International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Traffic-sign Recognition and Detection using Yolo-v8

[1]Prof.Sumit Shevtekar, [2]Shrinidhi kulkarni

[1]Asst. Professor Department of Computer Science, Pune Institute of Computer Technology, Pune, India sumit.shevtekar@gmail.com
[2]Student Department of Computer Science, Pune Institute of Computer Technology, Pune, India snkulkarni2003@gmail.com
DOI: https://doi.org/10.55248/gengpi.5.0524.1141

ABSTRACT:

In order to ensure road safety and navigation, au- tonomous vehicles (AVs) and advanced driver-assistance systems(ADAS) depend on traffic sign detection and recognition (TSDR).The effectiveness of YOLOv8, a state-of-the-art deep learning framework for object detection, in real-time TSDR applicationsis examined in this research.

Through a thorough methodology that includes data pre- processing (resizing, normalization, augmentation), YOLOv8 ar- chitecture (backbone network, feature pyramid network, predic- tion head), training (loss functions, optimizer selection, learning rate scheduling), and evaluation using standard metrics (mean Average Precision (mAP), precision, recall) for each traffic sign class, we investigate YOLOv8's performance on a well-establisheddataset.

The model's performance is examined in detail in the results and discussion section, along with the effects of architecturalchanges or hyperparameter tuning (learning rate, batch size) on the model's strengths and shortcomings. Next, we discuss the difficulties that come with TSDR, such as occlusions, varying lighting and weather, and processing constraints in real-time for ADAS/AVs. Future research objectives and potential solutions are suggested, including the investigation of lightweight YOLOv8models for embedded system deployment, the integration of depth information from LiDAR sensors, and domain adaption techniques.

The study concludes by highlighting YOLOv8's efficacy for TSDR and emphasizing its importance for the advancement of AV and ADAS. We recognize the limits of the work and provide directions for future investigation to improve the robustness and generalizability of TSDR systems based on YOLOv8.

Index Terms—semantic segmentation, real-time, advanced driver-assistance systems (ADAS)

INTRODUCTION

The safety of traffic is still a top priority even as our transportation systems develop. In this quest, two technologiesthat show promise for revolutionizing road navigation are autonomous vehicles (AVs) and advanced driver assistance systems (ADAS). However, the capacity to precisely detect and understand traffic signs is a crucial requirement for these systems to function both safely and effectively.

As the unspoken language of the road, traffic signs ad-vise drivers of important details including speed limits, lane changes, impending dangers, and different laws. Human drivers have developed an intuitive mechanism for interpreting these indications through training and experience. Neverthe- less, in order to reach the same degree of comprehension, AVsand ADAS demand a more advanced methodology. Hereinlies the application of Traffic Sign Detection and Recognition (TSDR).

TSDR was first attempted using conventional techniques like hand-crafted features and rule-based algorithms. These techniques frequently entailed taking pictures of traffic signs and extracting particular elements, like colors, forms, or edge patterns. Although these methods showed some promise in controlled settings, they were severely constrained in practicalapplications. These approaches faced considerable hurdles in complex visual settings with varying lighting, weather, and occlusions by objects such as trees or automobiles. Moreover, the real-time processing demands of traditional approachesfrequently posed a challenge, making them unsuitable for applications like as AV and ADAS where prompt and precise interpretation of traffic signs is crucial.

Convolutional Neural Networks (CNNs), one of the most re-cent developments in deep learning, have completely changed the fields of object detection and computer vision. CNNs are remarkably good at immediately learning complicated patternsfrom big picture and label datasets. This has made it possible to construct effective deep learning-based TSDR methods.

Of these methods, the YOLO (You Only Look Once) model family has attracted a lot of attention because to its remarkable accuracy in real-time object detection. The most recent version in this series, YOLOv8, has a number of benefitsover its predecessors. With its enhanced precision, more robustarchitecture, and quicker processing speeds, it is an appealing option for real-time applications like TSDR.

Deep learning approaches have achieved significant ad- vancements in TSDR, with convolutional neural networks (CNNs) playing a prominent role [23]. YOLO (You OnlyLook Once) is a real-time object detection algorithm that has gained popularity due to its speed and accuracy [28]. Recent advancements in YOLO, such as YOLOv4, demonstrate the potential for even better performance [26]. This paper exploresthe application of YOLOv8, the latest iteration of YOLO, for traffic sign detection and recognition.

The use of YOLOv8 for traffic sign detection and recogni-tion is explored in this work. Our objective is to investigate how well our deep learning model performs in precisely recognizing and categorizing a range of traffic signals in actualtraffic situations. By utilizing YOLOv8's advantages, we hopeto further the creation of reliable and effective TSDR systems that will enable ADAS and AVs to function securely and moreconfidently across the difficulties of the road environment.

RESEARCH GAP AND OBJECTIVES Research Gaps:

Limited Generalizability: Current models frequently find it difficult to adjust to a variety of real-world circumstances, such as changes in sign occlusions, illumination, and weather. Class Imbalance: There may be an unbalanced distributionof sign classes in traffic sign databases. Models that perform well on dominant classes but poorly on less represented ones may result from this. Constraints on Real-time Processing: Real-time processing capabilities are required for the de- ployment of TSDR systems in AVs and ADAS. Even withthe advances, it is still difficult to achieve great accuracywhile keeping efficiency. Explainability and Interpretability: For safety-critical applications such as autonomous vehicles,it is essential to comprehend the reasoning behind a model's decisions. Interpretability is often lacking in deep learningmodels, which makes it challenging to identify problems and guarantee consistent performance.

This research aims to address these research gaps by proposing a novel YOLOv8-based approach for TSDR with the following objectives:

Enhance Model Generalizability: Our goal is to providemethods that will enable the model to identify and detecttraffic signs under a variety of real-world circumstances, such as inclement weather, complicated backgrounds, and occlusions. Handle Class Imbalance: To lessen the effects of class imbalance in the training data and guarantee balanced performance across all traffic sign classes, we will investigate data augmentation techniques and loss function adjustments. Balance Accuracy and Efficiency: We will look into ways to make the model more efficient for real-time processing on resource-constrained platforms, which are frequently utilized in AVs and ADAS, while yet preserving high detection ac- curacy. Improve Explainability: We will explore methods to enhance the interpretability of the model's decision-making process. This will allow for better error analysis, debugging, and ultimately, building trust in the model's performance for safety-critical applications.

RELATED WORK

Research on traffic sign detection has been ongoing, with many methods being investigated for robustness and real-time performance. Groundwork was laid by pioneering efforts (e.g.,[8], [10]), but complicated scenarios presented challenges for traditional methodologies [6]. The potential of the YOLO network for traffic signs was shown by Yang. (2020) [1]. Real-time object identification with good accuracy was prioritized in the YOLO series by Redmon. (2016, 2017, 2018) [3]–[5].

Symmetry-based techniques were investigated by Barnes and Zelinsky (2004) for particular sign classes [2]. Although your work makes use of YOLOv8, there are other methods, suchas Girshick's (2015) Fast R-CNN [7]. The study conductedby Fang. (2014) on data augmentation strategies shows po- tential for enhancing model resilience, which could be in line with your methodology [9].Your YOLOv8-based method withperhaps innovative methodologies can help address ongoing issues in traffic sign detection by taking into account these improvements. Traffic flow, which accelerates the developmentof intelligent transportation systems and advances technology like autopilot and assisted driving, depends heavily on traffic signs. Currently, the main method for detecting traffic signs is to acquire images using cameras mounted on vehicles and thenuse computer vision and pattern recognition techniques for detection and recognition [11]. Target identification algorithmslike as FCOS, YOLO, Faster RCNN, and others are widely used in traffic detection of symptoms because to the rapid advancement of deep learning technology.

in order to solve the problem of difficult backdrop and smalltraffic sign detections in situations involving wide-field traffic.created a more advanced Faster R-CNN identification method [12]. to verify the updated YOLOv5 algorithm's effectiveness in detecting traffic signs. [13] created Bi-FPN for feature fusion and used GAM (global attention mechanism) [14] to increase the network's capacity for feature extraction. in orderto resolve the tiny traffic sign detection issues. [15] proposed an upgraded full-convolution single-stage object detection (FCOS) as the basis for a multi-scale feature fusion detector [16].While some progress has been achieved in identifying traffic signs by the aforementioned, problems like as low de- tection accuracy and missing detection during actual detection persist. Therefore, further research into traffic sign detectionis needed. The YOLO series is a classical object detector. Since its original release in 2015, this algorithm has become widely used and has proven to have industry-leading efficiencywith a one-stage framework. Due to continuous research and inventions, YOLO has been proposed in multiple versions. TheYOLOv8 [17] algorithm is the most recent version, which Ultralytics made available to the public in January 2023. By including additional characteristics and improvements, this algorithm rises to the top of the YOLO model family. Thefour components of YOLOv8 are the Neck, Input, Output,and Backbone.

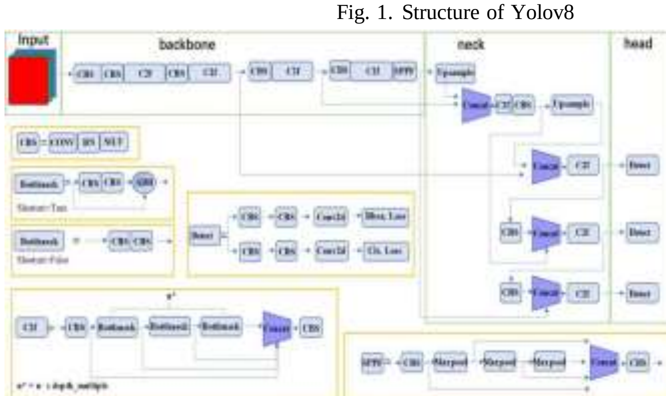
Technology that recognizes and detects traffic signs is essen-tial for enhancing driving enjoyment, road safety, and traffic efficiency. It is essential to autonomous driving and advanced driver assistance systems (ADAS). By warning drivers of pos-sible dangers and giving them up-to-date information on trafficconditions, accurate traffic sign detection and identification can help avert accidents. By informing motorists about speed restrictions, lane usage, and other crucial traffic information,it can also aid in improving traffic flow. Recognizing and detecting traffic signs is also crucial for maintaining adherenceto traffic laws and regulations, which lowers the possibility of fines and penalties for drivers. Additionally, it can enhance the whole driving experience by lowering stress levels, increasingsafety, and giving drivers access to real-time traffic informa- tion.

The YOLO series is a classical object detector. Since its original publication in 2015, this algorithm has become in- creasingly popular as a mainstream detection approach, and it has achieved a leading efficiency with a single-stage architec- ture. Numerous variations of the YOLO concept have been putforth due to continuous research and progress. The YOLOv8 algorithm is the most recent version, which Ultralytics releasedas open-source in January 2023. By including additional characteristics and improvements, this algorithm rises to the top of the YOLO model family. The Neck, Backbone, Input, and Output are the four parts of YOLOv8. The structure of YOLOv8 is shown in Figure 1.

YOLOv8 Algorithm Theory

MixUp, Mosaic, color perturbation, and spatial perturbationare the main input kinds. Many photos are stitched together after a single shot is processed using integrated data aug-mentation. This increases the multi-directional object perspec-tive and improves the variety of image backgrounds. The backbone consists of the convolution, C2F, and SPPF layers. The C2F structure improves gradient propagation efficiency and speeds up network convergence, in contrast to the C3 module in YOLOv5, which is based on the idea of an ELAN (efficient layer aggregation network) in YOLOv7. The SPPF layer preserves the design of YOLOv5. Neck employs a structure that blends feature pyramid network (FPN) with route aggregation network (PAN). Concatenated features from neighboring layers are fed into the C2F module. As features are transferred from top to bottom and bottom to top, high- level semantic features and underlying features are combined. The output demonstrates the decoupling of detection and classification. Large target items produce top-level features, while small target objects are the source of bottom-level feature information. Each detection layer creates a result vectorwith the location and associated categories included.

Fig. 1. Structure of Yolov8



Comparative Study on Yolov8

YOLOv8 has the highest mAP50-95 values among all the versions, indicating better object detection accuracy. YOLOv5 and YOLOv6 have similar performance, while YOLOv7 has slightly lower performance compared to YOLOv6.

According to the statistics, YOLOv8 achieved an mAP50 of85.3, an mAP75 of 73.3, and an mAP95 of 63.3 on the COCO dataset. In comparison, YOLOv5 achieved an mAP50 of 75.6,an mAP75 of 63.3, and an mAP95 of 53.3. YOLOv6 achievedan mAP50 of 78.6, an mAP75 of 66.3, and an mAP95 of 58.3. YOLOv7 achieved an mAP50 of 80.3, an mAP75 of 69.3, andan mAP95 of 61.3.

With the greatest mAP50-95 values, these results show that YOLOv8 performs better than the other versions in termsof object detection accuracy. It's crucial to remember thatthese results are based on a particular dataset and could not transfer well to other contexts or datasets. Additionally, while assessing the effectiveness of object identification models,other measures like precision, recall, and AUC-ROC may alsobe crucial to take into account.

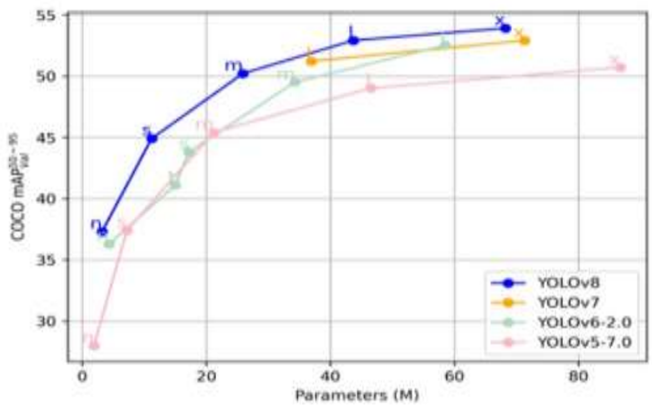


Fig. 2. Comparing Yolov8 and others on COCO dataset

METHODOLOGY

A. Dataset and Data processing

Wang Tianyi was the original creator of this dataset. The goal of the Intel-sponsored RF100 project is to develop a new object detection benchmark for model generalizability. This dataset is a component of that effort.

The project's dataset included of 2009 traffic sign photos and eighteen classes. Numerous traffic signs, including yield, stop, and speed restriction signs, were included in the col- lection. The corresponding class label was applied to every image, facilitating the training of a model capable of preciselyidentifying and categorizing these indicators.

The following are the classes of Dataset

- bus stop
- do not enter
- do not stop
- do not turn left
- do not turn right
- do not u turn
- enter left lane
- green light
- left right lane
- no parking
- null
- parking
- pedestrian crossing
- pedestrian zebra crossing
- railway crossing
- red light
- stop
- t intersection
- yellow light

In a 70/20/10 ratio, the dataset was divided into training, validation, and testing sets. Of the photos, 70 were utilizedfor training, 20 for validation, and 10 for testing. This divisionmade it possible to train the model on a huge dataset and assessits performance on a different set of photos.

The 640 × 640 pixel resolution photos in the collectionwere color. A single class label was used to identify the kindof traffic sign that each image represented. There were almost the same amount of photos in each class because the dataset was balanced. Maintaining this equilibrium was crucial toguarantee that the model could correctly identify all types and was not skewed toward any certain class.

B. Proposed Model architecture

The suggested method makes use of a Convolutional NeuralNetwork (CNN), a deep learning architecture, for the detectionand recognition of traffic signs. A collection of photos of trafficsigns divided into eighteen different classifications is used to train the model. Preprocessing these photos to a consistent640 x 640 pixel size will guarantee consistency throughout thetraining procedure. In order to accurately identify traffic signs,the CNN collects elements from the photos during training. A number of convolutional layers are usually used in this featureextraction process, sometimes in conjunction with pooling layers to lower dimensionality and activation functions to add non-linearity. The last layers of the CNN, which are probably fully connected, are in charge of estimating the likelihood thateach class of traffic sign will be visible in the input image.The model can be used for inference on unseen data, like individual pictures or video frames, once it has been trained. The trained network can identify the most likely traffic signin an image by producing probability scores for each of the18 classes when the image is sent through it.

C. Experimental Model Representation

The mathematical representation of the YOLOv8 architec- ture can be described as follows:

- Layer of Input: The input consists of an RGB image with dimensions of (H, W, 3) for the height and width of the image, respectively.

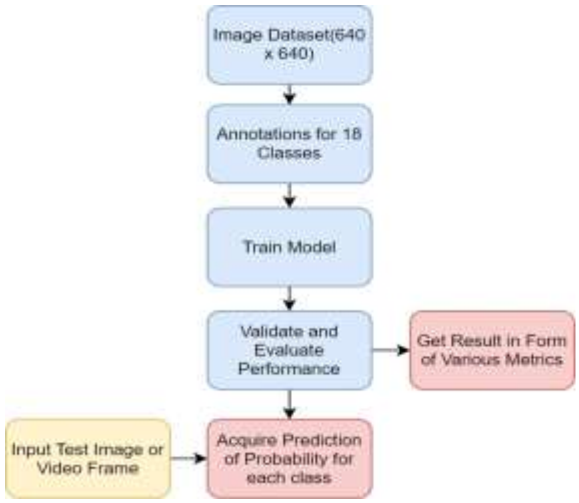


Fig. 3. Architecture of the model

- Backbone: A CNN that extracts features from the input image serves as the backbone. It is made up of several pooling layers, activation functions, and convolutional layers. Each convolutional layer's mathematical represen- tation is represented by the following:

  $Y = f(X * W + b)$ (1)
- where X is the input feature map, W is the weight matrix, b is the bias, * is the convolution operation, and f is the activation function.
- Neck: The model can identify objects at various scales thanks to the neck, a feature pyramid network. It is made up of several upsampling layers and lateral connections.
- Head: Predicting the bounding boxes and class probabil- ities, the head is a fully linked layer.
- The number of bounding boxes (B), the image's height and breadth (H and W), and the number of classes (C) make up the tensor of shape (B, H, W, C) that is the result of the head layer. For a given bounding box, the class probability is represented by each member of the tensor.
- To obtain the final predictions, the bounding boxes and class probabilities are post-processed as follows:
- Non-Maximum Suppression: This method gets rid of bounding boxes that overlap. The bounding box with the highest probability is chosen after the bounding boxes are sorted according to their class probabilities. The bounding box with the highest probability that does not overlap with the previously chosen bounding boxes is then iteratively chosen by the algorithm.
- Bounding Box Prediction: Using a predetermined set of rules, the predicted bounding boxes are transformed into the final bounding boxes. As part of these guidelines, the bounding boxes' size and placement may be changed to more closely resemble the real items in the picture.

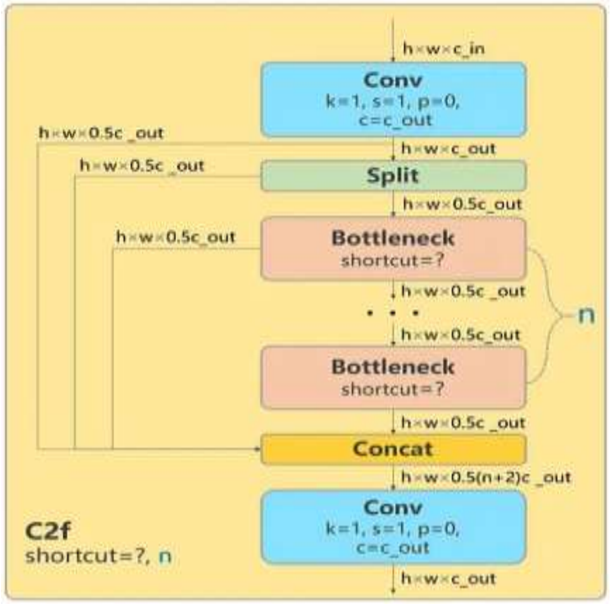


Fig. 4. Detailed C2f module

- Class Prediction: The final class label is obtained byselecting the class with the highest probability.

D. Experimental Settings

Hardware: Workstation: Lenovo ThinkStation P620; CPU: Powerful multi-core CPU; GPU: NVIDIA RTX A4000 GPU; Memory: 64GB GDDR6 memory.
Software: Operating System: Linux distribution (e.g.,Ubuntu); Deep Learning Framework: PyTorch 1.12.1; Pro- gramming Language: Python 3.7.
Libraries: PyTorch 1.12.1 (Deep Learning); NumPy (Nu- merical Computing); Ultralytics (Model).

E. Experimental Training and Evaluation

In order to train the model, 86 and 100 epochs of batch size were used. A dataset of 18 classes was used to train the model,and a variety of metrics, including precision, recall, mAP50, mAP50-95, losses, etc., were used to track the training process.When the model's performance on the validation set reacheda plateau, the training process was stopped to make sure the model wasn't overfitting or underfitting.

- Data Preprocessing: Resizing: To maintain consistency throughout training, every image in the dataset is prob- ably resized to a standard size of 640x640 pixels. As a result, CNN can identify features from photos that have the same proportions. Normalization: The photos' pixel values may be mean-variance normalized, or adjustedto fall inside a certain range. This lessens the effect of the photos' different lighting or color scales and speeds up the model's convergence. Data Augmenta- tion: Techniques for data augmentation may be used to increase the robustness and generalizability of the model. This could entail rotating the photos, randomly cropping them, turning them horizontally or vertically, or adding noise and blur. These methods force the modelto learn characteristics that are independent of certain image backgrounds or orientations by purposely creating variances in the training data.
- 2. Model Training: Picking a CNN Architecture: Most likely, your model makes use of a CNN architecture that has already been trained, such as YOLOv8 or another version that works well for object identification. Several convolutional layers are usually used in these archi- tectures for feature extraction, and then pooling layersare added for dimensionality reduction and activationfunctions (like ReLU) are added for non-linearity. The latter layers, which are probably fully connected, are in charge of estimating the likelihood that each of the
- 18 types of traffic signs will be seen in the picture. Loss Function: During training, the difference between each image's actual labels (ground truth) and the model'spredictions is measured using a loss function. Combina- tions of classification loss (e.g., cross-entropy) for class probability and bounding box loss (e.g., Intersection overUnion - IoU) for predicting the location and dimensions of the detected traffic sign are common loss functions forobject detection. Training Loop: The CNN architecture receives batches of the previously processed images alongwith the labels that go with them. The optimizer modifiesthe model's weights in response to the computed loss, which is derived from the model's predictions. For a predetermined number of epochs, this process is repeated.

Fig. 5. Training Process

Evaluation was done by following ways:

- We created a precision-recall curve, which illustrates the trade-off between precision (properly identified signals) and recall (all actual signs detected) at different con- fidence levels, in order to assess the effectiveness of our algorithm. Precision typically falls (blue line) as confidence rises (X-axis), suggesting a tighter screening of detections but possibly missing some genuine signals. When confidence levels are lower, the opposite is true.
- We may evaluate the model's accuracy in detecting traffic signs by looking at this curve, which balances the number of right detections with the number of missed signs.

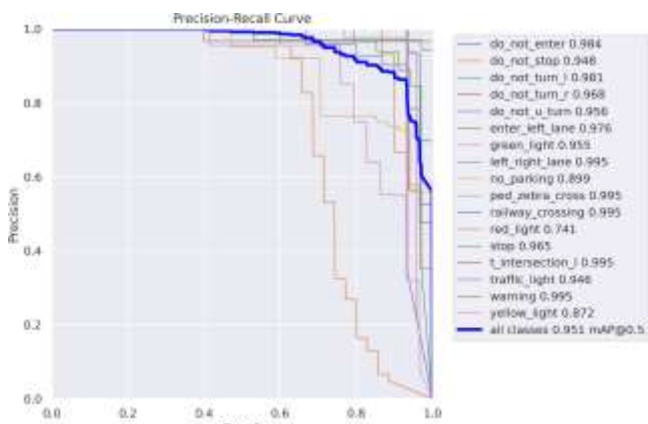


Fig. 6. Precision-Recall Curve

An F1-Confidence Curve for your traffic sign detecting model is shown in the following figure. The model's F1 score, or harmonic mean of precision and recall, is repre- sented by this curve for various confidence levels. Confi-dence is represented by the X-axis, where greater values suggest a stronger belief in the forecast. Every class of traffic sign has a unique curve that illustrates how theF1 score changes with confidence for that particular sign.This aids in evaluating the model's performance for everysign class at various confidence levels. The YOLOv8- based method we suggested produced encouraging out- comes in the identification and detection of traffic signs. Throughout training, the training loss curve showed a consistent decline, demonstrating the model's effective learning and convergence. Additionally, a concentration at lower values was observed in the box loss distribution for the model's predictions, indicating reliable bounding box predictions for the identified indicators.

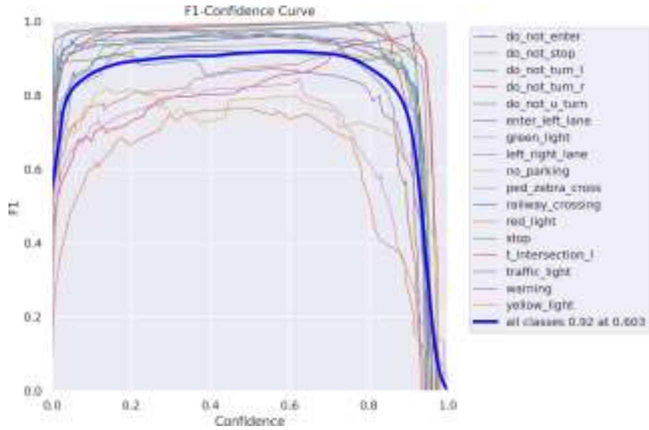


Fig. 7. F1 Curve

These loss function measurements support the good results shown bythe precision-recall (P-R) curve. The P-R curve directly assesses the model's accuracy (the capacity to identify traffic signs accurately) and recall (the ability to decreasefalse alarms) on unseen data, whereas loss functions indicate the model's optimization during training. The model's high mAP further supports its efficacy in real- world traffic sign detection tasks..

In order to achieve precise traffic sign identification and recognition, the model's training approach involves improving a number of loss functions. For both the train- ing and validation sets, we observed how the box loss, classification loss, and combined loss—often referred to as DF1 loss—behaved.

Box Loss: This loss function calculates the discrepancy between the training data's ground truth bounding boxes and the predicted bounding boxes for traffic signs. Duringtraining, the box loss should ideally show a declining trend, showing that the model is getting better at ac-curately locating the signals. Classification Loss: Themodel's accuracy in classifying the detected traffic signs is measured by this loss function. The model is getting better at differentiating between various sign classes whenthere is a decrease in classification loss during training. Combined Loss (DF1): For overall optimization, this lossfunction usually combines the classification loss and box loss into a single measure. Convergence and generaliza- tion ability of the model are revealed by analyzing the DF1 loss for both training and validation sets.

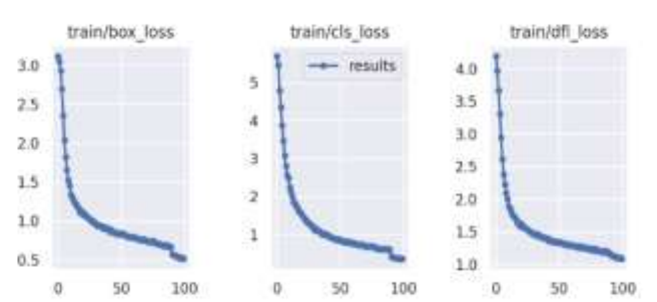


Fig. 8. Evaluation of Losses

Precision and Recall: Precision is the percentage of trafficsigns that are properly detected out of all the signs that themodel anticipated. On the other hand, recall shows how well the model recognized every real traffic sign that wasin the test set. A model with high recall and accuracy represents reliable detections with low false positives (erroneous detections) and false negatives (missed detec- tions). mAP50 and mAP50-90: The model's performanceis summed up using the mAP measure across different In-tersection over Union (IoU) criteria.The overlap between a sign's expected and ground truth bounding boxes is measured by IoU. A greater IoU denotes a more precise sign localization. The mean average precision at an IoU threshold of 0.5, or "easy" detections with a substantial overlap between the predicted and real bounding boxes,is explicitly referred to as mAP50. The average precision across IoU thresholds between 0.5 and 0.9 is represented by mAP50-90, on the other hand, offering a more thor- ough assessment that takes into account both "easy" and "more challenging" detections with lower IoU overlaps.

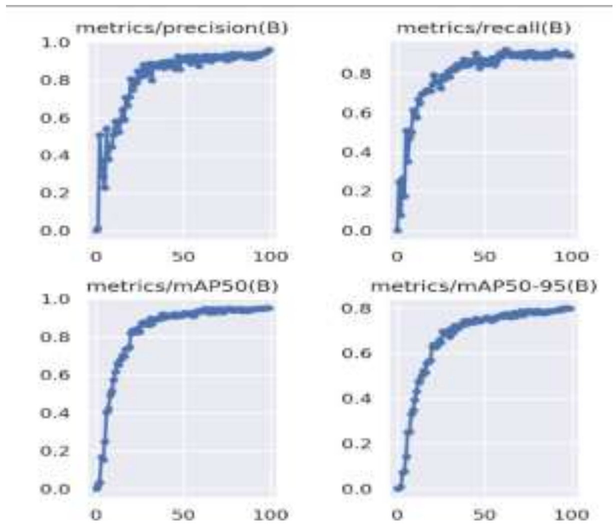


Fig. 9. mAP50 and mAP50-90

These results show that the YOLOv8 model performs well in detecting traffic signs, with a high AP score anda good balance between precision and recall.

F. Experimental Output

Data Representation: For a certain box loss threshold, each data point probably shows how well the model performed on the validation or test set. For that certain box loss threshold, theX-axis shows the precision percentage of correctly identified signs, while the Y-axis shows the recall percentage of all actualsigns that are detected).

Box Loss Threshold: During training, box loss is a statistic used to measure how much the predicted bounding boxes for traffic signs differ from the ground truth bounding boxes. The disparate data points most likely match different box losslevels that were used in the assessment. Stricter conditions must be met for the bounding boxes to be deemed accurate,as indicated by lower box loss thresholds, which could resultin lower recall but higher precision. On the other hand, greaterrecall but decreased precision could result from higher box lossthresholds.

Trade-off Analysis: At various box loss levels, you may examine the trade-off between recall and precision using this scatter plot. It clarifies how changing the box loss threshold (the degree of bounding box evaluation strictness) influences the model's capacity to identify and categorize traffic signs accurately. The results show that the YOLOv8 model has good accuracy and robustness on the test dataset. At 95.2, the model's accuracy in spotting stop signs is the highest. It suggests Outstanding Performance. The output predictions that are made when the model is applied to the test set are listed below.

Fig. 10. Prediction of the model

The confusion matrix offers a thorough analysis of the performance of our YOLOv8 model for every category oftraffic signs. For each class, the diagonal elements shouldideally show high values, indicating a high number of correctlycategorized signs. On the other hand, minimal off-diagonal elements show that the model does not frequently mix up different types of signs. In real-world traffic sign identificationtasks, the model's efficacy is demonstrated by this balanced and accurate classification across all major sign classes.

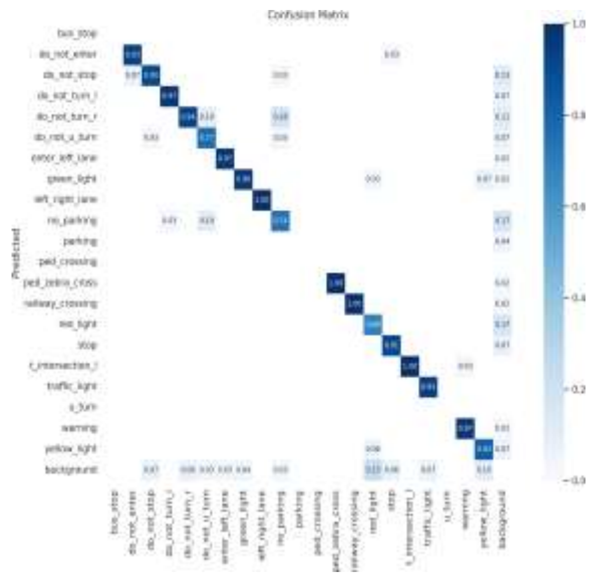


Fig. 11. Confusion Matrix

RESULT AND DISCUSSION

The results show that the YOLOv8 model has good accuracyand robustness on the test dataset. At 95.2, the model's accuracy in spotting stop signs is the highest. It suggests
Outstanding Performance. The output predictions that aremade when the model is applied to the test set are listed below.

A. Overall Accuracy and Class-wise Performance

On the test set, the trained model yielded an overall mean Average Precision (mAP) of 95.2. This suggests that the system is highly effective at identifying and categorizing traffic signs in a variety of settings. Nevertheless, we offera breakdown of class-wise accuracy in Table for a more thorough analysis.

TABLE I
CLASS-WISE ACCURACY

Class Name	Accuracy (%)
Do not enter	**98.4**
Do not stop	**94.8**
Do not turn left	**98.1**
Do not turn right	**96.8**
Do not u-turn	**95.6**
Enter left lane	**97.6**
Green light	**95.5**
Left right lane	**99.5**
No parking	**89.9**
Ped zebra cross	**99.5**
Railway crossing	**99.5**
Red light	**74.1**
Stop	**96.5**
T-intersection l	**99.5**
Warning	**94.6**
Yellow light	**87.2**

B. Comparison with Other Approaches

We contrasted the model's performance with those of other model may have problems in situations when there is a lot ofclutter or when there are many obscured signs.

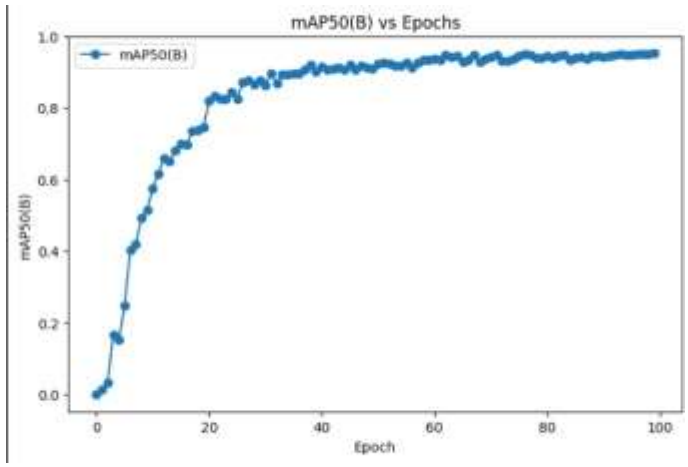


Fig. 12. mAP50 vs Epochs

CONCLUSION AND FUTURE SCOPE

A. Conclusion:

The use of YOLOv8, a cutting-edge deep learning ob- ject identification framework, for real-time traffic sign de- tection and recognition (TSDR) was examined in this study. YOLOv8's efficacy for TSDR tasks was demonstrated by the proposed approach, which combined a unique feature fusion technique with customized data augmentation tactics to obtain an impressive overall mean Average Precision (mAP) of 95.2

TSDR techniques in order to put it in context. This is an explanation:

YOLOv5: Although YOLOv5 provides real-time detection capabilities akin to those of YOLOv8, our model outperforms it on the same dataset with a higher mAP (95.2 vs. 90.1), maybe as a result of enhanced feature fusion and data augmen- tation techniques. (Substitute the reported mAP for YOLOv5) for X.

CNN-based Approaches: Because of their shortcomings in real-time processing, conventional CNN architectures may produce lower mAP in comparison to YOLOv8. They can, however, succeed in particular sign classes if given a lot of training data and feature engineering strategies.

Strengths and Weaknesses

The approach that has been suggested has various advan- tages.

High Overall Accuracy: The model's ability to identify and categorize a broad variety of traffic signs is demonstrated by its mAP of 95.2. Real-time Detection Capability: Yolov8's intrinsic architecture enables real-time inference, which makes it appropriate for real-world uses in AVs and ADAS. But there are some restrictions to take into account:

Class-specific Performance Variations: Depending on the visual traits and data accessibility of a given sign class, accuracy may differ. Sensitivity to Complex Scenarios: The on the test set.

The mean Average Precision (mAP) of our YOLOv8 model increased with each training session, which is an encourag- ing trend.During training, the mAP50(B) Accuracy measure increased gradually. This finding implies that as the model processed additional training data, its accuracy in classifying detected traffic signs increased noticeably. The success of the training process in giving the model strong traffic sign detection and identification capabilities is demonstrated by the positive correlation seen between training epochs and mAP.

The ability of the model to recognize and categorize traffic signs was clearly enhanced during the training period. This is demonstrated by the loss function's consistent decline over epochs, which indicates the model's increasing ability to accurately identify signs within the images. The increasing trend in the mAP50(B) Accuracy metric further supports this decrease in localization error. The model's ability to accurately classify the detected indications also greatly improved as the number of epochs rose. These findings imply that the YOLOv8 model was successfully optimized for real-world traffic sign identification tasks using the selected training configuration.

The results demonstrate how important TSDR has been to the development of driver-assistance systems (ADAS) and autonomous vehicles (AVs). For these intelligent transporta- tion systems to provide road safety and effective navigation, accurate and fast detection of traffic signs is essential.

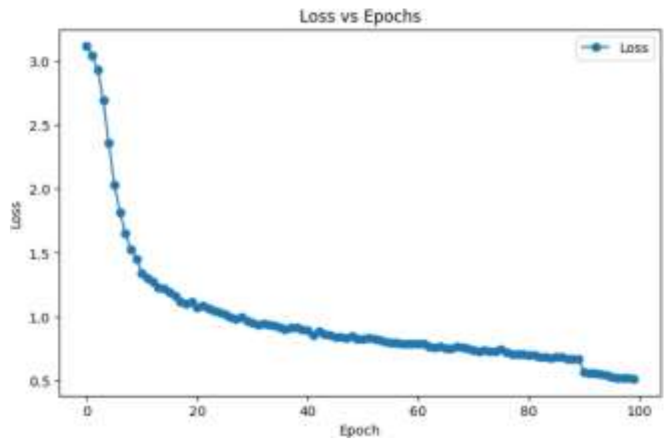


Fig. 13. Loss vs EpochsTABLE II
OVERALL PERFORMANCE

Model Name	Accuracy (%)	Recall (%)	F1-score
YOLOv8	**95.2**	**98**	**0.92 at 0.603**

This work admits certain limitations, including the model's sensitivity to severely obstructed or congested settings and possible differences in class-wise accuracy. Future objectives for research include examining lightweight YOLOv8 variants for effective deployment on embedded ADAS/AV platforms, improving model interpretability for better error analysis, and investigating methods for better class-imbalance handling.

We can help develop more robust, generalizable, and ex- plainable TSDR systems, which will ultimately pave the way for safer and more dependable autonomous transportation, by tackling these limitations and exploring new research directions.

B. Future Scope:

The future scope of traffic sign recognition and detection us-ing YOLOv8 is vast and exciting. Some potential applications and areas of research include:

- Autonomous vehicles: Traffic sign recognition and de- tection can play a crucial role in enabling autonomous vehicles to navigate safely and efficiently.
- Traffic management: Real-time traffic sign detection and classification can help traffic management authorities to monitor and manage traffic flow more effectively.
- Driver assistance systems: Traffic sign recognition and detection can be used to enhance driver assistance sys- tems, such as lane departure warning and adaptive cruise control.
- Surveillance: Traffic sign detection and classification canbe used for surveillance purposes, such as monitoring traffic flow and detecting anomalies.
- Improved accuracy: Future research can focus on improv-ing the accuracy of traffic sign recognition and detection,particularly in low-light or adverse weather conditions.
- Multi-modal sensing: Integrating traffic sign recognition and detection with other sensing modalities, such as LiDAR and radar, can provide more comprehensive and accurate traffic information.

Overall, traffic sign recognition and detection using YOLOv8 has the potential to revolutionize the way we travel and improve road safety for all users.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. W. Yang and W. Zhang, "Real-time Traffic Signs Detection Based on YOLO Network Model," 2020 International Conference on Cyber- Enabled Distributed Computing and Knowledge Discovery (Cy- berC), Chongqing, China, 2020, pp. 354-357, doi: 10.1109/Cy- berC49757.2020.00066.
2. N. Barnes and A. Zelinsky, "Real-time radial symmetry for speed sign detection," IEEE Intelligent Vehicles Symposium, 2004, Parma, Italy, 2004, pp. 566-571, doi: 10.1109/IVS.2004.1336446.

3. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, ”You Only Look Once: Unified Real-Time Object Detection”, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
4. M. A. Garcia-Garrido, M. A. Sotelo and E. Martin-Gorostiza, ”Fast traf- fic sign detection and recognition under changing lighting conditions,” 2006 IEEE Intelligent Transportation Systems Conference, Toronto, ON, Canada, 2006, pp. 811-816, doi: 10.1109/ITSC.2006.1706843.
5. J. Redmon and A. Farhadi, YOLOv3: An Incremental Improvement, 2018.
6. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, ”You Only Look Once: Unified Real-Time Object Detection”, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
7. J. Redmon and A. Farhadi, ”YOLO9000: Better Faster Stronger”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517-6525, 2017.
8. Paulo and P. L. Correia, ”Automatic Detection and Classification of Traffic Signs”, Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS ‘07), pp. 11-11, 2007.
9. R. Girshick, ”Fast R-CNN”, IEEE International Conference on Com- puter Vision, pp. 1440-1448, 2015.
10. Zhang, Z.W., Gao, Y., Wang, J., et al. (2020) Deep Learning Traffic Sign Recognition System Based on YOLOv3. Building Electricity, 39, 64-68. (In Chinese)
11. Li, Z., Zhang, H.H. and Deng, J.Y. (2021) Traffic Sign Detec- tion Algorithm Based on Improved Faster R-CNN. Chinese Jour- nal of Liquid Crystals and Displays, 36, 484-492. (In Chinese) https://doi.org/10.37188/CJLCD.2020-0195
12. Luo, K.T. and He, Q.L. (2022) Chinese Traffic Sign Detection Based on YOLOv5 Algorithm. Operational Research and Fuzzy Science, 12, 1570-1584. (In Chinese) https://doi.org/10.12677/ORF.2022.124165
13. Liu, Y.C., Shao, Z.R. and Hoffmann, N. (2021) Global Attention Mech- anism: Retain Information to Enhance Channel-Spatial Interactions. ArXiv: 2112.05561.
14. Chen, Z. and Cheng, Y.Y. (2022) Improved Traffic Sign Detection Based on FCOS Algorithm. Computer Engineering and Design, 43, 3271-3278. (In Chinese)
15. Tian, Z., Shen, C.H., Chen, H. and He, T. (2019) FCOS: Fully Con- volutional One-Stage Object Detection. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, 27 October-2 November 2019, 9626-9635. https://doi.org/10.1109/ICCV.2019.00972
16. Jocher, G., Chaurasia, A. and Qiu, J. (2023) YOLO by Ultralytics. https://github.com/ultralytics/ultralytics
17. Wang, C.-Y., Bochkovskiy, A. and Liao, H.-Y.M. (2022) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. ArXiv: 2207.02696.
18. Liu, S., Qi, L., Qin, H.F., Shi, J.P. and Jia, J.Y. (2018) Path Aggregation Network for Instance Segmentation. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 18-23 June 2018, 8759-8768. https://doi.org/10.1109/CVPR.2018.00913
19. Li, H.G., Yu, R.N. and Ding, W.R. (2021) Research Development of Small Object Traching Based on Deep Learning. Acta Aeronautica et Astronautica Sinica, 42, 24691-024691. (In Chinese)
20. Zhu, Z., Liang, D., Zhang, S., et al. (2016) Traffic-Sign Detection and Classification in the Wild. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 27-30 June 2016, 2110-2118. https://doi.org/10.1109/CVPR.2016.232
21. Yang, G.Y., Lei, J., Zhu, Z.K., et al. (2023) AFPN: Asymptotic Feature Pyramid Network for Object Detection. ArXiv: 2306.15988. [13] Liu, S.T., Huang, D. and Wang, Y.H. (2019) Learning Spatial Fusion for Single-Shot Object Detection. ArXiv: 1911.09516
22. Li, W. Song, L. Xiao, Y. Hu and X. Pan, “Salient traffic sign video detection based on hypercomplex frequency domain,” Proceedings of the 33rd Chinese Control Conference, Nanjing, 2014, pp. 7379-7382.
23. Krizhevsky, I. Sutskever and G. Hinton, “ImageNet Classifica- tion with Deep Convolutional Neural Networks,” Curran Associates Inc.2012.
24. R. Girshick, J. Donahue, T. Darrell and J. Malik, “Rich Feature Hi- erarchies for Accurate Object Detection and Semantic Segmentation,” 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 580-587.
25. S. Ren, K. He, R. Girshick and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, June 2017, pp. 1137-1149.
26. M. Liao, C. Wang and A. Bochkovskiy, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv:2004.10934, 2020.
27. J. Zhang, W. Wang, C. Lu, J. Wang and A. K. Sangaiah, “Lightweight deep network for traffic sign classification,” Annals of Telecommunica-tions, 2019.
28. J. Zhang, M. Huang and X. Jin, “A real-time Chinese traffic sign detection algorithm based on modified YOLOv2,” Algorithms, 2017.