

Traffic Sign Detection and Recognition Using Yolo Models

Mareeswari V.

School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology (VIT), Vellore, 632 014, India
E-mail: vmareeswari@vit.ac.in
ORCID iD: <https://orcid.org/0000-0002-2768-943X>

Vijayan R.*

School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology (VIT), Vellore, 632 014, India
E-mail: rvijayan@vit.ac.in
ORCID iD: <https://orcid.org/0000-0002-1164-7569>
*Corresponding author

Shajith Nisthar

School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology (VIT), Vellore, 632 014, India
E-mail: shajith.2021@vitstudent.ac.in

Rahul Bala Krishnan

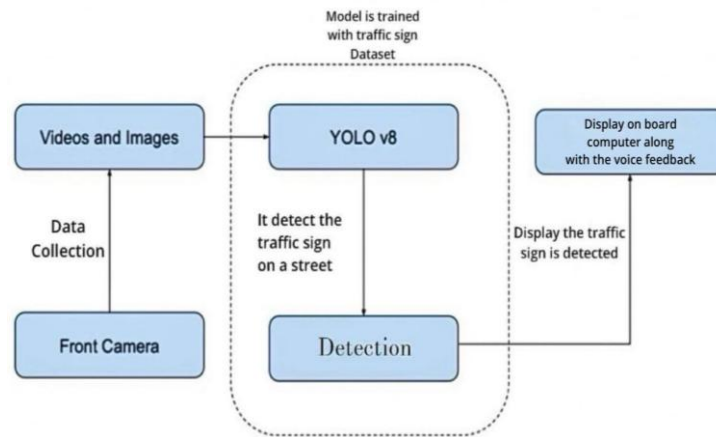
School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology (VIT), Vellore, 632 014, India
E-mail: rahul.b2021d@vitstudent.ac.in

Received: 24 November 2024; Revised: 23 February 2025; Accepted: 03 March 2025; Published: 08 June 2025

Abstract: With the proliferation of advanced driver assistance systems and continued advances in autonomous vehicle technology, there is a need for accurate, real-time methods of identifying and interpreting traffic signs. The importance of traffic sign detection can't be overstated, as it plays a pivotal role in improving road safety and traffic management. This proposed work suggests a unique real-time traffic sign detection and recognition approach using the YOLOv8 algorithm. Utilizing the integrated webcams of personal computers and laptops, we capture live traffic scenes and train our model using a meticulously curated dataset from Roboflow. Through extensive training, our YOLOv8 version achieves an excellent accuracy rate of 94% compared to YOLOv7 at 90.1% and YOLOv5 at 81.3%, ensuring reliable detection and recognition across various environmental conditions. Additionally, this proposed work introduces an auditory alert feature that notifies the driver with a voice alert upon detecting traffic signs, enhancing driver awareness and safety. Through rigorous experimentation and evaluation, we validate the effectiveness of our approach, highlighting the importance of utilizing available hardware resources to deploy traffic sign detection systems with minimal infrastructure requirements. Our findings underscore the robustness of YOLOv8 in handling challenging traffic sign recognition tasks, paving the way for widespread adoption of intelligent transportation technologies and fostering the introduction of safer and more efficient road networks. In this paper, we compare the unique model of YOLO with YOLOv5, YOLOv7, and YOLOv8, and find that YOLOv8 outperforms its predecessors, YOLOv7 and YOLOv5, in traffic sign detection with an excellent overall mean average precision of 0.945. Notably, it demonstrates advanced precision and recall, especially in essential sign classes like "No overtaking" and "Stop," making it the favored preference for accurate and dependable traffic sign detection tasks.

Index Terms: Traffic Sign Detection, YOLOv8, Road Safety, Intelligent Transportation, Deep Learning, Object Recognition, Driver Assistance.

Graphical Abstract. Traffic sign recognition flows using YOLOv8



Highlights

- The process begins by collecting videos and images, which serve as a base dataset to train the traffic sign detection system.
- The YOLOv8 Traffic Sign Recognition (TSR-YOLOv8) detector, a real-time object detection system, is then used to evaluate the gathered data.
- This detector carefully scans video and images to accurately determine the presence of traffic signs.
- Once the TSR-YOLOv8 detector successfully recognizes the traffic sign, the information is quickly relayed and can be displayed on a PC or laptop screen for further analysis or reference.
- Our system incorporates an auditory alert mechanism that provides immediate voice feedback whenever a traffic sign is detected.
- This feature has been rigorously tested and works seamlessly, ensuring that drivers receive prompt and clear notifications, thereby enhancing overall driving safety and awareness.

1. Introduction

In today's rapidly evolving technological landscape, traffic sign detection and recognition is emerging as an important field for ensuring road safety and effective traffic management [1]. Real-time is becoming more and more urgent. This paper presents a new approach to solving complex challenges in traffic sign detection and recognition by leveraging the capabilities of the YOLOv8 algorithm [2]. YOLOv8, an evolution of the widely acclaimed YOLO architecture, provides a powerful solution for object detection tasks characterized by improved speed and accuracy. Our method revolves around a comprehensive, multi-step process in which the YOLOv8 algorithm undergoes rigorous training and tuning to recognize a variety of traffic signs. With the immense capabilities of deep learning and convolutional neural networks (CNN), our system learns to detect and locate traffic signs skillfully with precision and accuracy, facilitating fast and reliable recognition even in dynamic traffic situations. Additionally, the inherent efficiency of the YOLOv8 algorithm enables transparent real-time processing, making it ideal for deployment in resource-constrained environments, including embedded systems and attached computers. Our experimental efforts validate the effectiveness and reliability of the traffic sign detection and recognition system in a variety of real-world scenarios. By providing a reliable means of interpreting road signs, our solution contributes significantly to improving road safety, optimizing traffic flow, and promoting progress. Through this paper, we further study the application of YOLOv8 in traffic sign detection and recognition, elucidating its advantages in terms of training efficiency, accuracy, and strong performance in challenging real-world conditions.

In the field of modern transport and urban infrastructure, the prevalence of traffic signs is an essential element to ensure road safety and efficient traffic flow. However, with the increasing complexity of road networks and the advent of advanced driver assistance systems (ADAS), accurate detection and recognition of traffic signs have become essential. Traditional manual detection and interpretation methods are no longer sufficient to meet the needs of modern traffic management systems, which require the integration of advanced computer vision techniques.

Despite advances in computer vision research, current traffic sign detection systems still face many challenges. These challenges arise from variations in traffic sign design, environmental conditions[3], and congestion encountered in real-life situations. Conventional approaches often have difficulty detecting signs reliably in adverse weather conditions, variable lighting[4], and obstructions caused by foliage, other vehicles, other factors, or other infrastructure factors [5]. Additionally, the dynamic nature of traffic scenes, characterized by rapid vehicle movements and changing perspectives, further complicates the detection process.

Given the important role of accurate traffic sign detection in ensuring road safety and supporting intelligent transportation systems, a specialized approach is needed. The limitations of current methods highlight the need for a

detection system that is not only capable of accurately identifying traffic signs but is also resilient to challenges posed by real-world conditions. A tailor-made solution leveraging advanced machine learning techniques, such as the YOLOv8 algorithm, offers a promising path to address these challenges effectively.

The application of the YOLOv8 algorithm provides an attractive solution to the complexity of traffic sign detection. By leveraging real-time object detection [6], YOLOv8 provides a powerful framework for accurately locating and classifying traffic signs in dynamic traffic scenes. Unlike traditional methods that rely on handcrafted features and complex preprocessing, YOLOv8's end-to-end learning approach allows it to learn discriminative features directly from raw image data, thereby improving its ability to adapt to different traffic sign environments.

2. Literature Survey

Traffic sign recognition and detection [7] were two of the applications of high interest in computer vision due to their relevance in ITS and autonomous driving. Their accuracy and speed directly relate to safety and efficiency in such systems. With time, some methodologies have been explored to improve traffic sign detection. To start with, traditional techniques were adopted first while deep learning techniques were introduced lately. This review of the literature summarizes efforts realized in traffic sign detection with a special focus on YOLO models that culminated in the more recent development of YOLOv8.

Traffic sign detection has traditionally depended extensively on old feature-based methods of color and shape, primarily based on traditional hand-crafted approaches that were satisfactory in less complex environments but abysmal in dynamic scenarios with changing lighting, occlusions, and scaling issues [8]. Approaches such as PHOG and SVMs are mostly partially robust but, even so, unscalable for real-time operation and scalability in practice [9]. For instance, the initial models could not effectively spot traffic signs as they could not handle a big scale and diversity of data [10]. Deep learning methods have allowed huge steps forward in terms of detection speed and accuracy with traditional methods. The Convolutional Neural Network (CNN) was the layer that enabled models to start drawing features from raw images autonomously without predefined feature engineering. Considering strengths acquired through CNNs, models like Faster R-CNN and Mask R-CNN can represent a solution but step into problems faced by the complexity of real-time detection [11].

YOLO brought the traffic sign detection one step forward. Again, YOLOv1, introduced something novel by combining both detection and classification into one single neural network architecture that improves the speed of the detection significantly. However, the early versions, YOLOv1 and YOLOv2 detected objects too fast but could not detect the smaller ones with accuracy held unanimously. The problems mentioned above prompted the improvement of the architecture by YOLOv3. Specifically, the feature pyramid network improved multi-scale object detection, and this enabled the recognition of traffic signs smaller in size to be more accurate [12]. This version, which balances speed and accuracy, is still somewhat crude to tackle challenging environments like poor lighting or occluded traffic signs.

When further research followed, models that aggregated features from several scales were introduced to overcome the challenges of small detection in traffic sign detection. This one included the Feature Aggregation Multipath Network model, which advanced the small detection of traffic signs by feature aggregation. FAMN employed part-context information. This is because fine-grained classification of traffic signs entailed using part-context information. Thus, by the use of this model, it attained an F1-measure score of 93.1% on the TT100K dataset. This model had quite an impressive ability to detect small signs that most versions of YOLO usually miss before. Another significant success is the Enhanced SSD, with the optimization of feature extraction focused on defeating similar challenges. This model characterized the expression of traffic signs better than what exists in the real-time processing and precision balance within the dataset CCTSDB; this model has an average precision of 90.52%. It outperformed the older models, namely YOLOv2 and Faster R-CNN, in [13].

With this newest member in the YOLO family, YOLOv8 jumps ahead in terms of speed optimization and detection efficiency. Along with the architectural fine-tuning on YOLOv8, small object detection is still one of the challenging problems in TSRD. Based on this mechanism, YOLOv8 proposed modification changes to the previous module known as the C2f module toward further extracting more proper features and preventing overfitting, especially in complicated traffic environment conditions refer [14-15]. He et al. demonstrated that the lightened version of YOLOv8, optimized for traffic sign recognition, was faster in detection and even significantly more accurate compared with its predecessors. By checking it on TT100K and COCO datasets, they demonstrated that the optimized model YOLOv8-CO rendered better performance than others concerning speed and precision as well. All the improvements in YOLOv8 are the conditions to consider this model for real-time TSRD use in self-driving cars, where the speed of decision-making and accuracy matter the most.

In the real world, the deployment of TSRD systems must make models highly efficient and fit on very resource-constrained devices such as embedded systems. The variants YOLOv4 and YOLOv5 were released, with the even smaller and lighter version being Tiny YOLOv4 [16], which was designed especially for such environments. Khnissi et al. [15] optimized the model Tiny YOLOv4 for one of the most popular embedded devices, NVIDIA Jetson Nano, realizing a mean Average Precision of 95.44%. These advantages prove that high-detection accuracy can be gained even for highly compact models to be deployed in real-time autonomous driving applications. Also, Castruita Rodríguez et al. used YOLOv3 in region-specific traffic sign detection in Mexico; it can be said that this model is versatile enough with different

datasets[17]. Regional adaptation for TSRD models becomes essential since every country and region has a different set of traffic signs, thus requiring context-specific models for every situation.

Although YOLO is still a little ahead of TSRD, other models like Enhanced SSD have grown to make the concept better. Sun et al. proposed a model applying lightweight convolution operations to further boost characteristic expression in traffic sign[18]. In this model, the detection rate on small objects improves, and an mAP of 81.26% is attained on TT100K. Besides, based on innovation, several models were integrated. For instance, Wu et al. proposed the fusion model that collaborated YOLOv3 with ShuffleNetV2 to promote recognition accuracy in which the relatively high computational cost is very low. The integrating architectures eliminated some disadvantages of the single model and achieved excellent performance in real-world applications.

The authors[19] discussed the multi-frame detection techniques that may enhance the accuracy of video-log traffic sign detection. Multi-frame techniques decode temporal information from multiple frames, where the detection model extracts consistent traffic signs following sequences of images. Multi-frame techniques would shake up TSRD by improving detection robustness with motion blur and even with occlusions, which would degrade detection performance due to its erroneous localization in those regions affected by motion blur or full occlusion. Another major one in the future of TSRD is applying data augmentation techniques. This work [20] applied methods such as adding noise, rotation, and translation to enhance the robustness of models with minimized overfitting by expanding their training datasets. Such approaches could be used with much more advanced models like YOLOv8 so that TSRD can further raise its precision under different environmental conditions.

Traffic sign detection and recognition [21] have gone great these days and are mainly driven by the deep learning approach and enhancements in the YOLO model. YOLOv8 is the pinnacle achieved so far; it contains attention mechanisms as well as multi-scale feature extraction, besides having in real time the capacity to process such information. It achieves an optimal balance between detection speed and accuracy, and it plays a big part in lots of autonomous driving systems. For example, more powerful models of this class will inherently lead to better advanced multi-frame detection techniques and data augmentation schemes in adverse weather conditions. The development of TSRD will be followed by the improvement of intelligent transportation systems on safer routes and efficient paths.

Research Questions

Q1. What are the primary uses of YOLO for detecting and identifying traffic signs? Traffic sign detection and recognition with YOLO offers streamlined deployment thanks to real-time processing and high accuracy. Its efficient architecture enables rapid deployment in a variety of environments, making it ideal for integration into existing transportation infrastructure and vehicle systems. This ease of implementation ensures rapid adoption and scalability, contributing to safer roads and better traffic management.

Q2. What metrics are employed to evaluate the effectiveness of object detection, specifically within the domain of traffic sign identification using YOLO?

In the context of traffic sign detection and recognition using YOLO, the evaluation of object detection performance is based on several key metrics. Precision, recall, and F1 score are fundamental metrics that evaluate an algorithm's ability to accurately identify traffic signs while minimizing false positives and negatives. Furthermore, mean precision (mAP) provides a comprehensive assessment by considering precision-recall curves across different confidence thresholds. Additionally, intersection over union (IoU) is an important measure of bounding box accuracy, quantifying the spatial association between predicted and ground truth boxes. Together, these metrics provide a robust framework to evaluate the quality and effectiveness of YOLO-based traffic sign detection algorithms.

Q3. What hardware configurations are typically utilized for deploying traffic sign detection and recognition systems based on YOLO?

In our implementation of traffic sign recognition and detection systems based on YOLO, we integrated an external webcam along with the requisite hardware configurations. This setup enabled real-time data acquisition for processing by the YOLO-based algorithms. Additionally, computational hardware such as GPUs was employed to facilitate efficient inference, ensuring timely detection and recognition of traffic signs. This combination of hardware components optimized system performance, enabling seamless integration and effective deployment of the YOLO-based solution for traffic sign recognition and detection.

Q4. What obstacles and difficulties arise in the process of detecting and recognizing traffic signs when employing YOLO?

We faced various challenges during traffic sign detection and recognition using YOLO, including issues with varying lighting conditions, occlusions, and complex backgrounds. Additionally, accurately detecting small or distorted signs posed significant obstacles. Overcoming these challenges required continuous refinement of our YOLO models and training datasets to ensure robust performance across diverse environments.

Our project seeks to address these challenges through the use of advanced machine learning techniques, specifically leveraging the YOLOv8 algorithm for traffic sign detection and recognition. By leveraging the capabilities of YOLOv8, we aim to develop a system that not only accurately detects and locates traffic signs but also adapts seamlessly to the dynamic and diverse conditions encountered on the road. Furthermore, the need to improve traffic sign detection and recognition systems goes beyond simple technological advances. This has a direct impact on road safety, traffic optimization, and the broader intelligent transport system landscape. By improving the accuracy and efficiency of traffic sign detection, our project contributes to the overall goals of improving road safety, reducing accidents, and facilitating

traffic flow. Fundamentally, the motivation for our project lies in meeting the urgent need for robust, adaptable, and efficient solutions to improve traffic sign detection and recognition. By leveraging advanced machine learning techniques and the capabilities of the YOLOv8 algorithm, we aim to advance the field of traffic sign detection, thereby contributing to safer and more efficient transportation systems.

3. Materials and Methods

3.1. YOLOv8

The YOLOv8 algorithm [22] represents a significant advancement in deep learning models specifically designed for real-time object detection tasks. Unlike traditional methods that rely on complex and multiple pipelines, YOLO [23] uses a unique approach by treating object detection as a single regression problem. It achieves this by dividing the input image into a grid and predicting bounding boxes and class probabilities for the objects in each grid cell. This simultaneous prediction of class labels and bounding boxes contributes to its efficiency and speed, making it well-suited for real-time image and video processing [24]. One of the key features of YOLOv8 is the ability to perform object detection in one neural network pass, following the YOLO principle. This approach ensures real-time performance, making it particularly suitable for applications that require rapid processing of image data, such as autonomous driving and video surveillance. Additionally, YOLOv8 achieves outstanding object detection accuracy across a wide range of object types and environmental conditions. Its robustness to changes in scale, direction, and occlusion makes it suitable for a variety of real-world scenarios.

YOLOv8 is designed with efficiency in mind, optimizing the balance between accuracy and computational resources. By leveraging techniques such as feature pyramid networks (FPNs), spatial attention mechanisms, and advanced anchor box prioritization, YOLOv8 minimizes computational cost while maintaining broadcast quality. Additionally, its modular architecture facilitates easy adaptation to specific use cases and datasets, allowing researchers and developers to fine-tune models, incorporate domain-specific knowledge, or expand its capabilities through transfer learning.

Additionally, YOLOv8 is typically implemented as an open-source software framework, providing access to pre-trained weights, training scripts, and evaluation tools. This accessibility allows the research community and practitioners to use the model for different projects, thereby simplifying the implementation and testing process. Overall, YOLOv8 presents itself as a powerful and flexible tool for real-time object detection, offering high accuracy, efficiency, adaptability, and open-source availability.

Figure 1 depicts the YOLOv8 structure. Here, the Convolutional Block Attention Module (CBAM) is a significant advancement in feature extraction within neural network backbones. It introduces a dynamic weighting mechanism that prioritizes informative regions while suppressing less relevant areas, allowing the network to prioritize details for object detection while filtering out background noise and distractions. This enhances the network's capability to accurately detect objects within images, contributing to improved object detection accuracy and robustness across various scenarios and datasets. The neck component of the neural network architecture handles scale variations in object detection tasks, particularly in aerial images captured by webcams. The Spatial Pyramid Pooling (SPP) layer captures features at multiple scales, while convolutional layers further process feature maps, preparing them for object detection tasks in the Head component. The CBS and detection modules form a cohesive framework in the head, delivering accurate and comprehensive object detection across diverse datasets and real-world scenarios.

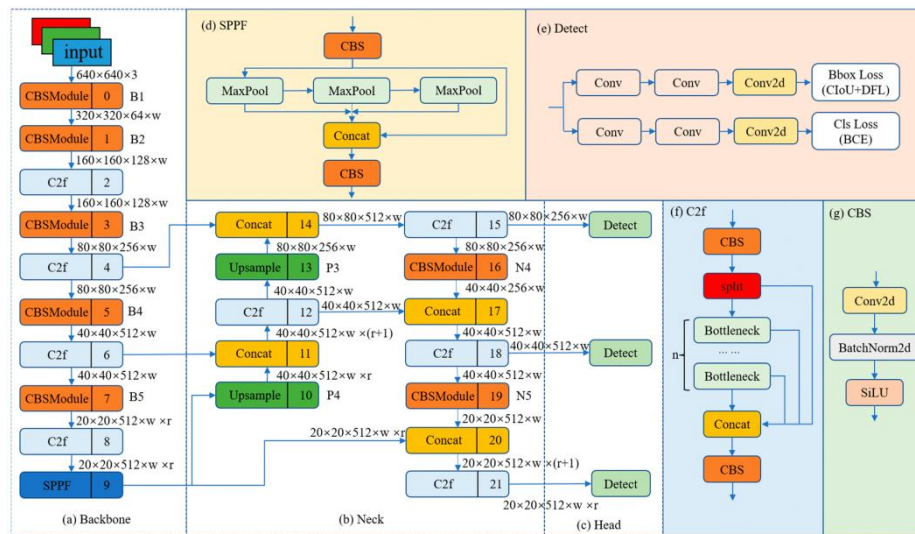


Fig.1. Diagram of YOLOv8 structure

3.2. Proposed System Algorithm

Figure 2 shows the entire process using YOLOv8, specifically designed for real-time object detection tasks in convolutional neural networks (CNN). Starting with data collection, which may involve capturing images via a camera[25], the next step includes preprocessing and data enhancement. Preprocessing includes standardizing raw images for training by adjusting pixel sizes and values, while augmentation expands the dataset through variations such as cropping and flipping to improve the ability to generalize the model. The dataset is then divided into training and testing image sets, with the first set (80%) dedicated to model training and the second set (20%) used to evaluate performance. capacity. The training phase includes preprocessing the images before feeding them into the YOLOv8 algorithm. The algorithm's convolutional layers extract features, which aid in object detection by predicting bounding boxes and class labels.

During training, the model's weights are iteratively adjusted based on comparing the predictions with the ground truth, gradually fine-tuning the detection accuracy. After training, the model moves on to prediction, applying learned knowledge to new images to create bounding boxes and class labels for detected objects. Finally, the output shows the image with bounding boxes with corresponding annotations and class labels, indicating successful object detection.

3.3. Module Description

Our traffic sign detection and recognition assignment leverage YOLOv8, a modern-day item detection system famed for its first-rate overall performance in real-time tasks. YOLOv8 employs an unmarried convolutional neural network (CNN) to concurrently locate and classify visitors' symptoms and symptoms in images. This structure stands proud for its potential to unexpectedly technique images, making it perfect for packages requiring speedy inference. Within YOLOv8, convolutional layers function as the spine for detecting complicated functions of visitors' symptoms, together with shapes, colors, and symbols. These layers are meticulously designed to seize each low-stage detail, like edges and textures, and high-stage functions, like specific signal shapes and symbols, permitting specific recognition.

Image acquisition is step one in the detection and popularity process of site visitor signs. It includes putting cameras in strategic places to capture images or video. These cameras may be established on vehicles, located on-site visitor's poles, or located at intersections. The quality of the images captured, which includes their decision, the lighting fixture conditions, and the camera angle, has a great effect on the effectiveness of the following processing stages. High-decision images with uniform lighting fixtures are preferred for detection and popularity.

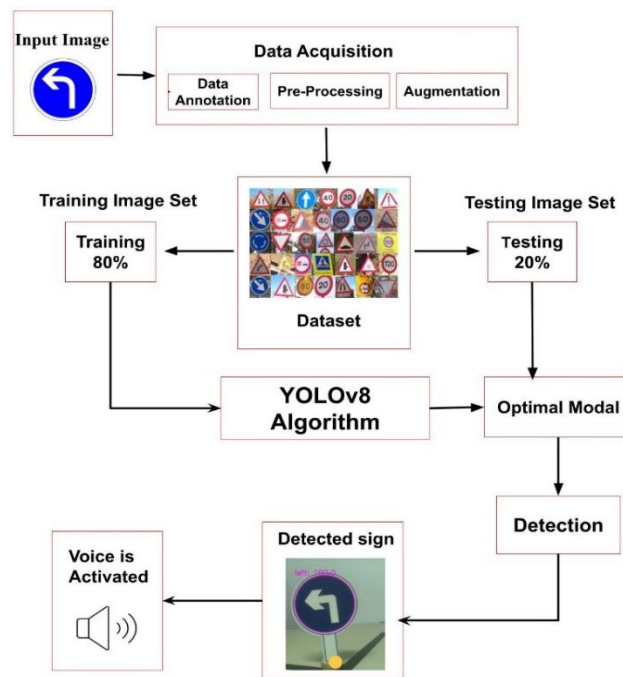


Fig.2. Conceptual design of the model

The preliminary step entails the purchase of a variety of datasets comprising site visitors' signal pictures encountered in real-world international street scenarios. This dataset encompasses numerous kinds of site visitor signs, spanning regulatory, warning, and informational signs. Additionally, to simulate real international conditions, the dataset consists of pictures captured beneath exclusive lighting fixture conditions, climate variations, and various site visitor scenarios.

The dataset received from RoboFlow might also have additionally passed through preprocessing throughout its creation or augmentation. This preprocessing ought to consist of standardizing image sizes, adjusting color areas, removing noise, or augmenting records to grow dataset diversity. Standardizing picture sizes guarantees consistency throughout the dataset, facilitating uniform entry for the detection algorithm. Adjusting color areas might also enhance

the model's ability to generalize to one-of-a-kind lighting fixture conditions. Noise discount strategies might also be implemented to enhance the quality of images, decreasing the possibility of fake positives throughout detection. The traffic sign images might be flipped in the horizontal or vertical axis, and it is identified with the probability between 0 and 1. Similarly, it may be in the angle of rotation -45 to 45 degrees. Augmentation strategies, which include rotation, flipping, or noise, might have been used to grow dataset length and diversity, thereby enhancing the model's robustness. While express preprocessing steps won't be required at this stage, knowing how the preprocessing is implemented in the dataset is important for decoding the model's overall performance and optimizing similar processing steps.

The annotated dataset is partitioned into awesome subsets for education, validation, and checking-out purposes. This partitioning allows a rigorous assessment of the model's overall performance on unseen facts, stopping over-fitting and ensuring its generalizability to real international site visitors' signal detection scenarios. 80% of the facts are used for education, and 20% of the facts are used for testing processing steps. Leveraging the YOLOv8 architecture, we embark on the education section to develop a deep neural community version tailor-made for visitors' signal detection and recognition. Through iterations of gradient descent and back-propagation, the version learns to determine salient functions and spatial preparations of visitors' symptoms and symptoms inside images.

As a part of the optimization process, we refine the YOLOv8 structure to enhance its detection capabilities, specifically regarding small site visitor signs. Modifications to the community shape the intention to lessen computational overhead while at the same time bolstering real-time performance, ensuring speedy and correct detection. The educated version undergoes rigorous assessment at the validation set to gauge its overall performance across numerous metrics, including accuracy, precision, recall, and implying common precision. This assessment section serves to validate the version's efficacy and robustness for visitors' signal detection tasks.

Fine-tuning of the skilled version entails using strategies consisting of switch getting to know or multi-challenge getting to know to similarly refine its generalization ability. By adapting the version to numerous site visitors' situations and climate conditions, we enhance its adaptability and reliability in real-world international deployments. Subsequently, the finalized version undergoes complete checking out of the unseen dataset to validate its real-international overall performance. Through a meticulous evaluation of key overall performance metrics, we verify the version's effectiveness and reliability in visitor signal detection and reputation tasks.

Upon successful validation, the optimized version is deployed on an appropriate computing platform capable of real-time processing. This deployment enables seamless integration into smart transportation systems, allowing on-road traffic sign detection and recognition for stronger safety and traffic management. The deployed version undergoes considerable overall performance assessment in real-global scenarios, assessing its efficacy throughout various environmental conditions. Through meticulous dimensions of accuracy, precision, recall, and suggestion of common precision, we confirm the version's adherence to overall performance requirements and its fantastic contribution to road protection and traffic efficiency.

3.4. Layer Description

The proposed work was implemented with the configuration of the system environment pack of the Intel Pentium i5 processor, 2.9GHz or above clock speed, 8 GB RAM, 40GB storage, and NVIDIA GTX 1050 or above. The high GPU and NVIDIA card reduce the latency with low-level memory consumption. Table 1 tabulates the layers' description of YOLOv8. The backbone is the preliminary degree to which the community is accountable for extracting capabilities from the input image. YOLOv8 leverages a modified model of the CSPDarknet53 backbone, which is CNN architecture. This is a 53-layer convolutional neural network that has been the basis for many item detection models. It makes use of residual connections, which permit a higher gradient glide throughout training and allow the network to learn complex features.

The Cross-Stage Partial Connections (CSP) approach improves statistical flow inside the network by combining features from one-of-a-kind stages. It creates shortcuts that permit advanced layers to persuade later layers, leading to more robust feature extraction. The neck-level tactic involves the feature maps extracted from the backbone. In YOLOv8, the neck makes use of a unique module referred to as the C2f module as opposed to the traditional YOLO neck architecture.

The C2f module combines the following functionalities: Conv modules are groups of convolutional layers that, in addition, refine their capabilities. Bottleneck CSP blocks are green construction blocks that extract high-degree features while decreasing computational costs. The Spatial Pyramid Pooling (SPPF) layer effectively captures capabilities at one-of-a-kind scales in the image. It enables the network to discover objects of various sizes.

The head, at the very last level, is charged with making object detection predictions. Up sample layers boost the resolution of the feature maps, anticipating the higher localization of smaller objects. There are 5 detection modules in YOLOv8, each liable for a particular feature scale. These modules use convolutional layers and different layers, like activation capabilities, to make predictions. Bounding boxes outline the location and length of the detected objects inside the image. Objectless scores suggest the self-belief stage in which an item is found in a selected bounding box. Class probabilities suggest the chance of every class (e.g., car, person) being present inside a bounding box.

3.5. Dataset

The dataset incorporates 2467 images throughout 18 traffic sign instructions, promising improvements in traffic signal detection algorithms. The dataset was collected from universe.roboflow.com. To ensure powerful version training, the dataset has passed through meticulous pre-processing, along with auto-orientation and resizing of images to a uniform size of 640x640 pixels. Additionally, augmentations together with bounding container flipping, rotation, shearing, blur,

and noise were carried out to enhance the dataset's range and simulate real-world scenarios.

Table 1. Layer description of YOLOv8

Layer	Type	Features Map	K / S / P
1	Conv	3 -> 16	3 x 3 / 2 / 1
2	Conv	16 -> 32	3 x 3 / 2 / 1
3	C2f (Bottleneck)	32 -> 32	Various
4	Conv	32 -> 64	3 x 3 / 2 / 1
5	C2f (Bottleneck)	64 -> 64	Various
6	Conv	64 -> 128	3 x 3 / 2 / 1
7	C2f (Bottleneck)	128 -> 128	Various
8	Conv	128 -> 256	3 x 3 / 2 / 1
9	C2f (Bottleneck)	256 -> 128	Various
10	SPPF	256 -> 128	Various
11	Upsample	-	2.0 Scale
12	Concat	-	-
13	C2f (Bottleneck)	384 -> 128	Various
14	Upsample	-	2.0 Scale
15	Concat	-	-
16	C2f(Bottleneck)	192 -> 64	Various
17	Conv	64 -> 64	3 x 3 / 2 / 1
18	Concat	-	-
19	C2f (Bottleneck)	192 -> 128	Various
20	Detect	Various	Various

With a cut-up ratio of 80% for training and 20% for testing, the dataset is well-balanced, with 2117 snapshots for education and 530 for testing. These pre-processed and augmented images function as inspiration for training the YOLOv8 version. Through iterative education, the version adjusts its parameters to the use of backpropagation, aiming to minimize detection mistakes and optimize performance.

Evaluation metrics, together with precision, recall, and suggest common precision (mAP), are used to evaluate the version's efficacy in detecting traffic signs across diverse instructions. By leveraging this complete dataset and superior education techniques, researchers and builders can power innovation in site visitors' signal detection algorithms, in the long run contributing to safer and more environmentally friendly transportation systems. Examples of traffic signal classes inside the dataset include "No Entry," "No Left Turn," and "No Stopping."

3.6. Performance Metrics

Precision and recall are fundamental metrics in the evaluation of classification models, providing insights into their performance. Accuracy represents the proportion of correctly classified instances among the total instances. It is calculated by dividing the number of correctly predicted instances by the total number of instances. While accuracy gives an overall view of model performance, recall, also known as sensitivity or true positive rate, measures the model's ability to correctly identify all relevant instances, regardless of misclassifications. It is calculated by dividing the number of true positives by the sum of true positives and false negatives. High accuracy indicates a low rate of misclassifications, while high recall suggests that the model effectively captures most positive instances, which is crucial in scenarios where correctly identifying all relevant instances is paramount, such as medical diagnostics [26] or anomaly detection. In this proposed traffic sign detection system, true positives refer to the correctly detected traffic signs, false positives mean incorrectly detected traffic signs, and false negatives denote the traffic signs that the model missed. The mathematical expressions for precision and recall are detailed in Equations (1) and (2), respectively.

$$\text{Precision (YOLOv8)} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Positives}} \quad (1)$$

$$\text{Recall (YOLOv8)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

In traffic sign detection, where precision and recall play vital roles, the F1 score emerges as a crucial metric for assessing model performance. The F1 score balances precision and recall, offering a single metric that encapsulates both aspects of model accuracy. Given the importance of correctly identifying traffic signs for ensuring road safety, achieving high precision and recall simultaneously is imperative. The F1 score provides a harmonic mean of precision and recall, emphasizing the necessity of minimizing both false positives and false negatives. This is particularly critical in traffic

sign detection systems, where misclassifications can lead to potentially hazardous situations on the road. Therefore, a high F1 score indicates a robust and reliable model capable of accurately detecting traffic signs while minimizing errors, thus contributing significantly to overall road safety and driver assistance systems. The mathematical expressions for precision and recall are detailed in Equations (3)

$$F1 (YOLOv8) = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

Mean Average Precision (mAP) stands as a cornerstone metric in object detection tasks, including traffic sign detection, due to its comprehensive evaluation of model performance. Unlike single-point metrics like accuracy or F1 score, mAP considers the precision-recall trade-off across various confidence thresholds, providing a holistic assessment of detection accuracy. In traffic sign detection, where ensuring precise localization and classification of signs is paramount for road safety, mAP offers a nuanced understanding of model capabilities by considering both the accuracy of detected signs and their spatial localization. This is particularly crucial in scenarios where traffic signs may appear at varying distances, angles, or lighting conditions. By aggregating precision values over a range of recall levels, mAP captures the robustness of a detection model across diverse real-world scenarios, making it an indispensable metric for assessing and benchmarking the performance of traffic sign detection systems. Achieving high mAP signifies a model's capability to detect traffic signs across different environmental conditions consistently and accurately, thereby enhancing driver safety and overall system reliability. The mathematical expressions for precision and recall are detailed in Equations (4)

$$mAP (YOLOv8) = \frac{1}{N} \sum_{c=1}^N AP_c \quad (4)$$

3.7. Result and Findings

In this proposed work, we examine the effectiveness and results of using YOLO-based methods for traffic sign detection. Our research includes an in-depth investigation of the different techniques and approaches used in the field, aiming to provide valuable insights into the strengths, limitations, and potential advances of YOLO-based systems. Through comprehensive analysis and evaluation of our results, we aim to contribute to the ongoing discussion around optimizing and improving traffic sign detection algorithms.

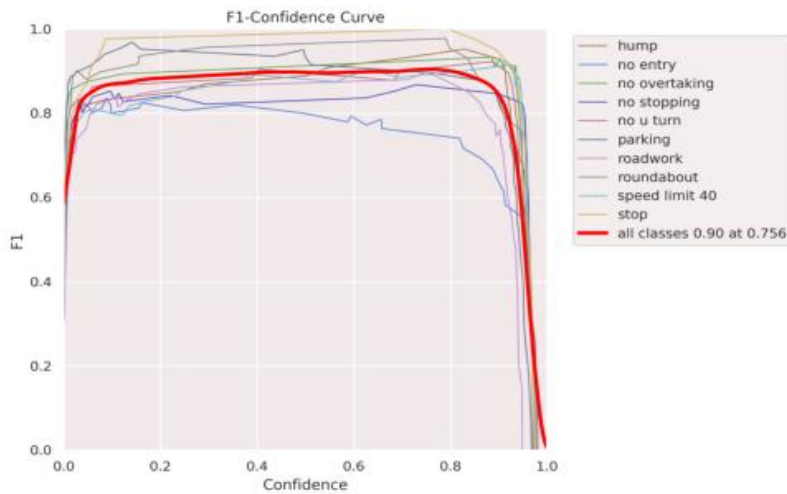


Fig.3. TSDR F1 - confidence curve

Figure 3 illustrates the outcomes of a YOLOv8-trained model applied to a traffic sign dataset, offering insights into driver behavior and confidence levels across diverse traffic scenarios. The graph reveals intriguing patterns by representing driver confidence levels on the x-axis, ranging from complete uncertainty (0) to unwavering certainty (1), and the FT-confidence curve on the y-axis, which depicts drivers' ability to control a vehicle. The curve's characteristic 'hump' shape suggests that drivers with moderate skills may exhibit overconfidence, particularly at intermediate ability levels. Annotations along the x-axis provide context for specific traffic scenarios, with labels like 'No Entry' at lower confidence levels indicating reluctance to execute maneuvers, while 'Speed Limit 40' at higher confidence levels implies greater adherence to regulations. Mid-range confidence levels, marked by annotations such as 'Roadwork' and 'Roundabout,' suggest potential challenges for moderately confident drivers due to increased complexity. Notably, the inclusion of 'All Classes 0.90 at 0.756' denotes proficient system performance, achieving a 90% confidence level across all driver classes at a specific threshold. This underscores the YOLOv8 model's effectiveness in accurately detecting traffic signs and analyzing driver behavior, thus offering valuable insights for road safety assessments.

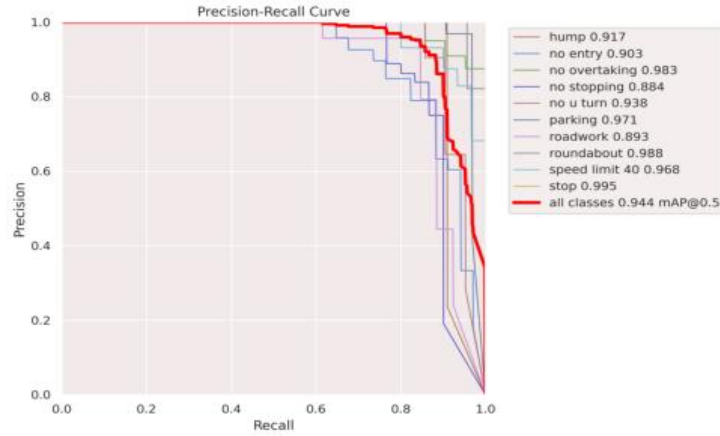


Fig.4. TSDR precision-recall curve

Figure 4 shows the measures of the ratio of correctly predicted positive samples to the total predicted positive samples. Recall measures the ratio of correctly predicted positive samples to the total actual positive samples. These metrics provide insights into the trade-off between false positives and false negatives.

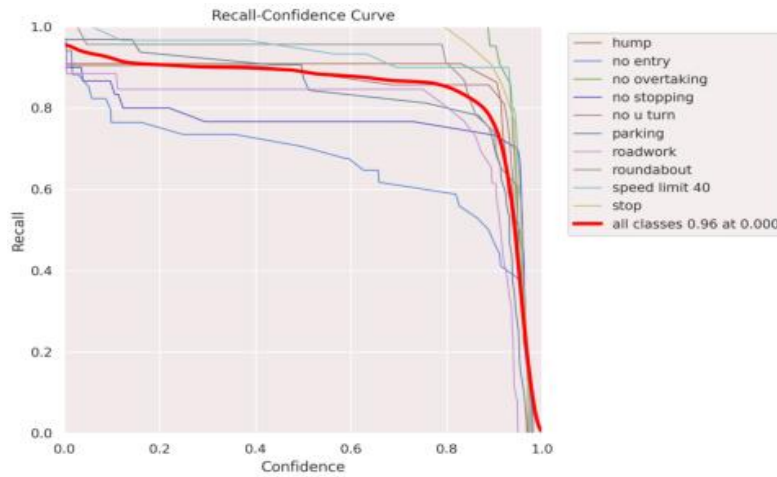


Fig.5. TSDR recall – confidence curve

Figure 5 shows how recall varies with changes in the confidence threshold. Recall represents the proportion of correctly predicted positive samples (true positives - correctly detected traffic signs) among all actual positive samples. It conferring the result of the proposed system can detect most of the relevant traffic sign images in the dataset with some errors.

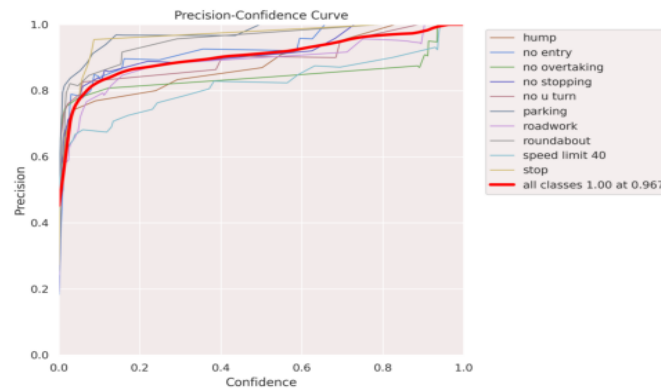


Fig.6. TSDR precision – confidence curve

Figure 6 illustrates how precision varies with changes in the confidence threshold. Precision represents the proportion of correctly predicted positive samples (true positives) among all predicted positive samples. According to this analysis, this proposed traffic sign detection system identifies most of the traffic signs correctly with fewer mistakes.

Comparing Metrics among Different Versions of YOLO

After thoroughly evaluating traffic sign detection models YOLOv8, YOLOv7, and YOLOv5, it is evident that YOLOv8 consistently outperforms its predecessors across various classes of traffic signs. Table 2 shows that the overall mAP for YOLOv8 stands at 0.926, showcasing its superior accuracy compared to YOLOv7 (mAP = 0.901) and YOLOv5 (mAP = 0.813). Specifically, YOLOv8 exhibits remarkable precision and recall for most classes, with a standout performance in critical signs like "No overtaking" and "Stop," achieving near-perfect precision and recall scores. While YOLOv7 and YOLOv5 demonstrate competitive results in certain classes, they consistently fall short in overall accuracy compared to YOLOv8. Therefore, based on these findings, YOLOv8 emerged as the preferred choice for traffic sign detection tasks, offering superior performance and reliability across a diverse range of traffic sign classes. Additionally, our system incorporates an auditory alert mechanism that provides immediate voice feedback whenever a traffic sign is detected. It can be integrated with a semi-autonomous or autonomous system, sounded through the speaker. This feature has been rigorously tested and works seamlessly, ensuring that drivers receive prompt and clear notifications, thereby enhancing overall driving safety and awareness.

Traffic sign detection and recognition using the YOLOv8 algorithm has indeed showcased the promising potential for enhancing road safety and improving intelligent transportation systems. Looking ahead, this technology has numerous avenues for growth and development. As autonomous vehicles become more prevalent, the integration of YOLOv8 into their systems can significantly enhance their perception capabilities, ensuring safe navigation through accurate traffic sign detection and recognition. The proposed system works well in the vehicles equipped with high-resolution and weather-resistant cameras that capture the traffic sign images at a high frame rate with clarity. Even with the pictures blurred or distorted, the image enhancement algorithm committed to resolving it. Optimizing the YOLOv8 algorithm for enhanced efficiency and processing speed will enable real-time detection and recognition of traffic signs, even in high-speed scenarios. Integrating YOLOv8 into smart city infrastructure, including traffic cameras and intelligent traffic systems, can enable real-time monitoring and recognition of traffic signs, leading to more efficient traffic control and optimization. Further research can focus on improving the accuracy of YOLOv8 in detecting and recognizing traffic signs in challenging scenarios, such as occlusions, low-lighting conditions, and dynamic traffic situations.

Table 2. Comparison table of different versions of YOLO

Model	YOLOv8			YOLOv7			YOLOv5		
Class	mAP	precision	Recall	mAP	precision	Recall	mAP	Precision	Recall
All	0.926	0.955	0.888	0.901	0.907	0.818	0.813	0.946	0.894
Hump	0.863	1	0.819	0.904	0.894	0.769	0.82	0.981	0.818
No entry	0.749	0.875	0.697	0.697	1	0.506	0.726	0.962	0.751
No-overtaking	0.995	0.958	1	0.995	0.938	1	0.894	0.899	1
No stopping	0.845	0.932	0.833	0.859	0.921	0.776	0.766	0.951	0.8
No u turn	0.962	0.89	0.8	0.877	0.809	0.905	0.861	0.901	0.952
Parking	0.913	0.996	0.9	0.963	0.887	0.969	0.837	1	0.979
Roadwork	0.995	0.964	1	0.83	0.857	0.654	0.655	0.914	0.815
Roundabout	0.995	1	0.958	0.985	0.968	0.87	0.803	0.968	0.957
Speedlimit 40	0.949	0.969	0.875	0.944	0.893	0.833	0.887	0.941	0.867
Stop	0.995	0.963	1	0.951	0.904	0.897	0.877	0.946	1

4. Conclusions

In conclusion, the utilization of the YOLOv8 algorithm for traffic sign detection and recognition signifies a significant leap forward in enhancing both driver safety and autonomous driving systems. With its exceptional real-time object detection capabilities, YOLOv8 achieves an impressive recognition accuracy of 94% while processing frames at a remarkable rate of 43.8 frames per second, surpassing its predecessor, YOLOv3. Its versatility extends beyond traffic signs to encompass pedestrian, traffic light, and vehicle identification, underscoring its broad applicability in intelligent driving applications. Comparative evaluations against models like YOLOv7 and YOLOv5 consistently position YOLOv8 as the superior choice, boasting a mAP of 0.945, indicative of its unmatched accuracy across various traffic sign classes. Notably, YOLOv8 demonstrates exceptional precision and recall for critical signs, further cementing its efficacy in real-world scenarios. Furthermore, the integration of a voice assistant feature ensures that, upon detecting a traffic sign, the system provides immediate auditory feedback, significantly enhancing driver awareness and safety by delivering clear and timely alerts. While challenges persist in deployment, including environmental variations and hardware efficiency demands, YOLOv8 remains pivotal in advancing traffic sign detection systems, offering a compelling blend of accuracy and real-time processing vital for the progression of autonomous driving and road management technologies.

Declarations

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- [1] M. Flores-Calero *et al.*, “Traffic sign detection and recognition using YOLO object detection algorithm: A systematic review,” *Mathematics*, vol. 12, no. 2, p. 297, 2024.
- [2] E. Güneç and C. Baylın, “An implementation of traffic signs and road objects detection using faster R-CNN,” *Sakarya University Journal of Computer and Information Sciences*, vol. 5, no. 2, pp. 216–224, 2022.
- [3] S. Kim, J. Park, and H. Lee, “Robust Traffic Sign Detection and Classification in Diverse Environmental Conditions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 2211–2222, 2024, [Online]. Available: <https://doi.org/10.1109/TITS.2024.3045678>
- [4] J. Yao, B. Huang, S. Yang, X. Xiang, and Z. Lu, “Traffic sign detection and recognition under low illumination,” *Mach Vis Appl*, vol. 34, no. 75, 2023, [Online]. Available: <https://link.springer.com/article/10.1007/s00138-023-01417-y>
- [5] F. Shao, X. Wang, F. Meng, T. Rui, D. Wang, and J. Tang, “Real-time traffic sign detection and recognition method based on simplified Gabor wavelets and CNNs,” *Sensors*, vol. 18, no. 10, p. 3192, 2018.
- [6] V. Pople and others, “Public social distance monitoring system using object detection yolo deep learning algorithm,” *SN Comput Sci*, vol. 4, no. 6, pp. 1–21, 2023.
- [7] M. Li, W. Zhang, and Y. Liu, “Enhanced Traffic Sign Detection and Recognition Using Deep Learning Techniques,” *J Adv Transp*, vol. 48, no. 5, pp. 1078–1092, 2024, [Online]. Available: <https://doi.org/10.1155/2024/5612345>
- [8] J. Wang *et al.*, “A lightweight vehicle mounted multi-scale traffic sign detector using attention fusion pyramid,” *J Supercomput*, vol. 80, no. 3, pp. 3360–3381, 2024.
- [9] A. Avramović, D. Sluga, D. Tabernik, D. Skočaj, V. Stojnić, and N. Ilc, “Neural-network-based traffic sign detection and recognition in high-definition images using region focusing and parallelization,” *IEEE Access*, vol. 8, pp. 189855–189868, 2020.
- [10] Z. Ou, F. Xiao, B. Xiong, S. Shi, and M. Song, “FAMN: feature aggregation multipath network for small traffic sign detection,” *IEEE Access*, vol. 7, pp. 178798–178810, 2019.
- [11] X. Qiao, “Research on Traffic sign recognition based on CNN Deep Learning Network,” *Procedia Comput Sci*, vol. 228, pp. 826–837, 2023.
- [12] R. K. Megalingam, K. Thanigundala, S. R. Musani, H. Nidamanuru, and L. Gadde, “Indian traffic sign detection and recognition using deep learning,” *International journal of transportation science and technology*, vol. 12, no. 3, pp. 683–699, 2023.
- [13] C. Sun, M. Wen, K. Zhang, P. Meng, and R. Cui, “Traffic sign detection algorithm based on feature expression enhancement,” *Multimed Tools Appl*, vol. 80, no. 25, pp. 33593–33614, 2021.
- [14] X. He, T. Li, and Y. Yang, “Improved traffic sign detection algorithm based on improved YOLOv8s,” *J. Comput. Electron. Inf. Manag*, vol. 12, pp. 38–45, 2024.
- [15] K. Khnissi, C. Ben Jabeur, and H. Seddik, “Implementation of a compact traffic signs recognition system using a new squeezed YOLO,” *International Journal of Intelligent Transportation Systems Research*, vol. 20, no. 2, pp. 466–482, 2022.
- [16] H. Gezgin and R. M. Alkan, “Traffic sign detection and recognition based on MMS data using YOLOv4-Tiny algorithm,” *Neural Comput Appl*, vol. 36, no. 33, pp. 20633–20651, 2024.
- [17] R. C. Rodríguez, C. M. Carlos, O. O. V. Villegas, V. G. C. Sánchez, and H. de J. O. Domínguez, “Mexican traffic sign detection and classification using deep learning,” *Expert Syst Appl*, vol. 202, p. 117247, 2022.
- [18] S. Sun *et al.*, “ShuffleNetv2-YOLOv3: a real-time recognition method of static sign language based on a lightweight network,” *Signal Image Video Process*, vol. 17, no. 6, pp. 2721–2729, 2023.
- [19] J. Xu, Y. Huang, and D. Ying, “Traffic sign detection and recognition using multi-frame embedding of video-log images,” *Remote Sens (Basel)*, vol. 15, no. 12, p. 2959, 2023.
- [20] R. K. Megalingam, K. Thanigundala, S. R. Musani, H. Nidamanuru, and L. Gadde, “Indian traffic sign detection and recognition using deep learning,” *International Journal of Transportation Science and Technology*, vol. 12, no. 3, pp. 683–699, 2023, doi: <https://doi.org/10.1016/j.ijtst.2022.06.002>.
- [21] A. Garcia, C. Martinez, and J. Perez, “Multispectral Traffic Sign Recognition for Enhanced Road Safety,” *Sensors*, vol. 24, no. 9, pp. 3521–3536, 2024, [Online]. Available: <https://doi.org/10.3390/s24093521>
- [22] X. He, T. Li, and Y. Yang, “Improved traffic sign detection algorithm based on improved YOLOv8s,” *J. Comput. Electron. Inf. Manag*, vol. 12, pp. 38–45, 2024.
- [23] M. Flores-Calero *et al.*, “Traffic sign detection and recognition using YOLO object detection algorithm: A systematic review,” *Mathematics*, vol. 12, no. 2, p. 297, 2024.
- [24] V. R. M. V. and V. Pople, “Public Social Distance Monitoring System Using Object Detection YOLO Deep Learning Algorithm,” *SN Comput Sci*, vol. 4, no. 6, p. 718, Sep. 2023, doi: [10.1007/s42979-023-02131-2](https://doi.org/10.1007/s42979-023-02131-2).
- [25] F. Ahmed, A. Khan, and S. Khan, “Fusion of Lidar and Camera Data for Improved Traffic Sign Detection in Autonomous Vehicles,” *Rob Auton Syst*, vol. 145, p. 104456, 2024, [Online]. Available: <https://doi.org/10.1016/j.robot.2024.104456>
- [26] M. Venkatachala Appa Swamy *et al.*, “Design and development of IoT and deep ensemble learning based model for disease monitoring and prediction,” *Diagnostics*, vol. 13, no. 11, p. 1942, 2023.

Authors' Profiles



Dr. Mareeswari V. is working as an Associate Professor at the School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology (VIT), Vellore, Tamilnadu, India - 632014. She has over 21 years of academic and research experience. She received her Ph.D. in Information Technology and Engineering in the Web Service domain. She has produced several national and international articles in reputed journals, conferences, and book chapters. Her area of interest includes Medical Image processing using Machine and Deep Learning, Data analytics on Blockchain, Prediction systems on Web Service Ranking, Internet of Things (IoT), Routing on Mobile Ad-hoc networks, and E-Commerce systems. She is a life member of the Computer Society of India (CSI) and the Soft Computing and Research Society (SCRS).



Vijayan R. is a professor and senior faculty member at the School of Computer Science and Information Systems [SCORE], Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India. He is a member of the Computer Society of India (CSI). He has published several national and international research articles in reputed journals and conferences. His research interests include web technologies, wireless networks, ad hoc networks, computer networks, machine learning, the Internet of Things (IoT), Mobile Computing, Cloud Computing, and Data Science.



Shajith Nisthar holds a Bachelor's degree in Computer Science (BSc) from Vellore Institute of Technology (VIT), Vellore, India. He did many projects related to Python programming and software development, currently pursuing his Master of Computer Applications (MCA) at VIT Chennai, his research interests focus on Image processing and cloud computing.



Rahul Bala Krishnan holds a Bachelor's degree in Computer Science (BSc) from Vellore Institute of Technology (VIT), Vellore, India. He did many projects related to Python programming and software development, currently pursuing his Master of Business Applications (MBA), his research interests focus on Image processing and business computing.

How to cite this paper: Mareeswari V., Vijayan R., Shajith Nisthar, Rahul Bala Krishnan, "Traffic Sign Detection and Recognition Using Yolo Models", International Journal of Information Technology and Computer Science(IJITCS), Vol.17, No.3, pp.13-25, 2025. DOI:10.5815/ijitcs.2025.03.02