generalized Character Sequence Detection*," *IEEE Access*, vol. 8, pp. 35185-35199, 2020, Article (CrossRef Link).
27. Techniques. J Inform Tech Softw Eng 8: 232. Article (CrossRef Link) K. Yogheedha, A. S. A. Nasir, H. Jaafa and S. M. Mamduh, "*Automatic Vehicle License Plate Recognition System Based on Image Processing and Template Matching Approach*", International

Conference on Computational Approach in Smart Systems Design and Applications *(ICASSDA)*, ,pp. 1-8, 2018. Article (CrossRef Link)

28. Y. Kulkarni, S. Bodkhe, A. Kamthe and A. Patil, "*Automatic number plate recognition for motorcyclists riding without helmet*", 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, 2018, pp. 1-6, Article (CrossRef Link)
29. M. T. Qadri and M. Asif, "*Automatic Number Plate Recognition System for Vehicle Identification Using Optical Character Recognition,*" International Conference on Education Technology and Computer, Singapore, pp. 335-338, 2009. Article (CrossRef Link)
30. M-A. Khan, M. Sharif, M-Y. Javed, T. Akram, M. Yasmin, T. Saba, "*License number plate recognition system using entropy based features selection approach with SVM*," vol. 12, issue 2, p. 200 – 209, February, 2018. Article (CrossRef Link)
31. M R Asif, Q Chun, S Hussain, M S Fareed,"*Multiple licence plate detection for Chinese vehicles in dense traffic scenarios*"IET Intelligent Transport Systems.doi:10.1049/ietits.2016.0008
32. K. Deb, M. K. Hossen, M. I. Khan and M. R. Alam, "*Bangladeshi Vehicle License Plate Detection method based on HIS color model and geometrical properties*," 2012 7th International Forum on Strategic Technology (IFOST), 2012, pp. 1-5, doi: 10.1109/IFOST.2012.6357650.
33. F. S. Khan, R. M. Anwer, J. van de Weijer, A. D. Bagdanov, M. Vanrell and AM. Lopez,"*Color attributes for object detection,*" 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 3306-3313, doi:10.1109/CVPR.2012.6248068.
34. Hidayatullah, Priyanto; Zuhdi, Miftahuddin."*Color-Texture Based Object Tracking Using HSV Color Space and Local Binary Pattern*" International Journal on Electrical Engineering and Informatics; Bandung Vol. 7, Iss. 2, (Jun 2015): 161-174
35. LA Elrefaei, MO Al-musawa, NA Al gohany,"*Development of An Android Application for Object Detection Based on Color, Shape, or Local Features*". The International Journal of Multimedia & Its Applications (IJMA) Vol.9, No.1, February 2017 DOI:10.5121/ijma.2017.9103



OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

A Comparative Analysis of Camera, LiDAR and Fusion Based Deep Neural Networks for Vehicle Detection

Shafaq Sajjad[1], Ali Abdullah[1], Mishal Arif[1], Muhammad Usama Faisal[1], Muhammad Danish Ashraf[2], Shahzor Ahmad[1]
[1]College of Electrical & Mechanical Engineering, NUST.
[2]Synapsify Systems.
* Correspondence: Shafaq Sajjad,shafaqsajjad95@gmail.com
Citation | Sajjad. S, Abdullah. A, Arif. M, Faisal. M. U, Ashraf. M. D, Ahmad. S, "A Comparative Analysis of Camera, LiDAR and Fusion Based Deep Neural Networks for Vehicle Detection". International Journal of Innovations in Science and Technology. Vol 3, Special Issue, pp: 177-186, 2021.
Received | Nov 15, 2021; Revised | Jan 27, 2022 Accepted | Jan 28, 2022; Published | Jan 29, 2022.

Self-driving cars are an active area of interdisciplinary research spanning Artificial Intelligence (AI), Internet of Things (IoT), embedded systems, and control engineering. One crucial component needed in ensuring autonomous navigation is to accurately detect vehicles, pedestrians, or other obstacles on the road and ascertain their distance from the self-driving vehicle. The primary algorithms employed for this purpose involve the use of cameras and Light Detection and Ranging (LiDAR) data. Another category of algorithms consists of a fusion between these two sensor data. Sensor fusion networks take input as 2D camera images and LiDAR point clouds to output 3D bounding boxes as detection results. In this paper, we experimentally evaluate the performance of three object detection methods based on the input data type. We offer a comparison of three object detection networks by considering the following metrics - accuracy, performance in occluded environment, and computational complexity. YOLOv3, BEV network, and Point Fusion were trained and tested on the KITTI benchmark dataset. The performance of a sensor fusion network was shown to be superior to single-input networks.
Keywords: Sensor fusion; object detection; 3D object detection; LiDAR point cloud; self-driving cars.

Introduction.

Object detection has taken primary importance in autonomous driving. At present, current perception systems utilize input data in the form of 2D images, point clouds, or a combination of both 2D images and LiDAR point clouds to achieve accurate 3D localization and detection of vehicles. Research in the field of object detection has produced mature algorithms for 2D images. The emergence of Region-based Convolutional Neural Network (RCNN) [1], Fast-RCNN [2], and Faster-RCNN

[3] removed the bottlenecks of large operating time and high computation power in 2D object detection. Different from region-based algorithms, multiple versions of the state-of-the-art object detector You Only Look Once (YOLO) [5, 6,7] have eased the task of predicting bounding boxes and class probabilities in 2D images.

LiDAR is a widely used sensor in obtaining distances between the object and the sensor. LiDAR emits an infrared laser beam to determine the distance via the time- of-flight principle. The wavelength of LiDARs exploited in self-driving cars is in class 1 eye-safe range. In general, LiDARs perform comparatively better in challenging weather conditions such as fog and rain as opposed to optical cameras. These sensors are also relatively more resilient to changes in ambient light conditions. While 2D LiDARs featuring an array of beams such as those from Sick or LeddarTech are typically manipulated in Intelligent Transportation Systems (ITS), self-driving cars make use of 3D LiDARs such as those from Ouster or Velodyne. These 3D LiDARs maneuver a rotating swivel that covers the entire field of view by scanning an array of laser beams across it. The infrared lasers are in the form of pulses and objects reflect these pulses hence distance information is obtained, yielding a 3D point cloud of the surrounding environment. Vertical resolution and angular revolution are key features that dictate the choice of a 3D LiDAR in an application. Currently, the common use of LiDARs is limited by their high cost.

For the challenge of 3D object detection using LiDAR point clouds, the computer vision community has developed several methods. These include point-cloud voxels [2,3], or transforming the 3D view of point-cloud into a top-down 2D view to exhibit objects [1]. Some other techniques focus on estimating the 6-DoF from a sequence of images. Point Net [9] architecture has garnered significant attention in the research cadre of autonomous vehicles. A variant of this architecture has been applied in Point Fusion [14] to devise an application-agnostic algorithm. However, point clouds do not output color information and, depending upon the resolution of the sensor, point clouds are more or less sparse [15].

Another approach to 3D object detection combines LiDAR data and 2D images. This method has been manifested to benefit from the complementary capabilities of cameras and LiDARs. In a conceptually simplistic approach, a 2D detection network has been utilized to make detections in the 3D point cloud [15]. This is achieved by fusing output from a 2D detector with a transformed 3D point cloud. More involved approaches include Point Fusion [14] and Multi-View 3D Network (MV3D) [12] where a region-based fusion approach has been proposed.

In this paper, we classified object detection algorithms based on input data. Three object detection networks have been identified in this regard: 1) YOLO v3 [7] for 2D images, 2) BEV detection for point clouds, and 3) Point Fusion [14] for fused data. We retrained these algorithms on KITTI [8] and specified metrics to assess the performance of these networks. The same dataset was put into service to evaluate all three networks to ensure the completeness of the comparison. Metrics for performance evaluation are accuracy, performance in occluded environment, and computational complexity.

Following are the main contributions of this work:
•We surveyed existing object detection methods for 2D images, point clouds, and sensor fusion networks.
•We chose representative methods in each category and retrained them on the KITTI dataset [8], and evaluate their performance based on accuracy, performance in occluded environment, and computational complexity.
•We conveyed an analysis of the results along with the advantages and disadvantages of each algorithm.

Related Work.

This section highlights outstanding object detection works for 2D images, point clouds, and sensor fusion. It also reviews the performance of each network in comparison with other networks.

2D Image Approaches

Object detection networks detect certain object classes within an image. Two main categories of state-of-the-art methods can be identified: one-stage methods and two stage-methods. YOLO [5], RetinaNet [18], and Single Shot Multibox Detector (SSD) [19] are one-stage methods that prioritize inference speed. On the other hand, detection accuracy takes precedence in two-stage networks as they first propose candidate regions having a high likelihood of staging the objects, and then score these regions to provide the final detections. Examples include Faster R-CNN [3] and Mask R-CNN [4]. The task of bounding box estimation has great importance in the object detection problem. In some previous works, box encoding is applied where center coordinates (x,y) and offset of the bounding box are considered. RCNN [1], Fast RCNN [2], Faster-RCNN [3], YOLO [5], YOLOv2 [6], and Mask R-CNN [4] wield this type of encoding method for bounding box with a slightly different loss calculation scheme. YOLO [5] bases detection on a regression model. Image is divided into a grid of size S*S and B number of bounding boxes, their confidence scores, and class probabilities are predicted for each cell. Predictions are then encoded as a tensor. As compared to RCNN [1], YOLO [5] offers a faster detection speed. However, there is a slight reduction in performance.

Point cloud Approaches

3D Fully Convolutional Network (FCN) extended the application of 2D FCN by applying it to point cloud data [13]. In some previous works, sophisticated segmentation algorithms have been applied to propose candidates [17]. Region Proposal Network (RPN) is a more recent method of candidate proposal. Complex-YOLO [10] proposed Euler-Region-Proposal Network (E-RPN) for pose estimation. Moreover, PointNet [9] is a ground-breaking contribution that consumes raw point cloud data and is compatible with several applications including part segmentation, object classification, and detection. Some other object detection algorithms transform point clouds into multiple views including Front View and Bird's Eye View (BEV) multi-view feature maps [12,15]. A similar approach is adopted by VeloFCN [16], where point-cloud is transformed into front view. Our algorithm also takes advantage of BEV transformation to perform object detection in point clouds.

Sensor Fusion Approaches

MV3D [12] is a sensor fusion network that efficiently deals with the limitations associated with the sparse nature of point clouds. It transforms point clouds into multiple views to make accurate 3D predictions. Moreover, the network is conveniently divided into two sub-networks: the first sub-network generates 3D candidate box proposals and the second sub-network fuses features from multiple modalities. This fusion framework rejects redundant features. A more conceptually simplistic approach is provided in [15] where 2D detections from a CNN are projected onto the 3D point cloud to obtain LiDAR point subset. A novel model-fitting algorithm then identifies the 3D bounding box based on generalized car models. Point Fusion [14] is a more recent contribution in 3D object detection that processes 3D point cloud data and 2D image data separately with PointNet [9] architecture and a CNN respectively. Information loss associated with BEV point clouds is mitigated in this algorithm.

Material and Methods.

In this paper, we focused on the car detection problem. KITTI [8] benchmark dataset was employed to retrain open-sourced algorithms. The 3D object detection task of the KITTI [8] dataset contains aligned 2D images and point clouds. Labels were available in the form of 2D and 3D bounding boxes. A total of 7418 point clouds and corresponding 2D images were adopted for training the networks from the KITTI [8] benchmark dataset. For 2D images, YOLOv3 [7] was evaluated and for point cloud data, a 2D projection approach to Bird's Eye View (BEV) was adapted as arrayed in MV3D [12] and 3D FCN [17]. Point Fusion [14] was retrained and evaluated as a framework for sensor fusion.

KITTI [8] dataset was obtained via VM Station Wagon mounted with number of different sensors including Velodyne HDL64 high precision Global Positioning System (GPS) inertial navigation system and RGB camera. A total of 6-hour drive data was obtained from driver viewpoint in [8]. Velodyne HDL64 rotates at 10Hz frequency with angular resolution of 0.09^{o}. It captures 1.3 million points points/second with 360^{o} horizontal and 26.8^{o}vertical field of view having range of 120m [8]. In this paper, we have focused on left camera RGB images, corresponding point clouds, and calibration files including the calibration details for velodyne to camera calibration.

Following method was adopted to implement and evaluate the performance of three object detection algorithms that are, YOLOv3, BEV network, and Point Fusion:

i. KITTI dataset was obtained from its website as it is an open-source dataset.
ii. The training dataset was split into training and validation dataset in the ratio 1:1. It was used to categorize three classes of objects namely car, pedestrian and cyclist.
iii. YOLOv3 was trained using RGB images only as it is a 2D detection network and does not require point cloud data.
iv. BEV network was trained using point clouds only as this framework makes detections in point clouds. Point clouds were projected into BEV to encode the information of density height and intensity. Firstly, the height feature was obtained by discretizing the point cloud into a 2D grid with a 0.1m resolution. Secondly, in every cell, the value of reflectance of every point having maximum

height was obtained. Thirdly, the density feature simply proclaimed the total number of points in a cell. By implementing these steps, BEV portrayal of point clouds was obtained for the dataset.

v. Point Fusion was trained using both RGB images and LiDAR point clouds as it is a sensor fusion network.

vi. These trained frameworks were then tested using test images available in the dataset.

Results and Discussion.

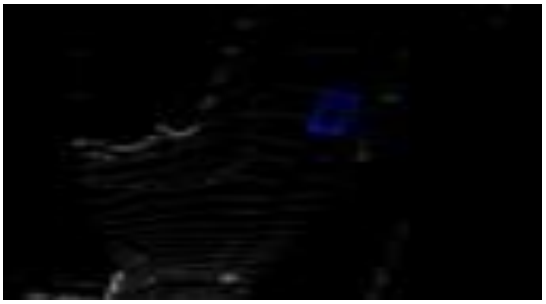
Table 1.shows a comprehensive comparison of models trained and evaluated on the KITTI benchmark dataset for car detection.

Table 1. Comparison of Object Detection Networks trained on KITTI

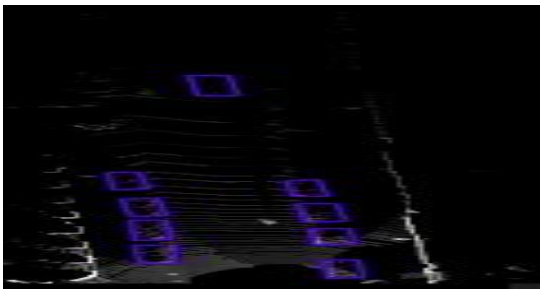
Method	Type of Input Data	Input Processing	No. of Stages	Average Precision (%)	Inference Time (sec)
YOLO v3	2D images	S*S grid	1	45	0.5
BEV	Point-clouds	2D projection	2	42	0.9
Point Fusion	2D images + point clouds	PointNet + ResNet	2	47.8	1.2

(a) YOLOv3 Detection Result

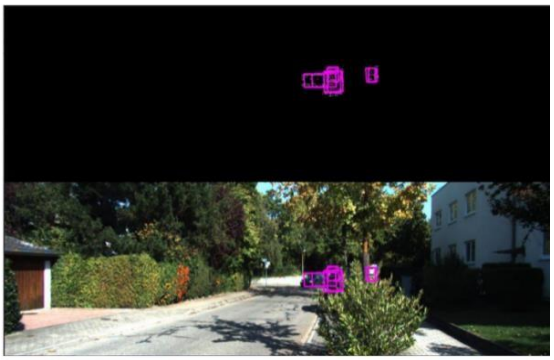
(a) YOLOv3 Detection Result



(b) BEV Detection Result

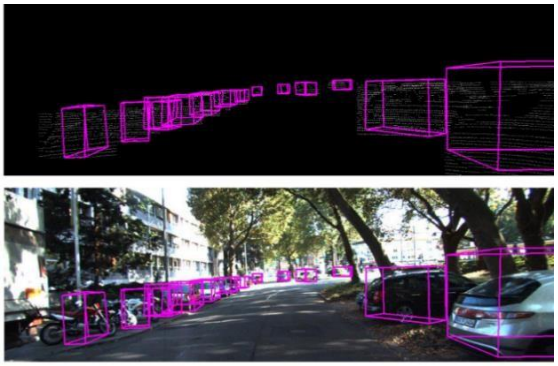


(b) BEV Detection Result



(c) Point Fusion Detection Result

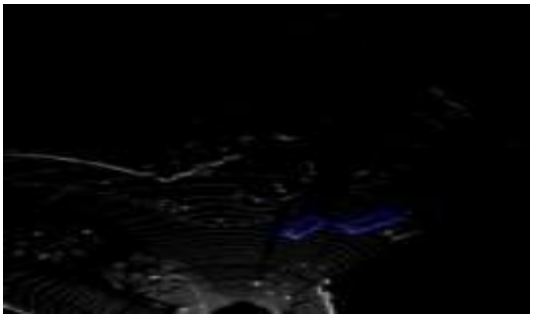
Figure 1.Object Detection Results for Three Implemented Algorithms. (a) YOLOv3: Partially Visible Vehicles not detected (b) BEV: Partially Visible Vehicles detected (c) Point Fusion: Partially Visible Vehicles detected



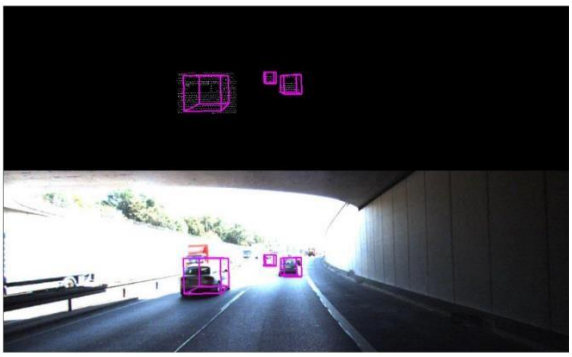
(c) Point Fusion Detection Result

Figure 2. Object Detection Results for Three Implemented Algorithms. (a) YOLOv3: Missed Detections in Occluded Environment (b) BEV: Occluded Vehicles Detected (c) Point Fusion: Maximum Number of Occluded Vehicles Detected

(a) YOLOv3 [7] Detection Result



(b) BEV Detection Result



(c) Point Fusion [14] Detection Result

Figure 3. Object Detection Results for Three Implemented Algorithms. (a)YOLOv3: Detections Missed (b) BEV: Detections Missed (c)PointFusion: No Missed Detections

Discussion.

Three metrics were chosen to evaluate the performance of each algorithm: 1) Accuracy, 2) Performance in Occluded Environment, 3) Computational Complexity.

Accuracy

Average precision (AP) was considered as a metric to determine the accuracy of detections. AP scores for implemented algorithms are listed in Table 1. As compared to YOLOv3 and BEV, Point Fusion gives higher AP. Projection losses associated with BEV reduced detection accuracy. On the other hand, YOLOv3 displayed reduced performance with KITTI dataset. Qualitative results depicted in Figure 1-3 revealed that sensor fusion gives the best detection results in all scenarios. Hence, the sensor fusion network proposed in Point Fusion overcomes the drawbacks associated with single sensor networks.

Performance in Occluded Environment

The closeness or merging of two factors such that one is completely or partially covered by the other is referred to as occlusion. Object detection in an obstructed environment is a beneficial indicator of algorithm performance due to the problem's complexity. As seen in Fig. 2, YOLOv3 and BEV missed most occluded objects whereas the sensor fusion network detected all occluded objects available in the scene. From the results delineated in Figure 1-3, it was observed that YOLOv3 gives poor performance, BEV gives an intermediate performance, and Point Fusion produces the most accurate results in an occluded environment. This proves that sensor fusion frameworks are suited for application in all types of scenarios.

Computational Complexity

YOLOv3 has a fully convolutional architecture comprised of 106 layers. It is the slowest network compared to BEV and Point Fusion; however, it is less sophisticated than many other detection networks. On the other hand, BEV is the least complex algorithm as it projected a 3D point cloud into a 2D point cloud using the method offered in MV3D and made detections using Faster-RCNN. Point Fusion lies between the other two algorithms in terms of computational complexity. Moreover, the

performance of Point Fusion was increased by the adoption of PointNet that processed point clouds in raw form. From qualitative results subdued in Figure 1-3 and AP unveiled in Table 1, it can be derived that the increased computational complexity of sensor fusion frameworks can be overlooked owing to their increased detection accuracy.

In Figure 1-3 qualitative results are set forth. In Point Fusion detection results, front views of corresponding point clouds were also appended to reveal comprehensive results. 3D detections from Point Fusion were projected on the point clouds to generate front view detections. When compared with BEV detection results in Figure 1-3, it became evident that Point Fusion also gives better performance when detections were made in point clouds. This performance improvement was justified by the fact that sensor fusion networks extract features from both 2D images and point clouds exploiting intensity, height, and density information. There is a partially visible vehicle in Fig. 1(a) that was not spotted by YOLOv3.The other two networks, on the other hand, caught the identical car, demonstrating that point clouds and sensor fusion are more capable of recognizing partially visible objects than 2D images.

Moreover, in Fig. 2(a), YOLOv3 missed several occluded objects whereas maximum occluded objects were detected by Point Fusion. This was an important observation as performance in an occluded environment is an important parameter to evaluate the performance of networks. While sensor fusion frameworks are computationally complex and have greater inference time as reported in Table 1, these challenges can be traded off for better performance and accuracy of detection.

Conclusion.

We provide a comparison of three object detection techniques based on the input data type in this paper. An image-only algorithm, a LiDAR-only method, and a sensor fusion framework are among them. The KITTI benchmark dataset is operated to test these object detection systems. Performance evaluation concerning three metrics – i.e., accuracy, performance in occluded environment, and computational complexity – show that the sensor fusion framework gives better overall performance than single sensor algorithms. Qualitative and quantitative results expressed also support the thesis that sensor fusion for object detection is more productive as compared to camera and LiDAR only algorithms. As part of future work, we intend to explore the performance improvements achievement due to sensor fusion in the context of overhead vehicle profiling for Intelligent Transportation Systems (ITS).

Acknowledgement. The manuscript has not been published to other journals and all authors contributed significantly to the research.

Author's Contribution. All authors contributed significantly towards the completion of this research and all authors are in agreement with the content of the manuscript.

Conflict of interest. There exists no conflict of interest for publishing this manuscript in IJIST as the manuscript has not been published or submitted to other journals. However, this research was presented in International Conference on Engineering & Computing 2021.

Project details. This work was supported by the Higher Education Commission of Pakistan (HEC) Technology Development Fund (TDF) grant TDF03-219 awarded to Dr. Shahzor Ahmad.

REFRENCES

1. G. Ross, *"Rich feature hierarchies for accurate object detection and semantic segmentation,"* in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.
2. G. Ross, "*Fast R-CNN*," in Proceedings of the IEEE Conference on Computer Vision (ICCV), 2015.
3 R. Shaoqing, *"Faster R-CNN: Towards real-time object detection with region proposal networks,"* in Proceedings of the IEEE Conference on Neural Information Processing Systems (NIPS), 2017.
4. H. Kaiming, "*Mask R-CNN*," in Proceedings of the IEEE Conference on Computer Vision (ICCV), 2017.
5. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *"You Only Look Once: Unified, Real-Time Object Detection,"*in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.
6. J. Redmon and A. Farhadi, *"YOLO9000: Better, Faster, Stronger,"*in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6517-652.
7. J. Redmon and A. Farhadi, *"YOLOv3: An Incremental Improvement,"* 2018.
8. A. Geiger, P. Lenz, and R. Urtasun, *"Are we ready for autonomous driving?The kitti vision benchmar suite,"* IEEE CVPR, 2012.
9. C. R. Qi, H. Su, K. Mo, and L. J. Guibas, *"PointNet: Deeplearning on Point Sets for 3D Classification and Segmentation,"* in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017
10. S. Martin, *"Complex-YOLO: Real-time 3d object detection on point clouds*," in Proceedings of the European Conference on Computer Vision (ECCV).
11. S. Song, and J. Xiao, *"Deep sliding shapes for amodal 3d object detection in rgb-d images,*" in In Proceedings of the IEEE Conference on Computer Visio and Pattern Recognition, 2016.
12. X. Chen, H. Ma, J. Wan, B. Li, and T. Xia., *"Multi view 3D object detection network for autonomous driving,"* IEEE CVPR, 2017.
13. B. Li, *"3D Fully Convolutional Network for Vehicle Detection in Point Cloud,"* in IROS,, 2016.
14. D. Xu, D. Anguelov, and A. Jain, *"Point Fusion: Deep Sensor Fusion for 3D Bounding Box Estimation,*" in Proceedings of the IEEE Conference on Computer Visio and Pattern Recognition,, 2018.
15. X. Du, M. H. A. Jr, S. Karaman, and D Rus, *"A General Pipeline for 3D Detection of Vehicles,*" IEEE ICRA, 2018.
16. B. Li, T. Zhang, and T. Xia, *"Vehicle Detection from 3d lidar using fully convolutional network,*" In Robotics: Science and Systems, 2016.
17. D. Nister, O. Naroditsky, and J. Bergen, *"Visual Odometry*," IEEE CVPR, 2004

18. T. Y. Lin, P. Goyal, and G. Ross, *"Focal Loss for Dense Object Detection*," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2999-3007, doi: 10.1109/ICCV.2017.324.
19. W. Liu et al., *"SSD: Single Shot MultiBox Detector*," 2016 ECCV

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Flow Analysis of Various Inlet Velocity Profiles on Indoor Temperature for Energy Conservation of HVAC System Using CFD

Atta ul Mannan Hashmi[1], Arshan Ahmed[1], Fahad Rafi Butt[2], Shahbaz Ghani[1] and Imran Akhtar PhD[1].
[1]Department of Mechanical Engineering, NUST College of Electrical and Mechanical Engineering,National University of Sciences and Technology, Islamabad, Pakistan.
[2]Research Associate, Digital Pakistan Lab, NUST College of Electrical and Mechanical Engineering,National University of Sciences and Technology, Islamabad, Pakistan.
* Correspondence: Atta ul Mannan Hashmi; atta44009@gmail.com.
Citation | Hashmi. A. M, Ahmad. A, Butt. F. R, Ghani. S and Akhtar. I, "*Flow Analysis of Various Inlet Velocity Profiles On Indoor Temperature for Energy Conservation of HVAC System Using CFD*". International Journal of Innovations in Science and Technology, Vol 3, Special issue, pp: 187- 196, 2021.
Received | Dec 14, 2021; Revised | Jan 25, 2022 Accepted | Jan 28, 2022; Published | Feb 06, 2022.

Energy conservation has been the most popular topic of the modern world. Heating, Ventilation and Air Conditioning (HVAC) systems consume approximately 10 % of the total energy of world. In order to improve the efficiency of HVAC systems, two dimensional (2D) room with inlet, outlet and heat source has been modeled. ANSYS Fluent has been used for numerical analysis of air flow in a 2D room. User Defined Functions (UDFs), which are coded in C language and hooked in ANSYS fluent, have been used for recording temperature variations and for heat generation within 2D room. Besides studying velocity fields and temperature distributions within indoor environment under specified boundary conditions, reference region for comparative analysis is also selected during Steady State (SS) numerical simulations. During transient analysis, temperature variations of a selected location are recorded for four different scenarios under varying inlet velocity profiles i.e. three for 0°, 30°, 60° angle with 1.3661 m/s velocity and fourth 0° with 2.7322 m/s velocity. Temperature profile of reference region after 1500 sec of transient simulations are compared with the steady state. Temperature profile of the scenario once the air is injected at 30° closely matched with the steady state temperature profile of the selected region. Time for attainment of SS temperature is also measured and compared after transient simulations. SS temperature value was attained twice, first at 240 seconds when the air was injected at 0° with 2.7322 m/s and secondly at 522 seconds when inlet air entered at velocity of 1.3361 m/s at 30°. The power

consumption by increasing the fan speed is much higher as compared to the power consumed for changing direction only.

Keywords: HVAC systems; energy conservation; computational fluid dynamics; user defined functions and heat generation.

Introduction

Study of indoor environment is relatively complicated owing to its dependence on velocity fields, temperature distribution and relative humidity maps. Study of indoor environment gets more complex once it comes to the presence of heat source within closed environment. HVAC systems are used in closed room for providing comfortable environment to the occupants. The fundamental phenomenon of convection governs the heat transfer within closed environment and keeps the temperature within comfortable zone. In the presence of heat source both natural and forced convections contributes towards heat transfer. Convection is a process of transferring heat from hotter to cooler medium by mass motion of fluids. In natural or free convection, hot air moves up and away from the hotter surface whereas in forced convection air is forced over the hot object for taking away heat and lowering its overall temperature. Heat transfer coefficient 'h' depends upon direction, magnitude of fluid flow and the temperature difference between the flowing fluid and the hotter object [1].

Over many years, various studies and analysis have been carried out to ascertain the air flow patterns and temperature gradient profiles within various geometrical boundaries. Major focus of the researchers have remained on forced convection rather than natural convection because heat transfer through the latter is considered negligible as compared to former. In 2D model, temperature distribution and velocity profile for free, forced and mixed convection with appropriate geometries and mesh refinement was presented by Horikiri et al. [2]. Although the CFD study was validated with experimental data for 2D empty room. But natural and mixed convections were achieved by keeping constant temperature boundary conditions only. Ni & Bai [3] presented the statistical research on air conditioning energy performances in data centers. HVAC systems constitute almost 38% of total energy consumed by data centers. Study also emphasized on the comfortable and optimum operating temperature range of 18°C~27°C which is also recommended by American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). Fang et al. [4] carried out Computational Fluid Dynamics (CFD) simulations in order to study temperature distribution of the air conditioned room both under steady and transient conditions. However, fixed inlet conditions and profiles were used for the analysis. Mesenhöller et al. [5] presented the effect of ventilation on the interior temperature distribution. Unsteady ventilation led to stronger mixing effects in the interior zone as compared to fixed and steady ventilation. Raczkowski et al. [6] validated CFD model with the experimental data for ensuring thermal comfort in a naturally ventilated rooms. Schiavon & Melikov [7] calculated the energy conservation of HVAC systems under different velocity inlet profiles. Wang & Chen [8] studied various turbulence models for the assessment of transitional flows within closed environment. The complexity of the experiment and simulation was gradually increased from isothermal forced convection in totally empty room to a room with box and then to a heated box.

Study of the literature reveals that for economizing the energy consumption by HVAC systems, air distribution and temperature contours in closed environment under various circumstances plays the most pivotal role. Modeling and controlling HVAC system's response under different conditions mainly depends upon the thermal comfort which includes, velocity field, temperature profile and relative humidity maps.

According to third fan law, just by doubling the initial velocity will result in 8 time's increase of power consumption by the fan only. Third law of the fan is stated below:-

$$P_2 = P_1 \left(\frac{U_2}{U_1}\right)^3 \quad (1)$$

Where P_2 is the new power consumption, P_1 is the initial power consumption, U_2 new RPMs of the fans, U1 the initial RPMs of the fan.

Model Equations

Velocity field and temperature distributions are governed by basic conservation laws i.e. mass, momentum and energy conservation laws. The corresponding equations are:-

Continuity equation.

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} + \frac{\partial \rho w}{\partial z} = 0 \quad (2)$$

Where ρ is density, u, v, w are velocity components in Cartesian coordinates system.

Energy equation.

$$\rho c \left(\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} + w\frac{\partial T}{\partial z}\right) = \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) + Q \quad (3)$$

Where c is specific heat capacity, T is temperature, k is Thermal conductivity.

Conservation of momentum

$$\rho \left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right) = -\frac{\partial p}{\partial x} + \rho g_x + \mu \nabla^2 u \quad (4)$$

$$\rho \left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}\right) = -\frac{\partial p}{\partial x} + \rho g_y + \mu \nabla^2 v \quad (5)$$

$$\rho \left(\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z}\right) = -\frac{\partial p}{\partial x} + \rho g_w + \mu \nabla^2 w \quad (6)$$

Where p is pressure, μ is dynamic viscosity, g is gravity, ∇^2 is Laplacian.

Methodology

Methodology adopted for numerical simulation is depicted in the following flow diagram:-

CAD Model	• CAD model of 2D Room.
Problem Setup	• CAD model imported in Ansys fluent • Meshing • Application of Boundary Conditions
Steady State Simulation	•For velocity and temperature distribution •Center line used as refernce for temperaure measurements. •Indentification of point on center line for further comparison during transient simulations.
Transient Simulation	•3 Cases for varying inlet air velocity angles and 1 case for increasing inlet velocity. •UDFs for temperature measurements. •UDMs for storing average temperatures. •UDFs coded in C language and hooked in Ansys fluent
Analysis	• Temperatures measured during transient simulations under varying inlet profiles were compared with steady state temperatures. • Time for attainment of desired temperatures compared.

Numerical Simulation

Geometry and Setup. A 3D cubic room [8] as depicted in Figure 1 has been used for analyzing velocity field and temperature distribution. In order to study the flow separation, a cubic box of half the dimensions of the room (1.22 x 1.22 x 1.22m) was used at the center. Cubic box was heated to study the effect of natural convection in a room. About 0.03m inlet and 0.08m outlet cross configurations were used for analysis.

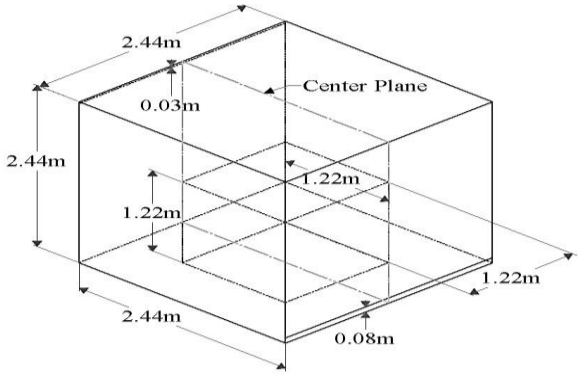


Figure 1. 3D CAD Model of Room under Consideration.

For the present study, center plane having inlet at top left and outlet at right bottom as shown in Figure 2.

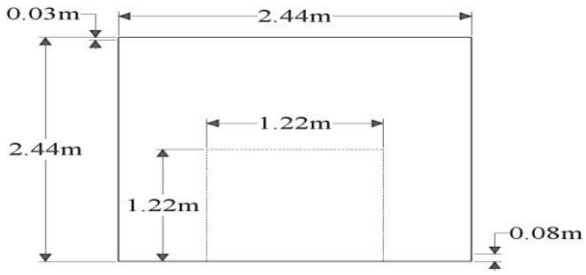


Figure 2. 2D CAD Drawing.

ANSYS fluent was used for numerical analysis. Complete 2D center plane of the room was discredited into 187,617 all quad cells of max size 0.005m. Inflation with first layer thickness of 0.0042m and growth rate of 1.05 was applied on all the edges with 5 layers as shown in Figure 3. k-ϵ Realizable with standard wall function turbulence model has been used for modeling turbulence. Semi Implicit Method for Pressure Linked Equations (SIMPLE) with second order upwind schemes have been set for pressure, momentum and energy whereas first order implicit scheme has been used in this analysis.

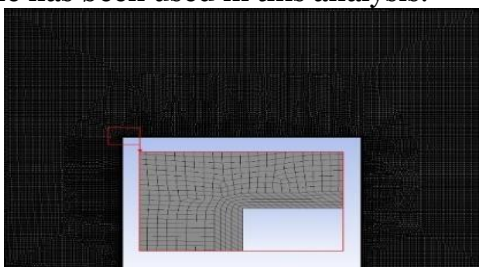


Figure 3. Mesh with Inflation at the edges.

Air was injected at 1.3661m/s with the constant temperature of 295.35K because same boundary conditions were used by reference [8]. A constant heat of 470.3W/m^3 was generated from the box to simulate actual scenario and to cater for buoyant forces effect. Outlet boundary condition of pressure-outlet with zero Pascal of gauge pressure at 323.15K was set at the 2D room oulet. Initial interior temperature of the complete room was set at 323.15K.

For numerical analysis, case was initially run for steady state in Ansys fluent for obtaining velocity field and temperature distribution profile in the room. Center line as shown in Figure 6 was used as a reference for temperature measurements. For transient analysis, User Defined Function (UDF) of heat source was used for heat generation by the box and UDF (Execute at the END) was used to calculate average temperature around the center line as identified during steady state analysis. Both direction and magnitude of inlet air velocity profiles were varied in transient analysis depending upon already calculated average temperature of any identified point or region during steady state study. Present study only discusses the effect of change in direction and magnitude of inlet velocity at a constant temperature of 295.35K. Total of 4 cases excluding steady state were analyzed, out of which 3 cases for constant inlet velocity at clockwise 0°, 30° and 60° with the top wall. Last case was analyzed with the doubled velocity at 0° and constant temperature of 295.35K.

Validation and Verification. Air distribution in a cubic room with the heated box was validated for constant inlet velocity of 1.3661 m/s and isothermal boundary conditions. Figure 4 depicts the heat generation by the box within a cubic room.

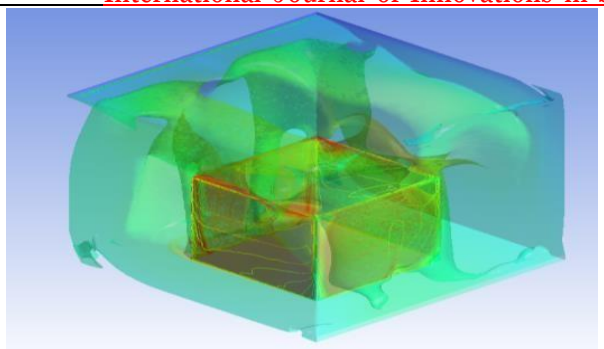


Figure 4. Temperature Iso-Surfaces inside a Cubic Room with Heat Generation.

The results from the current simulations are compared with Wang & Chen [8] for validation before carrying out further analysis. The comparison plot of normalized x-velocity with the normalized height is validated in the Figure 5.

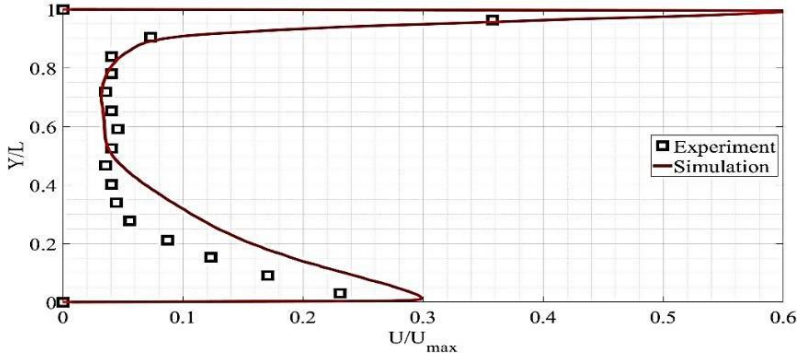


Figure 5. A Comparison of Present Simulation with the Experimental Data.

Results

In order to simplify the analysis, center line was selected and its variations were studied with changing inlet profile. A point (1.22, 1.503) on the center line was chosen which had the maximum gradient below which the effect of heat generation source takes over. The steady state temperature contours along with the chosen point and velocity magnitude contours are shown in Figure 6&7 respectively.

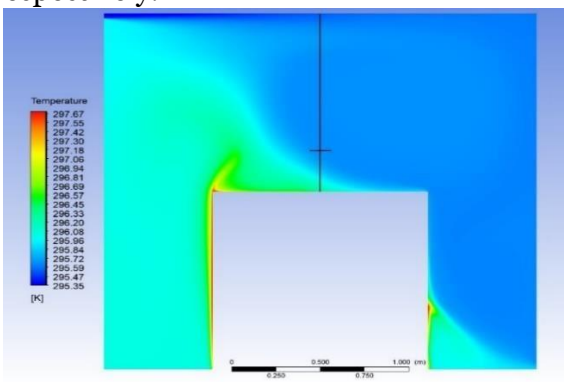


Figure 6. Steady State Temperature Profile.

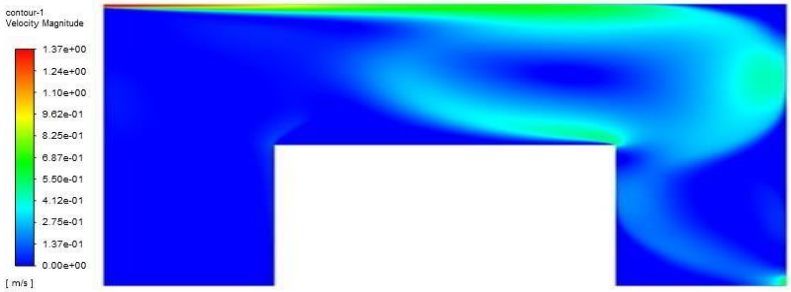


Figure 7. Steady State Velocity Magnitude Contours.

Temperature was measured at the center of each cell of the selected rectangular region around centerline as depicted in Figure 8. Average temperature of the region was calculated at the end of each iteration during steady state and after every time step during transient analysis. UDFs were separately coded and hooked in ANSYS fluent for recording live data during the simulation. Average temperature was stored separately in a User Defined Memory (UDM) for analysis purpose. Calculated average temperature can be used for continuously changing inlet velocity or temperature profiles. For simplicity and focusing on direct effect of inlet profiles on temperature distributions, UDF was only used for calculating average temperatures under different fixed inlet velocity profiles represented in Figure 8.

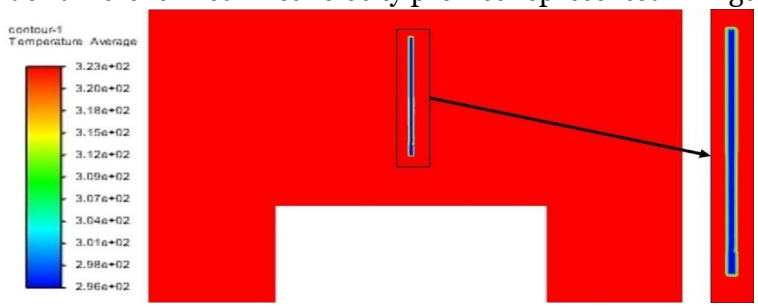


Figure 8. UDM Temperature Contours.

Transient Analysis. Transient analysis was run for 1500 seconds of physical time for three cases with constant velocity inlet (1.3661 m/s) at 0°, 30° and 60° of flow angles clockwise with the top wall. The transient response for the center line differed from the steady state response. The variation in the temperature profiles of center line during steady state and transient simulations (after 1500 seconds of physical time) are shown in Figure 9.

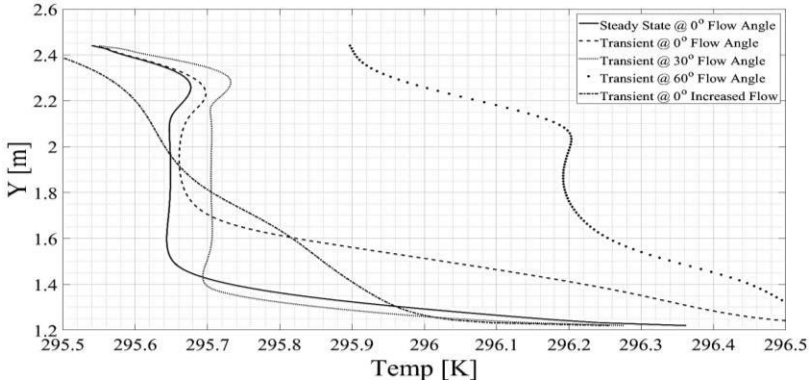


Figure 9. Center Line Temperature Variation Plots.

Solid line in the graph (Figure 9) represents steady state variations. Dashed, dotted and dashed-dot lines shows constant inlet velocities at 0°, 30°, 60° clockwise angles with the top walls whereas 'o' plot shows the velocity variations with doubled inlet velocity i.e. 2.7322 m/s. Center line graphs showed that overall profile of inlet velocity injected at 30°, matches with the steady state profile with further decrease in temperature below 1.428m of height. Although overall profile of flow injection at 30° was shifted towards right by 0.05K but the overall temperatures values below the height of 1.428m were lower than the steady state injection at 0°. The overall center line of height (1.5 ~2.16m) averaged temperatures of all the four cases along with the steady state bench mark as shown in Figure 10. The averaged center line temperature at 30° is 0.11K less than the inlet at 0° at same constant velocity.

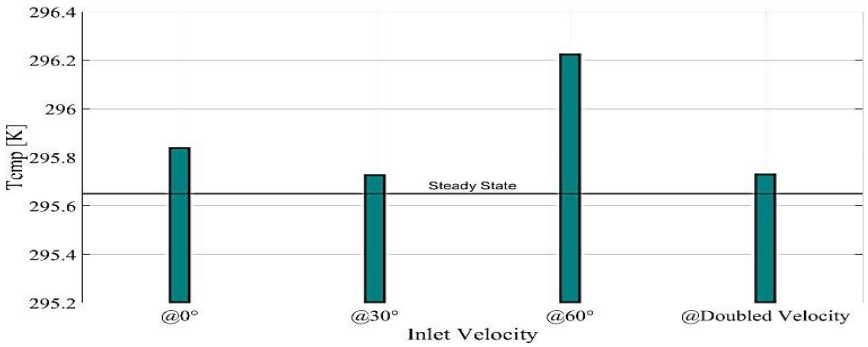


Figure 10. Center Line Averaged Temperatures.

Area averaged temperature of the complete room was also calculated for 1500 sec of transient simulation. Comparative analysis as shown in Figure 11 revealed that the overall averaged room temperature at 30° is 0.0273K less than the velocity inlet at 0°. Averaged area room temperature for doubled velocity (2.7322 m/s) inlet at 0°is 0.3044K less than velocity (1.3661 m/s) at 0°.

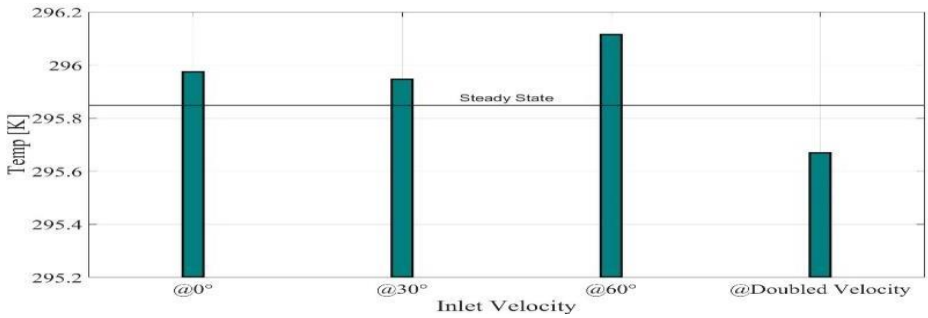


Figure 11. Area Averaged Temperatures of Room.

The minimum temperatures achieved at the selected point (1.22, 1.5) was also compared for better analysis. Although with the change in selection (point or region), the values of temperature may produce results much different from the one shown in Figure 12. But overall averaged room temperature and the time it takes to reach will contribute towards energy conservation.

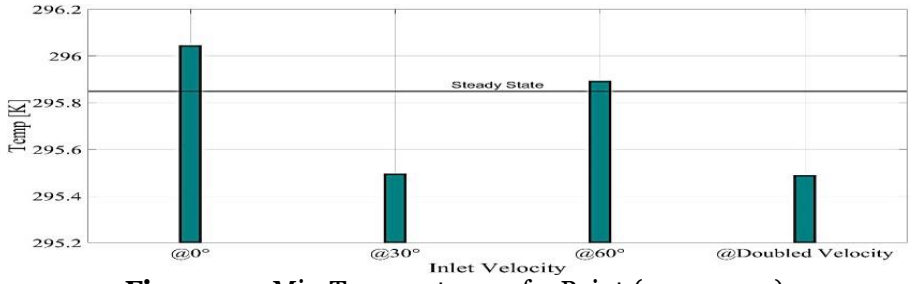


Figure 12. Min Temperatures of a Point (1.22,1. 503).

Discussion

After 1500 seconds of physical time transient simulations, SS point temperature value was achieved for only two transient simulation cases, first one was when the inlet air was injected at 30° with velocity (1.3661 m/s) and the second was once air was injected with doubled velocity (2.7322 m/s) at 0°. So, during transient simulations, time to attain desired temperature could be reduced either by changing angle or magnitude of the inlet air by keeping its temperature constant. Time can be further reduced even be making an appropriate combinations of both i.e. angle and magnitude of inlet air. But the energy requirement to change the direction was observed less than by increasing the velocity magnitude.

Comparison of transient temperature profiles at a chosen point (1.22, 1.5) for all four cases is shown in Figure 13. The transient study showed that the SS temperature value of the chosen point i.e. 295.65K could not be achieved at 0° inlet velocity at 1.3661 m/s however it was achieved at doubled the inlet velocity i.e. 2.73 m/s at 240 sec. SS temperature value of the selected point was also achieved in 522 sec once the constant temperature air was injected at 1.3661 m/s but at 30°.

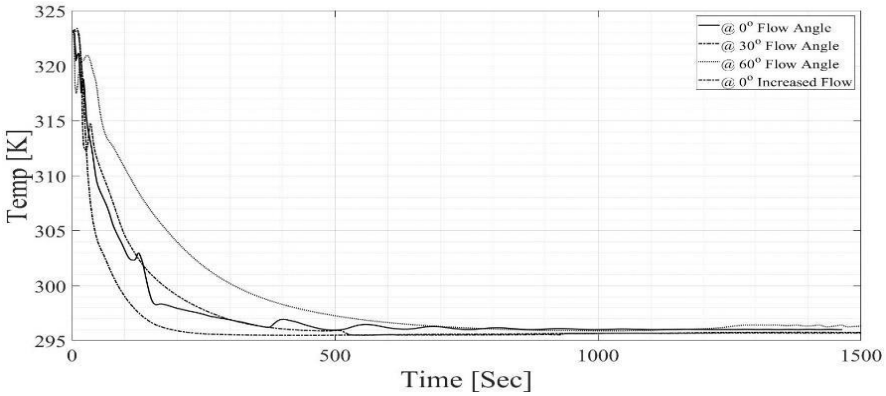


Figure 13. Transient Temperature Profiles

Doubling the velocity at inlet from the constant area will double the volume flow rate. According to first fan law, the RPMs is directly proportional to volume flow rate. In this study volume flow rate is doubled from 0.1 m^3 /s to 0.2 m^3 /s (i.e. velocity is doubled from 1.3661 m/s to 2.7322 m/s through area 0.0732 m^2) keeping other conditions constant and changing the velocity only. So, as per Equation 1, the power consumption by the fan will increase by 8 times.

Conclusion.

Varying the inlet velocity profiles in 2D room, the indoor velocity field and temperature distribution vary significantly as compared to open air. Although by changing velocity, magnitude will produce early and better cooling but more energy will be consumed. By doubling the velocity at inlet, the power consumption will increase 8 times of the initial fan power or the power being consumed just by changing the inlet direction. To achieve similar results without consuming extra energy, change in inlet velocity direction is one of the recommended option. However, change in inlet direction may take little more time than by increasing the velocity only. Steady State temperature of 295.65K is achieved either by doubling the velocity (2.7322 m/s) in 240 seconds or by injecting the air at 30° with 1.3661 m/s in 522 seconds.

Acknowledgement. This manuscript has not been published in other journals previously. This research is supported by CEME, NUST, Pakistan.

Author's Contribution. Atta ul Mannan Hashmi developed the theory and performed the numerical computations. Arshan Ahmed and Fahad Rafi Butt verified the analytical results. Shahbaz Ghani and Imran Akhtar analyzed effect of various inlet velocity profiles on indoor temperature for energy conservation of HVAC. All authors discussed the results in detail and contributed to the final manuscript.

Conflict of interest. The authors declare no conflict of interest for publishing this manuscript in IJIST.

REFRENCES

[1] https://www.iea.org/news/air-conditioning-use-emerges-as-one-of-the-key-drivers-of-global-electricity-demand-growth

[2] K. Horikiri, Y. Yao, and J. Yao, "Numerical simulation of convective airflow in an empty room," "Int. J. Energy Environ" 2011, vol. 5, no. 1, pp. 574–581.

[3] J. Ni and X. Bai, "A review of air conditioning energy performance in data centers," "Renew. Sustain. Energy Rev." 2017, vol. 67, pp. 625–640.

[4] P. Fang, T. Liu, K. Liu, Y. Zhang, and J. Zhao, "A simulation model to calculate temperature distribution of an air-conditioned room," "Proc. - 2016 8th Int. Conf. Intell. Human-Machine Syst. Cybern. IHMSC" 2016, vol. 1, no. 2, pp. 378–381.

[5] E. Mesenhöller, P. Vennemann, and J. Hussong, "Unsteady room ventilation – A review," "Build. Environ." 2020, vol. 169.

[6] A. Raczkowski, Z. Suchorab, and P. Brzyski, "Computational fluid dynamics simulation of thermal comfort in naturally ventilated room," "MATEC Web Conf." 2019, vol. 252, p. 04007.

[7] S. Schiavon and A. K. Melikov, "Energy saving and improved comfort by increased air movement," "Energy Build." 2008, vol. 40, no. 10, pp. 1954–1960.

[8] M. Wang and Q. Chen, "Assessment of various turbulence models for transitional flows in an enclosed environment (RP-1271)," "HVAC R Res." 2009, vol. 15, no. 6, pp. 1099–1119.

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

An Automated Framework for Corona Virus Severity Detection Using Combination of AlexNet and Faster RCNN

Muhammad Haris Munir[1], Rabbia Mahum[2], Muhammad Nafees[3], Muhammad Aitazaz[4], Aun Irtaza[5]
[1,2,3,4,5] Department of Computer Science, University of *Engineering* and Technology-Taxila,Taxila 47050, Pakistan
[1]mharis.munir@students.uettaxila.edu.pk,[2,*]rabbia.mahum@uettaxila.edu.pk,
[3]muhammad.nafees@students.uettaxila.edu.pk,
[4]muhammad.aitzaz@students.uettaxila.edu.pk, [5]aun.irtaza@uettaxila.edu.pk
Citation | Munir. M. H, Mahum. R, Nafees. M, Aitazaz. M, Irtaza. A, "An Automated Framework for Corona Virus Severity Detection Using Combination of AlexNet and Faster RCNN". International Journal of Innovations in Science and Technology. Vol 3, Special Issue, pp: 197-209, 2022.
Received | Dec 27, 2021; Revised | Feb 2, 2022 Accepted | Feb 4, 2022; Published |Feb 19, 2022.

Coronavirus has affected daily lives of people all around the globe. Lungs being the respiratory organ are the most affected by such a virus. Alternative techniques for diagnosing the coronavirus involving X-rays and CT scans of the chest have been proposed. The severity of the disease, on the other hand, is a crucial component in the patient's treatment. As a consequence, an automated approach to ascertain the severity of the coronavirus on the lungs is designed to decrease the impacts of the coronavirus on the lungs and practice the right treatment. In this manuscript, we proposed a deep learning-based model for identifying the severity level of coronavirus on the lungs which is further categorized in high, moderate, and low. We employed AlexNet for the disease detection and Faster RCNN for the severity level prediction based on the affected area of the lungs. The evaluation is assessed using X-rays and CT scans of the lungs. Total 1400 images have been employed for the training and performance evaluation of the proposed system. The metrics that we considered for the performance evaluation are accuracy, precision, recall, error rate, and time. The results showed that our proposed model attained about 98.4% accuracy and 98.15% precision.

Keywords: Classification; Detection; Deep Learning; Severity Detection and Disease Identification.

INTRODUCTION

An outbreak of the COVID-19 virus emerged in December 2019 in Wuhan, the capital of central China's Hubei province. Although it had a zoonotic origin however, its widespread human-human transmission has resulted in the spread of the virus all over

the world. This virus's emergence has been officially labeled as a global pandemic, and it has impacted the complete lifestyle around the world [1] [2]. The common symptoms of this virus are fever and cough with other non-specific symptoms like dyspnea, headache, muscle soreness, and fatigue. Initial analysis in Wuhan reveals the presence of some bilateral lungs opacities on chest CT scans in infected patients. Moreover, the lobular and sub-segmental area of consolidation is considered as one of the most important factors in the infected patients [3]. The chest CT scan is witnessed worst approximately 10 days after symptoms [4].

Due to advancement in Deep Learning (DL) techniques, various models of DL have been developed for the detection of corona virus. However, lab results are also considered for the validation of decisions by physicians. Transfer learning models have been employed i.e. GoogleNet, inception_V3 for virus detection that was trained on 320 images. The accuracy of this model was 82.5%. Xu, Xiaowei, et al. (2020) [5] have applied the VNET–IR–RPN model for both Segmentation and classification of CT scan images to detect the virus. Their model was trained upon 528 CT scan images. The overall accuracy was observed as 86.7%. Song, Ying, et al.(2021) [6] applied DRENet based upon pre-trained ResNet50, very robust to detect objects within images. About 777 images participated in this analysis to train their model and achieved an accuracy of 93%.

El Asnaoui, Khalid, and Youness Chawk (2020) [7] have done a comparative analysis using modern Deep Learning models (VGG-16, VGG19, DenseNet201, Inception_ResNet_V2, Inception_V3, Resnet50, andMobileNet_V2) to check which model does better detection and classification. Dataset of 6087 images of CT and X-rays scans were availed to train the models. Results proved that inception_Resnet_V2 and Densnet201 provided good results comparatively (92.18% and 88.09 % respectively). Rajaraman et al. [8] has developed a DL technique using X-ray images named as CXR images. He attained the 55% and 65% accuracies on test data using Twitter and Montreal CXR images. Gozes et al. (2020) [9] have built a model that enables covid patients to be distinguished from non-covid patients. Their approach generated a lung anomaly localization map as well as measurements. This approach was divided into two parts:

Part I: To detect nodules, for that 3D analysis was used. Small opacities were created using software, followed by localization and measurements.

Part II: The lung Crop stage is the first phase in Part II. A lung segmentation module was used to extract the lungs region of interest (ROI) (U-net architecture). The next step is to use ResNet50, a deep convolutional neural network model to detect coronavirus abnormalities. Narin et al. [10] Have built an autonomous and deep learning based technique for predicting Covid using X-ray images (2020). Three Deep Convolutions Architectures were used in the suggested technique. They have used a dataset comprising of 50 images of covid patient's X-ray and 50 images of healthy person's X-ray. The X-rays, as well as all other images, were scaled to 224x224 pixels. To get around the issue of a restricted number of datasets, the Transfer learning models were utilized by the authors. The DCNN was a transfer learning models that utilized the k-fold method and the transfer learning methodology and achieved an accuracy of 98%.

A deep learning classifiers system called "COVIDX-Net" was proposed by Hemdan et al. (2020) [11] to assist radiologists in automatically identifying Covid19. The framework created enables for the classification of Covid X-rays into covid positive and covid negative categories. They have used a set of data that included 50 X-ray images were divided into two groups: normal and Covid positive. Instances that are favorable, each consists of 25 X-rays images. The photographs were shrunk to 224x224 pixels. For the project, 80% of the photographs were used for testing and 20% for training. By using VGG19 and DenseNet models they have attained results, with F-1 scores of 89% and 91% for normal and covid positive patient, respectively.

Farooq and Hafeez (2020) [12] introduced CNN to distinguish Covid patients from other Pneumonia and healthy patients. They used COVIDX dataset [13]. They used 5941 chest radiography images in the dataset, gathered from 2839 patients. They employed a subset of the COVIDX dataset, which was separated into four batches. Covid with 48 images, Bacterial with 660 images, Viral with 931 images, and Normal with 1203 images the training stage, which was completed in three parts. The results showed that the suggested Covid Res-Net worked well. Shan et al. [14] have proposed a new DL model that can segregate and quantify infected areas in CT images of Covid patients. In order to assist radiologists in clarifying automatic annotation of each case, the authors used VB-Net Neural Network and (HITL) technique. They then employed evaluation metrics to analyze the model's performance. The CT scans were sorted into categories. The segmentation network has been trained using these CT images.

The segmentation findings were manually corrected by radiologists and used to feed the model as updated data. They iteratively created the model. The study (Xu et al., 2020) used deep learning algorithms to identify COVID-19 from Influenza-A viral pneumonia and healthy pictures. To classify Computed Tomography (CT) images, they employed multiple CNN. The procedure can be broken down into four steps: 1) The photos were pre-processed in order to extract the most useful pulmonary areas. 2) Multiple candidate picture cubes were segmented using a 3D CNN. 3) To distinguish the pictures patch into Covid, Influenza-A, and healthy, an image classification model was applied. 4) An overall analysis report for one CT sample was created using the noisy-or Bayesian function. For segmentation, the VNET-IR-RPN model was utilized, and for classification, the ResNet-18 model with the location attention mechanism model were used. They have achieved an accuracy of 86.7%.

El Asnaoui et al. (2020) compared contemporary DCNN architecture for binary categorization of pneumonia images using fine-tuned VGG16. They achieved accuracy of 96%. The authors of Zhang et al. (2020) [15] demonstrated a deep learning algorithm for detecting Covid in patients using X-ray pictures of chest. The model consisted of three parts: The main network, which is made up of 18 layers of residual convolutional neural networks (RCNN), is the first. It extracts high level features from an X-rays of the chest. The second is the classification head, which is responsible for generating the Pcls classification score. It was driven by the main network's extracted features. The anomaly

detection head is the third component, and it allows Pano to generate a scalar anomaly score. The choice was made based on a threshold T after calculating the categorization score and scalar anomaly score. The findings revealed that the sensitivity was high as 96% for T = 0.15.

In this research, we introduce a novel technique that will help the radiologist and the physicians to detect the presence of coronavirus and the severity of COVID-19 manifestation using the Deep Learning model. Hence, considering the sensitiveness of this disease, time does matter and any delay in the diagnosis and isolation of patients affects the treatment and prognosis and ultimately affects the control of the COVID-19 epidemic.

The main contributions of the proposed framework are as follows: We proposed an effective coronavirus severity detection technique through a hybrid deep learning model i.e. AlexNet [16] and FasterRCNN [17]. Our proposed framework is robust to the input images of lungs i.e X-ray and CT Scans while existing state-of-the-art techniques are based on either X-ray or CT scan images. Our detection method effectively extracts the attributes from the images using modified AlexNet and categorizes the severity using Faster RCNN into three levels i.e. moderate, high, and low.

Material and Methods.

This study proposes a two-step technique for the detection of the corona virus's severity. In the first step, AlexNet was availed for the classification of the images to detect the virus's affected images. In the second step, faster RCNN was trained on multi-class images to determine the severity of the coronavirus. We have made the following block diagram shown in Figure 1 using Microsoft Visio showing the complete working of proposed model.

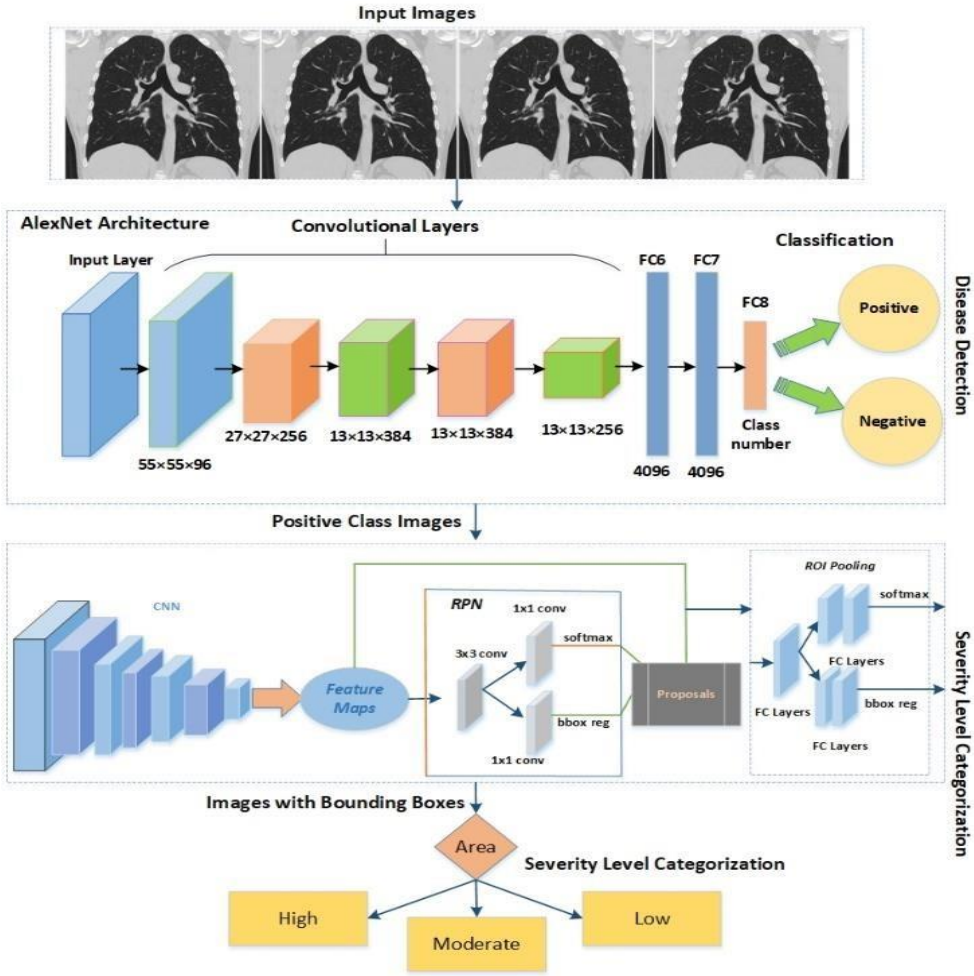


Figure 1. Architecture of the Proposed System.

AlexNet CNN for Detection of Corona Virus

We displayed an AlexNet Convolutional Neural Network (CNN) for corona virus identification in the proposed strategy. Five convolutional layers, two fully hidden layers (totally connected), and one output layer make up the CNN's total of 25 layers (fully connected). Moreover, three maximum pooling layers having size 3x3 were manipulated with a stride of 1 after the three convolutional layers i.e. first, third, and fifth. After each convolutional layer and the first two fully connected layers, we exploited Rectified Linear Unit (ReLU) layers. Normalization layers having window size 5 were adopted after the second and third layers. Dropout value of 0.5 was applied before the first and second Fully Connected (FC) layers. A softmax layer was employed after the last FC layer. Moreover, the AlexNet accepts the size of the image as 227x227, however, images from our dataset had a resolution of 340x320x3, therefore down-sampled to the size of 227x227x3 for the AlexNet algorithm as AlexNet architecture accept image's dimension of 227x227. The input image was convolved on the first convolutional layer with a filter of 11x11. The size of the window was reduced to 5x5 in the next layer. The

window size for the third, fourth, and fifth layers was set to 3x3. Fully linked layers with 4096 nodes were deployed after convolutional layers. Two nodes for the binary classification were added in the final FC layers: affected and non-effected. The parameter of learning rate was set to 0.01 and the batch size was set to 75.

Problem Formation

Let $I=\{I1,I2,\ldots\ldots,In\}$ indicate the total number of images obtained from the CT scans and X-rays. Images were classified into two classes i.e. Affected as $I_{eff}$and Non-Effected as I_{n_eff}through AlexNet. For severity detection, I_{eff} images were classified further into three classes such as I_{mod}, I_{hogh}, and I_{low} through Faster RCNN.

Faster RCNN for Disease Severity Detection

The Faster RCNN is a modified version of existing models such as RCNN and Fast RCNN, that operate the Edge Box to create the region proposals for the detection of objects. Faster RCNN exercises a Region Proposal Network (RPN) instead of the Edge Boxes technique to form the region proposals. Therefore, the complexity of Faster RCNN is reduced than existing models. The image sample and the position of the affected area of the image sample are the two types of input for the faster RCNN. Faster RCNN is trained on images that fall into one of three categories based on the area of the bounding box that depicts the damaged lungs.

Faster RCNN produces two outputs: 1) classified images based on training, and 2) bounding box coordinates.

Image Annotation

Image annotation is a process to label the classes for the classification in multi-class object recognition. In the proposed system, we have applied the labeling technique for annotation of the images and generated a manual box around the affected part of the lungs under the guidance of a medical expert. These annotations were saved in XML files that have three classes named "High", “Medium” and “low” for the affected part of the lungs and other properties such as x-minimum, y-minimum, x-maximum, y-maximum, width, and height according to the class. For all images, there exists an XML file that was employed to create the CSV file. At last, the training file is created to get benefit in a training phase.

Classification

For classification of the images, according to COVID-19 virus severity effect on lungs, we apply trained faster RCNN. The Faster RCNN is trained on images of the lungs with three severity levels based on the bounding box area. COVID-19-related training is based on photos with a bounding box around the afflicted region of the lungs.

Using factors such as height and width to show the diseased section of the lungs, we calculated the maximum area of the bounding boxes. The severity level is classified as mild if the region covers less than 25% of the lungs. The severity level is classed as medium if the area covers between 25% and 50% of the lungs. The level is classified as high if it is greater than 50%. Furthermore, the model complexity is lowered due to fewer bounding boxes in training photos. Hence, classified images of severity levels of patients, help the physicians to decide the further treatment. The severity level is defined as *S,* computed as described in equation 1.

$$S = \begin{cases} Low & if\ (area < 25\%) \\ Medium & if(area(25\% \leq area < 50\%) \\ High & if(area \geq 50\%) \end{cases} \quad 1$$

Dataset: The proposed model was evaluated on a dataset of 1400 unenhanced chest CT scans (700) [18] and X-ray images (700) [19] of lungs. The age range of patients was from 5 to 85 years. The dataset was divided into 350 non-covid images of X-ray, 350 non-covid images of CT Scans, 350 covid positive images of X-rays, and 350 covid positive images of CT Scans. Most of the patients were having co-existing diseases such as coronary heart disease, hypertension, diabetes, and pneumonia. Images were acquired of the inpatients having positive Reverse Transcription Polymerase Chain Reaction (RT-PCR) test, additionally convoyed by clinical symptoms. The range of scanning was from the apex to the lungs. Moreover, the Image format was DICOM consisting of 16-bits grayscale values with the size of 512x512. Therefore, all images were down-sampled to (227x227 & 224x224) exploited for the training and testing. We selected 60% samples for the training of AlexNet and Faster RCNN and 40% for the testing of the model. For the training of AlexNet and Faster RCNN, images were down-sampled to the resolution of 227x227x3 and 224x224x3 respectively. We altered various parameters in the training phase such as learning rate, epochs, and mini-batch size. We attained the best results for the CT Scan datasets as the images are more high-quality than the X-ray.

Experimental Evaluation.

In this section we evaluated the performance of the proposed model, the metrics that we considered were precision, recall, accuracy, f1-score, and error rates. Both precision (percentage of true positive) and recall (percentage predicted as true positive) are vital metrics for retrieving performance information of the model. We are employing these metrics to evaluate the performance percentage of classification of positive cases using our proposed model and also in our case positive class retrieval should be accurately performed than the negative class as we don't want to miss any covid positive case patient. Therefore, recall should be high as in critical medical cases we can compromise on precision and we cannot compromise on recall. Moreover, true positive (TP), false positive (FP) and false negative (FN) are employed in Precision and Recall as shown in figure 2 and 3. F1-score is the harmonic mean of both precision and recall. It is computed using FP and FN. More precisely, FP represents the number of images classified incorrectly as positive case and FN refers to the cases that are positive in reality however are classified as negative case.. It also provides us significant performance evaluation even in the case of unbalanced dataset. Furthermore, we have used accuracy as performance metric to compute the performance percentage of our proposed model over testing data. We have computed error rate to evaluate the percentage of our testing data that is misclassified.

Formulae of above-mentioned metrics are given below:

$$Precision = \frac{TP}{TP+FP}, \quad 2$$

$$Recall = \frac{TP}{TP+FN} \quad 3$$

$$F1\ Score = \frac{2*(Precision*Recall)}{(Precision+Recall)} \quad 4$$

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

$$Error\ rate = \frac{FP+FN}{TP+TN+FP+FN} \quad 6$$

Confusion matrix is a summary table that exhibited the performance of our mode in term of prediction of the examples of every class. Axes of below matrix shows predicted-labels vs actual labels. Entries of following table tells us the number of positive covid cases the proposed model has correctly predicted i.e. TP, how many positive covid cases it has falsely predicted i.e FP, number of negative covid cases it has predicted correctly i.e. TN, and how many negative covid cases it has falsely predicted i.e. FN. In the following confusion matrix, we have utilized 560 testing images of x-rays and CT-scans. These images include 300 images of positive class and 260 images of negative class. The positive class images showed various severity levels, however in this experiment we are considering results of binary classification. The results are exhibiting the classification performance of the proposed system that is how many cases were correctly classified as positive and negative cases. The confusion matrix is shown below.

We have made the following confusion matrix Microsoft Visio after results calculation.

Total=560	Positive	Negative
Positive	292	8
Negative	6	254

Figure 2. Confusion matrix for binary classification

In the second evaluation step, we computed the performance of our proposed algorithm on both datasets i.e. CT Scans and X-ray images. The performance of our proposed model over both datasets images such as X-rays and CT Scans is considerable. The accuracy of the proposed model is 98.3% over CT Scans images and 98.5% over X-ray images and error rates are 1.7% and 1.5% respectively. Furthermore, overall performance of our model is showing significance of the technique to use it in medical field. The accuracy of the model is 98.4%, precision is 98.15%, and recall is 98.3%. Other than the evaluation results, our proposed algorithm is a generalized model that can be trained on less number of images for classification as compared to traditional machine learning based models which makes it capable to employ it for other disease detection

problems. The detail of the results is reported in Table 1. The graph is shown in figure 3.

Table 1. Performance Evaluation on the Datasets.

Dataset(Image)	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	Error (%)
CT scans	98	98.1	98.05	98.3	1.7
X-ray	98.3	98.5	98.4	98.5	1.5
Average	98.15	98.3	98.22	98.4	1.6

In our second experiment, we trained Faster RCNN, and AlexNet on training data separately to compare results with the proposed hybrid model. Faster RCNN takes a training time of 150 minutes for training images while AlexNet takes 175 minutes. Moreover, AlexNet performs classification on the bases of visual features whereas Faster RCNN employs bounding box for the object detection and performs classification. In our scenario, Faster RCNN is trained over effected are of lungs in Xrays and CT Scans. The bounding box locations representing the affected area of lungs are fed to the model and then classification is achieved to assess the performance. The AlexNet algorithm achieves the 96% accuracy while Faster RCNN attains 96.8% accuracy. Furthermore, our proposed hybrid model achieves the 98.4% accuracy that shows the efficacy of the system. The comparative results are reported in Table 2.

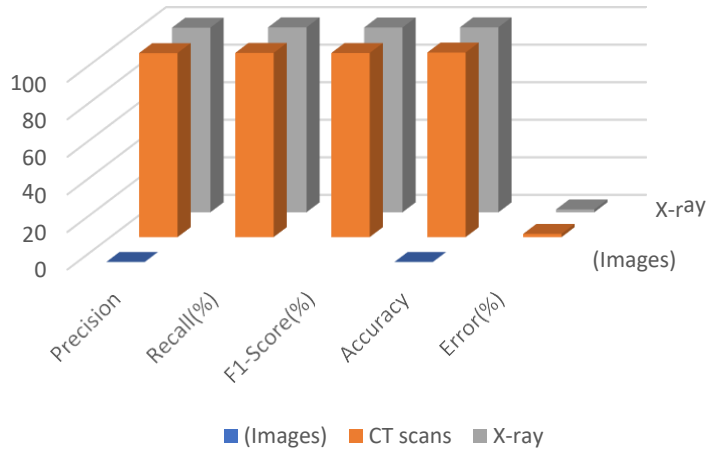


Figure 3. Performance Plot for the Proposed Model.

We have made the graph shown in figure 3 using Microsoft Excel sheets after results generation.

Table 2. Evaluation Metrics for the Proposed System.

Model	Precision (%)	Recall (%)	Accuracy (%)	Error (%)
AlexNet	92	93	96	4
Faster RCNN	94	94.8	96.8	3.2
The Proposed Model	98.15	98.3	98.4	1.6

In our last experiment, we compared the performance of our hybrid deep learning model against the various automated deep learning techniques i.e. Xception, ResNet_V2, DenseNet201, ResNet50, MobileNet_V2, and SqueezeNet employing the

same dataset for training and validation. We tuned training parameters for the proposed system such as initial learning rate of 0.001, batch size of 35, and epochs of 20. Our proposed algorithm takes minimum time for training i.e. 39,200 seconds. Additionally, for efficiency evaluation, we employed different datasets images i.e. CT scans and X-ray images of Covid and pneumonia patients. The algorithm takes the same time for the training of both datasets. In Table 3, the detection performance comparison is shown. From the reported results, it is shown that our hybrid model outperforms the other deep learning methods. The comparison is shown in figure 4. Furthermore, some of the positive class tested images are shown in figure 5.

Table 3. Comparative Analysis with the other Deep Learning Models.

Method of Covid-19 Detection	Accuracy (%)	Time(seconds)
Xception	95.70	49,820
ResNet_V2	93.18	48,921
DenseNet201	90	40,829
ResNet50	89.01	48,920
MobileNet_V2	91	41,293
SqueezeNet	92.3	43,232
The Proposed Model	98.4	39,200

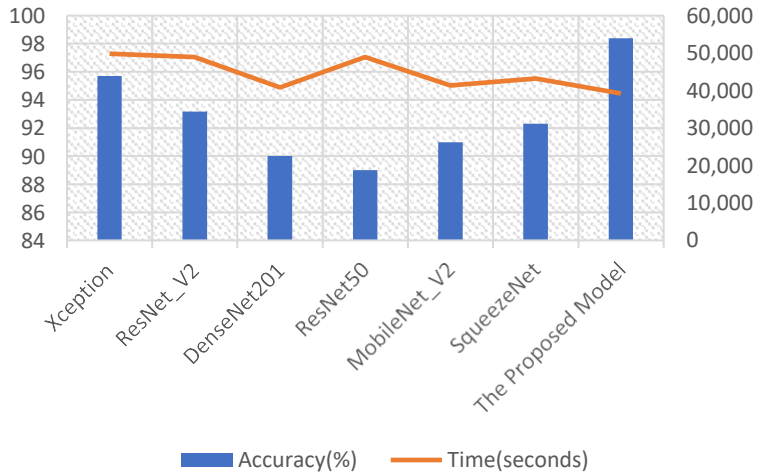


Figure 4. Comparison with the Existing DL Techniques.

Figure 4 is showing the effectiveness of our proposed model in comparison to already existing models.

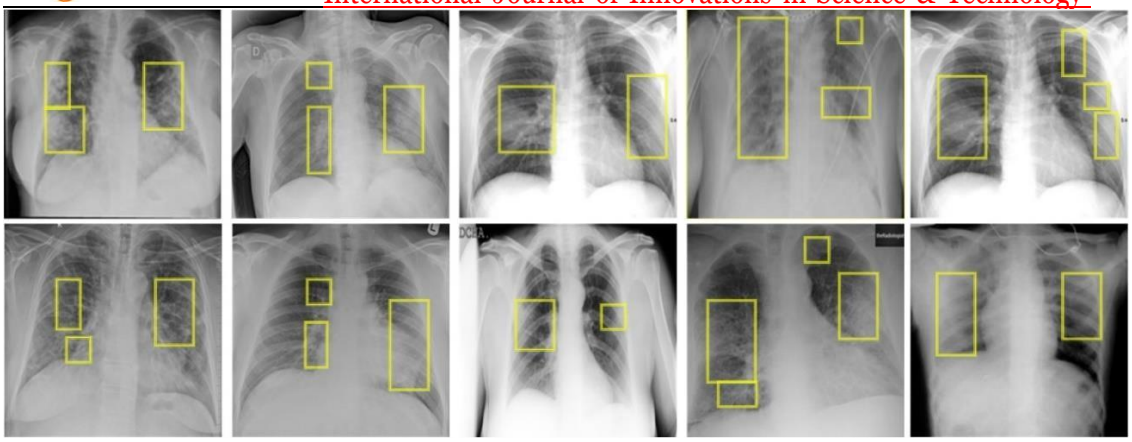


Figure 5. Detection results on test data

The images in figure above were generated after implementation of technique on the **input** images, then after that we have formatted/combined those images in Microsoft Visio.

We have employed various experiments to evaluate the results and efficacy of our proposed model. It is depicted from the results that our proposed algorithm outperformed the existing detection techniques of the corona virus detection. Our proposed model is based on transfer learning mechanism that is already trained over millions of images. Therefore, it can be utilized for other disease detection problems. The training time of our proposed system is also comparable to other existing techniques. Moreover, Faster RCNN algorithm extracts the most representative features from the images due to implication of bounding boxes, which is not a characteristic of the other existing classification models for corona virus detection.

Conclusion.

This study proposes an effective Covid-19 severity detection framework using a combination of deep learning models. We harnessed AlexNet deep learning model for coronavirus detection to filter the virus-affected images. Next, we employed the Faster RCNN algorithm to detect the affected area of the lungs due to the virus. The severity of the ailment was determined based on the area. X-rays and CT scans of the lungs are maneuvered to assess the examination. For the training and performance evaluation of the proposed system, a total of 1400 photos were employed. The average accuracy of our proposed model is 98.3 percent, indicating the framework's effectiveness. The performance of our proposed framework degrades to some extent when X-ray images of pneumonia were classified as Covid-19 positive. We are planning to overcome this restraint in the future.

Acknowledgement. We thank Dr. Syed Aun Irtaza (Chairman CS Dept. at University of Engineering and Technology, Taxila) for his kind support by supervising and reviewing our work and we also thank Rabbia Mahum (Lecturer CS Dept. at University of Engineering and Technology, Taxila) for her supervision and assistance in technique implementation and drafting. This research was supported by University of Engineering and Technology, Taxila.

Author's Contribution.

1. RabbiaMahum, Muhammad Haris Munir : Drafting and Implementation of Technique.
2. Muhammad Nafees, Muhammad Aitazaz : Validation of Results.
3. Dr. Syed AunIrtaza : Reviewed.

Conflict of interest. The Authors have no conflict of interests.

Project details. Not Applicable.

References

[1] Salehi, Ahmad Waleed, "Review on machine and deep learning models for the detection and prediction of Coronavirus," *Materials Today: Proceedings,* vol. 33, pp. 3896-3901, 2020.

[2] C. Ouchicha, "CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest x-ray images," *Chaos, Solitons & Fractals,* vol. 140, pp. 110245, 2020.

[3] Bhattacharya Sweta Maddikunta, Praveen Kumar Reddy Pham, Quoc-Viet Gadekallu, Thippa Reddy Chowdhary, Chiranji La Alazab, Mamoun Piran, Md Jalil, "Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey," *Sustainable cities and society,* vol. 65, pp. 102589, 2021.

[4] Pan, Feng, Ye, Tianhe, Sun, Peng, Gui, Shan, Liang, Bo, Li, Lingli, Zheng, Dandan, Wang, Jiazheng, Hesketh, Richard L, Yang, Lian, "Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia," *Radiology,* 2020.

[5] Xu, Xiaowei Jiang, Xiangao, "A deep learning system to screen novel coronavirus disease 2019 pneumoni," *Engineering,* vol. 6, no. 10, pp. 1122-1129, 2020.

[6] Song, YingZheng, ShuangjiaLi, LiangZhang, XiangZhang, XiaodongHuang, ZiwangChen, JianwenWang, RuixuanZhao, HuiyingChong, Yutian, "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," *IEEE/ACM transactions on computational biology and bioinformatics,* vol. 18, no. 6, pp. 2775-2780, 2021.

[7] El Asnaoui, Khalid, "Automated methods for detection and classification pneumonia based on x-ray images using deep learning," in *Artificial intelligence and blockchain for future cybersecurity applications*, Springer, 2021, pp. 257-284.

[8] Rajaraman, S., Siegelman, J., Alderson, P. O., Folio, L. S., Folio, L. R., & Antani, S. K., "Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays.," *Ieee Access,* vol. 8, pp. 115041-115050, 2020.

[9] Gozes, Ophir, Frid-Adar, Maayan, Sagie, Nimrod, Zhang, Huangqi, Ji, Wenbin, Greenspan, Hayit, "Coronavirus detection and analysis on chest ct with deep learning," *arXiv preprint arXiv:2004.02640,* 2020.

[10] Narin, Ali, Kaya, Ceren, Pamuk, Ziynet, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *Pattern Analysis and Applications,* vol. 24, no. 3, pp. 1207-1220, 2021.

[11] Hemdan, Ezz El-Din, Shouman, Marwa A, Karar, Mohamed Esmail, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055,* 2020.

[12] Farooq, Muhammad, Hafeez, Abdul, "Covid-resnet: A deep learning framework for screening of covid19 from radiographs," *arXiv preprint arXiv:2003.14395.*

[13] Wang, Linda, Lin, Zhong Qiu, Wong, Alexander, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *Scientific Reports,* vol. 10, no. 1, pp. 1-12, 2020.

[14] Shan, Fei, Gao, Yaozong, Wang, Jun, Shi, Weiya, Shi, Nannan, Han, Miaofei, Xue, Zhong, Shen, Dinggang, Shi, Yuxin, "Lung infection quantification of COVID-19 in CT images with deep learning," *arXiv preprint arXiv:2003.04655,* 2020.

[15] Zhang, Jianpeng, Xie, Yutong, Pang, Guansong, Liao, Zhibin, Verjans, Johan, Li, Wenxin, Sun, Zongji, He, Jian, Li, Yi, Shen, Chunhua, "Viral pneumonia screening on chest X-ray images using confidence-aware anomaly detection," *arXiv preprint arXiv:2003.12338,* 2020.

[16] Krizhevsky, Alex, Sutskever, Ilya, Hinton, Geoffrey E, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* vol. 25, 2012.

[17] Chen, Xinlei, Gupta, Abhinav, "An implementation of faster rcnn with study for region sampling," *arXiv preprint arXiv:1702.02138,* 2017.

[18] "GitHub, GitHub. 2020/11/06: https://github.com/last accessed 2020/11/06.".

[19] Kaggle, "kaggle, Kaggle. 2020/11/06: https://www.kaggle.com/, .".

OPEN ACCESS
International Journal of Innovations in Science & Technology
IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Visualizing Impact of Weather on Traffic Congestion Prediction: A Quantitative Study

Shahrukh Hussain[1], Usama Munir[1], Muhammad Salman Chaudhry[1]
[1] Dept. of Computer Science, FCC University Pakistan.
* Correspondence: Muhammad Salman Chaudhry|salmanchaudhry909@gmail.com.
Citation | Hussain. S, Munir. U, Chaudhry. M. S, "*Visualizing Impact of Weather on Traffic Congestion Prediction: A Quantitative Study*". International Journal of Innovation in Science and Technology. Vol 3, Special Issue, pp: 210-222, 2022.
Received | Dec 15, 2022; Revised | Feb 17, 2022 Accepted | Feb 18, 2022; Published | Feb 20, 2022.

A substantial amount of research has been done to develop improved Intelligent Transportation Systems (ITS) to alleviate traffic congestion problems. These include methods that incorporate the indirect impact on traffic flow such as weather. In this paper, we studied the impact of weather conditions on traffic congestion along with more spatial and temporal factors, such as weekdays/time and location, which is a different approach to this problem. The proposed solution uses all these indicators to estimate the flow of traffic. We evaluate the level of congestion (LOC) based on the traffic volume grouped in certain regions of the city. The index for the defined LOC indicates the traffic flow from "free -flowing" to "traffic jam". The data for the traffic volume count is collected from the Department of Transportation (DOT) for NYMTC. Weather conditions along with special and temporal information have an essential role in predicting the congestion level. We used supervised machine learning for this purpose. The prediction models are based on certain factors such as the volume count of the traffic at the entry and exit point of each street pair, particular days of the week, timestamp, geographical location, and weather parameters. The study is done on the major roadways of each of the four prominent boroughs in New York. The results of the traffic prediction model were established by using the Gradient Boosting Regression Tree (GBRT) which showed an accuracy of 97.12%. Moreover, the calculation speed was relatively fast, and it has stronger applicability to the prediction of congestion conditions.
Keywords: Gradient Boosting; Decision Tree Algorithm; Supervised Machine Learning; Traffic Congestion

INTRODUCTION

An enormous increase in urban traffic has been observed recent times, globally [11]. The overall process of modernization is speeding up, leading to the rapid growth of vehicular traffic on roads [12]. To cater the needs for a huge surge in traffic, urban road networks are becoming over complex [13]. Consequently, urban traffic problems are getting serious and traffic congestion is one of them [1]. In metropolitan cities, if the factors leading to the congestion are neglected or, congestion is not predicted and reported

properly to the users in time, it can lead the road networks to be paralyzed [14]. The early step to tackle the problem of congestion is to prevent it from happening [14]. Therefore, the establishment of traffic flow, forecasting with respect to the day and time of the day is conducive to the preparation of targeted preventive measures which serve as an early warning [15]. The usage of Intelligent Transportation Systems (ITS) to predict traffic-related information has gained popularity in the field of smart transportation. A well-designed ITS can estimate and inform drivers of the locations and time frame of congested road sections, thus giving them a warning to avoid taking that route [16]. Moreover, it can also provide a significant amount of information for authorities of large metropolitan areas in order to control the parameters of the traffic signal to reduce the Level of Congestion (LOC) [17].

Supervised machine learning models are highly effective and fast with training structured data [18]. However, performance and accuracy of the model is highly dependent upon the dataset since its correct input features and labeling are followed by minimum null values which define accuracy of a model in a real-world [18]. These models are expected to generate adequate results with precision as the datasets become more diverse [19].

In this research paper we have used the supervised machine learning models to estimate the traffic flow and congestion of recent times. The correlations between implicit traffic-related data and weather condition data define influence of values on each other. A detailed exploratory analysis was performed over important weather features that impact congestion on the roads the most, to unveil individual impacts over the traffic flow within the given route at a certain time and day of the week.

The objectives of this research were to examine and evaluate granular relationships between external factors (Weather and ToD (time of day) in our case) with "Traffic Counts" within an area by the use of Supervised Machine Learning Algorithms. The conclusion we reached enabled us to reach a concise evaluation of these 2 factors and paved the way for a future deep-dive into other factors to further quantify an expected traffic count within an area, based on those factors. The objectives were achieved as relationships were established.

BACKGROUND

Researchers from different domains have studied the problem of traffic flow and congestion using various techniques in the past. Statistical analysis is based on a variety of features that lead to the measurement of congestion of vehicles on roads such as motion of the vehicle, stationary time of the vehicle, the velocity of the vehicle, or the cluster of the vehicles within the selected segment of the road network. Data collection is the first step to solve the traffic problems. Various methods including GPS-based [2] and cellular-based [3] sensors installed in smartphones, vehicles, and roadsides, to gather data of geographical location and timestamp. Modern traffic solution are based or safe city cameras and drones etc. to extract the vehicle data from intersections, highways, and freeways. [9] Both the supervised and unsupervised machine learning algorithms have played an integral part in predicting the congestion based on the feature, labeled and unlabeled. In our study, we took the data from each entry and exit of each pair of streets to estimate the traffic congestion and to measure the influence of weather on the overall pattern of traffic flow The main objective of this research is to apply a supervised machine

learning model to predict the traffic flow and to examine the impact of weather condition on traffic congestion

RESEARCH QUESTIONS

The following are the research problems that we tackled in our study:

We have used the supervised machine learning model to estimate the traffic flow and congestion in the study. The correlations between implicit traffic-related data and weather conditions, define the influence of values on each other. A detailed exploratory analysis was performed over important weather features that impact congestion on the roads, to unveil individuals' impacts over the traffic flow within the given route at a certain time and day of the week.

RESEARCH METHODOLOGY

This study utilized various approaches to analyze and use supervised machine learning to produce adequate results with minimal error.

Approach

Supervised machine learning approach has been used in this research to study the problematic statement. With the experiments conducted on the dataset, we were able to make clear judgments, according to the supervised machine learning algorithms for fast and relatively accurate results. All the inputs (features) and outputs are labeled in the dataset which are required to train the model. It is also important to note that the supervised machine learning is mainly used to deal with two problem-sets: classification and regression. For accurate prediction of traffic congestion, the model was trained based on continuous values of the traffic count. Congestion level of traffic count was varied with the help of the classification model. The accuracy and precision of scaled traffic count is used in the research to define the Level of Congestion (LOC).

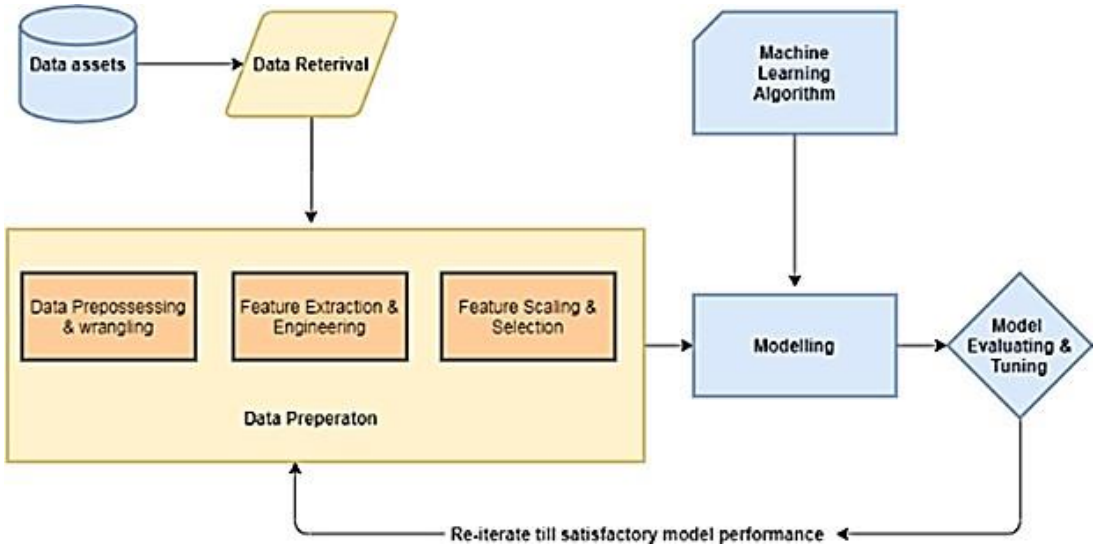


Figure 1. Supervised Machine Learning Pipeline

The basic pipeline can be seen in Figure 1. The workflow for the proposed system in this research diverse involves the ingestion of raw data that has been obtained from source and then applying the data processing techniques to wrangle, processing and engineering meaningful features and attributes from this dataset. All these data preparation

steps allow us to train the model which is later used in the testing phase and then for hyper parameter optimization. It helped us in evaluation of a machine learning algorithm that is suitable for the data. The features selected from the data were used in the deployment of the model. The regression- model was used to experiment with the dataset as data comprises of continuous values. The regression model includes independent variables pertaining to month, hour, and a dummy variable for weekends, the geographical clusters variable, and direction, as well as weather data.

Dataset Description

Regarding data acquisition, we took the monthly historical traffic data for the month of March 2018 from the Department of Transportation (DOT) for New York Metropolitan Transportation Council (NYMTC). This data was comprised of its four major Boroughs showing a decent distribution of values found within each Borough with no visible skewness or anomalies observed as stated in .Table 1

Table 1. Borough Distribution

No.	Borough	Number of Values	Percentage of Dataset
1	The Bronx	1368	27.94%
2	Queens	1152	23.53%
3	Manhattan	1320	26.96%
4	Brooklyn	1056	21.57%

. The geographical data of 22 streets in boroughs was retrieved using the Google Maps API and acquired the weather data from the external source of same geographical locations and timestamps in our dataset. Weather data acquired of same location had 9 attributes such as cloud cover, precipitation, dew point, relative humidity, precipitation cover, temperature, visibility, conditions, and wind chill.

RESULTS AND DISCUSSION

Exploratory analysis

Various methods and approaches were used for scaling, standardization, and normalization of the dataset for model training and testing to obtain satisfactory outcomes. The parameters comprise of the independent variables – Date, Hourly Time (e.g. 9-10 a.m.), Weather features which further comprise of Conditions, Precipitation, Cloud Cover and Visibility (scaled to obtain a convenient severity level between 0 and 3) to be used for our Regression Model. The dependent variable comprises of the parameter Traffic Count.

Scaling Traffic Count

Traffic count, being a continuous value within the data, comprises of values ranging from 0 to 3000 that need to be normalized to a better form to relate it with various features and to extract useful relations with them. For a vigilant representation of co-relations, Traffic Count Label was scaled and grouped by the respective start and end locations. Values were scaled using min-max scalar and scaled into 4 equal divisions from 0 to 3 and termed as Congestion Divisions. The approach behind the min-max scalar is to subtract the minimum value in the feature and then divide it by the range. The range is the

difference between the original maximum and original minimum as shown in the Equation 1. The advantage of using a scalar is that the shape of the original distribution is preserved.

$$x_{sc} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Using this min-max scalar, the Level of Congestion (LOC) based on traffic count has been defined and grouped into its respective Borough, which is also normalized. LOC was categorized into 4 discrete values: 0 represents low congestion level, 1 represents mild congestion level, 2 represents slightly high congestion level and finally, 3 represents high congestion level. The distribution of the overall traffic count is shown in the described range from 0 to 3.

Relations with weather features

Based on the domain knowledge, four relevant weather features (Conditions, Precipitation, Cloud Cover, Visibility) were selected that could influence the traffic count. Values were scaled to obtain a convenient severity level between 0 and 3, . The scaled divisions were made, based on domain knowledge related to that feature. Finally, values were compared with Scaled Traffic count to observe the strongest intersection of similar severity values to depict influence. The greater the number of overlapping of similar severities, the greater will be the influence.

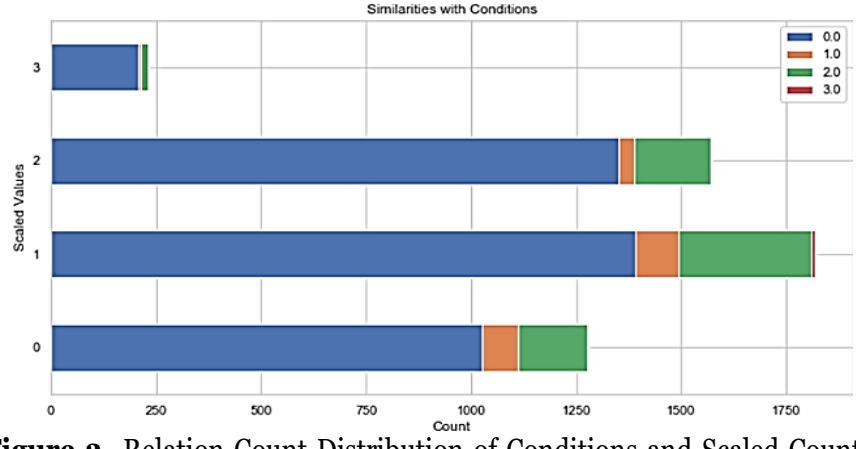


Figure 2. Relation Count Distribution of Conditions and Scaled Count

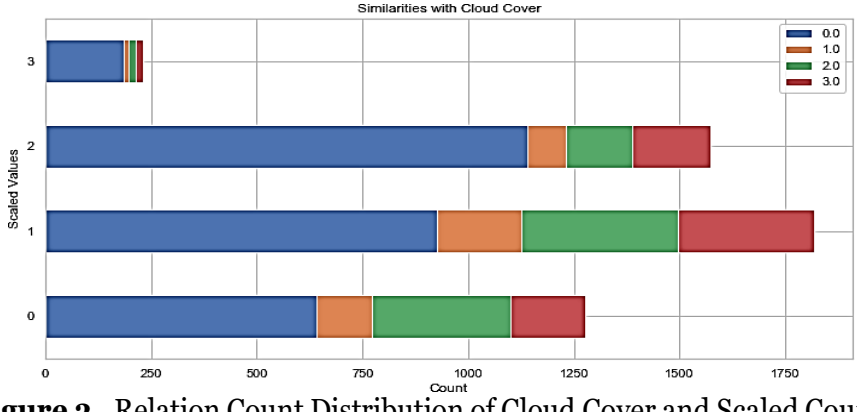


Figure 3. Relation Count Distribution of Cloud Cover and Scaled Count

Comparison of scaled severity values of weather features with Traffic Count showed that greatest intersections arose with Conditions and Cloud Cover as shown in Figure 2. And Figure 3. Further evaluation was made in the Feature Analysis section.

Relations with Time of Day

Peak times were to be analyzed from the entire dataset i.e., all given locations to depict where the greatest traffic count was observed in case of a Weekday and a Weekend, to evaluate the importance of 'Type of Day'.

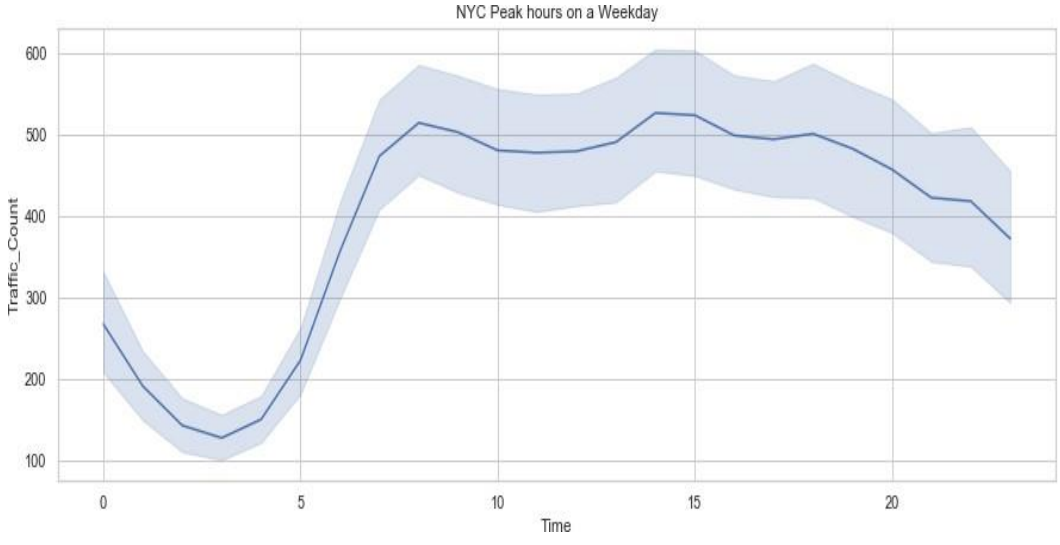


Figure 4. Peak hours on a Weekday

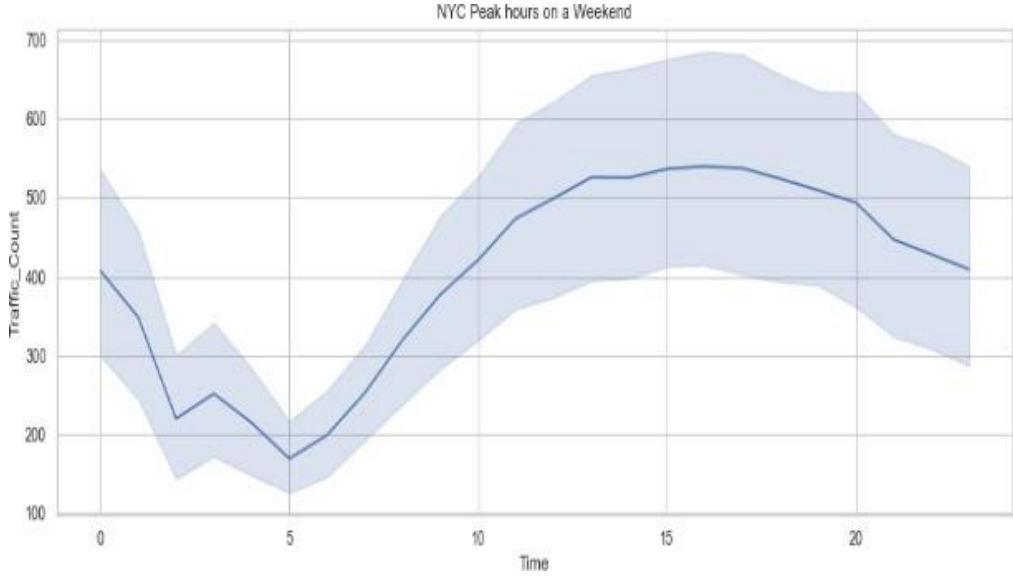


Figure 5. Peak hours on a Weekend

According to both figures, the peak hour ranges showed distinguishable variations on both Weekdays and Weekends. On a weekday, the peaks were found to be on 2-3 pm and 8-9 pm and for the weekends, these were found to be between 2-7 pm, indicating greater peak hours and traffic counts on Weekends.

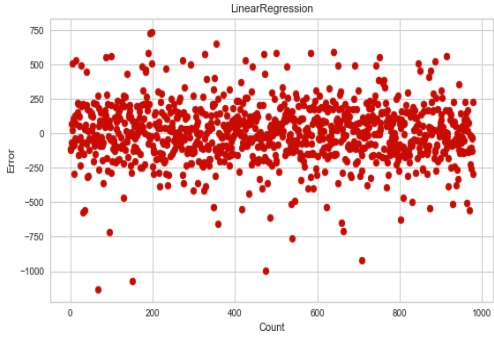
MODEL SELECTION

Base Accuracy

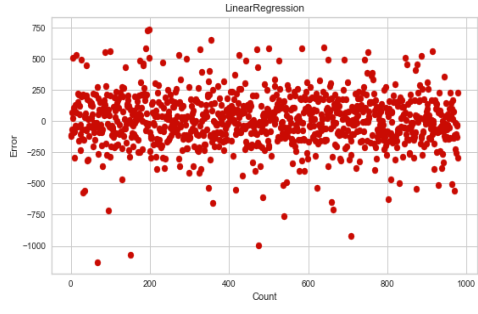
Since our target label i.e., Traffic count was observed continuous in nature, Regression Model was implemented as a means to provide the best coordination. Tested models comprised of Linear Regression, Lasso Regression, Decision Trees, Random Forest Regression (RF) [6], and Gradient Boosting Regression Tree (GBRT). Among all, Random Forest showed the best results with an accuracy of 96.14%.

Actual VS Predicted

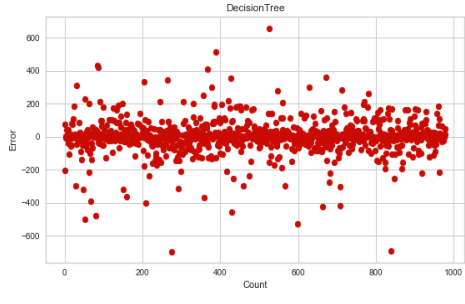
The accuracy Score of all the models was validated by relating their predicted values with the original targeted values by plotting them against each other and subsequently, visualizing the integrated result. The minimum difference in values i.e., which were found on densely populated sites closer to the baseline was considered as possessing least error, while the ones distant from them can be termed as anomalies as shown in Figure 6.



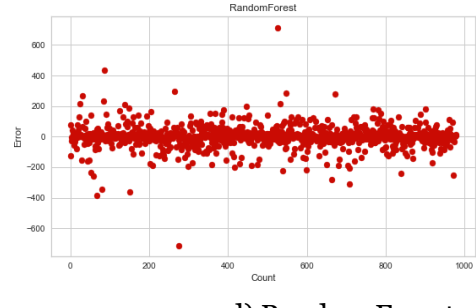
a) Linear Regression



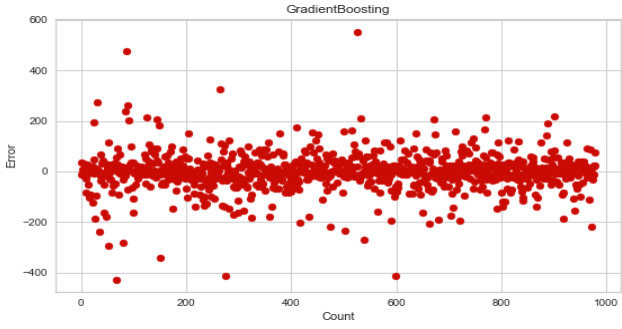
b) Lasso Regression



c) Decision Tree



d) Random Forest



e) Gradient Boosting Regression Tree

Figure 6. Model Accuracy VS Precision. Number of data values represented on x-axis and error on y-axis. Plotted coordinates represent the difference between actual and predicted values with density of points near base indicating greater accuracy.

Hyper parameter Optimization

After optimizing the hyper-parameters of all the regression models selected for this research, it was observed that the base accuracy of the Gradient Boosting Regression Tree was increased from 92.39% to 97.12% using the Grid Search CV and validation of the model was evaluated by the shuffle spilt validation. Overall, we observed an improvement of 5.12% in the model accuracy. The gradual increment can be observed in Table 2.

Table 2. Parameter-Tuning Accuracy

Accuracy	n-estimators	Max-depth	Min-sample-leaf	Max-sample-split
92.39%	100	3	1	2
94.38%	200	3	1	2
95.90%	100	5	1	2
92.83%	100	3	2	2
92.83%	100	3	1	5
96.22%	200	5	2	5
96.42%	400	5	2	5
96.62%	400	7	2	5
97.12%	400	7	5	5
96.67%	600	7	5	10

MODEL EVALUATION

Although after rigorous experimentation with optimized hyper-parameters of the selected models, the correct validation results were generated to evaluate the performance

of the model on training data. However, the fit of a proposed regression-based model should therefore be better than the fit of the mean model, so all models were evaluated via R2 error in which the ratio of the variance of the model and the total variance of target variable was taken and put forward a value between 0 and 1 with 1 being the best one. MAE was also used to identify the difference between the forecasted value and the actual value. The general definition of the R2 score can be seen in Equation 2.

$$R^2 \equiv 1 - \frac{SS_{res}}{SS_{tot}} \tag{2}$$

Where,

$$SS_{res} = \sum_i (y_i - \hat{y})^2 \tag{3}$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \tag{4}$$

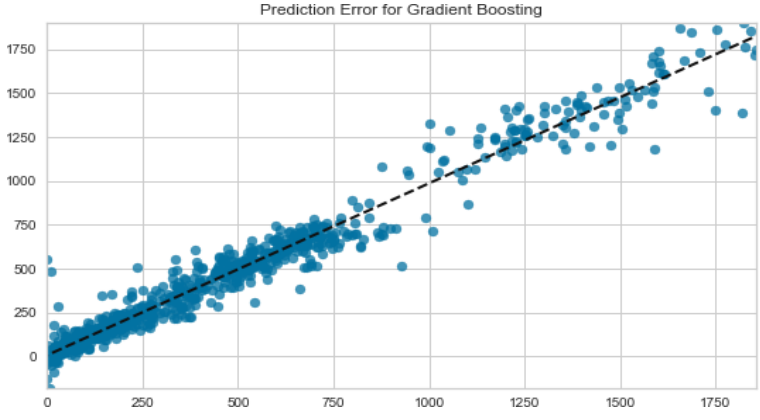


Figure 7. R2 Accuracy for Gradient Boosting

Feature analysis

Trained model was tested for co-relations among all the existing features which showed vibrant relations of Traffic Count with weather as standing out among all other weather features with the greatest feature importance, also observed in Figure 5, similar to our prior exploratory analysis in the section.

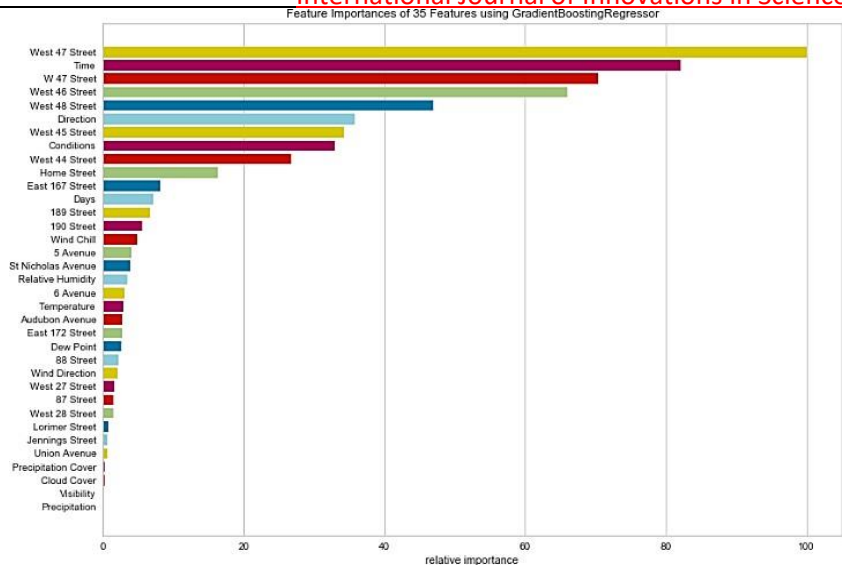


Figure 8. Feature Importance

Discussion

On successfully recording, the highest accuracy of our model GBRT, the predicted results were visualized via a heat map through the Folium library. The visualized points indicated the coordinates of the start and end of the street locations of the designated routes within the data. Furthermore, the intensity of the color of the marker of that location expresses the intensity of traffic flow in that particular route. It ranges into 4 severity levels of congestion with a color palette of blue, green, and yellow. The red color indicates a relatively high traffic flow and a darker blue color represents a relatively low traffic flow. Visualized results for a time duration of 7:00-8:00 am describe the peak hour for the traffic flow, leading to congestion, as seen in the Figure 6. With the help of the visualization, the traffic pattern can be studied and understood.

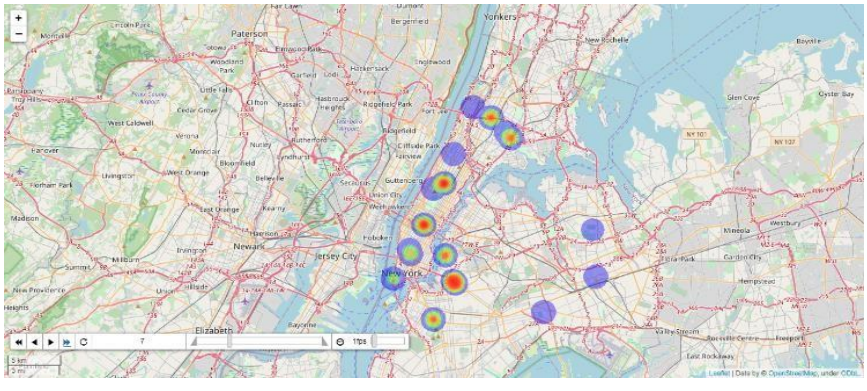


Figure 9. Visualized Results

CONCLUSION AND FUTURE SCOPE

After detailed analysis, the conclusion can be made that external factors like weather (in our case) do impact traffic congestion as a whole for the given dataset and that our Gradient Boosting Regression Tree model records the best accuracy score for

predicting Traffic Count parameters i.e. 97.12% considering all relevant features as mentioned in Table 3. On the contrary, the greatest feature importance examined among all the weather features was Weather Conditions, followed by Wind Chill.

Table 3. Model Accuracy

Model	Base Accuracy	After Hyper parameter Tuning
Linear Regression	75.23%	71.60%
Lasso Regression	75.22%	71.50%
Decision Tree	93.24%	94.10%
Random Forest	96.14%	96.80%
Gradient Boosting	92.39%	97.12%

In this paper, the model is based on regular day predictions. However, we look forward to the implementation of a more robust model that will also consider the Planned Special Events (PSEs), like festival holidays, social events like concerts, sporting events like cricket and football matches and so on. Moreover, seasonal changes may also affect the traffic flow adversely because of the ambiguities in weather they may bring. Therefore, we also look forward to work with this aspect. We believe our research will pave the way for greater opportunities in the field of data gathering and will help in developing a more stabilized road network of the city that is less prompted to traffic congestion.

Conflict of interest. The Authors have no conflict of interests.
Project details. Nil.

REFERENCES

[1] Sweet, M. “Traffic Congestion’s Economic Impacts: Evidence from US Metropolitan Regions”. 2013 SAGE journals-Urban Studies, vol. 51, pp. 2088-2110, 2013

[2] Thianniwet, Thammasak & Phosaard, Satidchoke & Pattara-atikom, Wasan. “Classification of Road Traffic Congestion Levels from GPS Data using a Decision Tree Algorithm and Sliding Windows”. 2009 World Congress on Engineering, vol. 1, 2009

[3] Jahangiri, A., &Rakha, H. A. “Applying Machine Learning Techniques to Transportation Mode Recognition Using Mobile Phone Sensor Data”. IEEE Transactions on Intelligent Transportation Systems, vol. 16, pp. 2406–2417, 2018

[4] Jayapal, C., & Roy, S. S. “Road traffic congestion management using VANET”, 2016 International Conference on Advances in Human Machine Interaction (HMI), 2016

[5] P. Chhatpar, N. Doolani, S. Shahani, and R. Priya, “Machine learning solutions to vehicular traffic congestion,” 2018 International Conference on Smart City and Emerging Technology (ICSCET), 2018.

[6] Y. Liu and H. Wu, "Prediction of road traffic congestion based on random forest," 2017 10th International Symposium on Computational Intelligence and Design (ISCID), vol. 2, 2017

[7] Chowdhury, B.,Kinhikar, M. &Alleema, N. N. "Road Traffic Prediction using Machine Learning". International Research Journal of Engineering and Technology (IRJET). Vol. 06, 2019

[8] M. M. Chowdhury, M. Hasan, S. Safait, D. Chaki, and J. Uddin, "A traffic congestion forecasting model using cmtf and machine learning," 2018 Joint 7th International Conference on Informatics Electronics and Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision and Pattern Recognition (icIVPR), 2018

[9] Huang, F.-R., Wang, C.-X., & Chao, C.-M. "Traffic Congestion Level Prediction Based on Recurrent Neural Networks". 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 2020

[10] Jia Lu, & Li Cao. "Congestion evaluation from traffic flow information based on fuzzy logic". 2003 IEEE International Conference on Intelligent Transportation Systems, 2003

[11] Sun Ye. "Research on Urban Road Traffic Congestion Charging Based on Sustainable Development". International Conference on Applied Physics and Industrial Engineering, Physics Procedia 24 pp. 1567–1572, 2012

[12] Ge Shi, Jie Shan, Liang Ding, Peng Ye, Yang Li, Nan Jiang. "Urban Road Network Expansion and Its Driving Variables: A Case Study of Nanjing City". International Journal of Environmental Research and Public Health, 16, pp. 2318, 2019

[13] Gao ZH, Chen ZJ, Liu YX, Huang K (2007) Study on the complex network characteristics of urban road system based on GIS. Proceedings of SPIE 6754:67540N

[14] S R Samal1, P Gireesh Kumar2, J Cyril Santhosh3, and M Santhakumar. "Analysis of Traffic Congestion Impacts of Urban Road Network under Indian Condition". IOP Conf. Series: Materials Science and Engineering, 2020

[15] Z. Yin, J. Wang and H. Lu, "A Study on Urban Traffic Congestion Dynamic Predict Method Based on Advanced Fuzzy Clustering Model," 2008 International Conference on Computational Intelligence and Security, 2008, pp. 96-100, doi: 10.1109/CIS.2008.194.

[16] Jarašūnienė, Aldona. (2007). Research into Intelligent Transport Systems (ITS) technologies and efficiency. TRANSPORT. 22. 61-67, 2010

[17] Wei-Hsun Lee and Chi-Yi Chiu. "Design and Implementation of a Smart Traffic Signal Control System for Smart City Applications". Sensors 2020, 20, 508; doi:10.3390/s20020508

[18] Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592-x

[19] Abdulraheem, Ajiboye & Abdullah Arshah, Ruzaini & Qin, Hongwu. (2015). Evaluating the Effect of Dataset Size on Predictive Model Using Supervised Learning Technique. International Journal of Software Engineering & Computer Sciences (IJSECS). 1. 75-84. 10.15282/ijsecs.1.2015.6.0006.

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Report Generation of Lungs Diseases from Chest X-ray using NLP

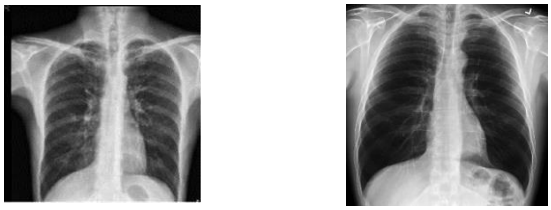
Iqra Naz[1], Shagufta Iftikhar[1], Anmol Zahra[1], Syeda Zainab Yousuf Zaidi[1]
[1] NUML, National University of Modern Languages, Islamabad. Pakistan
* Correspondence Author's Iqra Naz Email: bscs-ms18-id019@numls.edu.pk.
Citation | Naz. I, Iftikhar. S, Zahra. A, Zaidi. S. Z. Y, "Report Generation of Lungs Diseases from Chest X-ray using NLP". International Journal of Innovations in Science and Technology. Vol 3, Special Issue, pp: 223-233, 2022.
Received | Jan 26, 2022; Revised | Feb 10, 2022 Accepted | Feb 24, 2022; Published | Feb 26, 2022.

Pulmonary diseases are very severe health complications in the world that impose a massive worldwide health burden. These diseases comprise of pneumonia, asthma, tuberculosis, Covid-19, cancer, etc. The pieces evidence show that around 65 million people undergo thechronic obstructive pulmonary disease and nearly 3 million people pass away from it each year that makes it the third prominent reason of death worldwide. To decrease the burden of lungsdiseases timely diagnosis is very essential. Computer-aided diagnostic, are systems that supportdoctors in the analysis of medical images. This study showcases that Report Generation Systemhas automated the Chest X-Ray interpretation procedure and lessen human effort, consequently helped thepeople for timely diagnoses of chronic lungs diseases to decrease the death rate. This system provides great relief for people in rural areas where the doctor-to-patient ratio is only 1 doctor per 1300 people. As a result, after utilizing this application, the affected individual can seekfurther therapy for the ailment they have been diagnosed with. The proposed system is supposed to be used in the distinct architecture of deep learning (Deep Convolution Neural Network), this isfine tuned to CNN-RNN trainable end-to-end architecture. By using the patient-wise official split of the OpenI dataset we have trained a CNN-RNN model with attention. Our model achieved an accuracy of 94%, which is the highest performance.
Keywords: Attention, chest X-rays, classification, convolutional neural network, deep learning, natural language processing, pulmonary diseases, recurrent neural network, report generation.

INTRODUCTION

During past few years, the health of human beings has been brutally threatened by pulmonary diseases. Covid-19, Pneumonia, asthma and tuberculosis are the diseases from which millions of people suffer every year. These diseases are not often detected at early stages that leads patients towards the death. There are several diagnostic imaging modalities, for instance Chest X-ray (CXR) shown in Figure 1.

Computed tomography (CT) and magnetic resonance imaging (MRI), are executable nowadays to help in the clinical diagnosis of pulmonary diseases. So, for an examination at early stages chest X-rays are used because they are economical and easily accessible [1].



a) Normal CXR b) Abnormal CXR

Figure 1. Representation of Normal and Abnormal chest radiographs

Now a days massive amount of CXRs are produced worldwide [1]. Nevertheless, deficiency of diagnostic resources for instance trained radiologists are major hindrance for robust pulmonary diagnostics [2]. For the diagnosis of pulmonary diseases with massive authenticity is the tough duty even for radiologists. Moreover, there is a shortage of expert radiologists in under developed countries who are competent to read chest radiographs. As opposed to diagnosis with the help of humans, more convincing results with less diagnostic errors can be achieved by computer aided design (CAD) tools [3]. Therefore, researchers proposed different lungs disease detection systems and developing the programmed algorithms by using chest radiography for the detection of pulmonary diseases.

Many lungs disease detection and report generation systems exist like Thorax-Net [1] and Tie Net [4] is briefly discussed in related work section. However, there are still some weaknesses in those existing system. Previous researches have predominantly been enthused by the aspiration to create an automated diagnostic system that can boost the survival rate of patients. However, correctly interpreting the information is always a major challenge. The number of publications and researches in this field are more but progress toward the performance of systems is slow.

Recently, pre-trained language models are getting across-the-board achievement which inspired researchers to concentrate generating multimodal representation using Transformer-based pre-trained models [5]. To our best knowledge, developing an automated report generation system for pulmonary diseases is still a challenging task. This system will be a deep convolutional neural network with regularized attention that will classify the pulmonary diseases on the basis of CXR. Numerous revolutions in the applications of computer vision have been attained by Deep Learning methodologies. The tremendous success in the field of deep learning has provoked many investigators/researchers to employ DCNN. This model contains the branch of classification that determines the image representations and an attention branch works for acquiring distinguished attention maps for better performance of classification. We have used an attention branch, CheXNet [6] is used as classification branch then outputs of mutual branches are aggregated for identification of individual input. Conversely, an automatic radiological diagnosis and reporting system like this system is a crucial framework to build.

This paper enlightens a study assessing the viability of a CNN-RNN with attention-based report generation system to diagnosed disease in CXR. We estimated sensitivity,

specificity, and positive predictive value by comparing results of the automated report generation system with other related studies.

There have been recent efforts on pulmonary diseases that we discovered in this section. On the basis of Machine Learning and deep learning there are multiple pulmonary diagnosis schemes i.e. transformers, LSTM and CNN-RNN. Identification of lung disease predominantly depends on patterns of radiology [7]. Timely recognition of Lungs diseases can lead to its cure and stoppage of infectious development.

Tie Net is the model proposed by Xiaosong Wang et al [4]. It classifies the CXRs via image features and text embedding from associated reports. Their [4] auto-annotation framework achieved great accuracy that was over 0.9 in allocation of disease labels for their evaluation dataset that was hand-label [4]. In addition, they transformed the Tie Net into a CXR reporting system. However, it is noted that there are false predictions that can be distinguished in the image that were verified by radiologist but in the original report were not prominent in some way, specifies that to some extent this proposed network can associate the appearance of image with the description of text.

Mendonca, Eneida presented an examination evaluating the achievability of a Natural Language Processing (NLP) [8] based observing framework to screen for medical care related pneumonia in neonates. For the evidence of radiographic abnormalities, this detection system examined radiology reports. They evaluated affectability, explicitness, and positive prescient incentive by different outcomes and clinician's decisions. Affectability was 71% and particularity was 99% [8]. Their results showed that the computerized strategy was efficient, as their framework was NLP- based so there were few limitations as well **1)** Reports that contains incorrect punctuation leads to the errors in recognition of sentence boundaries and poor grammar. **2)** A number of occurrences that missed abbreviations. **3)** Misspelling of terms. Concisely, it does not assure that system will be worthwhile in a clinical setting because many steps were necessary before such a monitoring system which can actually be used.

Horry, Michael J [9] worked on Automated Lung Cancer Diagnosis, separated an independent lung nodule dataset into benign or malignant classes with good results, and deep learning techniques were applied. Maximum accuracy achieved with positive prediction was greater than 81%.

Wang, Hongyu proposed a model Thorax-Net Attention Regularized Deep Neural Network techniques were used in it [1]. By using the patient-wise official split of the Chest X-ray14 dataset Thorax -Net assessed contrary to three state-of-the-art deep learning models and attained maximum per-class accuracy (0.7876, 0.896) without using external training data. But still certain limitations were there like 1) Training dataset was imbalanced. 2) Deficiency of the annotations of pathological abnormalities, leads to the inaccuracy. 3) Many images contained one or more disease labels.

Lovelace et al. [10] developed a radiology report generation model that generates superior reports as measured by both standard language generation and clinical coherence metrics compared to competitive baselines and it utilizes the transformer architecture. Assessing radiological report is difficult as typically language generation metrics that assess image captioning systems cannot be easily examined. They extracted clinical information differentially obtained by their generated reports. By using standard natural language generation (NLG) objective LNLG they first train a report generation model and then further fine-tune the model to make it extra medically clear by presenting a further medical clear

objective. Though these metrics estimate language resemblances among produced and actual reports, they cannot calculate proficiency of clinically correct reports are the produced by the models.

Table 1. Comparison of both approaches

	Authors	Year	Dataset	Methodology	Results
1.	Wang, Xiaosong et al.	2018	Three datasets naming: ChestX-ray14, Hand-labeled, OpenI	Multi-level attention models integrated into an end-to-end trainable CNN-RNN architecture	90 % accuracy
2.	Wang, Hongyu et al.	2019	Chest X-ray 14 dataset	Attention Regularized Deep Neural Network	(78%, 89 %) accuracy in both experiments
3.	Horry, Michael J et al.	2020	LIDC-IDRI dataset	Deep learning techniques	greater than 81% accuracy
4.	Mendonça et al.	2005	Dataset from the National Institute of Nursing Research	NLP-based observing framework	Sensitivity was 71% and specificity was 99%
5.	Lovelace et al.	2020	MIMIC-CXR dataset	transformer architecture and natural language generation (NLG) objective LNLG	(90%, 97%) accuracy

Material and methods. In the following Section, a powerful pulmonary diagnostic strategy has been described on the basis of CNN-RNN architecture with attention mechanism. We have explored two methodologies that are trained on standard dataset OpenI [11] by Indiana University. We have predicted the results of our model trained on OpenI dataset for the predictability and responsiveness.

The proposed model comprises of the classification branch which learns the image representation and extracts the features and an attention branch pays much attention to probable abnormal regions. As a final point, to produce the results of diagnosis both branches are joined.

Dataset

Machine learning projects require a dataset to train and test the model and for that purpose we used OpenI (It is freely available dataset by Indiana University) [11]. Indiana University has assembled radiography dataset from multiple institutes that is publicly accessible. We acquired 3955 distinctive CXRS and related 3820 frontal and 3646 lateral images. These images were marked up with key concepts that comprises findings, diagnoses and body parts from the OpenI dataset. In our evaluation, we have selected 3820 distinctive images of front view and their parallel reports.

Initially, portable network graphics (PNG) format images were saved in dataset. Afterwards, rescaling of images to 1024 × 1024 size feature vector was performed. The dataset was divided into a training, testing and validation subset on the basis of patient ID. Identical patients' information appears only just once in any of the training or testing set. The training set contained 2506 images, testing set was comprised of 200 images and validation set consisted of 500 images.

Preliminaries of Deep Learning

For the proposed system we tested two different methodologies, CNN-RNN and CNN-RNN with attention mechanism.

Basic architecture of CNN

CNN key architecture comprises of the layers listed below:

CONVOLUTION LAYER

The convolution layer [12] units were connected with corresponding local patch units achieved by filtered prior layer. Above the locally weighted sums, feature map triggers the units which was done by Rectified Linear Unit (ReLU) [13].

POOLING LAYER

Features and patterns were integrated semantically with private feature map by this layer. From the former layer extreme or average of input features were calculated and used like feature map output.

FULLY-CONNECTED LAYER

A mesh was created by attaching each single unit with preceding layer. Generally, to extract features 2 or 3 stacks of the pooling and convolution layers were positioned before fully-connected layer [14].

SOFTMAX LAYER

This layer transforms each output class features into probabilities. Softmax layer's entire quantity of units is equivalent to whole output classes. In Softmax function equation (1) Softmax (b_p) and b_i symbolize probability and feature of p^{th} class correspondingly. Non-normalized probability measurement is e^{p}_{b} and for normalizing distribution of probability form output classes [15] $\sum_{q=1}^{m} e_q^b$ is used as in equation:

$$\text{Softmax}(b_p) = \frac{e_{bp}}{\sum_{q=1}^{m} e_q^b} \qquad (1)$$

For exhibiting non-linearity along with attaining quicker learning convergence [13] as an activation function ReLU can be used. To reduce inaccurate classification weight optimization of units is considered at learning stage. Stochastic gradient descent optimizer is usually used for Figured block propagation algorithm weights for applying gradient.

Basic architecture of RNN

Basic architecture of RNN contains following layers.

EMBEDDING LAYER

Embedding Layer allows a static form vector to be translated from each word. The resulting dense vector consumes real values as a substitute of just 0's and 1's. The fixed length of word vectors supports us to denote words in an improved way accompanied by reduced dimensions.

LSTM LAYER

LSTM layer [16] allows to state the several memory units inside the layer. Every unit or cell inside the layer has an inner cell state, and outputs a hidden state.

DENSE LAYER

It is used for outputting a prediction and is called Dense Layer [17]. It is a fully connected layer that usually follows LSTM layers.

Attention Mechanism

This proposed mechanism is an enhancement of the CNN-RNN models. The issue of context vector was bottleneck for encoder-decoder models which made it difficult to deal with long sentences. The solution to this problem was "Attention", which has greatly enhanced the quality of translation systems, presented by *Bahdanau et al.* [18] And Luong *et al* [19]. This technique lets the model concentrate on the appropriate regions where desired. Specifically, it forces an extra step in the reasoning procedure, by allocating different importance to features from different image regions it identifies salient regions. The model will find the values of alpha from the data. For this reason, the function is defined as:

$$_{qr} = f_{ATT}(_{Sr-1}, h_q) \qquad (2)$$

The significance of q_th decoding input, r_th output is captured by this quantity. q_th location demonstration is h_q and state of the decoder is s_r-1 is the till that point. These quantities are needed to define e_qt and f_ATT is a function.

Through all the inputs, at this instant (e_qr) to sum to 1 is the quantity that we need. By taking softmax the e_qr is transformed to a probability distribution.

$$\alpha_{qr} = \frac{\exp(e_{qr})}{\sum_{q=1}^{M} \exp(e_{qr})} \qquad (3)$$

e_qrs are softmax of Alphas. To produce the r_th output Alpha_qr symbolizes possibility of concentrating on q_th input. Here, the f_ATT function is defined:

$$e_{qr} = A^{T}_{arr} \tanh(B_{arr} S_{r-1} + C_{arr} h_q) \qquad (4)$$

Parameters which will be learned are A, B and C for determination of the value of e_qr while training.

Results and Discussion

In this section, the techniques we've implemented are presented. The first approach is encoder-decoder, encoder in our case is CNN and decoder is a RNN, and second approach is CNN-RNN with attention mechanism. Firstly, the entire OpenI dataset was cleaned and prepared in which we cleaned the text by converting all characters into lowercase, performed simple decontractions, removed punctuation from text except full stop because findings contain multiple sentences), eliminated numbers from findings, words that are ≤ 2 in length (e.g, 'is', 'to' etc), But the word 'no' is not be removed since it adds value) and multiple spaces or full stops or 'X' frequently existing in findings.

After preprocessing the entire dataset was split into training, testing and validation, with the observation of multiple data points of an individual. As, if there are 4 images of an individual with corresponding patient report. As we have to deal with pairs of CXRs, various

records related to particular person will be created. So, it's compulsory to divide the dataset on the basis of entities to keep away the problem of data outflow. By means of the distinctive 'Patient Id' feature dataset is split and then model is trained. The training, testing and validation set consists of 2506, 200 and 500 images respectively.

Our model generates a report containing one word at a time providing a combination of two images. Before feeding the text data into our model we applied Tokenizer on it. Model inputs are partial reports with Images as shown in Figure 2. Block diagram. Every image is converted into a static sized vector and then sent to the model as an input. We are doing this by means of **transfer learning**. About 121 number of layers' convolutional neural network CheXNet were trained with ChestX-ray 14 (biggest openly available dataset of chest radiography), which contains frontal X-ray of 100,000 images along with 14 diseases. For each image, to get the bottleneck features is mainly the goal here.

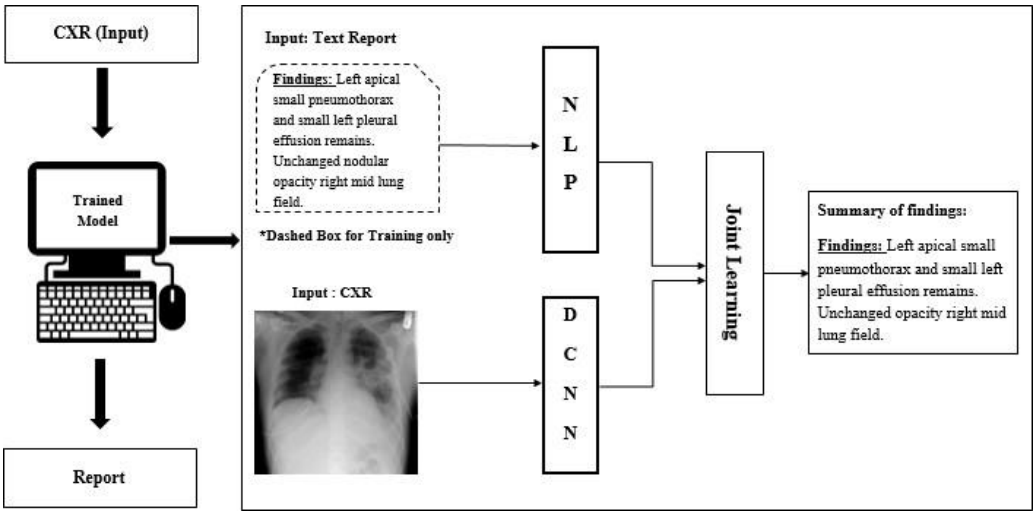


Figure 2. Flow of study

Encoder-Decoder model:

In both approaches we used different layers of CheXNet [6] for output but we have taken ReLU (avg_pooling) layer as output in our first approach CNN-RNN. We rescaled every CXR to (224,224,3) and then delivered it to the CheXNet through which we obtained 1024 size feature vector. For acquiring 2048 features both these features are concatenated. For training and validation, we formed two data generators (at a time produces BATCH SIZE no of data points).

As stated previously our encoder-decoder model is a word-by-word model (sequence-to-sequence). Image features and the partial sequences are taken as an input by the model for generation of following word in series. To train the model the input sequence is split into 2 pairs of input-output. And also, we have used the merge architecture in which at any point RNN is not opened to image vector. Instead, once the encoding of prefix is done through RNN in the entireness then image was presented to language model.

After that we used embedding layer, where each word was mapped on a 300 dimension demonstration via pre-trained GLOVE model [20] and made the parameter 'trainable=False' so while training weights never change.

To mask the zeros in order we created Masked Loss Function so model only acquires the desired words from report and the attention of model will not be given to them. We have used Categorical Cross entropy as our loss function as output words are One-Hot-Encoded. With a learning rate of 0.001, Adam optimizer was utilized also to examine the predictions of model two techniques were used that include beam and greedy search.

Encoder-Decoder with Attention:

After training the CNN-RNN model we did not achieve up to mark accuracy so we extended our model with attention mechanism. For the betterment of encoder-decoder models attention mechanism was proposed. The related chunks of input sequence are focused. We have used output of any of convolution layers (transfer learning) containing spatial facts as opposed to fully connected representation.

This model used same dataset as well split and tokenizer just like the CNN-RNN model as discussed previously. Teacher forcing method was implemented to train the model but we did not converted text into a sequence-to-sequence model this time. CheXNet was used but in our first approach CNN-RNN we have taken ReLU (avg_pooling) layer as output and here the output was bn (ReLU Activation) that was the final convolutional CheXNet layer. After merging the 2 images, the final output shape was (None, 7, 14, 1024). So encoder's input was (None, 98, 1024) after reshaping.

There are 2 parts, an encoder and a decoder of the model but now the decoder consists of an additional attention module, for instance one-step attention decoder. The sub-classing API of keras [21] was used that provided us an additional customizability and control over our architecture. Same loss function, learning rate of 0.001 and optimizer was used. Like previous model same prediction techniques (Beam and greedy search) and BLEU score [22] are utilized for evaluating the predictions. After training the model we achieved up to mark accuracy.

Table 2. Comparison of both approaches

Models	Activation Function	Loss Function	Optimizer	Learning Rate	No of epochs	Batch Size
CNN-RNN	softmax	Categorical Cross entropy	Adam	0.001	40	1
CNN-RNN with Attention mechanism	softmax	Sparse Categorical Cross entropy	Adam	0.001	10	14

Furthermore, In this section, we explored and analyzed both approaches based on CNN-RNN and CNN-RNN with attention. In this study, benchmark dataset OpenI is used for evaluating the efficacy of proposed methodology. Google Colab having 16 GB of RAM with GPU is used for training and the testing purpose. The experiments are performed on two

architectures presented in Table 2. The Figure below shows the training and validation accuracies respectively.

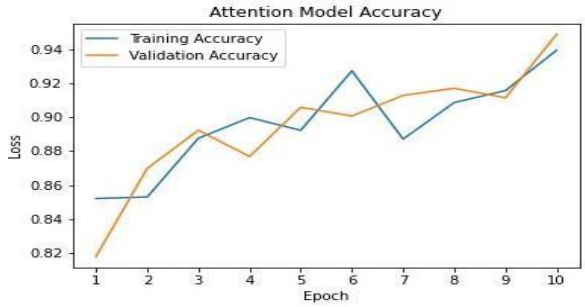


Figure 3. Accuracy Curves of Attention model

While training CNN-RNN with Attention mechanism we have achieved upto mark accuracy which was 96% with validation accuracy of 94% which can be seen in Figure 3. To examine these methods, beam and greedy search methods were used. The most possible word in each time-step is the output by Greedy search. On the other hand, at each time step the most probable sentence by multiplying probability of each word and getting the sentence that has the highest probability is showed by beam search. As compared to beam search greedy search is much rapid despite the fact that beam search generates correct sentences. To evaluate the prediction, we have used the BLEU score as the metric.

Conclusion. We have proposed the model designed for CAD for report generation of pulmonary diseases using CXR. By using the patient-wise official split of the OpenI dataset we have assessed this model in contrast to 2 state-of-the-art models of deep learning. Our model achieved an accuracy of 94%, which is the highest performance.

Acknowledgement. Acknowledgements are considered necessary.

Author's Contribution. Syeda Zainab Yousuf Zaidi supervised this paper. Anmol Zahra, Shagufta Iftikhar and Iqra Naz did the implementation and write up. All authors read and approved the final manuscript.

Conflict of interest. All authors declare that they have no conflicts of interest.

Project details. This paper is based on a project that is deployed on localhost and completed in November 2021.

REFERENCES

[1] H. Wang, H. Jia, L. Lu, and Y. Xia, "Thorax-Net: An Attention Regularized Deep Neural Network for Classification of Thoracic Diseases on Chest Radiography," *IEEE J. Biomed. Heal. informatics*, vol. 24, no. 2, pp. 475–485, Feb. 2020, doi: 10.1109/JBHI.2019.2928369.

[2] Z. Ul Abideen *et al.*, "Uncertainty Assisted Robust Tuberculosis Identification With Bayesian Convolutional Neural Networks," *IEEE access Pract. Innov. open Solut.*, vol. 8, pp. 22812–22825, 2020, doi: 10.1109/ACCESS.2020.2970023.

[3] Z. Huo *et al.*, "Quality assurance and training procedures for computer-aided detection and diagnosis systems in clinical use," *Med. Phys.*, vol. 40, no. 7, 2013, doi: 10.1118/1.4807642.

[4] X. Wang, Y. Peng, L. Lu, Z. Lu, and R. M. Summers, "TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays."

[5] J. Shi, C. Liu, C. T. Ishi, and H. Ishiguro, "3D skeletal movement-enhanced emotion recognition networks," *APSIPA Trans. Signal Inf. Process.*, vol. 10, pp. 1060–1066, Dec. 2021, doi: 10.1017/ATSIP.2021.11.

[6] M. Almuhayar, H. H. S. Lu, and N. Iriawan, "Classification of Abnormality in Chest X-Ray Images by Transfer Learning of CheXNet," *ICICOS 2019 - 3rd Int. Conf. Informatics Comput. Sci. Accel. Informatics Comput. Res. Smarter Soc. Era Ind. 4.0, Proc.*, p. 8982455, Oct. 2019, doi: 10.1109/ICICOS48119.2019.8982455.

[7] "Fundamentals of Diagnostic Radiology | William E. Brant, Clyde A. Helms | download." https://en.1lib.ae/book/2738744/017b45 (accessed Feb. 26, 2022).

[8] E. A. Mendonça, J. Haas, L. Shagina, E. Larson, and C. Friedman, "Extracting information on pneumonia in infants using natural language processing of radiology reports," *J. Biomed. Inform.*, vol. 38, no. 4, pp. 314–321, Aug. 2005, doi: 10.1016/J.JBI.2005.02.003.

[9] M. Horry *et al.*, "Deep mining generation of lung cancer malignancy models from chest x-ray images," *Sensors*, vol. 21, no. 19, Oct. 2021, doi: 10.3390/S21196655.

[10] J. Lovelace and B. Mortazavi, "Learning to Generate Clinically Coherent Chest X-Ray Reports," *Find. Assoc. Comput. Linguist. Find. ACL EMNLP 2020*, pp. 1235–1243, 2020, doi: 10.18653/V1/2020.FINDINGS-EMNLP.110.

[11] D. Demner-Fushman *et al.*, "Preparing a collection of radiology examinations for distribution and retrieval.," *J. Am. Med. Inform. Assoc.*, vol. 23, no. 2, pp. 304–310, Jul. 2015, doi: 10.1093/JAMIA/OCV080.

[12] Y. Jeon and J. Kim, "Active Convolution: Learning the Shape of Convolution for Image Classification," Accessed: Feb. 26, 2022. [Online]. Available: https://github.com/.

[13] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines."

[14] S. H. S. Basha, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, "Impact of fully connected layers on performance of convolutional neural networks for image classification," *Neurocomputing*, vol. 378, pp. 112–119, Feb. 2020, doi: 10.1016/J.NEUCOM.2019.10.008.

[15] N. M. Nasrabadi, "Book Review: Pattern Recognition and Machine Learning," *https://doi.org/10.1117/1.2819119*, vol. 16, no. 4, p. 049901, Oct. 2007, doi: 10.1117/1.2819119.

[16] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS 1 LSTM: A Search Space Odyssey," doi: 10.1109/TNNLS.2016.2582924.

[17] S. Wu, S. Zhao, Q. Zhang, L. Chen, and C. Wu, "Steel surface defect classification based on small sample learning," *Appl. Sci.*, vol. 11, no. 23, 2021, doi: 10.3390/app112311459.

[18] D. Bahdanau, K. H. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, Sep. 2014, Accessed: Feb. 26, 2022. [Online]. Available: https://arxiv.org/abs/1409.0473v7.

[19] M. T. Luong, H. Pham, and C. D. Manning, "Effective Approaches to Attention-based Neural Machine Translation," *Conf. Proc. - EMNLP 2015 Conf. Empir. Methods Nat. Lang. Process.*, pp. 1412–1421, Aug. 2015, doi: 10.18653/v1/d15-1166.

[20] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1532–1543, 2014, doi: 10.3115/V1/D14-1162.

[21] A. Gulli, A. Kapoor, S. Pal, O'Reilly for Higher Education (Firm), and an O. M. C. Safari, "Deep Learning with TensorFlow 2 and Keras - Second Edition," p. 646.

[22] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a Method for Automatic Evaluation of Machine Translation," *Proc. 40th Annu. Meet. Assoc. Comput. Linguist.*, pp. 311–318, 2002, doi: 10.3115/1073083.1073135.