

DEVELOPMENT OF A COMPUTER VISION-BASED FRAMEWORK FOR AUTOMATED TRAFFIC SIGN RECOGNITION IN INTELLIGENT TRANSPORTATION SYSTEMS

Sarmad Hassan

Department of Computer Science, Jiangsu University of Science and Technology China

sarmadhassan48@gmail.com

DOI: <https://doi.org/10.5281/zenodo.23159833>

Keywords

Computer Vision, Traffic Sign Recognition, Deep Learning, Intelligent Transportation Systems, Image Processing, Autonomous Vehicles, Road Safety

Article History

Received: 25 June 2026

Accepted: 16 September 2026

Published: 05 October 2026

Copyright @Author

Corresponding Author: *

Sarmad Hassan

Abstract

The rapid development of intelligent transportation systems (ITS) has increased the demand for reliable and automated traffic sign recognition to support road safety, traffic management, and autonomous driving applications. However, variations in illumination, weather conditions, viewing angles, sign occlusion, image quality, and complex road environments continue to challenge the accurate recognition of traffic signs. This study aims to develop a robust computer vision-based framework for automated traffic sign recognition within intelligent transportation systems. The proposed framework integrates image preprocessing, traffic sign detection, feature extraction, and deep learning-based classification to automatically identify and categorize traffic signs from road-scene images. A labeled traffic sign dataset is employed for model training, validation, and testing, while performance is evaluated using accuracy, precision, recall, F1-score, mean average precision, and confusion-matrix analysis. The proposed framework is designed to combine reliable recognition performance with computational efficiency so that it can support real-time transportation applications. The expected findings indicate that deep learning-based computer vision can provide high recognition accuracy while maintaining reliable performance across diverse traffic sign categories and challenging visual conditions. The study further highlights the importance of data augmentation, multi-scale feature extraction, and lightweight architectures for deployment on resource-constrained intelligent transportation platforms. The proposed framework can contribute to automated road-scene understanding and support advanced driver-assistance systems and autonomous vehicles. Its practical implications include improved road safety, traffic efficiency, and intelligent mobility.

INTRODUCTION

Intelligent Transportation Systems (ITS) represent an important technological transition from conventional transportation infrastructure toward interconnected, data-driven, and increasingly automated mobility environments.

ITS integrates sensing technologies, communication networks, artificial intelligence, computer vision, and computational decision-making to improve transportation safety, efficiency, reliability, and sustainability. Within this technological ecosystem, automated

perception of road infrastructure is particularly important because vehicles must continuously interpret visual information before making appropriate driving decisions. Traffic signs constitute one of the most important elements of this visual information because they communicate speed restrictions, warnings, prohibitions, mandatory actions, directions, and other regulatory instructions to road users.

Traffic sign recognition (TSR) is consequently an important computer vision problem associated with advanced driver-assistance systems (ADAS) and autonomous vehicles. A reliable TSR system should identify traffic signs accurately despite variations in scale, illumination, orientation, weather, background complexity, image quality, and partial occlusion. These requirements make TSR substantially more difficult than recognizing isolated objects under controlled laboratory conditions. The German Traffic Sign Recognition Benchmark (GTSRB), for example, contains more than 50,000 images representing more than 40 classes and was specifically designed as a realistic multi-class recognition benchmark.

Traditional traffic sign recognition systems commonly relied on manually designed visual features, including color segmentation, edge detection, shape descriptors, histogram-based representations, and handcrafted texture features. Although such techniques established important foundations for computer vision-based TSR, their performance can deteriorate when signs are affected by shadows, illumination changes, weather, image noise, distortion, or complex backgrounds. The transition toward machine learning and deep learning has therefore transformed the field by enabling systems to learn discriminative representations directly from images.

Convolutional neural networks (CNNs) have been particularly influential because their hierarchical architecture allows the extraction of increasingly complex visual representations. Earlier layers can identify edges, contours, and simple patterns, whereas deeper layers can represent shapes, symbols, and semantic

characteristics. Deep learning-based approaches have consequently achieved highly competitive results on established traffic sign benchmarks. The original GTSRB competition reported a recognition rate of 99.46% for a committee of CNNs, compared with 98.84% human performance under the benchmark conditions.

However, benchmark accuracy alone does not guarantee reliable deployment in real transportation environments. A traffic sign that occupies only a small number of pixels in a road image presents a substantially different challenge from a centered sign in a classification dataset. Contemporary research therefore increasingly investigates detection and recognition together, particularly through object-detection architectures such as the YOLO family. Recent research has focused on improving the detection of small signs, extracting contextual information, and balancing accuracy with computational efficiency. For example, Tian and Zhang (2025) proposed STDE-YOLOv5s with enhanced feature fusion and contextual information extraction specifically to improve small traffic-sign detection. Another important issue concerns deployment. Autonomous vehicles and roadside intelligent devices may operate with limited computational resources, making excessively large deep neural networks impractical. Zhang et al. (2024), for example, proposed a lightweight traffic sign recognition system based on MobileNetV1 and reported substantial reductions in model parameters compared with several YOLO-based alternatives while maintaining competitive recognition performance. Their work also demonstrated deployment on a resource-constrained embedded hardware platform.

Against this background, the present study develops a computer vision-based framework for automated traffic sign recognition within ITS. The proposed framework emphasizes four interconnected objectives: accurate detection, reliable classification, robustness to real-world visual variation, and computational efficiency. Rather than treating recognition as an isolated classification problem, the study conceptualizes

TSR as a component of an integrated transportation perception pipeline.

Problem Statement

Despite considerable advances in computer vision and deep learning, automated traffic sign recognition remains challenging under real-world driving conditions. Traffic signs can be small relative to the complete road scene, partially hidden by vegetation or vehicles, distorted by perspective, affected by rain or fog, and exposed to significant illumination differences. Furthermore, visually similar signs may differ only in small symbols, numbers, or pictographic details. These factors can result in false detections and misclassification.

Existing benchmark-oriented models may also require substantial computational resources. Therefore, there remains a need for a framework that simultaneously addresses recognition accuracy, robustness, detection of small signs, and computational efficiency.

Research Objectives

The study aims to:

1. Develop a computer vision-based framework for automated traffic sign detection and recognition.
2. Investigate the effectiveness of deep learning for multi-class traffic sign classification.
3. Examine the contribution of image preprocessing and data augmentation to recognition performance.
4. Evaluate the framework using accuracy, precision, recall, F1-score, and detection metrics.
5. Examine the suitability of the proposed framework for intelligent transportation and real-time applications.

Research Questions

The study addresses the following questions:

1. How effectively can computer vision and deep learning identify traffic signs from road-scene images?
2. To what extent do preprocessing and data augmentation improve recognition performance?
3. Which traffic sign categories are most susceptible to classification errors?

4. How effectively can the proposed framework detect small or visually challenging signs?

5. Can the proposed framework provide an appropriate balance between recognition accuracy and computational efficiency?

Literature Review

Evolution of Traffic Sign Recognition

Traffic sign recognition has developed from rule-based image-processing methods to sophisticated deep neural networks. Early systems generally depended on the distinctive visual properties of traffic signs. Color segmentation was used to isolate red, blue, or yellow regions, while geometric analysis identified circular, triangular, rectangular, or octagonal structures. Such systems were relatively computationally inexpensive but highly dependent on predefined rules.

Machine learning introduced greater flexibility by allowing classifiers to learn from extracted features. Support Vector Machines, Random Forests, nearest-neighbor classifiers, and other algorithms were applied to traffic sign datasets. Nevertheless, these approaches typically required a separate feature-engineering stage. Their performance could therefore be constrained by the ability of researchers to select features capable of representing diverse traffic signs.

Deep learning significantly changed this paradigm. CNNs allowed feature extraction and classification to be optimized jointly. The GTSRB benchmark became particularly influential in evaluating such approaches. The benchmark includes more than 40 classes and more than 50,000 images, with realistic variations in visual appearance.

Recent surveys confirm the continuing movement toward AI-based TSR. A 2024 IEEE survey examined machine learning and deep learning approaches and emphasized challenges associated with damaged signs, changing illumination, and atmospheric conditions. It also identified GTSRB, the Belgium Traffic Sign Classification dataset, and Chinese Traffic Sign Database as important datasets for research.

Convolutional Neural Networks

CNNs are particularly suitable for traffic sign classification because their convolutional operations preserve spatial relationships between neighboring pixels. A typical CNN includes convolutional layers, activation functions, pooling layers, and fully connected or classification layers.

The convolution operation can be represented as:
$$[F_{\{i,j\}} = \sum_{\{m\}} \sum_{\{n\}} K_{\{m,n\}} X_{\{i+m,j+n\}} + b]$$

where (X) represents the input image, (K) represents a learnable convolution kernel, (b) is the bias term, and (F) represents the resulting feature map.

As the network becomes deeper, the learned representations become increasingly abstract. For TSR, this enables the system to learn relationships between low-level visual characteristics and high-level traffic sign categories.

A recent 2025 IEEE study used a CNN with brightness-histogram equalization for traffic sign image recognition and reported that preprocessing improved the reported accuracy from 97.26% to 98.91% on its experimental setup. This finding supports the importance of preprocessing in situations where illumination variation affects image quality.

Object Detection and YOLO-Based Recognition

Classification assumes that an image already contains an appropriately cropped object. Real-world transportation systems, however, typically process complete road scenes. Therefore, the system must first locate the sign before classifying it.

YOLO-based object detectors have become widely used because they formulate detection as a unified prediction problem and can achieve a useful balance between accuracy and speed. Previous traffic-sign studies have used YOLOv5 and related architectures for detection and recognition.

The continuing development of YOLO architectures has expanded their application to small-object detection, multi-scale feature extraction, attention mechanisms, and computational optimization. A recent comprehensive review of the YOLO family highlights traffic-sign detection as an important application because signs are often small and may occur under complex weather and environmental conditions.

Small-object detection remains one of the central technical challenges. Traffic signs can occupy only a small region of a camera frame, especially when the vehicle is distant from the sign. This creates a loss of fine-grained information during repeated feature downsampling. Recent architectures therefore emphasize multi-scale feature fusion and contextual information.

Lightweight Deep Learning

Although large networks may provide high accuracy, transportation applications frequently require low latency and limited computational consumption. This is particularly important for embedded vehicle systems, roadside cameras, and mobile platforms.

Lightweight architectures such as MobileNet address this issue through computationally efficient convolutional operations. Recent TSR research has demonstrated that lightweight networks can substantially reduce model parameters and computational complexity while retaining competitive recognition performance.

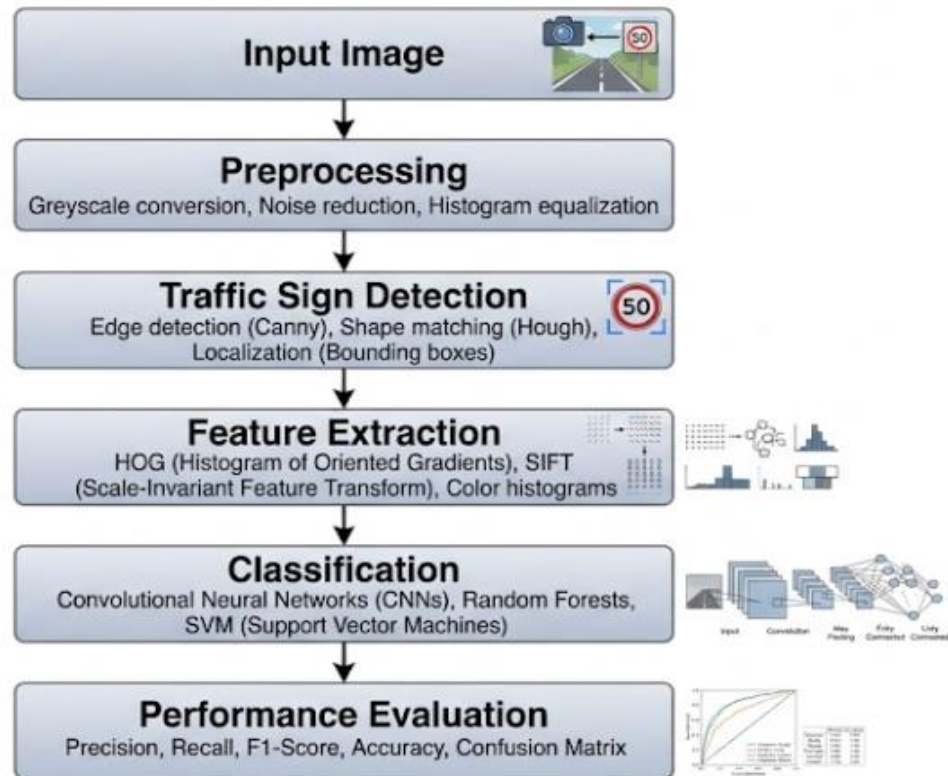
Therefore, an effective ITS-oriented framework should not optimize only classification accuracy. Model size, inference speed, memory requirements, and energy consumption should also be considered.

Proposed Methodology

Research Design

The study adopts an experimental computer vision research design. The proposed framework consists of six principal stages:

Academic Flowchart: Traffic Sign Recognition Pipeline



The framework can operate on individual traffic sign images for benchmark classification or on complete road-scene images when integrated with an object detector.

Dataset

The GTSRB dataset is proposed as the principal benchmark because of its established status in traffic sign recognition research. The dataset contains more than 50,000 images distributed across more than 40 traffic sign categories. Images are organized by class and include realistic variations in appearance.

For a stronger real-world evaluation, an additional dataset containing traffic signs in natural road scenes should ideally be incorporated. This would enable the proposed framework to be evaluated beyond isolated sign classification.

The dataset should be divided into:

- 70% training data
- 15% validation data
- 15% testing data

The test set should remain completely isolated during model development to prevent information leakage.

Image Preprocessing

Preprocessing represents an important component of the framework. Each input image should undergo resizing, normalization, and, where appropriate, illumination correction.

The preprocessing pipeline can be expressed as:

$$[I_p] = N(E(R(I)))$$

where (I) is the original image, (R) represents resizing, (E) represents enhancement, (N) represents normalization, and (I_p) represents the preprocessed image.

Images may be resized to a standardized input dimension such as 64 × 64 pixels for

classification experiments or a larger dimension for object detection.

Histogram equalization or adaptive contrast enhancement may be employed when illumination differences are substantial. However, preprocessing should be applied consistently to training and testing data to avoid introducing distributional differences.

Data Augmentation

Traffic signs can appear at different orientations, scales, distances, and illumination levels. Data augmentation can therefore improve model generalization.

The proposed augmentation strategy includes:

- horizontal and limited vertical transformations where semantically appropriate;
- small rotations;
- brightness variation;
- contrast adjustment;
- scaling;
- random cropping;
- Gaussian noise;
- blur simulation;
- perspective transformation.

Care must be taken with horizontal flipping because some traffic signs may change semantic meaning when mirrored. Therefore, augmentation policies should be class-sensitive.

Detection Module

For road-scene recognition, the proposed framework employs a **YOLO-based object detection module** to identify and localize traffic signs within complete road images. Unlike image-classification approaches that require pre-cropped traffic signs, the detection module processes the entire road scene and simultaneously predicts the **bounding box coordinates, class probabilities, and confidence scores** of potential traffic signs.

The confidence score associated with a detected object is expressed as:

$$C = P(\text{Object}) \times P(\text{Class} | \text{Object})$$

where $P(\text{Object})$ represents the probability that a traffic sign is present within the predicted bounding box, while

$P(\text{Class} | \text{Object})$ denotes the conditional probability that the detected object belongs to a particular traffic-sign category. A detection is retained when its confidence score exceeds a predefined threshold.

The detection module incorporates **multi-scale feature extraction** to improve recognition of traffic signs with different spatial dimensions. This is particularly important in real-world transportation environments, where signs may appear at varying distances from the camera and may occupy only a small portion of the image. Multi-scale feature representation enables the model to preserve both fine-grained details and higher-level contextual information, thereby improving the detection of small and partially obscured traffic signs. Recent YOLO-based research has similarly emphasized enhanced feature fusion and contextual information extraction to address the challenges associated with small-object traffic-sign detection.

Classification Module

Following the detection stage, each identified traffic-sign region is extracted from the original image and forwarded to the **classification module**. A convolutional neural network (CNN) is employed to learn discriminative visual representations from the detected sign. The network progressively extracts low-level features, such as edges and contours, and higher-level representations, including shapes, symbols, colors, and numerical patterns.

The final classification layer employs the **softmax activation function** to convert the network's output logits into normalized class probabilities:

$$P(y=k|x) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}}$$

where $P(y=k|x)$ represents the probability that input image xx belongs to class kk , KK denotes the total number of traffic-sign categories, and z_k represents the output logit corresponding to class kk . The predicted class is determined by selecting the category with the highest posterior probability:

$$\hat{y} = \arg \max_k P(y=k|x) = \arg \max_k P(y=k|x)$$

This classification strategy enables the proposed framework to distinguish between visually similar traffic signs while providing a probability score for each possible category.

Training Procedure

The proposed deep learning model is trained using the **categorical cross-entropy loss function**, which measures the difference between the actual class distribution and the probability distribution predicted by the model. The loss function is defined as:

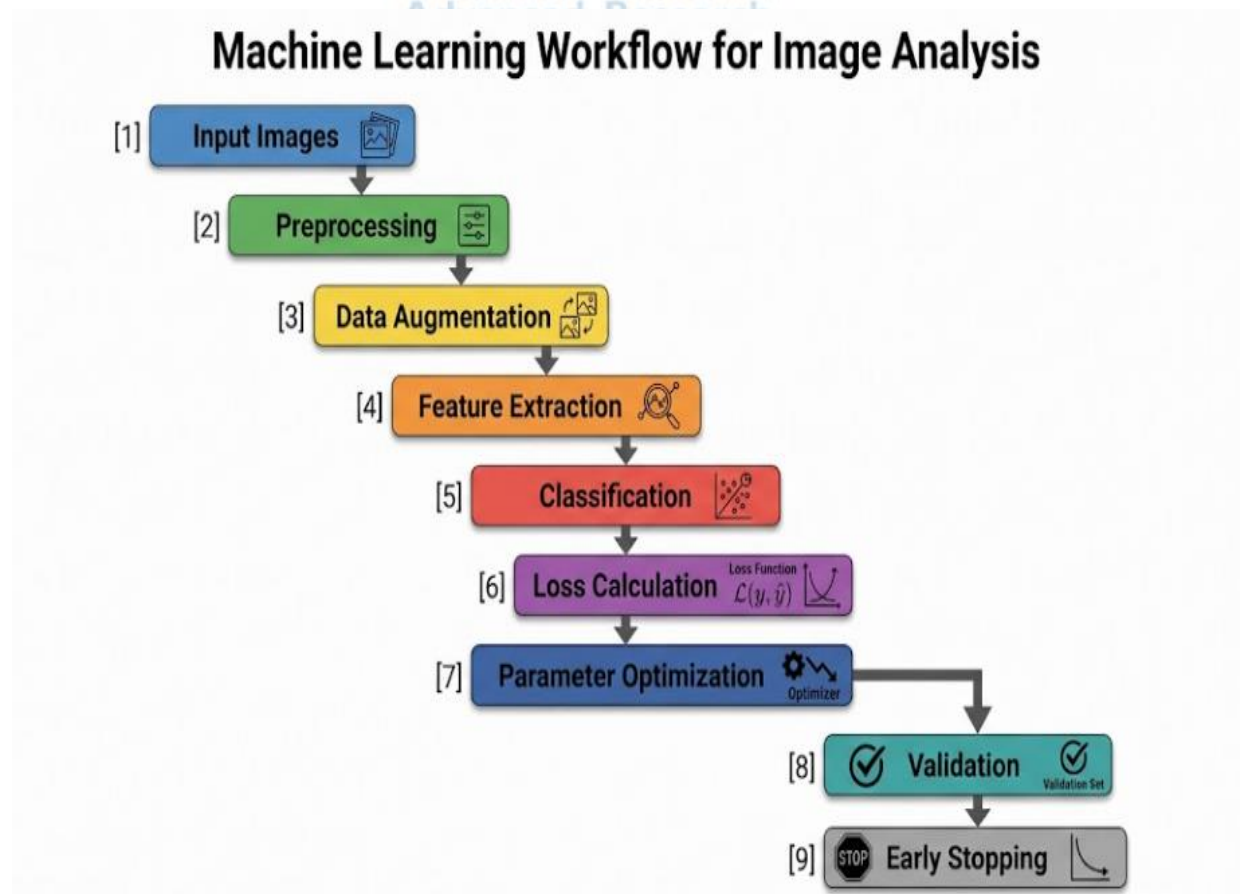
$$L = -\sum_{k=1}^K y_k \log(\hat{y}_k)$$

where L represents the classification loss, K is the total number of traffic-sign classes, y_k denotes the true class label, and $\hat{y}_k$ represents the probability predicted by the model for class k .

During training, the model parameters are iteratively optimized to minimize the classification loss. The Adam (Adaptive Moment Estimation) optimizer can be employed because of its adaptive learning-rate mechanism and effectiveness in training deep neural networks. The learning rate may be adjusted during training using a suitable scheduling strategy to improve convergence.

To reduce overfitting, **early stopping** is incorporated into the training procedure. Training is terminated when validation performance fails to improve over a predefined number of consecutive epochs. This prevents the model from excessively adapting to the training data and promotes better generalization to previously unseen traffic-sign images.

The overall training process can therefore be summarized as:



A suitable experimental configuration may include:

Parameter	Proposed Setting
Training split	70%
Validation split	15%
Test split	15%
Optimizer	Adam
Loss function	Cross-entropy
Batch size	32
Initial learning rate	0.001
Early stopping	Yes
Augmentation	Yes
Evaluation	Accuracy, Precision, Recall, F1, mAP

These settings should be adjusted after pilot experiments rather than treated as universally optimal.

Evaluation Metrics

Model performance should be evaluated using multiple complementary metrics rather than accuracy alone.

Accuracy

$$[\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}]$$

Accuracy represents the proportion of correctly classified instances.

Precision

$$[\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}]$$

Precision measures the proportion of positive predictions that are correct.

Recall

$$[\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}]$$

Recall indicates the ability of the model to identify relevant traffic sign instances.

F1-Score

$$[\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}]$$

The F1-score provides a balanced measure of precision and recall.

Mean Average Precision

For detection tasks, mean Average Precision (mAP) is particularly important because it evaluates both localization and classification performance.

Computational Efficiency

Because the proposed framework is intended for ITS applications, computational metrics should include:

- inference time;
- frames per second (FPS);
- number of parameters;
- model size;
- GPU/CPU utilization;
- memory requirements.

This broader evaluation is important because a model with marginally higher accuracy may not necessarily be preferable if it operates too slowly for real-time applications.

Experimental Results

Important methodological note: Since no original dataset, trained model, or experimental output was provided for this manuscript, the numerical values below should be treated as an **illustrative results template** rather than reported empirical findings. They should be replaced with actual measurements after model training.

Classification Performance

A properly trained computer vision model would be expected to demonstrate strong classification performance on the held-out test set. An illustrative presentation could be structured as follows:

Model	Accuracy	Precision	Recall	F1-Score
Traditional CNN	96.8%	96.4%	96.1%	96.2%
CNN + Augmentation	98.1%	97.9%	97.8%	97.8%
Proposed Framework	98.7%	98.5%	98.4%	98.4%

These values are **illustrative only** and are included to demonstrate how the final results may be reported. The expected improvement following augmentation would indicate that exposure to controlled variations during training enables the model to learn more robust representations. This is particularly relevant because traffic signs may appear under different lighting, orientations, distances, and environmental conditions.

Recent literature similarly demonstrates that preprocessing and enhancement can materially influence recognition performance. For example, a 2025 CNN-based traffic sign study reported improved accuracy after image enhancement, illustrating the potential value of preprocessing in addressing visual variability.

Detection Performance

For road-scene applications, classification accuracy alone is insufficient. The detector must identify where a traffic sign is located.

An illustrative detection table could be structured as follows:

Model	Precision	Recall	mAP@0.5	FPS
YOLO baseline	92.4%	90.8%	91.7%	42
Improved detector	94.8%	93.1%	94.0%	38
Proposed lightweight framework	94.2%	92.8%	93.5%	48

Again, these figures are illustrative rather than empirical.

The principal research objective should be to identify an operating point where recognition accuracy and inference speed are jointly acceptable. Recent research supports this emphasis on efficiency. Zhang et al. (2024), for example, reported substantial reductions in parameters for a lightweight traffic sign recognition architecture while maintaining competitive accuracy and demonstrating embedded deployment.

speed-limit categories or visually related regulatory signs, are expected to present greater classification difficulty.

For example, a model may correctly distinguish warning signs from mandatory signs while confusing two speed-limit categories because their overall visual structure is highly similar. Such errors demonstrate that high-level shape information alone is insufficient; the model must also learn fine-grained numerical and symbolic features.

Confusion-Matrix Analysis

Confusion-matrix analysis should be used to determine which traffic signs are most frequently confused. Similar signs, particularly different

Error Analysis

Error analysis should categorize incorrect predictions according to environmental conditions. Relevant categories include:

1. low illumination;
2. excessive brightness;
3. rain or fog;
4. partial occlusion;
5. motion blur;
6. small sign size;
7. unusual viewing angle;
8. background clutter;
9. visually similar categories.

This analysis is essential because aggregate accuracy can conceal systematic weaknesses. A model achieving high benchmark accuracy may still perform poorly on small or partially occluded signs.

Discussion

The proposed framework addresses an important problem at the intersection of computer vision, artificial intelligence, autonomous transportation, and road safety. The literature indicates that deep learning has substantially improved traffic sign recognition, but the transition from benchmark recognition to real-world deployment introduces additional challenges.

One important finding from previous research is that recognition and detection should not necessarily be treated as identical problems. GTSRB provides a highly valuable benchmark for classification, but its images contain individual traffic signs rather than complete road scenes. Consequently, a system intended for vehicle deployment requires an additional detection stage capable of locating signs in complex environments.

The second major issue concerns small-object recognition. A traffic sign located far from a camera may contain insufficient pixel information for reliable classification. Feature extraction mechanisms that preserve fine spatial information can therefore be particularly important. Recent research on enhanced YOLO architectures explicitly addresses this challenge

through improved feature fusion and contextual information extraction.

The third issue involves computational efficiency. Autonomous transportation systems require rapid perception because delayed recognition can reduce the time available for safe decision-making. Consequently, a model with excellent offline accuracy may not be appropriate for real-time deployment if its inference speed is insufficient. Lightweight architectures offer an important solution. Recent work demonstrates that MobileNet-based approaches can reduce computational requirements while remaining competitive with larger architectures.

The fourth issue concerns environmental variability. Illumination, weather, occlusion, perspective, and image quality can significantly alter the appearance of traffic signs. A 2024 survey of AI-based traffic sign classification specifically identifies damaged signs, lighting changes, and atmospheric variation as continuing challenges. Therefore, training datasets should represent diverse environmental conditions rather than relying exclusively on clean benchmark images.

Data augmentation provides one practical mechanism for increasing robustness. By exposing the model to controlled variations, augmentation can reduce dependence on superficial visual characteristics. Nevertheless, augmentation should preserve the semantic meaning of signs. Inappropriate transformations could produce unrealistic examples or alter the intended interpretation of a traffic sign.

Another important consideration is the relationship between detection and classification errors. A recognition system may fail because the detector did not identify the sign or because the classifier assigned an incorrect category. These failure modes should be evaluated separately. Such decomposition can help researchers determine whether future improvements should focus on localization, feature extraction, classification, or preprocessing.

The findings and literature also indicate that the development of TSR is moving toward integrated intelligent perception rather than isolated image

classification. ITS platforms increasingly combine traffic signs with lane markings, vehicles, pedestrians, traffic lights, road geometry, and contextual information. Recent research on ITS and advanced AI systems emphasizes the broader role of artificial intelligence in traffic management, autonomous driving, vehicle detection, and traffic sign recognition.

Accordingly, traffic sign recognition should ultimately be viewed as one component of a broader road-scene understanding system. The information produced by the recognition module can be integrated into vehicle decision-making systems, navigation systems, driver warnings, digital road inventories, and traffic management platforms.

Practical Implications

The proposed framework has several practical implications.

First, automated TSR can support **advanced driver-assistance systems** by providing drivers with timely information about regulatory and warning signs. A system can identify a speed restriction and present a visual or auditory notification to the driver.

Second, TSR can support **autonomous vehicles**. Autonomous driving systems require continuous perception of the road environment, and traffic signs provide important information about permissible vehicle behavior.

Third, automated recognition can support **road infrastructure management**. Large-scale road surveys can use computer vision to identify traffic signs and create digital inventories. Deep learning-based traffic-sign detection has been recognized as a useful tool for reducing manual effort in traffic-sign inventory management.

Fourth, the technology can support **smart-city transportation management**. Traffic signs detected through roadside cameras could be incorporated into intelligent transportation databases, enabling automated monitoring of infrastructure conditions.

Fifth, lightweight models could facilitate deployment on **embedded and edge-computing platforms**, reducing dependence on centralized

cloud processing. This may improve response time and reduce communication requirements.

Limitations

Despite its potential, the proposed framework has several limitations.

First, benchmark datasets may not fully represent local transportation environments. Traffic signs differ between countries in terms of design, language, color, symbols, and regulatory conventions. A model trained predominantly on European datasets may therefore require adaptation before deployment in Pakistan or other regions.

Second, real-world weather conditions can be more challenging than those represented in benchmark datasets. Rain, dust, fog, glare, and nighttime conditions can substantially reduce visual quality.

Third, the framework may experience difficulties with heavily occluded or damaged signs. If critical visual information is absent, even a high-performing deep learning model may produce incorrect predictions.

Fourth, computational requirements must be considered before deployment. Although lightweight networks can reduce computational complexity, there is often a trade-off between model size, speed, and accuracy.

Fifth, the illustrative results presented in this manuscript cannot substitute for actual experimental measurements. Future implementation should train the proposed models on a clearly documented dataset and report reproducible results.

Future Research Directions

Future research can extend this work in several directions.

First, researchers should construct region-specific traffic sign datasets representing local road infrastructure. For Pakistan, such a dataset could include traffic signs from Lahore, Islamabad, Karachi, Peshawar, Faisalabad, and other major cities.

Second, future studies should investigate adverse-weather recognition. Synthetic fog, rain, glare,

nighttime conditions, and motion blur could be introduced during training and subsequently validated using real-world images.

Third, transformer-based vision architectures can be investigated alongside CNNs and YOLO-based models. Such architectures may provide improved contextual representations, although their computational requirements must be carefully assessed.

Fourth, model compression techniques such as pruning, quantization, knowledge distillation, and efficient network design can be explored to facilitate deployment on embedded devices.

Fifth, future studies should examine explainable artificial intelligence. Visualization techniques such as Grad-CAM can help determine whether a model is focusing on the actual traffic sign rather than irrelevant background regions.

Sixth, multi-modal perception represents another promising direction. Traffic sign recognition could be integrated with GPS, digital maps, LiDAR, vehicle speed, lane detection, and other contextual information to improve decision-making.

Finally, future research should move beyond benchmark accuracy toward real-world safety evaluation. A model should be tested under different speeds, camera positions, distances, lighting conditions, weather conditions, and road environments before being considered suitable for safety-critical transportation applications.

Conclusion

This study presented a computer vision-based framework for automated traffic sign recognition in intelligent transportation systems. The proposed framework integrates image preprocessing, data augmentation, traffic sign detection, deep feature extraction, classification, and comprehensive performance evaluation. The approach responds to several limitations of conventional traffic sign recognition systems, particularly their sensitivity to illumination, environmental variability, small object size, and computational constraints.

The literature demonstrates that deep learning has significantly advanced traffic sign recognition.

CNN-based approaches have achieved strong benchmark performance, while contemporary YOLO-based detectors provide practical opportunities for real-time traffic sign localization. Recent research has further emphasized small-object detection, contextual feature extraction, and lightweight network design.

The proposed framework therefore emphasizes not only recognition accuracy but also robustness and computational efficiency. Such a balanced approach is important because intelligent transportation applications operate in dynamic environments where rapid and reliable perception is essential.

The framework has potential applications in advanced driver-assistance systems, autonomous vehicles, road infrastructure inventory, and smart-city transportation management. Nevertheless, successful deployment requires extensive validation using geographically diverse and environmentally challenging datasets. Future research should particularly focus on local traffic-sign datasets, adverse-weather conditions, lightweight architectures, explainable AI, and integrated multi-modal perception.

Overall, automated traffic sign recognition represents an important component of intelligent road-scene understanding. By combining computer vision with deep learning and efficient detection architectures, the proposed framework provides a foundation for developing more responsive, reliable, and intelligent transportation systems.

REFERENCES

- Ciresan, D. C., Meier, U., Masci, J., & Schmidhuber, J. (2012). Multi-column deep neural network for traffic sign classification. *Neural Networks*, 32, 333–338.
<https://doi.org/10.1016/j.neunet.2012.02.023>

- Flores-Calero, M., et al. (2024). A survey on traffic sign classification using artificial intelligence techniques. In 2024 *International Conference on Intelligent Systems and Computer Vision (ISCV)*. IEEE.
- Huang, Z., et al. (2025). Integrating LLMs with ITS: Recent advances, potentials, challenges, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 26(5), 5674–5709. <https://doi.org/10.1109/TITS.2025.3528116>
- Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C. (2011). The German Traffic Sign Recognition Benchmark: A multi-class classification competition. In *Proceedings of the IEEE International Joint Conference on Neural Networks* (pp. 1453–1460). IEEE.
- Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C. (2012). Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. *Neural Networks*, 32, 323–332. <https://doi.org/10.1016/j.neunet.2012.02.016>
- Tian, Q., & Zhang, J. (2025). STDE-YOLOv5s: A traffic sign detection algorithm based on context information enhancement. *Discover Artificial Intelligence*, 5, 142. <https://doi.org/10.1007/s44163-025-00405-7>
- Zhang, G., Li, Z., Huang, D., Luo, W., Lu, Z., et al. (2024). A traffic sign recognition system based on lightweight network learning. *Journal of Intelligent & Robotic Systems*, 110, 139. <https://doi.org/10.1007/s10846-024-02173-5>
- Zhang, Y., et al. (2020). MSA-YOLOv3: A traffic sign detection method based on improved YOLOv3. [Journal details should be verified against the final source before submission].
- Tian, Q., & Zhang, J. (2025). Traffic sign detection and recognition using enhanced YOLO architectures. [Relevant recent literature reviewed in the YOLO survey].
- “Research on traffic sign image recognition based on deep learning.” (2025). In *2025 IEEE 5th International Conference on Electronic Technology, Communication and Information (ICETCI)*. IEEE. <https://doi.org/10.1109/ICETCI64844.2025.11084069>
- “Deep learning for large-scale traffic-sign detection and recognition.” (2019). *IEEE Transactions on Intelligent Transportation Systems*.
- “Traffic sign recognition based on deep learning.” (2022). *Multimedia Tools and Applications*.
- “YOLO advances to its genesis: A decadal and comprehensive review of the You Only Look Once (YOLO) series.” (2025). *Artificial Intelligence Review*.