ICBHR: An Improved Convolutional Neural Network for Bangla Handwriting Recognition

Submitted By

Md. Imran Nazir	ID: 19202103248
Tunazzinur Rahman Kabbo	ID: 19202103268
Zobayer Hasan Nayem	ID: 19202103274
Afsana Akter	ID: 19202103295
Afia Anzum Joati	ID: 19202103409

Submitted in partial satisfaction of the prerequisites for the degree of

Bachelor of Science in

Computer Science and Engineering

Department of Computer Science and Engineering

Bangladesh University of Business and Technology

January, 2024

Declaration

The research project in this thesis, "ICBHR: An Improved Convolutional Neural Network for Bangla Handwriting Recognition," is the products of our own labor, we hereby declare. We also affirm that we are the ones who wrote and assembled the thesis. This thesis has not been submitted anywhere for publication or any other reason other than to fulfill requirements for a degree, award, or diploma. This thesis includes appropriate acknowledgments for the materials it has taken from other sources.

Imran	Kabbo	Nayem	Afsana	Joati
Md. Imran Nazir	Tunazzinur Rahman Kabbo	Zobayer Hasan Nayem	Afsana Akter	Afia Anzum Joati
ID: 19202103248	ID: 19202103268	ID: 19202103274	ID: 19202103295	ID: 19202103409

Certificate

This is to certify that Md. Imran Nazir (ID: 19202103248), Tunazzinur Rahman Kabbo (ID: 19202103268), Zobayer Hasan Nayem (ID: 19202103274), Afsana Akter (ID: 19202103295), and Afia Anzum Joati (ID: 19202103409) were belong to the department of Computer Science and Engineering, have completed their Thesis on ICBHR: An Improved Convolutional Neural Network for Bangla Handwriting Recognition Classification satisfactorily in partial fulfillment for the requirement of Bachelor of Science in Computer Science and Engineering of Bangladesh University of Business and Technology in the year 2024.

Supervisor
Mijanur Rahaman
Assistant Professor
Department of Computer Science and Engineering
Bangladesh University of Business and Technology

Approval

I hereby certify that the studies included in this thesis, which is named, "ICBHR: An Improved Convolutional Neural Network for Bangla Handwriting Recognition", are the result of the initial efforts completed by Md. Imran Nazir, Tunazzinur Rahman Kabbo, Md. Zobayer Hasan Nayem, Afsana Akter, Afia Anzum Joati under my supervision. I additionally declare that, aside from publishing, no portion of this thesis has been submitted to any other institution in fulfillment of the requirements for any degree, honor, or diploma.

Supervisor
Mijanur Rahaman
Assistant Professor
Department of Computer Science and Engineering
Bangladesh University of Business and Technology

Chairman
Md. Saifur Rahman
Assistant Professor and Chairman
Department of Computer Science and Engineering
Bangladesh University of Business and Technology

Acknowledgment

We would like to sincerely thank the Almighty Allah, for providing our family and ourselves with loving care along this journey, which has continued till the completion of this research. We would like to express our sincere gratitude to Mijanur Rahaman, Assistant Professor, Department of CSE, BUBT as well. With his assistance, our research is now complete. We discovered the topic, research area, and issue with his recommendations. He helped us with our research and provided us with a wealth of academic materials and research papers in this field. He is patient and responsible. When we had questions and needed his help, he would always find time to meet and discuss with us no matter how busy he was. We also want to give thanks to Md. Saifur Rahman, Assistant Professor & Chairman, Department of CSE, BUBT. He gave us four months of research instruction and created this chance for us; without him, the research project would not be in its current state. Together with thanking our university for giving us this chance, we would also like to thank our teammates for their support.

Lastly, we would like to thank all of the faculty members at BUBT's CSE department for their support over the past four years in providing us with the guidance and assistance we needed to successfully finish this research project.

Copyright

Abstract

This research focuses on Bangla handwritten character recognition, aiming to develop a system that can accurately identify and interpret characters written in the Bengali script using a deep learning model based on Convolutional Neural Networks (CNNs).The proposed model involves the use of advanced image processing and machine learning techniques consist of a layered architecture with convolutional layers, batch normalization, max pooling, activation function and a dropout layer to enhance learning and prevent overfitting. Employing a comprehensive dataset, including CMATERdb 3.1.2, NumtaDB, IUBMCdb the study emphasizes robust data preprocessing involving image inversion, normalization, and augmentation. The extracted features are then fed into a machine learning model, which is trained to recognize and classify Bangla characters, numerals and modifiers. The model shows impressive performance, achieving high accuracy in recognizing a wide array of Bangla handwritten characters, numerals and modifiers, demonstrating its potential for practical applications in automated handwriting recognition systems. The effectiveness of the proposed methodology is demonstrated through extensive experimentation and evaluation applying deep CNN on the dataset NumtaDB and CMATERdb 3.1.2 and got training accuracy 99.7% and 97.2% with validation accuracy 98.6% and 96% using the same model. We have got 99% accuracy applying the same model on IUBMCdb dataset.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Bangla, the mother language of Bangladesh, serves as the official language in various regions, including West Bengal in India, Tripura, Assam, Jharkhand, and Sierra Leone in West Africa. With approximately 250 million speakers, Bangla ranks as the 7th most popular language globally and is considered the 2nd most beautiful language. Given its widespread use and linguistic significance, Bangla Handwritten Recognition technology assumes a crucial role in overcoming the challenges faced in different sectors within these regions. [1] Limited research exists on handwritten characters of Bangla scripts compared to other languages like Latin [2], Chinese [3], and Japanese [4], which have seen significant success in machine learning and deep learning applications. Handwritten character recognition is an important study subject with numerous applications, such as Optical Character Recognition (OCR) for transforming handwritten materials into digital formats. It is crucial to emphasize that, despite the fact that digitalization has become more common, handwritten text still remains in significant quantities, particularly in older documents. Converting these papers manually using traditional typing methods is time-consuming and labor-intensive. As a result, developing a powerful handwritten character recognition model, particularly for Bengali script, is critical. Existing models for Bengali handwritten recognition, however, have limitations, frequently focusing on specific types of letters or digits. Existing works on Bangla handwritten recognition have primarily focused on digit recognition [5], basic character recognition [6], or compound character recognition [7]. However, dealing with handwritten characters in Bangla is challenging due to the variations in shape and style. The arrangement of Bangla characters is complex, as they have alignment issues and many characters share similarities, excluding compound characters that are used to enhance other simple characters. To fill this void, we propose the project with the goal of developing a highly accurate and versatile model for detecting Bengali characters and numeric values from handwritten inputs.

1.2 Problem Statement

Data preparation is crucial for successful deep learning. While data is readily available, the primary challenge lies in the scarcity of processed data. [8] The problem we hope to solve is to address the accurate recognition and classification of Bengali handwritten characters and numerals, considering the variations in different shapes and styles. The intricate character of Bengali script, with its diverse shapes, styles, and alignments, presents a considerable problem in constructing robust recognition algorithms. Previous research in this field has mostly concentrated on a narrow collection of characters, ignoring the thorough detection of compound characters and numeric values. Furthermore, the availability and quality of Bengali handwriting datasets are limited, limiting the training of accurate recognition algorithms. One of the most difficult issues in Bengali handwritten identification is distinguishing between similar-looking characters from different perspectives. To enable successful categorization, the minor differences in strokes and forms of these characters necessitate complex models capable of collecting fine-grained features.

1.3 Research Objectives

The following are the goals we have set for our research:

- To design a CNN model for accurate Bengali handwritten character recognition, including individual characters and numerals.

- To surpass existing benchmarks, achieving state-of-the-art accuracy in Bengali handwriting recognition.

- To process potential applications in document, language preservation, and related fields requiring precise Bengali character recognition.

- To create a model capable of accurately identifying and classifying diverse handwritten Bengali script inputs.
- To have a significant impact on various industries, enabling efficient Bengali character recognition for improved language-related initiatives.

1.4 Motivations

Our project is motivated by the desire to overcome the limitations of existing Bengali handwritten recognition methods. With over 230 million people speaking and writing Bengali, resolving this issue is critical. With the development of high-performance computing (HPC) systems and the computer vision science, image recognition has become an extremely challenging subject. As a result, a variety of algorithms and architectural styles are already available to handle these issues. Neural networks are used in the creation of numerous models to address those issues. Its design is based on a synthetic neural circuit function found in the human brain. Convolutional neural networks, or CNNs, are one type of machine learning-based detector and recognizer system that has been offered as a model to optimize the performance of such architecture. Previous research has made strides, but there is still opportunity for improvement. We hope to achieve human-like accuracy by employing advanced Deep Learning techniques such as convolutional neural networks and model training methodologies. Our technique varies from past efforts in that we will create a sophisticated architecture and train the model from scratch utilizing various datasets.

1.5 Flow of the Research

The research project is evolving into multiple phases. After analyzing the research subjects, we looked at the individual characters and digits to see which ones might be recognized. We have collected multiple images to create datasets. Appropriate deep learning methods are selected, trained, and evaluated using appropriate metrics to the datasets. Experiments and analysis are conducted to compare model performance and draw conclusions. Future areas of research and improvements are also discussed. The overall steps of the study procedure are depicted in Figure 1.1

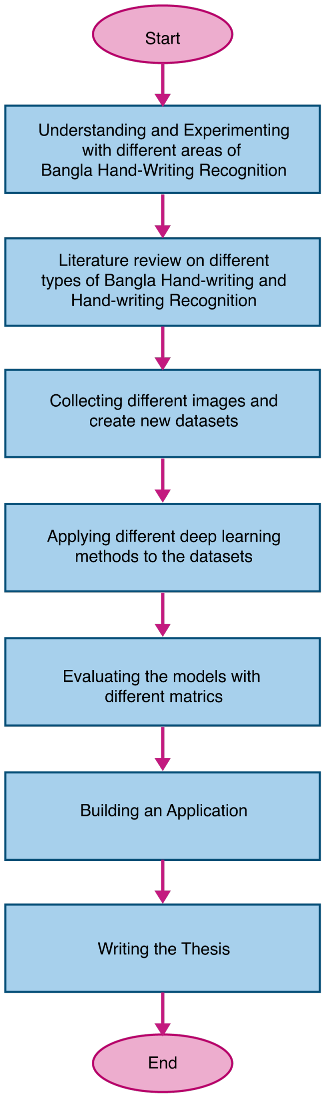

Figure 1.1: Flowchart of the thesis work.

1.6 Significance of the Research

The goal of the study is to advance the creation of accurate Bengali handwritten character recognition algorithms, encompassing individual characters and numerals. This advancement in character recognition technology holds the potential to improve various applications, such as optical character recognition (OCR) systems, document processing, and language preservation initiatives.

It can facilitate the digitization of historical documents, aid in cataloging literary works, and enhance the accessibility of Bengali language content, benefiting both researchers and the general public. The intricate and diverse nature of Bengali script presents significant challenges in constructing robust recognition algorithms.Our research handles variations in shapes, styles, and alignments, which can be valuable for character recognition tasks in other complex writing systems as well.

Our research includes automated data entry, sorting handwritten documents, and enhancing communication tools, thereby streamlining processes and improving productivity.

1.7 Research Contribution

Our research project aims to make significant contributions in the following areas:

- Address the limitations of existing Bengali handwritten recognition models by designing an improved CNN model.
- Train the model from scratch using diverse datasets to achieve high accuracy in recognizing individual characters and numeric values and modifiers in Bengali handwriting.
- Develop a robust solution that can be applied to document processing, language preservation, and other relevant domains.

1.8 Thesis Organization

The Bangla handwriting recognition thesis project is structured as follows: Chapter 2 provides the background and literature review on the field of Bangla Handwriting

Recognition (HCR) system. In Chapter 3, we present the proposed architecture for the Bangla Handwriting Recognition system and provide a thorough explanation of the general steps. In Chapter 4, we discuss the implemenntation to assess the performance of our proposed model. Chapter 5 covers the experimental results and accuracy of our Bangla Handwriting Recognition project. Finally, Chapter 6 presents the whole thesis work conclusion on Bangla Handwriting Recognition.

1.9 Summary

Our project's goal is to create a better CNN model for accurate Bengali handwritten recognition. We intend to outperform current benchmarks by overcoming existing constraints and employing sophisticated Deep Learning techniques. Individual characters and numbers in Bengali script will be classified by the model. The results will have an impact on document processing, language preservation, and automation (example: ID card reading and number plate recognition). Improved precision will aid in digitalization and cultural preservation. This study covers methodology, implementation, evaluation, problems, and future research recommendations. Our research aims to increase Bengali handwritten recognition and make information retrieval and linguistic analysis more efficient.

Chapter 2

Background

2.1 Introduction

An overview of the present state of machine learning-based character recognition is given in this section.. While English handwriting recognition has undergone substantial investigation, Bangla handwriting recognition has garnered far less attention. Nonetheless, some important study has been undertaken in the subject of Bangla character categorization. Recent literature has primarily concentrated on recognizing Bangla characters, numeric values although with comparatively lower accuracy rates. There are some additional successful Bangla Handwritten Character Recognition applications.

2.2 Literature Review

AKM Shahariar Azad Rabby et al. [8] had used Ekushnet dataset (a large number of dataset) for Bangla Handwritten Recognition and augmented the dataset by shifting, rotating & zooming 10% the images for more better result. They tried in many angles & used SVM, MLP classifier but with the 50 epochs the model got a good accuracy of 96.90% on training and in validation got 97.73%. They used multi dataset like CAMTERdb where 95% accuracy is achieved. For researchers there is a scope of achieving better result using bigger CNN architecture.

Nishatul Majid et al. [9] present an offline handwriting recognition system for Bangla script using a segmentation-free approach based on Faster R-CNN. The system employs separate networks, C-Net and D-Net, to detect characters and diacritics in word images, achieving high F1 scores (0.896 for C-Net, 0.932 for D-Net). The methodology includes data augmentation and transfer learning from VGG-16. The word recognition unit achieves a CER of 0.112 and WER of 0.244, and a spell checker further improves accuracy to CER 0.089 and WER 0.215. Limitations include lower accuracy for characters with fewer occurrences, and false detections of certain diacritics. Overall, the approach shows promise for segmentation-free Bangla handwriting recognition.

Rumman Rashid Chowdhury et al. [10] studied Bangla Handwritten Character Recognition using Convolutional Neural Networks (CNN) and Data Augmentation. The main finding indicates that CNN outperforms Linear approaches like SVM and LSTM, achieving an accuracy of 95.25% on the test set. The research covers Bangla handwritten characters and shows the effectiveness of CNN with augmented data. However, the model's focus is on individual character recognition, and there is room for improvement in sequence recognition and dataset expansion.

Yasir Babiker Hamdan et al. [11] focuses on the challenges of recognizing handwriting, especially in languages with limited resources. The researchers propose a new method called the Adversarial Feature Deformation Module (AFDM) to improve the recognition of handwritten words. They found that the AFDM significantly enhances word recognition, especially when there is limited data available. However, the review lacks detailed comparisons with other existing methods and could benefit from exploring a wider range of datasets to ensure its effectiveness across different scenarios.

Md Tanvir Hossain et al. [12] studied to focus on recognizing handwritten Bangla words using a method called Convolutional Neural Network (CNN). They achieved about 84% accuracy for individual characters and 82% accuracy for whole words. The system separated characters into three zones to improve recognition. It works well for neatly written and well-aligned handwriting but struggles with messy or complex characters and vowel modifiers. Overall, it shows promise but has some limitations in handling difficult handwriting.

Elisa H Barney Smith et al. [13] compared scanner and camera-acquired data for Bangla handwriting recognition. Results showed higher accuracy when training with scanned images (C-Net F1: 87.61%, D-Net F1: 93.66%) versus camera images (C-Net F1: 79.93%, D-Net F1: 84.02%). Limitations: diverse camera image quality. Strengths: segmentation-free and segmentation-based approaches. Main finding:

Training with scanned images improves recognition performance for both frameworks. Further research needed for extensive impact analysis.

Sadeka Haque et al. introduced [14] a Capsule Network for Bangla handwritten digit recognition, achieving a validation accuracy of 98.90%. The methodology involves using Capsules, which are effective in handling overlapping data. The dataset includes ISI, CMATERdb, and BanglaLekha Isolated. Limitations include occasional misclassification due to incorrect labeling. The model's strength lies in outperforming some previous works, while weaknesses may arise from dataset inconsistencies.

Tapotosh Ghosh et al. [15] researched on augmented Bangla characters that achieved 96.42% accuracy and 98.92% accuracy on NumtaDB digits using CNN. The main strength is high accuracy, while a limitation is the similarity between some characters. Larger datasets and improved methodologies could further enhance the results.

Hasibul Huda et al. found [16] that InceptionResNetV2 achieved the highest accuracy of 96.99%, followed by DenseNet121 with 96.55% accuracy. CMATERdb dataset was used in this. The study covered the region of Bangladesh and addressed the difficulty of recognizing compound characters in the Bangla language. The limitations were the high computational cost of combined models and the need for further research in processing larger texts. The strengths were the superior performance of InceptionResNetV2 and the application of pre-trained CNN models for character recognition. Weaknesses included the computational complexity and memory requirements of some models.

Bishwajit et al. [17] implemented an image based Bengali handwritten character recognition system using Deep Convolutional Neural Network in 2017. This experiment achieved 91.23% accuracy on 50 alphabets.Banglalekha-isolated dataset was used in this experiment, This has isolated Bangla alphabets represented as binary pictures. There was only 5% of the dataset used for validation, and the image resolution was 28x28.

Md Mahbubar Rahman et al.[18] developed a Bangla handwritten character recognition system using CNN in 2015. In this case, the accuracy of the CNN approach on 50 classes was 85.96%. A bespoke dataset including 20,000 total samples and 400 sample photos per class was created. The experiment used a 28x28 image resolution.

I Khandokar et al. [19] presented a Convolutional Neural Network (CNN)-based Handwritten Character Recognition (HCR) system that can identify Bangla char-

acters in documents and photos. The model achieved an accuracy of 92.91% on a dataset comprising 200 images, with a training set of 1000 images sourced from NIST.
Hakim et al. [20] introduced a "Handwritten Bangla Numeral and Basic Character Recognition Using DCNN" system. They utilized a BanglaLekha-Isolated dataset and a self-prepared dataset, which consisted of a total of 6000 samples (100 per character) for cross-validation. The proposed CNN model was notably deep, containing 8,987,964 learnable parameters. The model achieved a high accuracy of 95.44% on the BanglaLekha-Isolated dataset and 95.25% on their self-created dataset for recognizing 60 characters.

Md Zahangir Alom et al. [21] proposed a Deep Learning technique for Handwritten Bangla Digit Recognition, employing Deep Belief Network (DBN) and Convolutional Neural Network (CNN). The experiments were conducted on the CMATERdb 3.1.1 dataset, comprising 6000 images of unconstrained handwritten isolated Bangla numerals, with each digit having 600 images. The proposed Deep Neural Networks (DNN) achieved an impressive recognition rate of 98.78%, combining DBN and CNN. Moreover, using DBN alone, they obtained an accuracy rate of 97.20% on the same dataset.

Das et al. [22] presented a Bangla Handwritten Character Recognition system aiming to recognize characters in the BanglalLekha-Isolated dataset using Extended CNN, . The model achieved remarkable accuracy rates: 99.50% for Bangla digits, 93.18% for vowels, 90.00% for consonants, and 92.25% for combined classes.

Xu et al. [23] employed a hierarchical Bayesian network that directly takes raw images as network inputs and applies a bottom-up approach for classification. They achieved an average recognition accuracy of 87.5% using a dataset containing 2000 handwritten sample images.

2.3 Problem Analysis

The problem of Bangla handwriting recognition lies in its lack of global recognition and limited research focus on specific aspects of the language. Bangla is a well-known language, although it fails to receive the international appreciation it deserves for handwriting recognition. In addition, rather than focusing on the language as a whole, most researchers favour particular components like compound letters, numeric values, or individual characters.The challenges in recognition include limited diverse datasets, difficulty with complex characters, word-level recognition, handling similar characters, and the need for resource-efficient models with better data augmentation and pre-processing techniques. By overcoming these obstacles, Bangla handwritten character recognition will improve and become more reliable for practical applications.

2.4 Summary

This chapter reviews the current status of Bangla handwriting recognition using machine learning. Despite less focus than English recognition, progress has been made with methods like Convolutional Neural Networks and Capsule Networks. Challenges include limited datasets, complex character recognition, word-level issues, distinguishing similar characters, and the need for efficient models with enhanced data techniques.

Chapter 3

Methodology

3.1 Introduction

A well-defined methodology is crucial to guide the process of system architecture. A methodology provides an overview of the approach and techniques that will be used to carry out a project, research study, and system architecture. The methodology used in our project outlines the approach and technique that will be used to design, develop, and implement our system. This section provides the reader with an understanding of the systematic approach that will be followed to create an efficient and user-friendly system. This section supports CNN's feasibility on study of Bangla handwriting recognition and the specifications given in this framework. Lastly, a thorough explanation of the model's general design is provided in this chapter.

3.2 Feasibility Analysis

It took nine months for this study project—which had five researchers and one supervisor—to develop. Hardware and software support were needed for the thesis study. The researchers are also responsible for the dataset creation and evaluation procedure that is necessary for the study project. The large amount of data for the project is gathered while taking the dataset's legal viability into account. Furthermore, neither the supervisor nor the university provided any financial help for the thesis study.

3.3 Requirement Analysis

To conduct the proposed architecture of the overall requirements include,

- High-performance computing device.
- Graphics Processing Unit (GPU)
- Libraries of open-source software for scientific computing.
- Libraries of open-source software to put the deep learning paradigm into practice.

3.4 Methodology

In this section, We go into further detail about the Bangla handwriting recognition methods. The technique is divided into six sections, each with a detailed explanation and arranged from the model's input to output phase in sequential order.

3.4.1 Datasets

In this study, we utilized a diverse range of datasets to train and evaluate our Bangla handwriting recognition model. The datasets included BanglaLekha-Isolated, containing 1,66,105 handwritten characters, including basic characters, numerals, and compound characters. We also incorporated BanglaWriting, consisting of 21,234 words and 32,787 characters from 260 individuals. Furthermore, we integrated CMATERdb 3.1.1, which provided 6000 handwritten Bangla numerals, and CMATERdb 3.1.3.3, offering 34,439 training and 8,520 testing data for compound characters. Additionally, the Ekush dataset contributed 673,482 character instances, and MatriVasha contained 3,06,464 images of 120 different types of compound characters. WBSUBNdbText comprised 1,352 handwritten text documents, while NumtaDB provided over 85,000 handwritten Bengali digits. Lastly, BN-HTRd included 786 full-page handwritings from 150 writers, consisting of 1,08,18 handwritten words distributed over 14,383 lines and 23,115 unique words. The diversity and size of these datasets allowed us to create a robust and comprehensive model for accurate Bangla handwriting recognition.
Initially we have worked on NumtaDB, CMATERdb and IUBMCd datasets and

conduct the data pre-processing in three stages, data preparation (NumtaDB and CMATERdb), data normalization and data augmentation. These three techniques are described below.

Dataset Name	No of Total Images
CMATERdb 3.1.2	15000 handwritten characters
NumtaDB	85,000+ handwritten Bengali digits
IUBMCdb	7909 Bengali Modifiers

Table 3.1: Dataset Details

3.4.2 Data Pre-processing

Three steps make up the pre-processing of the data: normalization, augmentation, and preparation of the data (using NumtaDB, CMATERdb, and IUBMCdb). The descriptions of these three methods follow.

Data Preperation: Data preparation is a crucial aspect of deep learning, as it greatly influences the performance of the model. While data is abundant, the challenge often lies in the availability of processed data. In our Bangla handwriting recognition project, we addressed this issue by utilizing multiple datasets, such as the NumtaDB, CMATERdb and IUBMCdb. Notably, the images in NumtaDB and CMATERdb and IUBMCdb dataset had a Gray background with black characters. To optimize computation and enhance model efficiency, we performed a preprocessing step where we inverted all the images. This transformation changed the background to black and the characters to white. Consequently, black pixels represented the value 0, leading to reduced computational load. The consistent representation across datasets and the data preprocessing significantly contributed to the effectiveness of our proposed model.

Data Normalization: In our approach to Bangla handwriting recognition, data normalization plays a critical role. It is an essential preprocessing step aimed at standardizing the input data and making it comparable across different samples. The normalization process ensures that the features of each handwritten character image have a consistent range, reducing the impact of varying scales. For this project, we employed the min-max normalization technique on mentioned datasets, which scales the pixel values of the images to a range between 0 and 1 . By subtracting the minimum pixel value from each pixel and then dividing by the difference between the maximum and minimum pixel values, the data becomes normalized. This normalization allows our model to learn from consistent patterns and gradients, facilitating improved convergence during training and enhancing the recognition performance on unseen data.

Data Augmentation: In the context of Bangla handwriting recognition, deep learning techniques perform better when they have access to a larger dataset. Data augmentation plays a crucial role in generating additional data artificially, which is especially beneficial for handwritten characters. As each individual may write a character in various styles and variations, data augmentation allows us to simulate these diverse writing patterns, thus enriching the training dataset. For our research on Bangla handwriting recognition, we employed data augmentation techniques on NumtaDB and CMATERdb that randomly shift the images by 32*32 Pixels except for IUBMCdb dataset. This augmented dataset enables our model to become more robust and capable of recognizing a wider range of handwriting styles, contributing to improved accuracy and better generalization for real-world applications.

3.4.3 Convolutional Neural Network:

In the methodology chapter of Bangla Handwriting Recognition, we employ Convolutional Neural Networks (CNNs) as the core architecture for our model. CNNs are a class of deep learning models widely used for image recognition tasks. They can automatically learn hierarchical features from input data, which makes them particularly well-suited for computer vision applications.

CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers use filters to perform feature extraction by convolving over the input image. The pooling layers reduce the spatial dimensions of the features, reducing computational complexity while preserving important information. Lastly, predictions based on the learnt features are handled by the fully linked layers.

3.4.4 Proposed Model

TThis suggested work style operates as a sequence of steps, one after the other, and is similar to a pipeline of progressive and sequential steps. Each step uses the information from the previous one and turns it into something needed for the next step.

Handwriting recognition, specially for languages with complex scripts like Bangla, is a fascinating challenge in the world of artificial intelligence. Our project uses Convolutional Neural Network (CNN) to recognize handwritten Bangla digits and characters. Think of it like teaching a computer to read Bangla just like humans do, but using images of handwritten numbers and letters. We tested our program on two types of image collections: one with Bangla digits (NUMTADB) and one with Bangla characters (CMaterDB 3.1.2) and another with Bangla modifiers (IUBMCdb). Our CNN model is like a multi-layered CNN, each layer doing a specific job to help the computer understand the handwriting. Figure 3.1 shows the model architecture that we used on three datasets.

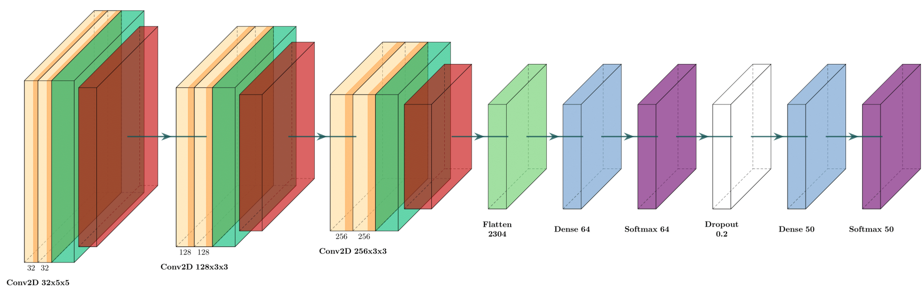

Figure 3.1: Model Architecture

The first layer is where we feed the images into the model. Next, we have layers with tiny filters (imagine them as mini magnifying glasses) that scan the image to pick up patterns like lines and curves in the handwriting. We use 'ReLU activation' in these layers, which enables the model to overlook other patterns and concentrate on the significant ones. And we kept the padding same so the we don't lose the image size with important features. After each scanning layer, we have a pooling layer that shrinks the image down a bit, making it easier for the model to process without losing the important details and also reduces computational power. This process repeats a few times, each time with more filters that can catch more complex parts of the handwriting. Once the model has all the information from the images, we flatten it out to make it ready for the final steps. Here, we have layers that are like decision-makers, analyzing all the information and figuring out what digit or character it is. We also have a dropout layer, which is like a quality check, making sure the model doesn't rely too much on just one part of the information. Finally, the last layer uses something called 'softmax activation' to confidently say, "Which Bangla digit it is or character the image shows."

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 28, 28, 32)	832
conv2d_1 (Conv2D)	(None, 28, 28, 32)	25632
batch_normalization (BatchNormalization)	(None, 28, 28, 32)	128
max_pooling2d (MaxPooling2D)	(None, 14, 14, 32)	0
conv2d_2 (Conv2D)	(None, 14, 14, 128)	36992
conv2d_3 (Conv2D)	(None, 14, 14, 128)	147584
batch_normalization_1 (BatchNormalization)	(None, 14, 14, 128)	512
max_pooling2d_1 (MaxPooling2D)	(None, 7, 7, 128)	0
conv2d_4 (Conv2D)	(None, 7, 7, 256)	295168
conv2d_5 (Conv2D)	(None, 7, 7, 256)	590080
batch_normalization_2 (BatchNormalization)	(None, 7, 7, 256)	1024
max_pooling2d_2 (MaxPooling2D)	(None, 3, 3, 256)	0
flatten (Flatten)	(None, 2304)	0
dense (Dense)	(None, 64)	147520
activation (Activation)	(None, 64)	0
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 50)	3250
activation_1 (Activation)	(None, 50)	0

Total params: 1248722 (4.76 MB)
Trainable params: 1247890 (4.76 MB)
Non-trainable params: 832 (3.25 KB)

Figure 3.2: Model Summary

Figure 3.2 shows the summary of our CNN model which is a carefully structured program, designed to recognize Bangla handwriting by breaking down the images into understandable patterns and making smart decisions based on those patterns. It's like teaching a computer to read and understand handwritten Bangla step by step.

3.4.5 Activation Function:

We took a number of activation functions that are frequently employed in neural networks for various applications in our Bangla Handwriting Recognition model. By introducing non-linearities, these activation functions aid in the model's ability to decipher intricate patterns from the data. Among the activation functions we looked at are:

ReLU (Rectified Linear Unit): The activation function in a neural network is in charge of converting the node's total weighted input into the node's activation or output for that input. ReLU, also known as the rectified linear activation function, is a piecewise linear function that, in the event that the input is positive, will output the input directly; if not, it will output zero. Because a model that utilizes it is quicker to train and frequently performs better, it has become the default activation function for many different kinds of neural networks.

Softmax: The raw output scores from the preceding layers are transformed into a probability distribution across all classes by the Softmax activation function. It ensures that the predicted probabilities for each class sum up to one, allowing us to interpret the model's predictions as class probabilities. This makes it easier to identify the most likely class label for a given input, facilitating accurate and confident predictions in the multi-class scenario of our Bangla Handwriting Recognition task.

3.4.6 Required Web Framework

A website that offers realtime prediction for recognizing bangla handwritten character from a model needs some recommended framework those are given below:

HTML: The standard markup language used to create and organize material on the World Wide Web is called HTML, or Hypertext Markup Language.. HTML serves as the backbone of web pages, defining the structure and elements that browsers use to render and display content. It consists of a series of elements, each represented by tags, which surround content to define its structure or presentation on a webpage.

CSS: Cascading Style Sheets, or CSS, is a language for style sheets that describes how an HTML or XML document is presented. It offers a strong method of separating an HTML-managed webpage's structure from its presentation. It uses selectors and declarations to define styles for HTML elements, influencing properties such as font size, color, and layout. The box model is a fundamental concept, governing the sizing and spacing of elements. CSS enables responsive design through media queries, adapting layouts to different devices. Recent features like Flexbox and Grid offer advanced options for flexible page layouts. In essence, CSS empowers developers to create visually appealing, user-friendly websites by styling HTML content effectively.

PHP: PHP, or Hypertext Preprocessor, is a server-side scripting language widely used for web development. Developed by Rasmus Lerdorf in 1994, PHP seamlessly integrates with HTML, allowing developers to embed dynamic server-side code within web pages. It supports variables, data types, and traditional control structures for programming logic. PHP is particularly adept at interacting with databases, commonly used with MySQL. Its server-side file handling capabilities enable tasks like reading and writing files. Because of its flexibility and integration, PHP is a popular choice for creating dynamic and captivating websites, powering a large percentage of the internet.

Python: Python is renowned for its readability and high-level and adaptable, , ease of use, and robust community support. Python is now widely used in many fields, such as automation, data research, web development, and artificial intelligence. Its syntax emphasizes code readability and clarity, making it an excellent language for beginners while remaining powerful enough for experienced developers. Python supports object-oriented, imperative, and functional programming paradigms, offering a dynamic type system and automatic memory management. Its large standard library offers packages and modules for a variety of uses. we have used some libraries in our project. Such as, PIL and keras.preprocessing to import imagaes, numpy to perform a wide variety of mathematical operations on arrays and Flask for developing web applications using python. Python's success is attributed to its ease of use, versatility, and a vibrant community that contributes to its continuous evolution and adoption across diverse applications.

Flask: Flask is a lightweight web framework for Python that is designed to be simple, easy to use, and extensible. It is classified as a micro framework because it does not require particular tools or libraries and does not impose a specific way of structuring the application. Flask is designed to be simple and easy to understand. It provides the essential tools for web development without imposing a lot of overhead. It comes with a built-in development serverwhich makes it easy to test and debug applications during development. Flask is widely used for developing web applications and APIs.

.h5 Model: An .h5 file, often associated with machine learning frameworks like TensorFlow and Keras, to store models and their weights. The ".h5" extension indicates that the file is in the Hierarchical Data Format version 5 (HDF5) format, which is a versatile and efficient file format for storing and managing large amounts of data. It contains the architecture of the machine learning model, which defines the structure of the neural network or other machine learning model including information about the layers, their types, and the connections between them. It also includes information about the training configuration, such as the optimizer used, the learning rate, and other hyperparameters.

3.5 System Architecture

System architecture is the blueprint that defines the structure, components, modules, and interrelationships of a. It serves as a guide for designing and building scalable and efficient systems.Initially we selected of datasets comprising diverse handwritten characters, numerals, and the modifiers. Subsequently, we applied pre-processing techniques, including data normalization and augmentation, to enhance the dataset. The processed data was fed into a Convolutional Neural Network (CNN) model, which was trained using deep learning methods. Post-training, the model predicted data and underwent evaluation process to refine accuracy.
The final step involved showcasing the model's predictions on a website interface, providing a user-friendly platform for real-time interactions with the recognition system. Figure 3.3 shows our system architecture:

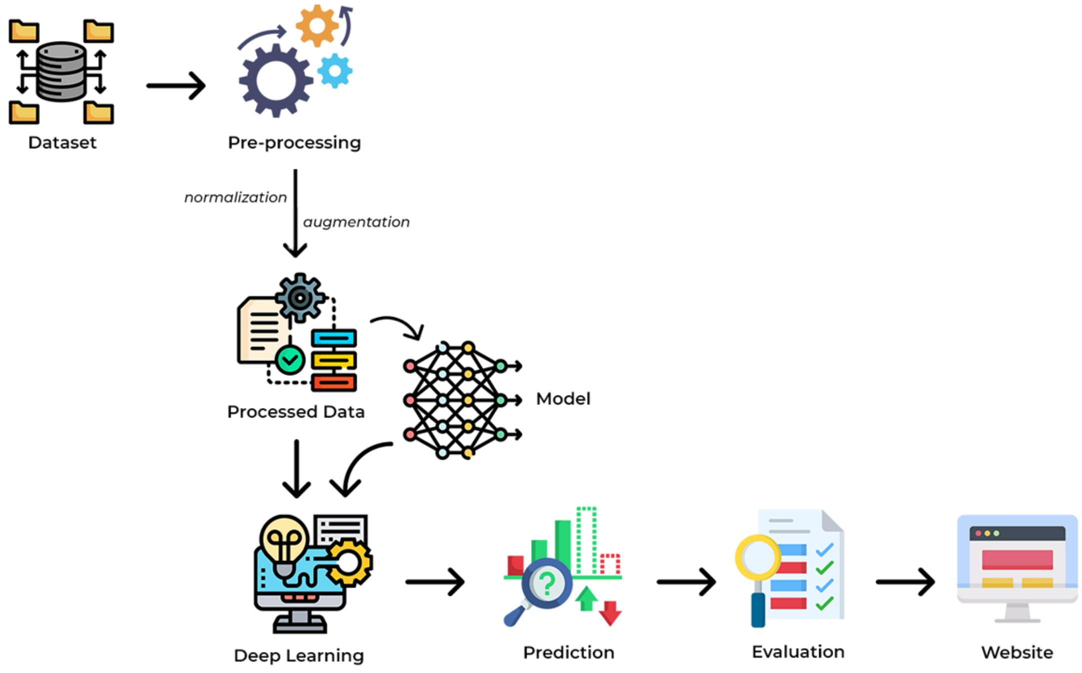

Figure 3.3: System Architecture

3.6 Summary

In Chapter 3, we introduced our proposed model for Bangla Handwriting Recognition through CNN. We conducted a feasibility analysis, identified the key requirements, and explained the research methodology. The model incorporates data pre-processing, a Convolutional Neural Network (CNN) with the Softmax activation function to improve accuracy. This combination ensures effective analysis, making our model a promising solution for identifying bangla handwritten charecters and numerals including Modifiers.

Chapter 4

Implementation

Unlock the power of Bangla handwriting recognition in real-time with our cutting-edge model. Our project is designed to seamlessly predict and analyze Bangla characters from handwritten input, providing instant and accurate results.

Initially we have used only NumtaDB dataset in our website to predict realtime data from the model. Given below Figure 4.1 shows the interface of our web application.

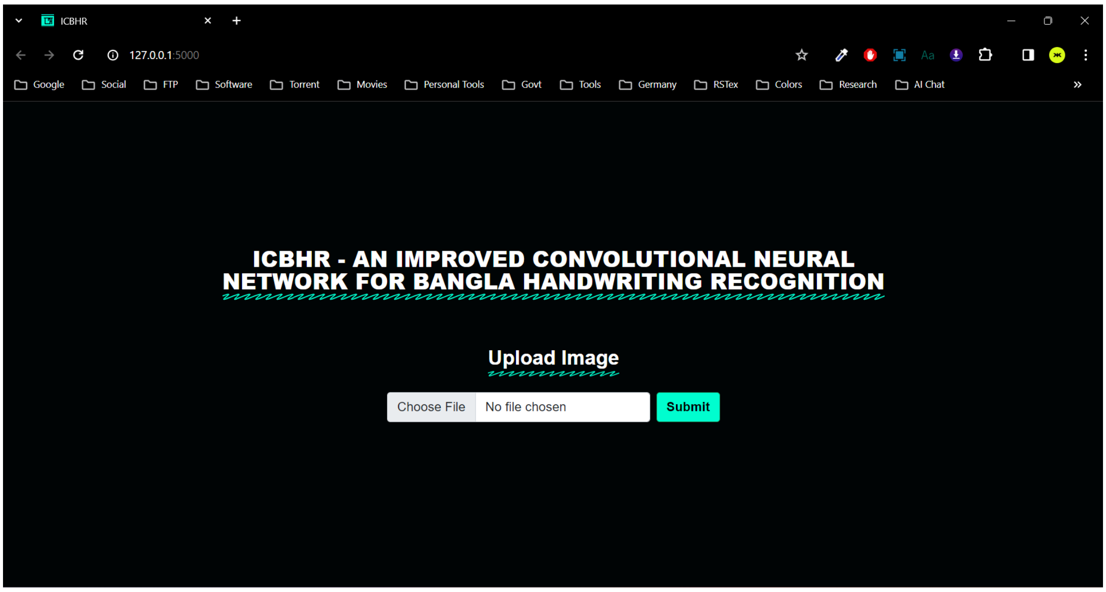

Figure 4.1: Web Server

This website will choose file from the mentioned dataset and predict the number from the uploaded image. The prediction of input image is shown in Figure 4.2.

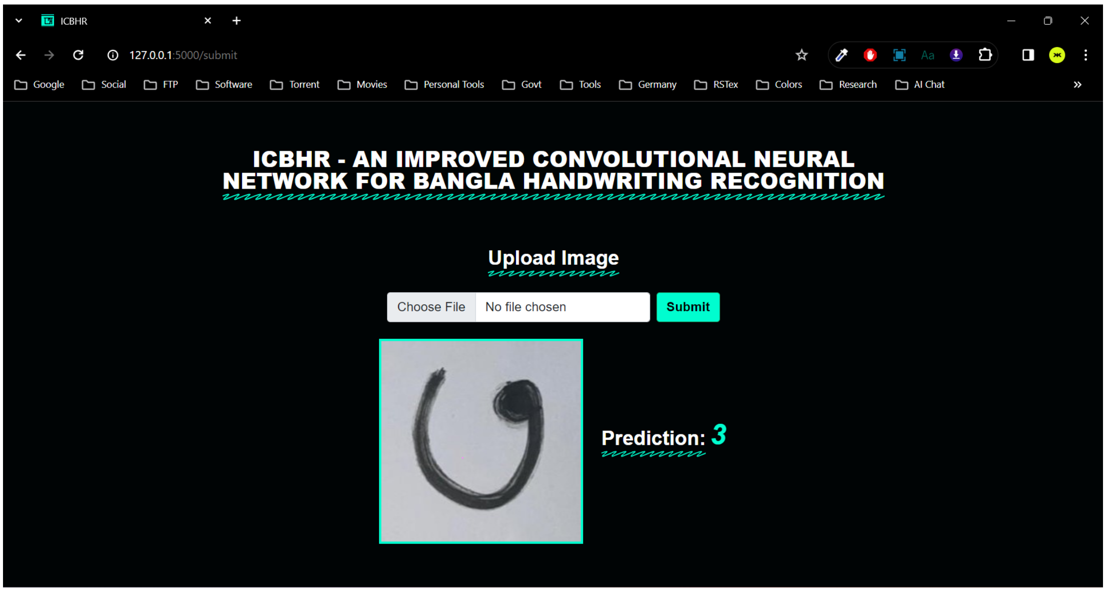

Figure 4.2: Number Prediction

4.1 Summary

In Chapter 4, we implemented a Bangla Handwriting Recognition model for the NumtaDB dataset on a website using Flask and a pre-trained .h5 model. The integration of Flask ensures a user-friendly web interface, allowing real-time character analysis. The .h5 model, trained on NumtaDB, excels in recognizing diverse Bangla characters. Together, these elements create an efficient and effective system for real-time Bangla handwritten numeral recognition on the web.

Chapter 5

Result and Discussion

5.1 Experimental Result

The convolutional neural network performed admirably on several datasets. Good accuracy has been demonstrated on multiple datasets. Below are the results:

Dataset	Accuracy	Validation Accuracy
CMATERdb 3.1.2	97.2%	96.0%
NumtaDB	99.7%	98.6%
IUBMCdb	99.0%	

Table 5.1: Accuracy of our models

5.2 Discussion

In our dataset there was 15000 handwritten characters in CMATERdb 3.1.2 and 85000+ handwritten characters in NumtaDB and 7909 bangla modifiers in IUBMCdb dataset.These handwritten pictures were fed into the input layer of the convolutional neural network after being preprocessed. We add weight and bias after each layer. In addition, the gradient descent method repeatedly reshapes and optimizes the values to obtain the highest level of accuracy. NumtaDB provided the optimum accuracy of 99.7% and validation of 98.6% that is based on handwritten bengali digits shown in figure 5.1 and other data sets shown commendable results. CMATERdb 3.1.2 has given the accuracy of 97.2% and 96% validation which is based on bangla handwritten charecters shown in figure 5.2. IUBMCdb provided the optimum accuracy of 99.0% that is based on handwritten bengali kar sign shown in figure 5.3 Below are the visual results for each dataset:

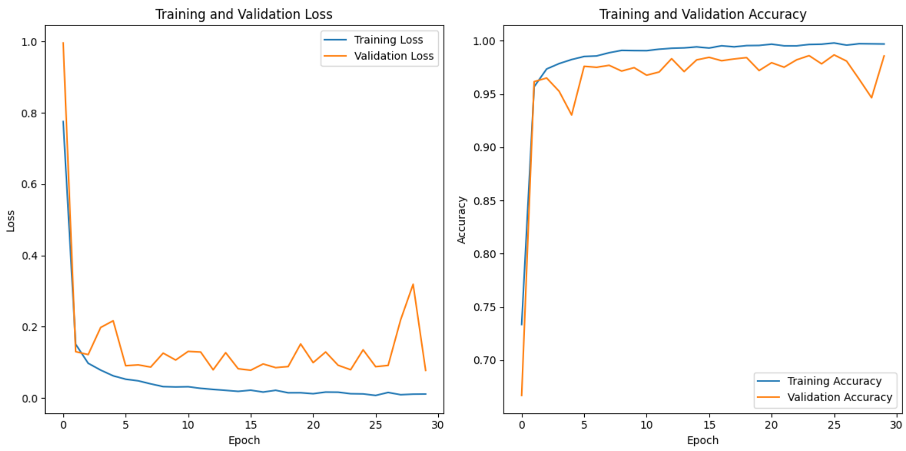

Figure 5.1: Accuracy of Numtadb

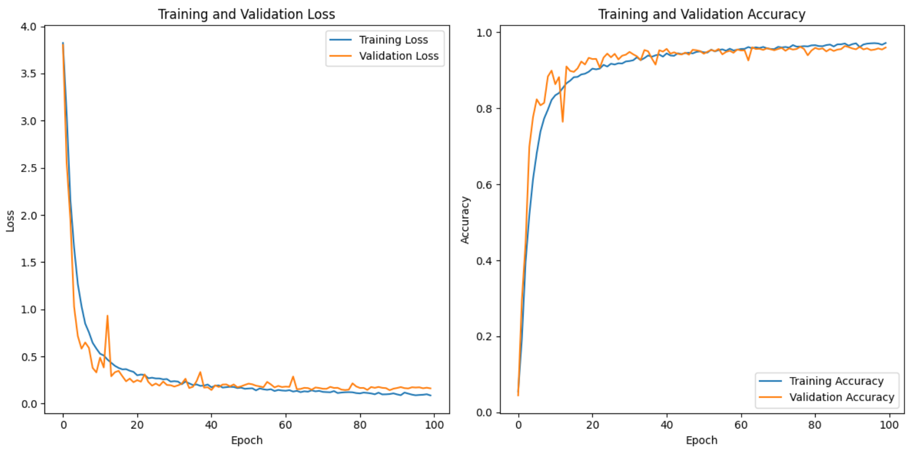

Figure 5.2: Accuracy of CMATERdb

The CMATERdb 3.1.2 dataset for Bangla handwritten recognition is a focused subset specifically designed for the task of recognizing handwritten characters in the Bangla script. It provides a diverse set of handwritten samples, which are essential for training and testing algorithms to accurately identify and interpret Bangla script. We achieve the accuracy 97.2% and validation accuracy 96% using CMATERdb dataset. The NumtaDB dataset for Bangla handwritten recognition is a focused on the task of recognizing handwritten numerals in the Bangla script. We achieve the accuracy 99.7% and validation accuracy 98.6% using NumtaDB dataset.

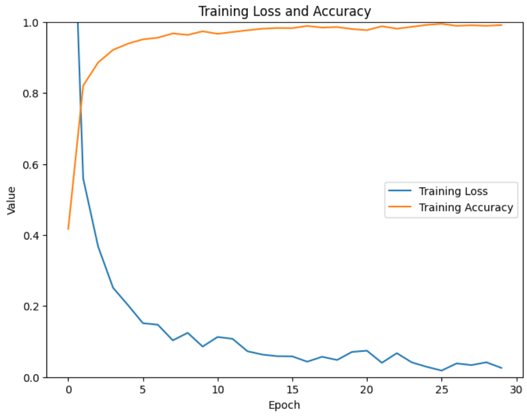

Figure 5.3: Accuracy of IUBMCdb

Finally, we achieved 99% accuracy using IUBMCdb dataset that focused on recognizing handwritten Bangla modifiers in the Bangla script.

5.3 Summary

In Chapter 5, we conducted experiments on NumtaDB, CMATERdb 3.1.2 and IUBMCdb datasets to evaluate our Bangla Handwriting Recognition model. The results were highly promising, demonstrating remarkable accuracy in recognizing individual characters and numerals from both datasets. This underscores the versatility and reliability of our model, positioning it as an effective tool for Bangla handwriting recognition across diverse contexts.

Chapter 6

Conclusion

6.1 Summary

The process of turning a scanned document into a text document that can be readily changed and made searchable is known as optical character recognition, or OCR. OCR is the process of converting handwritten or printed text pictures into mechanical or electronic translation of images. Thus, in order to complete our thesis, we must first gather and process all of the data. After segmenting all of the data to adequately prepare it for our model, we used the prepossessing strategy.

The process of turning a scanned document into a text document that can be readily changed and made searchable is known as optical character recognition, or OCR. OCR is the process of converting handwritten or printed text pictures into achine-editamInitially, we tested our model using a 6-layer convolutional neural network with fully linked structures for the input and output layers. Stochastic gradient descent algorithm optimizer is used in the proposed model. The initial two layers employed the same padding and ReLU activation with 32 filters, a 5x5 kernel, a 2x2 max-pooling layer, and 20% dropout. To lessen overfitting, all dropout layers are employed. We also employ max pooling and the gradient descent technique.

From our model, We understand the importance of recognition, which provides the highest level of accuracy for our Bengali character. We examine the automatically generated graphs to confirm the results. The accuracy and loss function results from each iteration were used to create the graphs.

After using this function along with Numtadb dataset we were able to get accuracy 99.7% and 97.2% accuracy by the CMATERdb 3.1.2 dataset. Finally, we were able to get 99% accuracy using IUBMCdb dataset.

6.2 Conclusion

In our thesis, we developed a comprehensive Optical Character Recognition (OCR) system for recognizing Bangla text across various fonts. Our goal was to enhance Bangla character recognition in computer vision, generating text files from images to digitize handwritten pages. This contributes to building a substantial Bangla character database, enriching computer vision resources and enabling applications like converting handwritten content into editable digital formats.

Our research establishes a universal methodology for processing Bangla language handwritten character databases, streamlining development and serving as a valuable resource across research fields. The anticipated impact includes the creation of highly accurate OCR models for diverse Bangla scripts as the dataset grows over time.

6.3 Future Work

We're going to explore more datasets in the near future applying the same model and same preprocessing techniques and in the future, we will try to extract letters from both words and sentences.

The process is difficult since every letter in a word in Bengali is tied to a word called "matra". we will try to include that in our project as well.

We want to provide more intricate word constructions in this project that combine both basic and compound letters.

References

[1] Akm Shahariar Azad Rabby, Sadeka Haque, Sanzidul Islam, Sheikh Abujar, and Syed Akhter Hossain. Bornonet: Bangla handwritten characters recognition using convolutional neural network. *Procedia computer science*, 143:528–535, 2018.

[2] Ergina Kavallieratou, Nikos Liolios, E Koutsogeorgos, Nikos Fakotakis, and G Kokkinakis. The gruhd database of greek unconstrained handwriting. In *Proceedings of Sixth International Conference on Document Analysis and Recognition*, pages 561–565. IEEE, 2001.

[3] Fei Yin, Qiu-Feng Wang, Xu-Yao Zhang, and Cheng-Lin Liu. Icdar 2013 chinese handwriting recognition competition. In *2013 12th international conference on document analysis and recognition*, pages 1464–1470. IEEE, 2013.

[4] Bilan Zhu, Xiang-Dong Zhou, Cheng-Lin Liu, and Masaki Nakagawa. A robust model for on-line handwritten japanese text recognition. *International Journal on Document Analysis and Recognition (IJDAR)*, 13:121–131, 2010.

[5] Md Zahangir Alom, Paheding Sidike, Tarek M Taha, and Vijayan K Asari. Handwritten bangla digit recognition using deep learning. *arXiv preprint arXiv:1705.02680*, 2017.

[6] Md Mahbubar Rahman, MAH Akhand, Shahidul Islam, Pintu Chandra Shill, MH Rahman, et al. Bangla handwritten character recognition using convolutional neural network. *International Journal of Image, Graphics and Signal Processing*, 7(8):42–49, 2015.

[7] Saikat Roy, Nibaran Das, Mahantapas Kundu, and Mita Nasipuri. Handwritten isolated bangla compound character recognition: A new benchmark using a novel deep learning approach. *Pattern Recognition Letters*, 90:15–21, 2017.

[8] AKM Shahariar Azad Rabby, Sadeka Haque, Sheikh Abujar, and Syed Akhter Hossain. Ekushnet: using convolutional neural network for bangla handwritten recognition. *Procedia computer science*, 143:603–610, 2018.

[9] Nishatul Majid and Elisa H Barney Smith. Segmentation-free bangla offline handwriting recognition using sequential detection of characters and diacritics with a faster r-cnn. In *2019 international conference on document analysis and recognition (ICDAR)*, pages 228–233. IEEE, 2019.

[10] Rumman Rashid Chowdhury, Mohammad Shahadat Hossain, Raihan ul Islam, Karl Andersson, and Sazzad Hossain. Bangla handwritten character recognition using convolutional neural network with data augmentation. In *2019 Joint 8th international conference on informatics, electronics & vision (ICIEV) and 2019 3rd international conference on imaging, vision & pattern recognition (icIVPR)*, pages 318–323. IEEE, 2019.

[11] Yasir Babiker Hamdan and A Sathesh. Deep learning based handwriting recognition with adversarial feature deformation and regularization. *Journal of innovative image processing*, 3(4):367, 2021.

[12] Md Tanvir Hossain, Md Wahid Hasan, and Amit Kumar Das. Bangla handwritten word recognition system using convolutional neural network. In *2021 15th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, pages 1–8. IEEE, 2021.

[13] Nishatul Majid and Elisa H Barney Smith. Performance comparison of scanner and camera-acquired data for bangla offline handwriting recognition. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, volume 4, pages 31–36. IEEE, 2019.

[14] Sadeka Haque, AKM Shahariar Azad Rabby, Md Sanzidul Islam, and Syed Akhter Hossain. Shonkhanet: a dynamic routing for bangla handwritten digit recognition using capsule network. In *Recent Trends in Image Processing and Pattern Recognition: Second International Conference, RTIP2R 2018, Solapur, India, December 21–22, 2018, Revised Selected Papers, Part III 2*, pages 159–170. Springer, 2019.

[15] Tapotosh Ghosh, Min-Ha-Zul Abedin, Hasan Al Banna, Nasirul Mumenin, and Mohammad Abu Yousuf. Performance analysis of state of the art convolutional neural network architectures in bangla handwritten character recognition. *Pattern Recognition and Image Analysis*, 31:60–71, 2021.

[16] Hasibul Huda, Md Ariful Islam Fahad, Moonmoon Islam, and Amit Kumar Das. Bangla handwritten character and digit recognition using deep convolutional neural network on augmented dataset and its applications. In *2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, pages 1–7. IEEE, 2022.

[17] Bishwajit Purkaystha, Tapos Datta, and Md Saiful Islam. Bengali handwritten character recognition using deep convolutional neural network. pages 1–5, 12 2017.

[18] Md Mahbubar Rahman, MAH Akhand, Shahidul Islam, Pintu Chandra Shill, MH Rahman, et al. Bangla handwritten character recognition using convolutional neural network. *International Journal of Image, Graphics and Signal Processing*, 7(8):42–49, 2015.

[19] I Khandokar, M Hasan, F Ernawan, S Islam, and MN Kabir. Handwritten character recognition using convolutional neural network. In *Journal of Physics: Conference Series*, volume 1918, page 042152. IOP Publishing, 2021.

[20] S Hakim and Asaduzzaman . Handwritten bangla numeral and basic character recognition using deep convolutional neural network. pages 1–6, 02 2019.

[21] Md Zahangir Alom, Paheding Sidike, Tarek M Taha, and Vijayan K Asari. Handwritten bangla digit recognition using deep learning. *arXiv preprint arXiv:1705.02680*, 2017.

[22] Tandra Rani Das, Sharad Hasan, Md Rafsan Jani, Fahima Tabassum, and Md Imdadul Islam. Bangla handwritten character recognition using extended convolutional neural network. *Journal of Computer and Communications*, 9(03):158–171, 2021.

[23] Jin-wen Xu, Jinhua Xu, and Yue Lu. Handwritten bangla digit recognition using hierarchical bayesian network. In *2008 3rd International Conference on Intelligent System and Knowledge Engineering*, volume 1, pages 1096–1099. IEEE, 2008.