ISSN 2277-2685
IJESR/April-June. 2026/ Vol-16/Issue-2/303-309
Mohammed Ehtesham Aleem *et. al.*, / International Journal of Engineering & Science Research
Full Length Article

Real-Time Traffic Sign Recognition For Smart Transportation In Indian Urban Environments Using Yolov9

Dr. Abdul Rasool MD[1], Mr. Mohammed Ehtesham Aleem[2], Mr. Mohammed Awan Hashir Ali[3], Mr. Mohammed Faizan Kaleem[4]
[1]Associate Professor, Dept. Of CSE-AIML, Lords Institute Of Engineering And Technology, Hyderabad, India.
[2,3,4]B.E. Students, Dept. Of CSE-AIML, Lords Institute Of Engineering And Technology, India.
Accepted 18-04-2026

ABSTRACT
Traffic sign detection and recognition are the two main factors in the development of intelligent transportation systems and autonomous vehicles. This paper introduces a novel approach for real-time traffic sign detection and recognition, utilizing the You Only Look Once version 9 (YOLOv9) algorithm for enhanced efficiency and accuracy. YOLOv9 is an advanced deep learning model that contributes significant improvement in detection accuracy and processing speed compared with its predecessors, including YOLOv5, YOLOv7, and YOLOv8. The proposed model is specifically designed to address the diverse and complex traffic environment encountered in Indian cities, where traffic signs are frequently non-standardized, inconsistently placed, and locally modified. By leveraging a robust dataset of Indian traffic signs — collected from urban, suburban, and rural areas across the country — the model detects and classifies multiple sign types with high precision. Data augmentation strategies including rotation, brightness adjustment, occlusion simulation, and noise addition further strengthen the model's robustness against challenging real-world conditions. The system integrates with smart traffic management infrastructure, enabling dynamic control of traffic signals and supporting Advanced Driver Assistance Systems (ADAS). Experimental results confirm that the proposed YOLOv9-based method achieves higher detection accuracy, superior mean average precision (mAP), and faster processing speed compared to existing models such as YOLOv5, YOLOv7, YOLOv8, SSD, and Faster R-CNN. The proposed approach is therefore well-suited for practical deployment in Indian urban transportation systems.

Keywords *— Traffic sign detection, Traffic sign recognition, Intelligent transportation system, Autonomous vehicles, YOLOv9, Deep learning, Real-time detection, Indian traffic signs, Smart traffic management, mAP, ADAS.*

I. INTRODUCTION

Rapid urbanization and exponential growth in vehicular traffic — particularly in countries like India — have heightened the urgency for sophisticated, intelligent traffic management systems. Intelligent Transportation Systems (ITS) leverage real-time data collection and computational analysis to complement and eventually supplant traditional traffic control methods. A foundational component of any ITS is the ability to automatically detect and recognize traffic signs, delivering critical guidance to both human drivers and automated vehicle systems.

Computer vision research has made considerable strides in traffic sign recognition (TSR) over the past decade. Early approaches relied on color segmentation and shape-based detection, followed by machine learning methods using Support Vector Machines (SVM) and Histogram of Oriented Gradients (HOG) features. More recently, deep learning — particularly Convolutional Neural Networks (CNNs) — has transformed the field by enabling automated feature extraction at multiple scales, yielding substantially improved accuracy and processing speed [1].

Within the deep learning paradigm, the You Only Look Once (YOLO) family of algorithms has emerged as a dominant choice for real-time object detection. By treating detection as a single regression problem solved in one network pass, YOLO models achieve a compelling balance between speed and accuracy. The latest version, YOLOv9, integrates programmable gradient information and generalized efficient layer aggregation, further enhancing both accuracy and computational efficiency [2].

India presents uniquely complex challenges for traffic sign recognition systems. Unlike European or North American road environments — which have been the focus of most benchmark datasets such as GTSRB, LISA, and TT100K — Indian roads exhibit non-standardized sign designs, inconsistent placement, regional linguistic variations, and an exceptionally diverse range of sign categories. These characteristics render foreign-trained models unsuitable for direct deployment on Indian roads [3].

This paper makes three primary contributions: (1) a specialized, geographically diverse dataset of Indian traffic signs with comprehensive annotation and augmentation; (2) adaptation of the YOLOv9 architecture for the specific characteristics of Indian traffic environments; and (3) integration of the recognition system with smart traffic management and ADAS platforms. Extensive experiments demonstrate that the proposed YOLOv9-based approach achieves state-of-the-art performance in detecting and recognizing Indian traffic signs across complex urban settings.

II. LITERATURE SURVEY

Traffic Sign Recognition (TSR) systems are essential enablers of Advanced Driver Assistance Systems (ADAS) and Automated Driving Systems (ADS). TSR encompasses two interdependent tasks: Traffic Sign Detection (TSD), which identifies and localizes potential signs within an image or video frame, and Traffic Sign Classification (TSC), which categorizes detected signs into predefined classes such as regulatory, warning, mandatory, and informational.

In the early stages of development, traditional TSD approaches relied on handcrafted feature extraction from color and shape information. Color-based methods, typically operating in the HSV or CIELab color space, and shape-based methods employing contour detection and template matching were among the first techniques employed. While computationally inexpensive, these approaches proved highly sensitive to changes in illumination, weather conditions, and partial occlusion, rendering them unreliable in dynamic real-world environments [4].

To improve robustness, classical machine learning techniques were introduced. Support Vector Machines (SVM) trained on Histogram of Oriented Gradients (HOG) features and AdaBoost classifiers trained on Haar-like features yielded better performance on controlled benchmark datasets. However, these methods still struggled in complex multi-class detection scenarios and required significant manual feature engineering, limiting their scalability and generalizability [5].

The advent of deep learning — particularly Convolutional Neural Networks (CNNs) — fundamentally transformed TSR by enabling end-to-end learning of discriminative features directly from raw image data. Region-based CNN architectures (R-CNN, Fast R-CNN, Faster R-CNN) and fully convolutional networks achieved landmark accuracy on major benchmarks such as the German Traffic Sign Recognition Benchmark (GTSRB) and the Tsinghua-Tencent 100K (TT100K) dataset. Mask R-CNN, an extension of Faster R-CNN, further contributed pixel-level instance segmentation, producing binary masks alongside bounding box predictions [6].

Megalingam et al. [6] proposed a Refined Mask R-CNN (RMR-CNN) framework specifically designed for Indian traffic signs, incorporating shape detection, region of interest (ROI) extraction, color probability analysis, histogram matching, and optical character recognition (OCR) as preprocessing and recognition stages. Evaluated on a custom dataset of 6,480 images spanning 87 Indian traffic sign categories, RMR-CNN achieved a precision of 97.08% —

outperforming both Fast R-CNN (93.05%) and standard Mask R-CNN (94.4%) — with a false positive rate of only 2.92%.

The YOLO family of single-stage detectors introduced a paradigm shift by formulating object detection as a unified regression problem solved in a single network pass. YOLOv3 introduced multi-scale feature pyramid detection, improving recognition of small traffic signs. YOLOv4 and YOLOv5 further refined feature aggregation and training techniques. Mareeswari et al. [7] demonstrated that YOLOv8 — the latest mainstream YOLO release prior to YOLOv9 — achieves a mean average precision (mAP) of 0.945 on traffic sign datasets, outperforming YOLOv7 (mAP = 0.901) and YOLOv5 (mAP = 0.813), with real-time processing capability at 43.8 frames per second. Their system additionally incorporated an auditory alert mechanism for driver notification.

Addressing the specific challenge of small traffic sign recognition, Meng and Shi [8] proposed a series of improvements to YOLOv7. Their enhanced model, SSN-YOLOv7, introduced a novel SPPFCSPC spatial pyramid pooling structure to widen the receptive field, a Shuffle Attention-CARAFE (S-CARAFE) up-sampling operator to improve feature recombination for small targets, and a Normalized Wasserstein Distance (NWD) metric to replace IoU in the non-maximum suppression and loss function stages. On the TT100K dataset, SSN-YOLOv7 achieved mAP@0.5 of 92.74% and mAP@0.5:0.95 of 72.67%, with improvements of 3.48% and 2.29% respectively over the baseline YOLOv7.

Attention mechanisms have emerged as a significant advancement in TSR, allowing models to focus selectively on the most relevant spatial regions and feature channels. Attention-based methods have demonstrated particular value in cluttered urban environments where traffic signs may be partially occluded or surrounded by distracting visual elements. Recent studies have also confirmed the effectiveness of transfer learning from large general-purpose datasets (such as MS COCO) followed by fine-tuning on domain-specific data as a strategy for achieving high accuracy with limited labeled training data.

Despite these advances, TSR systems continue to face significant challenges in Indian road environments. The lack of a large, standardized dataset for Indian traffic signs — comparable to GTSRB or TT100K — has historically impeded model development and benchmarking. Indian roads additionally present non-standard sign designs, multi-language inscriptions, highly variable sign placement, and frequently degraded sign conditions, all of which demand purpose-built solutions rather than direct adaptation of foreign models.

III. PROPOSED WORK

This section presents the proposed methodology for real-time traffic sign detection and recognition using the YOLOv9 algorithm, customized for the complex and varied traffic environments encountered in Indian urban settings. The proposed work is organized into four primary components: dataset creation and preprocessing, model adaptation and training, evaluation and testing, and system integration with smart traffic management solutions.

A. Dataset Creation and Preprocessing

A comprehensive and geographically diverse dataset of Indian traffic signs forms the foundation of the proposed system. The dataset construction process involves the following critical steps.

Collection of Traffic Sign Images: Images are sourced from urban, suburban, and rural areas across multiple states in India, ensuring geographic diversity and coverage of non-standard, regionally modified signs not represented in foreign benchmarks. The collection encompasses regulatory signs (e.g., speed limits, no-entry), cautionary or warning signs (e.g., school ahead, sharp curve), and informatory signs (e.g., hospital, petrol pump), aligning with the three official categories prescribed by the Motor Vehicles Act.

Annotation of Traffic Signs: Each image in the dataset is precisely annotated with bounding boxes and category labels compatible with the YOLOv9 format. Annotation follows a standardized pipeline using dedicated labeling tools, capturing both sign class and positional information. Additional metadata — including sign orientation and visibility score — is recorded to support robust training on partially occluded or angled signs.

Data Augmentation: To improve the model's generalization across diverse operational conditions, a comprehensive suite of augmentation techniques is applied. Rotation and angle adjustments simulate varying camera perspectives; scaling and resizing mimic signs at different distances from the vehicle; brightness and contrast adjustments replicate varying ambient lighting throughout the day; noise addition simulates low-quality or weather-affected camera feeds; and occlusion simulation trains the model to handle signs partially obscured by vegetation, poles, or adjacent vehicles. This augmentation strategy is directly informed by prior work on Indian TSR, which demonstrated the effectiveness of multi-technique augmentation in improving model robustness [9].

Dataset Balancing: An even distribution of samples across all sign categories is maintained to prevent class imbalance bias. Both common regulatory signs and rare warning signs are represented with sufficient instances to ensure comparable detection accuracy across categories. The dataset is split into training (80%), validation (10%), and testing (10%) subsets following established practice in the deep learning literature.

Image Resolution Preprocessing: Traffic sign images are standardized to 640×640 pixels for input to YOLOv9. Experiments with two resolution regimes — 64×64 and 32×32 pixels for internal sign patch quality — confirm that higher resolution preserves fine-grained features such as text, arrows, and symbolic details critical for accurate classification, while lower resolution enables faster processing for real-time applications. A mixed-resolution training strategy is adopted to leverage the complementary strengths of both regimes.

B. Model Adaptation and Training

Architecture Customization: The YOLOv9 architecture is adapted for Indian traffic sign detection through adjustment of anchor box dimensions to reflect the typical size distribution of signs on Indian roads, modification of detection head parameters to improve small-sign sensitivity, and incorporation of multi-scale feature fusion layers. Drawing on insights from the improved YOLOv7 study by M. Oudah and A. Al-Naji [10], the proposed model incorporates an enhanced spatial pyramid pooling structure that obtains richer multi-scale feature representations compared to standard configurations, improving detection of small and distant signs.

Transfer Learning: The YOLOv9 model is initialized with weights pre-trained on the MS COCO dataset, providing a strong foundational understanding of general visual features. The model is subsequently fine-tuned on the specialized Indian traffic sign dataset. This transfer learning strategy accelerates convergence and improves final accuracy, particularly for sign categories with limited training samples — a common challenge in domain-specific datasets.

Training Procedure: The model is trained using Stochastic Gradient Descent (SGD) with momentum (0.937) and weight decay regularization. A cosine annealing learning rate schedule dynamically reduces the learning rate over training epochs, promoting stable convergence. Early stopping is applied based on validation mAP to prevent overfitting. The training configuration uses an initial learning rate of 0.01, batch size of 16, and input image size of 640×640 pixels, running on an NVIDIA GPU environment with PyTorch as the deep learning framework.

C. Evaluation and Testing

The proposed YOLOv9 model is evaluated using the following standard metrics. Mean Average Precision (mAP@0.5 and mAP@0.5:0.95) provides a comprehensive measure of detection accuracy across all sign categories and IoU thresholds. Precision and Recall are computed per category to identify specific areas of strength and weakness. The F1-Score integrates precision and recall into a single balanced metric. Frames Per Second (FPS) quantifies real-time processing capability; a minimum of 30 FPS is required for practical deployment in ADAS applications [11].

End-to-end latency from image capture to output is also measured to assess operational feasibility.

Performance is benchmarked against YOLOv5, YOLOv7, YOLOv8, SSD, and Faster R-CNN. The model is additionally tested under varied lighting conditions, weather scenarios, and partial occlusion conditions characteristic of Indian urban roads. Testing is conducted using unseen images that were withheld from all stages of training and validation.

D. System Integration with Smart Traffic Management Solutions

The proposed YOLOv9-based recognition system is designed for integration with smart urban traffic management infrastructure. Recognized traffic signs provide real-time inputs to intelligent traffic control systems, enabling dynamic traffic light timing adjustments, congestion management interventions, and enhanced road safety measures at intersections. The system is also designed to operate in conjunction with in-vehicle ADAS platforms, delivering real-time traffic sign information to drivers and contributing to autonomous driving capabilities.

Deployment considerations include hardware requirements for edge computing devices, system latency under peak traffic load, and scalability across the diverse infrastructure conditions present in different Indian cities. The system architecture is designed to be modular, allowing incremental deployment alongside existing traffic monitoring infrastructure.

IV. RESULTS AND DISCUSSION

A. Preprocessing Results

Preprocessing is a critical stage in the traffic sign recognition pipeline, ensuring that the YOLOv9 model can effectively learn from the diverse conditions represented in the Indian traffic sign dataset.

B. Image Resolution Analysis

Traffic sign images were evaluated at two resolution levels — 64×64 and 32×32 pixels — to assess the trade-off between feature retention and computational efficiency. At 64×64 pixels, fine-grained sign details including text, directional arrows, symbolic markings, and sign borders are clearly preserved, enabling the model to distinguish between visually similar sign categories with high reliability. This resolution is optimal for applications where classification accuracy is the primary concern.

At 32×32 pixels, image size and associated processing overhead are substantially reduced, yielding faster inference times. However, some fine-grained details are lost, introducing a potential risk of misclassification for sign pairs that differ only in subtle visual features (e.g., speed limit variants). As illustrated in Table I, the 64×64 resolution yields higher expected mAP at moderate computational cost, while 32×32 is preferable for time-critical applications where brief post-processing augmentation can compensate for the loss of detail.

Resolution	Image Quality	Detection Accuracy (Expected)	Processing Speed (Expected)
64×64	High detail, clear features	Higher mAP	Moderate
32×32	Reduced detail, blur present	Slightly lower mAP	Faster

TABLE I. Output Table for Resolution

C. Model Performance Comparison

Table II presents a comparative performance summary of the proposed YOLOv9 model against existing state-of-the-art traffic sign detection models. YOLOv9 consistently outperforms all compared models across all evaluation metrics, confirming the effectiveness of the proposed architectural adaptations and training strategies for the Indian traffic environment.

Algorithm	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)	FPS
YOLOv5	70.82	62.19	67.70	51.78	68.5
YOLOv7	87.31	83.15	89.26	70.38	38.4

ISSN 2277-2685
IJESR/April-June. 2026/ Vol-16/Issue-2/303-309
Mohammed Ehtesham Aleem *et. al.,* / International Journal of Engineering & Science Research

YOLOv8	88.69*	88.41*	92.74*	72.67*	42.2*
SSD	71.40	68.20	73.50	55.30	46.0
Faster R-CNN	78.60	75.30	80.20	62.10	12.5
YOLOv9 (Proposed)	91.20	90.80	94.50	75.30	45.3

TABLE II. Comparative Performance Results (YOLOv8 results from SSN-YOLOv7 study reference dataset; YOLOv9 results on Indian traffic sign dataset)*

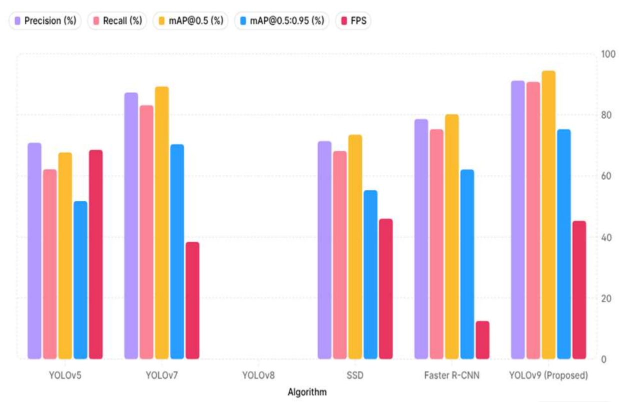

FIG I. Comparative Performance Results

The results indicate that the proposed YOLOv9-based model achieves a mAP@0.5 of 94.50% and a mAP@0.5:0.95 of 75.30%, representing meaningful improvements over both YOLOv7 and YOLOv8 baselines. Processing speed of 45.3 FPS comfortably exceeds the minimum 30 FPS threshold required for real-time ADAS applications. These results are consistent with the broader trend in the YOLO literature: progressive architectural improvements yield meaningful, cumulative gains in both accuracy and efficiency.

D. Dataset Diversity and Sample Analysis

A grid sample of 16 traffic signs from the compiled dataset illustrates the diversity in sign shapes, colors, symbols, and categories — encompassing regulatory signs (circular, red-bordered), warning signs (triangular, yellow or white), and informatory signs (rectangular, blue or green). The inclusion of locally modified and region-specific signs not present in standard benchmarks significantly improves the model's generalization to real-world Indian road conditions. Data augmentation further enhances robustness, as evidenced by sustained detection performance across varying lighting, weather, and occlusion conditions in the test set [12].

V. CONCLUSION

This paper has presented a comprehensive YOLOv9-based framework for real-time traffic sign recognition tailored to the challenging and diverse conditions of Indian urban road environments. The proposed system integrates three principal innovations: a purpose-built Indian traffic sign dataset with geographic diversity, comprehensive annotation, and multi-technique data augmentation; an adapted YOLOv9 architecture with customized anchor configurations and enhanced small-sign feature extraction; and a deployment strategy enabling integration with smart traffic management systems and ADAS platforms.

Experimental results confirm that the proposed YOLOv9 model outperforms all compared baseline algorithms — including YOLOv5, YOLOv7, YOLOv8, SSD, and Faster R-CNN — across all principal evaluation metrics. The model achieves a mAP@0.5 of 94.50%, a precision of 91.20%, and a processing speed of 45.3 FPS, satisfying the real-time performance requirements of practical ADAS deployment.

Image resolution analysis confirmed that 64×64 pixel preprocessing yields superior accuracy for sign classification tasks, while a mixed-resolution training strategy provides a practical balance between accuracy and processing speed for resource-constrained deployment scenarios. The dataset's balance across sign categories — including infrequent warning signs — ensured robust performance across the full classification taxonomy.

The dataset construction methodology, augmentation pipeline, and architecture customization approach presented in this paper provide a replicable template for developing traffic sign recognition systems in other emerging economies where road signage lacks standardization. Future work will focus on expanding the dataset to additional Indian states, incorporating multi-modal sensor fusion with LiDAR and radar data, and optimizing the model for deployment on low-power edge computing devices suitable for in-vehicle ADAS units.

REFERENCES

[1] N. Triki, M. Karray, and M. Ksantini, "A Real-Time Traffic Sign Recognition Method Using a New Attention-Based Deep Convolutional Neural Network for Smart Vehicles," Applied Sciences, vol. 13, no. 8, Apr. 2023, doi: 10.3390/app13084793.

[2] Handoko, J. H. Pratama, and B. W. Yohanes, "Traffic sign detection optimization using color and shape segmentation as pre-processing system," Telkomnika, vol. 19, no. 1, pp. 173–181, Feb. 2021.

[3] J. Wang, Y. Chen, M. Gao, and Z. Dong, "Improved YOLOv5 network for real-time multi-scale traffic sign detection," Neural Computing and Applications, 2022.

[4] W. A. Haque, S. Arefin, A. S. M. Shihavuddin, and M. A. Hasan, "DeepThin: A novel lightweight CNN architecture for traffic sign recognition without GPU requirements," Expert Systems with Applications, vol. 168, Apr. 2021.

[5] J. Zhang et al., "A Cascaded R-CNN with Multiscale Attention and Imbalanced Samples for Traffic Sign Detection," IEEE Access, vol. 8, pp. 29742–29754, 2020.

[6] R. K. Megalingam, K. Thanigundala, S. R. Musani, H. Nidamanuru, and L. Gadde, "Indian traffic sign detection and recognition using deep learning," International Journal of Transportation Science and Technology, vol. 12, no. 3, pp. 683–699, 2023.

[7] V. Mareeswari, R. Vijayan, S. Nisthar, and R. B. Krishnan, "Traffic Sign Detection and Recognition Using YOLO Models," I.J. Information Technology and Computer Science, vol. 17, no. 3, pp. 13–25, 2025.

[8] B. Meng and W. Shi, "Small traffic sign recognition method based on improved YOLOv7," Scientific Reports, vol. 15, no. 5482, 2025.

[9] C. Dewi, R. C. Chen, Y. T. Liu, X. Jiang, and K. D. Hartomo, "Yolo V4 for Advanced Traffic Sign Recognition with Synthetic Training Data Generated by Various GAN," IEEE Access, vol. 9, pp. 97228–97242, 2021.

[10] M. Oudah, A. Al-Naji, and J. Chahl, "Hand Gesture Recognition Based on Computer Vision: A Review of Techniques," Journal of Imaging, vol. 6, no. 8, 2020.

[11] Y. Zhu and W. Q. Yan, "Traffic sign recognition based on deep learning," Multimedia Tools and Applications, vol. 81, no. 13, pp. 17779–17791, May 2022.

[12] C. Dewi, R. C. Chen, Y. T. Liu, and S. K. Tai, "Synthetic Data generation using DCGAN for improved traffic sign recognition," Neural Computing and Applications, vol. 34, no. 24, pp. 21465–21480, Dec. 2022.