A Thesis Submitted to the Sylhet Engineering College for the Degree of
Bachelor of Science in **Computer Science and Engineering**

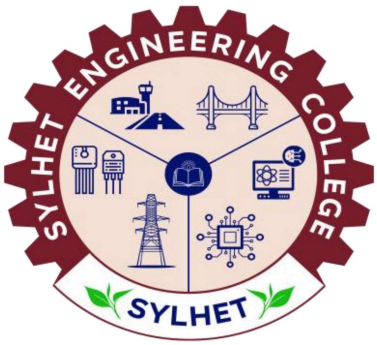
A Deep CNN-Based Approach for Recognizing Handwritten and Printed Meitei Characters

Submitted By-

Sayed Hanzala Abdullah
Registration No.:2019331560
&
Rayhan Mahmud Ansari
Registration No.:2019331566

Supervised By

Md. Abu Naser Mojumder
Associate Professor & Head
Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet



Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet
Affiliated with
Shahjalal University of Science & Technology (SUST)

Recommendation Letter from Thesis Supervisor

The thesis titled **"A Deep CNN-Based Approach for Recognizing Handwritten and Printed Meitei Characters"** submitted by the group mentioned below has been accepted as satisfactory in partial fulfillment of the requirements for the degree of B. Sc. in Computer Science and Engineering.

Group Members:

Sayed Hanzala Abdullah (2019331560)

Rayhan Mahmud Ansari (2019331566)

Signature of the Supervisor

Md. Abu Naser Mojumder
Associate Professor and
Head of CSE Department
Department of Computer Science & Engineering
Sylhet Engineering College

Certificate of Acceptance

The thesis is titled **"A Deep CNN-Based Approach for Recognizing Handwritten and Printed Meitei Characters"** submitted by **Sayed Hanzala Abdullah** and **Rayhan Mahmud Ansari**; Student ID. **2019331560** and **2019331566**; Session **2019-20**, to the Department of Computer Science and Engineering, Sylhet Engineering College, has been accepted as satisfactory in partial fulfilment of the requirement for the Degree of Bachelor of Science in Computer Science and Engineering and approved as to its style and contents.

BOARD OF EXAMINERS

_____

Internal

Nayan Kumar Nath
Lecturer
Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet

_____

Internal

Md Lysuzzaman
Lecturer
Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet

_____

Internal

Md. Rasel Ahmed
Assistant Professor
Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet

_____

Internal

Md. Nagrul Islam
Assistant Professor
Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet

_____

Chairman

Md. Abu Naser Mojumder
Head
Department of Computer Science and Engineering
Sylhet Engineering College, Sylhet

_____

Member (External)

Dr. Mohammad Shahidur Rahman
Professor
Department of Computer Science and Engineering
Shahjalal University of Science and Technology

Acknowledgements

First and foremost, we offer our deepest gratitude to the Almighty, whose boundless mercy and silent guidance have been our constant source of strength throughout this journey. Through every challenge and every success, His blessings have illuminated our path and sustained our resolve.

We would like to express our heartfelt appreciation to our respected supervisor, **Md. Abu Naser Mojumder**, for his invaluable support, continuous encouragement, and insightful feedback. His guidance has been pivotal in shaping the direction and quality of our research, and we are truly grateful for his mentorship.

Our sincere thanks also go to our respected teachers, **Md. Lysuzzaman** and **Nayan Kumar Nath**, whose dedication to teaching and depth of knowledge have left a lasting impact on our academic foundation. Their support and encouragement throughout our studies have been a source of inspiration.

We are especially thankful for the strong collaboration and mutual understanding we shared as thesis partners. The joint effort in completing our thesis, " **A Deep CNN-Based Approach for Recognizing Handwritten and Printed Meitei Characters**" was made possible through shared dedication, teamwork, and a commitment to learning.

Lastly, with all our love and gratitude, we acknowledge the endless support of our families. Their unconditional love, sacrifices, and constant prayers have been the backbone of our academic journey. Without their unwavering belief in us, this achievement would not have been possible.

A Deep CNN-Based Approach for Recognizing Handwritten and Printed Meitei Characters

by

Sayed Hanzala Abdullah, Rayhan Mahmud Ansari

Submitted to the Department of Computer Science & Engineering, in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science & Engineering

Abstract

This paper presents a deep convolutional neural network (CNN) approach for recognizing both handwritten and printed Meitei characters. Manipuri (Meitei), a significant Tibeto-Burman language, uses the unique Meetei Mayek script. Developing an accurate Optical Character Recognition (OCR) system for Meitei is essential for language preservation and enhancing digital accessibility. In this study, we address the challenge of recognizing both printed and handwritten Meitei characters by constructing the "MeiteiBD25" dataset. This dataset comprises 34,913 character samples across 53 classes—including 17,276 handwritten and 17,637 printed characters. A custom segmentation pipeline incorporating binarization, morphological operations, and contour detection was employed to extract individual characters from scanned documents. For classification, several pretrained CNN models—ResNet50, Xception, DenseNet121, and MobileNetV2—were fine-tuned and evaluated. The proposed method achieved a segmentation accuracy of 93% and a classification accuracy of 99.83%. This integrated segmentation-classification framework facilitates accurate recognition of both printed and handwritten Meitei characters, contributing significantly to the digital preservation and accessibility of the Manipuri language.

Keywords: OCR, Meitei, Classification, Segmentation, Printed character recognition, Handwritten character recognition, Text recognition.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Manipuri language, also known as Meitei, is a Tibeto-Burman language. It is the official language of Manipur and an additional official language in four districts of Assam [1]. There are native Meitei speakers mainly living in three different areas of the Sylhet Division in Bangladesh [2]. Hence, the development of machine learning and deep learning techniques for the recognition of the Manipuri language is crucial for enabling communication, improving accessibility, and preserving the language in the digital space.

Optical Character Recognition (OCR) systems digitize a language by converting printed or handwritten characters into a machine-understandable format [3]. OCR is an indispensable technology that enables numerous applications, such as document digitization and archiving, banking and financial services, assistive technology, and automated data entry [4]. Deep learning has revolutionized Optical Character Recognition (OCR) by enabling models to learn intricate patterns and structures in handwritten and printed text [5].

Developing a successful Optical Character Recognition (OCR) system requires multiple phases, including pre-processing, segmentation, feature extraction, classification, and recognition [6]. Meitei follows a disjoint writing style, so contours are used for segmentation.In this study, there are two non-public datasets considered for developing deep learning models. Those are printed character dataset Scanned from multiple books and handwritten dataset. These two datasets are merged together to create a custom dataset to train deep learning models. In this study, we will use several pretrained models such as Xception, DenseNet, MobileNetV2, and ResNet50 to classify characters.

1.1 Background

The rapid advancement of machine learning and deep learning technologies has transformed the field of Optical Character Recognition (OCR). Traditional OCR systems, which relied heavily on

manual feature engineering, often struggled with variability in fonts, styles, and handwritten text. In contrast, deep learning—especially convolutional neural networks (CNNs)—has provided powerful tools for automatically learning complex character patterns directly from raw image data.

Deep learning-based OCR systems have demonstrated strong performance across widely used scripts like Latin, Devanagari, and Arabic. These systems can now handle diverse handwriting styles, noisy inputs, and low-resolution scans with improved accuracy and efficiency. However, research targeting underrepresented languages and scripts, such as Meitei (Meetei Mayek), remains relatively limited—particularly in developing standardized datasets and robust models that address both handwritten and printed forms simultaneously.

Meetei Mayek, the official script used to write the Manipuri language, consists of a distinct set of characters including main letters (Iyek Ipi), associative symbols, semi-consonants (Lonsum Iyek), punctuation marks, and numerals. Its unique character shapes, disjoint writing style, and limited digital resources present significant challenges for OCR development.

While recent works have applied deep learning techniques to Meitei OCR, existing research generally focuses on either printed or handwritten characters in isolation. OCR systems capable of recognizing both handwritten and printed Meitei texts within a unified framework remain underdeveloped. Earlier methods based on traditional techniques—such as histogram projection, k-nearest neighbors (KNN), support vector machines (SVM), and handcrafted features like Histogram of Oriented Gradients (HOG)—also lack scalability and robustness, especially for handwritten text recognition.

Given this gap, our research aims to develop a deep convolutional neural network-based OCR system capable of recognizing both handwritten and printed Meitei characters. By creating a unified, comprehensive dataset and leveraging pre-trained models, this study seeks to improve recognition accuracy across both forms, contributing to practical applications such as document digitization, educational resources, and the digital preservation of Manipuri language materials.

1.2 Problem Statement

Despite the advancements in Optical Character Recognition (OCR) for widely used languages and scripts, Meitei (Meetei Mayek) remains significantly underrepresented in current OCR research

and technologies. Existing OCR systems lack support for the unique structural properties of the Meitei script, such as its disjoint character formation, complex symbol set, and stylistic variations across printed and handwritten forms.

Furthermore, most prior studies either focus solely on printed or handwritten characters, and often rely on traditional machine learning methods with handcrafted features, which are limited in their ability to generalize across diverse writing styles and degraded image conditions. Publicly available datasets for Meitei characters are scarce, and there is no established benchmark for comprehensive evaluation using modern deep learning techniques.

This research addresses these limitations by developing a deep convolutional neural network (CNN)-based OCR system tailored for recognizing both printed and handwritten Meitei characters. By constructing a custom dataset and leveraging pre-trained models, this study aims to improve the accuracy, robustness, and applicability of Meitei OCR systems. The goal is to bridge the gap between technological advancement in OCR and the digital inclusion of under-resourced languages like Manipuri.

1.3 Motivation

The Meitei script, or Meetei Mayek, is an essential part of the cultural and linguistic identity of the Manipuri-speaking population in India and Bangladesh. However, due to limited technological support and digitization efforts, the script remains largely absent from modern digital tools and systems. This creates barriers not only for communication and education but also for preserving the linguistic heritage of a community whose script and literature span centuries.

Building an accurate and efficient OCR system for Meitei characters can play a transformative role in several areas. It can assist in the digital archiving of historical manuscripts, enabling researchers and linguists to access and analyze rare Meitei texts. Educational platforms can incorporate such systems to help students learn the script interactively and improve literacy. Moreover, assistive technologies like text-to-speech tools can be developed for the visually impaired using recognized characters from both printed and handwritten sources.

From a research and technological perspective, tackling the challenges of Meitei OCR contributes to the broader goal of digital equity—bringing underrepresented languages into the scope of AI and machine learning advancements. It also creates a valuable benchmark for applying and evaluating deep learning architectures in low-resource and complex-script scenarios.

The motivation behind this research lies in both the technical challenge of building a robust deep learning-based OCR system for a unique script, and the social importance of promoting inclusivity, education, and cultural preservation through technology.

1.4 Research Area

This research lies at the intersection of computer vision, deep learning, and natural language processing (NLP), with a specific focus on Optical Character Recognition (OCR) for under-resourced scripts. It falls within the broader domain of artificial intelligence (AI) and pattern recognition, where image-based data is analyzed and interpreted through neural network-based techniques.

More specifically, the work contributes to the growing field of script recognition and digitization for low-resource languages, with emphasis on the Meitei (Meetei Mayek) script. By applying convolutional neural networks (CNNs) to both handwritten and printed character datasets, the research aims to enhance the performance of OCR systems for non-Latin scripts and promote language technology development for indigenous and less-digitized writing systems.

This study also contributes to applied research in multilingual computing, digital archiving, and document analysis, with potential impact in educational technology, cultural preservation, and inclusive AI.

1.5 Research Aim

The primary aim of this research is to develop an efficient and accurate deep learning-based Optical Character Recognition (OCR) system capable of recognizing both handwritten and printed Meitei (Meetei Mayek) characters. This involves constructing a custom dataset, applying and fine-

tuning pre-trained convolutional neural network (CNN) models, designing a custom CNN architecture, and evaluating model performance for robust character classification.

By achieving this aim, the research seeks to advance OCR technology for under-resourced languages, contribute to the digital preservation of Meitei script, and support real-world applications such as document digitization, educational tools, and assistive systems.

1.6 Research Objectives

To achieve the overarching goal of developing an effective OCR system for Meitei handwritten and printed characters, the specific research objectives are:

1. **Dataset Development:**
   To construct a comprehensive custom dataset comprising both printed and handwritten Meitei characters, sourced from scanned books and handwritten samples.
2. **Data Preprocessing:**
   To enhance image quality and consistency through preprocessing techniques such as grayscale conversion, noise reduction, and normalization, ensuring the images are suitable for subsequent segmentation and classification tasks.
3. **Segmentation Process Development:**
   To develop and implement a hierarchical segmentation pipeline for isolating individual characters, consisting of:
   - **Page-to-Line Segmentation**
   - **Line-to-Word Segmentation**
   - **Word-to-Character Segmentation**

   using techniques such as binarization, dilation, and contour detection.
4. **Model Fine-Tuning with Pre-trained CNNs:**
   To apply and fine-tune pre-trained convolutional neural network models—including Xception, DenseNet, MobileNetV2, and ResNet50—for feature extraction and character classification.
5. **Model Performance Evaluation:**
   To evaluate the performance of all models using standard classification metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis.

6. **Comparative Analysis:**
   To compare the performance of various pre-trained models to identify the most effective approach for Meitei character recognition.
7. **Contribution to Language Technology:**
   To support the development of language technology for the under-resourced Meitei script, enabling applications in digital archiving, educational resources, and assistive technologies.

1.7 Thesis Organization

This thesis is organized into four primary chapters following the introductory sections. Chapter 1 introduces the research topic, presenting the background, problem statement, motivation, research area, aim, objectives, and an overview of the thesis structure. Chapter 2 provides a comprehensive literature review, covering related work in both classification and segmentation of characters, particularly focusing on OCR techniques for Indic scripts and Meitei characters, and identifying existing research gaps. Chapter 3 details the methodology adopted in this study, starting with dataset preparation and the creation of the MeiteiBD25 dataset. This chapter also describes the hierarchical segmentation process, including line, word, and character-level segmentation, preprocessing steps, classification model development using MobileNetV2, evaluation metrics for both segmentation and classification, and the integrated pipeline for OCR. Chapter 4 presents and discusses the experimental results, including performance analysis of segmentation and classification models, followed by a discussion on findings. Finally, the conclusion and future work section summarizes the key contributions of the thesis and outlines potential directions for future research.

Chapter 2

Literature Review

For the purpose of providing the study's underlying theoretical explanation, this chapter reviews recent literature from journal publications, conference papers, and other sources related to Optical Character Recognition (OCR) and deep learning for Meitei script. It will explore how machine learning techniques have transformed character recognition tasks, the benefits they offer, and the challenges they pose.

The approaches developed for recognizing Meitei characters can generally be categorized into two process:

i. Segmentation

ii. Classification

The approach, segmentation, refers to the process of dividing an input image (such as a scanned page or handwritten text) into meaningful regions or segments, isolating individual characters or components from the background or adjacent characters. The goal of segmentation is to separate characters so that each can be individually recognized.

In contrast, classification involves identifying the category or class of a character by using a trained model that learns patterns from labeled datasets. This approach assigns each character image to one of the predefined character classes (letters, numerals, or symbols) based on features extracted during training.

Both classification and segmentation techniques typically utilize deep learning models that require large datasets for effective training. While classification directly predicts character labels, segmentation plays a crucial role in preprocessing by preparing accurate inputs for classification. One key limitation in segmentation is the complexity of accurately isolating characters in scripts like Meitei, which have disjoint writing styles and diverse fonts.

2.1 Related Work for Classification

[7] This study depicted that the use of a combination of 7*7 sized HOG_CNN features with a Three Nearest Neighbour(3NN) classifier can effectively recognize handwritten Meetei-Mayek characters.In this study,the dataset contain five thousand six hundred handwritten samples of 56 different classes. This study use HOG feature vectors to train a CNN to produce more efficient features that is further trained by the k-Nearest Neighbour classifier.This method achieved an accuracy of 98.714% for a HOG cell size of 7 × 7 and where k = 3.

[8] In this study working with two Indic scripts Bangla and Meitei Mayek and proposed a CNN based model . This gives a stable classifying object techniques for discovering this complex pattern using two datasets of those languages such as cMATERdb and ISI. Create a dataset called Mayek27. Get the accuracy of the proposed model for numeric 99.32% and for character 94.14%.

[9] This study uses transfer learning with pre-trained CNNs (VGG-16, VGG-19, and ResNet152-V2) to develop a printed character recognition system for the Meetei-Mayek language. Here VGG-16 outperforms other models, achieving a classification accuracy of 99.27%.

[10] This paper uses Isolated Handwritten Basic Character Recognition on Meetei Mayek. In this build a character dataset with a total of 8000 plus characters. This paper uses Transfer Learning on a pre-trained Convolutional Neural Network(CNN) used for deep convolutional features of Optical Character Recognition (OCR) for Meetei Mayek. This work uses K-Nearest Neighbour (K-NN) and Support Vector Machine (SVM) classifier models and gets accuracy of 96.63% and 97.55%.

[11] This study uses Optical Character Recognition(OCR) in handwritten Manipuri script. They prepared a neural network for this work. Data set contains 1000 sample half or it is used for training and the rest of it used for validation and testing. The accuracy of models are used in this paper is about 80%-85%

[12] This paper presents a pioneering effort in the recognition of Meitei Mayek's handwritten characters by introducing the first publicly available dataset of 60,285 samples and proposing a CNN-based recognition model. The authors address the lack of resources for this underrepresented script and achieve impressive accuracy (96.24% on the original dataset and 97.09% on the augmented dataset). The use of data augmentation to prevent overfitting and the detailed CNN architecture are notable strengths. The work is significant for preserving and digitizing the Meitei Mayek script, with potential applications in education and cultural preservation. However, further validation of diverse datasets and comparison with other state-of-the-art models could enhance its impact. Overall, it is a valuable contribution to OCR research for Indian scripts.

[13] In this study, the authors present a novel CNN-based framework for recognizing handwritten Meitei Mayek characters, addressing the lack of research and resources for this underrepresented script. A publicly available dataset of 60,000+ samples is introduced, alongside data augmentation techniques to enhance model robustness. The proposed CNN architecture achieves high accuracy (~96–97%), demonstrating its effectiveness. However, the study could be strengthened by comparing state-of-the-art models and exploring real-world scenarios (e.g., noisy documents). Overall, this work significantly advances OCR research for regional scripts and supports cultural preservation efforts.

2.2 Related Work for Segmentation

[14] This paper uses Optical Character Recognition (OCR) for segmentation for Meitei/Meetei Script. It uses 183 lines in a document 926 words and 4138 characters and gets an overall accuracy 95.6% by using their technique.

[15] This paper discusses methods for segmenting lines, words, and characters in Manipuri machine-printed text, a crucial step for developing an Optical Character Recognition (OCR) system. It details the segmentation process, including thresholding, noise removal, and projection-based methods for text segmentation. Additionally, the paper explores various recognition techniques such as Template Matching, Neural Networks, Statistical Classifiers, Hidden Markov Models (HMM), and Support Vector Machines (SVM). Despite the limited research on OCR for

Manipuri, this study provides a foundational framework. However, it lacks experimental validation, a benchmark dataset, and quantitative performance metrics, which limits reproducibility. Still, it serves as an essential step toward advancing Manipuri text recognition and future OCR development.

[16] The paper discusses methods for segmenting lines, words, and characters in Manipuri machine-printed text, which is crucial for Optical Character Recognition (OCR) systems. It outlines the segmentation process, including thresholding, noise removal, and projection-based methods for line, word, and character segmentation. The paper also explores recognition approaches such as Template Matching, Neural Networks, Statistical Classifiers, Hidden Markov Models (HMM), and Support Vector Machines (SVM). The research highlights that while OCR development for Manipuri is limited, this study provides foundational work for future advancements. However, the paper lacks experimental validation, quantitative results, and a dataset, limiting its reproducibility. Despite these limitations, the study offers a strong foundation for future research in Manipuri text recognition.

2.3 Research Gap

Despite significant advancements in OCR technology for many scripts, the Meitei script remains insufficiently explored, primarily due to the lack of a standardized dataset that integrates both printed and handwritten characters. Current segmentation techniques often struggle to generalize effectively across such diverse data, leading to common issues such as over-segmentation and under-segmentation. Consequently, the development of a robust segmentation pipeline capable of accurately handling combined datasets is essential. Furthermore, pre-trained convolutional neural network models have not been comprehensively benchmarked on datasets merging printed and handwritten Meitei characters, resulting in a limited understanding of the most effective architectures for this recognition task. Addressing these critical gaps will substantially contribute to the advancement of OCR research and practical applications for the Meitei script.

Chapter 3

Methodology

This chapter outlines the step-by-step methodology used to build a robust OCR system for recognizing handwritten and printed Meitei characters. The process is divided into three main phases: **dataset preparation, segmentation**, and **classification**.

3.1 Dataset Preparation

3.1.1 Meitei Characters Overview

The Meitei script, also known as Meetei Mayek, is the official script used for writing the Manipuri language. It has a unique set of characters, which are significantly different from scripts like Devanagari or Latin.

3.1.1.2 Character Set

Meetei Mayek consists of a total of 53 types of characters and signs.

3.1.1.3 Total Number of Meitei Script Symbols

Category	Number of Characters
Main Letters (IyekIpi)	27
Associative Symbols(CheitapMayek)	8
Semi Consonants(LonsumIyek)	8
Numeral (CheisingEyek)	10

Table 3.1 Types of Character

KOK	SAM	LAI	MIT	PA	NA	CHIL
TIL	KHOU	NGOU	THOU	WAI	YANG	HUK
UN	I	PHAM	ATIYA	JHAM	RAI	RAI
BA	JIL	DIL	GHOU	DHOU	BHAM	

Figure 3.1: Main Letters (IyekIpi)

(ot nap)	(inap)	(aatap)	(yetnap)
(sounap)	(unap)	(cheinap)	(nung)

Figure 3.2: Associative Symbols (CheitapMayek)

KOK LONSUM	LAI LONSUM	MIT LONSUM	PA LONSUM
NA LONSUM	TIL LONSUM	NGOU LONSUM	I LONSUM

Figure 3.3: Semi Consonants(LonsumIyek)

(ama)	(ani)	(ahum)	(mari)
(manga)	(taruk)	(taret)	(nipal)
(mapal)	(tara)		

Figure 3.4: Numeral (CheisingEyek)

3.1.2 Data Collection

We collected 200 pages from books written in the Meitei language to prepare the machine-printed character dataset. Additionally, 50 people from different age groups participated in creating the handwritten character dataset.

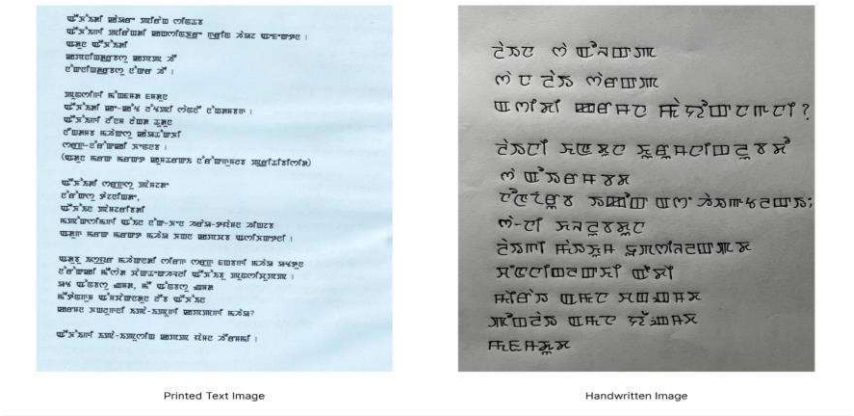


Figure 3.5: sample images

3.2 Segmentation Process

The segmentation process is designed to extract individual Meitei characters from document images, which may contain both handwritten and printed texts. A hierarchical, three-stage segmentation approach—line-level, word-level, and character-level—is adopted to ensure high accuracy in separating characters.

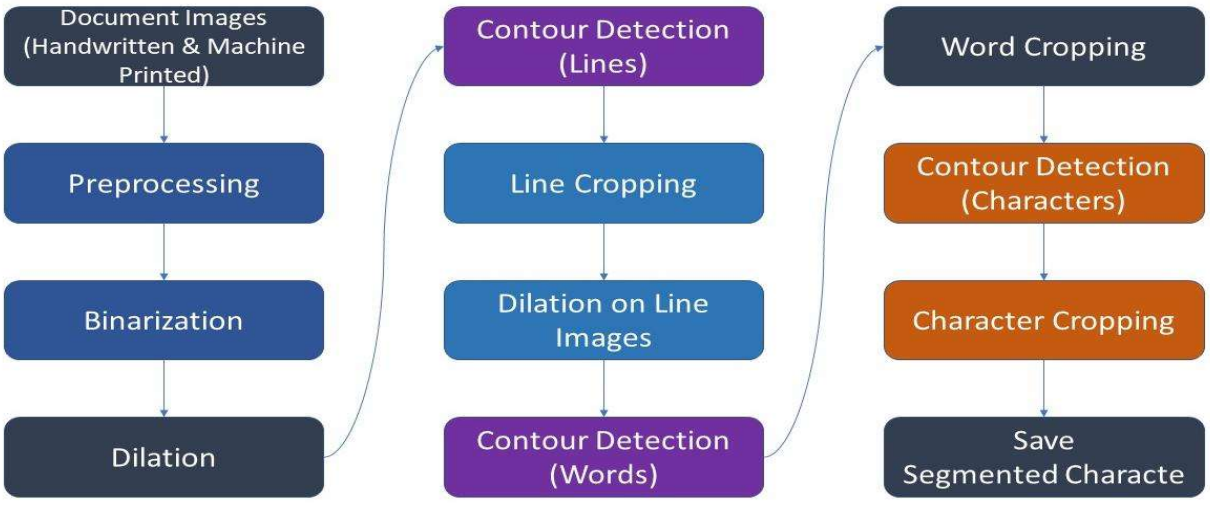


Figure 3.6: Methodology diagram for Segmentation

3.2.1: Preprocessing

The input document images are first converted to grayscale to simplify pixel values and then denoised using Gaussian blurring. This reduces background noise and enhances the clarity of character strokes. A fixed thresholding method is used for binarization, transforming the grayscale image into a binary format with clear contrast between text and background.

3.2.2: Line-Level Segmentation

A large horizontal morphological dilation is performed using a rectangular structuring element of size 20×4 pixels. This connects adjacent characters and words into larger text line blocks. Contour detection is then applied to locate and crop these continuous lines from the document image.

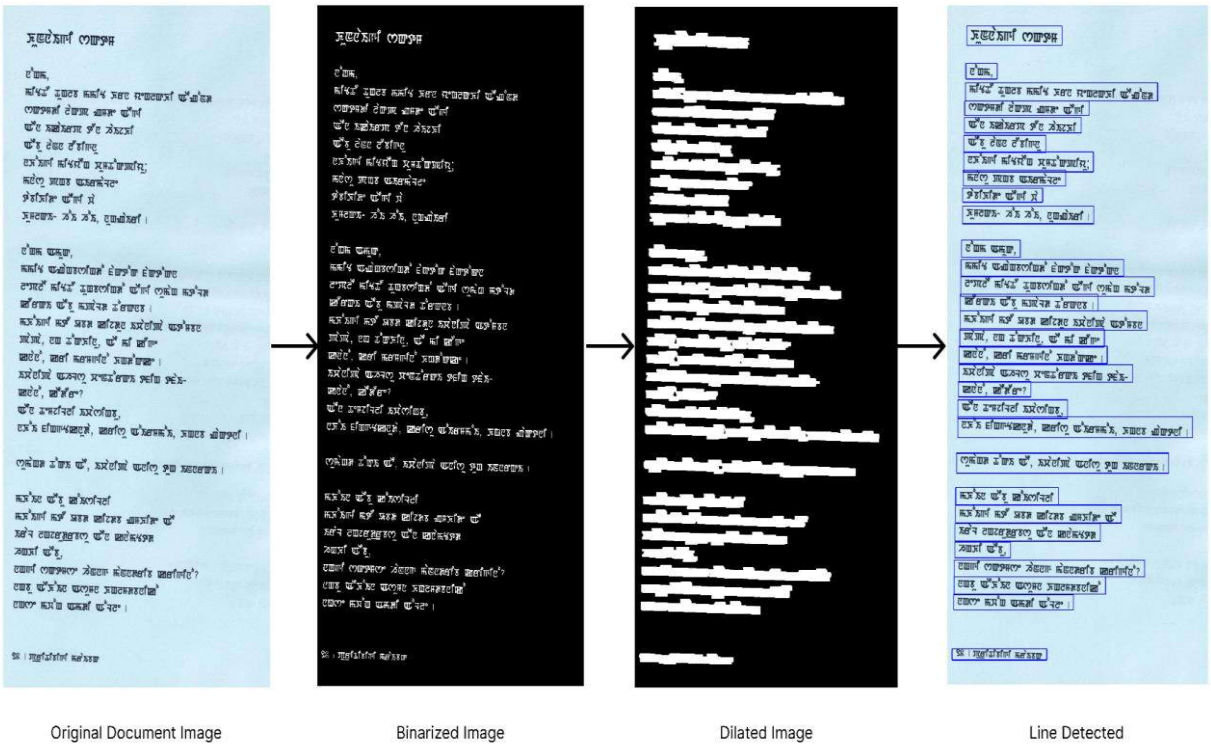


Figure 3.7: Line detection in printed document image

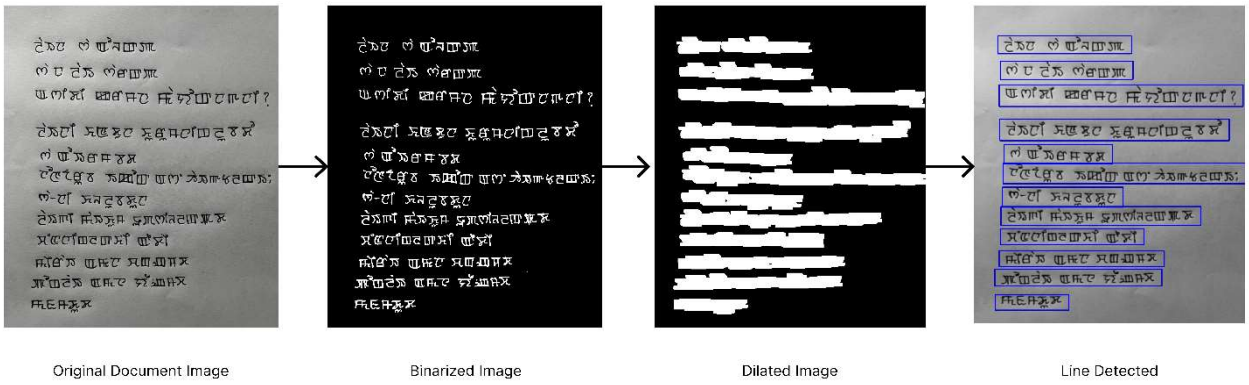


Figure 3.8: Line detection in handwritten document image

3.2.3: Word-Level Segmentation

Within each cropped line, a second round of dilation is applied with a smaller kernel (7×4 pixels) to group nearby characters into word-level regions without over-merging. Contour detection is again used to identify and extract word blocks from each line.

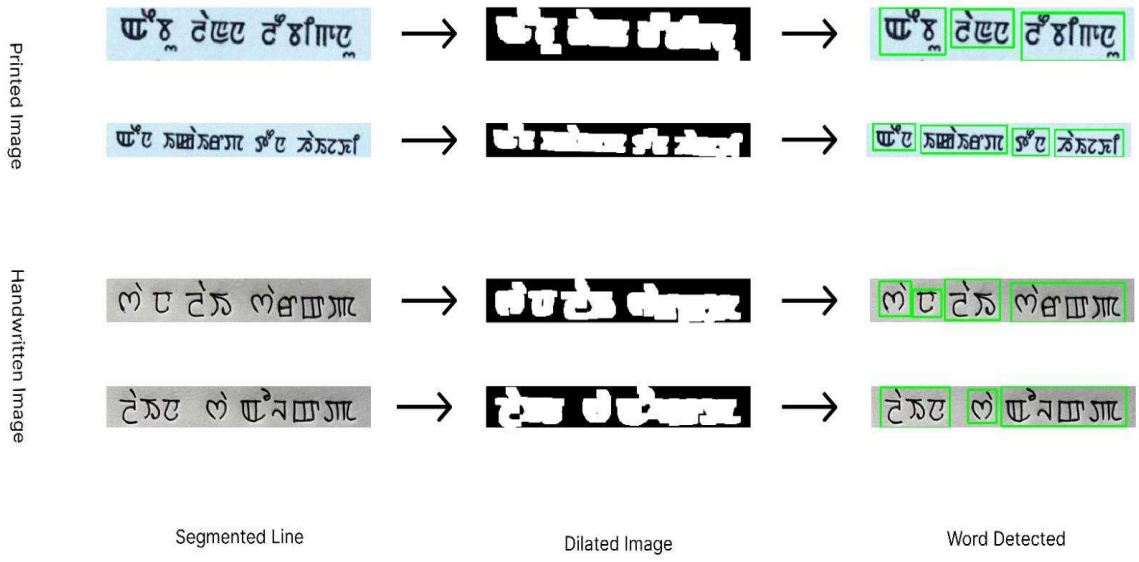


Figure 3.9: Word detection from line image

3.2.4: Character-Level Segmentation

The final segmentation step is applied directly to each word block. No dilation is used here to avoid merging individual characters. Contour detection is applied to isolate and extract each character. Noise or tiny elements that are not valid characters are removed using size-based filtering. The remaining characters are saved as individual images.

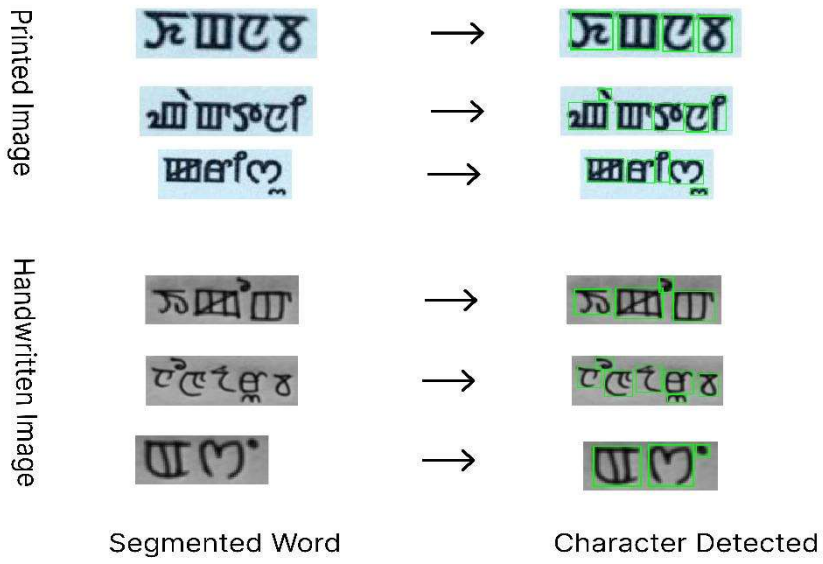


Figure 3.10: Character detection from word image

All extracted characters are manually verified and labeled into one of 53 unique classes that represent the complete Meitei script. This segmentation pipeline ensures clean, well-structured datasets ready for classification.

3.2.5 MeiteiBD25 Dataset

After applying the segmentation process, characters were extracted from the collected pages. These characters were then manually labeled into 53 distinct classes.

We annotated a total of 34,913 characters, comprising 17,276 handwritten and 17,637 printed samples. All characters are categorized into 53 classes.

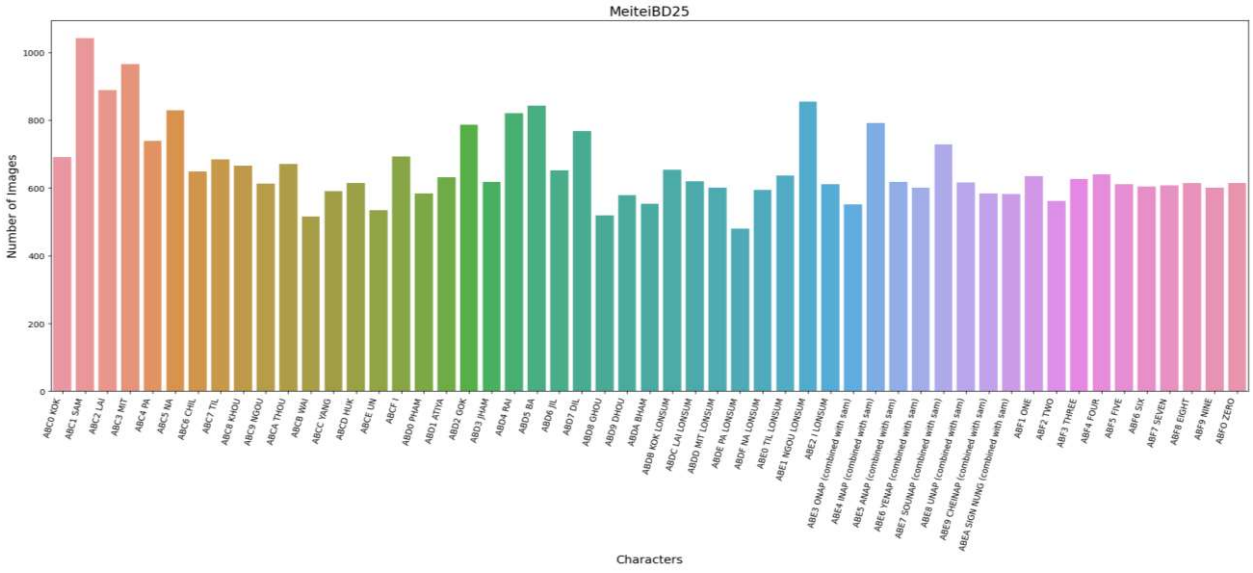


Figure 3.11: Data visualization for all class

3.3 Classification Process

To recognize the segmented Manipuri characters, this study utilizes deep learning-based image classification techniques. The classification process consists of data preparation, model selection, and training.

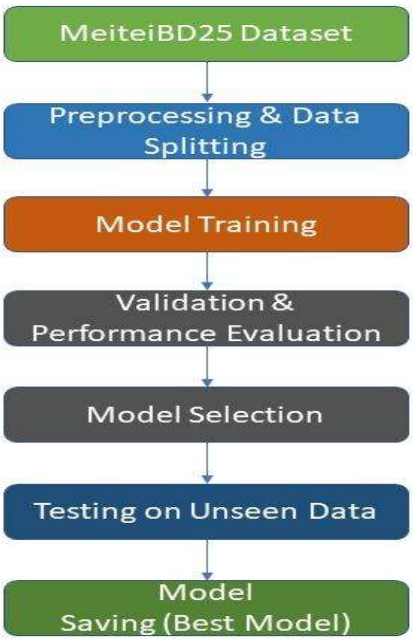


Figure 3.12: Methodology diagram for Classification

3.3.1 Dataset Preparation

The dataset used is MeiteiBD25, which contains grayscale images of segmented Meitei characters. Each image is converted to RGB and resized to 224×224 pixels to meet the input requirements of modern CNN architectures. The dataset is randomly split into three sets:

- **Training set:** 80% of the data
- **Validation set:** 10% of the data
- **Test set:** 10% of the data

Special care is taken to ensure that each class contains both printed and handwritten examples in all splits. The training set contains 27912 images, the validation set 3470 images, and the test set 3531 images. The class-wise distribution is shown in Figure 3.13.

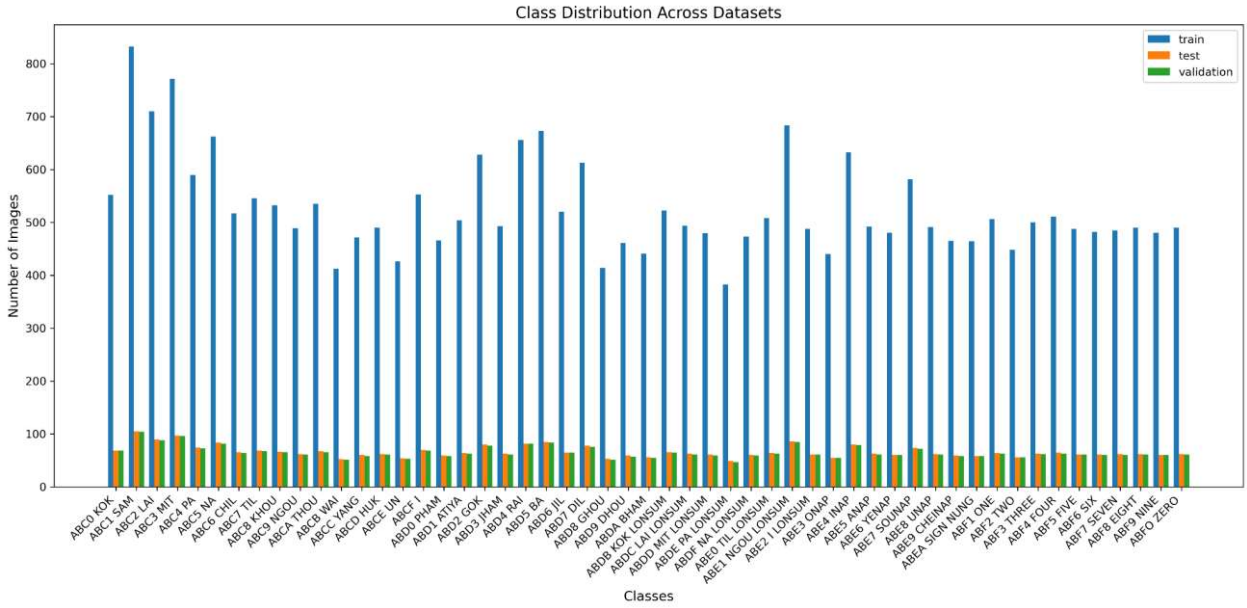


Figure 3.13: Distinct class count

3.3.2 Model Selection and Training

Four state-of-the-art CNN models were selected:

- MobileNetV2
- ResNet50
- Xception
- DenseNet

All models are initialized with ImageNet pre-trained weights and fine-tuned for the 53-class Meitei character classification task. The models use the Adam optimizer with a batch size of 32. The input image size for training is 224×224 pixels. Each model is trained for a different number of epochs and learning rate:

- MobileNetV2: 16 epochs with learning rate of 0.00001
- ResNet50: 10 epochs with learning rate of 0.000001
- Xception: 10 epochs with learning rate of 0.00001
- DenseNet: 10 epochs with learning rate of 0.000001

Cross-entropy loss is used as the loss function for all models.

3.3.3 Classification Model: MobileNetV2

MobileNetV2 is a lightweight CNN architecture designed for mobile and embedded devices. It utilizes depth-wise separable convolutions along with inverted residual blocks and linear bottlenecks to reduce computational cost while maintaining accuracy. Despite being compact, MobileNetV2 achieves competitive performance in character recognition tasks. Its efficient design makes it suitable for deployment on low-power devices or real-time OCR applications where speed and resource efficiency are critical.

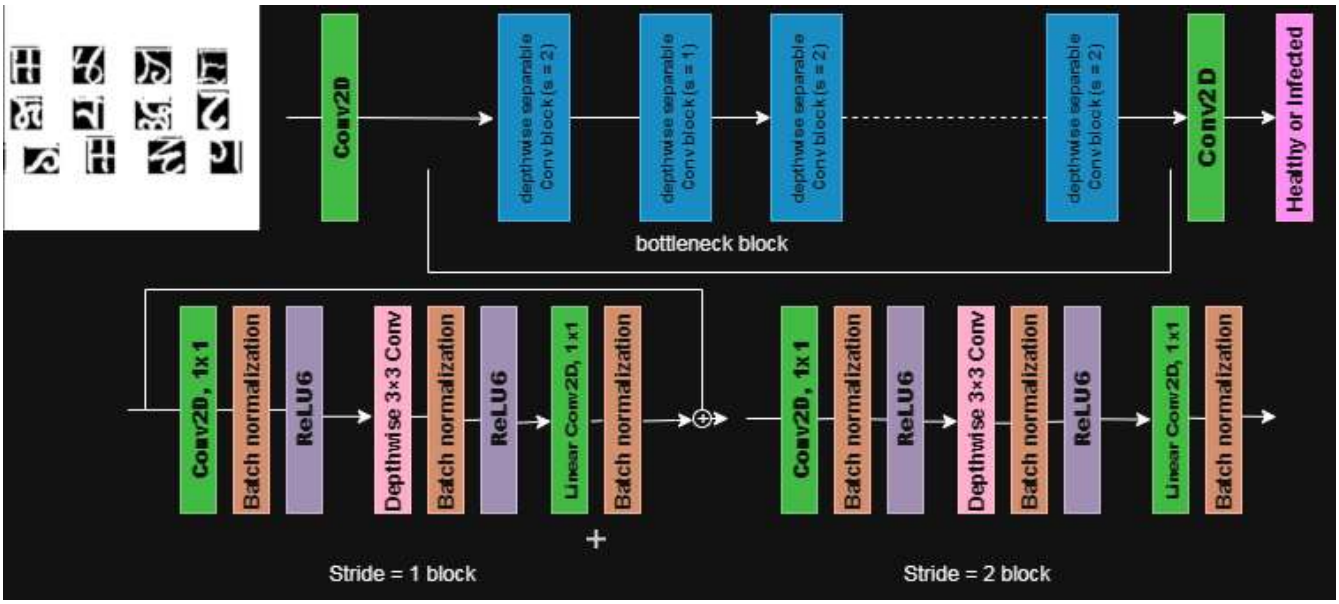


Figure 3.14: Architecture of MobileNetV2

3.4 Evaluation Metrics

This section details the evaluation criteria used to measure the performance of both the segmentation and classification modules of the OCR system.

3.4.1 Segmentation Evaluation Metrics

To assess the quality of the hierarchical segmentation process, the following metrics are used:

- **Word Accuracy (WA):**
  Measures the percentage of correctly segmented word regions compared to the ground truth.

$$Word\ Segmentation\ Accuracy = \frac{Segmented\ word\ count}{Actual\ word\ count}$$

- **Character Accuracy (CA):**
  Measures the percentage of correctly segmented individual characters, ensuring that the segmentation does not miss or incorrectly cut characters.

$$Character\ Segmentation\ Accuracy = \frac{Segmented\ character\ count}{Actual\ character\ count}$$

These metrics help in understanding whether the segmentation process is producing clean and accurate input for classification.

- **Over-Segmentation Rate (OSR)**

Over-segmentation refers to cases where a single character is mistakenly split into multiple parts. OSR provides insight into whether the segmentation process is too aggressive in identifying boundaries. The Over-Segmentation Rate (OSR) measures the proportion of such errors and is defined as:

$$\text{Over Segmentation Rate } = \frac{\text{Number of Over-Segmented Characters}}{\text{Total Number of Ground Truth Characters}} \times 100$$

Where:

- Number of Over-Segmented Characters refers to the total count of characters that have been incorrectly split into more than one segment.
- Total Number of Ground Truth Characters represents the actual number of correctly segmented characters based on manual annotation.

A lower OSR value indicates better segmentation performance, minimizing unnecessary splits within characters.

- **Under-Segmentation Rate (USR)**

Under-segmentation occurs when two or more distinct characters are incorrectly merged and treated as a single segment. USR indicates whether the segmentation process fails to correctly separate adjacent characters. The Under-Segmentation Rate (USR) quantifies this issue as:

$$\text{Under Segmentation Rate} = \frac{\text{Number of Under-Segmented Characters}}{\text{Total Number of Ground Truth Characters}} \times 100$$

Where:

- Number of Under-Segmented Characters represents the count of cases where multiple characters are merged into a single segment.
- Total Number of Ground Truth Characters refers to the actual number of distinct characters in the dataset.

A lower USR value signifies fewer merging errors and better character separation.

3.4.2 Classification Evaluation Metrics

The performance of the classification models is evaluated using the following standard metrics:

- **Accuracy:**

  The overall percentage of correctly classified character images.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

- **Precision:**

  The ratio of true positive predictions to all positive predictions, calculated per class and averaged.

$$Precision = \frac{TP}{TP + FP}$$

- **Recall:**

  The ratio of true positive predictions to all actual positives in the dataset, reflecting the model's ability to detect each class.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1 Score:**

  The harmonic mean of precision and recall, offering a balanced metric especially important for imbalanced datasets.

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Here, TP represents true positives, FP represents false positives, TN represents true negatives, and FN represents false negatives. These metrics allow for a comprehensive comparison of all four deep learning models and help identify the most suitable model for deployment.

3.5 Model Pipelining

To create a unified OCR system, we adopted a pipelining approach that combines segmentation and classification models. The segmentation model first processes the input document image—whether handwritten or printed—breaking it down into individual character images. These segmented characters are then passed to the classification model, which assigns the corresponding Meitei script label to each character using a trained CNN. The system's final output is a digitized version of the Meitei script text, suitable for applications such as digital archiving, translation, and educational tools. Figure 3.15 illustrates the overall system architecture after integration.

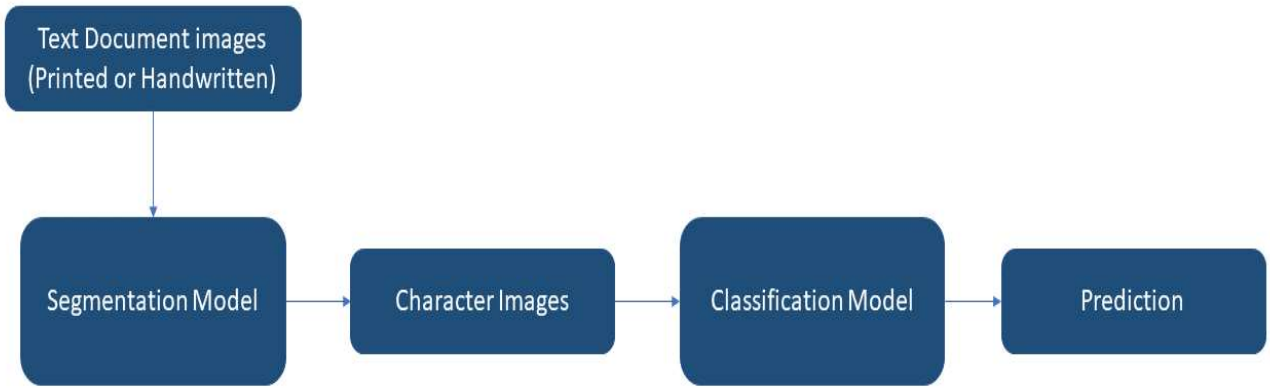


Figure 3.15: Pipelining the model

Chapter 4

Result and discussion

This chapter presents the results obtained from the segmentation and classification experiments, followed by a comprehensive analysis. The performance of the proposed OCR system is evaluated using the metrics described in the previous chapter.

4.1 Experimental Environment

All experiments were conducted on a personal workstation equipped with an Intel Core i5-12600K processor, 16 GB of RAM, and an NVIDIA GeForce RTX 3060 GPU, running on the Windows operating system. The deep learning models were implemented using Python 3.11.9 and the PyTorch 2.6.0 framework. This hardware–software configuration provided adequate computational resources to efficiently train deep learning models and perform real-time inference.

4.2 Analysis and Discussion Results

4.2.1 Segmentation model

Assessing the performance of character segmentation remains inherently challenging due to the absence of standardized evaluation metrics and the largely unsupervised nature of many segmentation methods. Unlike word segmentation—where evaluation can directly rely on ground truth annotations—character segmentation lacks universally accepted benchmarks for accuracy assessment. This necessitates the adoption of customized evaluation strategies tailored to the specific context and objectives of the segmentation task.

To evaluate the segmentation performance, two document images were randomly selected: one containing handwritten text and another containing printed text. In the printed document, the ground truth word count was 177, and the segmentation model successfully extracted 177 words, yielding a word segmentation accuracy of 100%. In the handwritten document, with a ground truth

word count of 61, the model extracted 56 words, resulting in a word segmentation accuracy of approximately 93%. Overall, the combined word segmentation accuracy was approximately 97%, demonstrating the model's robustness across varying text formats.

For character-level evaluation, 60 words images (30 handwritten and 30 printed) were randomly selected. In the printed word sample, 160 characters were present, while the model segmented 152, achieving a character segmentation accuracy of 95%. The Over-Segmentation Rate (OSR) was 1.25%, and the Under-Segmentation Rate (USR) was 1.875%. In the handwritten sample, 172 characters were present, and the model segmented 157 characters, yielding a character segmentation accuracy of 91.27%, with OSR at 0.5% and USR at approximately 4%. The overall character segmentation accuracy was approximately 93%.

Minor over-segmentation indicates occasional splitting of single characters into multiple fragments, especially within cursive or closely connected forms. Nonetheless, the system exhibits strong reliability in character isolation. Further improvements—such as adaptive thresholding or enhanced contour filtering—could potentially enhance segmentation precision.

4.2.2 Classification model

4.2.2.1 Performance of classification model on training

For the classification task, four CNN-based models—ResNet50, Xception, DenseNet, and MobileNetV2—were trained and validated. Among these models, MobileNetV2 achieved the best performance across multiple evaluation metrics, such as accuracy, precision, recall, and F1-score. Figure. 4.1 shows the loss curve for all those models in classifying the Manipuri character. The loss curves reflect how well each model reduces its errors throughout the training epochs. The MobileNetV2 model demonstrates a steep decline in loss, reaching a value nearly zero compared to the other models. At the time of termination of the training process, the model's loss was 0.0029, which indicated the strong capability of the model to generalize the training set of data. The consistent decline and closeness between training and validation loss indicate that the model is learning effectively while possessing good generalization with no sign of overfitting.

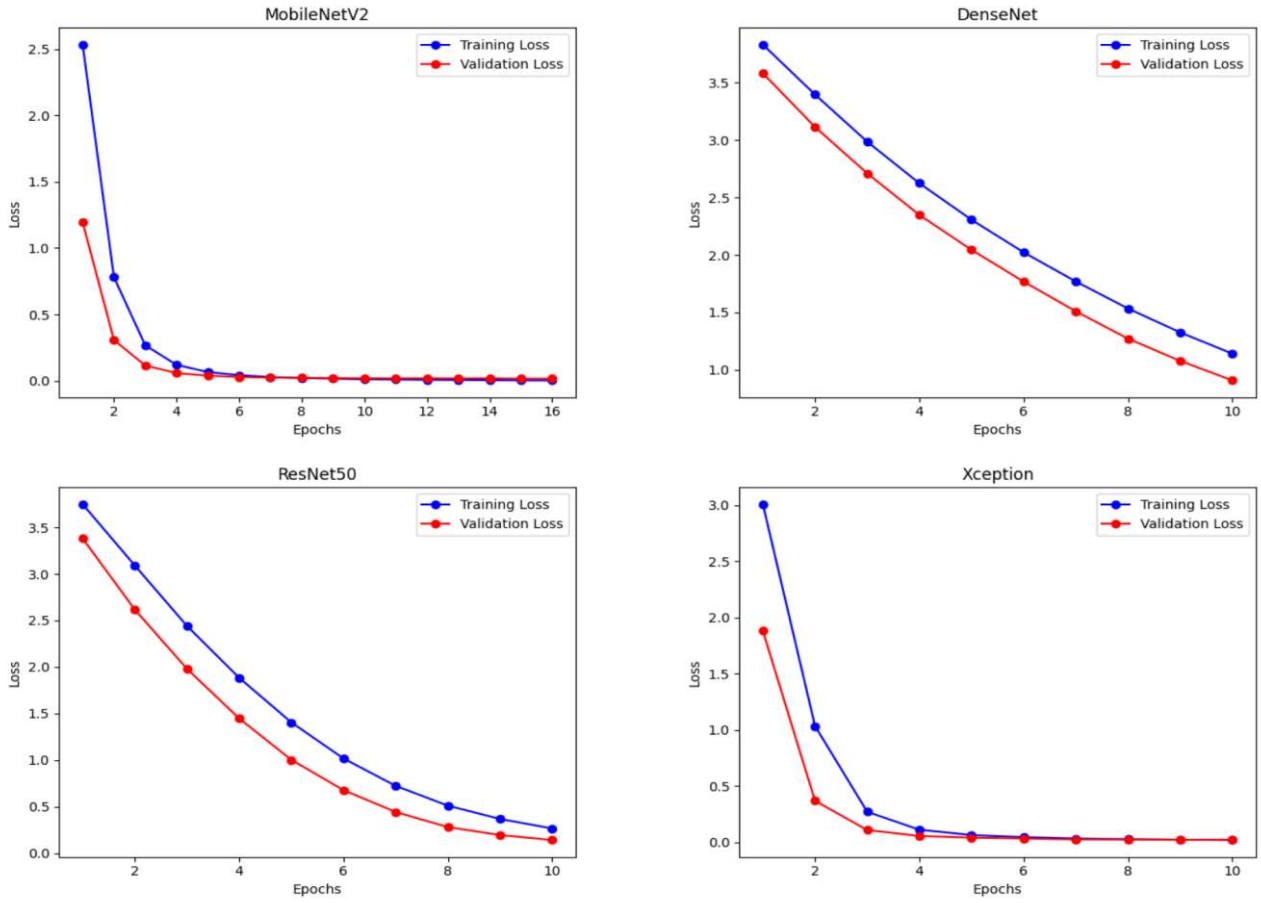


Figure 4.1: Models loss curve

Figure 4.2 shows the accuracy curve for all of those models for classifying the Manipuri character. The accuracy curves indicate the models' ability to generalize data over the training period. MobileNetV2 outperformed other models, as it obtained a training accuracy of 99.94% at the end of the training epochs. The validation accuracy mirrors the training accuracy, showing a similar increase over time. This parallel rise in both curves suggests that the model is not only performing well on the training data but is also generalizing effectively to unseen validation data. This indicates that the model is able to generalize to new, unknown data and is avoiding overfitting.

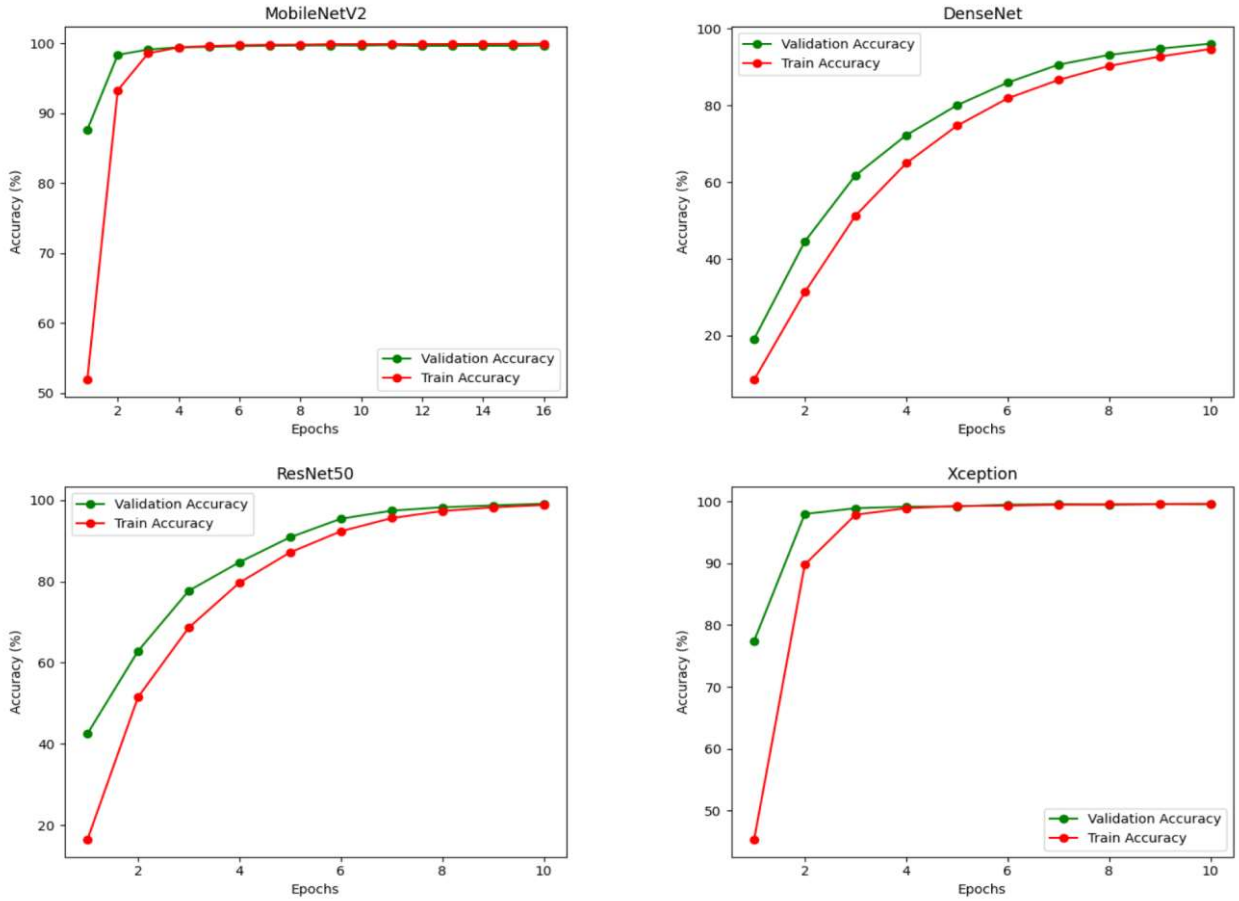


Figure 4.2: Models Accuracy curve

Table 4.1 represents the performance metrics—accuracy and loss (error)—for all models in training and validation. High accuracy and minimized loss on both validation and training data indicate that MobileNetV2 effectively fits the data and makes mostly correct predictions.

Models	Train		Validation	
	Accuracy	Loss	Accuracy	Loss
MobileNetV2	99.9463%	0.0029	99.7118%	0.0159
ResNet50	98.8105%	0.2654	99.1066%	0.1410
DenseNet	94.7012%	1.1411	96.0807%	0.9081
Xception	99.6310%	0.0187	99.5389%	0.0229

Table 4.1: Comparative Training Results for CNN Architectures

4.2.2.2 Performance on Test Dataset: v1

Table 4.2 compares the performance metrics (accuracy, precision, recall, and F1-score) of four deep learning models on the classification task. These results help to identify the most accurate, reliable, and generalizable model for real-world classification tasks like Manipuri character classification. MobileNetV2 clearly outperformed the other models across all metrics, with an accuracy of 99.92%, precision of 99.91%, recall of 99.92%, and an F1-score of 99.91%. These results indicate that the model is highly effective at correctly identifying positive cases (high recall), minimizing false positives (high precision), and maintaining a strong overall balance between these measures (high F1-score). It indicates that the model is robust, reliable, and well-optimized for the task in terms of both correctness and consistency.

Models	Accuracy	Precision	Recall	F1
MobileNetV2	99.83%	99.83%	99.85%	99.84%
ResNet50	99.43%	99.45%	99.41%	99.42%
DenseNet	99.63%	99.63%	99.64%	99.63%
Xception	99.75%	99.75%	99.75%	99.75%

Table 4.2: Comparative Training Results for CNN Architectures

Table 4.3 shows the class-wise precision, recall, and F1 score, which give a comprehensive overview of the model performance on each class. It demonstrates that MobileNetV2 achieved a high accuracy even at the individual character level, which is paramount to developing a reliable OCR system. It indicates the model's capability to effectively handle the complex structure of Manipuri scripts.

	precision	recall	f1-score	support
ABC0 KOK	1	1	1	69
ABC1 SAM	1	1	1	105
ABC2 LAI	1	0.9888	0.9944	90
ABC3 MIT	1	1	1	97
ABC4 PA	1	1	1	75
ABC5 NA	1	0.988	0.994	84
ABC6 CHIL	0.985	1	0.9924	66
ABC7 TIL	1	1	1	69
ABC8 KHOU	1	1	1	67
ABC9 NGOU	1	1	1	62
ABCA THOU	1	1	1	68
ABCB WAI	1	1	1	52
ABCC YANG	1	1	1	60
ABCD HUK	1	1	1	62
ABCE UN	1	1	1	54
ABCF I	1	1	1	70
ABD0 PHAM	1	1	1	59
ABD1 ATIYA	1	1	1	64
ABD2 GOK	1	0.9875	0.9937	80
ABD3 JHAM	1	1	1	63
ABD4 RAI	1	1	1	82
ABD5 BA	1	1	1	85
ABD6 JIL	0.9848	1	0.9923	65
ABD7 DIL	1	1	1	78
ABD8 GHOU	1	1	1	53
ABD9 DHOU	0.9833	1	0.9915	59
ABDA BHAM	1	0.9821	0.9909	56
ABDB KOK LONSUM	0.9701	0.9848	0.9774	66
ABDC LAI LONSUM	1	1	1	63
ABDD MIT LONSUM	0.9838	1	0.9918	61
ABDE PA LONSUM	1	1	1	49
ABDF NA LONSUM	1	1	1	60
ABE0 TIL LONSUM	1	1	1	64
ABE1 NGOU LONSUM	1	0.9883	0.9941	86
ABE2 I LONSUM	1	1	1	61
ABE3 ONAP	1	1	1	55
ABE4 INAP	1	1	1	80
ABE5 ANAP	1	1	1	63
ABE6 YENAP	1	1	1	60
ABE7 SOUNAP	1	1	1	74
ABE8 UNAP	1	1	1	62
ABE9 CHEINAP	1	1	1	59

ABEA SIGN NUNG	1	1	1	58
ABF1 ONE	1	1	1	64
ABF2 TWO	1	1	1	56
ABF3 THREE	1	1	1	63
ABF4 FOUR	1	1	1	65
ABF5 FIVE	1	1	1	61
ABF6 SIX	1	1	1	61
ABF7 SEVEN	1	1	1	62
ABF8 EIGHT	1	1	1	62
ABF9 NINE	1	1	1	60
ABFO ZERO	1	1	1	62
accuracy			0.9983	3531
macro avg	0.9982	0.9984	0.9983	3531
weighted avg	0.9983	0.9983	0.9983	3531

Table 4.3: Classification Report

Table 4.3 illustrates that the model generalizes new data well in every class despite the fact of how many samples each class has, as it achieved high macro scores of almost 99.8%. The weighted scores for precision, recall, and F1 score are almost 99.8%, which indicates that the model demonstrates exceptional and balanced classification performance. This balance performance is important, especially for OCR of the Manipuri language, as some classes may have fewer samples or the complex and versatile shape of printed and handwritten characters.

Figure. 4.3 represents the confusion matrix, which is an important tool for performance evaluation. It compares the actual and predicted classes to clearly show the model's accuracy and errors. It is clear from that MovileNetv2 achieves high classification capabilities, as there are very few mistakes it made on testing data which consists of both handwritten and printed characters.

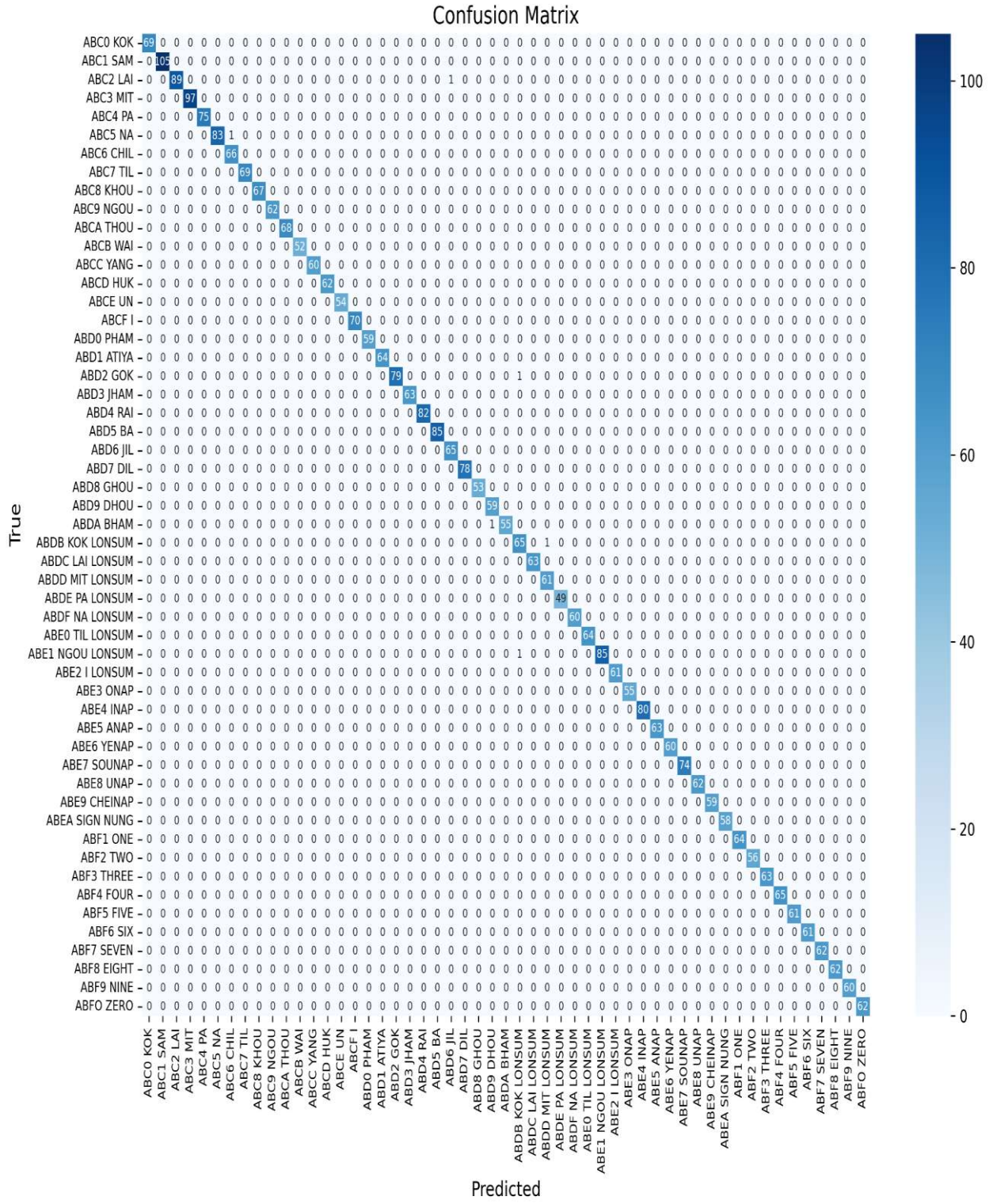


Figure 4.3: Confusion matrix

Table 4.4 compares how well MobileNetV2 performs on three types of testing datasets—a dataset consisting of both handwritten and printed characters, a dataset consisting only of handwritten characters, and a dataset consisting only of printed characters. Though the model was trained on the merged dataset, it shows the model's capability for classifying handwritten and printed characters separately. The model achieved 99.78% accuracy on handwritten characters and 99.92% on machine-printed characters. It indicates that the model has a strong capability to generalize Manipuri characters despite their shape variety for human writing or machine printing.

	Printed	Handwritten	Merged
Accuracy	99.78%	99.88%	99.83%
Support	1802	1729	3531

Table 4.4: Classification Results: Printed, Handwritten, and Merged Data

Chapter 5

Conclusion and Future work

Developing an OCR system for the Manipuri (Meitei) script remains a challenging task due to its under-researched nature and the absence of standard datasets. In this work, a deep learning-based recognition system was proposed, focusing on both handwritten and printed characters. A segmentation approach using morphological operations and contour detection was adopted, though alternative methods could be explored. Pretrained CNN models, including ResNet50, Xception, DenseNet121, and MobileNetV2, were fine-tuned for character classification, demonstrating the potential of transfer learning for this low-resource script. While the system shows promising performance, further improvements are needed to make it suitable for large-scale, real-world applications.

In particular, handwritten characters present more variability and segmentation challenges compared to printed forms. In future work, we aim to improve the current segmentation pipeline to better handle different handwriting styles and inconsistencies. Expanding the dataset with more diverse handwritten samples and exploring lightweight models for deployment are also priorities. This ongoing effort is intended to address current limitations and contribute towards a practical, scalable OCR solution for the Manipuri script.

References

[1] Chelliah, S., & Ray, S. (2002). Early Meithei manuscripts. In *Medieval Tibeto-Burman Languages* (pp. 59-71). Brill.

[2] Faquire, A. B. M. R. K., & Karim, R. (2010). Language situation in Bangladesh. The Dhaka University Studies, 67(2), 63-77.

[3]Memon, J., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). IEEE access, 8, 142642-142668.

[4] Asif, A. M. A. M., Hannan, S. A., Perwej, Y., & Vithalrao, M. A. (2014). An overview and applications of optical character recognition. Int. J. Adv. Res. Sci. Eng, 3(7), 261-274.

[5] Subramani, N., Matton, A., Greaves, M., & Lam, A. (2020). A survey of deep learning approaches for ocr and document understanding. arXiv preprint arXiv:2011.13534.

[6] Hamad, K., & Kaya, M. (2016). A detailed analysis of optical character recognition technology. International Journal of Applied Mathematics Electronics and Computers, (Special Issue-1), 244-249.

[7] Nongmeikapam, K., Wahengbam, K., Meetei, O.N. and Tuithung, T., 2019. Handwritten Manipuri Meetei-Mayek classification using convolutional neural network. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4), pp.1-23.

[8] Hazra, A., Choudhary, P., Inunganbi, S., & Adhikari, M. (2021). Bangla-Meitei Mayek scripts handwritten character recognition using convolutional neural network. Applied Intelligence, 51(4), 2291-2311.

[9] Dingari, V. V. S., Kosanam, G., Chavatapalli, D. S. S., & Devi, C. N. (2023, February). A Printed Character Recognition System for Meetei-Mayek Script Using Transfer Learning. In International Conference on Science, Technology and Engineering (pp. 519-528). Singapore: Springer Nature Singapore.

[10]Chingakham, N. D., Das, D., & Haobam, M. D. (2019). A Meetei Mayek Basic Characters Recognizer Using Deep Features. In Computational Intelligence, Communications, and Business Analytics: Second International Conference, CICBA 2018, Kalyani, India, July 27–28, 2018, Revised Selected Papers, Part II 2 (pp. 306-315). Springer Singapore.

[11]Laishram, R., Singh, P. B., Singh, T. S. D., Anilkumar, S., & Singh, A. U. (2014, December). A neural network based handwritten Meitei Mayek alphabet optical character recognition system. In 2014 IEEE International Conference on Computational Intelligence and Computing Research (pp. 1-5). IEEE.

[12]Hijam, D., & Saharia, S. (2018). Convolutional neural network based Meitei Mayek handwritten character recognition. In Intelligent Human Computer Interaction: 10th International Conference, IHCI 2018, Allahabad, India, December 7–9, 2018, Proceedings 10 (pp. 207-219). Springer International Publishing.

[13]Chanu, P. R., & Nickson, O. (2022). Manipuri Meitei Mayek Numeral Classification by Using HOG-Assisted Deep Learned Features. In ICT with Intelligent Applications: Proceedings of ICTIS 2021, Volume 1 (pp. 403-411). Springer Singapore.

[14]Khuman, Y. L. K., Devi, H. M., & Singh, K. N. (2018). Segmentation of printed Meitei/Meetei script documents. Digit. Image Process., 10(3), 40-44.

[15]Nath, K., Jelil, S., & Rahul, L. (2014, November). Line, word, and character segmentation of Manipuri machine printed text. In 2014 International Conference on Computational Intelligence and Communication Networks (pp. 203-206). IEEE.

[16]Chanu, P. R. (2022). Manipuri Handwritten Script Recognition Using Machine and Deep Learning. Machine Learning Algorithms for Signal and Image Processing, 129-137.