2023 26th International Conference on Computer and Information Technology (ICCIT)
13-15 December, Cox's Bazar, Bangladesh

2023 26th International Conference on Computer and Information Technology (ICCIT) | 979-8-3503-5901-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCIT60459.2023.10441040

Bengali License Plate Recognition: Unveiling Clarity with CNN and GFP-GAN

Noushin Afrin, Md Mahamudul Hasan, Mohammed Fazlay Elahi Safin, Khondakar Rifat Amin
Md Zahidul Haque, Farzad Ahmed, and Md. Tanvir Rouf Shawon
Department of Computer Science and Engineering,Ahsanullah University of Science and Technology, Dhaka, Bangladesh
noushinmaifa@gmail.com, mhtusar120725@gmail.com, fazlayelahisafin@gmail.com, khrifatamin@gmail.com,
jahidprototype2@gmail.com, farzad.cse@aust.edu, shawontanvir95@gmail.com

***Abstract*—Automated License Plate Recognition(ALPR) is a system that automatically reads and extracts data from vehicle license plates using image processing and computer vision techniques. The Goal of LPR is to identify and read the license plate number accurately and quickly, even under challenging, conditions such as poor lighting, angled or obscured plates, and different plate fonts and layouts. The proposed method consists of processing the Bengali low-resolution blurred license plates and identifying the plate's characters. The processes include image restoration using GFPGAN, Maximizing contrast, Morphological image processing like dilation, feature extraction and Using Convolutional Neural Networks (CNN), character segmentation and recognition are accomplished. A dataset of 1292 images of Bengali digits and characters was prepared for this project.**

***Index Terms*—ALPR, GFPGAN, CNN, ResNet-50**

I. INTRODUCTION

Automated License Plate Recognition (ALPR) technology has revolutionized the field of vehicle identification by automating the extraction of information from license plates. Through cutting-edge computer vision techniques, ALPR systems capture license plate photos, segment characters, and accurately identify them. In our study, we place primary emphasis on developing an innovative ALPR system that harnesses the power of deep learning and image processing techniques, specifically optimized for Bengali license plates. Number plates consist of digits, registered letters and city names. The digits serve as unique identifiers for each vehicle, while the registered letters indicate the vehicle type. The city name on the license plate provides information about the vehicle's registration location. Accurately identifying and interpreting these components allows us to extract specific details and understand the vehicle's identification, type, and origin.

However, identifying license plates from poor-quality images presents significant challenges. Issues such as poor lighting, diverse plate fonts, and variations in sizes and orientations complicate the precise segmentation and recognition of characters.

Bengali symbols and numerals are a little more complicated than English characters and digits, making it difficult to recognize number plates in Bengali. There have been prior studies on Bengali license plate recognition [1], [2] and digit recognition [3], [4], but none have taken into account the low image quality that is prevalent for street cameras, which are the primary source of license plate images.

Our research's main goal is to build advanced methods that can accurately recognize and evaluate the components of Bengali license plates, even in low-quality photographs. In addition, we utilize GFPGAN for image enhancement and leverage the state-of-the-art CNN and ResNet-50 models for the task of recognizing the digits and characters of the number plates. By showcasing the capabilities of deep learning and image processing, our research aims to pave the way for improved ALPR systems. These advancements will have widespread applications, benefiting numerous industries that rely on accurate vehicle identification. The results of our project will contribute to the development of ALPR systems with enhanced efficiency and precision, enabling more reliable and effective vehicle identification. The contribution of the paper can be summarized into the following points:

- Modern deep learning methods are used by us to perform Bengali license plate recognition in challenging environments, including Convolutional Neural Networks (CNNs) and pre-trained models like ResNet-50.
- We incorporate GFPGAN for image enhancement, further improving the quality of input images and enhancing recognition performance.
- A dataset of 1292 images of Bengali digits and characters was prepared for this project.
- We scored an excellent 88% average accuracy in accurately identifying the 30 various number plates tested by the SequenceMatcher class of difflib[1] library. The codes and dataset can be found at - https://github.com/mhtusher131/Bengali-License-Plate-Recognition__Unveling-Clarity-with-CNN-and-GFP-GAN?fbclid=IwAR0PLgVXkW3lJ8huKg__5vqDtK4qPOMv2u-I-iegGoibDJzvhNpxpPjTObJ4

[1] https://docs.python.org/3/library/difflib.html

979-8-3503-5901-5/23/$31.00 ©2023 IEEE

II. LITERATURE REVIEW

This literature review provides an overview of the advancements made in detecting number plates in Bengali as well as in other languages from poor-quality images.

Study on License Plate Recognition Using CNNs: These three papers [5]–[7] focus on license plate recognition systems that use convolutional neural networks (CNNs) and other computer vision techniques. Munna et al. [5] describes a systemfor utilizing a CNN to find and identify license plate, with a specific focus on Bengali characters. Suvon et al. [6] propose a real-time system for recognizing Bengali license plates using a combination of computer vision techniques and a CNN. Dhar et al. [7] present a method that uses edges to recognize license plates detection and a CNN for feature extraction and character recognition. These papers collectively demonstrate the effectiveness of using CNNs for license plate recognition, which has important applications in traffic control, surveillance, and tracking of stolen vehicles.

Study on Automatic License Plate Recognition for Vehicle Authentication in Bengali: The study by Nooruddin et al. [1] and Islam et al. [2], focuses on Bengali Automatic License Plate Detection and Recognition for Vehicle Authentication. The paper [1] by Nooruddin et al. uses a Random Forest model for the area-based detection and identification of license plates aspect ratio criteria. The paper [2] by Islam et al. introduces a unique four-step system including preprocessing, license plate processing, character recognition, and vehicle registration determination. This approach achieves a 96.1% character recognition score using techniques like character segmentation, edge detection, and template matching.

Study on Automatic License Plate Recognition for Vehicle Authentication in other languages: This study [8] by Kathirvel et al. presents a systematic number plate recognition (SNPR) methodology using YOLOv5s for automating the fine collection and enforcing speed limits on Indian roads. The proposed system achieves 98.2% accuracy on number plate identification task. The study [9] by Antar R et al. explores automatic license plate recognition for Saudi car plates. The experiment used 50 images, employing canny edge detection, threshold techniques, and horizontal projection for plate segmentation. OCR was applied to read English and Arabic characters separately,obtaining an accuracy of 96% for English and 92.4% for Arabic. Raj R et al. [10] implements an ANPR system using YOLOv5 for number plate detection. The paper [11] by kabiraj et al. [2023] offers a method for enhancing number plate detection and identification accuracy using Enhanced-Super-Resolution with Generative Adversarial Network (ESRGAN).

III. BACKGROUND STUDIES

GFPGAN: Generative Flow-Based Probabilistic Graphical Model [12] is a type of generative model that uses a probabilistic graphical model to generate data from a given set of input variables. It is based on the idea of using flow-based generative models to learn the underlying structure of data and then generate new data from it. GFPGAN has been used in various applications such as image generation, text generation, and music generation. It has also been used for anomaly detection and forecasting tasks. GFPGAN is an efficient and powerful tool for generating high-quality data from complex distributions.

ResNet-50: The motivation for developing deep neural networks stemmed from challenges in training deep networks, including the degradation of performance and vanishing gradients. ResNet, a significant architecture, emerged to tackle these issues, with ResNet-50 [13] as a specific variant featuring 50 layers and shortcut connections that enable direct information flow, addressing vanishing gradients and aiding deep network training while using residual blocks to capture intricate features effectively.

Performance Metrics: Recall is a measurement of the proportion of positive cases that the classifier properly predicted out of all the positive cases in the data, whereas precision is measured by the percentage of correctly predicted positive outcomes (TP). The harmonic mean of recall and precision is F1.

IV. DATASET

We have created the dataset for Bengali License Plate Detection from Poor Quality Images. The dataset creation process involved several steps to collect and organize the Bengali number plate images:

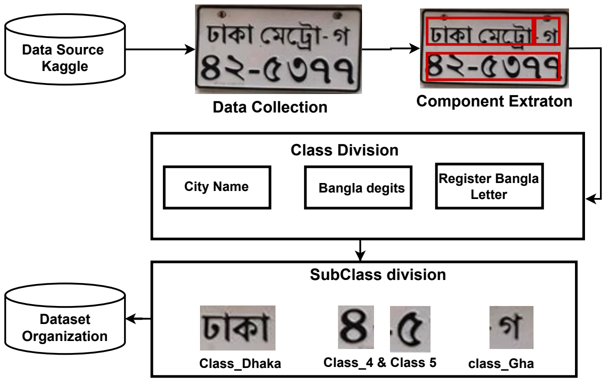

Figure 1: Navigating Dataset Development Pipelines-Strategies and Insights

Data Collection: A diverse collection of Bengali number plate images was gathered from various sources[2]. The images were collected to encompass a wide range of vehicle types, number plate designs, and environmental conditions[3].

[2] https://www.kaggle.com/furcifer/bangla-license-plates-synthetic
[3] https://github.com/mmshaifur/BLRPS

Component Extraction: From each collected image, the specific components of interest were extracted. These components included digits, registered letters, city names, and metro names. Manual image cropping techniques were employed to precisely isolate these components, removing any unnecessary background information.

Class Division: The extracted components were categorized into three main classes: digits, registered letters, and city names. This allowed for separate analysis and classification of each component type.

Subclass Division: To capture variations within each class, subclass division was implemented. This resulted in 10 subclasses for digits, 4 subclasses for registered letters, and 3 subclasses for city names. The subclasses were designed to represent specific patterns and variations found in the dataset.

Dataset Organization: The final dataset was organized into subclasses. The images were standardized to a size of 32x32 pixels with RGB color channels, ensuring consistency across the dataset.

A. Dataset Statistics

Table I presents the statistical information of the dataset. The dataset consists of three main classes: digits, registered letters, and city names. Each class is further divided into subclasses, as specified. The number of images per subclass is satisfactory, ensuring sufficient data for model training and evaluation. In our dataset the city names are ”Dhaka(ঢাকা)”,” Dhaka Metro (ঢাকা মেট্রো))”, ”Narayanganj(নারায়ণগঞ্জ)”, .The registered letters are Ka(ক), ga(গ),Jha(ঝ),and La (ল).The digits are from 0(০) to 9(৯). Bangladesh employs a diverse range of cities and registered letters on its number plates, reflecting the extensive coverage of these elements in its transportation system. These selections serve as the foundation for constructing a model that recognizes Bangladeshi number plates in low-quality images. Furthermore, the dataset's flexibility allows for potential expansions of additional cities and letters as required for the project's goals. Fig. 2 depicts some samples from our dataset.

Table I: Class distribution among train, validation and test set.

Class	Subclass type	Train	valid-ation	Test
Digits	0 (০)	90	35	22
	1 (১)	140	35	27
	2 (২)	130	35	23
	3 (৩)	130	35	22
	4 (৪)	100	35	23
	5 (৫)	140	35	24
	6 (৬)	90	35	26
	7 (৭)	120	35	20
	8 (৮)	110	35	21
	9 (৯)	1	45	21
registered	Ka (ক)	13	7	7
	ga (গ)	15	9	9
	jha (ঝ)	30	13	14
	La (ল)	18	12	12
City	Dhaka(ঢাকা)	18	11	12
	Narayanganj(নারায়ণগঞ্জ)	30	15	10
	Metro(মেট্রো)	18	12	10

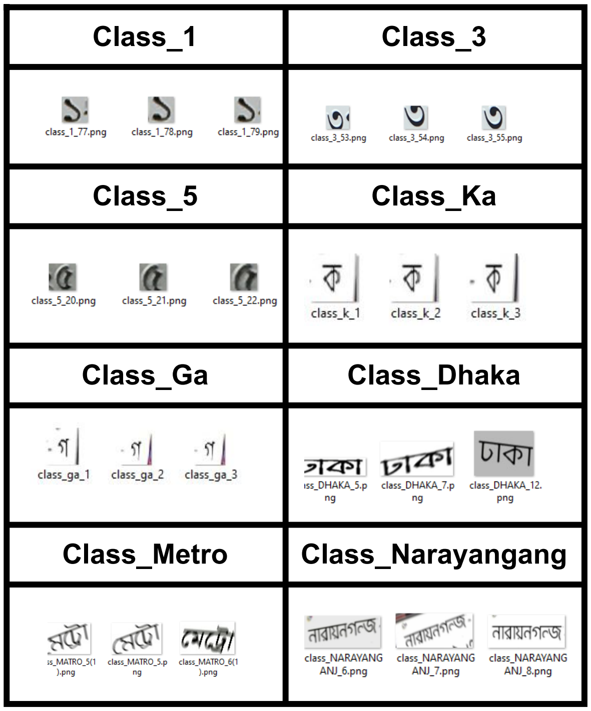

Figure 2: Illustration of the samples from our dataset.

V. METHODOLOGY

This section describes our methodology in brief. Fig. 3 shows the workflow of our methodology. The ten individual steps of our methodology are described elaborately.

1) Read the image: This is the initial phase of our suggested procedure. Python at this point captures and reads the car picture. The image is blurry and of poor quality, making it difficult to read without altering how the read image or test image is depicted.

2) Resize Image: We must resize the image to a predetermined size because the acquired image may be of any size.Using the provided method, the height of the car image is converted into 32*32 pixels.

3) GFPGAN: GFPGAN [12](Generative Focused Pixel-wise GAN). It is a type of generative model that can be used to generate new images from existing ones. GFPGAN uses a combination of convolutional neural networks (CNNs) and Using generative adversarial networks (GANs), high-quality images can be produced from inputs with low resolution..GFPGAN has been used in low-resolution blur images to increase the quality of the image by reducing the blurriness.

4) RGB to greyscale picture conversion: To be more precise,The image's red, green, and blue parts are separated and then combined to form a single grey image.

5) Grey-scale picture binarization: Using the Otsu method, the grey image is next transformed into a binary image. In this method, the conversion procedure is started after determining a threshold value.

6) Maximising Contrast: Maximising the contrast of an image is the process of increasing the difference between the lightest and darkest areas of an image. It is done by adjusting the brightness, contrast, and gamma levels of the image.

7) Morphological Operation: Dilation has been used to fill up the noise within the objects of the image.

8) Extracting the License Plate Region: In the suggested procedure, one of the key steps is the extraction of the license plate region. a) Multiplication of the images: To get an output that only contains the license plate region, we now multiply the dilated image by the grayscale image. b) Match contours: A technique for identifying the edges of numbers and letters in an image is called match contours. In order to do this, a set of predetermined shapes is matched up with the edges of the image's objects. This method is being used to classify, identify forms, and identify the items. The image's potential dimensions are obtained, the binary image's counter is found, and the coordinates of the enclosed rectangle are then provided. The contour's dimensions are then examined in order to separate the characters based on their size. After extracting each character using the coordinates of the surrounding rectangle, the output was prepared for categorization.

9) Character Segmentation: A license plate from Bangladesh contains two lines of text. The first line consists of a Bengali letter and one or two Bengali words. The second line contains six numbers as required by the BRTA's Retro-Reflective standard. Separations exist between the six digits. However, "Matra" is used to connect the alphabetic characters in a word (a line joins several characters). It is believed that the entire word "Matra" consists of one connected component. Here, linked component analysis is utilized to segment the characters. Following the binarization of the recovered license plate image, blue boundary boxes are created for each of the associated items in the shown image. After that, the characters are taken out of the plate and stored as a separate image. The recognition module receives the recently segmented characters and words.

10) Recognition of the characters: At this point, it should be possible to identify the characters that have already been split and saved as separate images. The CNN model is employed in the suggested method. The CNN model is trained on a sizable dataset of tagged images before being used to recognize segmented images utilizing datasets. The model was then able to learn the characteristics of each image and categorize them. With 1292 photos of Bengali characters and numbers, we trained the model. The model can then be used to categorize fresh photos after being trained. Images that have been segmented have been highly correctly sorted into their respective groups using CNN models.

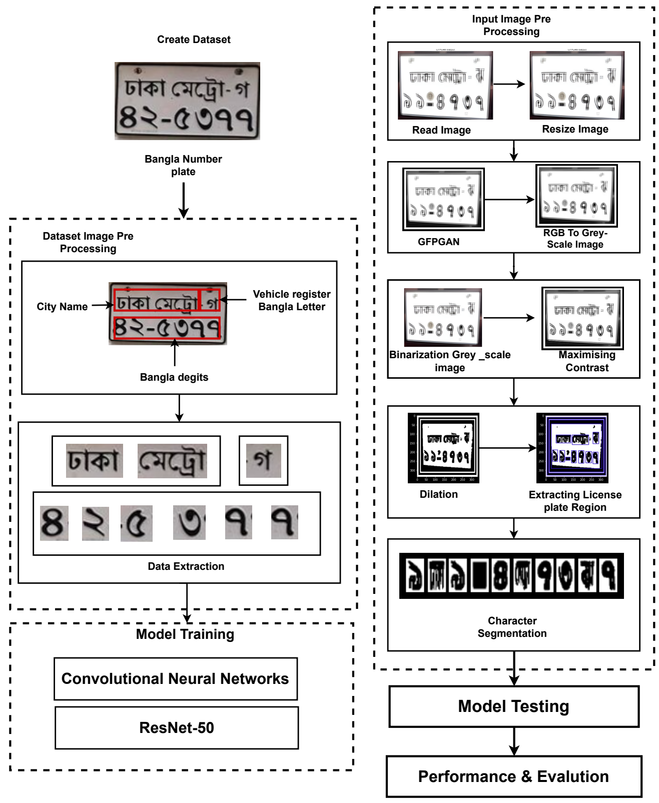

Figure 3: A visual guide to the proposed approach.

VI. RESULTS AND EXPERIMENTS

A. *Hyperparameter Settings*

In our instance, we used images with a size of $32, 32, 3$ to train the CNN and ResNet50 models. In both models 1292 images were used for training, 429 images were used for validation and 303 images were used to conduct tests. The learning rate was 0.0001 and the entire batch length as 1.

B. *Performance of the Proposed Models*

The general effectiveness of CNN and ResNet50 is given in Table II. We presented the confusion matrix of our best-performing CNN architecture in Fig. 4.

Table II: Performance of Different Architectures (CNN and ResNet50) on the Test Dataset

Architecture	Accuracy	Precision	Recall	F1-score
CNN	92.38	0.9667	0.9162	0.9476
ResNet50	87.50	0.9343	0.8827	0.8984

CNN: Table II shows that the recall is projected to be 0.9162 and the accuracy is assessed to be 92.38 percent for CNN. During the CNN training process, we achieved a validation accuracy of 92%. The model was trained over the course of 40 epochs, with the validation accuracy measured at 98.80% and the validation loss measured at 0.0303.The model was not developed further because overfitting was a possibility. The precision number, which can be seen as 0.9667, indicates false positive situations, but the recall value, which can be seen as 0.9162, reveals CNN's limitations in the case of false negative cases. Fig. 4 shows the confusion matrix of CNN for the testing samples.

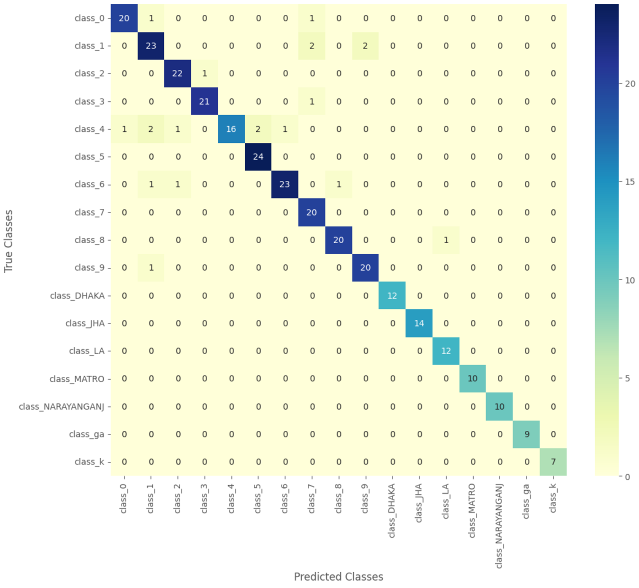

Figure 4: Confusion Matrix of the proposed CNN architecture.

ResNet50: The ResNet50 model demonstrates strong overall performance with an accuracy of 87.50 %, indicating its ability to correctly classify the majority of samples in the dataset. To train the model 30 epochs were used. The high precision value of 0.9343 signifies that the model is cautious in predicting positive instances, resulting in a low number of false positives. Moreover, the recall score of 0.8827 suggests that the model effectively captures a significant portion of the actual positive samples, reducing false negatives. The F1-Score of 0.8984 further validates the model's balanced performance between precision and recall. In summary, ResNet50 exhibits notable accuracy, precision, and recall, making it a reliable choice for the given classification task.

Nevertheless, there is a deficiency in ResNet50's ability to accurately detect Bengali characters on number plates. Figure 5 illustrates instances of incorrect predictions made by ResNet50. While the model is capable of accurately predicting Bengali numerals, it occasionally struggles to properly identify certain Bengali characters on the plates, specifically those used for city names and registration letters

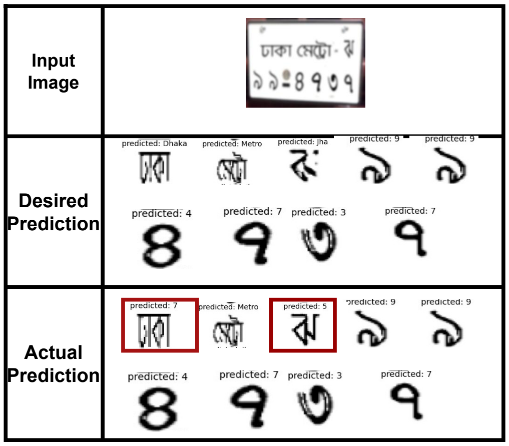

Figure 5: Unveiling Model Challenges- Exploring an Incorrect Prediction by ResNet50

Performance of our methodology: We utilized the SequenceMatcher class of difflib library which compares pairs of sequences of any type to evaluate our 30 testing samples (number plates), achieving an 88% accuracy on average.Some examples of SequenceMatcher performance are given in Table III. While the CNN-based predictions for the plates were generally robust, occasional plate noise led to the inclusion of some distorted images in our model's output (Fig. 6). Consequently, slight alterations in the desired output sequences occurred. As a result, the overall accuracy was somewhat affected.

Table III: Sample Data - Desired vs. Generated Output with Corresponding Accuracy

Desired Output	Generated Output	Accuracy
9dhaka94matro773jha7	9Dhaka984Matro773Jha7	0.95
2dhaka2matro404ka4	2dhaka2matro404ka4	1
5dhaka47818jha	553Dhaka47898Jha	0.86
5dhaka2315ka6	5Dhaka231526	0.84

C. *Experimental Analysis*

The accuracy of 88% shows that there is a fair amount of similarity between the machine-generated translation and the reference translations. This accuracy is calculated by taking the average of all the 30 samples (number plate) that were taken to evaluate the model. Table III shows some sample scores achieved by SequenceMatcher Library. Experimenting with various threshold values yielded diverse results. When employing a stringent 95% threshold, approximately 57% (17 out of 30) of the number plates achieved 95% accuracy or higher. Conversely, a 90threshold led to over 80% (24 out of 30) exceeding 90% accuracy, emphasizing broader accuracy without compromising quality

As we are detecting the Bengali digits and letters from poor-quality images by maximizing the image quality

Table IV: Comparison of our proposed approach with others.

Authors	Architecture	Resolution Enhancement	Accuracy
Proposed	CNN ResNet50	GFPGAN	88% 83%
Kabiraj et al. [11]	CNN	ESRGAN	84%
Haque et al. [14]	PSNR	ESRGAN	78%

using several image processing techniques, there remain some excessive noises in the image (Fig. 6). If these noises could be minimized the model would have achieved a higher accuracy.

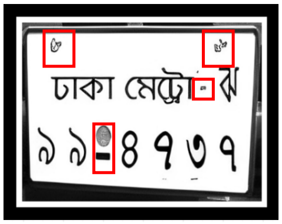

Figure 6: Uncovering Image Noise: How Garbage-Induced Distortions Affect Model Accuracy

We face some issues with the spacing of characters. In Bengali language ”Matra” is employed to connect the alphabet in a word (a line connects multiple characters). It is believed that the entire word ”Matra” consists of one connected component. Sometimes the matra gets disconnected but it still represents one connected component. It made it hard for the model to recognize these components. The image segmentation part is also challenging while solving these problems.

D. Performance Comparison

Table IV provides a brief comparison of various papers. Kabiraj et al. [11] uses increased super-resolution with ESRGAN and CNN for enhanced character recognition. Haque et al. [14] uses ESRGAN-based image reconstruction to obtain more accuracy in reading Bengali license plates. With the use of cutting-edge methods like GFPGAN, morphological processing, CNN, and RestNet50, our suggested system demonstrates Automated License Plate Recognition (ALPR) with digit recognition, generating better results.

VII. CONCLUSION AND FUTURE WORK

This research study presents a method for identifying vehicle license plates in Bangladesh, written in Bengali. CNN provides a wealth of features to enhance character recognition accuracy. The system was evaluated using 1292 images of Bengali digits and characters and achieved an accuracy of 88%. In terms of future scope, this system can be enhanced by adding other image processing methods such as noise reduction, perspective correction, and picture enhancement to further improve the precision of license plate recognition. To effectively control traffic and increase security, the system can also be connected with real-time surveillance systems. The system is adaptable for use around the world and can be expanded to recognize license plates from more nations and languages.

References

[1] S. Nooruddin, F. A. Sharna, and S. M. M. Ahsan, “A bangladeshi license plate detection system based on extracted color features,” in *2020 23rd international conference on computer and information technology (ICCIT)*, pp. 1–6, IEEE, 2020.

[2] M. A. Islam, G. M. Chowdhury, and M. N. M. Haque, “Bangla license plate detection, recognition and authentication with morphological process and template matching,” in *2021 2nd International Conference for Emerging Technology (INCET)*, pp. 1–6, IEEE, 2021.

[3] M. T. Rouf Shawon, R. Tanvir, and M. G. Rabiul Alam, “Bengali handwritten digit recognition using cnn with explainable ai,” in *2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–6, 2022.

[4] T. N. K. Khan, R. D. Dewanjee, N. J. S. Shorna, M. M. U. R. S. Sowad, and M. T. R. S. Shawon, “Evaluation of bengali handwritten mathematical equationusing convolutional neural network,” 2023.

[5] N.-A. Alam, M. Ahsan, M. A. Based, and J. Haider, “Intelligent system for vehicles number plate detection and recognition using convolutional neural networks,” *Technologies*, vol. 9, no. 1, p. 9, 2021.

[6] M. N. I. Suvon, R. Khan, and M. Ferdous, “Real time bangla number plate recognition using computer vision and convolutional neural network,” in *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 1–6, IEEE, 2020.

[7] P. Dhar, S. Guha, T. Biswas, and M. Z. Abedin, “A system design for license plate recognition by using edge detection and convolution neural network,” in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, pp. 1–4, IEEE, 2018.

[8] A. Kathirvel, S. P. Gandu, *et al.*, “Systematic number plate detection using improved yolov5 detector,” in *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, pp. 1–6, IEEE, 2023.

[9] R. Antar, S. Alghamdi, J. Alotaibi, and M. Alghamdi, “Automatic number plate recognition of saudi license car plates,” *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8266–8272, 2022.

[10] S. Raj, Y. Gupta, and R. Malhotra, “License plate recognition system using yolov5 and cnn,” in *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, pp. 372–377, IEEE, 2022.

[11] A. Kabiraj, D. Pal, D. Ganguly, K. Chatterjee, and S. Roy, “Number plate recognition from enhanced super-resolution using generative adversarial network,” *Multimedia Tools and Applications*, vol. 82, no. 9, pp. 13837–13853, 2023.

[12] F. Yin, Y. Zhang, X. Cun, M. Cao, Y. Fan, X. Wang, Q. Bai, B. Wu, J. Wang, and Y. Yang, “Styleheat: One-shot high-resolution editable talking face generation via pre-trained stylegan,” in *European conference on computer vision*, pp. 85–101, Springer, 2022.

[13] A. Patil, J. Ingale, A. Upadhye, S. Jaiswal, S. Kakade, and A. Bhatt, “Deep learning-based approach for indian license plate recognition using optical character recognition,” in *2023 IEEE Guwahati Subsection Conference (GCON)*, pp. 1–5, IEEE, 2023.

[14] N. Haque, S. Islam, R. A. Tithy, and M. S. Uddin, “Automatic bangla license plate recognition system for low-resolution images,” in *2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–6, IEEE, 2022.