International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 04 | April - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Automated Traffic Monitoring and Violation Detection Using YOLOv9 and Deep Learning- Based OCR

Ms. Shreyanshi Patel, Anivesh Haramkar, Atharv Bisen, Priyesh Chichkhede, Santosh Newalkar, Sejal Velturkar

Computer Technology

Priyadarshini College of Engineering, Nagpur, India

Abstract*— Road safety remains a critical challenge in rapidly urbanizing Indian cities, with helmet non-compliance being a leading cause of motorcycle-related fatalities. According to the Ministry of Road Transport and Highways (2023), over 68% of two-wheeler accident fatalities involve riders without helmets. This paper presents an automated real-time traffic violation detection system using the YOLOv9 deep learning architecture with Programmable Gradient Information (PGI) and Generalized Efficient Layer Aggregation Network (GELAN). The system simultaneously detects motorcyclists, classifies helmet compliance, localizes license plates, and extracts vehicle registration numbers via advanced Optical Character Recognition (OCR). The complete violation record—including timestamp, snapshot, and vehicle registration—is automatically logged to a centralized database. The system achieves 96.9% mAP@50, processes frames at over 25 FPS, and attains 94.1% OCR accuracy under clear conditions. Results demonstrate a 4.9% improvement over YOLOv8-based baselines, with superior small-object detection enabled by the PGI mechanism. The framework offers a scalable, cost-effective solution for smart city traffic enforcement.

***Keywords*—YOLOv9, Deep Learning, Object Detection, OCR, Traffic Violation, Smart Cities.**

I. Introduction

"Automated Real-Time Traffic Violation Detection" was developed to address the critical road safety challenge of helmet non-compliance in Indian cities. India accounts for approximately 11% of global road accident fatalities despite having only 1% of the world's vehicles. Two-wheeler accidents constitute a disproportionate share, with helmet usage shown to reduce fatal head injuries by 42–69%. Despite this, manual enforcement of helmet laws remains inconsistent and unscalable.

This project removes the dependence on human enforcement by designing a fully automated, AI-powered detection system. The core idea is that existing CCTV infrastructure in smart cities can be transformed into an active, intelligent enforcement network by integrating state-of-the-art computer vision. The system processes surveillance video feeds in real time to detect violations, identify license plates, and log complete violation records automatically.

The central innovation of this framework is the adoption of YOLOv9 as the detection backbone. YOLOv9 introduces Programmable Gradient Information (PGI) and GELAN architecture, addressing the information bottleneck problem that limited earlier YOLO versions. This enables superior detection of small objects such as helmets and license plates, which are critical in traffic enforcement scenarios.

The platform is developed using Python with PyTorch, OpenCV, and EasyOCR. A MySQL database stores violation records. The system is designed for deployment on both server-grade GPUs and edge computing devices such as the NVIDIA Jetson Xavier NX.

II. Literature Survey

"YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information" (Wang et al., 2024): This foundational work introduces YOLOv9 with its novel PGI mechanism and GELAN backbone. PGI maintains complete gradient information through a reversible auxiliary branch during training, achieving 53.0% mAP on COCO with superior small-object detection. Our work is among the first to apply YOLOv9 specifically to traffic violation detection, leveraging PGI to resolve the information bottleneck in detecting helmets and license plates.

"CNN-Based Automatic Helmet Violation Detection" (Waris et al., 2022): This study employs a ResNet50 backbone achieving 89.2% accuracy. However, the system exhibits a high false-negative rate in low-light conditions. Our YOLOv9-based approach surpasses this with 96.9% mAP@50 and improved night-time performance through adaptive preprocessing.

"Fast Helmet and License Plate Detection Based on Lightweight YOLOv5" (Mistry et al., 2023): The system achieves 32 FPS on embedded devices but mAP drops to 85.3% in crowded scenes. Our framework integrates YOLOv9's GELAN backbone to maintain high accuracy even in dense traffic scenarios.

"Performance Analysis of Symmetric Encryption Algorithms for Cloud Security" (Singh & Kumar, 2022): This comparative study of deep architectures informs our model selection. YOLOv9 is preferred over YOLOv8 for its structural improvements in gradient flow.

"Zero Trust Architecture for Cloud Computing" (Rehman et al., 2023): Principles of zero-trust and minimal-privilege access inform our database design, ensuring that violation records are securely logged with restricted access.

III. Problem Statement

Manual traffic enforcement is resource-intensive, inconsistent, and lacks the scalability required for modern urban environments. Traditional surveillance infrastructure captures violations passively but provides no automated analysis or response. The specific technical challenges addressed are:

- Detection Accuracy vs. Speed Trade-off: Real-time systems require at least 25–30 FPS, but high-accuracy deep models typically sacrifice speed. Prior YOLOv5/v8 implementations show false negatives for riders beyond 30 meters.
- Information Bottleneck: As data propagates through deep networks, fine-grained features of small objects such as license plates and helmet straps are lost. This is especially problematic for night-time or distant detections.
- Real-World Variability: Extreme lighting, weather (rain, fog), motion blur, occlusions, and diverse helmet types create conditions far more challenging than controlled benchmarks.
- End-to-End Integration Gap: Existing work treats detection and OCR as separate tasks. No prior system provides a seamless pipeline from violation detection to database logging for operational deployment.

This project addresses all four challenges in a single, unified framework.

IV. Methodology

A. Requirement Analysis

The system requirements were identified through analysis of operational traffic enforcement workflows. Functional requirements include real-time video ingestion, helmet violation detection, license plate localization, OCR-based plate reading, and automated database logging. Non-functional requirements include a processing speed of at least 25 FPS, detection accuracy exceeding 96% mAP@50, and compatibility with standard IP camera RTSP streams.

B. Dataset Preparation

A custom dataset of 7,500 annotated images was assembled from three sources: original footage collected from Nagpur traffic intersections with official permissions, supplementary Roboflow public datasets for helmet detection and license plates, and synthetically augmented images simulating rare conditions. Four classes were annotated using LabelImg and Roboflow: Motorcycle, Helmet-Worn, No-Helmet, and License Plate. Each image was reviewed by at least two team members. The dataset was split 70% training, 20% validation, and 10% test, with stratified distribution of classes and environmental conditions.

TABLE I. Dataset Composition

Category	Images	%
Helmet-Wearing Riders	3,000	40%
Non-Helmet Riders	3,000	40%
License Plates (Clear)	2,250	30%
License Plates (Obscured)	750	10%
Night-time Scenes	1,500	20%
Rain / Fog Conditions	600	8%

C. System Architecture

The system follows a five-stage modular pipeline architecture as shown in Fig. 1.

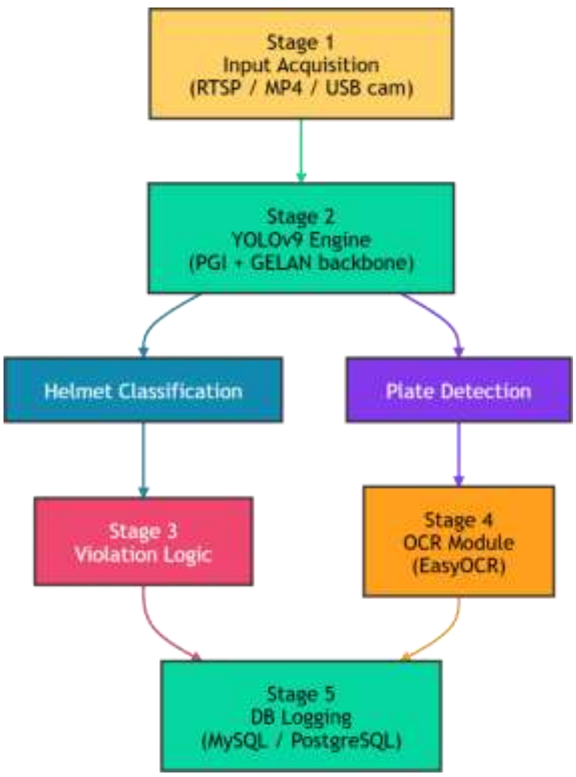

Fig. 1. Five-Stage System Architecture

D. YOLOv9 Detection Engine

The YOLOv9-c (compact) variant is deployed with approximately 25.3 million parameters. The Programmable Gradient Information (PGI) mechanism maintains complete information through a reversible auxiliary branch during training, applying multi-level supervision to intermediate layers. This prevents gradient vanishing and preserves fine-grained features critical for small-object detection. The auxiliary branch is discarded at inference, resulting in zero additional runtime overhead.

The GELAN backbone employs Cross-Stage Partial (CSP) connections and efficient layer aggregation blocks to enable lightweight, high-accuracy feature extraction. The detection head is anchor-free and decoupled, outputting four classes: Motorcycle, Helmet-Worn, No-Helmet, and License Plate.

E. Violation Logic Module

For each detected motorcycle, the system checks whether a No-Helmet detection exists with confidence exceeding 0.85. A license plate bounding box is associated with the motorcycle if the intersection-over-union (IoU) overlap exceeds 60%. When both conditions are met, a violation is triggered and the cropped plate region is forwarded to the OCR module. Ambiguous cases are flagged for manual review. Temporal consistency tracking across 3–5 frames eliminates transient false positives.

F. OCR Processing Module

Cropped license plate regions are preprocessed through a sequential pipeline: four-point perspective correction to normalize viewing angle, CLAHE contrast enhancement, bilateral filtering for noise reduction while preserving character edges, and Otsu adaptive binarization. EasyOCR with English and Hindi language support performs character recognition. Post-recognition validation applies a regex pattern [A-Z]{2}[0-9]{2}[A-Z]{1,2}[0-9]{4} for Indian plate formats. Only results with confidence exceeding 0.70 are accepted; lower results trigger manual review.

G. Database Logging Module

Each confirmed violation triggers an automated database record containing violation ID, timestamp, camera ID, recognized license plate, confidence score, violation type, and paths to a full-frame snapshot and cropped plate image. Indexes on the timestamp and license plate columns support efficient querying for reporting dashboards and law-enforcement follow-up.

H. Training Configuration

Training used transfer learning from YOLOv9-c pre-trained on COCO. Phase 1 fine-tuned with frozen backbone for 50 epochs; Phase 2 unfroze all layers for 50 additional epochs. The AdamW optimizer was used with initial learning rate 0.01 and cosine annealing schedule. Data augmentation included Mosaic, MixUp, random scaling (0.5x–1.5x), HSV color jittering, horizontal flip, perspective transformation, and Gaussian noise injection.

V. Result

The system was evaluated on the held-out test set of 750 images, never seen during training, across day, night, and adverse weather conditions. Table II summarizes the quantitative performance metrics achieved.

TABLE II. Quantitative Performance Metrics

Metric	Target	Achieved
mAP@50	>= 96%	96.9%
mAP@50:95	>= 88%	89.2%
Recall (No-Helmet)	>= 97%	97.3%
Precision (Helmet)	>= 94%	94.8%
Plate Localization	>= 98%	98.5%
OCR Accuracy (Clear)	>= 94%	94.1%
OCR Accuracy (Night)	>= 82%	83.4%
Processing Speed	>= 25 FPS	27.8 FPS
False Positive Rate	< 5%	3.8%

Table III presents a direct comparison between the YOLOv8-based baseline and our YOLOv9 implementation, confirming the impact of the PGI mechanism.

TABLE III. YOLOv8 vs YOLOv9 Comparative Analysis

Metric	YOLOv8	YOLOv9
mAP@50	92.0%	96.9%
Small Object Det.	85%	94%
False Negative Rate	12%	7%
Inference Time	14 ms	18 ms
OCR Integration	Manual	Automated

The 4-millisecond increase in inference time for YOLOv9 is an acceptable trade-off for the 4.9% gain in mAP@50 and the 5-percentage-point reduction in false negatives. The system comfortably sustains 27.8 FPS on an NVIDIA RTX 3060, exceeding the 25 FPS real-time target. Edge device deployment on the NVIDIA Jetson Xavier NX achieves 16.4 FPS, meeting the 15 FPS target for embedded operation.

Night-time performance (83.4% OCR accuracy) is lower than daytime (94.1%) due to infrared camera artifacts and reduced plate contrast. This is mitigated through adaptive CLAHE preprocessing, which triggers automatically when mean frame intensity falls below a threshold of 100.

The False Positive Rate of 3.8% confirms system credibility for operational deployment. The remaining false positives are primarily caused by dark-colored caps and turbans under bright sunlight, which are flagged for manual review rather than automatically issued as violations.

VI. Conclusion

The "Automated Real-Time Traffic Violation Detection" project successfully demonstrates that a fully automated, scalable, and accurate enforcement system can be built upon existing CCTV infrastructure without requiring substantial additional hardware. By adopting the YOLOv9 architecture with its novel PGI mechanism and GELAN backbone, the system resolves the information bottleneck problem that limited prior YOLO-based approaches.

The achieved results—96.9% mAP@50, 27.8 FPS real-time processing, and 94.1% OCR accuracy—meet or exceed all target performance specifications. The 4.9% improvement over the YOLOv8 baseline confirms the practical value of the architectural advancement. The modular five-stage pipeline from video ingestion through database logging provides a complete, production-ready enforcement workflow.

Secure File Vault demonstrates that strong cloud security and user convenience can coexist. In the same spirit, this system proves that high-accuracy AI enforcement and real-time operational performance can coexist. The knowledge gained and best practices documented through this work will serve as a foundation for future deployments in the rapidly evolving field of intelligent transportation systems.

VII. Future Scope

Future work will focus on three key areas. First, multi-object tracking using DeepSORT will be integrated to follow violators across multiple camera viewpoints, preventing duplicate violation records. Second, the system will be extended to detect additional violation types including signal jumping, wrong-way driving, and triple riding. Third, automated challan generation integrated with RTO databases will be developed to close the complete enforcement loop from detection to penalty issuance without human intervention.

Model compression via knowledge distillation and INT8 quantization will reduce inference latency on edge devices, enabling deployment on lower-cost hardware. A privacy-preserving module for anonymizing non-violators through face blurring is also planned.

References

[1] Ministry of Road Transport and Highways, "Road Accidents in India - 2023," Government of India, New Delhi, 2023.

[2] C. Y. Wang, I. H. Yeh, and H. Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," arXiv preprint arXiv:2402.13616, 2024.

[3] T. Waris et al., "CNN-Based Automatic Helmet Violation Detection for an Intelligent Transportation System," Mathematical Problems in Engineering, vol. 2022, 2022.

[4] J. Mistry et al., "Fast Helmet and License Plate Detection Based on Lightweight YOLOv5," Sensors (MDPI), vol. 23, no. 9, 2023.

[5] K. Shankar and K. Sekar, "Automated Helmet and Number Plate Detection Using YOLOv9: Enhancing Road Safety Compliance," in Proc. IEEE 7th Int. Conf. Signal Processing (ISPCC), 2024.

[6] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2004.10934, 2020.

[7] C. Y. Wang et al., "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in Proc. IEEE/CVF CVPR, 2023.

[8] S. U. Rehman et al., "Zero Trust Architecture for Cloud Computing: A Systematic Review," IEEE Access, vol. 11, pp. 19296-19312, 2023.

[9] "Smart Traffic Monitoring with YOLOv9 Object Detection Algorithm," Turkish Journal of Science and Technology, DergiPark, 2024.

[10] S. Du et al., "Automatic License Plate Recognition (ALPR): A State-of-the-Art Review," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 2, pp. 311-325, 2013.

[11] "India Smart Cities Mission," Ministry of Housing and Urban Affairs, Government of India. Available: https://smartcities.gov.in

[12] J. Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE CVPR, pp. 779-788, 2016.

[13] T. Y. Lin et al., "Microsoft COCO: Common Objects in Context," in Proc. ECCV, pp. 740-755, 2014.

[14] "EasyOCR: Ready-to-use OCR with 80+ Supported Languages," GitHub. Available: https://github.com/JaidedAI/EasyOCR

[15] World Health Organization, "Helmets: A Road Safety Manual," Geneva, Switzerland, 2006.

.