INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

# A Two-Phase Approach for Efficient Traffic Sign Detection and Recognition

**S Uma[1]*, S Prateeksha[2], V Padmapriya[3]**

**1** Associate Professor, Department of Computer Applications, B.M.S. College of Engineering, Bangalore, Karnataka, India
**2** Associate Consultant, CertPro, Bangalore, Karnataka, India
**3** Assistant Professor, Department of Computer Applications, B.M.S. College of Engineering, Bangalore, Karnataka, India

OPEN ACCESS

**Received:** 23-11-2023

**Accepted:** 21-02-2024

**Published:** 20-03-2024

**Citation:** Uma S, Prateeksha S, Padmapriya V (2024) A Two-Phase Approach for Efficient Traffic Sign Detection and Recognition. Indian Journal of Science and Technology 17(12): 1203-1212. https://doi.org/10.17485/IJST/v17i12.2985

* **Corresponding author**.

uma.mca@bmsce.ac.in

**Funding:** None

**Competing Interests:** None



Published By Indian Society for Education and Environment (iSee)

**ISSN**
Print: 0974-6846
Electronic: 0974-5645

## Abstract

**Objectives**: The objective of this study is to enhance the accuracy of traffic sign detection and recognition in modern intelligent transport systems, addressing real-time challenges under varying conditions. **Methods:** A two-phase approach is adopted. The first phase employs the You Only Look Once version 8 (YOLOv8) architecture to efficiently detect traffic signs under real-time conditions, considering variables like adverse weather and obstructions. Subsequently, the second phase employs a sequential convolutional network for precise recognition, utilizing the output from the first phase. This integrated method enhances traffic sign detection and recognition, contributing to road safety and efficient traffic management in complex transportation scenarios. **Findings:** The YOLOv8 architecture, utilized in Phase 1, demonstrated exceptional performance with a mean Average Precision (mAP) of 0.986 during validation. In Phase 2, the Convolutional Neural Network (CNN)-based recognition model achieved an impressive test accuracy of 98.7% on 463 test images, with a low-test loss of 0.1186, indicating consistent accuracy. The robustness of both models is confirmed by successful testing with three unseen images. YOLOv8 accurately detected and classified these images, while the CNN model correctly recognized them. These findings underscore the effectiveness of the two-phase approach in enhancing traffic sign detection and recognition, with significant implications for improving road safety and traffic management in real-world scenarios. **Novelty:** The novelty of this approach lies in its seamless integration of YOLOv8 for efficient traffic sign detection and a sequential convolutional network for accurate recognition, offering a significant advancement in addressing real-time challenges and contributing to enhancing road safety and traffic management in an increasingly complex transportation landscape.

**Keywords:** Traffic sign detection; Traffic sign recognition; Convolutional Neural Networks; YOLOv8; Object detection

# 1 Introduction

In today's world, ensuring road safety and efficient traffic management has become a paramount concern. Modern intelligent transport systems strive to address this concern by utilizing advanced technologies. One crucial aspect of these systems is the accurate detection and recognition of traffic signs, which play a pivotal role in guiding drivers and enhancing overall road safety[(1,2)]. Traffic signs are visual symbols, markings, or devices placed alongside roads and highways to convey important information and instructions to autonomous vehicles, drivers, pedestrians, and other road users[(3,4)]. These signs serve as a standardized communication system, using easily recognizable symbols, colors, and shapes to provide guidance on various aspects of road safety and traffic management. Traffic signs can communicate information about speed limits, directions, warnings of hazards, regulatory rules, and other essential guidance to help ensure safe and efficient transportation on roadways. These seemingly unremarkable road signs serve as a universal language, conveying vital information and guiding the choices and actions of autonomous vehicles as well as drivers, cyclists, and pedestrians on the road. However, detection and recognition are a significant challenge due to various factors such as adverse weather conditions, obstructions, and different lighting situations[(5)].

Utilizing computer vision, machine learning, and artificial intelligence to equip machines with the ability to identify, interpret, and react to traffic signs much like a human driver would be advantageous in this situation. This is known as automatic traffic sign detection and recognition. These systems are not mere conveniences but life-saving tools that operate silently in the background, ensuring that the right information reaches the right eyes at the right time[(6)]. They play a pivotal role in mitigating the risks of accidents, reducing congestion, and enhancing overall traffic management, ultimately contributing to the noble goal of road safety.

The need for advanced automatic traffic sign detection and recognition systems becomes even more pronounced when considering the challenges posed by adverse weather conditions, discoloration of signs, and various environmental factors. In recent years, the automotive industry has witnessed a rapid evolution toward Advanced Driver Assistance Systems (ADAS) and, eventually, autonomous vehicles. These technological advancements rely heavily on accurate and timely information from the surrounding environment, making the robust recognition of traffic signs crucial for their safe and efficient operation. In adverse weather conditions, such as heavy rain, snow, or fog, visibility is severely compromised, posing a significant challenge to both human drivers and automated systems. Automatic traffic sign detection becomes instrumental in providing an additional layer of safety by ensuring that critical information, like speed limits and warnings, is accurately conveyed to the vehicle's control system.

Moreover, the discoloration of traffic signs over time due to factors like sunlight exposure and environmental wear can lead to a decline in their visibility and legibility. This poses a considerable risk, especially in situations where immediate and precise recognition of signs is essential for making split-second decisions on the road. Adding to these challenges, the complexity of urban environments introduces further considerations. High-density traffic areas, intricate road networks, and dynamic traffic patterns amplify the difficulty of reliable traffic sign recognition. Automatic systems equipped with advanced computer vision capabilities are essential in navigating through such intricate scenarios. For non-ADAS vehicles, where human drivers still play a central role, the importance of automatic traffic sign detection is equally crucial. In situations where drivers may be fatigued, distracted, or facing challenging weather conditions, the system can act as a reliable safety net, reducing the likelihood of misinterpretation or oversight of vital traffic information.

If the challenges associated with automatic traffic sign detection and recognition are not effectively addressed, the consequences could be severe. Inaccurate or delayed recognition of traffic signs may lead to misinterpretations, increasing the risk of accidents, traffic violations, and overall road safety hazards. Inefficient traffic management, especially in congested urban areas, could result in heightened traffic congestion, longer commute times, and increased environmental pollution. Furthermore, as cities move towards smart infrastructure and interconnected transportation systems, the ability to integrate seamlessly with emerging technologies becomes crucial. Failure to address these challenges may not only compromise road safety but also hinder the realization of smart cities' potential, where efficient traffic management is integral to sustainable urban living.

Therefore, the development and implementation of robust automatic traffic sign detection and recognition systems are imperative not only to assist ADAS-equipped vehicles but also to enhance the safety and efficiency of the entire spectrum of road users. Failure to address these challenges may undermine the potential benefits of advanced transportation technologies and compromise the overarching goal of creating a safer and more streamlined road network.

The literature shows that many researchers have harnessed the power of deep learning and computer vision techniques to develop innovative approaches for traffic sign recognition. From CNNs specialized in circular sign recognition to novel architectures capable of simultaneously detecting and classifying signs, these works contribute to the ongoing efforts aimed at improving the accuracy, efficiency, and real-world applicability of traffic sign recognition systems. The literature survey was carried out to uncover the limitations by closely examining the methodologies, datasets, and the advantages and drawbacks associated with these studies. Understanding these limitations is crucial for gaining valuable insights and forming the basis for proposing innovative methods and solutions to advance the state of the art in traffic sign recognition.

In[7], the authors propose a traffic sign recognition technique based on CNNs with a focus on circular signs. They achieved a high test accuracy of 98.2% on the German Traffic Sign Recognition Benchmark (GTSRB)[8] by combining CNNs with pre-processing steps. However, the system is limited to circular signs and its resource requirements and sensitivity to input data quality are challenging. In[9], the authors introduce a Multi-Resolution Convolutional Neural Network (MR-CNN), a network designed for small traffic sign detection. They achieved a detection mean Average Precision (mAP) of 71.5% on the Tsinghua-Tencent 100K dataset by utilizing multi-scale deconvolution and contextual information. However, the system faces complexity and dataset diversity challenges. In[10], the authors provide a comprehensive review and analysis of traffic sign detection (TSD) methods, a crucial component of advanced driver-assistance systems and autonomous driving systems. It categorizes TSD methods into five groups: color-based, shape-based, color- and shape-based, machine learning-based, and Laser Imaging, Detection, and Ranging (LIDAR) based methods. The review discusses the strengths and limitations of each category, highlighting the effectiveness of machine learning-based approaches while noting challenges in balancing computational efficiency and accuracy, particularly with small and vague traffic signs. This work emphasizes the need for new datasets to address variations in traffic signs across different countries and challenges posed by nighttime and extreme weather conditions. While public datasets have demonstrated high performance, there's limited room for improvement, prompting exploration into more challenging scenarios and LIDAR-based solutions. However, LIDAR data collection remains a challenge for many researchers. In[11], Multi-Feature Fusion Single Shot Detector (MF-SSD) focuses on enhancing small traffic sign recognition through multi-feature fusion. It achieves precision scores of 28.8%, 67.5%, and 82.6% for small, medium, and large signs, respectively. Nevertheless, it lacks comprehensive comparisons, faces dataset dependencies, and needs optimization for real-time applications. In[12], the authors propose a computer vision system for speed limit traffic sign recognition, achieving a high-test accuracy of 98.97%. This system can assist drivers and enhance road safety but may face challenges in complex scenarios, sensitivity to environmental factors, and data annotation requirements. In[13], the authors propose the combination of YOLOv3 for object detection and CNNs for recognition achieves test accuracies of 89.56% and 86%, respectively. It excels in simultaneous sign location and classification but struggles with small objects and aspect ratios, requiring substantial computational resources. In[14], the authors propose a novel approach for traffic sign recognition in the context of the Internet of Vehicles (IoV) using Federated Learning (FL) and Spike Neural Networks (SNNs). They addressed the privacy concerns associated with centralized data collection by developing a FL framework that allows vehicles to collaboratively train a recognition model without sharing raw traffic sign data. A unique aspect of their approach is the use of SNNs, which offer advantages in terms of energy efficiency and robustness. They introduced a Neuronal Receptive Field Encoding (NRFE) method for efficient feature extraction from traffic sign images. Experimental results demonstrated the effectiveness of their Federated SNN (FedSNN) approach, especially in handling non-independently and Identically Distributed (non-IID) data distributions and noisy images. FedSNN showed promise for traffic sign recognition in IoV scenarios, achieving energy efficiency and robustness. However, there are still challenges related to communication overhead, device reliability, and security in the FL framework, which require further investigation. In[15], the authors address the challenge of traffic sign recognition in Pakistan using deep learning techniques. They began by highlighting the limitations of conventional image processing methods and emphasized the importance of CNNs in computer vision, with a focus on the availability of datasets. Since no traffic-sign dataset existed for Pakistan, they collected and annotated one, consisting of 359 different images. The methodology involved pre-training a CNN model on the German traffic sign dataset and fine-tuning it with the Pakistani dataset. They achieved a training accuracy of 54.8% for Pakistani traffic sign recognition, outperforming other techniques. However, limitations include the small size of the Pakistani dataset, which can affect generalization, and the need for more extensive data collection in the future to improve accuracy further. In[16], the authors introduce Traffic Sign Recognition with Spatial Attention (TSR-SA), a novel method for small traffic sign recognition, inspired by YOLOv4 and YOLOv5 object detection frameworks. The methodology enhances model performance through improved backbone features, receptive field blocks, refined detector grids, and a unique data augmentation technique called Random Erasing-Attention (RE-A). The Tsinghua-Tencent 100K (TT100K) dataset[17], known for its challenging small traffic sign scenarios, serves as the evaluation benchmark, where TSR-SA achieves an outstanding mAP of 90.2% while maintaining real-time processing. Despite its success, class imbalance within the dataset is recognized as a limitation, and the study highlights the need for research to address recognition in nighttime and extreme weather conditions. In[18], MSE-YOLOv4 simultaneously detects and classifies traffic signs, achieving mAP scores of 84.44% and 97.9% on their dataset and the German Traffic Sign Detection Benchmark (GTSDB)[19]. However, its performance may vary in complex scenarios, warranting further evaluation of diverse datasets and conditions. In[20], the authors present a methodology for traffic sign detection using the Ghost-YOLO model. The study utilized the TT100k dataset, which includes 9170 images with various traffic sign scenarios. Ghost-YOLO achieved an impressive mAP of 92.71%, showcasing its effectiveness in detecting traffic signs in complex environments. However, the model's limitations include the need for further optimization for mobile deployment and exploring image classification tasks. In [21], the authors propose an improved one-stage traffic sign detection algorithm called Attention-YOLOv4 to achieve real-time and high-accuracy detection. They evaluate the algorithm on the TT100K dataset, which contains 221 classes of traffic signs, and compare it with other methods such as Faster Region-based Convolutional Neural Network (R-CNN), Single Shot MultiBox Detector (SSD), Feature Pyramid Network (FPN), and YOLOv4. The proposed method achieves a balance between detection speed and accuracy, with a mAP of 86.1% at an Intersection over Union (IoU) of 0.5 and a detection speed of 29.1 ms per image. However, the study acknowledges there is still room for improvement in detection accuracy, suggesting that further enhancements may be explored. Additionally, the evaluation is primarily based on the TT100K dataset, and the method's performance in other real-world scenarios and datasets remains to be tested, potentially posing limitations in broader applications. A detailed survey on the use of various versions of YOLO for traffic sign detection and recognition has been conducted by the authors [22].

In scrutinizing the landscape of traffic sign recognition as portrayed by existing literature, several discernible gaps come to the forefront. Notably, a prevalent limitation involves the confined recognition scope observed in certain models. A number of studies within the surveyed literature tend to focus on specific types of traffic signs, such as circular or small signs, potentially impeding their adaptability to the diverse array of signs encountered in real-world scenarios. This limitation underscores the need for a more comprehensive approach that can adeptly handle the multitude of traffic sign variations present on roadways. Additionally, a pronounced gap is discerned in the diversity of datasets employed, with some studies falling short of adequately representing the real-world scenarios and variations inherent in traffic sign landscapes. This lack of diversity in datasets may hinder the models' ability to generalize effectively to practical situations, demanding a more inclusive approach in dataset selection.

In response to these identified gaps, the proposed methodology emerges as a robust solution that systematically addresses and rectifies these limitations. Foremost among the advantages is the integrated approach that combines the strengths of YOLOv8 for efficient traffic sign detection and a tailored CNN architecture for accurate recognition. This amalgamation directly addresses the confined recognition scope noted in the literature by enabling a multiclass classification system that accommodates the diverse range of traffic sign shapes and types. By embracing a holistic approach that integrates detection and recognition phases, the proposed system fills a crucial void left by models with narrow recognition scopes, offering a more inclusive and adaptable solution. Furthermore, the proposed work's emphasis on real-time processing serves as a direct response to the literature's gaps in efficient multiclass detection. Leveraging the YOLOv8 architecture, the methodology not only achieves high precision, recall, and mAP scores across various classes of traffic signs but also ensures timely detection which is a crucial aspect for practical applications [23–25]. The recognition of previously unseen images underscores the system's generalization capabilities, mitigating another gap identified in models that may lack emphasis on real-world scenarios. In essence, the proposed methodology not only identifies the gaps prevalent in the literature surveyed but, more importantly, actively contributes to bridging these gaps by presenting a comprehensive, real-world adaptable, and efficient solution to the challenges in traffic sign recognition.

A novel approach for the detection and recognition of traffic signs has been proposed to address various challenges, such as the variability in traffic signs, demanding environmental conditions, data collection and annotation complexities, and the need for robust generalization. Additionally, this approach aims to overcome real-time processing constraints, resource efficiency requirements, dataset bias, and class imbalance, while also addressing critical aspects of security and privacy and sensitivity to input data quality. Furthermore, the approach is designed to meet the unique demands of extreme conditions and enhance the capabilities of existing systems. This approach operates in two phases, aiming to improve accuracy and efficiency. This two-phase approach, building upon an established system, has the potential to revolutionize how traffic signs are detected and understood in real-time scenarios. The implementation combining YOLOv8 and CNN represents a pioneering achievement in this research endeavor. The proposed approach introduces a novel combination of YOLOv8 for rapid traffic sign detection and a dedicated CNN for accurate recognition. It seamlessly addresses real-world challenges, achieving robust performance with a mAP of 0.986 and a test accuracy of 98.7%. Its advantages lie in enhancing road safety, efficient traffic management, and scalability for future applications in diverse traffic scenarios.

This paper is structured into several key sections to present a coherent framework for efficient traffic sign detection and recognition. It commences with an abstract summarizing the research's objectives and contributions. Section 1, the introduction, contextualizes the importance of accurate traffic sign recognition in modern intelligent transport systems and a literature survey, providing insights into prior work and methodologies. In Section 2, the methodology elaborates on the proposed two-phase approach, detailing the use of YOLOv8 and CNN architectures, along with dataset information. Section 3, the results and discussion section critically evaluate the performance of the proposed method. Finally, in Section 4, the conclusion summarizes key findings and potential implications, and the paper closes with a list of references.

## 2 Methodology

The proposed algorithm performs traffic sign detection and recognition in two phases. In phase 1, traffic signs are detected using the existing YOLOv8 architecture. They are classified into 7 major classes based on color, shape, and image background i.e., Circular shape with a red border and white background as class 0, Circular with gray signs and white background as class 1, Circular with blue background as class 2, Circular with red background as class 3, Diamond as class 4, Triangle as class 5 and Hexagon as class 6. In phase 2, traffic signs are recognized using the newly designed CNN architecture. The Regions of Interest (ROIs) from the first phase are cropped to serve as the test data for the proposed CNN model.

### 2.1 Phase 1:Traffic Sign Detection using YOLOv8 architecture

In the first phase, the YOLOv8 architecture is employed for the detection of traffic signs. The GTSDB dataset was used and formatted to meet YOLOv8 standards which require the annotations in a text file for each image and a data configuration file containing paths to training and validation data. The implementation leverages the YOLOv8 architecture, specifically the variant YOLOv8n, to carry out the detection process. The model undergoes training for 20 epochs, utilizing images of dimensions 416 x 416 pixels. The hyperparameters used in this phase are shown in Table 1.

**Table 1. Hyperparameters for the YOLOv8 architecture**

| | |
|---|---|
| Initial learning rate | 0.0015 |
| Batch size | 32 |
| Momentum | 0.9 |
| Weight decay | 0.0005 |

1. **Initial Learning Rate:** A smaller learning rate allows for more precise updates, which is beneficial when fine-tuning a complex model like YOLOv8. The chosen learning rate of 0.0015 strikes a balance between training stability and convergence speed.
2. **Batch Size:** A batch size of 32 is used because it provides a good balance between computational efficiency and model convergence.
3. **Momentum:** A momentum value of 0.9 is used as it ensures that the gradient updates have a strong influence from the previous updates, aiding in faster convergence.
4. **Weight Decay (0.0005):** A weight decay value of 0.0005 is chosen to strike a balance between regularizing the model and allowing it to learn from the data effectively.

The results of this phase include ROI predictions, which are extracted and preserved for testing in the subsequent phase.

### 2.2 Phase 2: Traffic Sign Recognition using CNN architecture

In the second phase, a CNN-based architecture is designed and trained for traffic sign recognition. The dataset used for this phase is the GTSRB dataset. The dataset consists of images of traffic signs, each categorized into 43 classes representing different sign types. The following sections outline the essential stages of traffic sign recognition, encompassing data preprocessing, preparation, and the architecture of the CNN model.

- **Data Preprocessing and Visualization:** The GTSRB dataset undergoes preprocessing, where images are resized to a uniform 32x32-pixel resolution. These images are then converted into numerical arrays, and data is visualized to ensure data quality and realism.
- **Data Preparation:** In this phase, the image data is transformed into the float32 format, and labels are one-hot encoded. This preparation step ensures that the data is in a suitable format for input into the CNN model.
- **Dataset Splitting:** The dataset is partitioned into two equal parts, 50% for training, and 50% of the data is reserved for validation. This division facilitates effective model training while allowing for performance validation on a separate portion of the data.
- **Test Data Preparation:** Test data for the CNN model is acquired from Phase 1 and is subject to similar preprocessing and conversion steps to ensure compatibility with the model.

- **CNN Model Architecture:** The CNN model architecture is designed to capture and learn relevant features from the traffic sign images. The architecture includes convolutional layers, pooling layers, batch normalization, dropout layers, and fully connected layers, with a final softmax output layer for classification. In summary, the model processes input data through a sequence of layers, performing feature extraction, regularization, and classification using the softmax activation function.

Figure 1 shows the model summary of the proposed architecture.

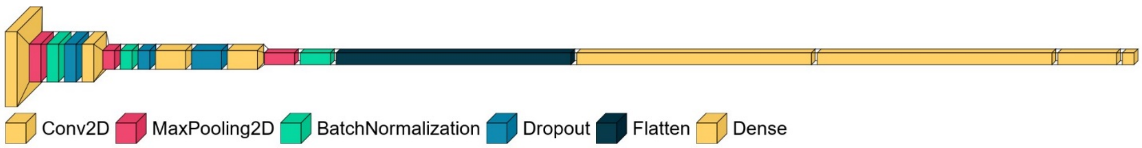


**Fig 1. The proposed architecture for object recognition**

- **Model Training and Evaluation:** The CNN model is compiled using the Adam optimizer and categorical cross-entropy loss function. It is then trained using the training dataset and validated using the validation dataset. The training process is repeated over 20 epochs to optimize the model's performance.

The CNN model architecture is designed to capture and learn relevant features from the traffic sign images.

The proposed methodology has been applied to the following two datasets:

**GTSDB:** The GTSDB dataset comprises 1526 annotated images classified into 43 classes which are organized into "train," "valid," and "test" subsets containing 1063, 308, and 155 images respectively. These images are annotated with precise details about traffic sign positions and categories. The dataset consists of real-world scenarios, including varied lighting and backgrounds, offering a benchmark to assess traffic sign detection algorithms. Annotations serve as ground truth for evaluating detection performance. Annotations are standardized using RoboFlow, ensuring consistent quality and dimensions. The dataset is sourced from RoboFlow.

**GTSRB:** The GTSRB dataset available on Kaggle comprises 39,209 images classified into 43 classes, designed to facilitate traffic sign recognition. This dataset is divided into two subsets: a training set and a validation set with an equal number of images. Each image corresponds to a different traffic sign, covering a wide range of sign types and variations.

## 3 Results and Discussion

### 3.1 Phase 1: Traffic Sign Detection using YOLOv8 architecture

In Phase 1, the YOLOv8 architecture is applied for traffic sign detection using the GTSDB dataset. Upon validation, the model achieved a mAP of 0.986, demonstrating its strong performance in traffic sign detection. The results are indicative of its robust performance, achieving a high precision (P) and recall (R) of 0.978. The mAP at IoU of 0.50 is 0.993, and the mAP at IoU ranging from 0.50 to 0.95 is 0.986. This performance is consistently observed across various classes of traffic signs. Particularly, class 0, class 2, class 5, and class 6 exhibit outstanding precision, recall, and mAP scores, demonstrating the model's accuracy in detecting and localizing traffic signs across multiple categories. The processing speed per image is efficient, with preprocessing taking 2.9ms, inference 132.1ms, loss calculation 0.0ms, and post processing 0.8ms. The results of this phase include ROI predictions, which are extracted and preserved for testing in the subsequent phase.

### 3.2 Phase 2: Traffic Sign Recognition using CNN architecture

The CNN model on the test dataset from phase 1 consisting of 463 images achieves a test accuracy of 98.7%. The test loss is calculated at 0.1186, indicating that the model's predictions are consistently close to the actual values. The evaluation was conducted across 15 batches, and each batch was processed efficiently in approximately 31-32 milliseconds per step. The Confusion Matrix for Traffic Sign Recognition using CNN Architecture is shown in Figure 2.

The Confusion matrix in Figure 2 provides a comprehensive summary of the performance of the CNN-based traffic sign recognition model. It visualizes the comparison between predicted traffic sign classes and their corresponding ground truth labels across the test dataset. Each row in the matrix represents the actual traffic sign class, while each column represents the predicted class. The diagonal elements of the matrix indicate the number of correctly classified instances for each traffic

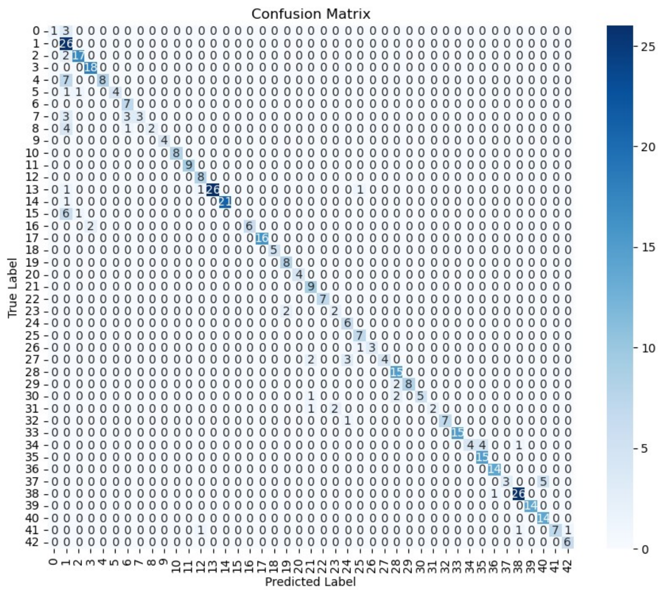


**Fig 2. Confusion Matrix for Traffic Sign Recognition using CNN Architecture**

sign class, while off-diagonal elements represent misclassifications. Notably, the CNN model is trained to recognize traffic signs across 43 distinct classes, encompassing a wide range of sign types and variations commonly encountered in road environments.

Table 2 consists of the predicted class, the actual class, and the count of instances for each scenario. In this matrix, there are 457 instances where the predicted class matches the actual class, representing true positives (TP) for each class. The sum of these values indicates the total number of correct classifications made by the model. Additionally, there are 6 instances where the predicted class does not match the actual class, indicating the total number of incorrect classifications. There were 6 instances in the test dataset that the model misclassified. This could be due to the complexity of the images, similarities between certain classes, or limitations in the model's ability to generalize.

**Table 2. Confusion Matrix of Traffic Signs With 43 Classes**

| | Predicted Class | Actual Class | Count |
|---|---|---|---|
| **0** | Total Correct | Total Correct | 457 |
| **1** | Total Incorrect | Total Incorrect | 6 |

In Table 3, the confusion matrix provides an overview of the model's performance for two classes, "Yield" (class 13) and "Keep right" (class 38). For Class 13, the model correctly predicts it as "Yield" in 29 cases. There were no cases where the model incorrectly predicted other classes as "Yield". For Class 38 "Keep right", the model correctly predicts it as "Keep right" in 27 cases. There were no cases where the model incorrectly predicted other classes as "Keep right".

**Table 3. Overview of the model's performance for Class 13 and Class 38**

| Confusion Matrix Overview | | |
|---|---|---|
| | **Predicted: 13** | **Predicted: 38** |
| **class: 13** | 29 | 0 |
| **class: 38** | 0 | 27 |

Furthermore, to evaluate the robustness and generalization capabilities of the models, three previously unseen images were tested on both the YOLOv8 and CNN models. In the case of YOLOv8, all three unseen images were successfully detected and classified with high accuracy, demonstrating the model's ability to handle real-world scenarios effectively. Similarly, when subjected to the CNN-based recognition model, all three unseen images were recognized correctly, further validating the model's capacity to accurately identify various traffic sign types.

Figure 3 shows the three unseen traffic signs with YOLOv8 detections classifying them as follows: the first sign is identified as class 3, indicating a circular shape with a red background, the second sign is classified as class 0, representing a circular shape with a red border and white background, and the third sign is recognized as class 2, which corresponds to a circular shape with a blue background.


**Fig 3. Detections using YOLOv8 architecture**

Figure 4 shows the three unseen traffic signs, which have been cropped from the results of Phase 1 and recognized using the CNN model and are identified as "No Entry," "Speed Limit (60km/h)," and "Keep Right."

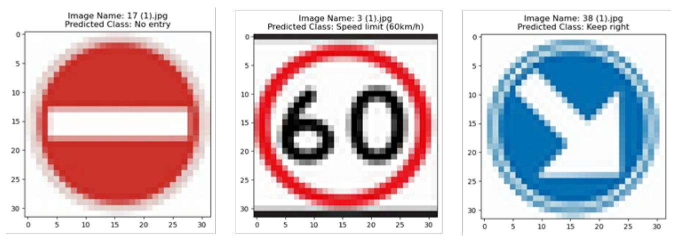


**Fig 4. Recognition using CNN architecture**

**Table 4. Comparison of the proposed model with the existing model**

| Metric | Proposed Model | Previous Model [12] |
|---|---|---|
| **Detection Algorithm** | YOLOv8 with CNN | YOLOv3 with CNN |
| **Test Accuracy (Detection)** | 0.986 | 0.895 |
| **Test Accuracy (Recognition)** | 98.7% | 86% |

To prove the validity of the proposed model, the results of the proposed approach are compared with the work carried out by the authors in [13]. Table 4 shows the results of the comparison. Both studies propose a two-phase approach for traffic sign detection and recognition. The datasets used in both papers include GTSDB and GTSRB ensuring a comprehensive evaluation of the proposed methodology. In terms of results, the proposed work, utilizing YOLOv8 for detection, achieved a mean Average Precision (mAP) of 0.986 during validation, demonstrating high precision and recall. The CNN-based recognition model in this work achieved an impressive test accuracy of 98.7%. In contrast, the previous work, which employs YOLOv3 for detection, achieved a test accuracy of 0.895 and 86% for detection and recognition, respectively. The results indicate that the proposed work outperforms in both detection and recognition accuracy. Compared to YOLOv3, the Anchor-Free detection of YOLOv8 avoids the limitations of choosing the best-fitting anchor for an object, improving the localization accuracy of images. Separate loss functions for objectness, class prediction, and bounding box prediction allow for better optimization and fine-tuning of each aspect, leading to more accurate predictions. YOLOv3 uses a single loss function whereas YOLOv8 uses separate loss functions for objectness, class prediction, and bounding box prediction resulting in more optimal performance trade-offs. Data augmentation techniques like MixUp and CutMix improve the model's robustness and generalizability, performing better on unseen data compared to YOLOv3. The proposed work yields superior results in terms of accuracy for both detection and recognition, providing a more robust and accurate system for traffic sign detection and recognition. The higher mAP in the detection phase and higher test accuracy in the recognition phase suggest that the proposed two-phase approach, incorporating YOLOv8 and a CNN-based architecture, is more effective in addressing the challenges of real-world traffic scenarios. This can contribute to enhanced road safety and traffic management, surpassing the performance of the previous work.

The results of the proposed two-phase approach for traffic sign detection and recognition showcase its effectiveness in real-world scenarios. The YOLOv8 architecture demonstrated exceptional performance in swiftly and accurately detecting traffic signs, as evidenced by a mean Average Precision (mAP) of 0.986 during validation. The subsequent CNN-based recognition model achieved an impressive test accuracy of 98.7% on 463 test images, emphasizing its proficiency in accurately identifying various traffic sign types. Moreover, when tested with three previously unseen images, both the YOLOv8 and CNN models successfully detected and recognized the traffic signs with high accuracy. This not only underscores the robustness of the integrated approach but also highlights its efficiency in handling new, real-world scenarios. The ability to accurately classify traffic signs in a shorter time frame is critical for ensuring the practical viability of the system, contributing to improved road safety and efficient traffic management in diverse transportation landscapes.

## 4 Conclusion

In conclusion, the proposed two-phase approach for the detection and recognition of traffic signs has the potential to revolutionize road safety and traffic management in real-time scenarios. Compared to the other versions of YOLO, the utilization of the YOLOv8 architecture results in swift and accurate detection. Further, the sequential convolutional networks result in enhanced recognition. The proposed approach also addresses the challenges posed by adverse weather conditions, obstructions, and varying lighting situations. The YOLOv8 architecture demonstrated robustness in traffic sign detection, achieving a remarkable mean Average Precision (mAP) of 0.986 during validation, indicating high precision and recall rates. Furthermore, the CNN-based recognition model exhibited outstanding performance, achieving a test accuracy of 98.7% on unseen traffic sign images, confirming its proficiency in accurately identifying various traffic sign types. These high levels of accuracy are pivotal in ensuring the reliability and effectiveness of the system in real-world scenarios, ultimately contributing to enhanced road safety and traffic management. Successful implementation of this system could lead to safer roads, informed drivers, and improved transportation network efficiency.

With further development and deployment, this approach has the potential to make a significant impact on road safety and traffic management worldwide. In the future, the proposed methodology can be applied to a larger dataset like TT100K to improve the algorithm's ability to detect and recognize traffic signs with large-size differences in traffic images with complex backgrounds, and the proposed methodology can be made into a single-phase methodology.

## References


1) Global Status Report on Road Safety . 2023. Available from: https://www.who.int/teams/social-determinants-of-health/safety-and-mobility/global-status-report-on-road-safety-2023.
2) Triki N, Karray M, Ksantini M. A Real-Time Traffic Sign Recognition Method Using a New Attention-Based Deep Convolutional Neural Network for Smart Vehicles. *Applied Sciences*. 2023;13(8):1–21. Available from: https://doi.org/10.3390/app13084793.
3) Lim XR, Lee CP, Lim KM, Ong TS, Alqahtani A, Ali M. Recent Advances in Traffic Sign Recognition: Approaches and Datasets. *Sensors*. 2023;23(10):1–17. Available from: https://doi.org/10.3390/s23104674.
4) Manual on Uniform Traffic Control Devices (MUTCD). . Available from: https://mutcd.fhwa.dot.gov/index.htm.
5) Munawar HS. Image and Video Processing for Defect Detection in Key Infrastructure. In: Machine Vision Inspection Systems;vol. 1. Wiley. 2020;p. 159–177. Available from: http://dx.doi.org/10.1002/9781119682042.ch7.
6) Temirgaziyeva S, Omarov B. Traffic Sign Recognition With Convolutional Neural Network. *Scientific Journal of Astana IT University*. 2022;12(12):14–23. Available from: https://doi.org/10.37943/12YZFG6952.
7) Sun Y, Xue B, Zhang M, Yen GG, Lv J. Automatically Designing CNN Architectures Using the Genetic Algorithm for Image Classification. *IEEE Transactions on Cybernetics*. 2020;50(9):3840–3854. Available from: https://doi.org/10.1109/TCYB.2020.2983860.
8) GTSRB - German Traffic Sign Recognition Benchmark. . Available from: https://www.kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign.
9) Liu Z, Du J, Tian F, Wen J. MR-CNN: A Multi-Scale Region-Based Convolutional Neural Network for Small Traffic Sign Recognition. *IEEE Access*. 2019;7:57120–57128. Available from: https://doi.org/10.1109/ACCESS.2019.2913882.
10) Liu C, Li S, Chang F, Wang Y. Machine Vision Based Traffic Sign Detection Methods: Review, Analyses and Perspectives. *IEEE Access*. 2019;7:86578–86596. Available from: https://doi.org/10.1109/ACCESS.2019.2924947.
11) Jin Y, Fu Y, Wang W, Guo J, Ren C, Xiang X. Multi-Feature Fusion and Enhancement Single Shot Detector for Traffic Sign Recognition. *IEEE Access*. 2020;8:38931–38940. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9007360.
12) Salimullina AD, Budanov DO. Computer Vision System for Speed Limit Traffic Sign Recognition. In: Computer Vision System for Speed Limit Traffic Sign Recognition. In2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). IEEE. 2022;p. 415–418. Available from: https://doi.org/10.1109/ElConRus54750.2022.9755744.
13) Karthika R, Parameswaran L. A Novel Convolutional Neural Network Based Architecture for Object Detection and Recognition with an Application to Traffic Sign Recognition from Road Scenes. *Pattern Recognition and Image Analysis*. 2022;32(2):351–362. Available from: https://doi.org/10.1134/S1054661822020110.
14) Xie K, Zhang Z, Li B, Kang J, Niyato D, Xie S, et al. Efficient Federated Learning With Spike Neural Networks for Traffic Sign Recognition. *IEEE Transactions on Vehicular Technology*. 2022;71(9):9980–9992. Available from: https://doi.org/10.1109/TVT.2022.3178808.

15) Nadeem Z, Khan Z, Mir U, Mir UI, Khan S, Nadeem H, et al. Pakistani traffic-sign recognition using transfer learning. *Multimedia Tools and Applications* . 2022;81:8429–8449. Available from: https://doi.org/10.1007/s11042-022-12177-8.
16) Chen J, Jia K, Chen W, Lv Z, Zhang R. A real-time and high-precision method for small traffic-signs recognition. *Neural Computing and Applications*. 2022;34(3):2233–2245. Available from: https://doi.org/10.1007/s00521-021-06526-1.
17) Traffic Sign Detection Benchmark. . Available from: https://cg.cs.tsinghua.edu.cn/traffic-sign/.
18) Zhang Y, Lu Y, Zhu W, Wei X, Wei Z. Traffic sign detection based on multi-scale feature extraction and cascade feature fusion. *The Journal of Supercomputing*. 2023;79(2):2137–2152. Available from: https://doi.org/10.1007/s11227-022-04670-6.
19) GTSDB Dataset - Test Split. . Available from: https://universe.roboflow.com/pfe-fvlva/gtsdb-dat55/browse?queryText=split%3Atest&pageSize=50&startingIndex=0&browseQuery=true.
20) Zhang S, Che S, Liu ZS, Zhang X. A real-time and lightweight traffic sign detection method based on ghost-YOLO. *Multimedia Tools and Applications*. 2023;82(17):26063–26087. Available from: https://doi.org/10.1007/s11042-023-14342-z.
21) Li Y, Li J, Meng P. Attention-YOLOV4: a real-time and high-accurate traffic sign detection algorithm. *Multimedia Tools and Applications*. 2023;82(5):7567–7582. Available from: https://doi.org/10.1007/s11042-022-13251-x.
22) Cui Y, Guo D, Yuan H, Gu H, Tang H. Enhanced YOLO Network for Improving the Efficiency of Traffic Sign Detection. *Applied Sciences*. 2024;14(2):1–15. Available from: https://doi.org/10.3390/app14020555.
23) Terven J, Córdova-Esparza DM, Romero-González JA. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*. 2023;5(4):1680–1716. Available from: https://doi.org/10.3390/make5040083.
24) Jiang P, Ergu D, Liu F, Cai Y, Ma B. A Review of Yolo Algorithm Developments. *Procedia Computer Science*. 2022;199:1066–1073. Available from: https://doi.org/10.1016/j.procs.2022.01.135.
25) Aboah A, Wang B, Bagci U, Adu-Gyamfi Y. Real-time Multi-Class Helmet Violation Detection Using Few-Shot Data Sampling Technique and YOLOv8. 2023. Available from: https://doi.org/10.48550/arXiv.2304.08256.