BS Project - FINAL REPORT
Comparative Analysis of Different Models and Techniques for Object Detection and Character Recognition

Mehul Paithane
Department of EECS
Indian Institute of Science Education and Research Bhopal
Telephone: +91 - 9685272624
Email: mehul18@iiserb.ac.in

***Abstract*—This paper uses a fraction of ANPR (Automatic Number Plate Recognition) and performs a comparative analysis of object detection and character recognition models. This system depends on many factors like: The type of model used and how robust that model is, to the type of camera used for image capturing. The weather conditions affect it as well. This paper shall only focus on 2-wheeler vehicles. Officials keep enforcing simple road laws like: Wearing Helmet while riding a 2-wheeler. To reinforce this, certain precautions are placed like the e-challan system that sends the driver a hefty fine if these laws are broken. To help make this possible, we take the help of different Machine Learning (ML) models like CNN , objection detection using openCV and YOLO object detection and, pair it with OCR (Optical Character Recognition) and densenet. This is a step towards making any city, smarter and making the job of law-enforcers more efficient and easy. We use the Dataset available at Kaggle - License Plate for training and testing the model. We get different accuracies for different models. This project will also tell the what the best combination of object detection with character recognition is.**

1. Introduction

Due to the rise in India's economic growth, we see a rise in the automobile industry and therefore a rise in the number of vehicles on the road. Here are a few statistics that indicate the same:
The table 1 below shows the Indian Automobile Market (IMA) by segments from 2017-2018 [1].

TABLE 1. INDIAN AUTOMOBILE MARKET AND MARKET SHARE (%) BY SEGMENT, 2017–2018

Type of vehicle	Total percent (%)
Commercial vehicles	3
Three wheeled vehicles	3
Four wheelers	13
Two wheelers	81

As a consequence of the increase in this number of vehicles, the number of auto-mobile related accidents have been increasing. Table 2 below shows that statistic [2].

TABLE 2. ROAD ACCIDENT STATISTICS OF INDIA (1990-2010)

Year	No. of road Accidents (in thousands)
1990	282.6
2000	308.3
2010	430.6

According to a news article, **6 two-wheeler drivers die every hour due to lack of helmet protection** [3]. To keep this under an over viewing system, we need a centralised structure that would hold the offenders responsible.
A successfull ANPR system requires additional hardware to maximize the accuracy like: condition of the number plate, no common standard for number plates, complex weather conditions, camera quality, camera mount position, day/night conditions and many more.
Due to this burst of automobiles, it is becoming harder to keep track of the offenders. A city works under certain rules and regulations that need to be followed. To enforce them certain methods and preventive measures are placed, challan system being one of them.
This project's main objective is to help law-officers' job by making things a little easier through catching defaulters using the help of machine learning and deep learning. The other important scope of this research project is to also perform comparative analysis of different ML techniques and see which one works the best. Even small improvements and comparisons of different models with results may help improve or change the system. This project will extract characters of licence plates using a different trained dataset. There are many pre-trained object - detection models available which can be changed a little to detect licence plates. General object - detection is mainly based on bounding box regression and a proper dataset with necessary annotations is required to train it.

2. Objectives

The objectives of this project are listed below:

1) The first objective is straight-forward, it is to make the job of Law-enforcers easier and efficient. Therefore the results have to be quick.
2) The Result of the end-semester report would be to compare different models in ML like openCV with OCR, use deep learning, densenet and see which one works better and would be the most efficient.
3) By geo-tagging the data, a spatial pattern can be extracted that shows hotspots around the city from whoch the data is collected.

3. Background

There have been many successfull attempts on ANPR, like the kaggle ANPR project in 2018 [4]. This system has been modified and improved upon from time to time.
A lot of papers on ANPR have been published that talk about it's versatility and efficiency like:

1) "Automatic Number Plate Recognition System (ANPR): A Survey" - Chirag Patel et al [5]. This paper concludes an overall overview of ANPR by using different size images of licence plates and different contrast extensions and median filtering of images. They use ANN to perform character segmentation.
2) "YOLO: Unified, Real- Time Object Detection" - Joseph Redmon et al [6]. This paper holds important information for us as it talks about YOLO and using this model for object detection. This model is able to detect objects in real time and is able to process 155 frames in a second.
3) "Object Detection with Deep Learning: A Review" - Zhong-Qiu Zhao et al [7]. This paper uses deep learning models like R-CNN for object classification and detection in 3 steps: informative region selection, feature extraction and finally classification.
4) "Compressed DenseNet for Lightweight Character Recognition" - Zhao Zhang et al [8]. This paper talks about using the advanced variant of C-RNN like dense convolutional network with temporal classification. To reduce the computing size of the weight blocks, they redesigned and redefined the internal features of the dense blocks.
5) "License Plate Detection and Recognition Using OpenCV–Python" - Vishal Sharma et al [12]. This paper talks about using openCV for image processing and computer vision for licence plate detection.
6) "Visualising spatial distributions" - M-J KRAAK. This paper focuses on GIS and how to visualise spatial data from different sources. Therefore manipulating and analysing the new data.
In India, VAHAN (National Vehicle Registry database of the ministry of road transport and Highways) and RTO hold the national vehicle registry database which can be linked to these models for further use.
The job of this paper is to do a compare these models based on their efficiency and rule out a winner.

4. Materials and Method

Data Acquisition:

1) **Helmet detection [9]:** This dataset consists of weights file trained with YOLOv3. It also consists of images of helmet with cfg files that can be used by openCV to detect helmets in the given images.
2) **Licence plate detection [10]:** Finding a dataset with annotations and images of 2-wheeler is difficult, so for licence plate we shall use the dataset with 4-wheelers.
3) **Spatial pattern:** For this part of the project, we will be using a dummy dataset and geotag those images.
    a) There is an existing dataset that contains images of vehicles and their annotation images of the license plates. This dataset is provided by the KAGGLE ANPR (Automatic Number Plate Recognition). This dataset consists a total of 433 images in PASCAL VOC format with bounding box annotations of the car license plates in the images.
4) **Character extraction [11]**: There are 2 datasets for this component of the code, one using CNN and one using densenet. Both are taken from kaggle. These datasets are somewhat similar: there are 35 classes of data ranging from all the alphabets and numbers and each class containing 1030 images. For pytesseract dataset, we use train and validation datasets that consists of 36 classes of 30 images in each.

The overall execution of this project is divided in 5 components, which we shall discuss in detail:

1) **Helmet detection:** I shall solely focus on 2 - wheelers in this project due their large number. Helmet detection would technically be the first objective and we would move further in case there is no helmet detected. This process will be achieved by using tensorflow and keras library and also by using YOLOv3 object detection model. About the models:
    a) YOLOv3: YOLOv3 is used for object-detection that uses features of CNN. It has 53 layers trained on imagenet and for the task of detection, 53 more layers are added.

2) **Licence plate detection:** A similar sort of model can be used for this purpose that was used in the previous component. As we want to do a comparison of different models for this, we would deploy different models to check their efficiency and rule out the best. About the models:
    a) CNN: we shall resize the dataset and split them into 2 subsections: train and test. The goal is to be able to train a CNN model and run it for 100 epochs - This would be enough to filter out the noise - which would be able to extract the images of the license plates from the images of the vehicles.
3) **Text conversion:** The next step would be to do character segmentation of licence plates using 2 different techniques and extract the text into printable format. This text may have some further uses. Also, we shall compare the techniques based on their accuracy of outputs. About the models:
    a) Densenet: We shall use pytorch, densenet uses blocks of neural network called dense blocks interconnected. Densenet is trained using imagenet, we shall only train the last 3 layers of the 4th dense block to train object detection.
    b) Pytesseract: We shall use openCV to extract the character bounding boxes by using functions like contouring, gaussian blur and thresholding. Then a sequential model is trained using tensorflow.keras, to predict segmented characters and compare it to the text output.
4) **Analysis of different models:** After deploying different models we shall see their accuracy and the amount of loss. They shall also judge them based on the time it took to run these models.
5) **Spatial pattern:** We shall plot the location of the people not wearing helmets in the city of bhopal and get hotspot regions of helmet defaulters.

5. Software and Tools

1) For dynamic programming experience, jupyter notebook and google colab is used as in some cases an external TPU (Tensor processing unit) is deployed for faster execution of the code.
2) For geo-tagging purposes, a software called Geosetter [13] and also Geoimgr is used [14].

6. Results

6.1. Object detection - licence plate: CNN

Let us look at the results for licence plate detection based on a CNN model are shown below: The given Model ran for 100 cycles (epochs) and through every cycle, the accuracy increased with the f-score and likewise the loss decreased. here are some of the images of the extracted licence plate numbers with bounding boxes:

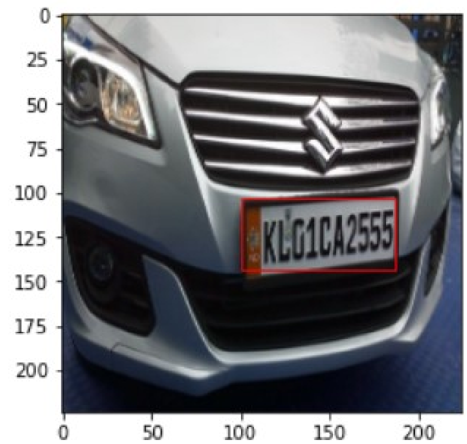

Figure 1. Image 2 of a car

Let us also see the mathematical accuracy of the model we have made:
The **Accuracy** of the CNN model is: **80.46 percent.**
The **Loss** of the CNN model is: **1.1 percent.** The table 3 below shows the f-score test results:

TABLE 3. (CNN MODEL TESTING RESULTS)

Loss	1.1231021 (%)
Accuracy	80.4597675 (%)

We have also analysed the relationship between the loss and accuracy variations as the cycles go on:

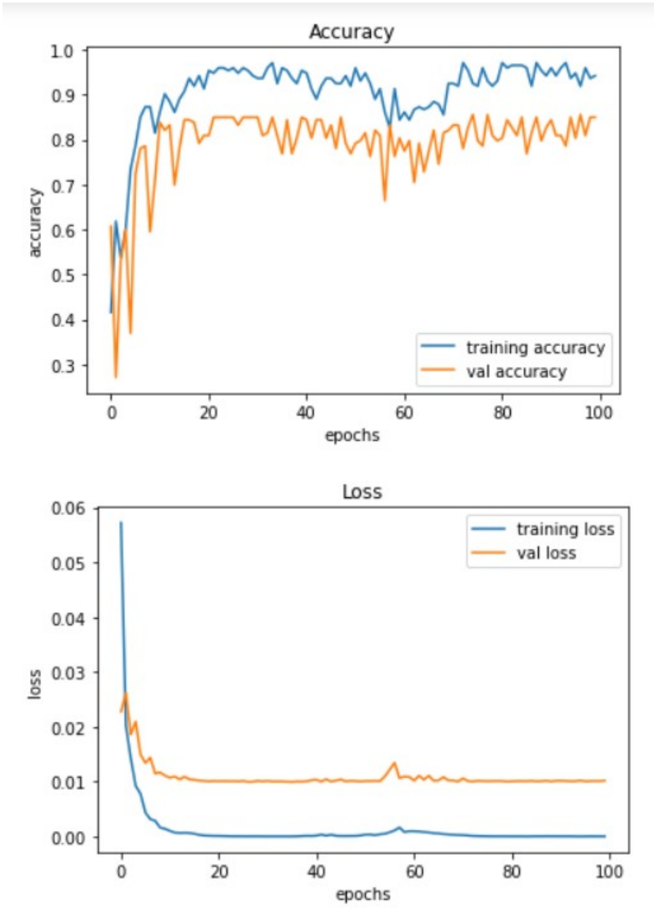

Figure 2. ACCURACY and LOSS vs epochs/cycles

6.2. Object detection - Helmet: YOLOv3

Let us first look at a sample image (from the kaggle dataset) that we would use to train and test this model on: After training the model by creating 2 networks from darknet for detecting 2 things:

1) Person on bike
2) Helmet

YOLOv3 model makes bounding boxes around the bike and the helmet and calculates the accuracy. We can see the final result
The accuracy is different for each image. For this random image, the helmet detection accuracy was: 91.54 (%) for helmet detection and 74.01 (%) for person on bike detection.

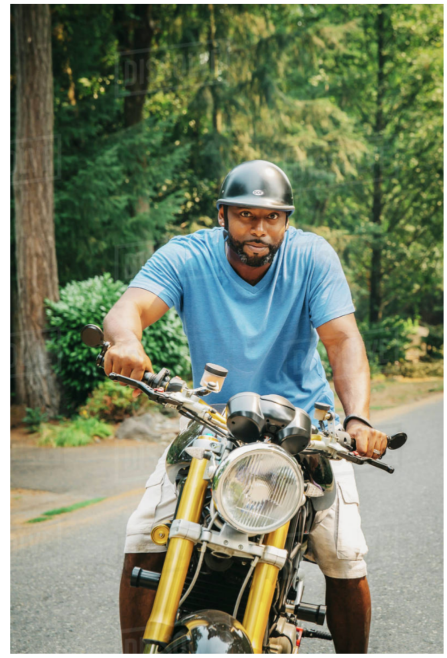

Figure 3. Sample image

Figure 4. YOLO output image

6.3. Character recognition - Pytesseract

We will call a test image to deploy pytesseract on by following the given steps:

1) Uploading a sample image:

Figure 5. PYT step 1

2) Grayscaling and contrasting the image:

Figure 6. PYT step 2

3) Using gaussian blur and contouring:

Figure 7. PYT step 3

4) Bounding boxes around objects and removing rectangular boxes:

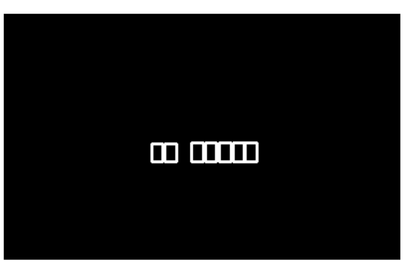

Figure 8. PYT step 4

Figure 9. PYT step 5

5) Now inserting the original image, on the bounding boxes, as can be seen above.
6) Segmenting and finding characters in the resulting images:

Figure 10. PYT step 6

7) The final result in text and the predictions are:

```
print(show_results())

PGMN112
```

Figure 11. Pytesseract results

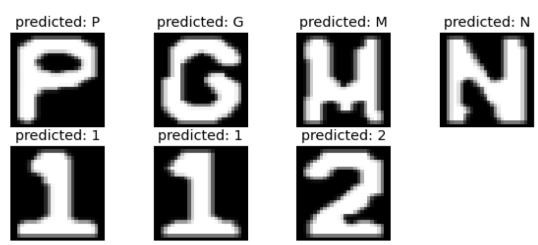

Figure 12. Predictions

The f1 score of the given model comes out to be: 0.9911.

6.4. Character recognition - Densenet

The accuracy of the densenet model comes out to be: **0.80**. Only the last 4 layers of the 4th denseblock was trained, therefore the accuracy is less.

Given below is the confusion matrix for the same:

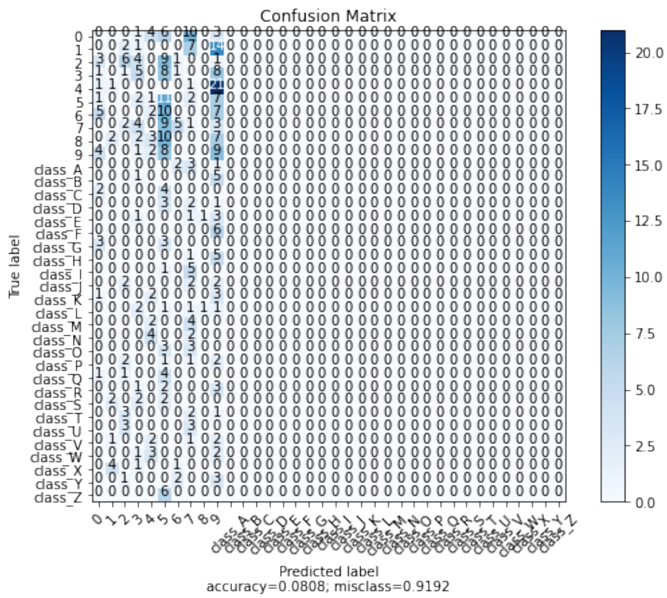

Figure 13. Confusion matrix

6.5. Character recognition - CNN

We train 2 models: training and testing. The accuracy and loss of the given 2 models are given below:

TABLE 4. CNN MODEL TRAINING AND TESTING ACCURACY

Model	Accuracy	Loss (%)
Training	0.965530	0.205412
Testing	0.991077	0.141119

Here is the graph showing the increase of accuracy of training and testing models:

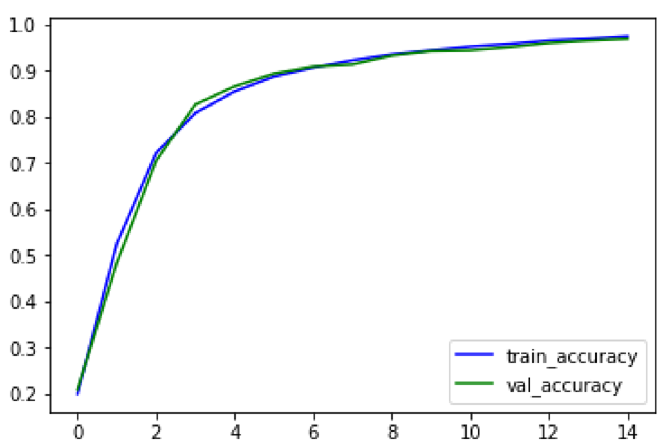

Figure 14. Training model accuracy increase

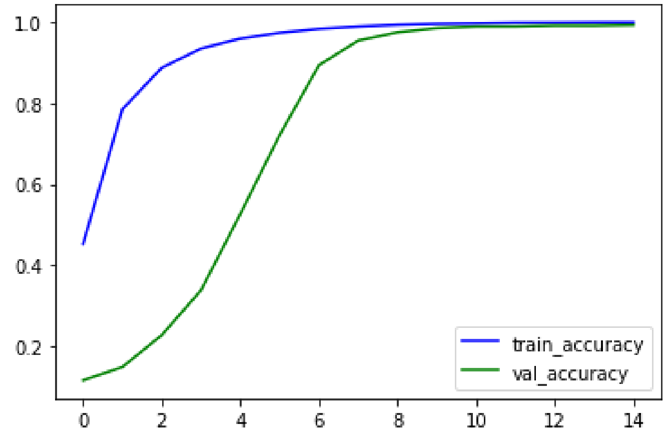

Figure 15. Testing model accuracy increase

6.6. Spatial pattern plotting

The data is taken from 3 different spots in Bhopal:

1) IISER bhopal gate.
2) Bairagarh.
3) DB mall.

Thye dummy dataset is then geotagged based on these 3 locations. Here are the results:

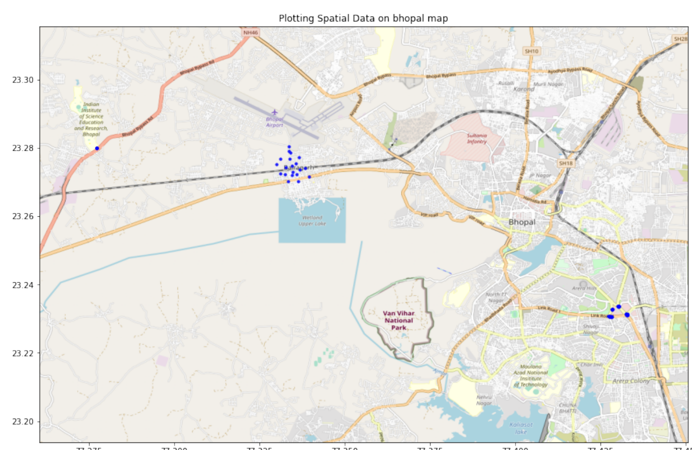

Figure 16. Spatial pattern output

The blue dots represent the number of helmet defaulters spread around bhopal city. the one dot in IISERB represents the images captured at the main gate.

Let us have a closer view of all the spatial patterns:

1) **IISERB gate:**

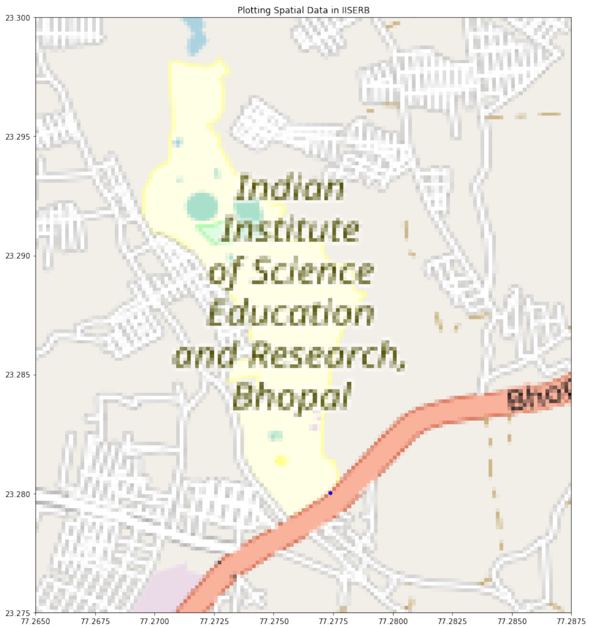

Figure 17. Spatial pattern output at IISERB gate

2) **Bairagarh:**

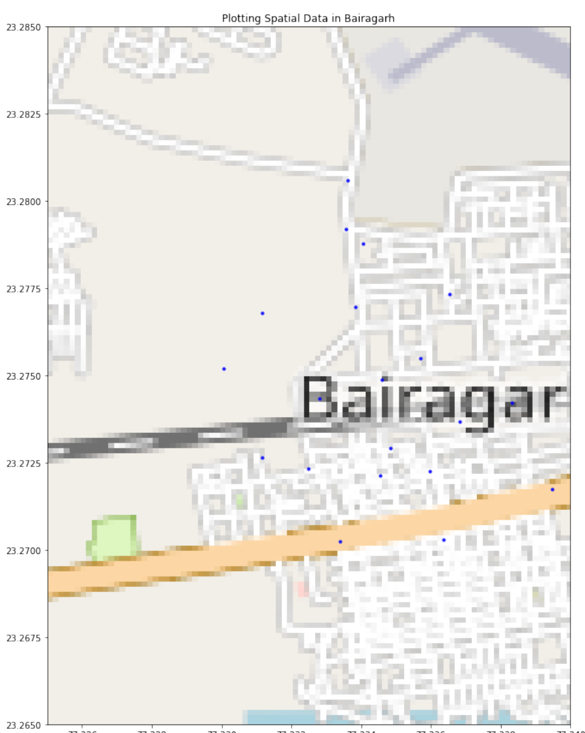

Figure 18. Spatial pattern output at bairagarh

3) **DB mall:**

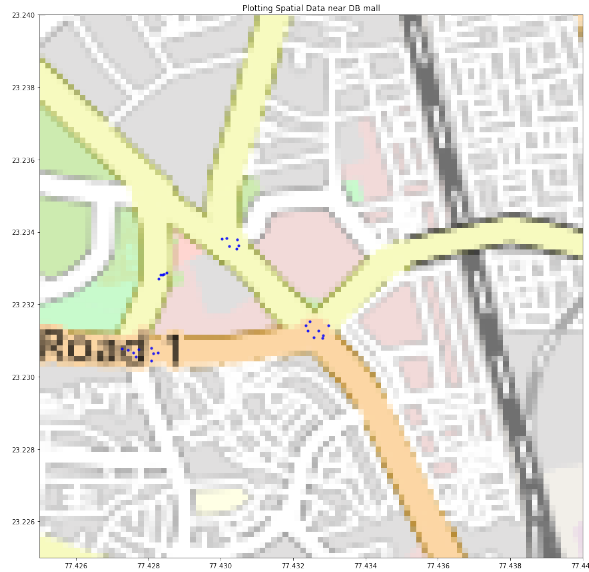

Figure 19. Spatial pattern output at DB mall

7. Discussions

This project can be used in many different ways and can have a bigger impact:

1) A portal can be created where citizens of any city can take pictures, while keeping their GPS location switched on. This can help officials keep track of the blind spots of the city. This would also help upgrade a city to smart city.
2) More accurate models can be created due to the non-existent template for licence plates.
3) Better high resolution cameras can be used for gathering dataset. Which can be used to train licence and text extraction models.

8. Conclusion

This system uses image processing techniques for recognition of the vehicle licence plate. The system works for wide ranges of licence plates and conditions.
We, can conclude that the YOLOv3 object detection model yielded the highest accuracy score of **0.9154**.
Similarly for text extraction and character recognition, we can conclude that openCV and pytesseract yielded the highest f1 score of **0.9911**. This report consists of preliminary results of licence plate detection models as well as character recognition models and compare them. It also consists of spatial pattern of helmet defaulters.

Acknowledgments

I am pleased to acknowledge **Dr. Vaibhav Kumar** for his invaluable guidance and feedback during the course of this project work. He has given me the proper guidance to help me understand important aspects of this project. I would also like to help Master Dhawal Patil for his assistance and support throughout this project.

References

[1] S. Miglani, “The growth of the indian automobile industry: Analysis of the roles of government policy and other enabling factors,” pp. 439–463, 2019.

[2] P. Mishra, “Vital Stats,” 03 2017. [Online]. Available: https://prsindia.org/policy/vital-stats/overview-road-accidents-india

[3] Prachi Salve, IndiaSpend.com, “In India, six two-wheeler riders die every hour in road accidents,” 12 2020. [Online]. Available: https://scroll.in/article/981989/in-india-six-two-wheeler-riders-die-every-hour-in-road-accidents

[4] K. Kaggle, “Car License Plate Detection,” 06 2020. [Online]. Available: https://www.kaggle.com/datasets/andrewmvd/car-plate-detection

[5] C. Indravadanbhai Patel, “ANPR : A survey,” 05 2013. [Online]. Available: https://www.researchgate.net/publication/236888959$_ANPR$

[6] J. Redmon, “YOLO: Unified, real- time object detection.” [Online]. Available: https://arxiv.org/pdf/1506.02640

[7] Z. Qiu-Zhao, “Object Detection with Deep Learning: A Review,” 04 2019. [Online]. Available: https://arxiv.org/pdf/1807.05511

[8] Jan 2019. [Online]. Available: https://arxiv.org/pdf/1912.07016