Image Processing based Traffic Surveillance

D K Chandrakanth
Dept. of E&CE
Siddaganga Institute of Technology
Tumakuru, India
1si20ec020@sit.ac.in

Siddharth Prabhu
Dept. of E&CE
Siddaganga Institute of Technology
Tumakuru, India
1si20ec090@sit.ac.in

Preetham G M
Dept. of E&CE
Siddaganga Institute of Technology
Tumakuru, India
1si20ec121@sit.ac.in

Mithun J B
Dept. of E&CE
Siddaganga Institute of Technology
Tumakuru, India
1si20ec124@sit.ac.in

A N Mukunda Rao
Dept. of E&CE
Siddaganga Institute of Technology
Tumakuru, India
anmukund@sit.ac.in

***Abstract*—Motorcycles being primary modes of road transportation globally, face significant risks when traffic rules such as failure to wear helmets, engaging in triple riding, or using mobile phones while driving are disregarded, increasing the likelihood of accidents and injuries, posing threats not only to those directly involved but to the entire community sharing the road. Traditional traffic law enforcement encounters challenges like resource constraints, inconsistent enforcement, limited coverage, and safety concerns for law enforcement officers. To tackle these issues, efforts are underway to revolutionize traffic law enforcement through an automated surveillance system driven by image processing technology. This proposed system extracts the license plate information of the bike when traffic rules violation such as triple riding, failure to wear helmet and mobile phone usage while riding occurs. This proposed system utilizes traffic videos as inputs and employs YOLOv8, an object detection model, to accurately identify bikes and subsequently trained to detect violations. Upon violation detection, an OpenALPR model extracts license plate of the bike and stores this secure data in the Firebase Cloud. This system is tested on Tumkur city road traffic and in college campus for 50 videos of total duration around 100 hours. The system achieves an accuracy of 99% in detecting bike riders, helmeted riders with 99% accuracy, no helmeted riders with 98% accuracy and detecting the instance of mobile phone usage with an accuracy of 98% and triple riding instance with 96%. The system achieved an overall accuracy in detecting the traffic offenders with 97%. Also the system detects the license plate information of the violated vehicle with an accuracy of 93%. Such technology aids law enforcement in consistently enforcing traffic rules, thereby promoting enhanced road safety and orderliness, with the ultimate goal of creating safer roads for everyone.**

***Index Terms*—YOLOv8, OpenALPR, Image Processing, Object Detection, Vehicle Tracking, Traffic Surveillance.**

I. Introduction

To address the pressing issue of road safety amid escalating accidents due to the rapid increase in vehicular ownership, an innovative solution, the implementation of a robust traffic surveillance system integrating state-of-the-art image processing technology is proposed. Advancements in image processing technology have paved the way for sophisticated surveillance infrastructure capable of real-time monitoring and automated enforcement. Notably, India observed a significant 12% surge in road accidents alongside an 18% increase in vehicle ownership in 2023 [Courtesy: Indiatoday.com].

To tackle this challenge, the development of a traffic surveillance system utilizing image processing technology with objectives (1) To extract and record the License plate information of the bike which violates the traffic rules such as (i) Triple riding (ii) Riding without helmet (iii) Mobile phone usage while riding (2) License plate extraction of car is proposed. This system autonomously monitors traffic violations. Upon detecting such violations, the system extracts license plate number of the bike and records it along with specific violation type and time stamp. This data is securely stored in a comprehensive records of traffic offenses in Firebase cloud. License plate extraction of car is also implemented in this project keeping zebra crossing violation detection in future.

II. Literature review

Mallela et al. introduced an automatic system for detecting triple riding and speedrule violations, utilizing YOLOv3 trained on the COCO dataset to identify violators [1]. Apoorva Saumya et al. developed a machine learning-based system for detecting bike riders without helmets and those engaged in triple rides from traffic surveillance videos using YOLOv3, identifying riders with bounding boxes and counting the number of riders per bike [2]. Dahiya et al. introduced a method for automatically detecting motorcyclists using object segmentation and background subtraction, followed by distinguishing helmeted and non-helmeted riders using a binary classifier [3]. Dasgupta et al. employed a deep neural network with YOLOv3 to identify multiple motorcycle riders and detect helmets, using bounding boxes for rider detection and localization, and a softmax classifier for predicting helmet presence [4]. Mistry, Jimit et al. developed a CNN-based system for automatic detection of helmeted and non-helmeted motorcyclists, using YOLOv2 for object detection and OpenALPR for license plate extraction from non-helmeted images [5]. H. Yasar et al. A neural network-based Cascade Object Detector developed on MATLAB detects mobile phone usage in drivers, with an accuracy of 75%[6]. Waqar Riaz et al. introduced an automatic license plate recognition model utilizing the AOLP dataset for license plate detection, character extraction, and recognition, training YOLOv3 for license plate extraction [7].

The major challenges to be addressed in this proposed project are:

1) Input Video Quality: Surveillance cameras often capture video in low resolution, making it difficult to accurately detect and classify objects such as individuals riding triple or without helmets or using mobile phones while riding. In low light condition the degrade in video quality, challenges for object detection and classification is complex.
2) Object Occlusion and Clutter: In crowded traffic environments, objects of interest, such as individuals riding triple or without helmets, may be occluded by other vehicles or cluttered backgrounds.
3) Variability in Appearance: Individuals may wear different types of clothing or headgear, further complicating detection and classification.
4) Mobile Phone Usage Detection: Identifying riders using mobile phones while operating vehicles adds an additional layer of complexity. Mobile phones can be held in various positions and orientations, and their appearance may change based on factors such as screen size and case design.
5) License plate extraction: Variable speed of vehicles, languages of number plate & mostly non-uniform letter on number plate.

An attempt is made to address the above challenges using an AI enabled night vision camera positioned at around 12 feet height, YOLOv8 object detection model along with the efficient algorithm .

III. System Overview

Fig.1 depicts the algorithm of the proposed system,

1. Receive a video feed of a traffic scene as the algorithm input with 4MP at 30 fps.
2. YOLOv8 model is used for vehicle detection, specifically focusing on identifying motorcycles and cars within the scene.
3. If a motorcycle is detected, proceed to further analysis for potential violations. Employ the YOLOv8 model for person detection to identify individuals on the motorcycle.
4. Check for violations related to motorcycle riders, such as no-helmet, triple riding and mobile phone usage while riding. Employ the trained YOLOv8 model for each specific violation detection.
5. If a car is detected employ the YOLOv8 model to extract the license plate.
6. If violations are detected, capture the violators using the trained YOLOv8 model for violation recognition.
7. Extract license plate information using OpenALPR model for the identified violators.
8. Store the extracted license plate information along with violation type and time stamp in the Firebase Cloud database for further reference.
9. Proceed to the next frame for continuous monitoring and analysis.

This algorithm integrates YOLOv8 models for vehicle, person, and violation detection, along with a OpenALPR model for license plate recognition, ensuring comprehensive monitoring and enforcement traffic regulations.

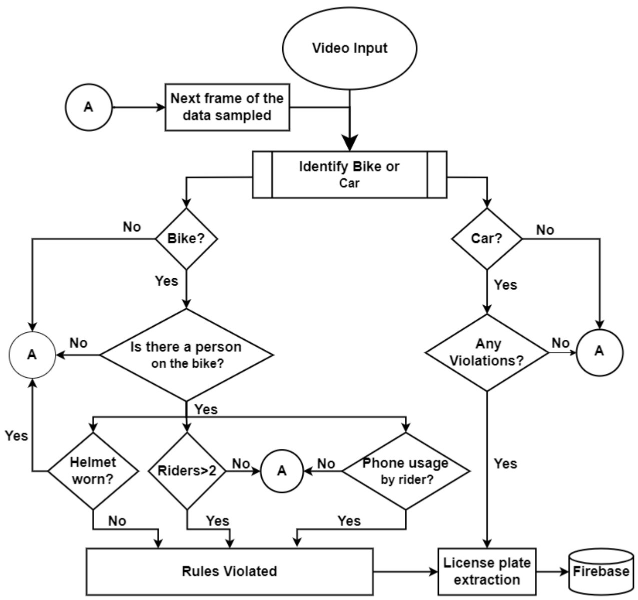

Fig. 1. Flowchart of the Proposed System considering both bike and car.

IV. Methodology

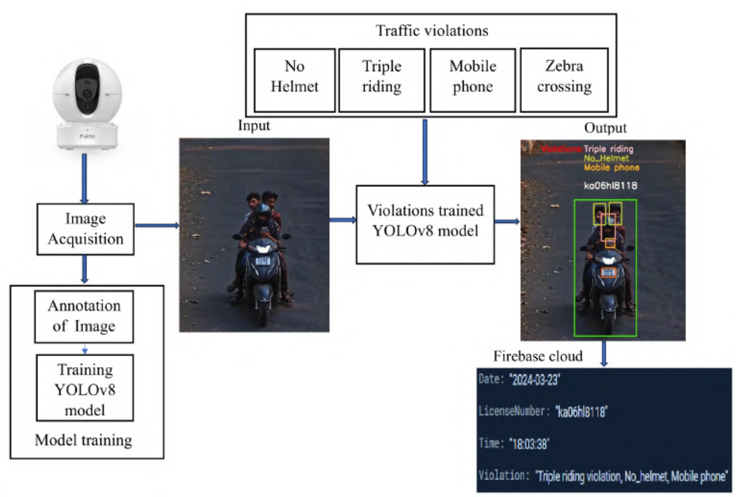

Fig. 2. Block Diagram of the Proposed System

The systems block diagram is depicted in Fig.2, during training stage three strategically positioned cameras capture traffic flow from multiple angles. These cameras feature highresolution capabilities to optimize performance. Captured images undergo resizing to streamline processing time without compromising accuracy. The proposed algorithm utilizes YOLOv8 for vehicle detection in the video feed of a traffic scene, with a focus on motorcycles and cars. For motorcycles, it employs YOLOv8 for person detection to identify violations such as no-helmet use, triple riding, and mobile phone usage. Similarly, the algorithm utilizes YOLOv8 to detect cars and its license plate information is extracted using OpenALPR model in future this could be used along with zebra crossing violation detection model.

A. Dataset gathering:

A comprehensive dataset comprising 2000 images was collected for training a custom object detection model to detect traffic violations. Testing was conducted using traffic video recordings from around the institution and main roads. Images for training were sourced from various databases on the web and captured manually, ensuring a diverse set of scenarios and angles. This approach provides a robust training and testing dataset, enhancing the model ability to accurately detect violations in real-world traffic conditions.

B. Annotation of Images:

The dataset was annotated into six classes: 'Car', 'Helmet', 'LP' (license plate), 'Mobile', 'No Helmet', and 'P_Bike'. Bounding boxes were manually drawn around objects of interest, such as persons, bikes, mobile and helmets, in each image. Annotation was done using the YOLO format with the Roboflow tool, resulting in YOLOv8 formatted labels. The images were resized to 640x640 for training the YOLOv8 model, which was used to detect all six classes for comprehensive traffic violation detection.

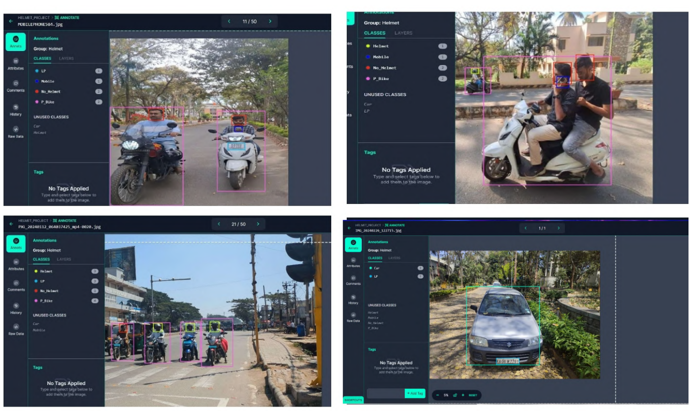

Fig. 3. Labeling on Roboflow for P_Bike, No_helmet, LP, Mobile, Helmet & Car.

C. Model Training:

The training process involved developing specialized YOLOv8 models for multiple detection tasks, including identifying individuals, bikes, helmets, mobile phone usage, car and license plates. Images and their associated labels were carefully arranged and then uploaded to Google Drive for training the models using Google Colab. The training was conducted over 150 epochs using the YOLOv8n model. After completing the training, configuration files containing crucial parameters and weights files were obtained. These files are essential for executing future prediction tasks, ensuring the effective deployment and utilization of the trained models.

D. Prediction:

In YOLOv8, the prediction process involves dividing the input image into a grid of cells, each responsible for predicting bounding boxes, objectness scores, and class probabilities. Anchor boxes help predict bounding box dimensions at different scales. For each grid cell, YOLOv8 predicts bounding boxes center coordinates, width, height, objectness scores (using the sigmoid function), and class probabilities (using softmax).

After predictions, non-maximum suppression removes redundant or overlapping boxes, keeping the most confident ones. The final prediction includes bounding boxes with class labels and confidence scores, representing detected objects.

The process can be summarized in the following steps:

1. Feature Extraction: The image is resized (e.g., 640x640) and fed through the network, extracting features.

2. Grid Division: The extracted features are divided into a grid (e.g., 13x13 cells).

3. Bounding Box Center and Size Prediction: For each cell, the network predicts a set number of bounding boxes with:
• Class probabilities: The likelihood of each class (person, car) for each box.
• Bounding box coordinates: These represent the box's location and size relative to the cell, calculated using equations:
- Center coordinates (tx, ty):

$$bx = \sigma(tx) + cx$$

$$by = \sigma(ty) + cy$$

- Width and height (tw, th):

$$bw = pw \cdot e^{tw}$$

$$bh = ph \cdot e^{th}$$

where:
- bx, by are the center coordinates of the bounding box.
- bw, bh are the width and height of the bounding box.
- tx, ty, tw, th are the outputs of the neural network.
- cx, cy are the coordinates of the top-left corner of the anchor box.
- pw, ph are the dimensions of the anchor box.

4. Decoding Predictions: Predicted attributes are decoded to obtain absolute coordinates and dimensions within the original image.

5. Non-Maximum Suppression (NMS): If multiple boxes predict the same object, NMS removes redundant boxes, keeping only the one with the highest confidence score.

6. Confidence Score Calculation: It is used to indicate the model's certainty that a bounding box contains an object and that the object belongs to a specific class is shown in eq.1

$$\text{Confidence Score} = \sigma(\text{Pr(Object)} \times \text{IoU}) \tag{1}$$

where:
- $Pr(Object)$ is the probability that an object exists in the bounding box.
- IoU(Intersection over Union) formula is used to calculate the overlap between two bounding boxes. It is shown in eq.2

$$\text{IOU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} \tag{2}$$

These equations and steps are used in YOLOv8 to predict bounding boxes, objectness scores, and class probabilities for objects detected in an image, contributing to its efficiency and accuracy.

E. Detection:

In the video processing pipeline, the initial stage involves employing the YOLOv8 model for bike detection. Upon successful bike detection, the system proceeds to utilize YOLOv8 for identifying individuals. Subsequently, the model scrutinizes potential violations like lack of helmet, mobile phone usage, and instances of triple riding. Upon detecting any violations, the YOLOv8 model crops the license plate of the violating vehicle, which is then inputted to the license plate recognition model for extraction. In cases where no vehicle is detected, the system smoothly transitions to analyzing the next frame, ensuring continuous monitoring and comprehensive assessment of the video footage.

Precision, recall, and accuracy are important metrics for evaluating the performance of an object detection system. They are calculated as follows:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives + False Positives}} \tag{3}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}} \tag{4}$$

$$\text{Accuracy} = \frac{\text{True Positives + True Negatives}}{\text{Total Number of Cases}} \tag{5}$$

These metrics can help assess the effectiveness of the system in detecting violations and can be used to optimize the performance of the system.

V. Software Description

A. You Only Look Once(YOLO)

YOLO (You Only Look Once) has become a leading solution in computer vision, especially for object detection. Its unique methodology allows real-time object detection in images, videos, or live feeds, providing quick and accurate recognition by analyzing entire images at once. The YOLOv8 architecture demonstrates its innovative approach, surpassing traditional algorithms in object detection challenges.

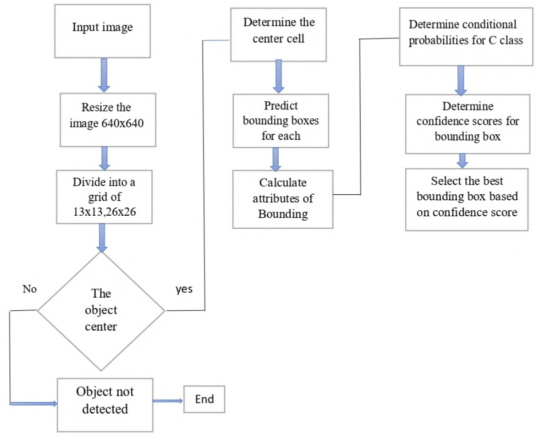

Fig. 4. Workflow of YOLOv8

VI. Results and Discussion

In results and discussion, the following abbreviations are used:
- NHV: No helmet violation.
- MPV: Mobile phone violation.
- TRV: Triple riding violation.
- LP : License plate.
- HD : Helmet detection.

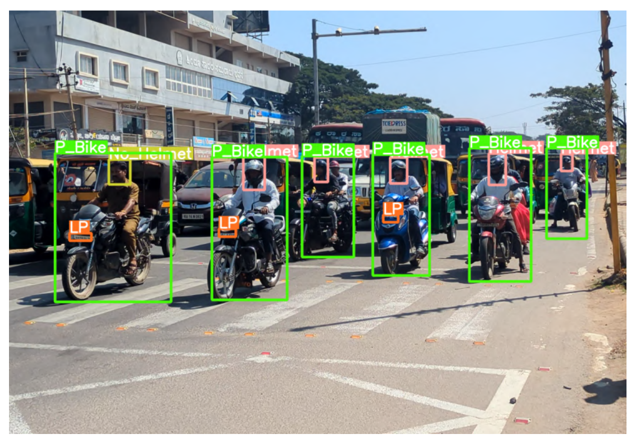

Fig. 5. Detecting Helmeted and No helmeted riders in a traffic scenario.

Fig. 6. No violation (6a) & extracting the LP for NHV scenario (6b).

Fig. 7. No violation (7a) & extracting the LP for a two people violating the helmet rule (7b).

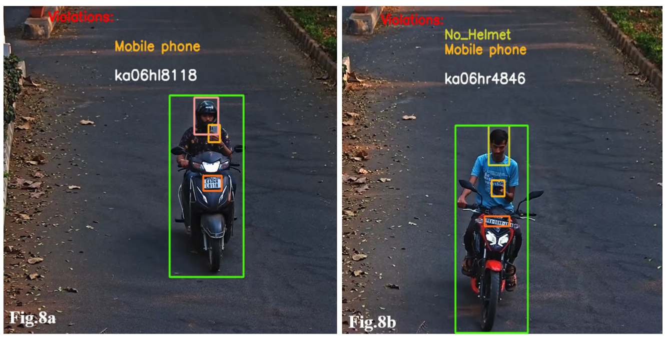

Fig. 8. Extracting LP for MPV (8a) and NHV with MPV (8b).

Fig. 9. Extracting the LP for TRV & NHV (9a) and TRV (9b).

Fig. 10. Extracting the LP for NHV (10a) & for TRV, NHV with MPV (10b).

Fig. 11. Extracting LP for car (11a) (11b)& No violation of bike (11b).

Fig. 12. Rule violators information in Firebase cloud.

The final merged model was tested on approximately fifty videos depicting traffic scenarios in Tumkur and on the scenarios in our college campus, considering 20 possibilities in the videos. The models performance was as follows:
Motorcycle detection accuracy is 99%. Helmet and No helmet detection accuracy is 99%. Mobile phone usage detection accuracy is 98%. Triple riding case detection accuracy is 96%. License plate extraction accuracy for both bikes and cars is 93%. The system has achieved an overall accuracy of 97% in detecting the traffic offenders.

Case 1: Multiple vehicles with helmet and no helmet riders in traffic scenario (Fig.5).
case 2: No violation- single rider with helmet (Fig.6a, Fig.11b), two riders with helmet (Fig.7a).
Case 3: Violation & license plate extraction- no helmet (Fig.6b, Fig.7b, Fig.10a), mobile usage (Fig.8a), mobile usage & without helmet (Fig.8b), triple riding & without helmet (Fig.9a), triple riding (Fig.9b), triple riding, no helmet & helmet (Fig.10b).
Case 4: License plate extraction of car (Fig.11a, Fig.11b).

The traffic rule violators' information such as license plate of the vehicles, violation type, date & time is recorded in Firebase cloud as shown in Fig.12.

Additionally the confusion matrix for motorcycle detection is presented in Table I. Table II shows the confusion matrix for helmet and no helmet detection. Table III displays the confusion matrix for mobile phone usage detection. Table IV illustrates the confusion matrix for triple riding detection. The confusion matrix for violations detection is depicted in Table V, and Table VI showcases the confusion matrix for license plate extraction.

TABLE I
CONFUSION MATRIX FOR MOTORCYCLE DETECTION

Actual Class %	**Predicted Class %**	
	Motorcycle	Other
Motorcycle	99	1
Other	1	99

TABLE II
CONFUSION MATRIX FOR HELMET AND NO-HELMET DETECTION

Actual Class %	**Predicted Class %**	
	Helmet	No Helmet
Helmet	99	1
No Helmet	2	98

TABLE III
CONFUSION MATRIX FOR MOBILE PHONE DETECTION

Actual Class %	**Predicted Class %**	
	Mobile Phone	No mobile phone
Mobile Phone	98	2
No mobile phone	2	98

TABLE IV
CONFUSION MATRIX FOR TRIPLE RIDING DETECTION

Actual Class %	**Predicted Class %**		
	1 rider	2 riders	More than 2 riders
1 rider	99	1	0
2 riders	2	98	1
More than 2 riders	1	3	96

TABLE V
CONFUSION MATRIX FOR VIOLATION DETECTION

Actual Class %	**Predicted Class %**	
	Violation	No Violation
Violation	97	3
No Violation	3	97

TABLE VI
CONFUSION MATRIX FOR LICENSE PLATE EXTRACTION

Actual Class %	**Predicted Class %**	
	LP match	1 to 2 Number Mismatch
LP match	93	7
1 to 2 Number Mismatch	7	93

VII. CONCLUSION

The system is designed to automatically detect traffic violations using the YOLOv8 model, effectively identifying infractions such as riding without a helmet, triple riding, and using a mobile phone while riding a bike. Demonstrating exceptional accuracy, the model achieves a 99% detection rate for bikes and riders. It distinguishes helmeted riders with 99% accuracy and non-helmeted riders with 98% accuracy, while also identifying instances of more than two riders on a bike with 96% accuracy. Additionally, it detects mobile phone usage with a accuracy of 98%. Furthermore, the system recognizes the license plate information of the vehicle with an accuracy of 93%. Overall, the system achieved 97% accuracy in detecting traffic rule violators. These results highlight the system's significant contribution to precise traffic offender detection and hence enhancing the road safety.

The model also extracts license plates for cars, which could potentially be extended to identify violators of zebra crossing rules in the future.

REFERENCES

[1] M. N. Chakravarthy, R. Volety, and N. R.K., "Detection of the triple riding and speed violation on two-wheelers using deep learning algorithms", *Multimedia Tools and Applications*, Springer, 2021.

[2] A. Saumya, V. Gayathri, K. Venkateswaran, S. Kale and N. Sridhar, "Machine Learning based Surveillance System for Detection of Bike Riders without Helmet and Triple Rides," in *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2020.

[3] K. Dahiya, D. Singh and C. K. Mohan, "Automatic detection of bike-riders without helmet using surveillance videos in real-time," in *2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, 2016.

[4] M. Dasgupta, O. Bandyopadhyay and S. Chatterji, "Automated Helmet Detection for Multiple Motorcycle Riders using CNN," in *2019 IEEE Conference on Information and Communication Technology*, Allahabad, India, 2019.

[5] J. Mistry, A. Aashishk, and K. P. Upla, "An automatic detection of helmeted and non-helmeted motorcyclists with license plate extraction using CNN," in *International Electrical Engineering Congress (iEECON)*, Chonburi, Thailand, pp. 1-4, Dec. 2014.

[6] H. Yasar, "Detection of Driver's mobile phone usage," in *2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, Manila, Philippines, 2017.

[7] W. Riaz, A. Azeem, G. Chenqiang, Z. Yuxi, Saifullah and W. Khalid, "YOLO Based Recognition Method for Automatic License Plate Recognition," in *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, Dalian, China, 2020.

[8] H. Shi and D. Zhao, "License Plate Recognition System Based on Improved YOLOv5 and GRU," *IEEE Access*, vol. 11, pp. 10429-10439, 2023.

[9] H. Lin, J. D. Deng, D. Albers and F. W. Siebert, "Helmet Use Detection of Tracked Motorcycles Using CNN-Based Multi-Task Learning," in *IEEE Access*, vol. 8, pp.162073-162084, 2020.

[10] X. -H. Wu, R. Hu and Y. -Q. Bao, "Block-Based Hough Transform for Recognition of Zebra Crossing in Natural Scene Images," in *IEEE Access*, vol. 7, pp. 59895-59902, 2019.