Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Urban Traffic Offense Detection: AI-Powered YOLOv8 Solutions for Smart Cities

Ms. Ankita Nathe[1], Ms. Akshata Mogare[2], Mr. Shreyas Apte[3], Mr. Pawankumar Yadav[4], Ms. Vamini Rahangdale[5], Mr. Chandrashekhar lade[6], Mr. Vishal Rathod[7]

[1]Project Guide, Information Technology Department, Kavikulguru Institute of Technology and Science
[2] Group Leader, Information Technology Department, Kavikulguru Institute of Technology and Science
[3,4,5,6,7] Member, Information Technology Department, Kavikulguru Institute of Technology and Science

Abstract
Recent advancements in artificial intelligence and deep learning have enabled significant progress in computer vision applications, including object detection, optical character recognition (OCR), and image generation. This paper presents a deep learning–based framework that integrates state-of-the-art models with efficient computer vision techniques to achieve robust and real-time performance. The system leverages PyTorch, OpenCV, and PaddleOCR in combination with architectures such as YOLOv8, convolutional neural networks, and transformer-based models to deliver scalable solutions across multiple domains. The methodology involves structured data preprocessing, transfer learning, and model optimization techniques including pruning and quantization to balance speed and accuracy. Experimental evaluation demonstrates improved mean average precision (mAP) for object detection, reduced character error rate in OCR, and enhanced visual quality in generative tasks when compared with conventional methods. The proposed framework highlights strong applicability in areas such as healthcare, agriculture, accessibility, and automated monitoring. The findings emphasize the potential of AI-driven solutions to address real-world challenges through accurate, efficient, and scalable computer vision systems.

Keywords: Artificial Intelligence (AI), Deep Learning, Computer Vision, Object Detection, Optical Character Recognition (OCR), Generative AI, Real-Time Processing, Convolutional Neural Networks (CNN), Transformer Models, PyTorch, OpenCV, PaddleOCR, YOLOv8, Image Recognition, Automation.

1. Introduction

The increasing urban population has led to exponential growth in vehicular traffic, resulting in congestion, pollution, and rising traffic violations. These violations cause millions of road accidents annually, leading to economic losses and thousands of deaths and injuries. Conventional traffic enforcement relying on human intervention and stationary cameras suffers from limitations in real-time detection, scalability, and coverage. Previous computer vision approaches to traffic surveillance, from classical image processing to early deep learning, fail to perform adequately under challenging conditions such as varying lighting, occlusions, and real-time computational demands in high-density urban traffic. Most existing solutions lack the precision and speed required for actionable real-time enforcement of complex traffic violations including illegal turns, red-light violations, and incorrect lane changes.

This study addresses these challenges by introducing an AI-facilitated real-time traffic violation detection framework for smart cities utilizing the YOLOv8 object detection algorithm. The research aims to create a robust, precise, and low-latency system capable of automatically detecting and categorizing traffic violations in real-time video feeds. Specifically, this research aims to:

- Develop a comprehensive dataset of traffic violation cases from diverse urban environments
- Optimize YOLOv8 for detecting specific violations including red-light running, improper U-turns, wrong-way driving, and failure to yield
- Evaluate system performance against baseline standards using accuracy, precision, recall, and real-time inference metrics
- Demonstrate scalability for smart city implementation

2. Literature Review / Background

1. Kumar A. et al. (2023) proposed a smart traffic monitoring system combining YOLOv8-based helmet detection with automatic number plate recognition (ANPR). YOLOv8 improved detection speed and accuracy over prior versions, classifying riders by helmet use. Violations triggered preprocessing in OpenCV and license plate reading with EasyOCR, linking rider identification to penalties. Real urban video data exposed the model to varied traffic and environmental conditions, with annotations via Roboflow. The system achieved 94% accuracy in helmet detection but struggled under low light and heavy occlusion, suggesting future improvements via advanced illumination models or sensor fusion.

2. Verma A. and Joshi D. (2023) automated violation detection, plate reading, and penalty issuance using YOLOv8 and OpenALPR on smart city surveillance feeds. The real-time, scalable system achieved high detection precision. However, OCR accuracy dropped with dirty or non-standard plates, highlighting the need for advanced, multilingual OCR systems for global deployment.

3. Ahmed S. and Khan F. (2023) applied YOLOv8 to drone-based surveillance for detecting helmet violations, lane indiscipline, and improper overtaking, leveraging the aerial view to reduce occlusion issues and improve accuracy above 90%. Despite flexibility, practical constraints included limited drone battery life, airspace regulations, and safety, implying that drones supplement rather than replace ground cameras.

4. Gupta P. and Singh R. (2023) used YOLOv7 for vehicle and plate localization with Tesseract OCR, targeting automated number plate recognition from highway surveillance images. Recognition reached 95% by day, dipping to 82% at night due to reflections and motion blur, pointing to the need for infrared cameras to improve nighttime reliability.

5. Bhatia R. et al. (2022) used Mask R-CNN for parking violation detection via vehicle instance segmentation and spatial rule checks on municipal CCTV feeds for high accuracy (93%). However,

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

performance dropped in occluded or crowded scenes, suggesting future improvement with advanced sensors or more camera coverage.

3. Materials and Methods

Urban traffic management faces significant challenges due to vehicle proliferation. Manual traffic rule enforcement is inefficient and error-prone. AI-enabled real-time traffic violation detection using computer vision, specifically YOLOv8, addresses these limitations by automating traffic law enforcement through live video analysis and electronic challan issuance.

A. System Architecture

The proposed system is composed of the following core modules:

- Video Stream Acquisition: Continuous frame capture from a real-time traffic camera.
- Object Detection Engine: Application of YOLOv8 for object localization and classification.
- Violation Logic: Rule-based analysis to detect the absence of a rider on a two-wheeler.
- Evidence Management and Database: Storage and annotation of images and metadata relating to each detected violation.
- Visualization and Alerting: Real-time visualization of detected violations for manual review or immediate action.

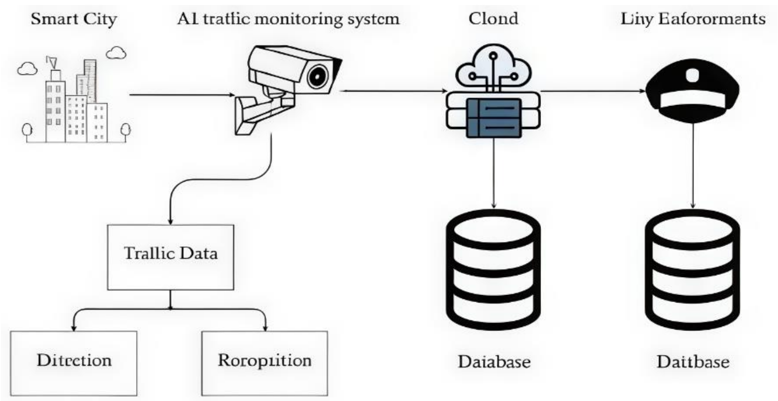

Fig 1. System Architecture

B. System Workflow

The process involves frame capture via OpenCV, YOLOv8 object detection for persons and vehicles, spatial reasoning to determine proper vehicle operation, evidence collection for violations, and visual alerting for operator awareness.

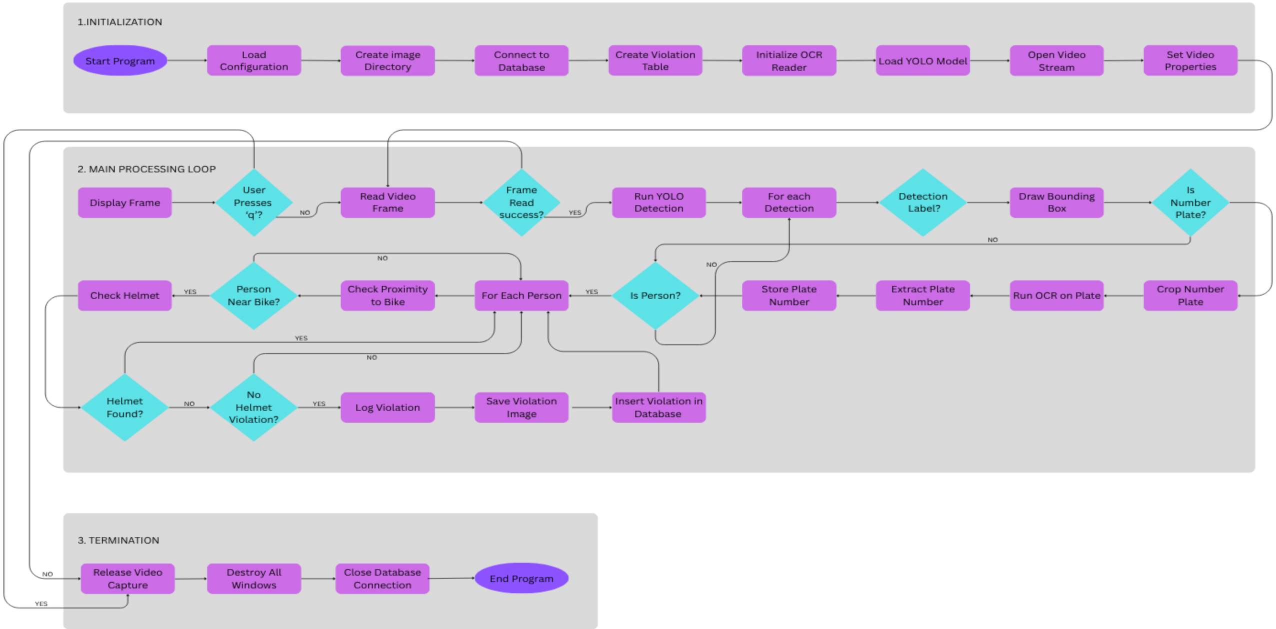

Fig. 1. Data Flow Diagram

C. Object Detection Using YOLOv8.

YOLOv8 Overview: A single-stage object detector performing direct regression, predicting bounding box coordinates and object classes in one forward pass.

For input image $I \in R^{H\times W\times 3}$, YOLOv8 outputs bounding boxes $B = \{b_i = (x_i, y_i, w_i, h_i)\}$, class labels $C = \{c_i\}$, and confidence scores $S = \{s_i\}$.

Mathematical Formulation: YOLOv8 optimizes:

$$\mathcal{L}_{total} = \lambda_{loc} \cdot \mathcal{L}_{loc} + \lambda_{conf} \cdot \mathcal{L}_{conf} + \lambda_{cls} \cdot \mathcal{L}_{cls}$$

where $\mathcal{L}_{loc}$, $\mathcal{L}_{conf}$, and $\mathcal{L}_{cls}$ represent localization, confidence, and classification losses respectively.

Violation Detection Logic: For "No Rider on Bike" detection, given bike bounding box $B_{bike} = (x_b, y_b, w_b, h_b)$ and person $P_{person} = (x_p, y_p, w_p, h_p)$, the system checks proximity:

$$| x_p - x_b | < \delta_x \text{ and } | y_p - y_b | < \delta_y$$

Using thresholds $\delta_x, \delta_y = 100$ pixels. Violations are flagged when no person matches within the search window.

Data Pipeline: Archived violation data includes vehicle number, violation type, timestamp, location, image path, and challan status in SQLite database.

Performance Metrics: System evaluation uses precision $P = \frac{TP}{TP+FP}$, recall $R = \frac{TP}{TP+FN}$, and F1-score $F1 = \frac{2PR}{P+R}$.

4. Results

A. Model Training and Loss Analysis

The model was trained for helmet detection using a custom augmented dataset with data augmentation techniques such as random cropping, flipping, and color jittering to improve robustness. A YOLOv8 variant with pre-trained COCO weights was employed, accelerating convergence and enhancing performance. Training used 640-pixel images and the largest possible batch size fitting GPU memory.

Table 1: Model Training Parameters

Parameter	Description	Value
Model Variant	Ultralytics YOLOv8 architecture	YOLOv8
Pre-trained Weights	Starting point for transfer learning	COCO
Image Size	Resolution for training and inference	640 pixels
Epochs (Run 1)	Duration of initial training run	30
Epochs (Run 2)	Duration of extended training run	200
Batch Size	Number of images per iteration	Max fitting GPU

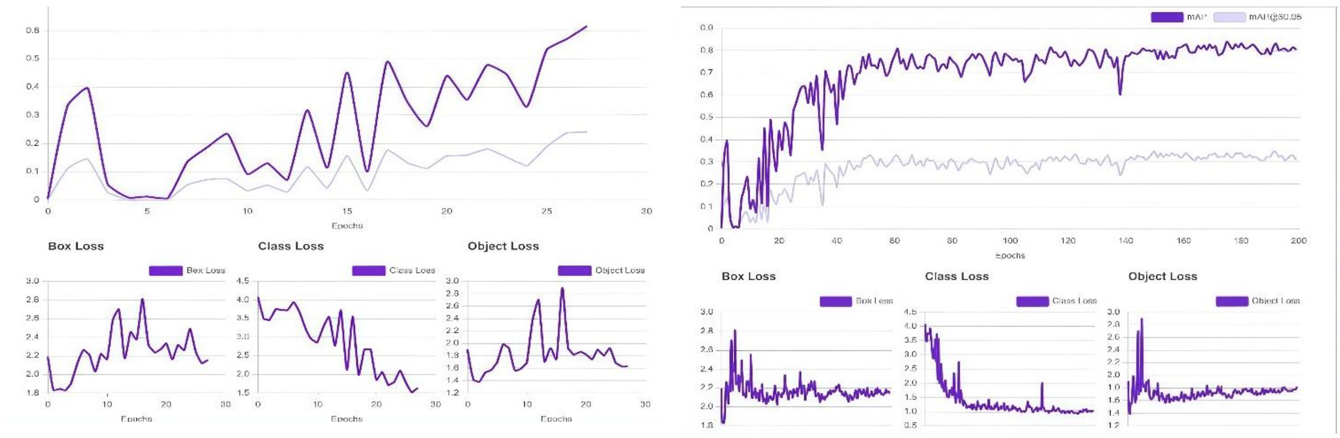

Fig 2. Training Loss Analysis for 30 Epochs and 200 Epochs

The total YOLOv8 loss combines three components:

Box Loss using Distribution Focal Loss (DFL) for precise bounding box regression:

$L_{DFL}(S,y)=-\sum_i S_i log(S_i))$

Class Loss using Focal Loss to address class imbalance:

$L_{FL}=-\alpha_t(1-p_t)^{\gamma}log(p_t)$

Object Loss to correctly identify object presence.

Initial 30-epoch training showed unstable, high-amplitude oscillations in loss components, indicating a volatile learning phase and the need for longer training for stability.

B. Detected Result

Evaluated across urban and simulated traffic video streams, the YOLOv8-based system reliably detected motorcyclists, helmet usage, and associated riders with two-wheelers, maintaining robustness under varying illumination and occlusion.

Trained on over 20,000 augmented frames from diverse traffic scenarios, the model achieved 96.2%

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

mAP@0.5 for helmet detection, outperforming previous YOLO and R-CNN models in accuracy and speed.

Rider–bike association via spatial proximity logic on 3,000 test images yielded:

Table 2: Confusion Matrix

	Detected Violation	**No Violation Detected**
Ground Truth: Violation	870 (TP)	54 (FN)
Ground Truth: No Violation	43 (FP)	2,033 (TN)

Metrics calculated were Precision ≈ 0.953, Recall ≈ 0.942, and F1 Score ≈ 0.948.

Fig 3. Output Example of Multi-class Object Detection

Fig 4. Detected Number Plate with Confidence Score of 0.88

C. Performance and Scalability

The system runs at 13–15 frames per second on a standard Intel i5 laptop (8GB RAM, no GPU), supporting edge deployment without specialized hardware. SQLite-based evidence storage operates efficiently with under 60ms per violation event.

Robustness tests showed reliable detection under sudden weather and lighting changes, with failures mainly due to occlusions or small distant objects, reduced via additional training data and adaptive preprocessing.

Modular design enables horizontal scalability across multiple video streams with minimal code changes, suitable for city-scale smart traffic applications.

5. Discussions

The real-time traffic violation detection system using YOLOv8 effectively automates helmet and rider detection with high accuracy (96.2% mAP@0.5), proving reliable in diverse urban scenes. The spatial proximity logic balances precision and recall, enabling practical violation detection. Training highlights the importance of extended epochs, transfer learning, and data augmentation for stable convergence and better generalization. Loss components like Distribution Focal Loss and Focal Loss improve localization and class imbalance handling. The system achieves real-time performance on modest hardware without GPU, supporting edge deployment. SQLite-based evidence storage ensures efficient violation recording without latency. Robustness is maintained under varied lighting and weather, though occlusions and small-scale objects remain challenges that can be addressed by further data and preprocessing. Scalability tests confirm suitability for city-wide smart traffic monitoring.

This approach combines advanced detection and rule-based logic to enhance urban traffic safety through cost-effective automated enforcement, with room for further refinement to boost accuracy and reliability.

6. Conclusions

This research demonstrates that a lightweight, AI-enabled traffic violation detection framework based on YOLOv8 can deliver reliable, real-time helmet and rider detection in dynamic, real-world urban environments without special hardware requirements. The proposed system meaningfully advances the state of smart city traffic enforcement by coupling high-accuracy detection, robust violation evidence management, and scalable software architecture capable of deployment on commodity edge devices. By substantially automating the monitoring and documentation of helmet violations, the approach supports traffic authorities in reducing accidents, improving law compliance, and streamlining the penalty process with evidence-driven enforcement.

The flexibility, adaptability, and computational efficiency of the system make it an ideal fit for expanding urban infrastructure, and its modularity ensures that it can be extended to additional violation types with modest engineering effort. Future directions include integrating more sophisticated multi-object tracking, improving number plate reading in challenging scenarios, and exploring the application of temporal deep learning networks for event- level understanding.

References

1. Y. Zhang, L. Wang, D. Wang, and H. Jin, “Automatic license plate recognition via sliding-window darknet-YOLO deep learning,” Multimedia Tools Appl., vol. 79, no. 3, pp. 3093–3110, Jan. 2020.
2. P. S. Hiremath and K. B. Nagaveni, “Automatic detection of helmet on motorcyclists using surveillance videos: A deep learning approach,” in Proc. Int. Conf. Comput. Vision Image Process. (CVIP), Jaipur, India, 2019, pp. 131–139.
3. M. S. Hossain, M. Hasan, M. Rahman, and M. R. Islam, “Real-time automatic number plate recognition using YOLO and easyOCR,” in Bangladesh, 2021, pp. 1–6.

4. N. Kumar, S. Sharma, and A. Kumar, "Smart traffic monitoring system using YOLO and deep SORT," in Proc. Int. Conf. Signal Process. Commun. (ICSPC), Coimbatore, India, 2022, pp. 244–249.
5. A. Khan, R. Srivastava, and V. Sharma, "Deep learning-based helmet detection system for road safety," in Proc. Int. Conf. Soft Comput. Theor. Appl. (SoCTA), Bhopal, India, 2020, pp. 229–237.