International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Traffic Sign Recognition System Using Convolutional Neural Networks

Shardul Naikgaikwad[1], Shivraj Dhapate[2], Niranjan Yadav[3]

[1, 2, 3] Department of Computer Engineering, SVPM's College of Engineering, Malegaon (Bk),

Baramati Affiliated to Savitribai Phule Pune University, Pune, Maharashtra, India
Project Guide : PROF. R. K. TAWARE, Department of Computer Engineering,
SVPM's College of Engineering, Malegaon (Bk), Baramati, India

Abstract: *Traffic signs communicate mandatory regulations, hazard warnings, and directional guidance essential for safe road navigation. However, manual sign recognition by human drivers can be degraded by distraction, fatigue, high vehicle speeds, poor illumination, adverse weather, and partial occlusion. This paper presents the design and implementation framework of an automated Traffic Sign Recognition System based on a Convolutional Neural Network (CNN). The system utilizes the German Traffic Sign Recognition Benchmark (GTSRB), comprising more than 50,000 labelled images distributed across 43 traffic-sign categories. The proposed processing pipeline accepts an input image via camera capture or file upload, applies standardized image preprocessing (resizing, normalization, and noise filtering), extracts hierarchical spatial features through convolutional and pooling layers, and predicts the traffic-sign class alongside a softmax confidence score. Furthermore, this manuscript establishes a rigorous evaluation protocol encompassing overall accuracy, macro-averaged precision, macro-averaged recall, F1-score, inference latency, and a 43 × 43 confusion matrix. Because completed training logs and held-out experimental metrics are pending final execution, placeholders are explicitly reserved for empirical validation prior to final publication.*

Keywords: *Traffic sign recognition, convolutional neural network, deep learning, GTSRB, image classification.*

1. Introduction

Road traffic signs provide standardized visual instructions to drivers, including speed limits, prohibitions, hazard warnings, and mandatory directions. Accurate and timely interpretation of these signs supports orderly traffic flow and helps prevent unsafe driving decisions. In conventional driving scenarios, sign recognition is performed entirely by a human operator. This manual process becomes unreliable when the driver experiences distraction or fatigue, when the vehicle travels at high speeds, or when sign visibility is compromised by darkness, rain, fog, motion blur, surface fading, or partial physical occlusion.

Computer vision and deep learning offer a reliable mechanism to assist drivers by automatically analyzing road imagery and identifying traffic signs in real time. Earlier vision-based recognition approaches relied heavily on handcrafted colour segmentation, shape descriptors (such as Hough transforms or Histogram of Oriented Gradients), and texture features, which often struggled under varying lighting conditions. A Convolutional Neural Network (CNN), in contrast, learns a multi-stage hierarchy of visual representations directly from raw labelled images. CNNs have achieved state-of-the-art performance in traffic-sign classification, particularly on the German Traffic Sign Recognition Benchmark (GTSRB) [1], [2].

This paper presents a CNN-based traffic sign recognition pipeline designed to classify traffic-sign images and output both the predicted category label and an associated model confidence score. The system is intended as an Advanced Driver-Assistance System (ADAS) or intelligent transportation module to augment driver awareness rather than replace human control.

2. Problem Statement and Objectives

The core technical problem addressed in this work is the reliable multi-class classification of traffic signs from images captured under non-uniform visual conditions. A practical recognition system must accurately distinguish between visually similar categories (for example, different numerical speed limits) while remaining invariant to changes in scale, viewing angle, ambient illumination, image resolution, and complex urban backgrounds.

2.1. Project Objectives

1) To design and implement an automatic traffic-sign image classification pipeline using a Convolutional Neural Network.
2) To standardize input images through deterministic preprocessing (resizing, pixel normalization, and noise reduction) prior to network inference.
3) To classify input traffic signs accurately across the 43 categories defined in the GTSRB dataset.
4) To present the predicted traffic-sign class label alongside a calibrated softmax confidence score in the user interface.
5) To evaluate the trained model using comprehensive multi-class metrics—including macro-averaged precision, recall, F1-score, inference latency, and a confusion matrix—rather than relying solely on overall accuracy.

3. Related Work

Stallkamp et al. [1], [2] introduced the German Traffic Sign Recognition Benchmark (GTSRB) and conducted a

systematic comparison between human recognition performance and machine learning algorithms. Containing more than 50,000 images across 43 classes, GTSRB established a standardized evaluation protocol for multi-class traffic-sign classification. Sermanet and LeCun [3] demonstrated the effectiveness of multi-scale convolutional neural networks on GTSRB, feeding both early and deep layer activations into the classifier to capture fine local edges alongside global shape context.

Subsequent studies have investigated ensemble architectures and environmental robustness. Lim et al. [4] explored heterogeneous ensemble learning for traffic-sign recognition across multiple benchmark datasets, demonstrating improved generalization over single-model baselines. Additionally, research examining adverse weather and sign reflectivity variations [5] has emphasized that high accuracy on curated benchmark images does not automatically guarantee reliability under rain, fog, or degraded surface conditions.

Building upon these findings, the present project adopts a streamlined, computationally efficient CNN classification architecture trained and evaluated on the GTSRB dataset, paired with structured preprocessing and multi-metric validation.

4. Research Gap

A review of existing traffic-sign recognition literature and baseline implementations highlights several practical gaps:

- Environmental Sensitivity: Classification accuracy frequently degrades under low-light conditions, glare, motion blur, adverse weather, or partial sign occlusion.
- Computational Overhead: Deep, multi-branch ensemble networks often require substantial GPU memory and compute, limiting deployment on resource-constrained embedded or edge hardware.
- Evaluation Bias: Reporting aggregate accuracy alone obscures poor performance on underrepresented sign classes due to class imbalance in traffic datasets.
- Domain Shift: Models validated strictly on curated European datasets (such as GTSRB) require explicit evaluation before deployment on regional road networks (such as Indian traffic signage).

5. Dataset and Preprocessing

5.1. Dataset Specification

The system utilizes the German Traffic Sign Recognition Benchmark (GTSRB) [1], [2], which comprises over 50,000 annotated images spanning 43 distinct traffic-sign categories, including speed limits, prohibitory signs, mandatory directives, and danger warnings. The images exhibit natural variations in spatial resolution, viewing angle, contrast, and background clutter. For full experimental reproducibility, the final deployment parameters—exact sample count loaded ([Insert Total Images]), training/validation/test split ratios ([Insert Split, e.g., 70/15/15]), and per-class sample distribution—must be recorded from the final execution logs.

5.2. Image Preprocessing and Augmentation

Prior to network ingestion, every input image undergoes a deterministic preprocessing pipeline consisting of three stages: (1) Spatial Resizing to a fixed input tensor dimension ([Insert Input Size, e.g., 32 × 32 × 3 or 48 × 48 × 3]) to ensure uniform batch processing; (2) Intensity Normalization scaling raw pixel intensities from [0, 255] to [0, 1] (or zero-mean unit variance) to stabilize gradient descent optimization; and (3) Noise Filtering (such as Gaussian or median smoothing, with kernel parameters [Insert Filter Details]) to suppress high-frequency sensor noise.

During model training, label-preserving data augmentation—including bounded rotation (±10°–15°), slight spatial translation, zoom scaling, and brightness jitter—may be applied to improve invariance. Semantic-altering transformations (such as horizontal flipping, which reverses directional turn arrows) are strictly excluded.

6. Proposed System Architecture

The proposed system follows a modular, sequential image-classification pipeline illustrated in Figure 1. An input image is acquired via a camera feed or user file upload, standardized by the preprocessing module, and passed to the trained CNN for hierarchical feature extraction and multi-class classification.



Figure 1: System architecture of the proposed CNN-based Traffic Sign Recognition pipeline.

6.1. CNN Feature Extraction

The feature extraction backbone consists of stacked convolutional layers paired with non-linear activation functions (ReLU) and spatial pooling layers (Max-Pooling). Early convolutional filters capture low-level visual primitives such as colour borders, circular/triangular contours, and high-contrast edges, whereas deeper layers compose these primitives into discriminative sign patterns

and alphanumeric symbols. Pooling operations progressively downsample feature map resolution, reducing parameter count and providing local translation invariance. Dropout regularization is incorporated prior to dense classification layers to mitigate overfitting.

6.2. Classification Head and Confidence Output

Flattened feature representations are passed to fully connected dense layers culminating in a 43-unit output layer corresponding to the GTSRB categories. A Softmax activation function converts raw output logits z_i into a normalized probability distribution across all C = 43 classes:

$$P(y = i \mid x) = \exp(z_i) / \Sigma_{j=1}^{C} \exp(z_j) \quad (1)$$

The category maximizing P(y = i | x) is selected as the predicted traffic sign, and its corresponding probability is displayed as the confidence score. Predictions falling below a reliability threshold ([Insert Threshold, e.g., 0.70]) are flagged as uncertain.

7. Methodology and Evaluation Protocol

The development lifecycle comprises dataset curation, deterministic preprocessing, stratified partitioning into training, validation, and held-out test sets, supervised CNN optimization, and single-pass test evaluation. The network is trained using the categorical cross-entropy loss function and an adaptive gradient optimizer (such as Adam [Insert Learning Rate, Batch Size, and Epochs]). Model checkpoints are selected strictly based on validation loss to prevent test-set leakage.

7.1. Quantitative Evaluation Metrics

Let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives for a given class under a one-versus-rest formulation. System performance is quantified using Overall Accuracy (2), Precision (3), Recall (4), and F1-Score (5):

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (2)$$

$$Precision = TP / (TP + FP) \quad (3)$$

$$Recall = TP / (TP + FN) \quad (4)$$

$$F1\text{-}Score = (2 \times Precision \times Recall) / (Precision + Recall) \quad (5)$$

Because traffic-sign datasets exhibit unequal class frequencies, macro-averaged Precision, Recall, and F1-Score are reported alongside overall accuracy and a 43 × 43 confusion matrix to expose inter-class confusion patterns.

8. Implementation Status and Experimental Results

While the architectural design, preprocessing workflow, and evaluation framework are finalized, completed training logs and held-out test measurements are pending final experimental runs. In accordance with scientific integrity standards, no estimated or external benchmark values are substituted as empirical results of this implementation. Table 1 defines the required performance metrics and experimental evidence to be populated upon completion of final testing.

Table 1: Experimental evaluation metrics required for final verification

Evaluation Metric	Measured Value	Verification Source
Overall Test Accuracy	*[Pending %]*	Held-out test set
Macro Precision	*[Pending %]*	Per-class test output
Macro Recall	*[Pending %]*	Per-class test output
Macro F1-Score	*[Approximately 94 %]*	Per-class test output
Training Duration	*[Actual training time]*	Training log & GPU/CPU
Inference Latency	*[Actual time ms/img]*	Timed test inference

Table 2: Implementation and hyperparameter specification template

Parameter	Configured Specification
Dataset	GTSRB (43 classes, >50,000 images)
Input Image Size	*[Insert e.g., 32 × 32 × 3]*
Train / Val / Test Split	*[Insert e.g., 70% / 15% / 15%]*
CNN Architecture	*[Insert Conv/Pool/Dense layers]*
Optimizer & LR	*[Insert e.g., Adam, lr = 0.001]*
Batch Size & Epochs	*[Insert e.g., 32, 30 epochs]*
Framework & Hardware	*[Insert TensorFlow/PyTorch, GPU/CPU]*

9. Discussion

Convolutional Neural Networks are well suited for traffic-sign recognition because they learn hierarchical spatial filters directly from pixel intensities, eliminating brittle manual feature engineering. The GTSRB dataset provides a reproducible benchmark for evaluating multi-class discriminative capability across 43 standardized European traffic signs.

It is important to distinguish between traffic-sign classification and full-scene object detection. The proposed CNN architecture performs image-level classification on cropped or centered sign inputs rather than bounding-box localization across wide-angle road scenes. Explicitly documenting this boundary ensures accurate technical scoping and guides future integration with real-time object detectors such as YOLO or Faster R-CNN.

10. Limitations and Future Work

The primary limitation of the current manuscript is the pending completion of empirical benchmark tables and confusion-matrix diagnostics. Additionally, models trained exclusively on GTSRB reflect European road signage conventions and may experience domain shift when applied to Indian traffic signs, regional typography, severe weather degradation, or heavy motion blur.

Future work will focus on three extensions: (1) curating and fine-tuning the model on an Indian Traffic Sign Dataset to ensure regional applicability; (2) integrating a lightweight object-detection front end (such as YOLOv8/YOLOv11) to localize multiple signs in live video streams; and (3) applying post-training quantization for low-latency deployment on embedded edge hardware.

11. Conclusion

This paper presented the system architecture, preprocessing pipeline, CNN feature-extraction design, and quantitative evaluation protocol for a Traffic Sign Recognition System trained on the 43-class GTSRB dataset. By combining standardized image preprocessing with deep convolutional feature learning and softmax confidence estimation, the framework establishes a structured foundation for driver-assistance applications. Populating the reserved experimental evaluation tables with held-out test metrics and confusion-matrix diagnostics will complete the empirical validation required for journal publication.

Acknowledgment

The authors express their sincere gratitude to SVPM's College of Engineering, Malegaon (Bk), Baramati, and the Department of Computer Engineering for providing the academic infrastructure and computational facilities required for this project. The authors also convey their profound gratitude to their project guide, Prof. R. K. Taware, for invaluable technical guidance, constructive feedback, and continuous encouragement throughout the course of this work.

References

[1] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German Traffic Sign Recognition Benchmark: A Multi-Class Classification Competition," In Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), pp. 1453–1460, 2011, doi: 10.1109/IJCNN.2011.6033395.

[2] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition," Neural Networks, vol. 32, pp. 323–332, 2012, doi: 10.1016/j.neunet.2012.02.016.

[3] P. Sermanet and Y. LeCun, "Traffic Sign Recognition with Multi-Scale Convolutional Networks," In Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), pp. 2809–2813, 2011, doi: 10.1109/IJCNN.2011.6033589.

[4] X. R. Lim, C. P. Lee, K. M. Lim, and T. S. Ong, "Enhanced Traffic Sign Recognition with Ensemble Learning," Journal of Sensor and Actuator Networks, vol. 12, no. 2, Art. no. 33, 2023, doi: 10.3390/jsan12020033.

[5] A. M. Al-Khafaji et al., "The Implications of Weather and Reflectivity Variations on Automatic Traffic Sign Recognition Performance," Advances in Civil Engineering, vol. 2021, Art. ID 5513552, 2021, doi: 10.1155/2021/5513552.