Ttitle of the Thesis

Improved Deep Learning Based Model for Vehicle Plate Detection, Recognition, and Authentication

Submitted By

Md. Mahmudul Hasan Suzan

ID: 191-16-410

Supervised By

Mr. Abdullah Bin Kasem Bhuiyan

Lecturer

Department of Computing & Information System

Daffodil International University

Daffodil International University

Dhaka, Bangladesh

Submission Date: 14 January 2023

APPROVAL

This thesis titled " **Improved Deep Learning Based Model for Vehicle Plate Detection, Recognition, and Authentication"**, Submitted by **Md. Mahmudul Hasan Suzan**, ID No: **191-16-410**, to the Department of Computing & Information Systems, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computing & Information Systems and approved as to its style and contents. The presentation was held on 14/01/2023.

BOARD OF EXAMINERS

Mr. Md Sarwar Hossain Mollah — **Chairman**
Associate Professor and Head
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

Mr. Md. Mehedi Hasan — **Internal Examiner**
Lecturer
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

Mr. Syed Tangim Pasha — **Internal Examiner**
Lecturer
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

Dr. Saifuddin Mr. Tareeq
Professor & Chairman
Department of Computer Science and Engineering
Dhaka University, Dhaka

External Examiner

DECLARATION

I, **Md. Mahmudul Hasan Suzan**, hereby declare that the work in this dissertation titled " **Improved Deep Learning Based Model for Vehicle Plate Detection, Recognition, and Authentication"**; I have done this thesis under the supervision of Mr. **Abdullah Bin Kasem Bhuiyan**, Lecturer, Department of Computing and Information System (CIS) of Daffodil International University. I am also declaring that this thesis or any part of there has never been submitted anywhere else for the award of any educational degree like B.Sc., M.Sc., Diploma or other qualifications.

Supervised By

14.1.23

Mr. Abdullah Bin Kasem Bhuiyan
Designation: Lecturer
Department of CIS
Daffodil International University

Submitted By

Suzan

Md. Mahmudul Hasan Suzan
ID: 191-16-410
Department of CIS

ACKNOWLEDGEMENT

Firstly, I would like to thank Allah for blessing me and giving me patience and power to complete my thesis work.

I thank my thesis advisor, **Abdullah Bin Kasem Bhuiyan, Lecturer**, the Department of Computing and Information System, Daffodil International University, Dhaka, for his informational and motivational support. I completed this research project due to his consistent support, patience, guidance, and insightful advice.

I would like to express my heartiest gratitude to our department head **Md. Sarwar Hossain Mollah** has supported and encouraged me during the challenges of university and this thesis work. Not least, our lecturer, **Md. Mehedi Hassan** for his valuable motivation and support in my university life. In addition, I would like to extend my deepest gratitude to the other faculty of the Department of Computing and Information System at Daffodil International University in Dhaka and the staff there.

DEDICATION

This work is dedicated to my parents, Md. Hafiz Uddin Mollah and Asia Begum have always loved me without conditions and whose heroic example has inspired me to work assiduously toward my goals.

ABSTRACT

In recent years computer vision models have made our daily life easy in various ways, especially in reducing roadside problems. Many research works are already completed to achieve the goal of automated road surveillance. But these models' actual implementation has failed due to the poor accuracy of the model and other relevant factors. This paper presents an improved model to detect, extract, recognize and validate Bengali license plates from vehicles. In order to recognize vehicle plates more accurately and for various uses, including automated vehicle monitoring, roadside assistance, toll collection, parking management, etc., we implemented a Yolo-based CNN model to detect Bangla license plates and mask R-CNN for recognition of license characters. A total of 6528 images were used in training our model. Based on roadside test images, the experiments can detect at a rate of 98.2%, recognition of 95.6%, and a validation rate of 100%, respectively.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Because of the country's growing population and increasingly congested roads, Bangladesh has seen a considerable rise in the number of automobiles and their usage. Large cities such as Dhaka and Chittagong are seeing increased traffic infractions due to improper management and traffic offences. It is becoming more challenging to effectively respond to various motor vehicle accidents in today's world due to the rapid expansion of the transportation sector and the dramatic rise in the number of cars on the road. The use of artificial intelligence and computer vision in the administration of transportation systems might help find solutions to these issues. However, developing nations like Bangladesh face many obstacles to adopting it. Better traffic monitoring, speed control, automated toll collection, parking lot management, license number validation automatically, and a wide variety of other vehicle-related services are all possible thanks to the automatic LPR system's capabilities. However, to our knowledge, the Bangladeshi transportation industry has not yet implemented any of these systems. There is still significant manual human power in recognizing license plates in Bangladesh.

Even though the ALPR method is used in other developed countries, Bangladesh struggles with several obstacles. Support for the Bengali language is unavailable on these systems since it is language-dependent, which is why the commonly used systems are language-dependent. Identification based only on the digital image attributes of an object cannot be relied upon because of the complex lexical structure of the Bangla language.

1.2 Research Aim

The primary goal of this research is to offer a straightforward technique for using a deep learning model to recognize the Bengali license plate used on the front and back sides of registered vehicles. In the paper, we proposed more improved model for detecting license plates that could provide real-time inference speed and accurate detection and recognition of license plates every time. In addition, one of our primary focuses was to prepare a large dataset and developing a dependable model that can be assembled and run quickly. Because we aim to make ALPR systems more accessible to a wider audience, this allowed us to launch the system more quickly.

1.3 Research Objectives

The roadside problem in Bangladesh due to the high number of vehicles is not unknown to us. The authorities fail to control road violations due to enough manpower. Thanks to the advanced artificial intelligence techniques, that brings revaluation on e-traffic management. So, in the future, we aim to build a more improved AI-based system for e-traffic management based on the finding of this research. The goals of this paper can be summed up as follows:

- To find the best model for Bengali license plate detection, recognition, and authentication.
- To design and develop a prepossessing less model to accurately recognize license plates.
- To Provide a comparison of already published work and our work.
- The accuracy of this work-related model will be improved in the future.

1.6 Research Scope

This research focuses on detecting and recognizing Bengali license plates from vehicles' front and rear regions. The geographical location and dataset covered in the study in Bangladesh. This research's comparison and literature review are only for any published machine learning-related work.

CHAPTER 2

INITIAL STUDY

2.1 Problem Background

There are more than 4.47 million automobiles that have been registered in Bangladesh. It is not simple to manually follow the movement of automobiles, particularly in urban regions with a high volume of vehicles due to the significant population density. The problem of people driving their vehicles at unsafely high speeds and the problem of people using stolen or forged license plates present a significant barrier to efficient traffic management. In addition, there are a lot of toll booths on highways all around the country, and most of them still use the manual input technique. This resulted in a significant backlog on the roadways, a typical occurrence in Bangladesh.

2.2 Rational of the study

Many research has been done on the following topic, but most research did not get the expected result. We reviewed many related papers in the literature review, and most of the work was done using several software tools like matplotlib. Also, these studies lack data and confidence in their models. So gathered more raw data, labeled them for improved training, and then compared it with other neural network techniques. As a result, we used 6528 images for training purposes of the model

2.3 Research Questionaries

There were a lot of challenges during data collection and study. The big challenge was configuring the model properly for training. So, to minimize the risk in the later part, we noted down some research questionaries.

- What is the source of data?
- What is the data collection methodology will be used?

- How much preprocessing is required for the collected data?
- What will be the excellent ratio for train-test data?
- Will the model produce a satisfactory result in the end?

2.4 Contribution

Our main contribution to the work is:

- The first implementation of Bangla character recognition model using mask r-CNN deep learning method and detectron2.
- We prepared a publicly annotated data set of over 6500 images.
- Methodology for collecting non-synthetic real-world data sets of low-quality license plate photos with minimum prepossessing.
- Achieved the highest precision, recall performance among the all-other techniques.

2.5 Overview of Standard License Plates

The Bengali alphabet and numbers are used on license plates for vehicles. The digital license plate's BRTA format is divided into two lines. The first and second lines contain the "city - vehicle class" and "class number - vehicle number" information.

The rules that the BRTA developed stipulate that the number of license plates for a vehicle must be written in Bangla script on a flat plate made of steel or aluminium with dimensions of 524 millimetres by 112 millimetres. Depending on the type of vehicle, the license plates must also follow a specific color scheme; nevertheless, these color schemes have no bearing whatsoever on the technique that has been provided here.

Words and alphabets written in Bangla can be found on the upper line, while numerals written in Bangla can be found on the below line. The first word printed in Bangla on the top line indicates the name of the city or metropolitan area where the vehicle is registered. The second term is "Metro," which is an abbreviation for the word metropolitan. The third one is the vehicle class, which can be denoted by a specific letter of the Bangla alphabet. Fig.1 and Table 1, we visualized the characteristics of Bangladesh license plates, This could refer to a car, motorcycle, or other types of vehicle. The following four numbers on the bottom line make up the vehicle's unique identification number, which has a total of six digits in it. This number can only be found on that vehicle. The digits of the vehicle class number represent the vehicle's classification. The following numerals are allowed to be displayed on a vehicle's license plate: 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 (R. Islam et al., 2020) .

Table 1: BANGLADESHI LICENSE PLATE CATEGORIES

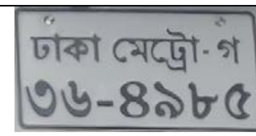
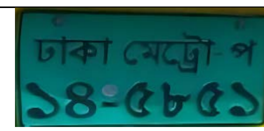
Vehicle Category	**License Plate Background**	**Color of Character**	**License Plates**
Private	White	Black	
Commercial	Green	Black	

Fig 1: Vehicle number plate of Bangladesh

2.6 Abridged Methodology

There has been some research on Bangla LPR systems in recent years. Furthermore, current ALPR systems have a wide range of applications, and numerous strategies have been developed in this field. A full LPR system can be deployed using a variety of methods. Some popular deep learning models like CNN and YOLO are the most popular.

We'll briefly review the ALPR systems' most popular techniques. The literature review section provides the descriptive approach and a review of these models.

2.7 Layout of the Report

In CHAPTER 1, we provided the introduction to this study with research objectives and scope.

In CHAPTER 2, we discussed the problem background, research questionaries, and a view of the abridge methodology.

In CHAPTER 3, we will provide an overview of the existing models of both our country and globally with a comparison of some related studies.

In CHAPTER 4, we will show the detailed methodology used for this research, including the data collection method and preprocessing.

In CHAPTER 5, we will analyze the experiment result and some of the metrics used to evaluate the mode.

CHAPTER 6, We will include the conclusion and discuss some future works.

CHAPTER 3

LITERATURE REVIEW

3.1 Overview

The process of detecting and recognizing vehicle license plates is done in many different ways worldwide. Most of the works in this field are centered on various image-processing methods and Neural Networks. Most research papers on the topic break down number plate recognition algorithms into three stages: the first stage involves the number plate region extraction, the second stage consists of the characters segmentation from the plate, and the third stage involves the recognition of characters. This section presents some previous research on license plate recognition that is significant to the proposed approach used internationally and for our country.

3.2 Global Practice for ALPR System

In (Wen et al., 2011) the authors used a new binary method to locate the Japanese license plate at 97.16\% and achieved 93.54% in character recognition using SVM integration with different character features. They used image processing techniques to recognize multi-line Japanese plate numbers in their work. C.N. (Huang et al., 2009) used a new adaptive image segmentation technique for license plate detection and achieved 96% accuracy. Still, during the recognition, they used a two-layer Probabilistic Neural Network (PNN), and the model was able to detect 89.1% of characters from plates. Moreover, (Jiao et al., 2009) proposed a faster and more accurate multi-style LP recognition model using quantitative parameters like plate rotation angle, plate line number, character type, and format. They achieved 95.5% detection and 97.1% recognition with their systems.

Numerous strategies, like the edge detection method, have been put forth by researchers to locate the license plates. A system using Support Vector Machines (SVM) was developed by the authors of (Kim et al., 2000), and they report a fantastic average character recognition rate of 97.2% for their invention. The architecture, however, was created specifically for Korean license plates and did not react to license plates from other countries. (Du et al., 2012) went over various techniques for automatic license plate recognition. According to the traits of each process stage, they have categorized additional methods. They divided the various algorithms into groups according to their advantages, disadvantages, recognition rates, and processing times. They have also offered a succinct set of guidelines for how future work on ALPR should be conducted, including how to handle challenges and other similar issues.

In (Zheng et al., 2005), (Sanyuan et al., 2004), (Sarfraz et al., 2003), (Kanayama et al., 1991) used the Sobel filter to detect edges. The image shows license plate edges due to the color shift between the plate and the car body. Horizontal edge detection yields two horizontal lines, vertical edge detection yields two vertical lines, and simultaneous edge detection yields a rectangle.

3.3 Bangla ALPR System

3.3.1 Neural Network-Based Approach

(Rahman et al., 2019) Implemented a Bangla license plate recognition system based on CNN and achieved a recognition rate of 89%. They put their research into practice using a deep learning framework built in MATLAB. (Ghosh et al., 2011) worked with neural network and image processing techniques. To extract the license plate, they used Sobel filters, and to get rid of irrelevant objects, they used morphological operations. They used linked component analysis for segmentation. Similarly, (Joarder et al., 2012) also implemented the same system with the feed-forward network. Both models could recognize license plates at 84% and 84.16%, and recognition was 80% and 92% accurate, respectively.

3.3.2 Image Processing Based

The author (M. A. Islam et al., 2021) proposed a morphological operations-based character recognition using the template matching technique and segmentation using connected component analysis. Despite the lack of data, they were still able to detect plates with an accuracy of 94 and recognize characters with a 95.1 accuracy. However, the system's flaw is that it takes a very long time to produce an average output of 2.21 seconds. (Baten et al., 2014) used linked component analysis and Bangla "matra" to segment license plate characters. They then recognized Bangla characters and integers using word template matching. Using "matra" to segment characters simplifies the method. Their algorithm recognized license plates in 1.3 seconds and recognition in 0.67 seconds. On the other hand, (R. Islam et al., 2020) used Morphological operation and histogram analysis for character segmentation. With the help of extracted Histogram of Oriented Gradient (HOG) features, they also used connected component analysis and bounding box technology for character recognition to achieve 94.6% character extraction accuracy and 91% ROI extraction accuracy.

3.4 Literature Review Summary

The systems presented by (Huang et al., 2009; Jiao et al., 2009; Kim et al., 2000; Sarfraz et al., 2003; Wen et al., 2011), the authors, reported that while the model's overall performance is adequate for detecting license plates, its recognition component has lower accuracy. The systems could only handle one line of large characters, though. However, Bangladeshi license plates have two lines. Even if the license plate had two lines of characters, the maximum accuracy obtained by (M. A. Islam et al., 2021) for plate detection was 94%, and her success rate for character recognition was 95%. So, by this summary, it is clear that a more accurate and precise model for this type of work is needed for future work.

CHAPTER 4

METHODOLOGY

4.1 Overview

On completing the literature review in Chapter 2, we find that the main problem with detecting or recognizing Bangla license plates is the lack of an enriched public data set and more processing phases. It is tough to compare these methods with each other as there is no proper data set to make a good benchmark. Therefore, we had to experiment with these methods to the best ability to get close to the original work with our data set. With that experience, we proposed a new method for detecting and recognizing the Bangla license plate.

4.2 Research Framework

We choose to use a Yolo custom-trained model for license plate detection and an improved deep learning-based mask-r-CNN network with detectron2 for character recognition. The Yolo model outperformed any object detection algorithm and generalized better in different scenarios. The main advantage of the Yolo model is it uses an end-to-end neural network that makes predictions of bounding boxes and class probabilities all at once. YOLO achieves state-of-the-art results, beating other real-time object detection algorithms by a large margin (Bandyopadhyay, 2022).

On the other hand, detectron2 is a modular object detection library enriched with many deep learning models such as Faster R-CNN, Mask R-CNN, RetinaNet, and DensePose. Detectron2 supports object detection with boxes and instance segmentation masks with a combination of semantic and instance segmentation. Fig 1.1 shows our Block diagram of the proposed method.

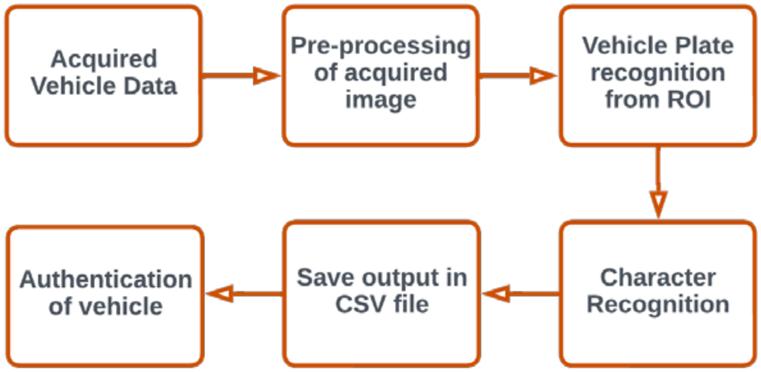


Fig 2: Proposed method Block diagram

4.3 Raw Data Acquisition

The data sample used in this research is collected from the roadside with 108mp Xiaomi mi 11 cameras. To construct a model that is as accurate as possible, we need photographs that feature a variety of settings, including both lighting and background elements. So, our dataset for both models consists of day and night data.

4.4 Acquired image preprocessing

This is the very first stage of the strategy that we have suggested. During this stage, the OpenCV program takes a picture of the vehicle and analyzes it.

In this research work, we focused on minimizing the preprocessing time and process for the model. That's why sufficient preprocessing was used for this work.

4.4.1 Cropping image

For better training and to reduce training complexity, the model must have a specific ROI labeled. Keeping it in mind, we first cropped the interested ROI(Region Of Interest) from images using a snipping tool in a rectangular shape. At the same time, the aspect ratio of the cropped image width and height was the same.

4.4.2 Resize Image

In the previous stage, the images were cropped into a rectangular shape. However, the image that was recorded may be any size. We will need to resize it to a preset size to conserve space and generalize it for the model. The height and width of the vehicle image are both shrunk to 416 pixels in this approach, but the aspect ratio is maintained throughout the process. Other than no other preprocessing is required for our proposed model.

4.5 Data Augmentation

Augmentation is essential for training the model in various conditions and facets. A substantial amount of model overfitting to the training data can be avoided by augmenting the image classification process. Nevertheless, setting augmentation parameters when augmenting a vehicle plate data set is laborious. Both YOLOv5 and detectron2 afford us the luxury of full integration with Augmentations, a well-known open-source image augmentation package. We enhanced the training of our AI models with this package to make them even more effective (ultralytics, 2021). This has automatically applied Albumentations transformations during YOLOv5 training.

4.6 Data Labeling

Data labelling is one of the important parts of supervised machine learning. Depending on the model, the data labeling techniques may vary. But the main concept for labeling images is that we label the images with our label class using any data labeling software. There are many popular data labeling programs available. In this part, we talked about the techniques for labeling data used for this research.

4.6.1 About LabelMe

We used LabelMe for labelling and annotating our dataset images in this research. LabelMe is a free, open-source software for image annotation tool that allows labeling objects and their location images. This software is available on GitHub (Wada, 2016).

LabelMe is designed to be very easy to use and can export annotation in popular formats like Yolo, coco, XML, etc.

4.6.2 Labeling process

At this stage, we first imported our dataset folder into label me, and then, using the polygon shape tool, we annotated every image with its appropriate label. The label class was written in English to make it easier for everyone to understand. LabelMe allowed us to save the output format automatically.

For training our model 1, we export one YOLO labeling text file per image. Each image object has a BBox annotation in the text file. Normalized to image size, annotations are 0–1. The format is:

< object-class-ID> <X center> <Y center> <Box width> <Box height>

For training of model 2, we exported the annotated images in coco JSON format to save the labels and metadata. Images and associated object detections are included in the format's labeled dataset. The format is:

{"categories": [ ], "images": [ ],"annotations": [ ]}

4.6 Dataset

For our data set, we used 108-megapixel cameras to capture the image. The images had the common Joint Photographic Experts Group (JPG) and PNG formats and were in color mode. The data set consists of day data for 2200 vehicles plate and 587 data for the night. After the augmentation, the dataset had a total of 6528 images; Every image consists of multiple instances

4.7 Proposed Model

As mentioned in subsection 4.1, we proposed two models for this research work. Yolo object detection model for license plate detection and mask rCNN for Bengali character recognition in the license plate. The Yolo uses a neural network to analyze the entire image, split it into sections, and forecast bounding boxes and probabilities for each component in order to identify the vehicle's license plates. Additionally, this guarantees that the object detection algorithm only identifies each object once. The main advantage of Mask R-CNN over other CNN algorithms for character recognition is that it has an additional branch for predicting segmentation masks on each Region of Interest (RoI) pixel-by-pixel. We will go into more detail about the suggested model in this section.

After finishing the literature review in CHAPTER 3, we concluded that the most significant challenge with detecting or recognizing was. No publicly available dataset is enhanced with Bangla license plates. As a result, we were required to conduct experiments with these methodologies to the best of our abilities in order to get our dataset as close as possible to the original work. To locate the license plate, our pipeline begins with capturing the image, then it's preprocessing it into a resolution of 416 by 416 pixels, and finally, its input into our first model. The first model uses a predetermined threshold value to locate any possible license plates; once a plate is found, it is extracted from the input image and cropped to remove any surrounding background. Following that, the second model is used to finish character segmentation, and based on the results, we can recognize the text on the plate. At last, the number on the driver's license is checked against the database already in place to ensure its authenticity. Fig 3 we depict the proposed flowchart for the models.

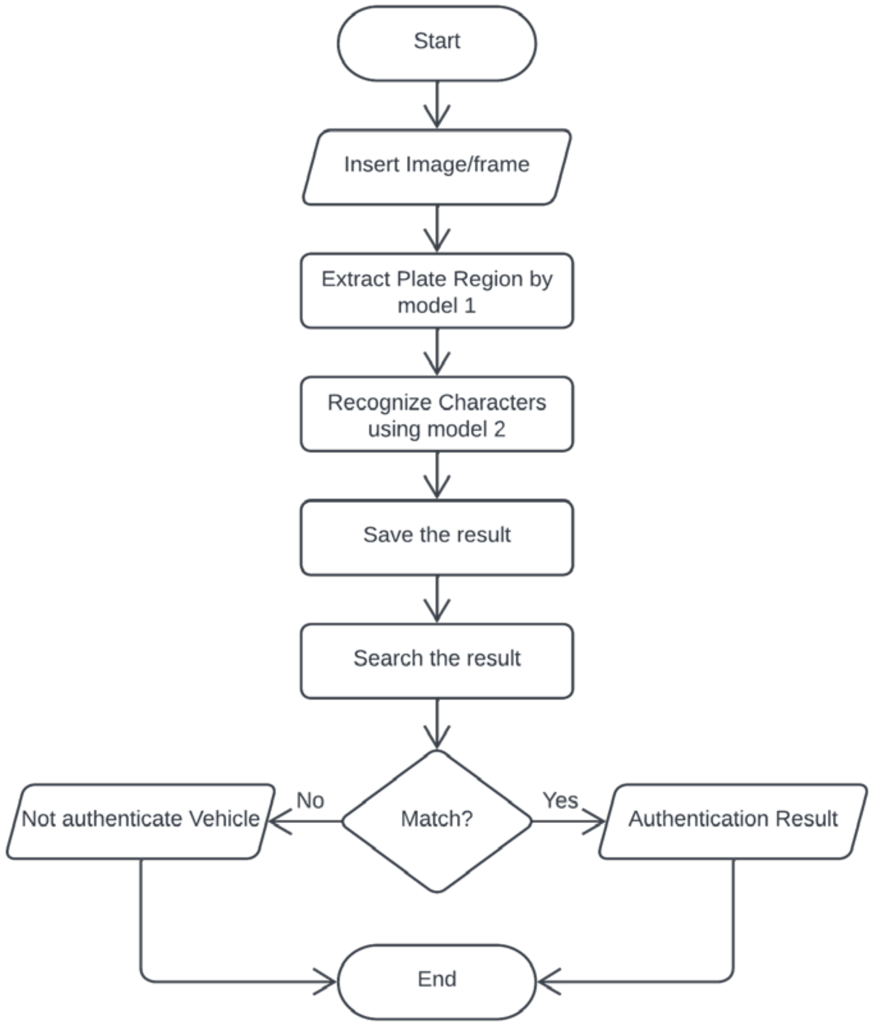


Fig 3: Proposed method flowchart diagram

4.8.1 Number Plate Detection

For our research, detecting vehicle plates is the first problem from images/video. It is also one of the major modules of this work. Our preferred convolutional neural network model with YOLOv5. We chose YOLO because it has a number of advantages over some other CNN models in the literature for our system.

YOLO can accurately detect and localize license plate images in real-time. While working on some images, other image processing methods did not scale for the whole dataset and were not as accurate as YOLO for this task. YOLO also removes background noise from data. We adopted YOLO for these reasons.

In order to extract rich and informative features from an input image, YOLO v5 makes use of the CSP — Cross Stage Partial Networks as its backbone. In figure 4, we show the primary architectural components of this model. The activation function known as Leaky ReLu is utilized in the middle/hidden layers, whereas the sigmoid activation function is utilized in the detection layer that comes last. One possible definition of relu activation is as follows: f(x) = max(0.1x, x)

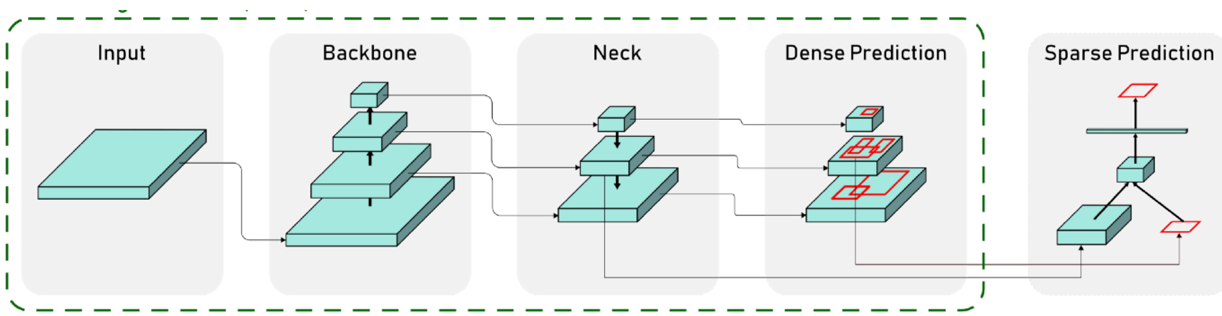


Fig 4: Architecture of Yolov5

YOLO is designed to achieve a minimum value for the sum squared error across all five parameters (x, y, w, h, and Ci). The location of the object's center will be determined using the x and y coordinates. It is anticipated that the width and height of the object will be relative to the image. Last but not least, the confidence ci is calculated by taking the intersection over union (IOU) of the anticipated and ground truth boxes. The class probability for any object that was found is denoted by the symbol Pi(c). The following is one possible definition for this model's loss function:

$$LOSS = L_{classification} + L_{confidence} + L_{CIoU}$$

$$L_{classification} = \sum_{i=0}^{s^2} \ell_i^{obj} \sum_{j=0}^{B} \left[\left(p_i(c) - \hat{p}_i(c) \right)^2 \right]$$

$$L_{confidence} = \sum_{i=0}^{s^2} \sum_{j=0}^{B} \ell_i^{obj} \left[\left(C_i - \hat{C}_i \right)^2 \right] + \lambda_{noobj}$$

$$\sum_{i=0}^{s^2} \sum_{j=0}^{B} \ell_i^{noobj} \left[\left(C_i - \hat{C}_i \right)^2 \right]$$

This model 1 has 127 convolutional layers. For detecting the license plate, we trained our model with image sizes of 416 by 416 pixels and 30 epochs, Batch size 16. This model was trained with a total of 2611 datasets. Fig.5 & 6 depicts the training configuration and batch output, respectively.

```
[ ] for fold in range(NUM_FOLD):
        print('FOLD NUMBER: ', fold)
        !python train.py --cache \
                         --img {416} \
                         --batch {16} \
                         --epochs {30} \
                         --data data_fold_{fold}.yaml \
                         --weights yolov5s.pt \
                       #  --save_period 10 \
                         --project model-folds \
                         --name yolov5s-e-100-img-fold-{fold}
        print('############################################################################################\n')
train: Scanning /content/gdrive/MyDrive/Colab Notebooks/coco2/Yolo_cross/dataset_folds_0/labels/train... 2187 images, 0 backgrounds, 0 corrupt: 100% 2187/2187 [
train: New cache created: /content/gdrive/MyDrive/Colab Notebooks/coco2/Yolo_cross/dataset_folds_0/labels/train.cache
train: Caching images (1.1GB ram): 100% 2187/2187 [00:02<00:00, 1089.22it/s]
val: Scanning /content/gdrive/MyDrive/Colab Notebooks/coco2/Yolo_cross/dataset_folds_0/labels/valid... 547 images, 0 backgrounds, 0 corrupt: 100% 547/547 [00:44
val: New cache created: /content/gdrive/MyDrive/Colab Notebooks/coco2/Yolo_cross/dataset_folds_0/labels/valid.cache
val: Caching images (0.3GB ram): 100% 547/547 [00:00<00:00, 1084.86it/s]

AutoAnchor: 4.66 anchors/target, 1.000 Best Possible Recall (BPR). Current anchors are a good fit to dataset ✅
Plotting labels to runs/train/exp32/labels.jpg...
```

Fig 5: Model1 training configuration

```
Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
28/29      1.93G    0.01995   0.007786          0         28        416: 100% 137/137 [00:11<00:00, 11.92it/s]
           Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 18/18 [00:02<00:00,  6.22it/s]
             all        547        627      0.985      0.981      0.993      0.763

Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
29/29      1.93G    0.01963   0.007874          0         26        416: 100% 137/137 [00:11<00:00, 11.87it/s]
           Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 18/18 [00:03<00:00,  4.62it/s]
             all        547        627      0.989       0.98      0.992      0.775
```

Fig 6: Model1 batch training sample

In Fig 7, we visualized a sample of a training batch of model 1. This figure shows the detection of labels during training. As we had only 1 class in Yolo, there are 0 shown as class labels. Later, we visualized the training loss with other metrics in Chapter Result.

Fig 7: Sample of a training batch

4.8.2 Character Recognition

After detecting every possible license plate with the YOLO model, the model returns all detected cropped number plates from a frame. Then the model passes the result to our second model for character recognition.

We used mask R-CNN for this work. For segmenting and recognizing the license plate, we trained our model for Max_Iter value of 5000, a learning rate of 0.0001, and course, BATCH_SIZE_PER_IMAGE of 128 and 3917 images of the number plate, which consist of several instances of characters. The figure depicts the complete model configuration for this model. The three main components of a mask R-CNN model are the backbone network, the region proposal network, and Box Head. The backbone network extracts features from the input image. In this case, we use the Feature Pyramid Network (FPN) type of backbone. So, it pulls features of different sizes to make better guesses about anchor boxes of different sizes. In Fig 8, we depict the main components of a mask R-CNN model.

The Region Proposal Network finds object regions from multi-scale features. The objectness scores and anchor deltas are proposed (centres and sizes relative to the picture size). The Box Head layer receives standard inputs from extracted features and box proposals via a RoI Pooling layer. Box Head predicts bounding box locations and classifications. We trained our model for 30 epochs at 0.0001 to segment and recognize license plates.

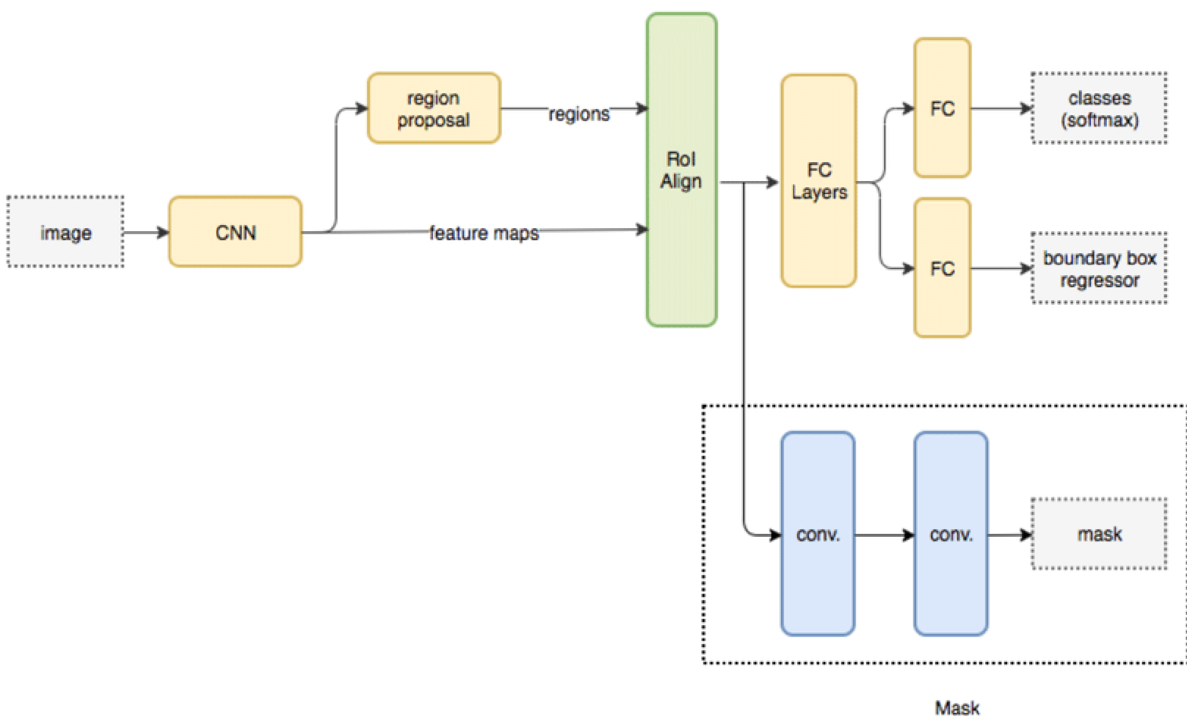


Fig 8: Architecture of the original Mask R-CNN framework

4.9 Authentication of the vehicle

The authentication of the vehicle is carried out using the approach that was suggested. We aim to locate a match between the vehicle's license plate number and the database. In section IV, we referred to a database that had already been predefined with the numbers of some registered automobile license plates. A CSV file is used to store the data contained within this database. Then, utilizing a program designed specifically for this purpose, we compared the database CSV file to the CSV file containing the testing vehicle number shown as Fig 9. If there is a match, the result will be as shown in Figure 16; if it is not, the system will recognize the car as an unauthorized one, as shown in Figure 17.

Licence Number	Owner	Contact	License Validity
Comilla-Ha 14-4113	Md Mahmudul Hasan Suzan	1963423770	31/12/2022
Dhaka Metro-Ga 23-0005	Asif Hasan	1642742857	31/12/2023
Dhaka Metro-Ma 11-4432	Abk Bhuiyan	1400141494	31/12/2024
Dhaka Metro-Ma 17-2197	Md Anisuzzaman	1900141499	31/12/2024

Fig 9: Sample Database of the registered Vehicles

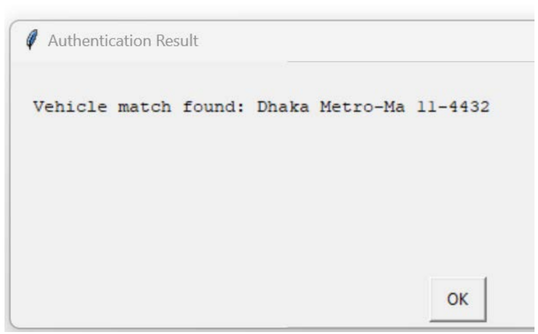


Fig 10: Positive Validation of Vehicles detection

Fig 11: Negative Validation of Vehicles detection

4.10 System Requirements:

For this research study and result analysis, we used software and hardware tools, which made out work smoother and faster. We used the following things in our research:

Hardware Requirements:

- Operating System (Above Windows 7)
- Hard Disk Minimum 500GB
- Ram Minimum 8GB+

Tools Requirements:

- Google Collab
- Python 3.8+ (coding language)
- Libraries: NumPy, Pandas, TensorFlow, Scikit-learn, detectron2.
- LabelMe for Data annotation
- Windows command prompt for testing.
- Microsoft Excel for saving results.

CHAPTER 5

EXPERIMENTAL ANALYSIS

In this section, we will discuss the experimental setup, findings, analysis, and discussion of the results, as well as the results discovered in previous studies. In addition to this, the findings that were discovered in earlier studies will be discussed.

5.1 Experimental Setup

All the experiments were conducted in the Google collab notebook environment utilizing photographs acquired from real-life conditions and under various illumination situations, such as in the afternoon, during cloudy and sunny weather, and in overcast conditions. The approach was applied to photos of automobiles, buses, and trucks from various metropolitan locations.

5.2 Dataset

516 testing images of vehicles with k-folding were used for the model 1 testing. Every license plate has a first line with two words and a letter and a second line with six numbers, totaling 4644 words and characters (alphanumeric symbols) that have been manually annotated with bounding boxes. The mask rCNN model, on the other hand, was assessed for 777 images.

5.3 Evaluation Metrics

There are many evaluation metrics available for evaluating the model's performance. We used k-fold cross-validation results with the metrics of recall, precision, mAP.5, and mAP.5:.95.

5.4 Result Analysis

Bangla characters were first taken out of photographs of license plates. The second phase involved training and testing on the recovered characters to ascertain the recognition accuracy. If the license plate is deformed, the algorithm will not perform properly (blurry image, unclear plate, etc.). In the event of license plate detection, the algorithm will not function properly if the contrast and brightness of the photos are subpar. Table 5 displays some distorted photos and the result of license plate detection and character extraction utilizing these images.

Following this, you will find a detailed description of the most important phases of the experimental results:

Table 2: CAPTURED NUMBER PLATES CHARACTERISTICS

Attribute of Vehicle Plate	**Characteristics**
Shape	Rectangular
Type	Double line (alphabet & numeric digits)
Background	Green/ Black
View	Front, rear
Distance	Maximum 10m
Skewed/Titled	Up to 10 degrees

5.4 Performance Evaluation

547 images of the license plate were used during testing. Compared to the front images, the angle for the left and right images was roughly 45 degrees. We used distinct evaluation metrics to assess the model. We used mAP, precision, and recall for the Yolo model to measure our performance. The result is shown below: In Fig 6, we visualized an example of a testing batch of model 1.

5.4.1 Vehicle Plate Detection

Our proposed model1 for vehicle plate detections performed outstanding results. This model has produced very good results. The model results are shown below in the Fig 10, Fig 11, Fig 12.

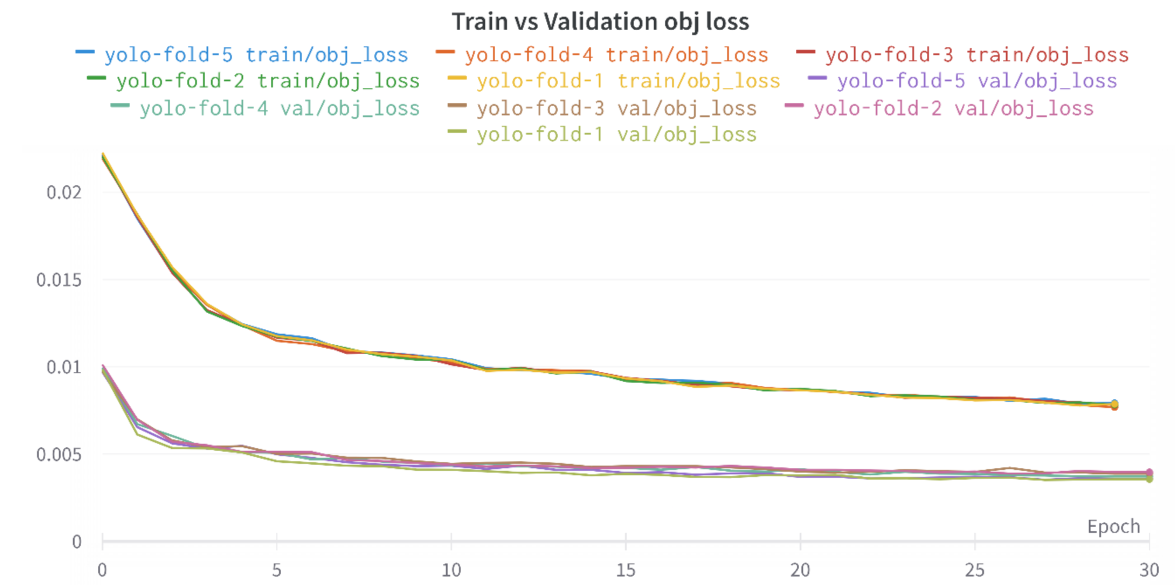


Fig 12: Model1 Object loss result

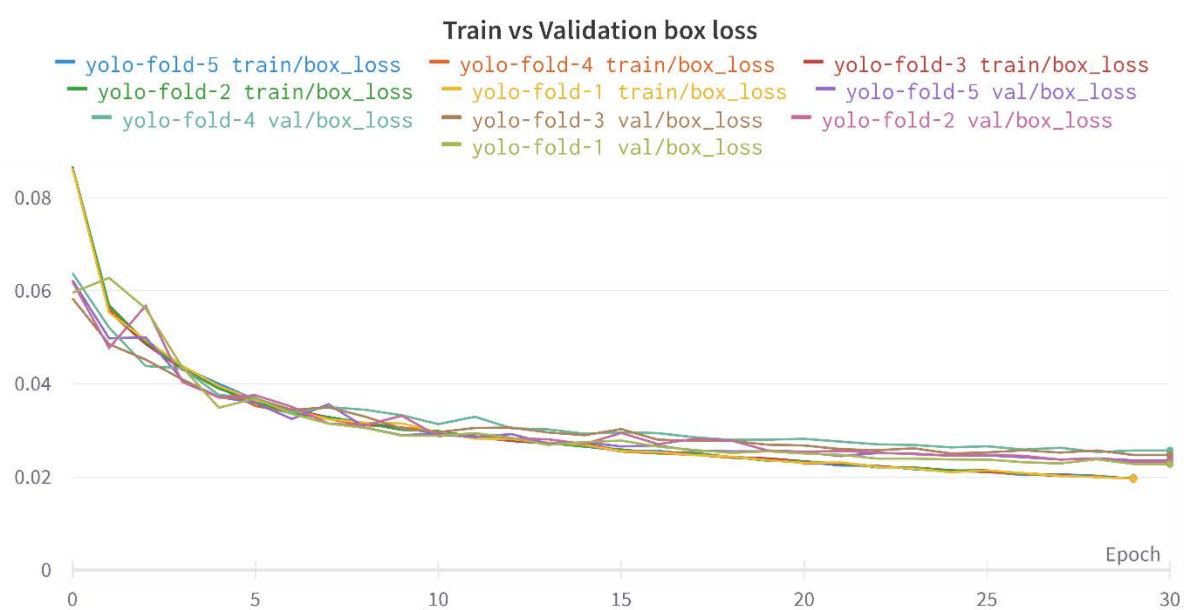


Fig 13: Model 1 box loss result

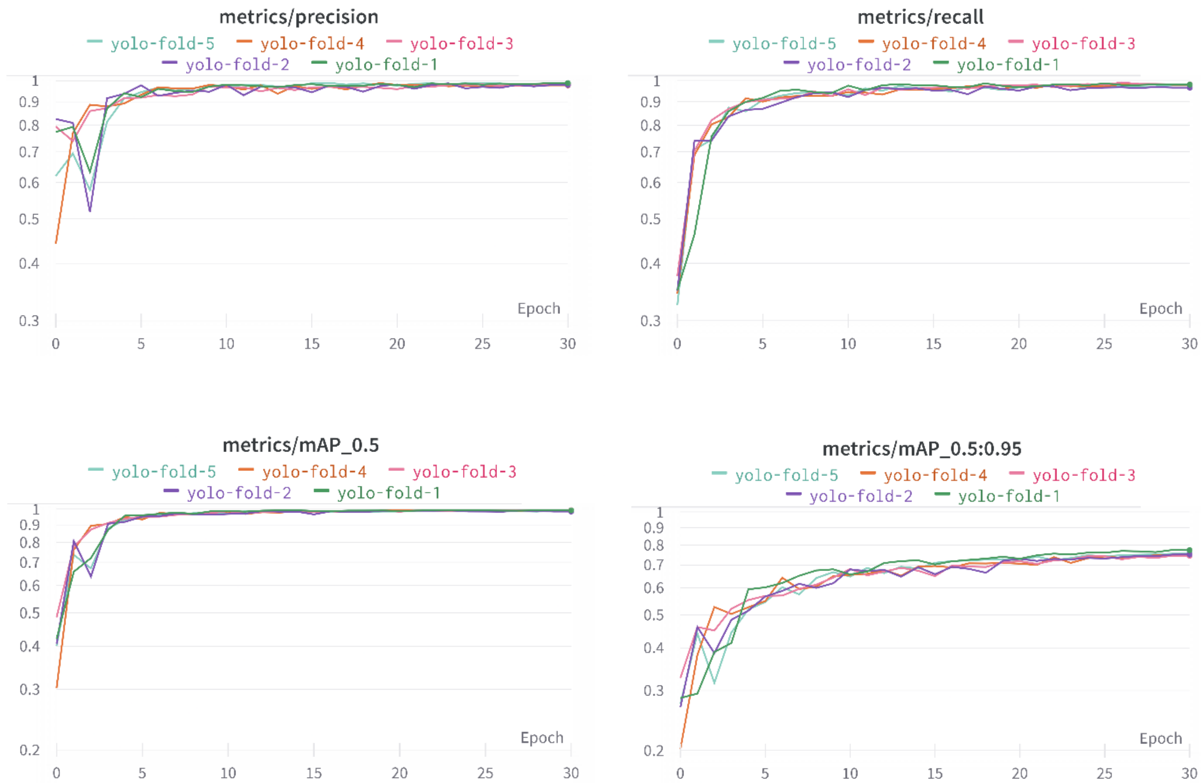


Fig 14: Yolo model performance evaluation

Table 3: K FOLD CROSS VALIDATION TEST RESULTS

Fold	Training Data	Images	Instances	Precision	Recall	mAP50	mAP50-95
1			627	0.989	0.981	0.992	0.774
2			625	0.979	0.966	0.984	0.751
3	2064	547	623	0.982	0.982	0.994	0.745
4			623	0.978	0.980	0.993	0.749
5			620	0.985	0.977	0.990	0.755

The following result is shown in Table 3, and the above graphs. Based on the 5 k-fold cross-validation result, It has been observed that the model produced a very good precision result of .989 in fold 1, top recall of .982 in fold 3, best mAP50 in folding 4 and best mAP50-95 in folding 1. Following this result, it is clear that using this trained Yolo model will produce excellent results in real-life use. We also observed that the model could detect multiple vehicle plates from a single frame, as shown in Fig 1.

Fig 15: Yolo validation sample

Fig 16: Multiple plate detection result

5.4.2 Character Recognition

Following the completion of the training, the model is automatically saved as a weight file. After that, the model can be loaded with the data from this file, and predictions can be made. For inference, the DefaultPredictor class will be used instead of the DefaultTrainer. The same type of evaluation metrics is also used to evaluate the model2. The evaluation result on 777 images has also produced good character recognition results. Fig 15, Fig 16, Fig 17, we visualized the model performance of the mask R-CNN model.

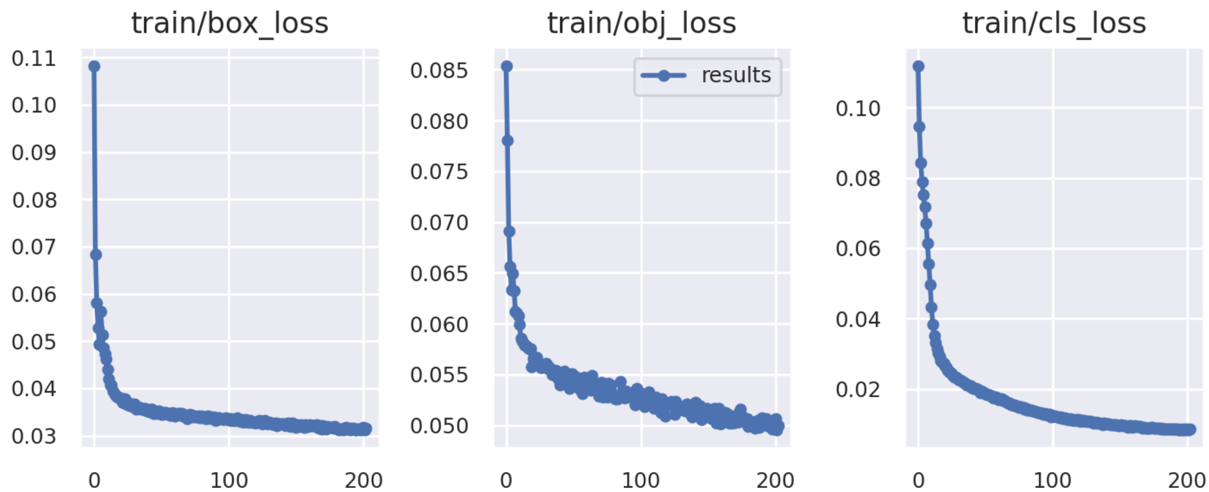


Fig 17: mask R-CNN training loss

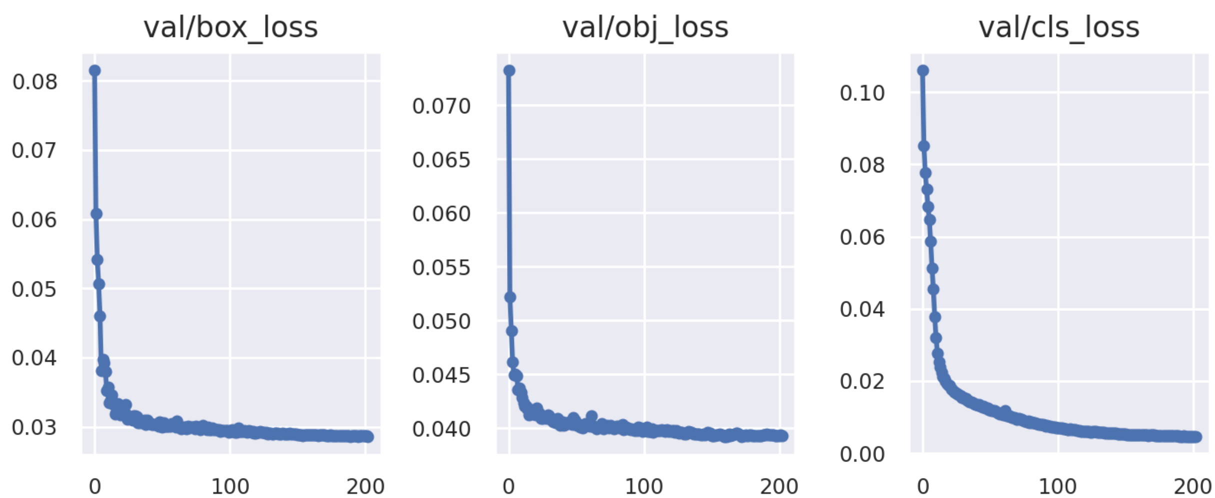


Fig 18: mask R-CNN validation loss

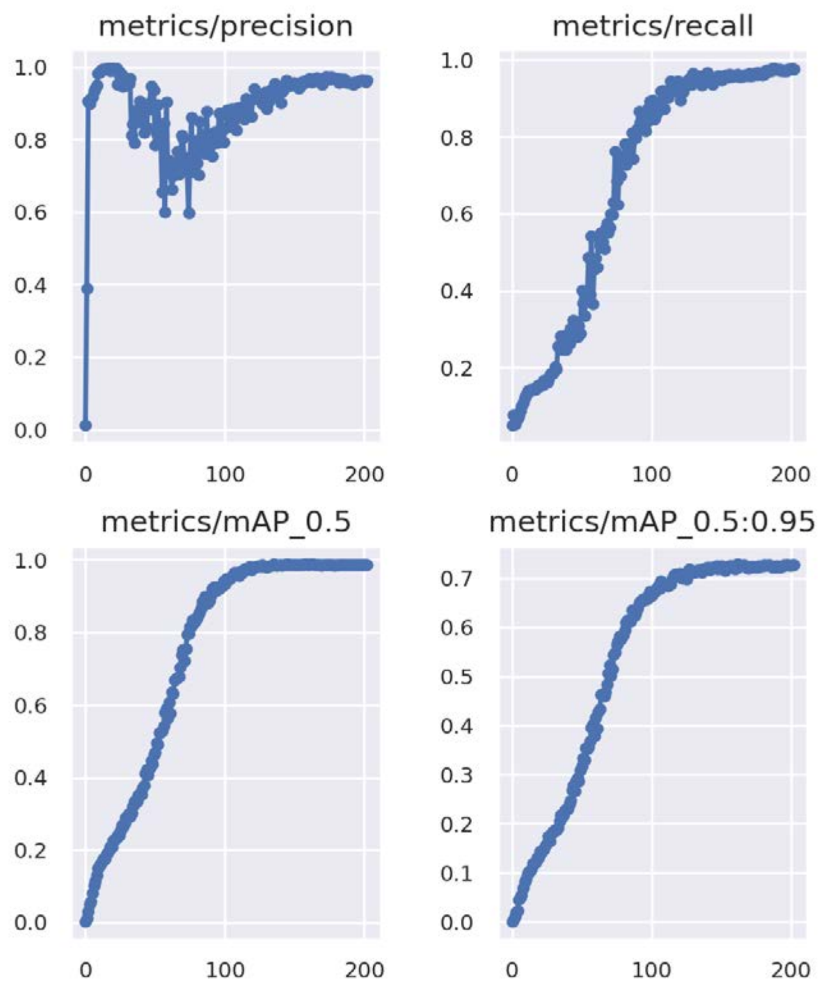


Fig 19: mask R-CNN model performance evaluation

Table 4: VALIDATION RESULT OF MASK RCNN

Model	Training Data	Validation Data	Precision	Recall	mAP_0.5	mAP_0.5:0.95
Mask R-CNN	3140	777	.965	.956	.989	.760

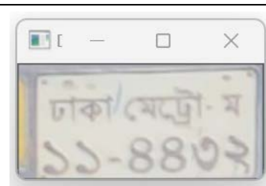
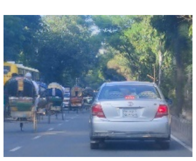
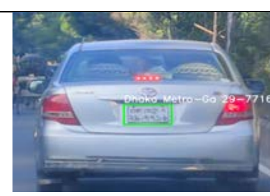
Following the Table 4, we can see the mask R-CNN model also achieved the highest precision of .965, Recall value of .956, mAP50 of .989 and mAP50-95 is .760. The amazing result of this metric is that mAP results in the threshold value 50-95 is .760. It means the model most of the detection is close to 100 AP. Fig 18 depict the combined output of the two model used.

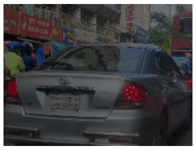
```
Photo width,height: 3554,2318. Detected plates: 2
License Detection: 1. Model1 prob: 0.90, Model2 results: Dhaka Metro-Ga 36-4985
License Detection: 2. Model1 prob: 0.86, Model2 results: Dhaka Metro-Ga 23-0005

C:\Users\mahmu\yolo_tess>
```

Fig 20: Combined Output of the Models

Table 5 has detailed outputs of some of the test images used. This table reflects our model's overall output and performance in various test conditions like blurry, low light ,etc.

Table 5: RESULT OF EXTRACTION, RECOGNITION AND AUTHENTICATION OF NUMBER PLATES

Test Image	**Environment**	**Extracted plate with model1**	**Recognition**	**Authentication**
	Bright			
	Skewed			
	Blur			

	Distant & Noisy			
	Low Light			

The models' output is saved in a CSV file for later use. The CSV file sample is shown in Table 6. This file keeps track of the vehicle plate number, the Date it appeared for detection, the location of the detection and the URL of the stored plate.

Table 6: RESULT OUTPUT SAVED IN DATABASE

Vehicle No	**Date Appeared**	**Location**	**License Plate**
Dhaka Metro-Ga 36-4985	1/13/2023	Dhaka	./model2_files\74fc84fc-3c92-4143-b46d-9e1fdfe08d84.jpg
Dhaka Metro-Ga 23-0005	1/13/2023	Dhaka	./model2_files\b9a7247e-1834-4d2d-9346-a7c0410dae28.jpg
Dhaka Metro-Ga 29-7716	1/13/2023	Dhaka	./model2_files\a01f98ab-dea2-43f2-9e4c-ab129584e350.jpg
Dhaka Metro-Ma 17-2197	1/13/2023	Dhaka	./model2_files\40deed44-8650-4932-9a6e-a4f5020e638c.jpg
Dhaka Metro-Ga 39-4564	1/13/2023	Dhaka	./model2_files\0b5127e0-566c-41e1-8fa4-7b967be26ba7.jpg

As mentioned in the literature review chapter, the proposed models did not achieve the expected results for Bengali license plates. Table 7, we compared these techniques to our model.

Table 7: COMPARISON OF OUR MODEL

Author	Method used	Total Data	Plate Detection	Character Recognition	License Validation
(Rahman et al., 2019)	SVM, RF, Deep CNN	1750	Not used	88%	Not used
(Joarder et al., 2012)	Edge analysis, SVM with template matching	N/A	80%	92%	Not used
(M. A. Islam et al., 2021)	Template matching & Connected Component Analysis	N/A	94%	95%	Not used
(Baten et al., 2014)	Template matching, Connected Component Analysis	N/A	Not used	N/A	Not used
(R. Islam et al., 2020)	Morphological operation and SVM	630	91%	94%	Not used
(Rokibul et al., 2016)	Genetic algorithms, Sobel filter, CCA	119	94%	84%	Not used
Our Model	YOLO, mask R-CNN	6528	98.2%	95.6%	100%

CHAPTER 6

CONCLUSION & FUTURE WORK

Although many proposed solutions are available for detecting and recognizing license plates, most models fail in terms of success rate in detection and speed. Most of the work has been done with the help of software tools and image-processing techniques, which is also very time-consuming. We proposed YOLOv5 for vehicle plate detection in various conditions and a new, improved model mask-R-CNN character recognition. In the proposed method, an effective and fast algorithm is developed and implemented to detect, extract, recognize and authenticate Bangla license plates. Our model has outperformed other models used for Bangladeshi license plate detection and recognition. In addition, we collected a massive amount to help this type of future work improvement.

However, there are always many improvements left for this type of work. We believe a more diversified dataset for training our deep neural network model will yield better results for our test dataset. We can categorize more vehicle detection tasks without changing our model with a diversified dataset. In future work, We will also combine the speed estimation model to provide a complete system that automates road surveillance and e-traffic monitoring.

REFERENCES

[1] Bandyopadhyay, H. (2022). YOLO: Real-Time Object Detection Explained. https://www.v7labs.com/blog/yolo-object-detection

[2] Baten, R. A., Omair, Z., & Sikder, U. (2014). Bangla license plate reader for metropolitan cities of Bangladesh using template matching. 8th International Conference on Electrical and Computer Engineering, 776–779.

[3] Du, S., Ibrahim, M., Shehata, M., & Badawy, W. (2012). Automatic license plate recognition (ALPR): A state-of-the-art review. IEEE Transactions on Circuits and Systems for Video Technology, 23(2), 311–325.

[4] Ghosh, A. K., Sharma, S. K. D., Islam, M. N., Biswas, S., & Akter, S. (2011). Automatic license plate recognition (alpr) for bangladeshi vehicles. Global Journal of Computer Science and Technology.

[5] Huang, Y.-P., Chen, C.-H., Chang, Y.-T., & Sandnes, F. E. (2009). An intelligent strategy for checking the annual inspection status of motorcycles based on license plate recognition. Expert Systems with Applications, 36(5), 9260–9267.

[6] Islam, M. A., Chowdhury, G. M., & Haque, M. N. M. (2021). Bangla License Plate Detection, Recognition and Authentication with Morphological Process and Template Matching. 2021 2nd International Conference for Emerging Technology (INCET), 1–6.

[7] Islam, R., Islam, M. R., & Talukder, K. H. (2020). An efficient method for extraction and recognition of bangla characters from vehicle license plates. Multimedia Tools and Applications, 79(27), 20107–20132.

[8] Jiao, J., Ye, Q., & Huang, Q. (2009). A configurable method for multi-style license plate recognition. Pattern Recognition, 42(3), 358–369.

[9] Joarder, M., Alam, M., Mahmud, K., Tasnuva, A., Kawser, M., & Ahamed, B. (2012). Bangla automatic number plate recognition system using artificial neural network.

[10] Kanayama, K., Fujikawa, Y., Fujimoto, K., & Horino, M. (1991). Development of vehicle-license number recognition system using real-time image processing and its application to travel-time measurement. [1991 Proceedings] 41st IEEE Vehicular Technology Conference, 798–804.

[11] Kim, K. K., Kim, K. I., Kim, J. B., & Kim, H. J. (2000). Learning-based approach for license plate recognition. Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No. 00TH8501), 2, 614–623.

[12] Rahman, M. M. S., Mostakim, M., Nasrin, M. S., & Alom, M. Z. (2019). Bangla license plate recognition using convolutional neural networks (CNN). 2019 22nd International Conference on Computer and Information Technology (ICCIT), 1–6.

[13] Rokibul, Md., Hossain, S., Roy, S., Alam, N., & Jahirul, M. (2016). Line Segmentation and Orientation Algorithm for Automatic Bengali License Plate Localization and Recognition. International Journal of Computer Applications, 154(9), 21–28. https://doi.org/10.5120/ijca2016912220

[14] Sanyuan, Z., Mingli, Z., & Xiuzi, Y. (2004). Car plate character extraction under complicated environment. 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583), 5, 4722–4726.

[15] Sarfraz, M., Ahmed, M. J., & Ghazi, S. A. (2003). Saudi Arabian license plate recognition system. 2003 International Conference on Geometric Modeling and Graphics, 2003. Proceedings, 36–41.

[16] ultralytics. (2021). Augmentation. https://docs.ultralytics.com/FAQ/augmentation

[17] Wada, K. (2016). labelme: Image Polygonal Annotation with Python.

[18] Wen, Y., Lu, Y., Yan, J., Zhou, Z., von Deneen, K. M., & Shi, P. (2011). An algorithm for license plate recognition applied to intelligent transportation system. IEEE Transactions on Intelligent Transportation Systems, 12(3), 830–845.

[19] Zheng, D., Zhao, Y., & Wang, J. (2005). An efficient method of license plate location. Pattern Recognition Letters, 26(15), 2431–2438.

191-16-410

ORIGINALITY REPORT

19%	15%	14%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	4%
2	www.researchgate.net Internet Source	4%
3	Md. Ashraful Islam, Gulam Mahfuz Chowdhury, Md. Niaz Morshedul Haque. "Bangla License Plate Detection, Recognition and Authentication with Morphological Process and Template Matching", 2021 2nd International Conference for Emerging Technology (INCET), 2021 Publication	2%
4	phamvanvung.github.io Internet Source	1%
5	link.springer.com Internet Source	1%
6	"Proceedings of the International Conference on Big Data, IoT, and Machine Learning", Springer Science and Business Media LLC, 2022 Publication	1%

7	Ambareesh Laxman Shirwal, Prasenjit Deb, Chandana S Bhat, Adithya Herle PS, Abhiram Naik T. "Ad-hoc Network Based Surveillance Robot", 2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2022 Publication	<1 %
8	nmbu.brage.unit.no Internet Source	<1 %
9	Shuai Zhao, Dong Ming Zhang, Hong Wei Huang. "Deep learning–based image instance segmentation for moisture marks of shield tunnel lining", Tunnelling and Underground Space Technology, 2020 Publication	<1 %
10	romisatriawahono.net Internet Source	<1 %
11	Mohammad Rokonuzzaman, M. Abdullah Al Amin, Masum Howlader Kazi Main Uddin Ahmed, Muhammad Towhidur Rahman. "Automatic vehicle identification system using machine learning and robot operating system (ROS)", 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), 2017 Publication	<1 %
12	aran.library.nuigalway.ie	

	Internet Source	<1%
13	biopen.bi.no Internet Source	<1%
14	Mahmudol H. Tusar, Md. T. Bhuiya, Md. S. Hossain, Anika Tabassum, Riasat Khan. "Real Time Bangla License Plate Recognition with Deep Learning Techniques", 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), 2022 Publication	<1%
15	dspace.muhas.ac.tz:8080 Internet Source	<1%
16	www.mdpi.com Internet Source	<1%
17	res.mdpi.com Internet Source	<1%
18	www.neuroquantology.com Internet Source	<1%
19	C.N.E. Anagnostopoulos, I.E. Anagnostopoulos, V. Loumos, E. Kayafas. "A License Plate-Recognition Algorithm for Intelligent Transportation System Applications", IEEE Transactions on Intelligent Transportation Systems, 2006	<1%

Publication

20	e-space.mmu.ac.uk Internet Source	<1 %
21	openconf.spsr.ase.ro Internet Source	<1 %
22	Nazmus Saif, Nazir Ahmmed, Sayem Pasha, Md. Saif Khan Shahrin et al. "Automatic License Plate Recognition System for Bangla License Plates using Convolutional Neural Network", TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), 2019 Publication	<1 %
23	Redowan Chowdhury, Fazleh Rabby, M. Samiur Rahman, Md. Abdur Razzak. "Identification of Unauthorized Vehicles by License Plate Recognition Through Image Processing", 2021 5th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), 2021 Publication	<1 %
24	1library.net Internet Source	<1 %
25	www.atdm2010.org Internet Source	<1 %
26	"Computer Vision – ECCV 2018", Springer Science and Business Media LLC, 2018	<1 %

Publication

27	assets.researchsquare.com Internet Source	<1 %
28	scholar.archive.org Internet Source	<1 %
29	thescipub.com Internet Source	<1 %
30	www.imedpub.com Internet Source	<1 %
31	www.semanticscholar.org Internet Source	<1 %
32	Myat Nyein Chan, Thuzar Tint. "A Review on Advanced Detection Methods in Vehicle Traffic Scenes", 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021 Publication	<1 %
33	Advances in Computer Vision and Pattern Recognition, 2014. Publication	<1 %
34	Tuhin Hossain, Ahmed Ainun Nahian Kabir, Md. Ahasun Habib Ratul, Abdus Sattar. "Sentence Level Sentiment Classification Using Machine Learning Approach in the Bengali Language", 2022 International	<1 %

	Conference on Decision Aid Sciences and Applications (DASA), 2022 Publication	
35	cps-vo.org Internet Source	<1%
36	d197for5662m48.cloudfront.net Internet Source	<1%
37	www.science.gov Internet Source	<1%
38	www.scribd.com Internet Source	<1%
39	www.zionmarketresearch.com Internet Source	<1%
40	Khalil Mustafa Ahmad Yousef, Bassam Jamil Mohd, Yusra Abd-Al-Haleem Al-Khalaileh, Ahlam Hani Al-Hmeadat, Bushra Ibrahim El-Ziq. "Automatic License Plate Detection and Recognition for Jordanian Vehicles", Advances in Science, Technology and Engineering Systems Journal, 2020 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		