Optical Character Recognition for Handwritten Bangla Documents

By

**MD. Sajid Hasan**

Roll: 1607006

&

**Avi Pal**

Roll: 1607054



**Department of Computer Science and Engineering**

**Khulna University of Engineering & Technology**

**Khulna 9203, Bangladesh**

**April, 2022**

Optical Character Recognition for Handwritten Bangla Documents

By

**MD. Sajid Hasan**

Roll: 1607006

&

**Avi Pal**

Roll: 1607054

A thesis submitted in partial fulfillment of the requirements for the degree of

“Bachelor of Science in Computer Science & Engineering”

**Supervisor:**

**Dr. Sk. Md. Masudul Ahsan**

Professor \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Department of Computer Science and Engineering Signature

Khulna University of Engineering & Technology

Khulna, Bangladesh.

Department of Computer Science and Engineering

Khulna University of Engineering & Technology

Khulna 9203, Bangladesh

April, 2022

****Acknowledgment****

All the praise to the almighty God, whose blessing and mercy succeeded us to complete this thesis work fairly. After that, we humbly acknowledge the valuable suggestions, advice, guidance and sincere co-operation of Dr. Sk. Md. Masudul Ahsan, Professor, Department of Computer Science and Engineering, Khulna University of Engineering & Technology, under whose supervision this work was completed with success. His intellectual advice, encouragement, and direction instill confidence in us, and scientific study necessitates a great deal of effort in learning and applying, as well as a broad viewpoint on problems from various angles. We would like to express our heartfelt gratitude to all of the faculty members, officials, and staff members of the Department of Computer Science and Engineering for their unwavering support in completing this project. We also want to thank the authorities of the Central Computer Center, Khulna University of Engineering & Technology for allowing us to use the High Definition Visual Computing and Multimedia Laboratory for training our dataset. Last but not least, we want to express our gratitude to our friends and family members for their unwavering support.

**Abstract**

As the use of computers has increased in Bangladesh in the twenty-first century, document processing has become important in a wide range of fields, including businesses, schools, hospitals, and many other industries. It takes a long time to read a handwritten paper and type it into a computer. An optical character recognition system can scan a document and extract text from it thus making the job of humans much easier. While there are multiple OCR systems available in the software industry, it is difficult to find a reliable similar system for Bangla. This system aims to detect and recognize handwritten characters from an image. Scanned images are preprocessed using various image processing techniques such as adaptive thresholding, canny edge detection, Hough transform, morphological operation etc. A custom dataset is built with 2100 images each containing a single character and another 1016 images each containing multiple characters. A total of 67 classes are used among which 50 characters, 10 numerals and others were supportive characters. Single Shot Detector is used to train the custom dataset. The dataset is trained using a transfer learning approach. The single character images are used to train the model at first. The final model is built by training this learned model with images containing numerous characters. Another CNN based character detection model is used to verify the result. To develop the CNN model, a sequential and a hybrid model are tested. The sequential model is used for verification because while the hybrid model performs slightly better, the sequential model has fewer computational cost. The system has performed admirably in given test data. Average macro precision, recall and accuracy are 88.26%, 80.43% and 72.46% respectively for character recognition. While average weighted precision, recall and accuracy are 83.99%, 84.18% and 72.98%. A major limitation of this system is that it does not detect compound characters. The system can be further improved with a rich dataset containing compound characters along with the normal characters.

Contents

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | **PAGE** |
|  | Title Page | | | i |
|  | Acknowledgement | | | ii |
|  | Abstract | | | iii |
|  | Contents | | | iv |
|  | List of Tables | | | vi |
|  | List of Figures | | | vii |
| **CHAPTER I** | | | **Introduction** | |  |
|  | | | 1.1 Background | | 1 |
|  | | | 1.2 Motivation | | 1 |
|  | | | 1.3 Problem Statement | | 2 |
|  | | | 1.4 Bangla OCR | | 2 |
|  | | | 1.5 Objectives  1.6 Scope of the Thesis  1.7 Organization of the Thesis | | 3  3  3 |
| **CHAPTER** **II** | | | **Literature Review** | |  |
|  | | | 2.1 Introduction  2.2 Previous Works | | 5  5 |
|  | | | 2.2.1 Image Processing Based Approaches | | 5 |
|  | | | 2.2.2 Deep Learning Based Approaches | | 8 |
|  | | | 2.3 Discussion | | 11 |
| **CHAPTER** **III** | | | **Proposed Methodology** | |  |
|  | | | 3.1 Introduction  3.2 Related Terms  3.2.1 Convolutional Neural Network  3.2.2 MobileNetV2  3.2.3 Single Shot Detector | | 12  12  12  14  16 |
|  | | | 3.3 Working Procedure  3.3.1 Processing Input  3.3.2 Character Segmentation and Recognition  3.3.3 Character Recognition  3.3.4 Post-processing  3.4 Conclusion | | 19  19  24  24  27  28 |
| **CHAPTER IV** | | | **Experimental Results** | |  |
|  | | | 4.1 Introduction | | 29 |
|  | | | 4.2 Experimental Setup | | 29 |
|  | | |  | 4.2.1 Hardware | 29 |
|  | | |  | 4.2.2 Software | 30 |
|  | | | 4.3 Dataset | | 30 |
|  | | | 4.4 Evaluation Metrics | | 32 |
|  | | | 4.5 Results | | 35 |
|  | | |  | 4.5.1 Experimental Results of Word Detection | 35 |
|  | | |  | 4.5.2 Experimental Results of Character Recognition  4.5.3 Experimental Results of Character Detection and Recognition  4.5.4 Performance Evaluation in Word Level | 36  40  45 |
|  | | | 4.6 Displaying Result  4.6.1 Postprocessing  4.6.2 Rearranging  4.6.3 Word Correction using Dictionary  4.7 Conclusion | | 45  45  46  47  47 |
| **CHAPTER V** | | | **Conclusion** | |  |
|  | | | 5.1 Summary | | 48 |
|  | | | 5.2 Limitations | | 48 |
|  | | | 5.3 Future Work | | 49 |
|  | | | 5.4 Conclusion | | 49 |
| References | | |  | | 50 |

List of Tables

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Description** | **Page** |
| 3.1  4.1 | Demonstration of Levenshtein distance between two Bangla words.  Each class, labels and their occurrences for proposed dataset. | 27  31 |
| 4.2 | Sample confusion matrix of a multiclass classification model with 25 images. | 33 |
| 4.3  4.4  4.5 | Performance evaluation of word detection from images.  Comparison between sequential and hybrid model.  Performance evaluation in word level. | 35  40  45 |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Description** | **Page** |
| 1.1 | Sample input and target output of the OCR system. | 2 |
| 1.2 | Flow diagram of overall OCR system. | 2 |
| 3.1 | Operations of a convolutional layer. | 13 |
| 3.2(a)  3.2(b) | Operation of pooling layers.  Visualization of ReLU activation function. | 13  13 |
| 3.3 | The architecture of the MobileNetv2 network. | 14 |
| 3.4 | Depthwise convolution transforms a 12123 image to an 883 image using three kernels. | 15 |
| 3.5 | Pointwise convolution reduces a three-channel image to a single-channel image. | 15 |
| 3.6 | Pointwise convolution using 256 kernels that results in a 256-channel image. | 15 |
| 3.7 | Architecture of a Single Shot MultiBox Detector (SSD). | 16 |
| 3.8 | Default bounding boxes of SSD. | 17 |
| 3.9 | Matching technique of SSD architecture. | 18 |
| 3.10 | Selecting the right bounding box using non-max suppression. | 19 |
| 3.11 | Flowchart of the OCR system for handwritten Bangla documents. | 20 |
| 3.12 | Input and output images of some preprocessing steps. | 21 |
| 3.13 | Sample image showcasing noise reduction. | 23 |
| 3.14 | Sample image showcasing word detection. | 23 |
| 3.15  3.16  3.17  3.18 | Architecture of the sequential CNN model.  Architecture of the hybrid CNN model.  CNN layers visualization for sequential model.  Word correction using dictionary. | 25  25  26  28 |
| 4.1 | CMATERDB 3.1.2 dataset classes. | 30 |
| 4.2 | Sample annotated images of dataset. | 32 |
| 4.3 | Sample images with detected words. | 36 |
| 4.4 | Some sample output from test data for classification model. | 37 |
| 4.5 | Precision of each class for sequential and hybrid model. | 38 |
| 4.6 | Recall of each class for sequential and hybrid model. | 39 |
| 4.7 | F1-score of each class for sequential and hybrid model. | 39 |
| 4.8 | Detected characters from sample image with bounding boxes. | 41 |
| 4.9 | Unsuccessful cases for character detection and recognition. | 42 |
| 4.10 | Precision of each class based on test data. | 43 |
| 4.11 | Recall of each class based on test data. | 43 |
| 4.12 | F1-score of each class based on test data. | 44 |
| 4.13 | Precision, recall and accuracy (%) of character detection and recognition. | 44 |
| 4.14  4.15 | Result of the OCR system after rearranging.  Result after word correction using dictionary. | 46  46 |

CHAPTER 1

Introduction

1.1 Background

Electronic voting (e-voting) has become a popular alternative to conventional paper-based voting system due to its convenience and ability to speed up the election process. Along with the merits there also comes detrimental issues. With the increasing dependency on technology, the security of e-voting systems has become a critical issue. Cyberattacks, hacking, and data breaches have all threatened the integrity of e-voting systems and raised concerns about their reliability. A decentralized system is thought to be a solution to these problems.

Blockchain technology, with its decentralized, secure and transparent nature, can be treated as a solution to enhance the security of e-voting systems. A blockchain-based e-voting system can provide a secure and tamper-proof record of every vote cast, ensuring the integrity and accuracy of election results. Hence the importance of secure and trustworthy voting systems cannot be overstated. Moreover, the decentralized nature of blockchain eliminates the need for a central authority to manage the voting process, reducing the risk of manipulation.

1.2 Motivation

Elections serve as the cornerstone of democratic societies and their integrity must be maintained to ensure that the will of the people is accurately reflected. In recent years, the widespread adoption of electronic voting (e-voting) has raised concerns about the security and reliability of these systems. Hence the importance of secure and trustworthy voting systems cannot be overstated. In light of these concerns, there is a growing interest in the development of secure and transparent e-voting systems that can guarantee the accuracy and integrity of election results. This is where Blockchain will play a vital role. The utilization of blockchain in e-voting systems has the potential to revolutionize the way elections are conducted. This study aims to examine the feasibility of using blockchain technology in e-voting systems and propose a practical and scalable solution for its implementation. This thesis will focus on addressing the security concerns surrounding e-voting systems and to contribute to the development of secure and transparent voting systems that can be relied upon.

1.3 Problem Statement

Though E-voting systems have provided a really good platform for conducting elections but the chances of it getting hacked/manipulated are huge which cannot be neglected. It mainly occurs due to the centralization of authority/database. The attackers can easily attack and change the contents of the database these days. This is where a decentralized system like Blockchain will come in handy to ensure the integrity of the election. A blockchain-based e-voting system can provide a secure and tamper-proof record of every vote cast, ensuring the integrity and accuracy of election results. However, there is a lack of practical and scalable solutions for the implementation of blockchain technology in e-voting systems. The main aim is to guarantee the accuracy and integrity of election results. The study will examine the challenges faced by current blockchain base e-voting systems and analyze the potential of blockchain technology in enhancing the security of e-voting systems.

Top of Form

1.4 Objectives

In this thesis, a practical and highly scalable solution for the implementation of blockchain in e-voting systems is proposed. The main objectives are as follows:

* To review the existing e-voting systems and the challenges they face in terms of security and reliability.
* To develop a blockchain based e-voting system for conducting elections
* To develop a multi-authority based administration system
* To include additional cryptography in the current system to increase the security layer of the blockchain system even more
* To evaluate the proposed solution through simulations and experiments to demonstrate its effectiveness in enhancing the security of e-voting systems.

1.5 Scope of the Thesis

As it was mentioned earlier, the primary objective of this thesis is to develop an efficient and highly scalable blockchain based e-voting system. The proposed system allows the inclusion of multiple authorities in the case of permitting the start of the election. The system is optimized in such a way so that it can handle high peers in less Ethereum gas cost. This study will also provide valuable insights into the future of secure and transparent voting systems. The main vision was to research and develop a voting system which will reflect the will of the people.

1.6 Contribution of the Thesis

To counter the drawbacks of e-voting systems, current blockchain based e-voting systems have already been in discussion and developed. But they also have some limitations. The limitations our proposed technique is able to overcome are as follows:

* Exclusion of single authority based e-voting system which removes the partiality and also gains the trust of the voters.
* Introduction of multi-authority permission based system with El Gamal cryptography
* Reduced gas cost

1.7 Organization of the Thesis

The rest of the thesis is organized as follows:

**Chapter 2** discusses about the earlier works regarding blockchain based e-voting systems. This includes the works done in other contexts with potentially applicable in our scenario. A brief discussion about their limitations is also given in this chapter.

**Chapter 3** discusses about the theoretical knowledge necessary for our development in details. The relevant technologies and tools which have connections with our proposed method is discussed in details.

**Chapter 4** describes our proposed technique in details. This includes how different components of the model works with the technique as a whole. This also describes the implementation process of our system and the prototyping we have developed to simulate the e-voting system.

**Chapter 5** focuses on the experimental studies of our proposed system. It also analyses our system with other existing systems and explains the improvements we have done over the existing designs.

**Chapter 6** draws the conclusion of our proposed system. It also addresses some limitations and future works which can be done for improving the system

CHAPTER 2

Literature Review

2.1 Introduction

The field of electronic voting (E-voting) has been an active area of research for several decades due to the importance it carries in every aspect of a nation. Many decentralized ways have been proposed to eliminate the shortcomings of the e-voting system. But they did not perform as per expectation. On the other hand, the need for secure and transparent voting systems has become increasingly important in recent years as the use of electronic systems has become more widespread. Then comes Blockchain. In this section, several researches related to our work is discussed.

2.2 Previous Related Works

Kshetri et al. [1] conducted a study on the use of blockchain technology for secure and transparent voting systems. Their proposed method found that blockchain technology could be used to create a tamper-proof record of voting data and to ensure the confidentiality and integrity of the voting process. The study concluded that blockchain technology could be a promising solution for enhancing the security of E-voting systems. Though this work was not suitable for high-scale implementation, but its theoretical knowledge laid grounds for future implementation.

Xu et al. [2] conducted a study on the security and privacy of blockchain-based E-voting systems. The authors proposed such a blockchain based system in which it could be used to prevent vote manipulation and to ensure the anonymity and privacy of voters. The study concluded that blockchain technology has the potential to significantly enhance the security and privacy of E-voting systems.

Su et al. [3] A secure and efficient e-voting system based on blockchain" is a research paper that proposes a secure and efficient electronic voting system using blockchain technology. The authors aim to address the security and privacy concerns of traditional e-voting systems through the use of blockchain, which offers a tamper-proof and transparent platform for secure online voting. The proposed system includes several security mechanisms such as encrypted voting records, digital signatures, and access control to ensure the confidentiality and authenticity of the voting process. The authors also present a theoretical analysis of the system's security and efficiency.

Efe et al. [4] proposed a blockchain-based e-voting system that provides security, transparency and accountability in political elections. The proposed system utilizes the characteristics of blockchain such as immutability and decentralization to prevent vote tampering and guarantee the confidentiality of voters. The results of the system are analyzed and compared to traditional e-voting systems, showing its advantages in terms of security and transparency.

Ali et al. [5] created a small scaled public network for e-voting. The system was developed in Ethereum Blockchain network. They deployed smart contracts to run several operations. They made sure of transparency, eligibility, provability in the platform.Though their work was remarkable but they had some limitations. Their system was designed only for a small scaled networks. They used custom private networks for testing out the system. The users in the network were forced to download and store all of the networks in their system to run the whole process.

Gupta et al. [6] worked on the design and implementation of a blockchain-based E-voting system. The authors found that blockchain technology could be used to create a secure and transparent voting system that could be easily audited and verified.

Chiara et al. [7] created an online auction system based on Ethereum smart contracts. In their work they used smart contracts to provide transparency and avoid cheating auctioneers. Though their work has provided a way for implementing smart contracts in the blockchain but their work is used only for a small number of people.

Dhar et al. [8] presents a blockchain-based e-voting system that aims to improve the security and efficiency of the voting process. The authors propose a system that uses blockchain technology to ensure the transparency, integrity, and confidentiality of the voting process. The system also includes features such as digital signature, encryption, and smart contracts to further enhance its security. The authors evaluated the performance of their proposed system and showed that it is capable of providing a secure and efficient e-voting solution.

Fan et al. [9] presents an anonymous blockchain-based e-voting system with privacy protection. The authors focus on the use of blockchain technology to enhance the privacy and security of electronic voting systems. The proposed system is designed to protect the anonymity of voters and to prevent unauthorized access to the voting process. The authors provide a detailed analysis of the system's security and privacy features, and they evaluate the performance of the system using a variety of performance metrics. The results of the study indicate that the system is able to provide high levels of security and privacy while maintaining a high level of efficiency and scalability.

González-Bañuelos et al. [10] proposed a better secure system with enhanced security features. The authors proposed a blockchain-based e-voting system that incorporates enhanced security features to prevent tampering and ensure transparency. The system uses a combination of digital signatures and encryption to ensure the authenticity and confidentiality of the voting process. Due to combinations of digital signatures and encryption being used, the system is more secure than most of the system. The authors carried out a series of experiments to evaluate the performance of their proposed system and showed that it provides a high level of security and transparency compared to traditional e-voting systems. But the major drawback of this method is the complexity and computation cost behind the whole operation.

Pour et al. [11] proposed a blockchain-based secure e-voting system using homomorphic encryption. The authors aimed to address the security concerns of traditional e-voting systems such as tampering and manipulation of voting results, privacy issues, and lack of transparency. In this study, the authors used homomorphic encryption to encrypt the vote data, which allows for the computation to be performed on encrypted data without decrypting it. This approach offers higher security and privacy compared to traditional encryption methods. The authors used the blockchain technology to create a secure and transparent voting system where each vote is recorded as a transaction on the blockchain, and the results are made publicly available. The study demonstrated the feasibility of the proposed system, and the results showed improved security, privacy, and transparency compared to traditional e-voting systems.

2.3 Discussion

In this chapter, out of many methods, some of them are mentioned and discussed. The reviewed systems mostly emphasized on the theoretical perspective of blockchain being implemented for e-voting system. Some of them laid the foundation for future implementations. Then in those implementations, variety of works were visible. Many authors proposed different kinds of implementation techniques in different kinds of platforms. Some were able to produce high security in trade of high computation cost. Some provided the basic property of security needed for their purpose. Most of the systems are not highly scalable. Some used custom networks which were forced to download in the local storage.

But one thing is noticeable in most cases that in all the systems which are implemented requires an authority to start the whole process and execute the whole process. But that users’ account could be hacked or manipulated. This doubt could lead to confusion among the voters. Though the whole system is decentralized, but this sense of doubt could be eliminated if multiple authorities would exist in the system and they would grant permission to that requested person. Then that systems’ security would have enhanced even more as it would require multiple layers of permission for the person to achieve authority.

**CHAPTER 3**

**Related Technologies**

3.1 Blockchain

Blockchain is a distributed P2P (peer to peer) technology that provides a secure and transparent way to store and transfer data. The ledgers used in this system are immutable. So they cannot be tampered or manipulated. Though it was originally created as a core component of cryptocurrency, but now it has been adapted in various scenarios. In a blockchain network the data are stored in a block. Each block is distinguished from other by its unique hash. Each hash is generated using the SHA 256 algorithm. The blocks are linked together to form a chain. Each block consists of data(transactions), nonce, block number, hash of its previous block. When a new block is added to the network, the previous hash is included in the new block which as a result creates an unalterable link between blocks. When an attacker will try to manipulate, say change a blocks’ content, then due to the characteristics of the hashing algorithm, the block will have a new hash. But the next blocks will show error here as the next blocks previous hash value will not match. As a result, the attacker will fail to change the system. Figure 3.1 shows a general structure of the blocks in a blockchain network.

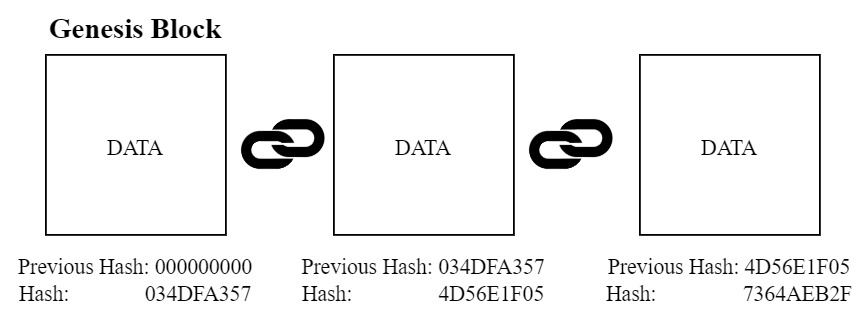


Figure 3.1: Structure of Blockchain

Blockchain mainly consists of 5 pillars. Together they form a structure which turns out to be unalterable, unbreakable. The 5 pillars are as follows:

* **Hash Cryptography:** We human beings have unique fingerprints which are used in various purposes of our life. The fingerprints are different from each other. That’s why we can use them as an identifier. Similarly to distinguish each blocks in the blockchain network SHA-256(Secure Hashing Algorithm) is used. It is a hashing algorithm implemented for creating unique hexadecimal numbers. The hash value will be of 256 bits (64 characters). This algorithm is useful because the numbers it creates are “One-Way” which means once we have created the number from the algorithm, we cannot turn it back to the original number. The numbers are also “Deterministic”, “Fast Computable”. The hashing technique also has the “Avalanche Effect” which means change in digit will result in change in many places in the hashed number. A general overview of a Hash Function is shown in Figure 3.2.

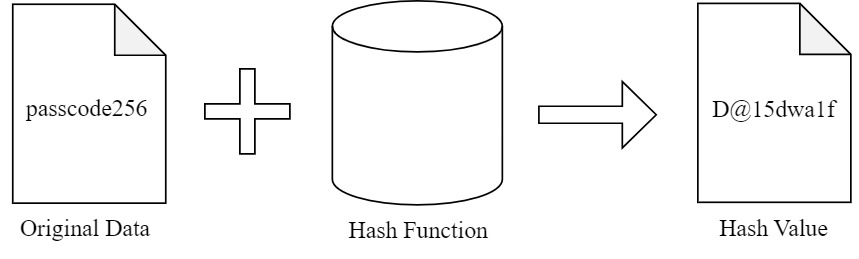
****

Figure 3.2: Mechanism of a Hash Function

* **Distributed P2P (Peer to Peer) Network:** The blockchain network consists of several nodes(devices). The data are stored in all of the designated nodes in the network. If one block is added to one of the nodes then that block is added to the rest of the nodes in the network too. But adding a block to a network is not that easy as to add a block it has to abide by some strong set of rules. Only if they pass all of them then they can be added to the network.
* **Immutable Ledger:** The blocks’ contents are considered as ledgers.The ledgers which will reside in the blockchain network are immutable. The reason is that say a block is added to the network. Now an attacker tries and succeeds to change the block and also the rest of the blocks in the network. Now the thing which will happen is as the blockchain is a distributed peer to peer network, the other nodes(devices) will see the change in the blocks with their copy. When majority of the nodes in the network will see this change, due to the consensus protocol they will change the corrupted copy of blocks with their own copy of blocks.
* **Mining:** In blockchain, the blocks mainly contains data in the form of transactions. The miners work hard to add these blocks(with transactions) into a network. They require high-powered technologies which is used to solve highly complex mathematical problems. Say a target is set. The miners then try to create a block achieving that target. When they will be able to create a block achieving the target, they will be able to add that block into the network. Thus they will be rewarded.These problems are “Hard to solve, Easy to verify”.
* **Consensus Protocol:** In blockchain, there consists many consensus protocols like PoW(Proof of Work), PoS(Proof of Stake) etc. Consensus Protocols are used to maintain the harmony in the blockchain network. Due to it, it protects the network from the attacker. It also provides a solution for the “Competing Chain” problem. Say in one instance, two different blocks are added to two different nodes in the network. The copy of blocks will propagate to its nearby nodes. Now the nodes will have two options to accept. They will now wait until the next block gets added. When the next block will be added by one of nodes. Then the network will accept that nodes’ copy of blocks and reject the other nodes’ block. That block will then be recognized as “Orphan Block”.

Therefore, it can be justified that due to all of the above characteristics, blockchain can be implemented for storing and accessing different kinds of information in a secure way. An overview of Data being encapsulated, transferred and stored is shown in Figure 3.3.

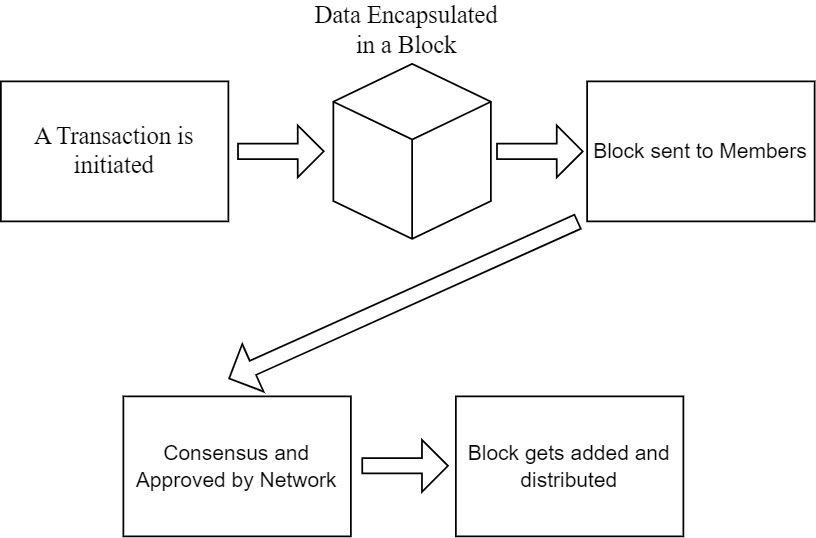


Figure 3.3: Data Flow Diagram of a Blockchain Network

3.2 Ethereum

Ethereum is a decentralized blockchain platform that provides a powerful infrastructure for the creation of decentralized applications (dApps). It was introduced in 2013 by Vitalik Buterin and has since become one of the most popular blockchain platforms, with a wide range of use cases and applications. Ethereum extends the capabilities of blockchain beyond the simple storage and transfer of data, by allowing for the creation of smart contracts. Smart contracts are written on top of this platform which are cryptographic blocks that store value and can only be unlocked if certain conditions are met. So developers are able to create automated, transparent, and tamper-proof applications on the Ethereum network. Ethereum also has its own cryptocurrency, Ether (ETH), which is used to power transactions and pay for the computational resources needed to run dApps on the Ethereum network. This allows developers to create decentralized applications that are not controlled by any central authority, providing greater security, transparency, and decentralization. So like data(transactions), smart contracts are also added to the blocks. Once it gets added to one block, it gets copied to other blocks too.

Ethereum also has a thriving developer community and a strong ecosystem of tools and resources for developers. This includes the Ethereum Development Network (EDN), which provides developers with a test environment for developing and testing dApps, and the Ethereum Virtual Machine (EVM), which is the runtime environment for executing smart contracts on the Ethereum network. EVM can be considered as a computer inside our computer. More on this is discussed in the later section. The potential of Ethereum to support complex applications and its strong developer community make it an attractive choice for the development of secure and transparent e-voting systems.

3.3 Smart Contract

A smart contract is a self-executing contract that automatically executes when specific conditions are met. It is a piece of code that is stored on the blockchain and runs on the blockchain network, providing a secure and transparent way to enforce the terms of a contract without the need for intermediaries. The process of smart contract’s various activities with a blockchain network is shown in Figure 3.4. It is based on a Turing-complete language called Solidity. It is different from bitcoin-script because it contains loops. Smart contracts are basically programs which executes when certain conditions are met. Thus the smart contracts are deterministic. When they are deployed in the network they obtain their own address. With help of smart contract, the data can be stored and transferred.

Smart Contracts are basically programs. So a question might come into mind that programs can be manipulated to harm the device, the program might also contain some coding error which might lead to infinite loop problem. But all of these problems are solved in EVM. As we mentioned earlier, EVM is like a computer inside of a computer. So all of the programs/smart contracts running on the machine is encapsulated. So anything done inside there will not have any effect on the actual computer(device). This allows the devices on the blockchain network to successfully run the smart contracts. Secondly comes the concept of “Ethereum Gas Cost”. When we try to drive a car, we drive it using gas. To purchase gas we need to bear some cost. That concept is used here too. As we said Smart contracts are basically programs, so inside a program there are many operations like logical operations, bitwise operations, mathematical operations etc. All of these operations require some amount of ethers to get executed. This greatly motivates the developers to write code efficiently so that least amount of ether is used. Say if infinite loop problem occurs, then for each loop operation the contract will cut ether. So there will come a time when the ether will run out. That’s how the program will not get hung or stuck. This concept of ethereum gas cost also motivates the developers to avoid writing big blocks of codes.

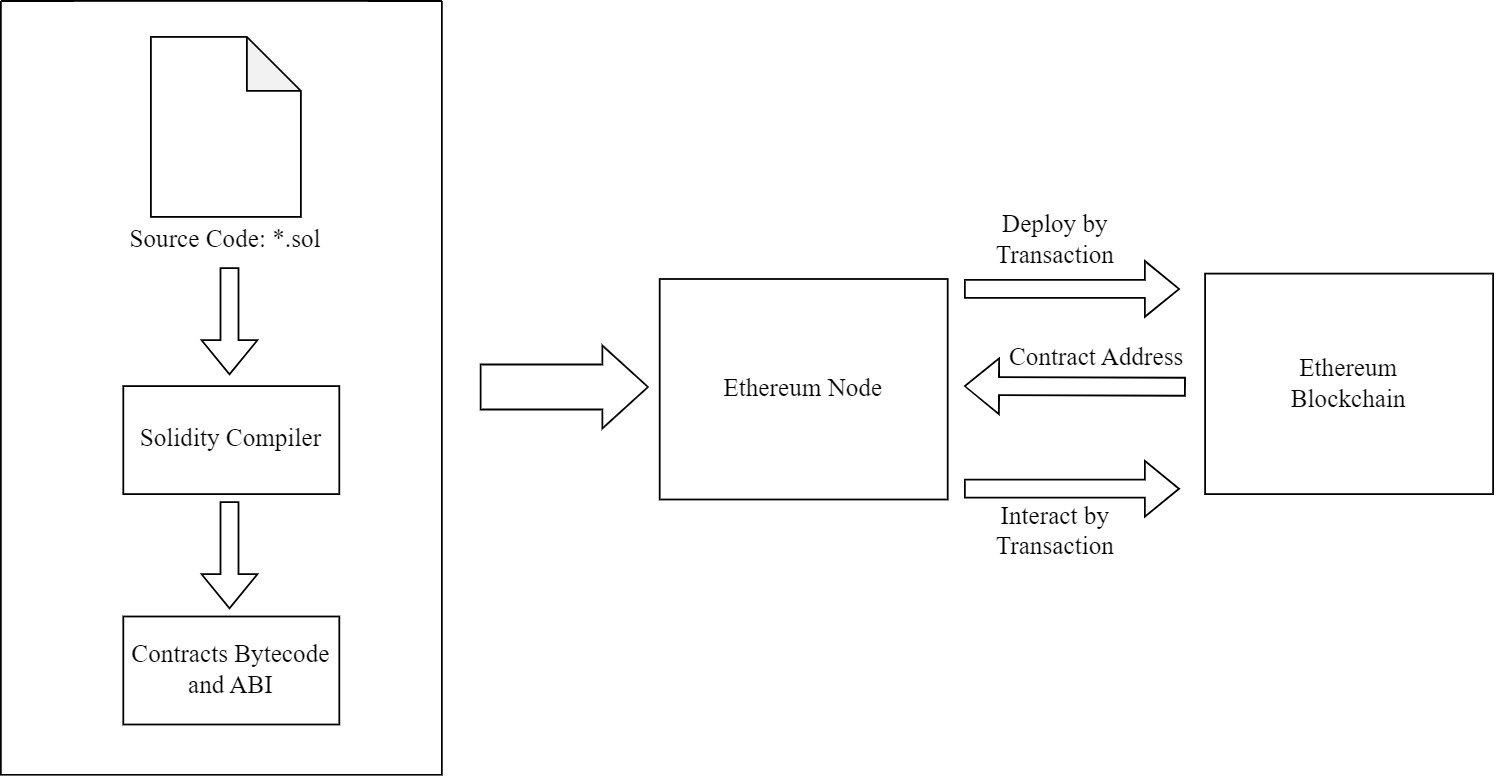


Figure 3.4: The process of smart contract's development, deployment, and interaction with the blockchain

3.4 Solidity

Solidity is a high-level programming language that is used to write smart contracts for the Ethereum blockchain. It was created specifically for the Ethereum platform and is designed to be easy to use and understandable for developers, while still providing a high degree of security and reliability. Solidity is a statically typed, contract-oriented programming language, meaning that developers can create and deploy self-executing contracts on the Ethereum blockchain using Solidity code. The contracts written in Solidity are compiled into machine-readable code, which is then executed by the Ethereum Virtual Machine (EVM). The contracts are structurally similar to the Classes of object-oriented programming. Solidity defines unique variables like “msg”, “block”, “tx” and others that are always present in the global namespace and have properties that allows users to access information about an invocation-transaction and the blockchain. Modifiers are a feature of Solidity that is very useful. Modifiers are enclosed unit of code that augments functions in order to change the flow of execution of the code. The purpose of this method is to eliminate conditional routes in function bodies using the condition-oriented programming (COP) paradigm. Modifiers are used to quickly change the behavior of functions and are specified after the function name in a whitespace-separated list. The modifiers body is the new function body, with ‘\_’ substituted with the original function body. There also consists “Events” which are fired in Solidity. User interfaces and applications can listen for those events without incurring significant costs. Solidity is also designed to be modular and scalable, making it easy for developers to create and deploy complex contracts on the Ethereum network. It includes a number of libraries and tools, such as the Remix development environment, that provide developers with a comprehensive development environment for writing, testing, and deploying smart contracts on the Ethereum network.

3.5 Remix IDE

Remix is a web-based integrated development environment (IDE) for writing, testing, and deploying smart contracts on the Ethereum blockchain. It is a powerful tool for developers that provides a comprehensive development environment for writing, testing, and deploying smart contracts written in Solidity. Remix is designed to be user-friendly, with a simple and intuitive interface that makes it easy for developers to write and test their smart contracts. It includes a number of features that streamline the development process, such as syntax highlighting, code completion, and error checking. One of the key benefits of Remix is its ability to provide a safe and secure environment for testing and deploying smart contracts. Developers can write and test their contracts in a simulated environment, allowing them to identify and fix any errors or vulnerabilities before deploying the contract to the Ethereum network. This helps to ensure the security and reliability of the contract, reducing the risk of bugs or vulnerabilities that could compromise the contract once it is deployed. If the programs run correctly then they can be assured that it will run okay in the main network also. Remix also provides a number of tools and resources that help developers to optimize their contracts for performance and efficiency. For example, it includes a number of performance profiling and debugging tools that help developers to identify and resolve performance issues, ensuring that the contract runs smoothly and efficiently on the Ethereum network. Remix has its own network like Remix VM (London). Apart from that it allows users to run using their Metamask Account. The metamask account will be connected to the main ethereum network or local blockchain network (ganache). Every user will be attached with an account also known as public/wallet address. Through this accounts certain users will be able to deploy smart contracts and run certain operations on the network. For each deploy, operations certain amount of ether will be cut from their account. Thanks to Remix’s really nice user friendly interface, the users will be able to see their account, amount and the functions available for them to execute in the network.

**3.6 Metamask**

MetaMask is a browser extension and mobile wallet that enables users to interact with decentralized applications (dapps) on the Ethereum blockchain. It provides a secure and user-friendly interface for users to store, manage, and send Ether (ETH) and other Ethereum-based tokens, as well as interact with dapps on the Ethereum network. With the help of it the user will be able to connect to the ethereum network with the browser. It allows users to interact with the dapps and smart contracts without the need of running a full ethereum node on their device. This also lets them engage with the network without understanding the technical details of the underlying technology. Metamask uses a hierarchichal deterministic wallet which generates a unique set of public and private keys for each user. It provides the users with a seed phrase that can be used to recover their funds in case they lose it. It supports not only Ethereum mainnet but also various testnets. It allows users to experiment with different blockchain environments. Of all wallets available till date, metamask is one of the most secure wallets out there. Till date it has no records of getting overridden or getting hacked. Hence it is a favored choice for both individual users and developers. Figure 3.5 shows the process of usage and interaction of Metamask with a blockchain network.

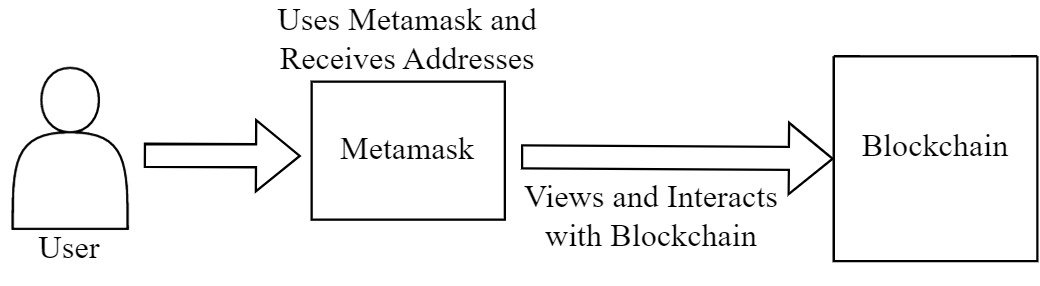


Figure 3.5: The process of using and interacting blockchain via Metamask

**3.7 Ganache**

Ganache is a personal/local blockchain for Ethereum development that is used to test and deploy smart contracts. This is simulator which gives an essence of real blockchain network that runs on a user's machine. It provides a simulated Ethereum environment for developers to test their smart contracts before deploying them to the main Ethereum network. Ganache offers a multitude of features that make it an indispensable tool for Ethereum development. One such feature is the automatic generation of test accounts with assigned balances, which enables developers to perform simulations and test transactions effortlessly. Additionally, Ganache features a web-based graphical user interface that allows developers to monitor the blockchain's state and keep track of the performance of smart contracts. Through Ganache the users will be able to keep track of the whole network. They will be able to see how many blocks are created and how they are linked, see if the block contains any smart contracts, see which address(account) sent transactions to which address. They will also be able to check other details like the transaction hash, block hash etc. It allows the developers to

control the speed and behavior of the blockchain which leads them to simulate different scenarios and test their smart contracts in a controlled environment. This feature helps to identify any issue and fix them before deploying them in the main network. Therefore, it increases the overall reliability and security of the system.

**CHAPTER 4**

**Proposed Methodology**

**4.1 Introduction**

In this chapter our proposed method for the blockchain based e-voting system is discussed in details. Our method consists of different components which are essential to run the system in a correct manner. Our proposed system is divided into three phrases. The phrases are as follows:

* Pre-Election Phrase
* Election Phrase
* Post-Election Phrase

All of the above mentioned phrases happens sequentially one after another. The whole process starts with Pre-Election Phrase. A certain administrator with certain rights initiates the process. All of these Phrases will be discussed in details in the following sections.

**4.2 Pre-Election Phrase**

During this phrase different works like proposing an election, adding candidates, authorizing voters, setting time-limit for election etc. are done. All of these are completed by a certain person who plays the role of the authority in the system. But that certain authority does not get the rights/permission right away. Our proposed system consists of multiple authority who propagates the permissions linearly among them until the final targeted authority gets the right. This phrase is explained in later sub-sections.

**4.2.1 Linear Authorization Chain for Permission Management and Resource Access**

Initially an Ethereum Based Blockchain network is deployed. We used REMIX IDE to write, test and deploy our smart contracts. The process for a linear authorization chain for permission management would involve the following steps:

* **Request for Access:** An user initiates a request for obtaining the permission to be the main authority to start the process of the election.
* **Initial Approval:** The first authority in the authorization chain would receive the request and review it to determine whether it is valid. If the request is deemed valid, the authority would then grant initial approval.
* **Forwarding to Next Authority:** The approved request would then be forwarded to the next authority in the chain for further review.
* **Review and Approval:** Each subsequent authority in the chain would review the request and either approve or reject it based on their level of authority.
* **Final Approval:** The final authority in the chain would have the final say on whether to approve or reject the request.
* **Access Granted:** If the request is approved by all authorities in the chain, access would be granted to the user.
* **Access Denied:** If the request is rejected by any of the authorities in the chain, access would be denied to the user.

The requests which are passed onto the authorities will be in an encrypted way so that the authorities will have to earn the ownership. We used El-Gamal cryptography in our system for encryption and decryption purposes. Figure 4.1 presents the flow of permission in the Linear Authorization Chain.

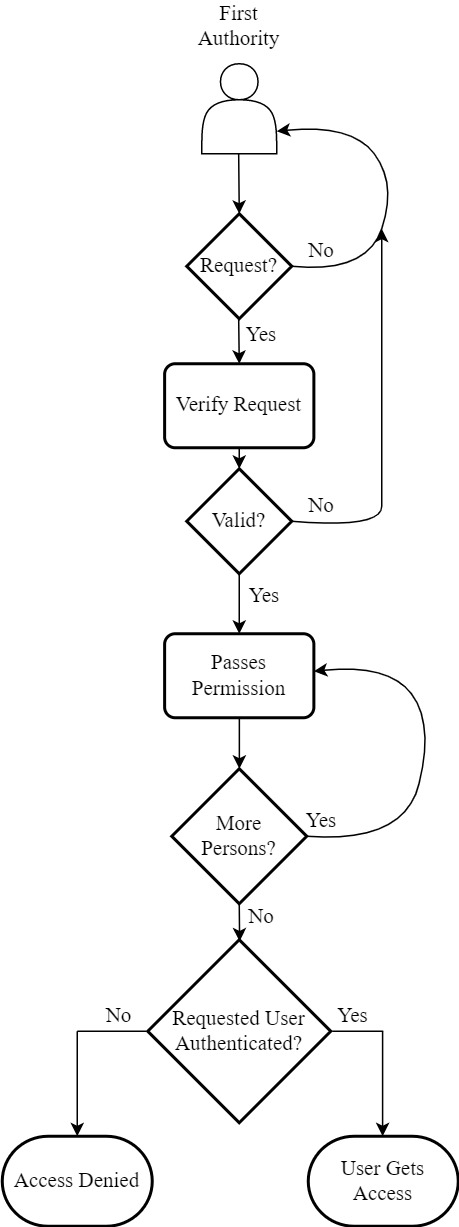


Figure 4.1: Flowchart of Linear Authorization Chain for Permission Management

**Algorithm 1:** Algorithm for initiating and requesting for access

1. procedure **openRequestAccess**(\_req\_owner, \_req\_key)
2. if (open\_req = false, request = false) then
3. req\_owner = \_req\_owner
4. req\_key = \_req\_key
5. open\_req = true
6. END
7. procedure **reqAccess**(\_key)
8. if (request = false, open\_req = true, req\_owner = msg.sender) then
9. request = true
10. open\_req = false
11. END

**4.2.2 Principle of El-Gamal cryptography and its implementation in the system**

**Principle of El-Gamal cryptography**

El-Gamal encryption is a public-key cryptosystem. It uses asymmetric key encryption for communicating between two parties and encrypting the message. It is a discrete logarithm problem

**Public and Private Key Generation:**

El-Gamal system consists of two public keys and one private key. Say the public keys are denoted as “y” and “g” and the private key is denoted as “x”. Figure 4.2 helps to show the process of key generation in a simpler fashion. Steps to generate them are as follows:

* Choose a large Prime Number, p
* Now generate g from p. g is the generator which is also the first encryption key (public key)
* Choose a private key, x where it meets the condition: 1 < x < p
* Generate second encryption key (public key) using,

y = gx mod p

* So the public key will be {g,y} and private key will be x.

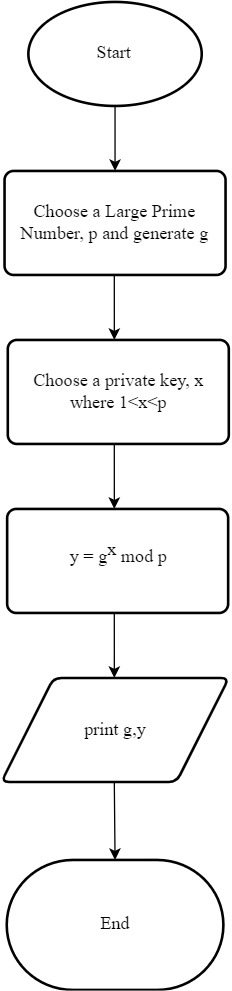


Figure 4.2: Flowchart of Key Generation in El-Gamal

**Encryption:**

In the process of encryption, two cipher-texts will be generated for one single message. The process is shown below:

C1 = gr mod p

C2 = m\*yr mod p

Here,

r = secret random integer of the sender

m = the message of the sender

C1 = first cipher-text

C2 = second cipher-text

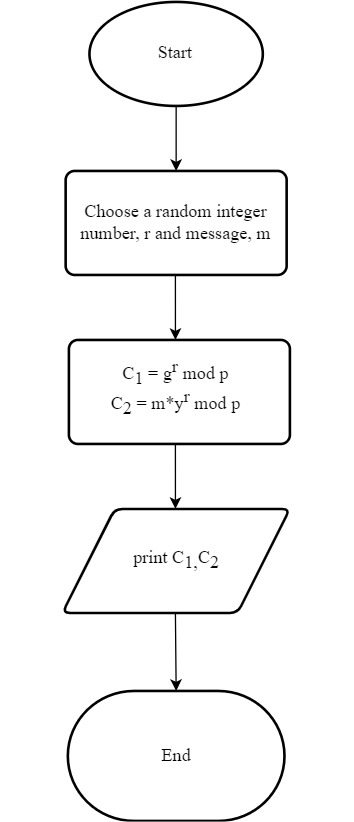


Figure 4.3: Flowchart of Encryption in El-Gamal

**Decryption:**

In this process, the cipher-texts will be used to a message. The process is shown below:

=

This is how the decryption will be done in case of El-Gamal cryptography to get the message

which was intended for the receiver.

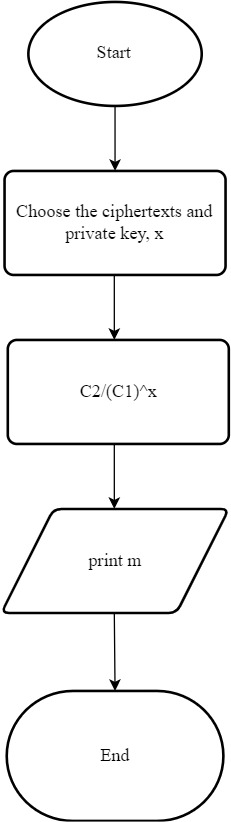


Figure 4.4: Flowchart of Decryption in El-Gamal

**Implementation of El-Gamal in our system**

The first authority will pass the ownership to the next authority and along with that he will also pass down a secret token. The token will be converted into two cipher-texts. Though the cipher-texts will be available publicly but only the second authority in line would be able to decrypt it as the token will be encrypted with the public key of the second authority. Only the second authority will be able to decrypt it using his private key. This is the principle of public key cryptography. So the steps are as follows:

* After receiving the request, the first authority will pass the ownership to the second authority. Along with it the first authority will also pass down a token which will be encrypted into two cipher-texts.
* The second authority will decrypt the cipher-texts and get a token. To gain ownership the second authority will then match the token with the original token of the first authority.
* If it matches, then the second authority will be granted the ownership of the smart contract. Then the second authority will do the same previous processes for transferring the ownership to the third authority.
* The transfer goes on till final authority.
* Ultimately when the user who requested in the first place will earn the ownership then he will start the pre-election phrase of our proposed system.

This encrypted, linear authorization process enhances the security of the permission

management system, as the confidentiality of the request and each step of the approval process would be protected by encryption. This ensures that sensitive information is protected even as it is transmitted between different authorities in the chain.

**Algorithm 2:** Linear Authorization Chain for Permission Management

1. procedure **transfer\_ownership** (\_newOwner, \_h1, \_h2, \_message)
2. if (\_newOwner != address(0), \_newOwner != owner, freq\_of\_owner=0)
3. nextOwner = \_newOwner
4. hidden\_number1 = \_h1
5. hidden\_number2 = \_h2
6. message = \_message
7. Inclusion of that owner in the ownerlist
8. owner\_count++
9. END
10. procedure **accept\_ownership**(\_checkMsg, \_owner\_count)
11. if(msg.sender=nextOwner, \_owner\_count=owner\_count, message=\_checkMsg)
12. owner=nextOwner
13. Inclusion in ownerlist
14. nextowner=address(0)
15. hidden\_number1=0
16. hidden\_number2=0
17. END
18. procedure **transferOwnershiptoAnotherParty**(\_address, \_h1, \_h2, \_msg)
19. if(\_address!=address(0), address!=owner, freq\_of\_owner=1)
20. Inclusion in ownerlist
21. nextowner=\_address
22. hidden\_number1=\_h1
23. hidden\_number2=\_h2
24. msg=\_msg
25. ownercount++
26. END

**4.2.2 Proposal of Election and Adding Candidates**

Following the previous event when the user will finally obtain the permission to start the pre-election phrase, he will be considered as the main authority now. He will then express a proposal for election. He will give an appropriate name for the election. Then comes the process of adding verified candidates. He will be provided a list of valid candidates forehand to eliminate the need of verifying them. The process of adding candidates during the pre-election phase of our proposed system consists of following steps:

* The current authority will set the number of candidates.
* After receiving the list, the current authority will add the candidate with an ID. This ID is used to distinguish a candidate with other candidates.
* The smart contract will keep checking if candidates with same ID occurs or not. Adding candidates with same ID will result in penalty.
* The smart contract will also keep checking if the number of candidates the authority wants to add, exceeds the number of candidates set previously in the first step.

After following the above processes, it can be confirmed that the candidates who participate are well-informed and eligible to compete in the election. Figure 4.5 can be observed to visualize the whole process.

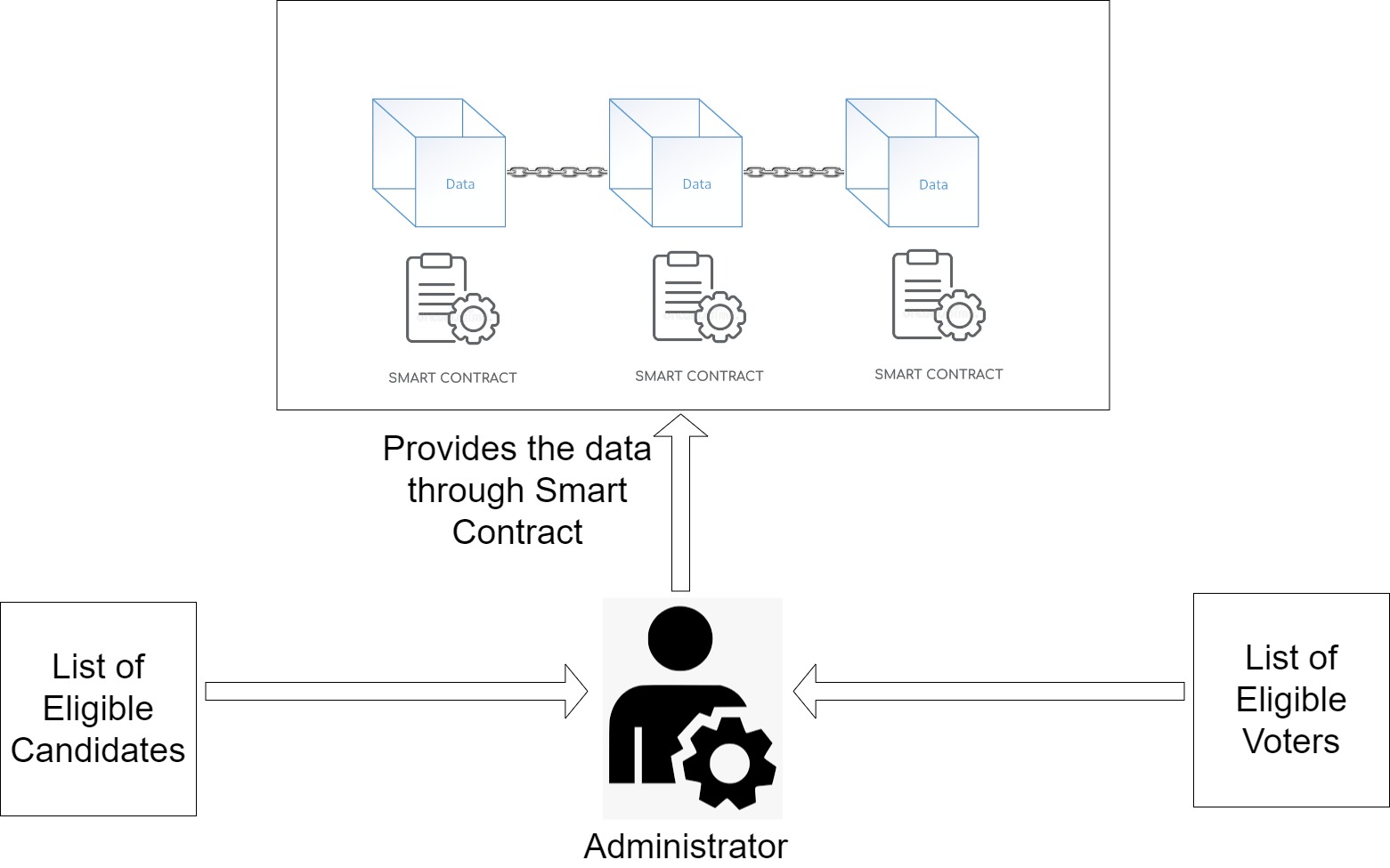


Figure 4.5: Block Diagram of Pre-Election Phrase

**Algorithm 3:** Algorithm for Adding Eligible Candidates

Here next\_candidate is set to zero initially which will act as the index pointer for the array.

1. procedure addCandidates(\_name, \_id)
2. if (next\_candidate >= candidateSize) return
3. if (\_id=candidate.id) return
4. candidate.name=\_name
5. candidate.voteCount=0
6. candidate.id=0
7. if (next\_candidate+1 <= candidateSize) then next\_candidate++

**4.2.3 Addition of Eligible Voters**

After adding eligible candidates, the authority will now move on to adding eligible voters. The authority will be presented with a list of eligible voters with their public address. Using the Remix IDE the authority will authorize them. The authority will insert the addresses into the smart contract’s function for authorizing voters. After the execution of the function the voter will be authorized for voting. But the voter will not be able to vote until the authority starts and changes the state of system.

**4.2.4 Setting Time-limit and Starting the Election**

Following the event of adding candidates and voters, the authority will now have to set a time-limit for the system. Otherwise the system would lose its importance. Setting time is a bit different in Smart Contracts than regular programs. After setting the time-limit the election process will start and the state will change to “Election” state in the smart contract.

**4.3 Election Phrase**

After starting the election, the authority will now observe the Election Phrase. During this phrase the registered eligible voters will be able to interact with the blockchain network with their metamask wallet as they are registered with their public address of the metamask. Figure 4.6 visualizes the whole phrase in a minimalistic form. This phrase consists of the following steps:

* **Use of Metamask Wallet:** The voters in our system will need to have the access of Metamask wallet. Through this wallet the eligible voters will cast the vote against their chosen candidate. To perform different operations in the blockchain network using the smart contracts, a fee is required. Normally a very small amount of ETH(Ethereum) will be cut from the voters metamask wallet. By using the wallet each user will get his own unique address which will not be recognizable by the outsiders. So the voters’ identity will not get revealed.
* **Hash Generation of the Vote:** When the voter is casting the vote, he will be asked to provide a secret key along with the candidate index. Later, the address of the voter, given candidate index and the provided secret key will be passed down into the SHA-256 algorithm to create a “Vote Hash”. This will be used later in the case of vote verification
* **Increment of Vote Counts of Candidates:** When the voters cast their vote against their chosen candidate, the vote count of that candidate is incremented. It will be updated over the whole blockchain network.
* **Verification of Vote:** When the voters will cast their vote, they might want to verify if their vote is correctly cast or not. So they will call the function of the smart contract for it. They will provide their address, their chosen candidate’s index and the secret key they passed during their vote. These inputs will be passed down into the SHA-256 algorithm to create a hash and then will be matched with the “Vote Hash” of the voter which was created before.
* **Elimination of Double Voting:** Double or multiple voting of a voter is also eliminated from the system. The smart contract in our system is developed in such a way that once a voter has cast their vote, they will not be able to cast another vote. Instead if they try to cast a vote, they will receive penalty.

After completion of this phrase, the voters will have cast their vote against their chosen candidate. But they would have to do it in a certain time-limit. Otherwise they will not be able to cast their even if they were eligible.

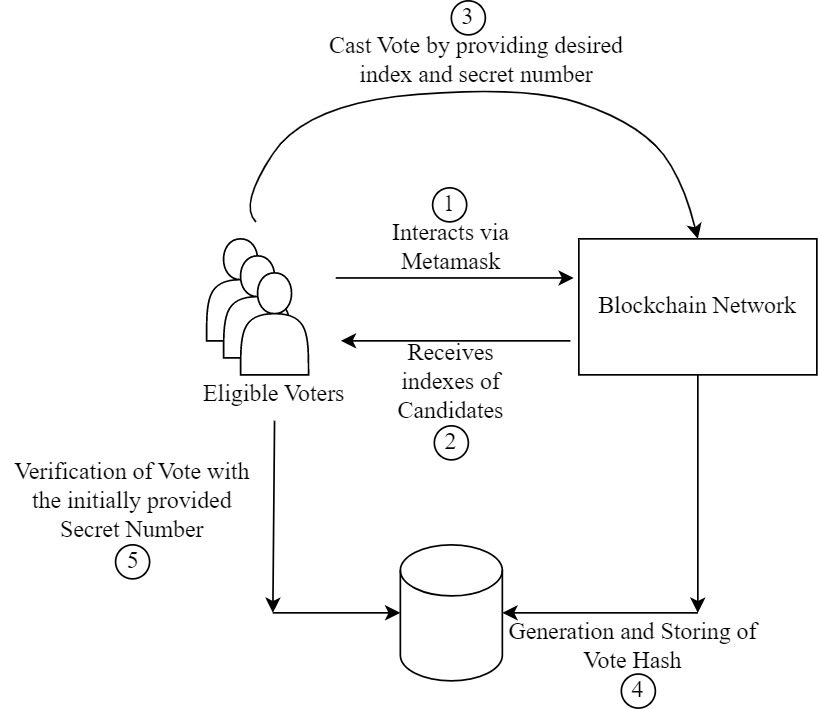


Figure 4.6: Block Diagram of Election Phrase

**Algorithm 4:** Algorithm for Voting and Verifying Vote

1. procedure **vote**(\_candidateIndex, \_snum)
2. if (timeExpired()) then show\_result()
3. else checkif (voter.voted=false, voter.authorized=true) then
4. candidates.votecount += 1
5. totalVotes += 1
6. set voter.voted = true
7. voter.voteHash = sha256(msg.sender, \_candidateIndex, \_snum)
8. END
9. procedure **verify**(\_candidateIndex, \_snum)
10. if (voter.voted=true) then
11. \_voteHash = sha256(msg.sender, \_candidateIndex, \_snum)
12. checkif (voter.voteHash=\_voteHash) return if true or not
13. END

**4.4 Post-Election Phrase**

When this phrase will start, the state of the smart contract will change to “End” state. Now this phrase can be obtained in two ways. One and the normal way is due to expiration of time. When the Election Phrase was being initiated, the authority had to set a certain time. Now time will pass by and when it reaches the end, the smart contract will immediately change its state to “Post-Election Phrase”. Then no further voting will be allowed.

Another way this phrase can be reached is due to authority. If the authority detects something unusual or sees something which might lead to some disaster, then he can terminate the election. But permission of other authorities is required for this operation to execute. This is a very unusual way to end the election. But this is for emergency case only.

Finally, at this phrase, the voters will be able to see the results. They would be able to check out which candidate won, which candidate got how many votes. At this phrase, our proposed system will wrap up.

**4.5 System Components**

**4.5.1 Smart contracts**

Smart contracts play the most crucial role at any Ethereum-based blockchain system. It does the same in our proposed system too. Due to its non-deterministic, verifiable and self-executable pieces of code, the blockchain is kept running and functional. Our proposed model relies heavily on the functionalities of smart contracts as it does most of the work in the blockchain. As stated earlier, from start of the process to end of the process everything is carried out in the smart contract. Our proposed system is able to do all of the works in single contract in an efficient manner.

At first when the contract is deployed, the first authority waits for the request of the user who wants to get the permission to start the election. When request is received, the other works start. Our proposed method which introduced “Linear Authorization Chain”, then gets happened. That too in Encrypted way. We have used El-Gamal cryptography for this purpose. The cryptography is done in a separate python file to lessen the computation cost of the smart contract. As a result, less ethereum gas fee will be required. Later when the user ultimately gets the permission to initiate and conduct the election, then he will do all of the requested works as stated earlier to run an election. In the same contract, when the Election Phrase will start, only the eligible voters will be able to operate also. They will have the opportunity to cast and verify their votes. Later they will be able to see the result.

All of the above works are done in a single smart contract in our system. Hence it is optimized and programmed in such way that minimum amount of gas cost (eth) is required.

**4.5.2 Blockchain Network Nodes**

The system we have proposed on the Ethereum platform uses proof-of-work mechanism to verify and maintain the validity of the blockchain. As stated earlier, this requires a network of nodes(devices) or miners to keep the system running. The miners participate and try to solve the problem which is stated. In exchange of finding the proof-of-work consensus, the miners accept Ethereum cryptocurrencies which is commonly referred to as Ethereum Gas cost in the blockchain.

Blockchain provides highly secure and scalable system. But this comes at a cost. In ethereum based blockchain system, GAS cost is required to run. The computational cost is determined by the gas cost. When a system is in development stage, it is important to keep in mind that the cost should be as low as possible by optimizing the program. While developing our proposed system we had to keep this matter in consideration.

**4.6 Summary**

After thoroughly evaluating our proposed model, we have created a functional e-voting system prototype that is capable of fulfilling all the necessary requirements of an e-voting system. Our system also incorporates a "Linear Authorization" method to enhance security in the permission management process, as all requests and approval steps are protected by encryption. This Ethereum-based prototype operates transparently and has been tested on the Ganache local blockchain runtime environment. The entire system is designed to run within a single smart contract written in Solidity and optimized for minimal Ethereum gas cost. The Metamask plugin is utilized to link the Ganache nodes to the Remix IDE, where different accounts serve various roles in the operation of the system.

**CHAPTER 5**

**Experimental Results**

4.1 Introduction

The results from all the methods that were pursued along the course of this thesis work are represented in this chapter. The chapter is divided into multiple sections and subsections. The rest of the chapter is organized as follows: Section 4.2 presents useful information about the used hardware, and the software, respectively, for future reproducibility. Section 4.3 provides a detailed overview of the used dataset. Section 4.4 represents the various Evaluation metrics used in the segmentation and classification phase. Section 4.5 presents experimental results for word detection, character recognition and character segmentation.

4.2 Experimental Setup

Useful information about the hardware used in the capturing and model training process as well as the software information regarding the toolkit versions, OS information, etc. are discussed in Experimental Setup section.

4.2.1 Hardware

Many hardware are used for the successful completion of the project. Without the correct hardware, the software may not run efficiently or at all. Useful information about the hardware used in the database creation and later processing and detection tasks are provided below.

* + - **Computer**
* Processor: 2.8GHz dual-core Intel Core i7 (Turbo Boost up to 3.3GHz) with 4MB shared L3 cache.
* RAM: 16GB 1600MHz DDR3L.
* ROM: 1TB APPLE SSD SM1024F.
* GPU: Integrated Intel Iris 5100 with 1536 MB VRAM.

4.2.2 Software

Various types of software are used to make the preprocessing, training and testing steps trouble-free. Useful information regarding the used OS information, programming language information, toolkit versions, etc. are provided below.

* Operating System: Windows 10.
* Programming Language: Python 3.7.
* Application type: Python Script.
* IDE: Spyder v3.3.0.
* User Interface: Flask 2.1.1
* Libraries: Tensorflow 2.7.0, Cudnn 8.1, Cuda 11.2, Scikit-learn v0.20.0, OpenCV v3.4.2, Pandas v0.23.4, Numpy v1.15.3, Seaborn v0.9.0, jupyter notebook v5.6.0.

4.7 Conclusion

Following the preprocessing phase, words are detected using morphological transformation. Detected words are segmented and detected using MobileNetV2. At first, a dataset with 2100 images each having a single character is trained. The trained model is then trained again with 1016 images. This time the images contained more than one character. Precision, Recall and Accuracy are found to be 88.26%, 80.43% and 72.46% respectively. To improve the performance of the model, another sequential CNN model is trained that detects isolated handwritten characters with an accuracy of 94.7%. Detected characters from the first model is passed on to the second model to check whether the detection result is right or wrong. Despite the difficult nature of the input data, experimental results suggest that this strategy is promising.

CHAPTER V

Conclusion

5.1 Summary

Optical Character Recognition systems are one of the applications of Computer Vision that has revolutionized education, banking and other business industries. OCR is commonly used technology in banking, healthcare, communication and education sectors. OCR for Bangla language can reduce the hidden cost and time for manual data entry in all government and private offices in Bangladesh. Although, the results of OCR needs to be re-edited as it can’t give fully correct results. In this thesis, an optical character recognition system for handwritten Bangla documents have been developed. A custom dataset has been created for character detection and recognition from words. Words have been segmented from documents by some preprocessing steps, Hough transform and morphological transformation. Hence, the proposed system has segmented and classified characters from words using single shot detector. The classification results are slightly improved using proposed CNN model.

5.2 Limitations

The limitations of developing OCR for handwritten Bangla documents are as follows:

* There are 50 normal characters, 10 number characters and 171 compound characters in Bangla language. It is hard to classify each character correctly because of the vast amount of characters.
* In Bangla language, there is no space between characters in each word. General image processing techniques can’t segment handwritten characters correctly. Though, Deep learning based segmentation technique proposed by this thesis overcomes this limitation.
* For creating a generalized system, the size of the dataset plays a huge role. For developing a complete OCR for handwritten Bangla documents, the dataset size needs to be huge. Although, a custom dataset has been created, dataset that is of large size can’t be made.
* This OCR performed poorly for some real world handwritten Bangla documents. As handwritten texts have much more diversity, small dataset is playing a big role here.
* It failed to preprocess those images which have dark background and light color handwriting.

5.3 Future Work

The following tasks can be performed in the future:

* A large dataset of character segmentation and classification can be built which has diverse handwriting.
* Currently, proposed dataset has 67 labels for characters. Number of labels can also be increased by including compound characters.
* The dataset can be trained with proposed SSD model for detection and classification.
* R-CNN, Faster R-CNN, SPP-net, YOLO and other object detection models can be used to get a comparison analysis of results.

5.4 Conclusion

Optical Character Recognition for handwritten Bangla documents is developed that takes handwritten text image as input and converts editable texts as output. Handwritten document image is segmented into lines and words with Hough Transform and Morphological Transformation. SSD-MobilenetV2 detection model is proposed to detect and recognize characters from word images. The proposed methodology is capable of segmenting and recognizing handwritten characters from images. Hopefully optical character recognition system for handwritten Bangla documents can help move us towards a more digital Bangladesh.

**REFERENCES**

|  |  |
| --- | --- |
| [1] | A. Bishnu and B. Chaudhuri. “Segmentation of Bangla handwritten text into characters by recursive contour following”. In: *Proceedings of the Fifth International*  *Conference on Document Analysis and Recognition*. ICDAR’99 (Cat. No. PR00318).  IEEE. 1999, pp. 402–405. |
| [2] | U. Pal, A. Belaïd, and C. Choisy. “Water reservoir based approach for touching numeral segmentation”. In: *Proceedings of Sixth International Conference on Document Analysis and Recognition*. IEEE. 2001, pp. 892–896. |
| [3] | V. Sharda and A. C. Kishan. “Skew Detection and Correction in Scanned Document Images.” PhD thesis. 2009. |
| [4] | E. A. Sekehravani, E. Babulak, and M. Masoodi. “Implementing canny edge detection  algorithm for noisy image”. In: Bulletin of Electrical Engineering and Informatics 9.4  (2020), pp. 1404–1410. |
| [5] | D. Duan, M. Xie, Q. Mo, Z. Han, and Y. Wan. “An improved Hough transform for line detection”. In: *2010 International Conference on Computer Application and System Modeling (ICCASM 2010).* Vol. 2. IEEE. 2010, pp. V2–354. |
| [6] | Y. Lu and M. Shridhar. “Character segmentation in handwritten words—an overview”. In: *Pattern recognition* 29.1 (1996), pp. 77–96. |
| [7] | S.-W. Lee, D.-J. Lee, and H.-S. Park. “A new methodology for gray-scale character segmentation and recognition”. In: *IEEE transactions on pattern analysis and machine intelligence* 18.10 (1996), pp. 1045–1050. |
| [8] | N. Dave. “Segmentation methods for hand written character recognition”. In: *International journal of signal processing, image processing and pattern recognition* 8.4 (2015), pp. 155–164. |
| [9] | J. Zhou and D. Lopresti. “Extracting text from WWW images”. In: *Proceedings of the Fourth International Conference on Document Analysis and Recognition*. Vol. 1. IEEE. 1997, pp. 248–252. |
| [10] | L. Likforman-Sulem, A. Hanimyan, and C. Faure. “A Hough based algorithm for extracting text lines in handwritten documents”. In: *Proceedings of 3rd international conference on document analysis and recognition.* Vol. 2. IEEE. 1995, pp. 774–777. |
| [11] | H. Wang and J. Kangas. “Character-like region verification for extracting text in scene  images”. In*: Proceedings of Sixth International Conference on Document Analysis and Recognition*. IEEE. 2001, pp. 957–962. |
| [12] | M. Z. Alom, P. Sidike, M. Hasan, T. M. Taha, and V. K. Asari. “Handwritten Bangla  character recognition using the state-of-the-art deep convolutional neural networks”.  In: *Computational intelligence and neuroscience* 2018 (2018). |
| [13] | A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. J. Wu, and A. Y. Ng. “Text detection and character recognition in scene images with unsupervised feature learning”. In: *2011 International Conference on Document Analysis and Recognition.* IEEE. 2011, pp. 440–445. |
| [14] | S. Roy, N. Das, M. Kundu, and M. Nasipuri. “Handwritten isolated Bangla compound  character recognition: A new benchmark using a novel deep learning approach”. In: *Pattern Recognition Letters* 90 (2017), pp. 15–21. |
| [15] | A. Fardous and S. Afroge. “Handwritten isolated Bangla compound character recogni-  tion”. In: *2019 International Conference on Electrical, Computer and Communication*  *Engineering (ECCE).* IEEE. 2019, pp. 1–5. |
| [16] | Ø. D. Trier, A. K. Jain, and T. Taxt. “Feature extraction methods for character recognition-a survey”. In: *Pattern recognition* 29.4 (1996), pp. 641–662. |
| [17] | A. Dutta and S. Chaudhury. “Bengali alpha-numeric character recognition using cur-  vature features”. In: *Pattern Recognition* 26.12 (1993), pp. 1757–1770. |
| [18] | B. Purkaystha, T. Datta, and M. S. Islam. “Bengali handwritten character recognition  using deep convolutional neural network”. In: *2017 20th International conference of*  *computer and information technology (ICCIT).* IEEE. 2017, pp. 1–5. |
| [19] | T. Ghosh, M. M.-H.-Z. Abedin, S. M. Chowdhury, Z. Tasnim, T. Karim, S. S. Reza, S. Saika, and M. A. Yousuf. “Bangla handwritten character recognition using MobileNet V1 architecture”. In: *Bulletin of Electrical Engineering and Informatics* 9.6 (2020), pp. 2547–2554. |
| [20] | T. Islam and R. I. Rasel. “Real-time bangla license plate recognition system using faster r-cnn and ssd: A deep learning application”. In: *2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON).* IEEE. 2019, pp. 108–111. |
| [21] | A. Chaudhuri, K. Mandaviya, P. Badelia, and S. K. Ghosh. “Optical character recognition systems for English language”. In: *Optical Character Recognition Systems for Different Languages with Soft Computing*. Springer, 2017, pp. 85–107. |
| [22] | M. B. Bora, D. Daimary, K. Amitab, and D. Kandar. “Handwritten character recognition from images using CNN-ECOC”. In: *Procedia Computer Science* 167 (2020), pp. 2403–2409. |
| [23] | S. Abdullah, M. M. Hasan, and S. M. S. Islam. “YOLO-based three-stage network for Bangla license plate recognition in Dhaka metropolitan city”. In: *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE. 2018, pp. 1–6. |
| [24] | S. Acharya, A. K. Pant, and P. K. Gyawali. “Deep learning based large scale handwritten Devanagari character recognition”. In: *2015 9th International conference on software, knowledge, information management and applications (SKIMA).* IEEE. 2015, pp. 1–6. |
| [25] | T. C. Wei, U. Sheikh, and A. A.-H. Ab Rahman. “Improved optical character recognition with deep neural network”. In: *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA).* IEEE. 2018, pp. 245–249. |
| [26] | K. O’Shea and R. Nash. “An introduction to convolutional neural networks”. In: *arXiv*  *preprint arXiv:1511.08458* (2015). |
| [27] | Prijono, B., 2018. Student notes: Convolutional neural networks (cnn) introduction. *Indoml. com.[Online]. Available: https://indoml. com/2018/03/07/stud ent-notes-convolutional-neuralnetworks-cnn-introduction*. |
| [28] | D. L. Elliott. *A better activation function for artificial neural networks.* Tech. rep. 1993. |
| [29] | Agarap, A.F., 2018. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*. |
| [30] | Seidaliyeva, U., Akhmetov, D., Ilipbayeva, L. and Matson, E.T., 2020. Real-time and accurate drone detection in a video with a static background. *Sensors*, *20*(14), p.3856. |
| [31] | Wang, C.F., 2018. A basic introduction to separable convolutions. *Towards Data Science*, *13.* |
| [32]  [33] | W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. “Ssd: Single shot multibox detector”. In: *European conference on computer vision*. Springer. 2016, pp. 21–37.  D. Bradley and G. Roth. “Adaptive thresholding using the integral image”. In: *Journal*  *of graphics tools* 12.2 (2007), pp. 13–21. |
| [34] | W. Rong, Z. Li, W. Zhang, and L. Sun. “An improved CANNY edge detection algorithm”. In: *2014 IEEE international conference on mechatronics and automation*.  IEEE. 2014, pp. 577–582. |
| [35]  [36]  [37]  [38]  [39] | P. V. Hough. *Method and means for recognizing complex patterns*. US Patent 3,069,654.1962.  Yujian, L. and Bo, L., 2007. A normalized Levenshtein distance metric. *IEEE transactions on pattern analysis and machine intelligence*, 29(6), pp.1091-1095.  N. Das, K. Acharya, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri. “A benchmark  image database of isolated Bangla handwritten compound characters”. In: *International Journal on Document Analysis and Recognition (IJDAR)* 17.4 (2014), pp. 413–431.  S. Alam, T. Reasat, R. M. Doha, and A. I. Humayun. “Numtadb-assembled Bengali handwritten digits”. In: *arXiv preprint arXiv:1806.02452 (2018).*  https://github.com/pritomsaha/Context-sensitive-Bangla-spell-checker |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |