International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-Jun-2026, ISSN (Online): 2395-566X

Traffic Sign Recognition

Prof. P.S. Togrikar, A.N.Jamdade, P.R.Shirke, H.J.Phadtare
E & TC Engineering Department S. B. Patil College of Engineering

Abstract- Traffic Sign Recognition (TSR) is an essential component of Advanced Driver Assistance Systems (ADAS) and intelligent transportation. This paper presents a cost-effective IoT-based TSR system using an ESP32-CAM for image acquisition and a backend server for processing. Due to limited edge-device capability, images are transmitted via a Telegram Bot for remote inference using the YOLOv3 deep learning model trained on the GTSRB dataset. To enhance robustness under real-world conditions such as occlusion and varying illumination, preprocessing techniques like CLAHE and data augmentation are applied. The system returns annotated results through a Telegram interface and a local GUI. Experimental results demonstrate high accuracy and reliable performance, validating the effectiveness of the proposed approach. The system also shows strong performance under partially occluded conditions, improving real-world applicability. Furthermore, the proposed architecture ensures low-cost deployment and scalability for smart transportation systems. This work highlights the potential of integrating IoT with deep learning for practical and accessible traffic monitoring solutions.

Keywords- Traffic Sign Recognition, YOLOv3, ESP32-CAM, IoT, GTSRB, Deep Learning, Object Detection, ADAS.

I. INTRODUCTION

With the rapid development of intelligent transportation systems, improving road safety and enabling automation have become major areas of focus. Traffic Sign Recognition (TSR) is an important technology used in Advanced Driver Assistance Systems (ADAS) and autonomous vehicles. It helps in detecting and understanding traffic signs such as speed limits, stop signs, and warning signs, which play a critical role in guiding drivers and controlling vehicle behavior. Accurate recognition of these signs improves decision-making, reduces accidents, and enhances overall traffic management.

Traditional TSR systems were mainly based on image processing techniques such as color detection, shape analysis, and feature extraction methods like Histogram of Oriented Gradients (HOG). These methods required manual feature design and often failed in real-world conditions such as poor lighting, weather variations, motion blur, and partial occlusion. Due to these limitations, their performance was not reliable for practical applications.

In recent years, deep learning techniques have significantly improved object detection and classification tasks. Convolutional Neural Networks (CNNs) and models like YOLO (You Only Look Once) provide high accuracy and fast detection, making them suitable for real-time TSR applications. Among these, YOLOv3 offers a good balance between speed and performance, especially for detecting small objects like traffic signs.

However, deploying such computationally intensive models on low-cost embedded devices is a major challenge due to limited processing power and memory. To address this issue, this paper proposes a hybrid IoT-based system. An ESP32-CAM module is used as an edge device to capture images, while a backend server performs image processing and traffic sign detection using the YOLOv3 model. The captured images are transmitted to the server through a network connection, and the processed results are returned to the user.

To improve system performance under real-world conditions, image preprocessing techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) are applied to enhance image quality. In addition, data augmentation methods, including synthetic occlusion, are used during model training to improve robustness against partially hidden traffic signs. These techniques help the system maintain high accuracy even in challenging environments.

The proposed system is designed to be low-cost, scalable, and efficient, making it suitable for applications such as smart transportation, traffic monitoring, and low-cost robotic systems. Although the system introduces some delay due to

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-Jun-2026, ISSN (Online): 2395-566X

network communication, it provides a practical solution where on-device processing is not feasible.

II. LITERATURE REVIEW

Traffic Sign Recognition (TSR) has been widely studied in the field of computer vision and intelligent transportation systems. Over the years, different approaches have been developed, ranging from traditional image processing methods to advanced deep learning techniques.

Early TSR systems were based on classical image processing and machine learning methods. These approaches typically involved multiple steps such as color segmentation, shape detection, feature extraction, and classification. Techniques like Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and Support Vector Machines (SVM) were commonly used. Although these methods were computationally efficient, they were highly dependent on lighting conditions, color consistency, and clear visibility of traffic signs. As a result, their performance degraded significantly in real-world environments with noise, occlusion, and varying illumination.

With the advancement of deep learning, Convolutional Neural Networks (CNNs) have become the preferred approach for TSR. CNN-based models automatically learn important features from images, making them more robust to variations in lighting, scale, and orientation. Initially, classification-based approaches were used where the system first detected a region of interest and then classified the traffic sign. However, these methods required separate stages for detection and classification, increasing complexity and processing time.

To overcome these limitations, object detection models such as R-CNN, Fast R-CNN, and Faster R-CNN were introduced. These models improved detection accuracy by generating region proposals and then classifying them. Although they achieved high accuracy, their computational cost made them less suitable for real-time applications.

More recently, one-stage object detection models like YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector) have gained popularity due to their speed and efficiency. These models perform detection and classification in a single step, making them suitable for real-time systems. Among them, YOLOv3 is widely used because of its ability to detect objects at multiple scales, which is especially useful for identifying small traffic signs in complex environments.

Several studies have applied YOLO-based models for traffic sign detection and achieved high accuracy. However, many of these approaches require high-performance hardware and are not suitable for low-cost embedded systems. In addition, challenges such as partial occlusion, motion blur, and poor lighting conditions are still areas of concern.

In this work, a hybrid IoT-based approach is used to address these challenges. A low-cost edge device is used for image capture, while computationally intensive tasks are performed on a backend server. Furthermore, techniques such as image enhancement and data augmentation are used to improve robustness under real-world conditions. This approach provides a balance between performance, cost, and practicality..

III. SYSTEM ARCHITECTURE

The proposed system follows a hybrid IoT-based architecture that separates image acquisition and processing tasks. This design allows the use of a low-cost embedded device for capturing images while performing computationally intensive operations on a more powerful backend system. The overall architecture ensures a balance between cost, performance, and efficiency.

The system mainly consists of three components: the edge device, the communication network, and the backend server.
The edge device is based on the ESP32-CAM module, which is responsible for capturing images of the surrounding environment. It is a low-cost microcontroller with an integrated camera and wireless communication capability. The ESP32-CAM captures images at regular intervals or on user request and prepares them for transmission.

The captured images are then sent to the backend server through a network connection such as Wi-Fi. This communication layer plays a key role in transferring data between the edge device and the processing unit. Since the ESP32-CAM has limited processing power and memory, it does not perform any complex image processing or deep learning tasks.

The backend server acts as the main processing unit of the system. It receives the images from the edge device and performs preprocessing operations such as noise reduction and contrast enhancement. After preprocessing, the images are passed to the YOLOv3 model for traffic sign detection and classification. The model identifies traffic signs and generates outputs in the form of bounding boxes and labels.

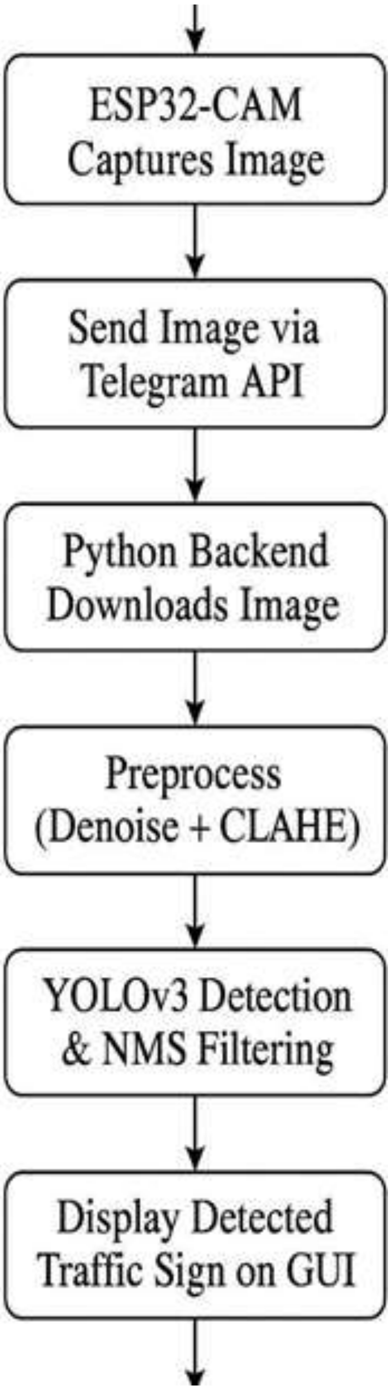


Fig.1. Block Daigram

Once the processing is completed, the results are sent back to the user through a graphical user interface (GUI). The GUI displays the annotated images and provides an easy way to monitor the system output. This makes the system user-friendly and suitable for practical applications.

The overall data flow of the system can be summarized as follows: image capture by the ESP32-CAM, transmission to the server, preprocessing and detection using the YOLOv3 model, and display of results on the user model This architecture ensures efficient utilization of resources whilemaintaining high detection accuracy. such as noise reduction and contrast enhancement. After preprocessing, the images are passed to the YOLOv3 model for traffic sign detection and classification. The model identifies traffic signs and generates outputs in the form of bounding boxes and labels.

The proposed system is scalable and can be extended by adding more edge devices or improving the backend processing capabilities. It provides a flexible and cost-effective solution for traffic sign recognition in real-world environments.

IV. RESULT AND DISCUSSION

To evaluate the performance of the proposed system, a combination of quantitative and qualitative analyses was conducted. The primary objective was to assess the accuracy, robustness, and latency of the YOLOv3 -based when deployed on an embedded edge platform interfaced with the interface ESP32-CAM and the Telegram-based backend.

A. Experimental Setup

The YOLOv3 model was trained using the German Traffic Sign Recognition Benchmark (GTSRB) dataset.

For testing, two separate datasets were used:

Standard Test Set: A clean subset of 2,000 images from the GTSRBdataset not used during training.

2. Obscured Test Set: A modified version of the same dataset in which 50% of the images were synthetically occluded with mud and water spots to simulate real-world conditions.

B. Performance Metrics

To evaluate performance, three standard object detection metrics wereused:

- Precision (TP / (TP + FP)) – Represents the accuracy of positive detections.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-Jun-2026, ISSN (Online): 2395-566X

- Recall (TP / (TP + FN)) – Measures the ability to detect all relevant objects.
- Mean Average Precision (mAP@0.5) – Primary evaluation metric indicating detection accuracy across all classes with IoU $\geq$ 0.5

C. Quantitative Results

The model's performance on both test sets is summarized in Table 2.

Test Dataset	Precision (%)	Recall (%)	mAP@0.5 (%)
GTSRB – Standard Test Set	97.2	96.5	96.8
GTSRB – Obscured Test Set	91.5	89.8	90.3

Table 2: Model Performance on Standard and Obscured Test Datasets.

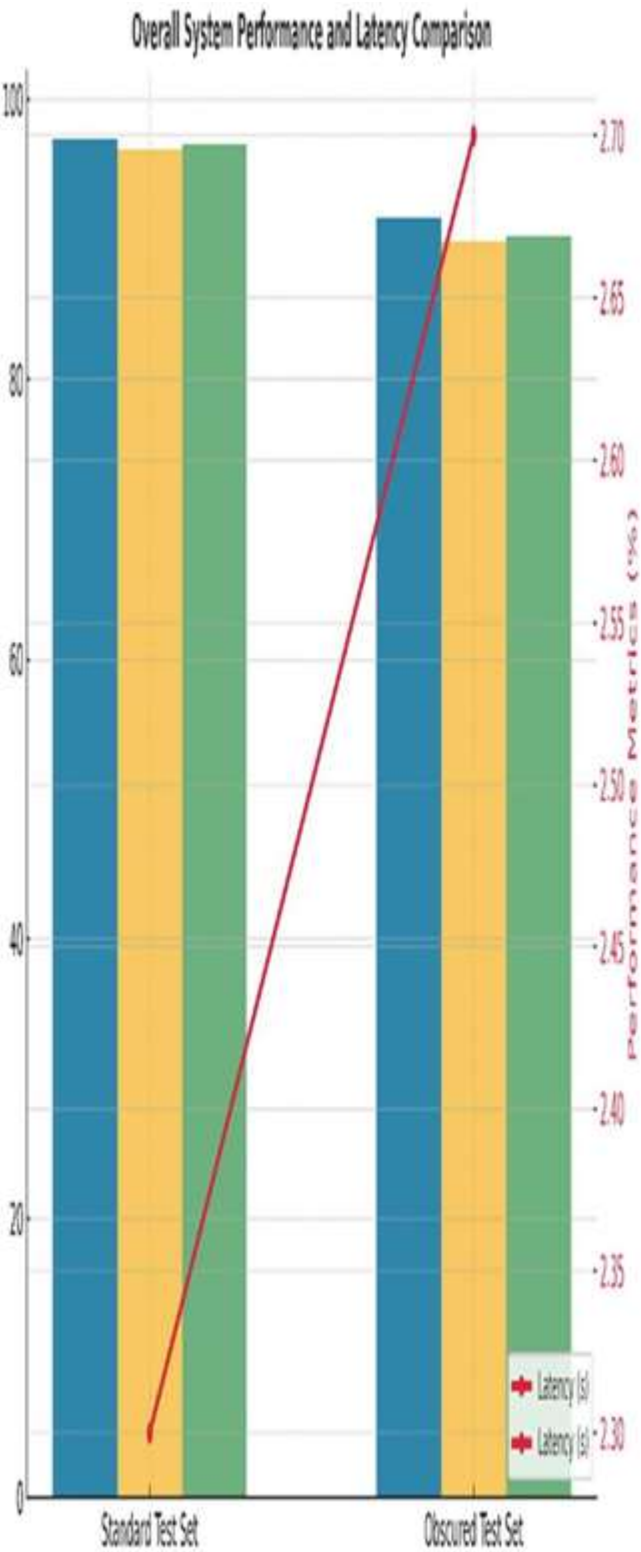
The results clearly show that while the performance dropped slightly on obscured data, the model still maintained high accuracy, confirming its robustness.

D. Discussion of Results

The YOLOv3 model achieved a mean Average Precision (mAP) of 96.8% on the standard test set, demonstrating excellent accuracy and reliability. When tested on the obscured dataset, the mAP decreased moderately to 90.3%, proving that the system remains resilient even under partially degraded visual conditions.

The overall end-to-end latency from image capture to displaying detection results was approximately 2 –3 seconds, primarily due to network transmission and CPU-based inference time. This confirms that while the system is suitable for semi-real-time operation, it could further benefit from edge-based inference acceleration (e.g., using NVIDIA Jetson or Tensors).

E. Graphical Presentation used:



Graph.1. Overall system performance graph

F. Error Analysis

A detailed analysis of failure cases on the obscured dataset revealed specific error patterns:

- **False Negatives (FN):** Primarily occurred when more than 40% of the sign surface was covered by occlusion, especially if the central symbol was hidden.
- **False Positives (FP):** Rare, but occasionally caused by background objects (e.g., circular logos or road patterns) that resemble traffic signs.
- **Misclassification:** Common in speed limit signs where only the numeric value was obscured, leading to confusion between similar signs like “50” and “80.”

These insights indicate that future improvement can be achieved using context-aware training or occlusion-specific data augmentation.

V. CONCLUSION

In this paper, we presented a comprehensive and integrated system for robust Traffic Sign Recognition, successfully addressing the critical challenge of identifying signs under adverse conditions. The core of our innovation lies in the synergistic fusion of accessible IoT hardware—specifically the ESP32-CAM—with a powerful deep learning model, YOLOv3, interconnected via a reliable cloud-based communication channel using a Telegram bot. This architecture proves to be not only cost-effective but also remarkably effective in solving a real-world problem that poses a significant threat to road safety.

Our quantitative evaluation yielded compelling results that validate the system's design. The model achieved an impressive mean Average Precision (mAP) of 96.8% on the standard German Traffic Sign Recognition Benchmark, confirming its high accuracy under ideal conditions. More importantly, when faced with a more challenging dataset featuring synthetically obscured signs, the system demonstrated exceptional resilience, maintaining a high mAP of 90.3%. This result is a direct testament to the robustness of the YOLOv3 architecture and the effectiveness of our targeted data augmentation strategy, proving that the system can reliably function even when visual information is degraded.

The project successfully demonstrates a practical and integrated solution that brings together embedded systems, cloud communication, and deep learning. While the end-to-end latency of 2.5 seconds, primarily due to network dependency, limits its current application to semi-real-time scenarios, the system provides a strong proof-of-concept. Ultimately, this work contributes a valuable and scalable blueprint for developing intelligent transportation systems that enhance situational awareness and driver safety in the complex and often unpredictable road environments of today learning. The strong quantitative results, achieving over 90% mAP even on obscured signs, validate our methodology.

REFERENCES

1. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.
   J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv preprint, arXiv:1804.02767, 2018.
2. S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
3. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, “SSD: Single Shot MultiBox Detector,” European Conference on Computer Vision (ECCV), pp. 21–37, 2016.
4. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, “The German Traffic Sign Recognition Benchmark: A Multi-Class Classification Competition,” International Joint Conference on Neural Networks (IJCNN), pp. 1453–1460, 2011.
5. D. Ciresan, U. Meier, J. Masci, and J. Schmid Huber, “A Committee of Neural Networks for Traffic Sign Classification,” IJCNN, pp. 1918–1921, 2011.
6. M. Haloi, “Traffic Sign Classification Using Deep Inception Networks,” arXiv preprint, arXiv:1511.02992, 2015.
7. B. Novak, V. Ilić, and B. Pavković, “YOLOv3 Algorithm with Additional Convolutional Neural Network Trained for Traffic Sign Recognition,” Proceedings of the Telecommunications Forum (TELFOR), pp. 420–423, 2018.