BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

EEE 312 (July 2023)

Digital Signal Processing I Laboratory

Final Project Report

Section: B2 Group: 03

IMAGE TO TEXT CONVERTER OF BANGLA LETTERS USING MATLAB

Course Instructors:

Dr. Maruf Ahmed, Assistant Professor
Abdullah Jubair Bin Iqbal, Part-Time Lecturer

Signature of Instructor: _____

Academic Honesty Statement

"In signing this statement, We hereby certify that the work on this project is our own and that we have not copied the work of any other students (past or present), and cited all relevant sources while completing this project. We understand that if we fail to honor this agreement, We will each receive a score of ZERO for this project and be subject to failure of this course."	
Signature: _____ **Full Name: MD. KAIDUL ISLAM** **Student ID: 2006111**	**Signature:** _____ **Full Name: SADAF SHAFIN CHOWDHURY** **Student ID: 2006113**
Signature: _____ **Full Name: MD. ATIQ SHAHRIAR SHIAM** **Student ID: 2006114**	**Signature:** _____ **Full Name: SHOHAN AHMED RIFAT** **Student ID: 2006116**

Table of Contents

ABSTRACT

The project "Image to Text Converter of Bangla Letters using MATLAB" addresses the need for efficient conversion of handwritten or printed Bangla text into editable text format. Leveraging image processing techniques and machine learning algorithms, the system enhances the quality of input images, extracts relevant features from Bangla characters, and trains convolutional neural networks (CNNs) to accurately recognize Bangla characters. The developed tool offers a user-friendly interface enabling users to upload images and obtain corresponding textual outputs. The methodology involves image pre-processing to standardize input images, feature extraction to capture shape, texture, and structural elements, and the use of machine learning models for character classification. The project aims to achieve high accuracy and performance, evaluated against benchmark datasets and real-world inputs. References from academic literature and related research provide valuable insights and methodologies for the development of the system. This project contributes to the advancement of optical character recognition (OCR) technology for the Bangla language, offering practical solutions for text digitization and analysis in various domains.

1 INTRODUCTION

The "Image to Text Converter of Bangla Letters using MATLAB" project addresses the challenge of converting handwritten or printed Bangla text from images into editable formats. It utilizes advanced image processing and machine learning techniques tailored for Bangla script recognition. The objective is to enable seamless conversion of Bangla text, regardless of handwriting styles, noise, or variations. The project aims to enhance accessibility to Bangla language content, foster digital inclusion, and streamline text extraction processes. Methodologies include image pre-processing, feature extraction, and training convolutional neural networks (CNNs). A user-friendly interface facilitates easy interaction, while references from academic literature inform the project's development. Ultimately, the project aims to advance OCR technology for Bangla, promoting digital innovation and collaboration in Bangla-speaking communities.

2 LITERATURE REVIEW

The development of an Image to Text Converter for Bangla letters using MATLAB is informed by existing research and advancements in the fields of image processing, machine learning, and optical character recognition (OCR) technology. This literature review explores key studies and methodologies relevant to Bangla character recognition and OCR systems.

1. Bangla Handwritten Character Recognition: A Review" by S. Ahmed, M. Uddin, et al. (International Journal of Scientific & Engineering Research, 2015):

- This comprehensive review provides insights into the challenges and approaches in Bangla handwritten character recognition. It discusses various techniques such as feature extraction, classification algorithms, and dataset availability.

2. Offline Handwritten Bangla Character Recognition Using Convolutional Neural Network" by M. R. Hossain, M. M. Rahman, et al. (2017 4th International Conference on Advances in Electrical Engineering, 2017):

- The paper presents a study on offline handwritten Bangla character recognition using convolutional neural networks (CNNs). It discusses preprocessing steps, feature extraction techniques, and the design of CNN architectures tailored for Bangla characters.

3.Bangla Optical Character Recognition (OCR) Using Deep Learning Based Image Classification Algorithms | IEEE Conference Publication | IEEE Xplore:

- This research explores the application of deep learning-based image classification algorithms for Bangla OCR. It investigates the effectiveness of convolutional neural networks (CNNs) and other deep learning models in accurately recognizing Bangla characters from images.

4. A Complete Bangla Optical Character Recognition System: An Effective Approach | IEEE Conference Publication | IEEE Xplore:

- The paper presents an effective approach to Bangla OCR system development, focusing on preprocessing techniques, feature extraction methods, and machine learning algorithms. It discusses the challenges specific to Bangla script recognition and proposes solutions for improved accuracy and performance.

These studies highlight the importance of robust preprocessing techniques, effective feature extraction methods, and the utilization of machine learning algorithms, particularly CNNs, for accurate Bangla character recognition. They underscore the need for dataset curation, model optimization, and evaluation metrics to assess the performance and reliability of OCR systems.

Informed by the insights from these studies, the proposed project aims to leverage MATLAB's image processing toolbox and machine learning capabilities to develop an efficient Image to Text Converter for Bangla letters. By integrating state-of-the-art techniques and methodologies, the project seeks to contribute to the advancement of OCR technology for the Bangla language, addressing the unique linguistic and cultural nuances of Bangla script recognition and promoting digital accessibility and inclusivity in Bangla-speaking communities.

3 PROBLEM FORMULATION

The "Image to Text Converter of Bangla Letters using MATLAB" project seeks to address the challenge of accurately converting handwritten or printed Bangla text from images into editable text formats. This task presents several key problems:

1. **Recognition of Bangla Characters**: The Bangla script possesses unique characteristics and intricacies that pose challenges for automated recognition. Variations in handwriting styles, diverse character shapes, and the presence of noise in images complicate the accurate identification of Bangla characters.
2. **Image Quality and Pre-processing**: Input images may vary in quality, resolution, and lighting conditions, affecting the performance of character recognition algorithms. Image pre-processing techniques are needed to enhance image quality, reduce noise, and standardize input formats for improved accuracy in character recognition.
3. **Feature Extraction and Representation**: Effective feature extraction methods are essential for capturing relevant characteristics of Bangla characters, such as shape, texture, and structural elements. Identifying discriminative features that distinguish between different characters is crucial for accurate classification.
4. **Machine Learning Model Training**: Developing and training machine learning models, particularly convolutional neural networks (CNNs), requires careful optimization and fine-tuning to achieve high accuracy and robust performance across diverse datasets and real-world scenarios.
5. **User Interface Design**: A user-friendly graphical interface is necessary to facilitate seamless interaction, enabling users to upload images easily and obtain corresponding textual outputs in an intuitive manner. The interface design should prioritize usability, accessibility, and efficient input-output functionality.
6. **Evaluation and Performance Metrics**: Assessing the accuracy, reliability, and performance of the image-to-text conversion system against benchmark datasets and real-world inputs is crucial for validating its effectiveness and identifying areas for improvement.

4 IMAGE PRE-PROCESSING

The proposed system acquires isolated Bangla handwritten characters and digits images from camera or scanner or pre-stored file in real-time. The system converts these acquired Pre-processing techniques are applied to standardize the format and improve the quality of input images. This may involve operations such as:

- **Noise Reduction:** Techniques such as median filtering or Gaussian blurring are used to reduce noise and improve image clarity.
- **Binarization:** Converting the image to binary format, where pixels are either black or white, helps separate foreground (text) from background.
- **Contrast Enhancement:** Adjusting the contrast and brightness levels can enhance the visibility of text and improve overall image quality.
- **Normalization:** Standardizing the size, orientation, and alignment of text regions within the image facilitates more accurate character recognition.

Mathlab Code

```
% Open file selection dialog
[filename, pathname] = uigetfile();

% Check if a file was selected
if ~isequal(filename,0)
```

```
  % Read the image
  img = imread(fullfile(pathname, filename));
else
  disp('No image was selected.');
end
gray = im2gray(img);
size(gray)
gsadjust = imadjust(gray);
H=fspecial("average",3);
gsfilter=imfilter(gsadjust,H,"replicate");
SEdisk = strel("disk",4);
Ibg = imclose(gsfilter,SEdisk);
Ibgsub =  Ibg - gsfilter;
BW=~imbinarize(gray) ;
BW=BW.*255;
img = imresize(BW, [28, 28]);
% imshow(BW);
save('Processed_img.mat','img');
```

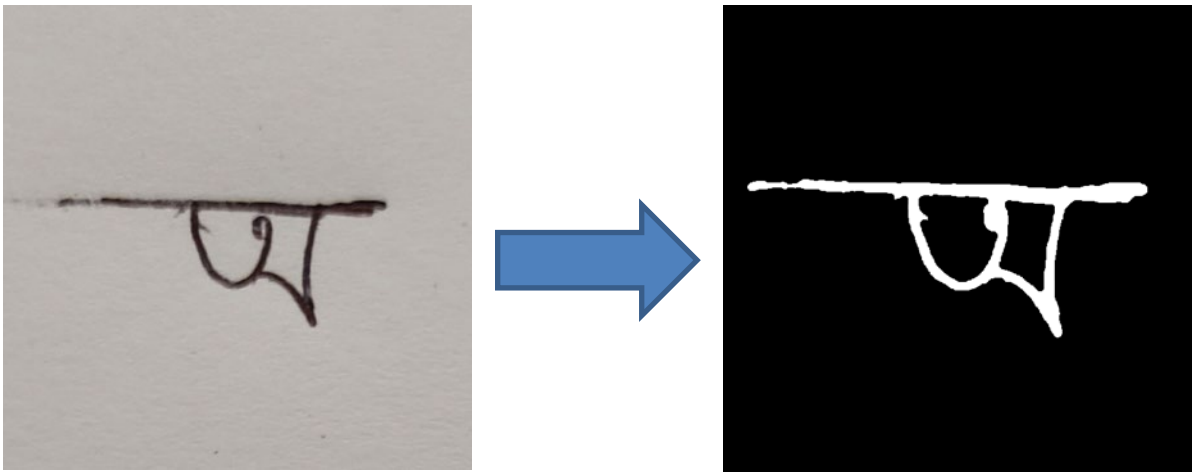


Figure : Image Pre-processing

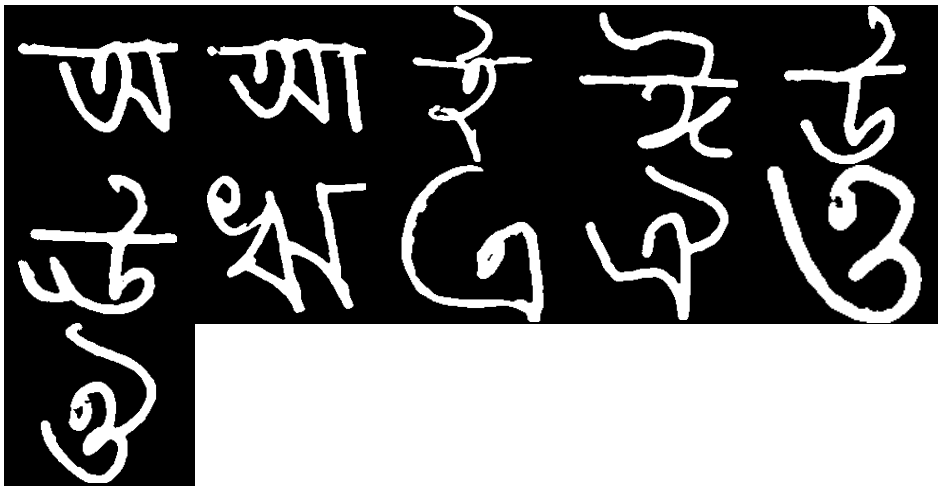
5 DATA PREPARATION

The dataset "Bangla Lekha isolated" comprises handwritten Bangla numerals, basic characters, and compound characters. It was collected from diverse geographical locations across Bangladesh and encompasses samples from various age groups. The dataset is rich in variability, reflecting the diversity present within the Bangla writing system.

The "Bangla Lekha isolated" dataset offers a valuable resource for researchers and practitioners interested in Bangla handwriting analysis, character recognition, and related classification tasks. Its diversity and breadth make it suitable for a wide range of applications, contributing to advancements in Bangla language processing and understanding handwritten documents in Bangla script.

Table 1. Dataset formations, number of classes, number of training, validation, and testing samples, and whether dataset is balanced or imbalanced.

formations	number of classes	number of training	validation	testing samples	balanced
'png'	*50*	*40000*	*10000*	*5000*	*Yes*



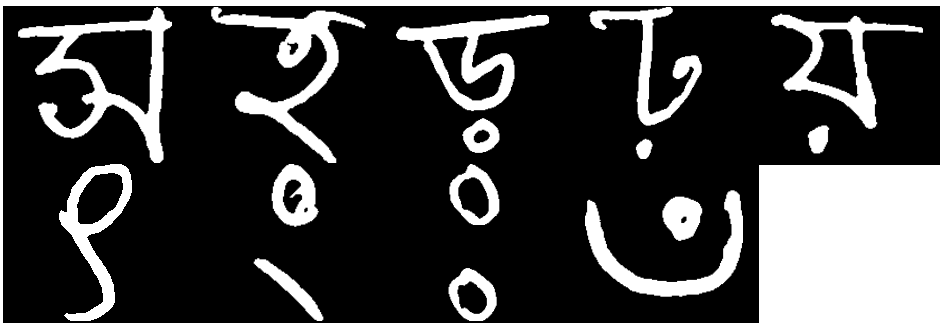


Figure : Bengali Basic Characters Samples from Bangla Lekha isolated Dataset.

6 NEURAL NETWORKS

The neural network for Optical Character Recognition (OCR) is a deep learning model designed to recognize and interpret characters from images. It comprises multiple layers of interconnected neurons that process pixel values from input images and learn to extract features relevant to character recognition.

The neural network architecture typically includes convolutional layers, which perform feature extraction by applying filters to input images to detect patterns such as edges, corners, and textures. Pooling layers are used to reduce spatial dimensions and extract key features, while fully connected layers learn higher-level representations of the input data.

During training, the neural network learns to map input images to corresponding character labels through a process called backpropagation, where errors are propagated backward through the network, and model parameters are adjusted to minimize the difference between predicted and actual outputs.

Once trained, the OCR neural network can accurately classify and recognize characters from new images, enabling automated conversion of scanned documents, handwritten text, or printed text into editable digital formats. This technology finds applications in document digitization, text analysis, and data extraction across various industries. The **convolutional neural network**, also recognized as **CNN**, is one of the most common forms of deep neural networks which is generally exercised for interpreting optical images of various research domains.

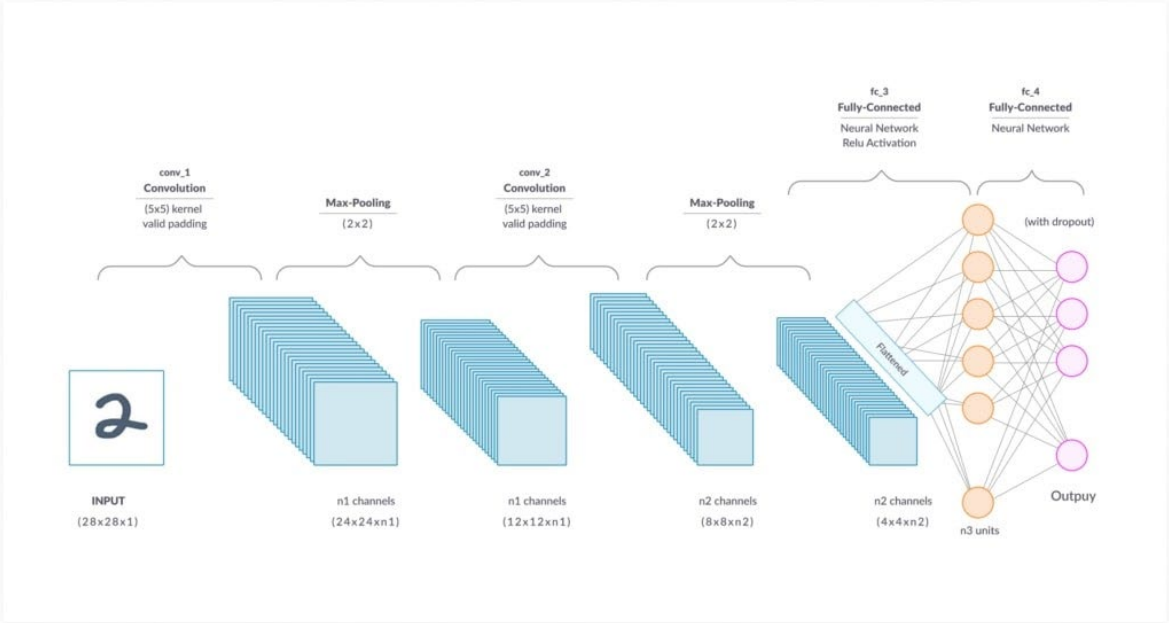
6.1 Deep Convolutional Neural Network

Deep learning has become a staple technology in the field of computer vision, with convolutional neural networks (CNNs) emerging as a prevalent architecture for processing optical images across diverse research domains. CNNs are extensively utilized for tasks such as object or person identification in images and videos, image classification, recommender systems, natural language processing, medical image processing, and financial analysis.

Inspired by the interconnectedness observed among neurons in the animal cortex, CNNs are specifically designed for image recognition and identification in various research domains. Their ability to learn hierarchical representations of visual features makes them particularly well-suited for tasks requiring complex pattern recognition and analysis in large datasets.

Through the application of convolutional layers, pooling layers, and fully connected layers, CNNs excel at extracting relevant features from input images and making accurate predictions

or classifications. Their versatility and effectiveness have propelled CNNs to the forefront of deep learning methodologies, driving advancements in computer vision and enabling innovative applications across numerous domains.



CNN Architecture: Types of Layers

Convolutional Neural Networks (CNNs) typically consist of several types of layers, each serving a specific function in the network architecture. The main types of layers in a CNN include:

1. **Convolutional Layer**: This layer applies convolutional filters to the input data, extracting features through spatial detection across the input volume. Convolutional layers are responsible for learning patterns and features within the input images.
2. **Pooling Layer**: Also known as downsampling or subsampling layers, pooling layers reduce the spatial dimensions of the feature maps generated by the convolutional layers. Common pooling operations include max pooling and average pooling, which help in reducing computational complexity and controlling overfitting.
3. **Fully Connected Layer**: Also referred to as dense layers, fully connected layers connect every neuron in one layer to every neuron in the next layer. These layers are typically found towards the end of the CNN architecture and help in making predictions based on the learned features.

Apart from these main types, CNN architectures may also incorporate other types of layers such as:

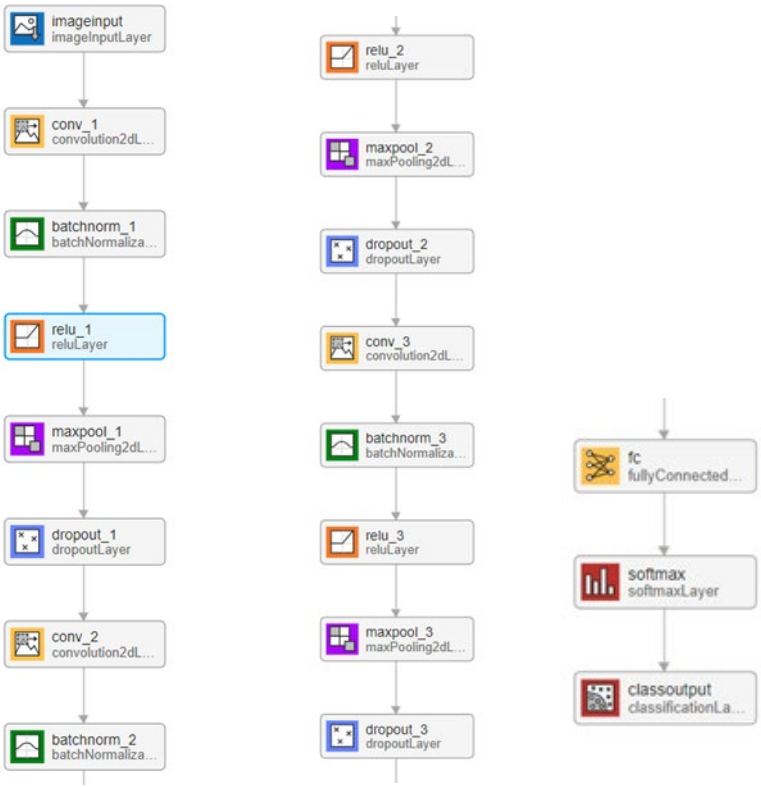
4. **Activation Layer**: Activation layers introduce non-linearity to the network by applying an activation function to the output of the previous layer. Common activation functions include ReLU (Rectified Linear Unit), sigmoid, and tanh.
5. **Normalization Layer**: Normalization layers normalize the activations of the previous layer to ensure stable training and improved convergence.
6. **Dropout Layer**: Dropout layers randomly drop a percentage of neurons during training, helping prevent overfitting by encouraging robust feature learning.
7. **Batch Normalization Layer**: Batch normalization layers normalize the activations of the previous layer over mini-batches, further stabilizing training and improving convergence.

These layers work together to learn hierarchical representations of input data, extract relevant features, and make predictions or classifications for tasks such as image recognition, object

detection, and image segmentation.

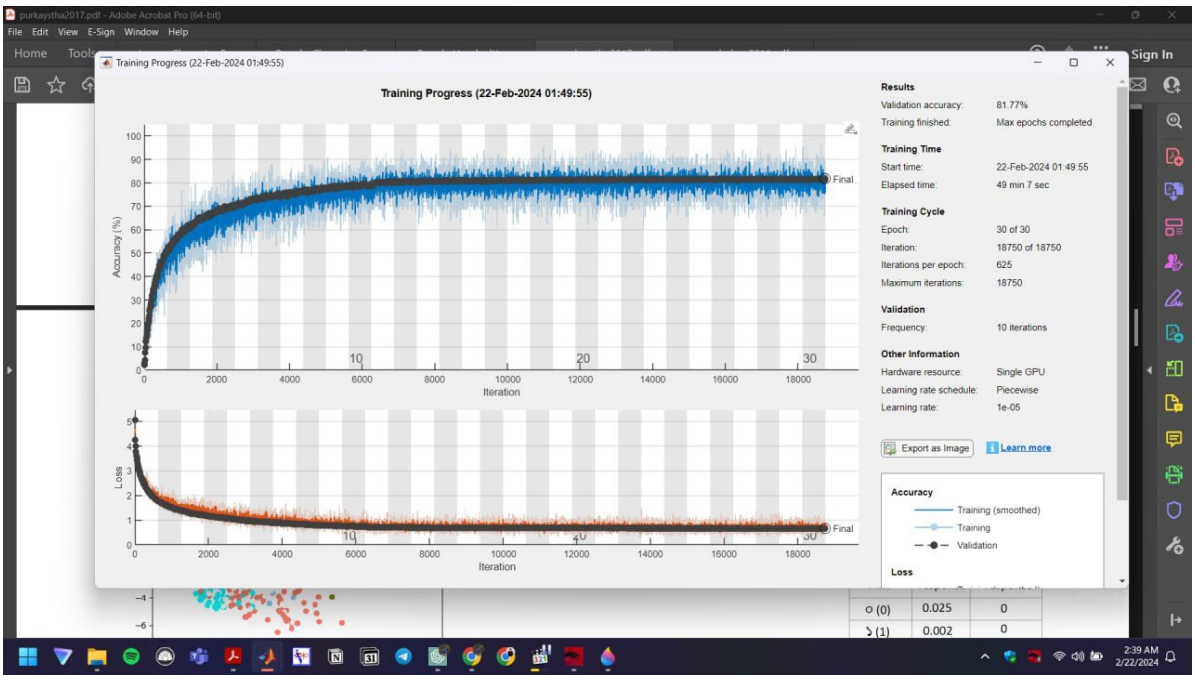
6.2 Proposed Adopted CNN Architecture

The following is a MATLAB generated diagram for our CNN.



6.3 Training A Network

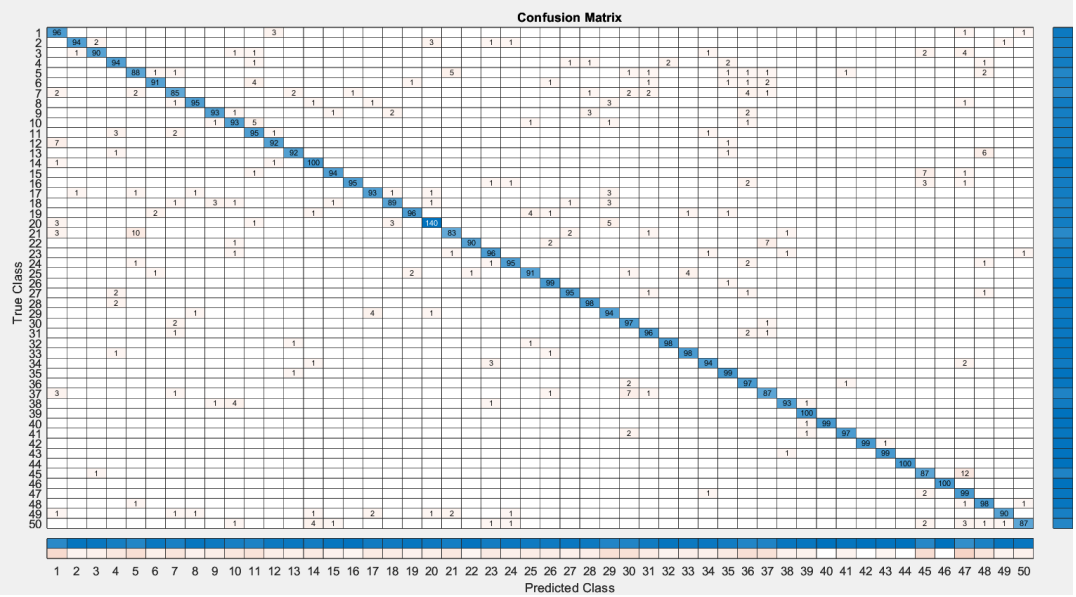
We trained the network multiple times using different epochs ,learning rates and batch sizes until We got the satisfactory results.



6.4 Testing the Network

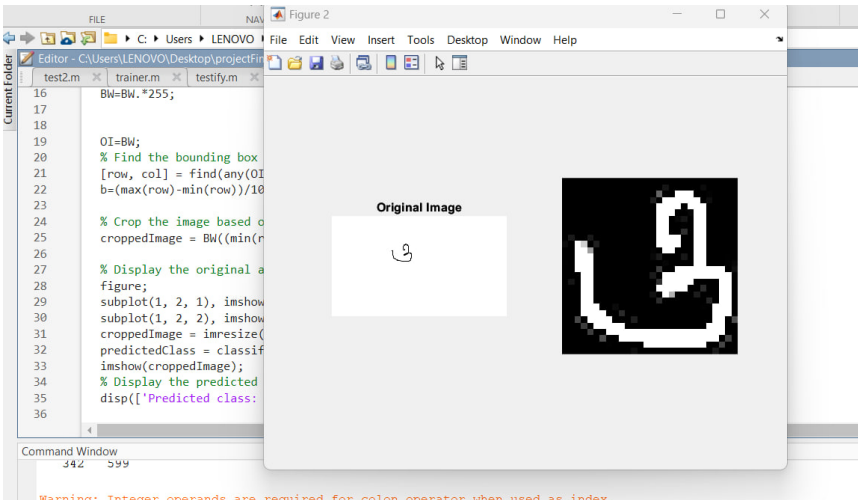
Now it was time to see how our model performs on unseen data. We used a dataset of 100 images per class to testify and the results were satisfactory. The accuracy rate was 93%

Here is the confusion chart for our testing. It shows that for letters with similar shapes our model showed error.



7 IMAGE PREPROCESSING REVISITED

Now with our model trained and tested we wanted to use the model for identifying handwritten letter Given as input from user. And the input can be of very different sizes and shapes. I may be tilted or it may have noise. So we tried to develop a more generalized image preprocessing algorithm.



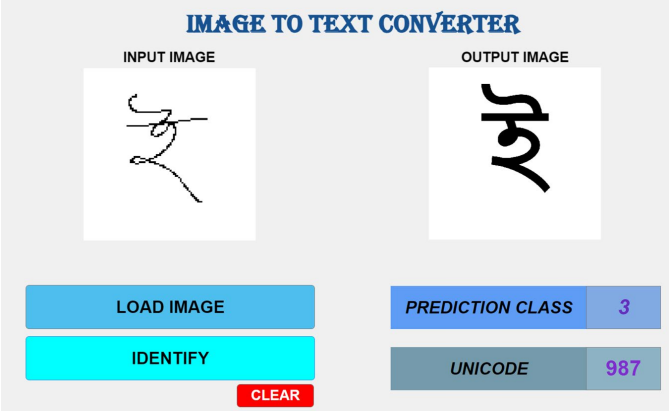
8 UNICODE IMPLEMENTATION

We also integrated the UNICODE for each letter so that it may be used for use across any platform

9 DEVELOPING A GUI

We focused on making the process user friendly. So we developed a GUI.
Our GUI contains the following components:

1. A button to load the images
2. A Button to start the Identification
3. Pop up to the class of the input image via a standard image
4. UNICODE of the letter.



NOVELTY

The novelty of the "Image to Text Converter of Bangla Letters using MATLAB" project lies in its focus on addressing the specific challenges and requirements of converting handwritten or printed Bangla text into editable text formats using advanced image processing and machine learning techniques. Several aspects contribute to the novelty of this project:

1. **Bangla Script Recognition**: The project specifically targets the recognition and interpretation of the Bangla script, which possesses unique characteristics and intricacies compared to other languages. By focusing on Bangla text conversion, the project addresses the need for tailored solutions to support digitalization efforts and accessibility in Bangla-speaking communities.
2. **Integration of MATLAB**: Leveraging MATLAB's powerful image processing toolbox and machine learning capabilities, the project offers a comprehensive and

integrated platform for developing and deploying the image-to-text conversion system. MATLAB provides a versatile environment for prototyping, testing, and optimizing algorithms, facilitating rapid development and deployment of the solution.

3. **Diverse Dataset Collection**: The project utilizes a dataset containing handwritten Bangla numerals, basic characters, and compound characters collected from diverse geographical locations and demographic groups within Bangladesh. This dataset diversity enhances the robustness and generalization capabilities of the recognition system, accommodating variations in handwriting styles and regional preferences.
4. **Customized Image Pre-processing Pipeline**: The project implements an efficient image pre-processing pipeline tailored to enhance the quality and standardize the format of input images containing Bangla text. By incorporating noise reduction, binarization, and normalization techniques, the pipeline prepares the images for accurate character recognition and classification.
5. **Machine Learning Model Optimization**: Through the training and fine-tuning of machine learning models, such as convolutional neural networks (CNNs), the project aims to achieve high accuracy and robust performance in recognizing Bangla characters. Model optimization techniques, including data augmentation, transfer learning, and hyperparameter tuning, are employed to enhance the model's ability to generalize and adapt to diverse inputs.
6. **User-friendly Interface**: The development of a user-friendly graphical interface enables seamless interaction with the image-to-text conversion system, allowing users to upload images and obtain corresponding textual outputs in an intuitive manner. The interface design prioritizes usability, accessibility, and efficient input-output functionality to cater to diverse user needs and preferences.

Overall, the project's novelty lies in its holistic approach to addressing the specific challenges of Bangla text conversion, leveraging advanced technologies, and methodologies to develop a robust, accurate, and user-friendly solution for digitalizing Bangla text content. By advancing the state-of-the-art in Bangla script recognition and optical character recognition (OCR) technology, the project contributes to enhancing digital accessibility, language processing, and information retrieval capabilities in Bangla-speaking communities.

LIMITATIONS OF TOOLS

The limitations of hardware, dataset availability, and Unicode support in MATLAB present significant challenges for the project. These limitations impact various aspects of the project's development and performance:

1. **Hardware Limitations**: Limited computational resources restrict the size and complexity of neural network architectures that can be trained effectively. This constraint can hinder the exploration of larger models or more extensive training regimes, potentially limiting the accuracy and robustness of the character recognition system.
2. **Dataset Constraints**: The availability and diversity of datasets containing handwritten or printed Bangla text are limited, particularly in terms of size, quality, and representation of different writing styles and variations.
3. **Unicode Support in MATLAB**: MATLAB's support for Unicode characters, particularly those specific to the Bangla script, are not available for handling text processing and manipulation tasks effectively.

Addressing these limitations require exploring alternative approaches, tools, and strategies.

FUTURE WORK

The lays the foundation for several potential avenues of future work and enhancements. Some key areas for future exploration and development include:

1. **Enhanced Image Pre-processing Techniques**: Investigate advanced image pre-processing techniques to further improve the quality and clarity of input images. This could involve exploring deep learning-based approaches for noise reduction, contrast enhancement, and text localization to handle diverse image conditions more effectively.
2. **Integration of Advanced Feature Extraction Methods**: Explore advanced feature extraction methods, including deep learning-based feature representations, to capture more discriminative characteristics of Bangla characters. Investigate the use of convolutional neural networks (CNNs) for end-to-end feature extraction and classification, potentially improving recognition accuracy and robustness.
3. **Expansion of Character Classes and Dataset Diversity**: Expand the dataset to include a broader range of Bangla characters, including ligatures, diacritics, and less common characters. Collect samples from additional geographical regions and diverse demographic groups to improve dataset diversity and generalization capabilities of the recognition system.
4. **Fine-tuning and Optimization of Machine Learning Models**: Experiment with different architectures and optimization strategies for machine learning models, such as CNNs, to further enhance recognition performance. Explore techniques such as transfer learning and data augmentation to improve model generalization and adaptability to new datasets and scenarios.
5. **Multi-language Support and Cross-script Recognition**: Extend the system to support recognition of multiple languages and scripts, facilitating multilingual text conversion and enabling cross-script recognition. Investigate techniques for handling script variations, transliteration, and language-specific nuances to develop a more versatile and inclusive text conversion tool.
6. **User Interface Enhancements and Accessibility Features**: Incorporate user feedback and usability studies to enhance the user interface, making it more intuitive, accessible, and user-friendly. Implement features such as real-time feedback, error correction mechanisms, and support for alternative input methods to improve user experience and accommodate diverse user needs.
7. **Integration with Text Processing and Language Understanding Tools**: Explore integration with text processing and natural language understanding tools to enable advanced text analysis, language translation, and semantic understanding capabilities. Integrate with existing language processing frameworks and APIs to extend the functionality of the system and support diverse applications and use cases.
8. **Evaluation and Benchmarking Against Real-world Data**: Conduct rigorous evaluation and benchmarking of the system against real-world data and scenarios to assess its performance, accuracy, and robustness. Collaborate with domain experts and end-users to validate the system's effectiveness and identify areas for improvement and refinement.

By exploring these avenues of future work and incorporating advancements in image processing, machine learning, and user interface design, the "Image to Text Converter of Bangla Letters using MATLAB" project can continue to evolve and provide valuable solutions for text digitization, language processing, and information retrieval in Bangla-speaking communities and beyond.

REFERENCES

1. Ahmed, S., Uddin, M., et al. "Bangla Handwritten Character Recognition: A Review." International Journal of Scientific & Engineering Research, 2015.
2. Hossain, M. R., Rahman, M. M., et al. "Offline Handwritten Bangla Character Recognition Using Convolutional Neural Network." 2017 4th International Conference on Advances in Electrical Engineering, 2017.
3. IEEE Xplore. "Bangla Optical Character Recognition (OCR) Using Deep Learning Based Image Classification Algorithms." IEEE Conference Publication.
4. IEEE Xplore. "A Complete Bangla Optical Character Recognition System: An Effective Approach." IEEE Conference Publication.
5. Alif, M. A. R., Ahmed, S., & Hasan, M. A. (December 2017). Isolated Bangla handwritten character recognition with convolutional neural network.
6. https://www.mdpi.com/2076-3417/11/15/6845#:~:text=Proposed%20BengaliNet%20Architecture,the%20name%20of%20the%20architecture.
7. Fully Connected Layers in Convolutional Neural Networks – IndianTechWarrior
8. https://paperswithcode.com/dataset/banglalekha-isolated

These references provide valuable insights and methodologies for the development of the project, informing its approach and contributing to advancements in Bangla script recognition and optical character recognition (OCR) technology.