2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE) | 978-1-6654-8910-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICRAIE56454.2022.10054306

Machine Learning Based Approach for Traffic Rule Violation Detection

Kruti Lavingia
Nirma University
Ahmedabad, India
kruti.lavingia@nirmauni.ac.in

Mihirsinh Vaja
Nirma University
Ahmedabad, India
18bce251@nirmauni.ac.in

Pooja Chaturvedi
Nirma University
Ahmedabad, India
pooja.chaturvedi@nirmauni.ac.in

Ami Lavingia
Sal College of Engineering
Ahmedabad, India
ami.lavingia@sal.edu.in

***Abstract* - The goal of this paper is to design an automated system model to monitor the violation of traffic rules, specifically the number of people sitting on a two-wheeler. Typically, in areas near the security offices, people tend to follow the rules, but in areas where no one is watching, people violate the rules. In our case of an organizational campus, if there are three people traveling on a two-wheeler but when they encounter a security guard, one of the persons gets down and walks ahead of the guarded area and then again sits back on the vehicle. In such cases, efficient methods are required to monitor the violation of specified traffic rules without human intervention. For the above-mentioned challenge, a deep learning-based solution is provided where the process starts with object recognition using YOLOv3 (You Only Look Once) model, using which a person sitting on any particular vehicle is identified based on a minimum threshold distance. Also, for the distance calculation, a depth estimation algorithm which helps us in finding the 3-D distance between objects from a 2-D image is implemented. Moreover, the number plate of the vehicle violating the above-mentioned rule is identified for easy identification of the person violating the rule. The proposed approach is implemented on a real time video streaming dataset. The simulation results show the efficiency of the proposed approach in terms of accuracy, precision and recall as 91%, 86% and 94% respectively.**

Keywords - YOLOv3, Deep Learning, Object Detection, Depth Estimation, Triple Riding

I. INTRODUCTION

In 2021, the average number of daily accidents in India was 1214 and a total of 3.54 lakhs road accidents were reported in 2020. Two Wheeler's constitute nearly 25% to 33% in road crash death [1]. Two-wheeler being one of the most common modes of transport in India, but due to less safety, makes traveling risky. So, to prevent such casualties the Government has imposed various rules and violation of these rules is a punishable offense by law. Following the traffic rules is a mandatory requirement for everyone because a crash not only damages the person who is not following the rules but also causes damage to the people around the accident spot leading to a major loss. Due to these reasons the fine for traffic rule violation has also been increased recently by the government. Normally people follow rules in the guarded area in order to avoid fine if caught by the guard. But when no one is watching people tend to violate the rules. So, it is necessary to detect such violations without human intervention by placing CCTV cameras on critical areas where there is no human monitoring and the usage of that area is more.

The different sections of the paper are: Section 2 related work 3 model and algorithms 4 proposed methodology 5 Simulation results and analysis 6 conclusion and future work 7 references.

II. LITERATURE SURVEY

Object detection and identification from a real time streaming data set is considered as a challenging problem in the field of computer vision and image processing. The traditional approaches are often not proved efficient in correct identification and classification of objects, so machine learning and deep learning approaches have been considered as of much significance. The deep learning-based approaches can be categorized into two types: one stage detectors and two stage detectors. In the one stage-based detection framework, each position on the image is considered as significant to be considered as either background or image object. But the two stage detector approaches consist of two steps: proposal generation which aims to identify the position on the image which can be considered as target object or point. In the next step, the feature vectors are analyzed using machine learning based approaches. The one stage detection approaches convergence is faster.

In [2], authors have proposed a real time object detector which considered the object detection as a regression problem which considers the input image comprising of grid cells where each grid cell is a candidate position for one or more objects. It was implemented using the OpenCV and keras packages in python. The limitation of YOLO was that it could only two objects at a location and it makes it difficult

978-1-6654-8910-2/22/$31.00 ©2022 IEEE

to detect the objects in the crowded scenarios. Several other models have used in the field of object detection such as Single Shot detectors (SSD) which improves the detection by using the concept of anchor box in each grid cell [3]. The various deep learning-based object detection approaches have been proposed in the literature [4, 5]. To overcome the limitation of basic YOLO model, several versions of YOLO has been proposed in the literature as YOLOv2 [6] and YOLOv3 [13]. YOLOv2 improves the accuracy of object detection by employing the deep convolutional neural network architecture which can be pre trained for high resolution images hence can detect objects more efficiently. YOLOv3 is an improved version of YOLOv2 model as it considers the real time video image as a whole. The advantage of this model is faster convergence and high accuracy as compared to other machine learning models.

A. Traffic rule monitoring approaches

Yange Li et al. [7] have implemented a helmet detection system which is a deep learning-based model which uses Mobile-Net algorithm. The application of the same model can be in Traffic CCTV cameras as well as the industry area. Shashank Chandak et al [8] uses various Convolution Neural Network (CNN) models for identification of helmets worn or not along with the crosswalk violation detection. For that they're using YOLO object detection algorithm for vehicle detection, Helmet worn or not is detected using CNN (Convolutional Neural Network) and Crosswalk violation is checked using Instance Segmentation by Masked-RCNN and finally Optical character recognition (OCR) is done for number plate identification. H. R. Mampilayil et al. [9] proposed the automated detection of vehicles and some techniques to recognize the vehicles moving in the wrong lane in a one-way traffic. The technique used was based on that, for each frame from a video input, for each block median of flow is determined. If the distance between the current block and the corresponding image is more than 2.57 then the vehicle is considered to be moving in the wrong direction. Pooya Sagharichi et al. [10] focuses on Automated License Plate Recognition where noise is removed using FMH filter then Background filtering and then canny edge detection for finding the location of the number plate. Then OCR algorithm can be used on the number plate to find the string value from the image of the cropped number plate. Dr. Rekha PM et al. [11] for detecting triple riding uses HAAR-Cascade for Object Detection using which they are applying face detection and then extracting various features of eyes, nose and mouth using which helmet detection will also take place and depending on the number of faces, triple riding is detected. Nikhil Chakravarty et al. [12] focused on distance based triple riding detection which starts with the object detection and then taking a Euclidean distance to check whether the person is sitting on that vehicle or not. Research is majorly done in order to detect whether the helmet is worn or not. This work was also focused on an extended variation of research done by Nikhil Chakravarty [12] with a variation of including a Depth Estimation for finding the distance between the objects along with the detection of the number plate if the violation is detected.

III. Models and Algorithm

The different models used in the proposed approach are discussed in this section as follows:

A. YOLOv3 for identification of objects

YOLO stands for You Only Look Once and is built using a CNN called DarkNet-53 consisting of 53 convolutional layers. YOLO can process images and do object detection in real time with 45 frames per second. The limitation of this model lies in the detection of smaller objects [13].

In case of multiple object detection Non-Maximal Suppression (NMS) is considered to be efficient. In case of multiple objects, this method provides the feature to include the boundary box around the each identified object as shown in the Fig.1. [14].

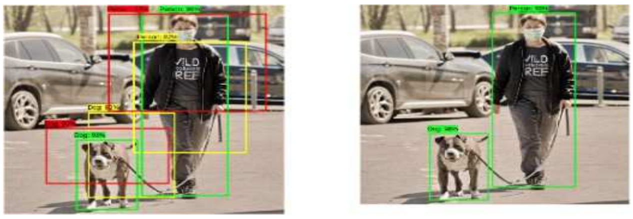

Fig. 1. Result of application of Non- Maximal Suppression [10]

B. Depth Estimation Algorithm

When a person is close to the camera, he looks bigger and when he is away from the camera, he looks smaller. So, the depth estimation algorithm is applied, which will generate a disparity map and disparity values of the pixel will be used to find the depth. Depth estimation algorithm is implemented by using one of the sub variants of Monodepth2 which is mono+stereo model. The disparity value for a pixel is calculated using the equation 1.

$$disp = disp * \text{width} * \text{scaling factor} \quad (1)$$

where the value of image width and scaling factor are considered as 416 and 0.54 respectively. The depth of the object is calculated using focal, baseline and disparity value using the equation 2.

$$depth = focal * baseline/disp \qquad (2)$$

Figure 2 and figure 3 shows the result of application of Depth Estimation on Image.

Fig. 2. Input image for Depth Estimation

Fig. 3. Result of Generated Disparity Map

C. Automatic Number Plate Recognition (ANPR)

Number Plate Recognition is also an important part of the traffic rules detection system. The coordinates of the violated vehicles are used to crop the image of a particular vehicle. Then the number plate is extracted from the image. Fig 4 shows a flow Graph for AutomaticNumber Plate Recognition.

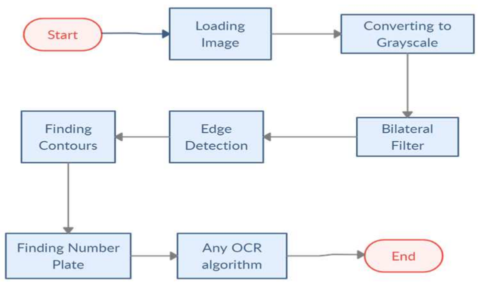

Fig. 4. Flow Graph for ANPR.

In the first step, the image is converted into grayscale format. Then a bilateral filter is applied to remove the noise [17]. After that, the edge detection is applied on the image. The results of the filtering and edge detection operations are as shown in the figure 5.

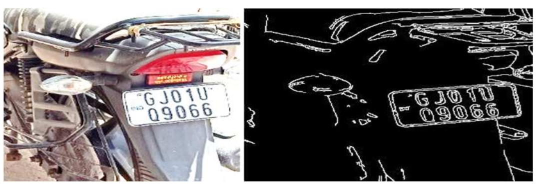

Fig. 5. The results of filtering and edge detection

Once all the edges are detected, the contours are found, where contours mean a simple line/curve joining continuous points with the same intensity, in other words it finds the polygon. Once the polygons are marked, all the contours of the polygons with the area of the polygon as key are sorted. In the next step iterate over the sorted contours which will also contain the coordinates of the corners of the polygon. During any iteration, if the count of the coordinates or the number of coordinates is equal to 4 then it means the polygon with 4 corners i.e. rectangle and almost all the number plates are rectangular in shape.

Now with the coordinates of the number plate, masking operation is applied to the image in order to remove the background image other than the number plate, and then crop that part of the image so only the number plate is left. After that any Optical Character Recognition (OCR) algorithm can be applied. In the proposed model, we have used **easyocr** to extract text from the image because string format of the vehicle number is required in order to use it in future to access the details from the database [18].

IV. METHODOLOGY

The different tasks performed to identify the vehicles having more than two persons are shown in the form of flow graph in figure 6.

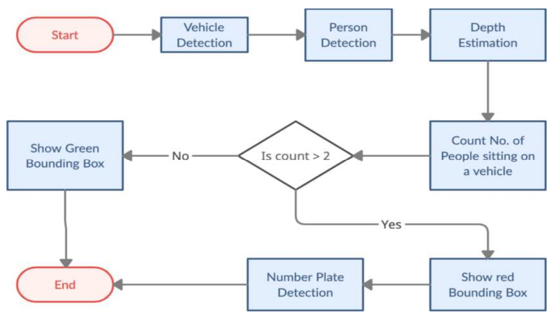

Fig.6. Flow Chart of the Proposed Approach

The major steps performed in the proposed approach are as:

A. Preparing the dataset of input image from the real time video processing

The frames are collected and stored in the form of images from the real time video captured from the CCTV or other high-resolution cameras.

B. Identification of two-wheeler and person

In the next step the vehicles and person are identified from the collected images. The result of the vehicle and person detection is as shown in the Fig. 7. It can be seen from the figure that the vehicle detected are bounded by the green boundary box and the person detected are bounded by the purple boundary box.

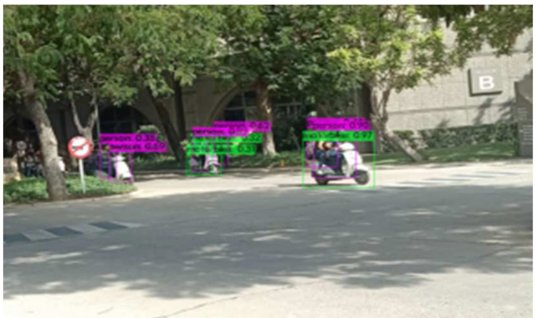

Fig. 7. Result of detection of two-wheeler and person

C. Calculation of the distance between the person and the two-wheeler

The real time video consists of frames containing the information in the three-dimensional space, so we will need three parameters for the calculation of the distance as: i. X coordinate, ii. Y coordinate and Depth (Luminance). The distance between the vehicle and the person is calculated using the Euclidean distance formula as shown in the equation 3.

$$Distance = \sqrt{(L_1 - L_2)^2 + (x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (3)$$

To determine the depth parameter, the inbuilt model mono depth2 in depth estimation algorithm is used. The mono depth2 model generates the heatmap like image in the form of disparity map as shown in the figure 8.

D. Identify whether the person is sitting on vehicle or not

If the pixel distance is less than the threshold value, then it can be inferred that the person is sitting on the vehicle. If the count of the person sitting on the vehicle is more than two, then it is reported as traffic rule violation [16].

E. Determination of traffic violation

Depending on the number of persons sitting on a vehicle, the boundary box is determined. If the number of persons sitting is less than two, then the boundary box is created in green color otherwise the color of the boundary box is red. The results of the simulation are as shown in the figure 8.

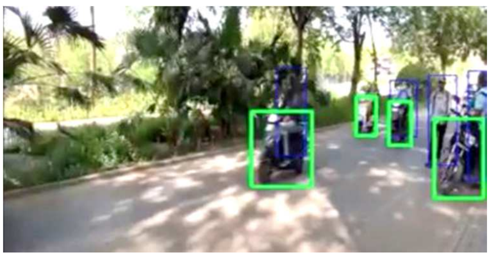

Fig. 8. Bounding Box for Vehicles.

F. Number Plate Recognition

After the identification of the vehicle violating the traffic rule, the number plate of the vehicle is cropped from the image of the vehicle as shown in the figure 9.

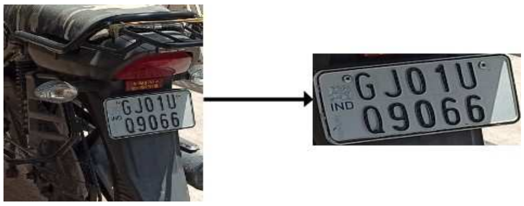

Fig. 9. Recognition of number plate localization.

The Optical character recognition algorithm is then implemented on the cropped to recognize the characters of the number plate [18].

V. Simulation Results and Analysis

In this section we will discuss about the implementational aspect of the proposed methodology.

A. Simulation Setup

The performance of the proposed methodology is implemented using the keras package and OpenCV in python language. The dataset is collected by capturing the frames from a real time video streaming.

B. Results discussion and analysis

a. Depth estimation results

The depth estimation algorithm is applied to determine the distance between the vehicle and the person. The results of the application of the depth estimation algorithm are as shown in the figure 10 and figure 11. Figure 10 displays the two-person standing little behind and Fig. 11 displays two person standing next to each other. It can be seen from the results that the distance between the person is found to be 70.29 units without using the depth estimation algorithm whereas with the application of depth estimation algorithm, the distance is obtained as 95.46 units.

Fig. 10. Result of distance estimation for the person standing behind another person.

It can be seen from the results that the distance between the person is found to be 59.22 units without using the depth estimation algorithm whereas with the application of depth estimation algorithm, the distance is obtained as 61.51 units.

Fig. 11. Result of distance estimation for the person standing next to each other.

b. *Confusion Matrix*

Confusion matrix is a graphical approach to represent the efficiency of the machine learning algorithm and is defined using the four parameters as: True Positive (TP), True Negative (TN), False Positive (FP) and FalseNegative (FN). [19-21]

The FP metric represents the number of instances when the rider is safe and the model also predicted the result as safe or no violation, whereas the FP represents the number of instances when the rider has violated the traffic rule but the model predicted as not safe or rule violation. The TN metric represents the number of instances when the rider violated the traffic rule and the model also detected violation. The FN metric represents the number of instances when the rider is safe but the model predicted violation. The results of the proposed model in terms of TP, FP, TN and FN are as shown in the figure 12 to figure 15 respectively.

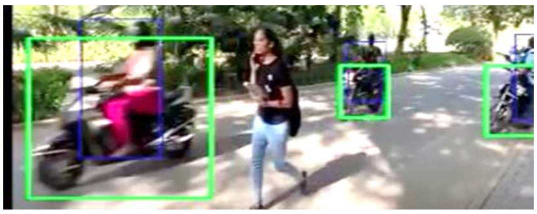

Fig. 12. Result of the True Positive for the proposed model

Fig. 13. Result of False Positive for the proposed model

Fig. 14. Result of TN for the proposed model

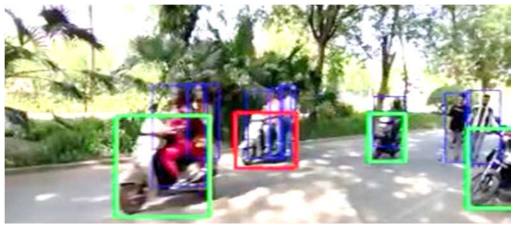

Fig. 15. Result of FN for the proposed model

The confusion matrix obtained using the parameters can be represented as shown in the Table 1:

Table 1. Result of confusion matrix

	Actual Values		
Predicted Values		**Safe**	**Violation**
	Safe	102	7
	Violation	17	34

c. *Accuracy*

This metric represents the number of images which have been correctly classified and is determined using the equation 4:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

$$Accuracy = 136/150 \Rightarrow 90.6\%$$

d. *Recall*

This metric represents the percentage of classes predicted correctly out of all positive classes and is determined using the equation 5:

$$Recall = TP / (TP + FN) \quad (5)$$

$$Recall = 102/119 \Rightarrow 85.7\%$$

e. *Precision*

This metric represents the percentage of correctly predicted positives out of all predicted positives and is determined using the equation 6.

$$Precision = TP / (TP + FP) \quad (6)$$

$$Precision = 102/109 \Rightarrow 93.57\%$$

VI. CONCLUSION AND FUTURE WORKS

The paper proposed an approach to detect the violation of triple riding rules in our college campus. It aims to automate the violation detection through a sequence of steps which start from the vehicle detection and up to automatic number plate detection. The results show the efficiency of the proposed model in terms of accuracy, precision and recall as 91%. 86% and 94% respectively. Our scope in this paper is limited up to the detection of triple riding, but in future all the rules which are possible to integrate in the same code including helmet detection, wrong lane detection can be implemented.

REFERNCES

[1] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: CVPR, 2016.

[2] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C.Berg, SSD: Single shot multibox detector, in: ECCV, 2016.

[3] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, in: ICCV, 2017.

[4] H. Law, J. Deng, Cornernet: Detecting objects as paired keypoints, in: ECCV, 2018

[5] J. Redmon, A. Farhadi, Yolo9000: better, faster, stronger, in: CVPR, 2017.

[6] A. Tonge, S. Chandak, R. Khiste, U. Khan and L. A. Be-woor, "Traffic Rules Violation Detection using Deep Learning", 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1250-1257.

[7] Yange Li, Han Wei, Zheng Han, Jianling Huang, Weidong Wang, "Deep Learning-Based Safety Helmet Detection in Engineering Management Based on Convolutional Neural Networks", Advances in Civil Engineering, vol. 2020.

[8] A. Tonge, S. Chandak, R. Khiste, U. Khan and L. A. Bewoor, "Traffic Rules Violation Detection using Deep Learning," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1250-1257.

[9] H. R. Mampilayil and R. K., "Deep learning-based Detection of One Way Traffic Rule Violation of Three Wheeler Vehicles," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1453-1457.

[10] Ha, Pooya Sagharichi and Mojtaba Shakeri. "License Plate Automatic Recognition based on edge detection." 2016 Artificial Intelligence and Robotics (IRANOPEN) (2016): 170-174.

[11] Dr. Rekha P M, Likitha G, Nishant Kumar, Alisha Simran, 2020, Study of Different Approaches to Develop an Automatic Traffic Rule Violation Detection System, INTERNATIONAL JOURNAL OF ENGINEERING RSEARCH & TECHNOLOGY (IJERT) NCAIT-2020 (Volume 8 Issue 15).

[12] Mallela, Nikhil Chakravarthy & Volety, Rohit & Ramalingam, Srinivasa Perumal & Rk, Nadesh. (2021). Detection of the triple riding and speed violation on two-wheelers using deep learning algorithms. Multimedia Tools and Applications. 80.

[13] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv.

[14] Joseph Redmon, Santosh Divvala, Ross Girshick, & Ali Farhadi. (2015). You Only Look Once: Unified, Real-Time Object Detection

[15] Singh, A. (2020, August 3). Selecting the Right Bounding Box Using Non-Max Suppression (with implementation). Analytics Vidhya. Retrieved May 2, 2022, from https://www.analyticsvidhya.com/blog/2020/08/selecting-the-right-bounding-box-using-non-max-suppression-with-implementation.

[16] License Plate Detection and Recognition Using OpenCv And Pytesseract. (2021, December 31). Section.io. Retrieved May 2, 2022, from https://www.section.io/engineeringeducation/license-plate-detection-and-recognition-using-opencv-and-pytesseract/

[17] H. Ahamad, N. Zaini and M. F. A. Latip, "Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI," 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2020, pp. 113-118, doi: 10.1109/ICCSCE50387.2020.9204934.

[18] Python ANPR with OpenCV and EasyOCR in 25 Minutes — Automatic Number Plate Recognition Tutorial. (2020, December 13). YouTube. Retrieved May 2, 2022, from Python ANPR with OpenCV and EasyOCR in 25 Minutes-Automatic Number Plate Recognition Tutorial.

[19] Pandey, N., Lavingia, A. and Lavingia, K., 2015. Analyzing the performance of optical system using different modulation techniques. International Journal of Advance Engineering and Research Development, 2(12).

[20] Dakwala, A. and Lavingia, K., 2016. A Machine learning approach to improve the efficiency of Fake websites detection Techniques. International journal of computer Science and Communication (IJCSC) Vol, 7, pp.236-243.

[21] Lavingia, K.R. and Mehta, R., 2022. Information Retrieval and Data Analytics in Internet of Things: Current Perspective, Applications and Challenges. Scalable Computing: Practice and Experience, 23(1), pp.23-34.