RideSafeAI: Real-Time Helmet and Multi-Rider Violation Detection Using YOLOv11 and Intelligent Tracking

K.Abhishikth[1], K.sivaramakrishna[2]

[1]Pursuing Computer Science, Andhra Loyola Institute of Engineering and Technology, Vijayawada - 12

[2]Professor, Head Of the Department of CSE(AIML), Andhra Loyola Institute of Engineering and Technology, Vijayawada – 12

---***---

ABSTRACT

The increasing use of two-wheelers has led to a rise in traffic violations such as helmet non-compliance and triple riding. Existing vision-based systems rely on single-frame detection, resulting in poor temporal consistency and high false positives. This paper proposes a real-time traffic violation detection framework integrating YOLOv11-based object detection with multi-object tracking for context-aware analysis. The system detects motorcycles, riders, and helmet usage while maintaining consistent identities across frames. Helmet violations are identified through head region analysis, and triple riding is detected by estimating rider count per vehicle. Additionally, Automatic License Plate Recognition (ALPR) is incorporated for enforcement. Experimental results demonstrate improved detection consistency, reduced false positives, and real-time performance. The proposed system offers a scalable solution for intelligent traffic monitoring and automated enforcement.

KEYWORDS— *traffic violation detection, helmet compliance detection, triple riding detection, YOLOv11, multi-object tracking, computer vision, deep learning, automatic license plate recognition (ALPR), intelligent traffic monitoring, smart transportation systems*

I. INTRODUCTION

The rapid growth of urbanization and the increasing reliance on two-wheelers as a primary mode of transportation have significantly contributed to traffic congestion and safety violations. Among these, helmet non-compliance and triple riding are major causes of severe injuries and fatalities in road accidents. Despite the presence of surveillance systems, effective and automated monitoring of such violations remains a challenging task, particularly in densely populated and dynamically changing traffic environments.

Recent advancements in deep learning and computer vision have enabled real-time traffic monitoring through object detection models such as YOLO. However, most existing systems rely on single-frame detection approaches, which lack temporal consistency and often produce false positives due to occlusions, motion blur, and complex traffic scenarios. Moreover, accurately associating multiple riders with a single vehicle and detecting compound violations remains a significant limitation in current methodologies.

To address these challenges, this paper proposes a real-time traffic violation detection framework that integrates YOLOv11-based object detection with multi-object tracking algorithms such as DeepSORT and ByteTrack. The proposed system ensures consistent identification of riders across frames, enabling reliable rider-to-vehicle association.

Helmet violation detection is performed using region-based head analysis, while triple riding is identified by estimating the number of riders linked to a single tracked motorcycle.

Additionally, the system incorporates Automatic License Plate Recognition (ALPR) to support automated enforcement.

II. LITERATURE SURVEY

Recent developments in intelligent transportation systems have leveraged deep learning and computer vision techniques for automated traffic monitoring and violation detection. Convolutional Neural Network (CNN)-based object detection models, particularly the YOLO family, have been widely adopted due to their capability for real-time performance [1]–[3]. Several studies have utilized YOLOv3 and YOLOv4 for detecting vehicles and identifying helmet usage among riders [1], [2], [9], [10]. While these approaches achieve high detection accuracy, they primarily rely on single-frame analysis, which limits their robustness in dynamic and occluded traffic environments.

To improve temporal consistency, multi-object tracking algorithms such as DeepSORT have been introduced [4]. DeepSORT combines motion prediction with appearance-based feature matching to maintain object identities across frames. Similarly, ByteTrack enhances tracking performance by effectively associating low-confidence detections, thereby improving robustness in crowded scenes [5]. However, most existing works focus on detection and tracking as separate tasks, lacking an integrated framework for comprehensive violation analysis.

Helmet detection has been addressed using region-based classification methods that focus on the head region of riders [9], [10]. Although these methods provide reasonable accuracy, their performance is often affected by variations in lighting conditions, camera angles, and partial occlusions. Furthermore, limited research has been conducted on detecting complex violations such as triple riding, which requires accurate rider-to-vehicle association and multi-frame validation.

In addition, Automatic License Plate Recognition (ALPR) systems have been widely used for vehicle identification and enforcement through Optical Character Recognition (OCR) [11], [12]. However, their integration with real-time violation detection systems remains limited, reducing their effectiveness in automated traffic enforcement applications.

In summary, existing approaches exhibit limitations such as lack of temporal consistency, high false positive rates, and inability to detect compound violations. To address these challenges, this work proposes a unified framework that integrates real-time object detection, multi-object tracking, and rule-based violation analysis for accurate and reliable traffic violation detection.

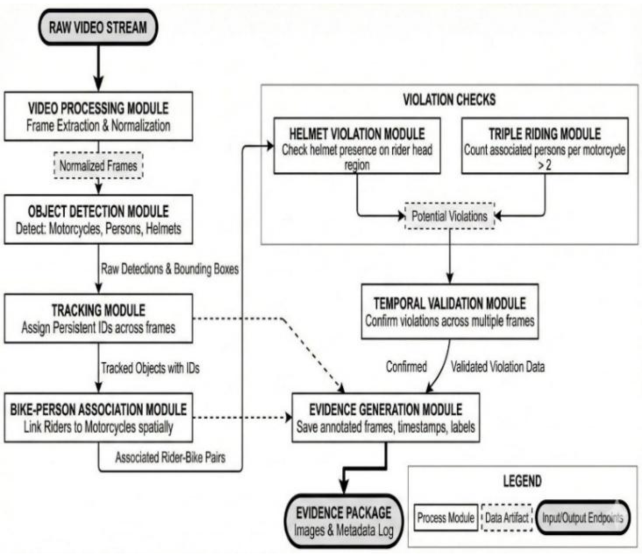

Fig.1 – System Architecture

III. PROPOSED WORK

This paper proposes a real-time traffic violation detection system that integrates deep learning-based object detection with multi-object tracking and rule-based analysis to identify helmet non-compliance and triple riding among two-wheeler users. The overall framework is designed to process live or recorded video streams and generate reliable violation reports with supporting evidence.

A. System Overview

The proposed system consists of four major components: (i) object detection, (ii) multi-object tracking, (iii) violation detection, and (iv) automated enforcement. Initially, video input is captured through surveillance cameras or uploaded recordings and processed frame-by-frame. Each frame is passed through the object detection module to identify motorcycles, riders, and helmets. The detected objects are then tracked across consecutive frames to maintain identity consistency. Based on the tracked information, violation detection rules are applied to identify helmet violations and triple riding. Finally, the system generates violation records along with extracted license plate information.

B. Object Detection using YOLOv11

The system employs the YOLOv11 model for real-time object detection due to its high accuracy and low latency. YOLO follows a single-stage detection architecture that predicts bounding boxes and class probabilities directly from input images [1]–[3]. It is trained to detect relevant classes, including motorcycles, riders, and helmets. Each input frame is processed to generate bounding boxes with associated confidence scores. Non-Maximum Suppression (NMS) is applied to eliminate redundant detections.

C. Multi-Object Tracking

To overcome the limitations of frame-wise detection, multi-object tracking is integrated using DeepSORT and ByteTrack algorithms. DeepSORT utilizes motion information through a Kalman filter and appearance features for robust identity assignment [4]. ByteTrack enhances tracking performance by associating both high-confidence and low-confidence detections, thereby improving tracking accuracy in dense traffic scenarios [5]. This tracking mechanism assigns unique IDs to each detected object and maintains their identities across frames, enabling consistent rider-to-vehicle association and reducing duplicate detections.

D. Helmet Violation Detection

Helmet violation detection is performed at the rider level by analyzing the head region of each detected rider. A spatial mapping approach is used to associate riders with their corresponding motorcycles. The presence or absence of a helmet is determined based on detection results within the head bounding region [9], [10]. To improve reliability, the system applies temporal validation by verifying helmet status across multiple consecutive frames, thereby reducing false positives caused by occlusions or detection errors.

E. Triple Riding Detection
Triple riding is identified by estimating the number of riders associated with a single tracked motorcycle. Using tracking IDs and spatial proximity analysis, multiple riders are linked to their respective vehicles. If the number of riders exceeds the permissible limit, a violation is flagged. Multi-frame verification is performed to ensure that the detected rider count is consistent over time, improving detection robustness in dynamic traffic conditions.

F. Automatic License Plate Recognition (ALPR)
For enforcement purposes, the system integrates an Automatic License Plate Recognition module. When a violation is detected, the region containing the license plate is extracted from the frame and processed using Optical Character Recognition (OCR) techniques [11], [12]. The extracted license number is stored along with violation details such as timestamp, type of violation, and visual evidence.

G. Violation Detection Engine
A rule-based violation engine is implemented to identify both individual and compound violations. The system checks for conditions such as:

- Helmet absence for one or more riders
- Triple riding on a single motorcycle
- Combined violations (e.g., triple riding with no helmets)

The engine ensures that violations are validated across multiple frames before generating alerts, thereby improving accuracy and reducing false detections.

IV. REQUIREMENT ANALYSIS

A. Hardware Requirements
The system requires a computer with at least 8 GB RAM and a multi-core processor (Intel i5 or above). A GPU is recommended for faster processing. Video input can be obtained from CCTV cameras or pre-recorded sources.
B. Software Requirements
The system is implemented in Python using libraries such as YOLOv11 for object detection, PyTorch for model execution, OpenCV for video processing, and DeepSORT/ByteTrack for tracking [4], [5]. OCR/ALPR is used for license plate recognition [11], [12].
C. Functional Requirements
The system detects motorcycles, riders, and helmets, tracks objects across frames, identifies violations (helmet and triple riding), and generates violation reports.
D. Non-Functional Requirements
The system ensures real-time performance, accuracy, scalability, and reliable operation under varying traffic conditions.

V. RESEARCH AND METHODOLOGY

The proposed system follows a deep learning-based methodology for real-time detection of traffic violations from video streams. The workflow consists of video acquisition, object detection, multi-object tracking, violation analysis, and automated enforcement. Initially, video input is captured from CCTV cameras or pre-recorded sources and processed frame-by-frame using OpenCV. Each frame is preprocessed and passed to the YOLOv11 model, which detects motorcycles, riders, helmets, and license plates with bounding box localization. YOLO-based models follow a single-stage object detection approach for real-time performance [1]–[3]. Non-Maximum Suppression (NMS) is applied to remove redundant detections and retain high-confidence outputs.

To ensure temporal consistency, multi-object tracking is performed using DeepSORT and ByteTrack [4], [5]. These algorithms assign unique IDs to detected objects and maintain their identities across frames, enabling accurate association between riders and their respective motorcycles.

Helmet violation detection is carried out by analyzing the head region of each rider. The system determines whether a helmet is present or absent based on detection within this region [9], [10]. Triple riding is identified by counting the number of riders associated with a single tracked motorcycle using spatial proximity and tracking IDs. A temporal validation mechanism is applied across multiple frames to reduce false positives.

Furthermore, an Automatic License Plate Recognition (ALPR) module is integrated to extract vehicle registration numbers using OCR techniques [11], [12]. Detected violations are recorded along with timestamps and visual evidence, enabling automated monitoring and enforcement.

Dataset
The proposed system utilizes multiple publicly available datasets to ensure robust performance across diverse traffic scenarios. The COCO (Common Objects in Context) dataset is used for general object detection, particularly for identifying motorcycles and persons, enabling the model to learn complex visual features from large-scale annotated data [1]. A specialized motorcycle helmet detection dataset is employed to train the helmet classification module, allowing accurate identification of riders with and without helmets under varying conditions [2]. Additionally, a human crowd dataset is incorporated to enhance detection and tracking performance in dense traffic environments with

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 04 | Apr 2026 www.irjet.net p-ISSN: 2395-0072

occlusions [3].
For license plate recognition, the UFPR-ALPR dataset is used to train and evaluate the Automatic License Plate Recognition (ALPR) module, enabling accurate extraction of vehicle registration numbers [4]. Furthermore, the Indian Driving Dataset (IDD) is included to provide real-world traffic scenarios with diverse road conditions, lighting variations, and vehicle types, thereby improving the model's generalization capability for deployment in Indian traffic environments [5].

VI. RESULTS AND ANALYSIS

The performance of the proposed traffic violation detection system was evaluated using standard object detection metrics, including precision, recall, and mean Average Precision (mAP). The model was trained and validated over multiple epochs, and the convergence behavior is illustrated through training and validation loss curves.
The training results show a consistent decrease in box loss, objectness loss, and classification loss, indicating stable learning and effective feature extraction. Similarly, validation losses follow a decreasing trend, demonstrating good generalization capability and minimal overfitting. Precision and recall values increase steadily during training, approaching values close to 1.0, which indicates high detection accuracy and reliability.
The model achieves an overall **mAP@0.5 of 0.994**, demonstrating excellent detection performance across all classes. Class-wise performance shows high precision values for helmet (0.995), no-helmet (0.994), rider (0.994), and license plate (0.995), indicating consistent accuracy across different object categories.

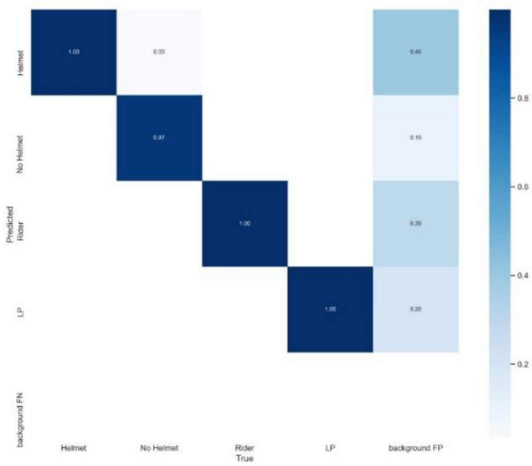

Fig.2 – Confusion matrix for Multi-Class detection

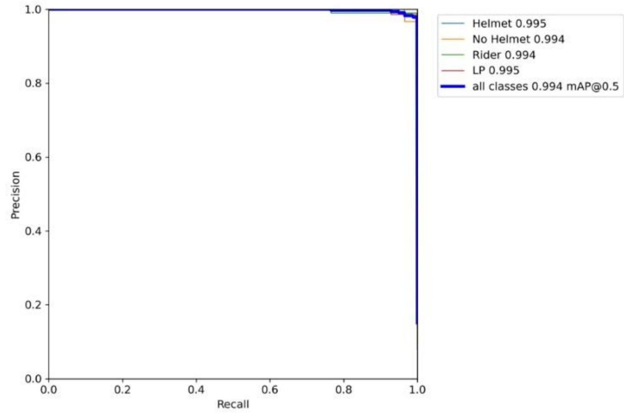

Fig.3 - Precision–Recall Curve

The confusion matrix further confirms the robustness of the model, with strong diagonal values representing correct classifications and minimal misclassification between classes. This highlights the model's ability to accurately distinguish between helmet and non-helmet cases, which is critical for violation detection.

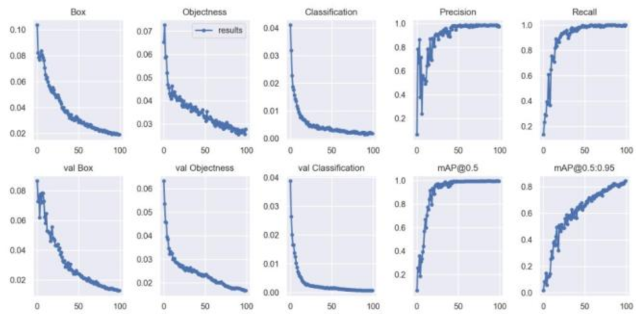

Fig.4 - Training and Validation Loss and Performance Metrics over Epochs

The precision-recall curve illustrates near-perfect performance, with curves closely approaching the top-right corner, indicating both high precision and recall. Additionally, the system demonstrates stable performance across varying traffic conditions, including occlusions and multiple object scenarios.

Overall, the integration of YOLOv11 with multi-object tracking and temporal validation significantly reduces false positives and improves detection consistency compared to traditional frame-based approaches. The results validate that the proposed system is capable of real-time, accurate, and reliable traffic violation detection suitable for practical deployment.

VII. CONCLUSION

This paper presented a real-time traffic violation detection system for identifying helmet non-compliance and triple riding among two-wheeler users using deep learning and multi-object tracking. By integrating YOLOv11-based object detection with tracking algorithms such as DeepSORT and ByteTrack, the proposed system ensures accurate rider-to-vehicle association and improved temporal consistency. The incorporation of a rule-based violation engine and ALPR module enables automated detection and enforcement with reliable evidence generation.

Experimental results demonstrate high detection accuracy, reduced false positives, and robust performance under varying traffic conditions. The system achieves near-perfect precision and recall, highlighting its effectiveness for real-world deployment.

Overall, the proposed framework offers a scalable and efficient solution for intelligent traffic monitoring and contributes toward enhancing road safety in smart city environments.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to their project guide and faculty members for their continuous support, valuable guidance, and insightful suggestions throughout the development of this work. Their encouragement and expertise greatly contributed to the successful completion of this project.
conducting this research. Special appreciation is extended to peers and colleagues for their constructive feedback and assistance during various stages of the project.

REFERENCES

[1] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
[2] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.
[3] G. Jocher et al., "YOLOv5," *GitHub repository*, 2020.
[4] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proc. IEEE ICIP*, 2017, pp. 3645–3649.
[5] Y. Zhang, P. Sun, Y. Jiang, et al., "ByteTrack: Multi-object tracking by associating every detection box," in *Proc. ECCV*, 2022.
[6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017.
[7] W. Liu et al., "SSD: Single shot multibox detector," in *Proc. ECCV*, 2016, pp. 21–37.
[8] R. Girshick, "Fast R-CNN," in *Proc. IEEE ICCV*, 2015, pp. 1440–1448.
[9] R. Wang, J. Zhang, and Y. Ma, "Helmet detection based on improved YOLOv3," *IEEE Access*, vol. 8, pp. 212889–212897, 2020.
[10] L. Yu, Z. Han, and Y. Yang, "Deep learning-based helmet detection for safety applications," *J. Computing in Civil Engineering*, vol. 34, no. 6, 2020.
[11] H. Li, P. Wang, and C. Shen, "Towards end-to-end car license plate detection and recognition with deep neural networks,"*IEEE Trans. Intelligent Transportation Systems*, vol. 20, no. 3, pp. 1126–1136, 2019.
[12] S. Zherzdev and A. Gruzdev, "LPRNet: License plate recognition via deep neural networks," *arXiv preprint arXiv:1806.10447*, 2018.
[13] M. Lakhani, S. Dave, and J. Patoliya, "IoT-based adaptive traffic management system," in *Proc. IEEE Int. Conf.*, 2018.
[14] A. Yadav et al., "Adaptive traffic management using machine learning," *Journal of Electrical Technology*, vol. 4, pp. 221–226, 2019.