Paper ID: 512

BENGALI LICENSE Plate Detection and Text Recognition using Deep Learning

Emran Hasan[1], Md. Miraiz Hossain[2], S. M. Fahim Faisal[3], Kazi Naimur Rahman[4]
[1-4]Department of Mechatronics and Industrial Engineering, Chittagong University of Engineering and Technology, Bangladesh
E-mail: u1809011@student.cuet.ac.bd, miraizhossain@cuet.ac.bd, fahimfaisal@cuet.ac.bd, naimur@cuet.ac.bd

Abstract

License plate recognition (LPR) is a critical technology for various applications such as traffic management, parking control, and law enforcement. This study presents a novel approach for detecting Bengali license plates, focusing on the unique challenges posed by the Bengali script. Though the LPR system is much more saturated for other languages, the LPR system has not made any notable progress in the Bengali-speaking nations or states in tackling their most concerning traffic management with insufficient road safety measures. To overcome these challenges, this research will use cutting-edge deep learning technology, especially the You Only Look Once (YOLO) object detection approach, to greatly enhance the accuracy, robustness, and overall performance of the LPR system. Advanced pre-processing, data augmentation, and post-processing methods are also explored to further improve system performance. Evaluation with real-world data, including diverse environmental conditions, shows the system's impressive robustness. The V7 model has the highest accuracy of 94.17%, whereas in the case of the F1 score, the V8 model has the score of 97%, which was the best among all three.

Keywords: YOLO, Bengali License Plate Recognition, Deep Learning

1. Introduction

Despite having the sixth-largest population in the world, the LPR system has not made any notable progress in the Bengali-speaking nations or states in tackling their most concerning traffic management with insufficient road safety measures. In Bangladesh, issues with traffic management and car ownership identification are significant. Most of the time, it is challenging to track down the drivers or owners of the cars that break the law or inadvertently repair the road. A traffic police officer must also spend a lot of time physically inspecting every vehicle's license plate. Therefore, the demand for an automatic license plate identification system to address these issues is great. So, a fast and robust LPR system is very important to identify stolen vehicles and support criminal investigations. In order to alleviate traffic congestion and transportation issues, it can also improve traffic management, speed toll collecting, and simplify parking procedures. LPR can also contribute to increased security, notably by regulating access to residential and protected locations. However, in order to exploit the technology's advantages while preserving individual rights, it is crucial to handle privacy and data security issues responsibly. This calls for appropriate rules and moral behavior.

In this study, an entire method is provided for recognizing license plates on automobiles in Bangladesh. This system is separated into two distinct phases. The identification and location of license plates is the initial stage. In the second stage, it is put into practice a reliable and quick segmentation algorithm designed exclusively for Bangladeshi license plates. The YOLO model is utilized for both cases.

2. Literature Review

License plate detection is a very common and saturated field. But it is very unfortunate that, despite being a very largely populated country, development on Bengali license plate detection isn’t developed that well yet. However, some of the works have been done and gained pretty good accuracy, although all contain some shortcomings as well. A comprehensive license plate identification system for Bangladeshi vehicles was proposed by Sohaib Abdullah et al. [7]. In the initial stage, they employed YOLOv3 to recognize a license plate from a picture of a car. The second stage uses this observed license plate to perform character detection and localization. However, as they didn't use other metropolitan license plates, this method is only for the Dhaka Metropolitan region. They therefore missed the "Dhaka" at this point. They concentrated on the final character of the first line, which designates the vehicle type, and the six numbers in the second line of the license plate. For character and digit recognition, they finally implemented ResNet-20. In the digit and character recognition stage, they had 92.7%

accuracy. Another license plate character recognition model was proposed by M M Shaifur Rahman et al. [8], which has very good accuracy. However, they simply used a CNN model to implement the character recognition portion, and the model's accuracy was 88.57%. For training, they used about 1700 photos. They had to manually identify and crop the image of the license plate from the vehicle because they were only focused on the recognition task. To create their training dataset, they also had to manually crop the characters. So, creating the dataset required a lot of labor. They collaborated with 16 courses, but they avoided mentioning the class names. However, they divided the segments into smaller ones, such as the multi-character "Metro" and "Dhaka" portions. In 2014, Raiyan et al. [9] proposed a template matching approach for recognizing license plates. But they only concentrated on the task of character and digit recognition. They took advantage of the "matra" peculiar Bangla character characteristic. Here, they largely used image processing methods. They presented some test cases but came to no broad conclusions, making it image-specific. But there was no mention of any dataset. The experiment lacks generalizability because they only used a small number of photos. Furthermore, they made no mention of a license plate detection system. Using image processing and machine learning, Prashengit et al. [10] suggested a license plate recognition system for Bangladesh. They unveiled a dataset of 2800 photos of license plates. They identified the license plate from a picture using a process called shape validation in image processing. For tilt correction, they next employed image processing once again. The characters on the license plate were then divided using related components. They took the connecting parts out of the license plate, each of which stands for a different character. With the Adaboost classifier, they put the recognition task into practice. They employed the histogram of gradient (HOG) and local binary pattern (LBP) as their two major features for recognition. Their suggested method has a 97.3% accuracy rate, which appears to be the most recent cutting-edge model. However, this system mainly relies on a number of image processing methods, such as HOG, LBP, tilt correction, ROI extraction, morphological analysis, shape detection, and shape validation. This system requires some time as a result. An end-to-end method for recognizing Bangladeshi license plates was proposed by Nazmus Saif et al. [11]. However, their study hasn't specifically referred to data from their dataset. Therefore, it is exceedingly difficult to assess or compare their work. Furthermore, they asserted that their dataset had an accuracy of 99.5%. They made a rather naive choice for this kind of challenge by using the YOLOv3 for both license plate detection and recognition. However, we replicated their model to assess their claim before they were finished. Unfortunately, utilizing YOLOv3 on our dataset, we were unable to attain that precision. We believe they encountered this issue as well and added it to the restriction. In conclusion, their dataset was not diverse enough and could not be used for real-life circumstances, which is why their model had such high accuracy. A sensor-based method for vehicle recognition has also recently been suggested.

3. Methodology

In this study, the overall license plate detection process is done in two steps. Firstly, by detecting the license plate from the object, and secondly, text extraction from the license plate. The whole process is now given in a flowchart in “Fig. 1.". At first, the license plate is detected in real time from video streaming. This video is taken from any camera and feeds into a deep learning model frame by frame. After detecting the license plate, the detected frame is fed into another deep learning model for text extraction and recognition. Therefore, two separate models have been used: Model 1 is used for license plate detection, and Model 2 is used for text recognition. Here three different versions of YOLO v5, v7, and v8 have been used to train the custom model. To detect a desired object by any deep learning model, someone has to go through several steps. First, build a custom model; second, train the model; and at last, test the model. To build and train a custom model, the following steps are required:

Training images collection

Collecting training data in this case was images of vehicles with license plates on them, and it was the most crucial part of the project. The acceptability, robustness, and accuracy of the project depend massively on it. Some images were collected for the training purpose, but it wasn't enough to recognize vehicles from all over the country. So, the dataset is taken from Roboflow, which includes all the districts of Bangladesh, and merged it with my own images. By doing so, the acceptability and recognition capability of the project have increased.

Images pre-processing

The input images must be pre-processed, before them to feed into the model. All the images are resized to a standard input size and the pixel values are normalized in a specific range compatible for the model.

Augmentation

Roboflow is a powerful tool for data augmentation. It enhances the dataset by creating variations of your original images, such as rotations, flips, implementing noise and blur, and brightness adjustments. These augmented images help improve model generalization and robustness, leading to better performance during training and testing.

Annotated image dataset

Annotation was a challenging task in this work. AutoDistill is used to annotate huge data automatically. But it lacks control over the wrong annotated image. Then annotation of images is done by Roboflow, over which we had much control in annotating images.

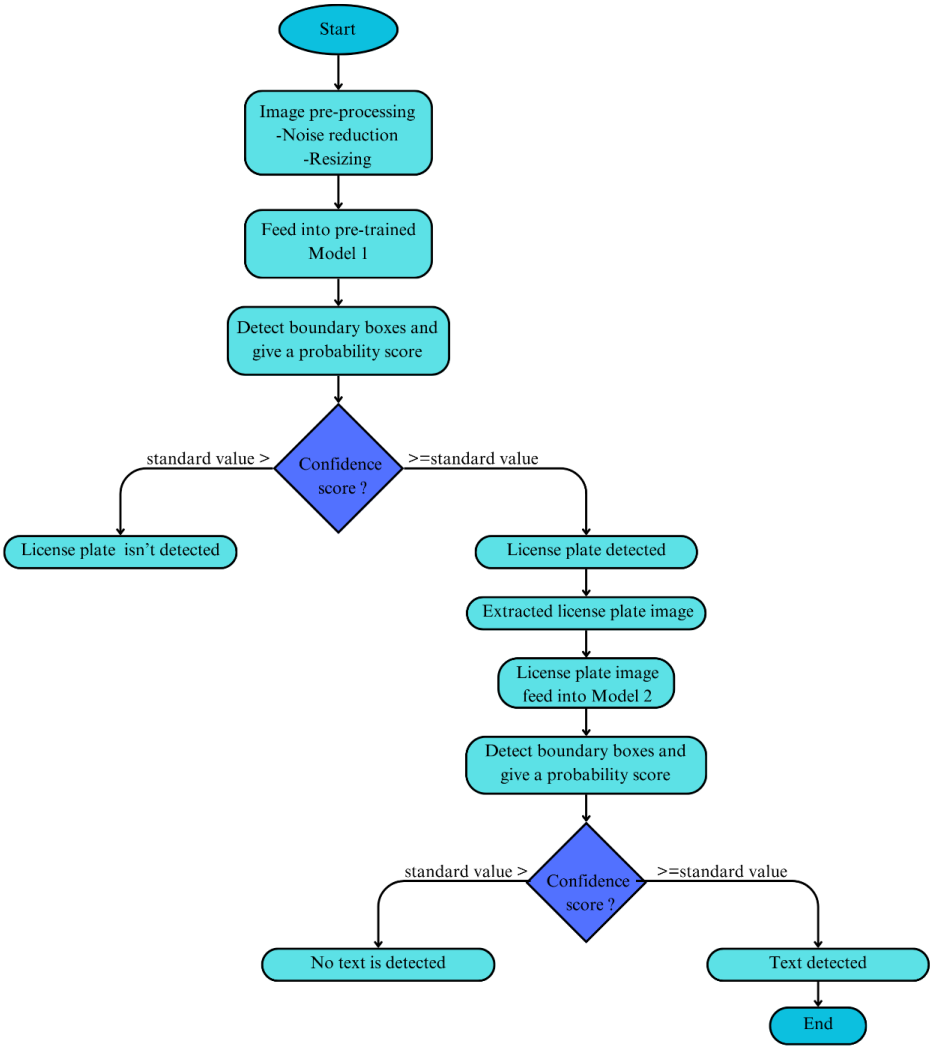


Fig. 1. Workflow of the proposed model

Export the dataset

This step includes converting the annotated dataset for the respective model for training purposes. Then we exported data directly to Google Colab for training.

Training the model

Training the model mirrors training the model by iterating through the training dataset. For training, CLI is used instead of a Python script as it was much easier in Google Colab. To train the dataset, various deep learning frameworks like TensorFlow and PyTorch can be used. During training, the model learns to predict the object with a bounding box and gives probabilities based on the annotated dataset.

Evaluation

A distinct dataset called the testing and validation set will be used to evaluate the model once it has been trained. The effectiveness of the model to identify the object will be evaluated using a range of criteria, including accuracy, precision, recall, and F1-score. This will provide informative information regarding the model's efficiency and general performance, helpful in establishing its sustainability for detecting the intended object

Fine-Tuning

Based on the evaluation we tuned some parameter like epoch size, batch size & confidence score during training period for better result.



(a) (b) (c)

Fig. 2. YOLO first takes an input image; YOLO framework will divide the input image into grids; YOLO then predicts the bounding boxes and measure the probabilities of each box containing object

Workflow

At first a camera will capture a frame from the live stream, which will be the input image for the system. This image will be the input for the first model. Model-1 will detect the license plate. Then the detected license plate would be cropped according to the coordinates of the detected bounding box. This cropped image of the license plate will be extracted and then fed into the second model for the text extraction.



(a) (b) (c) (d)

Fig. 3. Input image (a); Detected number plate (b); Extracted number plate (c); Detected text (d)

From the "Fig. 3" it is seen that the second model has detected the text within the license plate with a good confidence score. Now the text will be classified and the classes would be rearranged as the license plate based on the co-ordinate of the bounding box. Then the extracted text will be compared with the dictionary for the related information for the vehicle.

4. Results and Implementation

A total of three different versions of the YOLO model have been trained. This section will provide some evaluation matrix to evaluate the model performance and to choose the best version for this project. From the tables mentioned above, we can say that the V7 model has the highest accuracy. But in the case of the F1 score, the V8 model has the best score among all three. Now the question arises: Which version to choose, and what is the reason? Well, it's V8. First of all, V5 lags behind in both cases in front of the other two. For the other two, if we consider accuracy, it's V7, and in the case of F1 score, it's V8. So, both models have some advantages over the other. But V8 offers some key improvements over the V7 model. V8 is based on more lightweight architecture and focused on speed and efficiency. So, the V8 model would be more suitable among them.

Table 1. Accuracy and F1 score of all three models

Model	Accuracy	F 1- score
YOLOv5	90.47%	0.93
YOLOv7	94.17%	0.95
YOLOv8	92.31%	0.97

In this section, a real-life implementation of this work is created. Here, mobile is used as the webcam to capture frame from the real world.

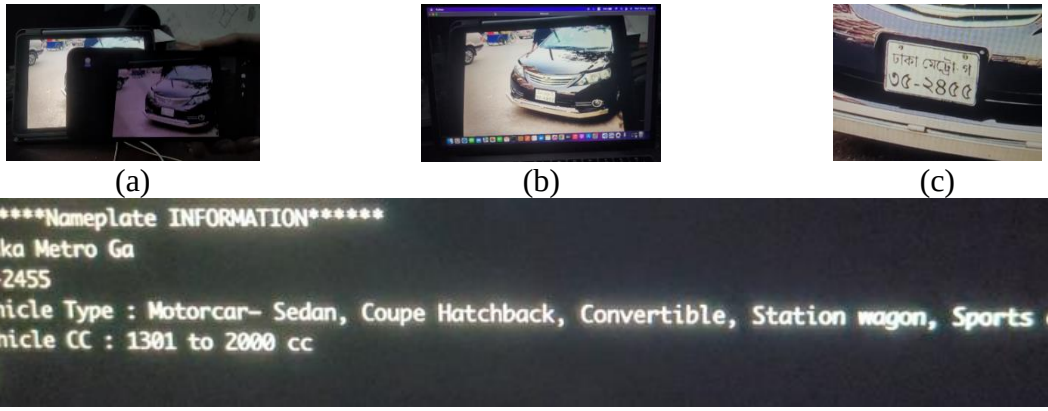


Fig. 6. Webcam capture (a); Live stream from the webcam (b); One specific frame from the live stream (c); Final output as a text with Some specific information about the vehicle (d)

6. Conclusion

In this study on building an LPR system using deep learning, there is potential to alleviate traffic congestion and transportation issues. Additionally, it can speed up toll collection, simplify parking procedures, and enhance traffic management. By controlling entry to homes and other protected areas, LPR can also help to strengthen security. However, it is essential to address privacy and data security issues appropriately in order to take advantage of the technology's benefits while upholding individual rights.

Although the library of vehicle information used in this project was just a demo to demonstrate the project, it has no real meaning unless we get the library from BRTA. At least we can create our custom dictionary for our own purpose and get the project to be applied with it.

References

[1] "BRTA standard license plate | Download Scientific Diagram." https://www.researchgate.net/figure/BRTA- standard-license-plate_fig2_338096474 (accessed Nov. 03, 2023).
[2] "Vehicle registration plates of Bangladesh - Wikipedia." https://en.wikipedia.org/wiki/Vehicle_ registration_plates _of _ Bangladesh (accessed May 18, 2024).
[3] "What Is Deep Learning?|How It Works, Techniques & Applications - MATLAB & Simulink."https:// www.mathworks.com/discovery/deep-learning.html (accessed Nov. 03, 2023).
[4] "Basic CNN Architecture: Explaining 5 Layers of Convolutional Neural Network|upGrad blog."https:// www.upgrad.com/blog/basic-cnn-architecture/ (accessed Nov. 03, 2023).
[5] "Timeline of You Only Look Once (YOLO) variants. | Download Scientific Diagram." https:// www.researchgate.net/figure/Timeline-of-You-Only-Look-Once-YOLO-variants_fig1_369379818 (accessed May 18, 2024).
[6] "YOLOv5 | PyTorch." https://pytorch.org/hub/ultralytics_yolov5/ (accessed May 18, 2024).
[7] M. M. Sarif, T. S. Pias, T. Helaly, M. S. R. Tutul, and M. N. Rahman, "Deep Learning-Based Bangladeshi License Plate Recognition System," *4th Int. Symp. Multidiscip. Stud. Innov. Technol. ISMSIT 2020 - Proc.*, 2020, doi: 10.1109/ISMSIT 50672.2020.9254748.
[8] M. M. Shaifur Rahman, M. Mostakim, M. S. Nasrin, and M. Z. Alom, "Bangla license plate recognition using convolutional neural networks (CNN)," *2019 22nd Int. Conf. Comput. Inf. Technol. ICCIT 2019*, 2019, doi: 10.1109/ICCIT48885.2019.9038597.
[9] R. A. Baten, Z. Omair, and U. Sikder, "Bangla license plate reader for metropolitan cities of Bangladesh using template matching," *8th Int. Conf. Electr. Comput. Eng. Adv. Technol. a Better Tomorrow, ICECE 2014*, pp. 776–779, 2015, doi: 10.1109/ICECE.2014.7026925.
[10] P. Dhar, M. Z. Abedin, R. Karim, Fatema-Tuj-Johora, and M. S. Hossain, "Bangladeshi license plate recognition using adaboost classifier," *2019 Jt. 8th Int. Conf. Informatics, Electron. Vision, ICIEV 2019 3rd Int. Conf. Imaging, Vis. Pattern Recognition, icIVPR 2019 with Int. Conf. Act. Behav. Comput. ABC 2019*, no. May, pp. 342–347, 2019, doi: 10.1109/ICIEV.2019.8858580.
[11] N. Saif *et al.*, "Automatic License Plate Recognition System for Bangla License Plates using Convolutional Neural Network," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2019-Octob, no. December, pp. 925–930, 2019, doi: 10.1109/TENCON.2019.8929280.
[12] "Yolo Framework | Object Detection Using Yolo." https://www.analyticsvidhya.com/blog/2018/12/practical-guide-object-detection-yolo-framewor-python/ (accessed Nov. 04, 2023).