# *Traffic Rules Violation Detection using Deep Learning*

[1]Aniruddha Tonge
*BE – Computer, VIIT*
Pune, India
aniruddha.tonge@viit.ac.in

[2]Shashank Chandak
*BE – Computer, VIIT*
Pune, India
shashank.chandak@viit.ac.in

[3]Renuka Khiste
*BE – Computer, VIIT*
Pune, India
renuka.khiste@viit.ac.in

[4]Usman Khan
*BE – Computer, VIIT*
Pune, India
usman.khan@viit.ac.in

Dr. L. A. Bewoor,
Assistant Professor, Department of Computer Engineering VIIT, Pune
laxmi.bewoor@viit.ac.in

***Abstract*—— In order to ensure safety measures on roads of India, the identification of traffic rule violators is highly desirable but challenging job due to numerous difficulties such as occlusion, illumination, etc. In this paper we propose an end to end framework for detection of violations, notifying violators, and also storing them for analyzing and generating statistics for better decision making regarding traffic rules policy. In the proposed approach, we first detect vehicles using object detection which is performed using YOLO, and then accordingly each vehicle is checked against appropriate violations viz. not wearing a helmet, violation of crosswalks. Helmet violation is detected using a CNN (Convolutional neural network) based classifier. Crosswalk violation is detected using Instance Segmentation by Mask R-CNN architecture. After violations are detected, vehicle numbers are obtained of respective violators using OCR, and violators are notified. Thus an end to end autonomous system will help enforcing strong regulation of traffic rules.**

***Keywords— Computer Vision, Deep Learning, Convolutional Neural Networks, Image Classification, Object Detection, Image Segmentation***

## I. INTRODUCTION

Between 2001 and 2017, only due to road accidents, india has experienced about 20.42 lakh fatalities and 82.30 lakh people had been injured. According to the data given by the ministry, DIU has found that two-wheelers are hit the most in 2017 road accidents. Two-wheeler accidents make up 33 percent of the total deaths in road accidents. 4 two-wheeler users lose their life every hour just because of not wearing a helmet.[1]

Two-wheeler is a highly common mode of transport, but due to less protection, there is a high risk involved. It is highly advisable for bike-riders to use a helmet to reduce the danger involved. Observing the importance of the helmet, governments have made riding a bike without a helmet a punishable offense and have instructed to prosecute the violators. However, current video surveillance technologies are passive and require substantial human assistance. Typically, these systems are inefficient due to the presence of human intervention where efficiency tends to decline over a long period of time due to human fatigue. Also, manual intervention for capturing helmet violation is inefficient because out of many two-wheeler riders violating only a single vehicle rider can be caught at a time. Four wheelers and two-wheeler riders violate crosswalk at signals. Manual intervention for detecting crosswalk violations turns out to be inefficient and would lead to traffic jams due to the interruption of the flow of traffic. Detecting crosswalk violations through CCTV cameras would be highly beneficial as the flow of traffic wouldn't interrupted and multiple violations can be detected simultaneously.

**Contributions.** In this paper, we propose an end to end system for detection of traffic rules violations namely not wearing helmet and cutting the crosswalk using advanced computer vision techniques like object detection, image classification, and image segmentation on CCTV footages. Vehicles are first detected using Object Detection Algorithm (YOLO), then the detected vehicles are checked against violations. Helmet violation is captured with the help of a CNN based Image Classifier. Crosswalk violation by a vehicle is determined using Image Segmentation employing Mask R-CNN architecture. Then if any violation(s) are identified, the vehicle number of the violator is obtained using OCR and a notification is sent to the violator. The proposed methodology focuses on an end to end autonomous system right from capturing the image frames to penalizing the violator.

## II. LITERATURE SURVEY

The existing system in India used for the detection of violations and punishing the violator is manual. The CCTVs at the signals record the traffic all day long. These real-time footages are monitored by the traffic police team in the control room and when a violation is observed by the team, the license number plate is noted and a screenshot is taken for proof. Then the violation is entered into the database and the violation notification is sent to the violator thus charging them fine. As this process is manual, it becomes a tiring job for the team to continuously monitor the screens and not let any violation go undetected through such huge traffic in the country. Hence automation of this process is necessary. This paper will discuss various methods that are already proposed by different researchers.

The paper focuses on the following major modules:

A. Vehicle Detection

B. Helmet Classification

C. Crosswalk Violation at Traffic Signals

D. License Plate Recognition

### *A. Vehicle Detection*

The first step after obtaining an image frame from the video is to detect all the vehicles in the image. Prem Kumar Bhaskar, et al.[8] proposed to develop an algorithm for vehicle recognition and tracking using Gaussian mixture model and blob detection methods. First, it differentiates the foreground from the background in frames by learning the features that represent the background. To detect the moving object correctly and remove the noise, some morphological operations have been applied. However, the drawback of this method is that the blob detection requires successive image frames for detection. Kunal Dahiya, et al. [9] proposed an approach that first detects all objects from surveillance videos using background subtraction and object segmentation. Then Features are extracted from it using methods like Local Binary Patterns [LBP], Histogram of Oriented Gradients [HOG], Scale Invariant Feature Transform [SIFT], and finally, a binary SVM classifier is used to classify the object as bike rider or not. Asım, et al in [15] proposes that fast RCNN is another approach of vehicle detection, in which the input image is given to CNN to produce a convolutional feature map, the region of interest feature vector is generated which predicts the class of the given region.

### *B. Helmet Classification*

After the objects i.e. the vehicles are detected, the next important task is to identify whether riders are wearing a helmet. J. Chiverton[2] proposed an approach by considering the reflective properties of a material. In this approach, the approximate region of the head of the rider is isolated, which is used as a feature that is derived from the histogram operations applied to the region. It is observed that the top of the heads is brighter than the bottom of the helmet. These properties are then used as an input feature to SVM with a linear kernel to classify the riders whether the rider is wearing a helmet or not. An issue with this approach is that the methodology is based on the concept of reflection which may vary for different surfaces and also ample amounts of light should be present to have a significant difference between upper and lower surfaces for detection. Gomathi, et al. [3] proposed a way for classification of the helmet using IR sensors, Arduino microcontroller, power supply, and RF Transmitter. The IR sensors will help in monitoring the riders. However, this approach involves overhead of hardware which may need to be replaced from time to time making this approach to have a high maintenance cost. The possibility of hardware being damaged is also high.

### *C. Crosswalk Violation at traffic signals*

One of the violations other than not wearing a helmet is vehicles that have cut the crosswalk at signals. Aaron Christian P. Uy, et al. in [5], proposes a method of detecting whether the crosswalk is cut or not by using Genetic Algorithm. In this methodology, at first, the obscurity of the vehicle is found with respect to the background image after which Genetic Algorithm is applied to detect violations if any. One of the drawbacks of this approach is that the background image is to be maintained to detect violations and if there is a change in the background then the background image must be updated accordingly i.e. constant updating of the background image is needed.

The proposed system by Samir Ibadov, et al. in [4], detects the crosswalks using the alpha channel method in canonical form which uses colors to mark the crosswalk and then detect the vehicles by faster R-CNN and then finally compare the frames and check if they overlap or not to classify whether the crosswalk is cut by the vehicle at the signal or not. One drawback of the alpha channel method is that it cannot represent semi-transparency. That's why color keying cannot produce a smooth blend on the edge of an opaque or transparent area.

### *D. License Plate Recognition*

After detecting the vehicles using object detection, and classifying them according to the violations, the number plate of the corresponding vehicle has to be recognized for storing violations and further processing. Amey Narkhede, et al. in [6] proposed an approach involving Edge detection, Hough transformation, and K-Nearest neighbor. The Canny Edge detection is used to find the boundaries of the objects, then Hough transformation is used for feature extraction in which lines are detected, this is then followed by the K-Nearest Neighbor which recognizes the segmented characters. Ana Riza, et al.[7] proposed an approach in which plates are recognized using Canny Edge Detection followed by Contour Detection and evaluation using mathematical formulas. After detecting the edges within a vehicle image, the contours are located and rectangular contours are searched.To detect valid plates, parameters like aspect ratio, area, and the length of the diagonal of the contour are considered. Pooya Sagharichi Ha, Mojtaba Shakeri in “License Plate Automatic Recognition based on Edge Detection” indicates a way to recognize the licence plate by the use of Canny edge detection. Subsequently, character extraction is carried out using template matching[14]

One major drawback using canny edge detection is, it detects all the edges in the image and we need to then carry out further processing to select only those edges which actually belong to the number plate. This becomes computationally expensive and may also produce incorrect results.

Above are the different proposed methods but currently, there is no end to end system used for automatic detection of violations. We therefore propose an end-to-end autonomous system that would use computer vision concepts to detect road traffic violations through existing CCTV cameras.

## III. PROPOSED METHODOLOGY

In this section, we will explain the proposed end to end system. This system consists of four major components. The system architecture is shown in figure 3.1. In the proposed system we first give an image frame from CCTV footage as input to the system then perform vehicle detection using Object Detection. Object Detection is carried out using YOLO (you look only once) for detecting vehicles in the image frame. After detecting vehicles, individual vehicles are cropped out using the coordinates obtained from bounding boxes given by Object Detection Algorithms. Now an individual vehicle is checked against different violations. Violations included in this proposed system are Helmet Violation (Two wheeler rider not wearing a helmet) and Crosswalk Violation (Violating Zebra Crossing). Two-wheeler vehicles will be checked against Helmet and Crosswalk Violation and Four Wheeler vehicles will be checked against Crosswalk Violation. Helmet violation is detected using CNN (Convolutional Neural Network) based classifier which works well on visual data. Crosswalk violation can be detected using Mask RCNN where instance segmentation helps in comparing coordinates of the bottom of tyre of the vehicle with those of detected crosswalk. Once violation(s) are detected for a vehicle then the number plate of the corresponding vehicle is detected using Object Detection. Again YOLO is used for detection of the number plate of a vehicle. OCR (Optical Character Recognition) is used to obtain license number from the number plate. Vehicle users are notified with associated violations and violations are inserted into the database. The database can be used to obtain statistical analysis on traffic rules violations that previously occurred. Now we will look at each individual module in detail.

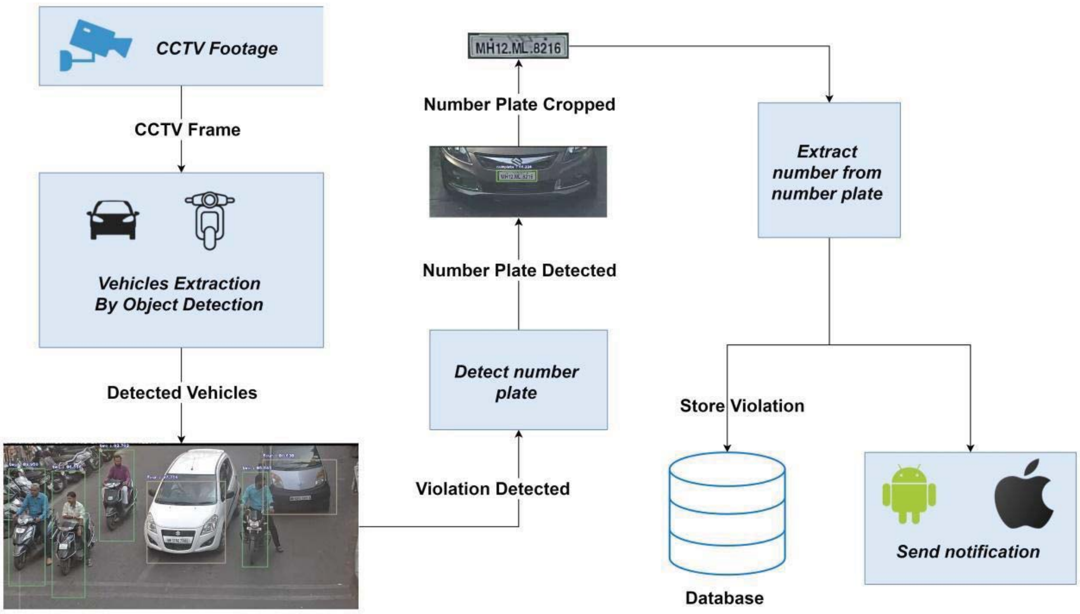


**Figure 3.1** Architecture Diagram for proposed system

Our proposed methodology for each module is given below:

### A. Vehicle Detection

When the signal goes red, a frame is obtained from the CCTV camera. Now, the frame is passed as an input to the object detection module for vehicle detection. We are using YOLO (You Only Look Once) for this purpose. YOLO is an object detection algorithm [10]. YOLO predicts the coordinates of bounding boxes directly using fully connected layers on top of the convolutional feature extractor. Predicting offsets instead of coordinates simplifies the problem and makes it easier for the network to learn. YOLO is better than other object detection algorithms like R-CNN, fast R-CNN, and faster R-CNN as they use pipelines to perform the detection of objects which incorporates multiple steps. Hence these algorithms are slow to run and hard to optimize as each individual component should be trained separately. YOLO can do this task with a single neural network, thus outperforming the above algorithms.

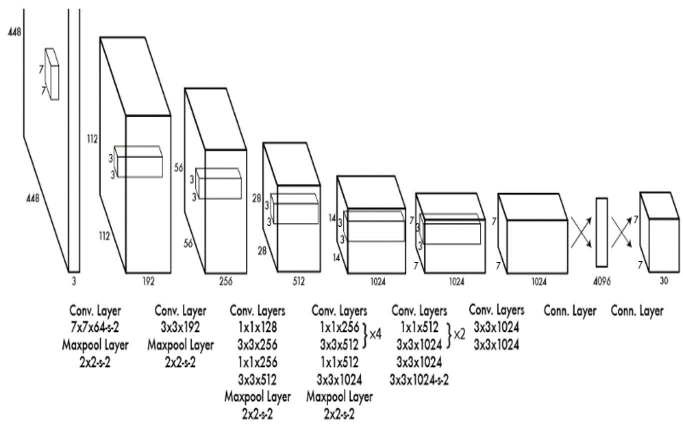


**Figure 3.2** Architecture of YOLO [10]

Figure 3.2 shows the architecture of YOLO which has 24 convolutional layers and 2 fully connected layers. 1x1 alternate convolutional layers reduce the space of the features from the layers preceding them [10].

The pre-trained model of YOLO[12] detects vehicles and persons as separate objects. Hence we built a custom model that detects the vehicles as well as the riders as a single object within the same bounding box as shown in figure 3.3. The custom model was built by using weights from the pre-trained model and the concept of transfer learning. Vehicle detection is performed using object detection.

**Figure 3.3** Object detection with YOLO

In fig. 3.3, it is seen that the objects are detected and classified in "two" and "four" classes. The green bounding box represents class "two" and the pink bounding box represents class "four".

*B. Helmet Classification*

After vehicles are detected from images, Two-wheelers are considered for the Helmet Classification task. Individual vehicles are cropped from the cctv frame with the help of corresponding bounding box, if the vehicle belongs to two wheeler class they are further cropped where we consider the upper half of the individual vehicle (upper 50% of vehicle frame). Now the upper half is considered as statistically, heads (helmets) are located in the upper half of the image. By considering the upper half of the image computational complexity of the system could be reduced. This upper half is given as input to the classifier, which outputs whether the image belongs to helmet or non-helmet class. Fig 3.4 depicts the flowchart for helmet classification.

The classifier used for helmet classification employs a CNN architecture as shown in fig. 3.5. It is a 12 layered architecture that includes various layers such as 5 Convolutional layers using ReLu as an activation unit, 4 pooling layers, and a single dense layer using softmax for classification into two classes. CNN is preferred over other methods as it is better at extracting visual features from image data.

Fig. 3.6 represents the feature maps that are generated from output of convolutional layer. These feature maps illustrate that the CNN learns common hidden features and structures among helmets and heads in the training set while training, thus being able to distinguish between helmet and non-helmet class.

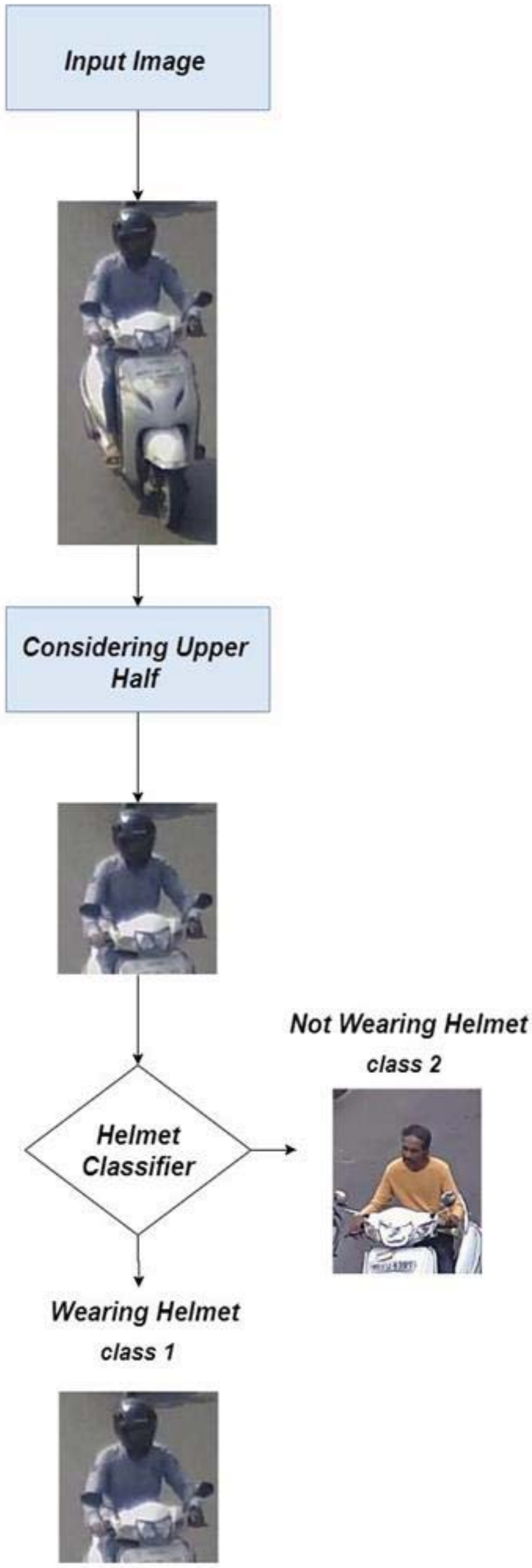


**Figure 3.4** Helmet classification using CNN. Input image is classified as "wearing helmet" or "not wearing helmet".

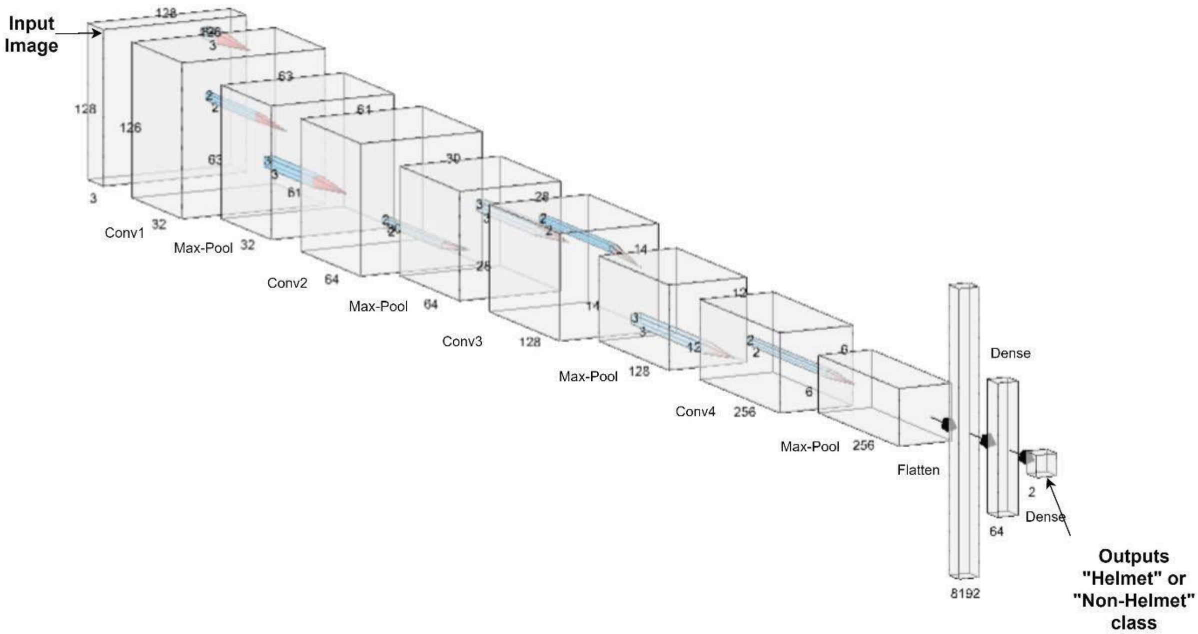


**Figure 3.5** Architecture of CNN for Helmet Classification

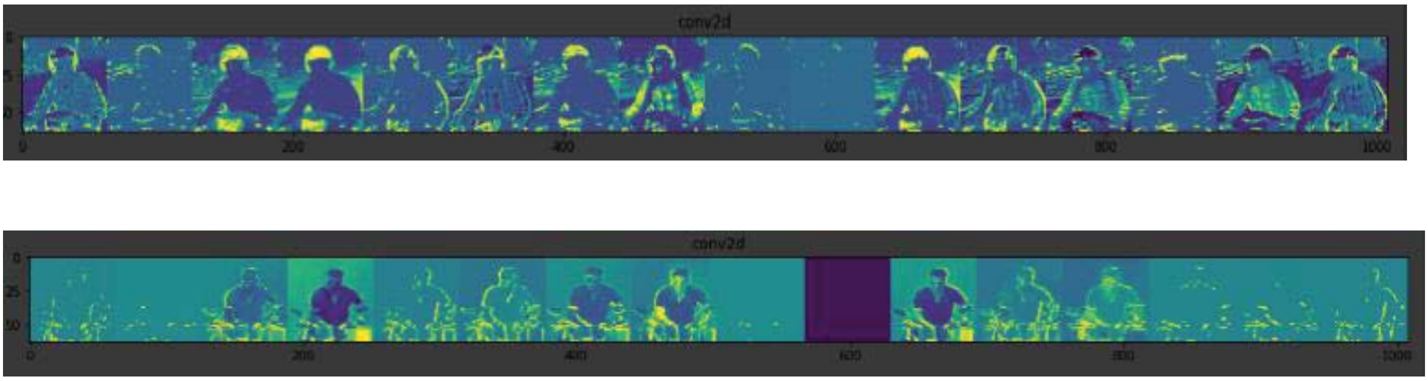


**Figure 3.6** Visualization of the trained representation by CNN for the classification of helmet and non-helmet class

### *C. Crosswalk violations at traffic signals*

Crosswalk violation is detected using instance segmentation. Mask R-CNN is an extension to faster R-CNN which includes a segmentation mask along with bounding boxes of objects segmented [11]. Segmentation gives us a pixel-wise mask for each object. Instance segmentation model was trained for two classes namely vehicle and crosswalk. Training mask R-CNN requires significant amount of time, hence we made use of pretrained mask R-CNN weights [13].After segmentation is performed we obtain a mask for each vehicle and zebra in the image along with their bounding boxes.

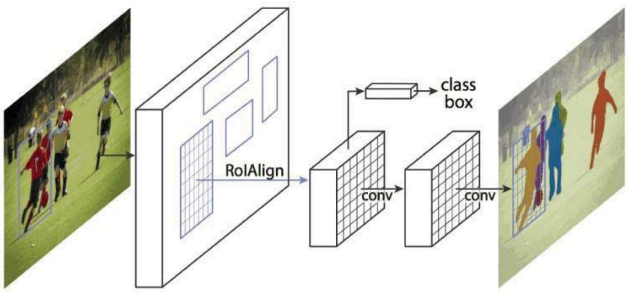


**Figure 3.7** The Mask R-CNN framework for instance segmentation [11]

Crosswalk violation occurs if a vehicle is present on a crosswalk (Zebra crossing). First, we obtain coordinates of the bottom of the tyre with the help of the bounding box of the respective vehicle. Now we check if the obtained coordinates overlap with the crosswalk, this can be achieved by comparing coordinates obtained with those of the bounding box of the crosswalk. We assume top-left corner of image frame as origin (0,0), Considering $(Zx_1,Zy_1)$ and $(Zx_2,Zy_2)$ as coordinates of bounding box of crosswalk of top-left corner and bottom-right corner respectively, $(Vx_1,Vy_1)$ and $(Vx_2,Vy_2)$ as coordinates of bounding box of a particular vehicle of top-left corner and bottom-right corner respectively. We compare $Zy_1$ with $Vy_2$, if $\mathbf{Vy_2 > Zy_1}$ where $\mathbf{Vy_2}$ will be y coordinate of bottom of tyre, then a violation has said to be occurred. Crosswalk violation is carried out for each vehicle by checking if the coordinate of the bottom of tyre lies within the crosswalk. It can be observed from fig 3.8 that a four-wheeler (masked yellow) and 2 two-wheelers (marked as green and purple) have violated crosswalk and their coordinates of the bottom of tyre lie over the crosswalk.

**Figure 3.8** Instance segmentation using Mask R-CNN

### *D. License Plate Recognition*

After detecting a violation of traffic rules, the task to be carried out is to get the vehicle number from the vehicle that violated traffic rules. First, we localize the number plate of the vehicle using Object Detection. Localization of number plates is carried out using YOLO[10] again. Localized license plates of the vehicles that violated any traffic rule are cropped out. The vehicle number is extracted by performing OCR (Optical Character Recognition) on the cropped number plate. In OCR character segmentation is carried out followed by template matching. Characters are segmented individually and then compared with characters in the database using template matching. Once the vehicle number is obtained the violation is stored in the database and the corresponding vehicle user is notified about the same through SMS or email.


**Figure 3.9** shows the detection of the number plate of the vehicle, then the characters are read using OCR

## IV. DATASET

The dataset used was prepared from cctv footage of various signals obtained from pune city's traffic police department. Dataset was generated from the cctv footage by obtaining image frames at a regular interval of approx 15 secs to avoid repetition of images. The images consisted of both sparse and dense traffic conditions, in order to include diversity in data.

TABLE I

DATASET USED FOR DIFFERENT MODULES

| Dataset | No. of images | No. of classes | Name of classes |
|---|---|---|---|
| Vehicle Detection | 525 | 2 | two, four |
| Helmet Classification | 1079 | 2 | helmet, non-helmet |
| Crosswalk Violation | 200 | 2 | vehicle, crosswalk |
| Number Plate Detection | 450 | 1 | numplate |

A total of 525 images were taken for vehicle object detection which consisted of various types of vehicles including bikes, cars, buses, and rickshaws. Out of 525, 420 images were considered for training and 105 for testing. A total of 1079 images were considered for the helmet classification, 492 of which were people wearing helmet and 587 without helmet. The images for helmet classifier were obtained by applying vehicle detection on the vehicle detection dataset and then selecting the vehicles with two wheeler class. Out of 1079 images 864 were considered for training and 215 for testing. For detection of zebra violation which employs image segmentation, 200 images were considered in which polygons were created around the objects to highlight its pixels and help prepare the masks for training mask-rcnn network. Out of 200,150 images were considered for training and 50 images for testing. Number. Number plate detection dataset was prepared by labelling individual vehicle's number plate considering both two wheelers and four wheelers. It consisted of total of 450 number plates, out of which 350 were used for training and 100 were used for testing.

## V. RESULT

In this section, we present experimental results of our modules viz. vehicle detection, helmet classifier and Image Segmentation. The experiments are carried on Ubuntu 18.04 machine consisting of i5 6th gen processor, 12 gb of ram and 256 gb SSD. Python 3.65 was used for evaluation of results along with libraries such as opencv-python for image processing, tensorflow as a Deep Learning framework (for training and prediction), numpy for mathematical operations and matplotlib for visualization.

1. Vehicle Detection:

Vehicle detection is performed using YOLO algorithm, which was trained on 420 images consisting of both the classes, two and four. This was evaluated against 105 test images considering metric as average precision per class and also calculating the mean average precision. The results for which are as follows:

TABLE II

RESULTS OF VEHICLE DETECTION

| Class | Average Precision |
|---|---|
| Four | 0.9406 |
| Two | 0.9594 |

The mean average precision(mAP) for the model is 0.95.

2. Helmet Classification:

Helmet Classifier performs image classification using CNN consisting of five convolutional layers with ReLu activation units, 4 max-pooling layers, and one fully connected dense layer with final softmax unit for classification into two classes. The proposed model architecture is evaluated using metrics such as precision, recall, accuracy, and F1-score on a test-set consisting of 215 images where 98 images belonged to helmet class and 117 belonged to non-helmet class.

TABLE III

RESULTS OF HELMET CLASSIFICATION

| Metric | Value |
|---|---|
| Precision | 0.8780 |
| Recall | 0.8925 |
| Accuracy | 87% |
| F1 score | 0.8852 |

3. Crosswalk Violation:

Crosswalk Violation employs use of mask-rcnn network [11]. This network is trained on 150 images with 2 classes namely vehicles, and crosswalk(zebra) using the concept of transfer learning. This network was then evaluated against 50 images each consisting of all of the 3 classes. Metric used for evaluation is mean Average Precision(mAP) is computed on a test-set consisting of 50 images.

TABLE IV

RESULTS OF CROSSWALK VIOLATION

| Metric | Value |
|---|---|
| mean Average Precision (mAP) | 0.96 |

4. Number Plate Detection:

Number plate detection is performed using YOLO again which was trained on 450 images . This was evaluated against 100 test images considering metric as average precision. The results for which are as follows:

TABLE V

RESULTS OF NUMBER PLATE DETECTION

| Class | Average Precision |
|---|---|
| numplate | 0.9305 |

## VI. CONCLUSION

The proposed approach employs the use of concepts such as YOLO, CNN, Mask R-CNN and OCR for automatic detection of traffic rules violations. It achieves the desired goal with precision and ease, but requires high computational power as it makes use of concepts like image segmentation and object detection. The advantage of proposed system is it has the ability to catch more number of violations as compared to the human intervened system. In addition, the proposed methodology provides an end to end autonomous system, when implemented would prove an upper hand in detecting violations. Thus strict regulations of traffic rule violations can be implemented resulting in better road safety and bring awareness among vehicle users.

## VII. FUTURE SCOPE

Future work for the proposed system can have a broad perspective. The scope can be extended to detect other violations of traffic rules that occur. This can include whether a four-wheeler driver is wearing a seat belt or not, detection of more than two persons riding a two-wheeler, detection of vehicles traveling against allowed direction (vehicles driving the wrong side), etc. Another issue in India is not all license plates have standardized format thus detection of fancy number plates could be a big help in standardizing number plates thus strengthening traffic rules discipline.

## ACKNOWLEDGMENT

This project was supported by the Police Commissioner, Pune. Thanks to the Traffic Police Department, Pune for providing us with dataset in form of CCTV footage of different signals across the city and explaining us the existing system in use for detection of traffic rules violations, which was a significant help in designing and developing proposed system.

## REFERENCES

[1] https://www.indiatoday.in/diu/story/two-wheeler-death-road-accidents-helmets-states-india-1602794-2019-09-24

[2] J. Chiverton, "Helmet presence classification with motorcycle detection and tracking," Intelligent Transport Systems (IET), vol. 6, no. 3, pp. 259–269, September 2012

[3] Gomathi, et al, "Automatic Detection of Motorcycle without helmet using IOT", South Asian Journal of Engineering and Technology

[4] Samir Ibadov, et al, "Algorithm for detecting violations of traffic rules based on computer vision approaches", MATEC Web of Conferences 132, 05005 (2017)

[5] Aaron Christian P. Uy, et al, "Automated Traffic Violation Apprehension System Using Genetic Algorithm and Artificial Neural Network", 2016 IEEE Region 10 Conference

[6] Amey Narkhede, et al, "Automatic Traffic Rule Violation Detection and Number Plate Recognition", IJSTE - International Journal of Science Technology & Engineering | Volume 3 | Issue 09 | March 2017

[7] Ana Riza, et al, "Localization of License Plates Using Optimized Edge and Contour Detection Technique ", 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017

[8] Prem Kumar Bhaskar, et al, "Image Processing Based Vehicle Detection and Tracking Method ",Universiti Teknologi PETRONAS Seri Iskandar, 31750, Perak, Malaysia

[9] Kunal, Dahiya ; Dinesh, Singh ; and C. Krishna, Mohan "Automatic Detection of Bike-riders without Helmet using Surveillance Videos in Real-time". 2016 International Joint Conference on Neural Networks (IJCNN)

[10] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection"

[11] Kaiming He, Georgia Gkioxari, Piotr Doll´ar, Ross Girshick, "Mask R-CNN", Facebook AI Research (FAIR) ,24 JAN 2018

[12] http://pjreddie.com/media/files/yolo.weights

[13] https://github.com/matterport/Mask_RCNN/releases/download/v2.0/mask_rcnn_coco.h5

[14] Pooya Sagharichi Ha, Mojtaba Shakeri, "License Plate Automatic Recognition based on Edge Detection"

[15] Abdullah Asım, et al, "A Vehicle Detection Approach using Deep Learning Methodologies"