



## Research on Traffic Sign Detection Based on Computer Vision

Yuhu Shi<sup>1,\*</sup>, Shufang Lv<sup>1</sup>, Ting Yan<sup>1</sup> and Wangxin Li<sup>1</sup>

<sup>1</sup> Information Technology Institute of Shanxi Finance & Taxation College, Taiyuan 030024, Shanxi, China

**SUMMARY:** *In order to tackle the difficulties presented by large - scale fluctuations, densely packed small targets, and category imbalance in traffic sign detection within complex urban traffic situations, this article puts forward an enhanced detection framework for traffic sign feature augmentation. This framework is built upon the YOLOv8n baseline model. When dealing with the CCTSDB dataset, the Star attention module and the ASFF adaptive multi - scale feature fusion module are incorporated. These modules enhance feature representation for occluded and small targets. For the TT100K dataset, coordinate attention and a four-head detection structure are adopted, alongside EMA weight smoothing to boost generalization. Tests on CCTSDB and TT100K show the refined model achieves 0.982 mAP@0.5 and 0.876 mAP@0.5:0.95 on CCTSDB. On TT100K, mAP@0.5 increases from 0.775 to 0.808, with the best F1 score of 0.75. Its performance approaches Faster R-CNN while preserving YOLO's real-time advantage. This method provides an efficient and reliable traffic sign detection solution for vehicle-assisted driving.*

**KEYWORDS:** *Traffic sign detection; YOLOv8; attention mechanism; fusion of features across multiple scales; real-time detection*

## 1 Introduction

Identifying traffic signs is one of the main perception functions for autonomous driving and intelligent driving support systems. The accuracy and real - time performance of this identification are directly related to driving safety. Thanks to the rapid progress of deep learning technology, object detection algorithms based on convolutional neural networks have achieved remarkable progress in the field of traffic sign recognition. However, traffic signs in real complex traffic scenarios are often faced with multiple challenges, including large scale variation, dense small objects, unbalanced category distribution, complex illumination conditions and occlusion interference. Current general detection models are unable to attain an ideal equilibrium between detection precision and inference effectiveness. The high - performance detection of traffic signs in complex scenarios still stands as a crucial issue that demands immediate resolution [1].

Taking YOLOv8n as the baseline model, this paper proposes an improved detection framework with feature enhancement for traffic signs. In order to meet the detection criteria for three types of signs in the CCTSDB dataset, the Star attention mechanism and the ASFF adaptive multi - scale feature fusion elements are utilized to enhance the feature representation of occluded and tiny targets. Regarding the TT100K dataset, which has a large number of categories and a significant quantity of small targets, the coordinate attention

\*syheducn@163.com

<https://doi.org/10.65102/is2026978>



For the promotion of multi-scale detection work, research workers put their emphasis on attention mechanism approaches and feature fusion methods. In the year 2026, Xiao together with Zhao put forward YOLO-TCS, that puts together multi-layer feature mixing and attention modules for promoting compatibility to multi-size traffic signs and obtains hopeful outcomes on public data sets. Wei et al. (2025) have made design of Autotrinet YOLO, it is a triple attention framework which can strengthen feature representation from channel, spatial and scale dimensions, thus it improves robustness for complex backgrounds. In the aspect of feature fusion, Xu and Xiao (2026) have put forward PST-YOLO, which uses polynomial progressive scale fusion to make effective aggregation of multi-scale features. It promotes the detection of small objects and meanwhile keeps a structure which is of light weight. Liang and his research companions (2025) have put forward YOLO-DPDG which possesses a dual-pooling dynamic grouping mechanism. This one is specially optimized toward small-scale and long-distance traffic signs, therefore it effectively alleviates the problem of ultra-small target missing detection. Ren Hongge together with his work partners (2025) have brought forward the HCF - YOLO model. Through utilizing mixing channel combination and assisting bounding box return methods, the location accuracy of multi - scale objects is further promoted. The aforementioned researches prove that attention mechanism and multi-scale feature integration are effective methods to enhance traffic sign detection effect. Even so, the current methods still have space to promote in the study of adaptive fusion weight and the cooperative optimization of many detection heads.

Under the push of demands for model light weight and real-time deduction, the research about traffic sign examination for special scenes such as bad weather and complex light environment has also obtained obvious advancement. In the year of 2026, the researcher Ziyin Zhang and his work companions put forward the TS-YLO arithmetic method. This algorithm is the one that is designed for the detection of small traffic signs in bad driving situations like rainy and foggy weathers. This method strengthens the importance of features in images that have low quality, and hence promotes the stability of detection in environments that are extreme. Qing Guo and other researchers (2025) have put forward the YOLO-SDW model. Based on skip connection and dynamic convolution design, it realizes model lightweighting while maintaining detection accuracy, and adapts to the computing resource limitations of vehicle-mounted terminals [13]. In 2025, Peizhi Wang and colleagues developed the YOLO - AEF framework. By implementing adaptive enhancement and fusion tactics, this framework addresses the degradation issue of traffic signs in demanding situations. Moreover, it efficiently mitigates the interference arising from changes in illumination and partial obstruction [14]. Shengsheng Tu et al. (2025) proposed the TSS-YOLO model. Focusing on the demand for fast and timely detection of small traffic signs, it achieves stable detection at a high frame rate by simplifying the network structure and optimizing the detection head design [15]. These studies have verified the feasibility of optimizing the YOLO model under specific constraints. However, most of the existing work focuses on the detection of signs in a single scenario or a specific category, and still lacks sufficient adaptability to large-scale traffic sign datasets with numerous categories and extremely unbalanced sample distribution, such as TT100K.

## 3 Methodology

### 3.1 Baseline YOLOv8 Detection Model

YOLOv8 is a prevalent single - stage object detection algorithm within the YOLO family. It boasts excellent detection precision and inference velocity, and satisfies the deployment

criteria for real - time traffic sign detection on in - vehicle terminals [16]. The entire model is composed of three key modules:

**Backbone:** The CSPDarknet architecture is employed. By means of convolution and C2f modules, multi - scale feature extraction is carried out. This enables the efficient separation of shallow detailed features and deep semantic features.

**Neck:** It applies the PAN-FPN structure to complete top-down semantic feature transmission and bottom-up detail feature fusion, and outputs multi-scale feature maps.

**Head:** The system employs a decoupled detection head framework to isolate the classification and regression branches. At the same time, it utilizes the Anchor - Free detection model, which lessens the complexity of fine - tuning the hyperparameters of the anchor boxes and boosts the adaptability for detecting small objects [17].

This paper selects YOLOv8n as the baseline model. The core basic formulas of its detection task are as follows:

(1) Intersection over Union, IoU (Core Evaluation Metric)

It is used to measure the overlap degree between the predicted bounding box and the ground-truth bounding box, and serves as the core computational basis for the localization accuracy of object detection. The formula is defined as follows:

$$IoU = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|} \quad (1)$$

where: The predicted bounding box of the model is indicated, the manually annotated ground - truth bounding box is represented, and the pixel area within the bounding box is referred to.

(2) Total Loss Function

The training loss of YOLOv8 is composed of three elements: loss for bounding box regression, loss for classification, and distribution focal loss. The main formula is presented as follows:

$$L_{total} = \omega_1 L_{boz} + \omega_2 L_{cls} + \omega_3 L_{dfj} \quad (2)$$

where:

The bounding box regression loss is denoted, and it serves to improve the localization precision of the predicted bounding boxes.

The classification loss is represented, and it is utilized to enhance the recognition accuracy of traffic sign categories.

$L_{dfj}$  represents the distribution focal loss, which is applied to improve the localization stability of small and blurry targets;

$\omega_1, \omega_2, \omega_3$  represents the weight coefficient of each loss term. In this paper, the default settings of YOLOv8 are adopted, which are 7.5, 0.5 and 1.5 respectively.

## 3.2 Design of Attention Module for Traffic Sign Feature Enhancement

Addressing the issues of intense background interference and the challenge of extracting effective features from traffic signs in intricate traffic situations, this article incorporates a lightweight attention mechanism to boost the feature output of the backbone network. It suppresses invalid background features and strengthens the effective semantic and detailed features of traffic signs. Different attention modules are adapted according to the characteristics of the two datasets respectively [18].

### (1) Star Attention Module

In order to fulfill the detection needs of three kinds of traffic signs within the CCTSDB

dataset, this research paper incorporates the Star attention module to refine the backbone network. It realizes feature interaction in channel and spatial dimensions through a multi-branch structure, and enhances the feature extraction ability for occluded and small-scale traffic signs[19]. The core feature enhancement formula is as follows:

$$F_{out} = F + \alpha \cdot M_e \cdot F + \beta \cdot M_s \cdot F \quad (3)$$

where:

The input feature map is denoted, and the enhanced output feature map is represented.

$M_e$  represents the channel attention weight, which is used to strengthen the valid channel features related to traffic signs;

$M_s$  represents the spatial attention weight, which is used to accurately locate the spatial position of traffic signs and suppress background interference;

$\alpha, \beta$  represents the learnable weight coefficient, which is used to balance the contribution of the two attention branches.

(2) Coordinate Attention Module

Considering the features of the TT100K dataset, such as a large number of categories, slight disparities among fine - grained classes, and a high percentage of small targets, this paper presents the CA coordinate attention module. It embeds positional information into channel attention to accurately capture the coordinate position and subtle features of traffic signs[20]. Its core feature enhancement formula is as follows:

$$F_{out} = F \cdot M_h \cdot M_w$$

where:

The input feature map is represented, and the enhanced output feature map is denoted.

$M_h$  represents the vertical attention weight, which is used to capture the height position information of the target;

$M_w$  represents the horizontal attention weight, which is used to capture the width position information of the target.

### 3.3 ASFF Adaptive Multi-scale Feature Fusion Module

Addressing the issues of significant scale variations in traffic signs and the low detection precision of small and medium - sized targets, this research incorporates the ASFF (Adaptive Spatial Feature Fusion) module during the model optimization process for the CCTSDB dataset. Instead of the conventional feature splicing operation used in the baseline model, this module is employed to tackle the problem of inadequate multi - scale feature fusion [21]. The key formula for adaptive weighted fusion is presented as follows:

(1) Fusion Weight Normalization

$$\alpha^l = \frac{\exp(w^l)}{\exp(w^1) + \exp(w^2) + \exp(w^3)} \quad (5)$$

where:  $l \in \{1,2,3\}$  represents three feature maps of different scales,  $w^l$  denotes the weight parameter of the  $l$ -th layer feature learned by the model,  $\alpha^l$  represents the normalized fusion weight, which satisfies  $\alpha^1 + \alpha^2 + \alpha^3 = 1$ .

(2) Multi-scale Feature Fusion

$$F_{\text{ASFF}} = \alpha^1 \cdot \tilde{F}^1 + \alpha^2 \cdot \tilde{F}^2 + \alpha^3 \cdot \tilde{F}^3$$

where:  $\tilde{F}^l$  denotes the  $l$ -th layer feature map after spatial size alignment,  $F_{\text{ASFF}}$  denotes the final fused output feature map. The model can automatically learn the optimal fusion weights of multi-scale features, strengthen the feature representation of small and medium targets, and improve the detection robustness across all scales.

### 3.4 Detection Head Structure Optimization

To address the difficult detection of ultra-small targets and long-tailed categories in the TT100K dataset, this paper extends the original 3-detection-head structure of the baseline model to a 4-detection-head structure, so as to realize full-coverage detection of traffic signs at all scales[22]. The core multi-scale prediction formula is as follows:

$$P_k = \text{Head}_k(F_k), \quad k \in \{1,2,3,4\}$$

where:

The number indicates the serial number of the detection head. The newly introduced detection head is designed for ultra - tiny targets and is associated with the characteristic featuring downsampled high - frequency information. Meanwhile, the detection heads for small, medium, and large targets are represented by, and respectively.

$F'_k$  denotes the input feature map of the corresponding scale,  $\text{Head}_k(\cdot)$  denotes the prediction network of the  $k$ -th detection head,  $P'_k$  denotes the output detection results, including category, bounding box and confidence score.

The newly added detection head for ultra-small targets can completely retain the detailed features of tiny traffic signs. It efficiently addresses the issue of undetected ultra - tiny targets in the baseline model. Moreover, it mitigates the problem of low detection precision that results from an inadequate number of samples for long - tailed categories.

### 3.5 EMA Weight Smoothing Training Strategy

To improve the generalization ability of the model and alleviate the overfitting problem during training, this paper introduces the EMA (Exponential Moving Average) weight smoothing strategy to optimize the weight update in the model training process. This strategy does not increase the model inference time and is compatible with real-time detection requirements[23]. The core weight update formula is as follows:

$$\theta_t^{\text{ema}} = \gamma \cdot \theta_{t-1}^{\text{ema}} + (1 - \gamma) \cdot \theta_t$$

where:

$\theta_t$  denotes the instantaneous model weight after the  $t$ -th training iteration,  $\theta_t^{\text{ema}}$  denotes the EMA smoothed weight updated after the  $t$ -th iteration,

$\gamma$  denotes the smoothing coefficient, which is set to 0.9998 in this paper. A larger value leads to smoother weight updates and stronger anti-noise capability.

With this strategy, the final inference weights of the model are the weighted average of historical weights throughout the training process, rather than the instantaneous weights at the end of training. It has the ability to efficiently mitigate the oscillations during the training process and enhance the model's ability to generalize in unencountered testing situations.

## 4 Experimental Results and Analysis

### 4.1 Experimental Environment and Dataset Construction

All experiments were conducted on a unified hardware and software platform. The hardware includes an Intel Xeon Silver 4214R CPU, an RTX 3090 GPU with 24GB VRAM, and 64GB RAM. The software environment runs on Ubuntu 20.04 with CUDA 11.7 and cuDNN 8.4, using PyTorch 1.12.0 and Python 3.8.

Two public traffic sign detection datasets are adopted in the experiments, namely CCTSDB (Chinese City Traffic Sign Database) and TT100K (Tsinghua-Tencent 100K). The CCTSDB data set consists of approximately 15,000 traffic scene pictures. These pictures can be divided into three main categories of traffic signs: warning, limiting, and guiding. These images possess a broad scope of resolution values. Furthermore, they put forward complex problems such as non-uniform illumination and object blocking. The TT100K dataset contain approximately 100000 pictures and include over 200 kinds of traffic signs, which has the features of obvious category unbalance and great scale change. According to the official division, the two data sets are all divided into a training set, a verification set, and a test set. The proportion of these collections is approximately seven shares for the training collection, two shares for the validation collection, and one share for the test collection.

### 4.2 Experimental Parameter Settings and Evaluation Metrics

The table below presents the experimental parameters. All models adopt identical settings for a fair comparison. We use the momentum-based SGD optimizer with a momentum of 0.937 and weight decay of  $5e-4$ . The learning rate adopts cosine annealing decay, starting at 0.01 and gradually decreasing to 0.0009. All models are trained for 50 epochs with a batch size of 16. Input images are resized to  $640 \times 640$  pixels, and Mosaic augmentation is applied to improve generalization.

Table 1 Experimental parameter configuration

Parameter Name	Parameter Value
Optimizer	SGD
Momentum	0.937
Weight Decay	5.00E-04
Initial Learning Rate	0.01
Learning Rate Strategy	Cosine Annealing
Training Epochs	50
Batch Size	16
Input Image Size	$640 \times 640$
Data Augmentation	Mosaic, random flipping and others

The evaluation tools use the conventional indices often employed in object detection. These include Accuracy, Recall Rate, F1-measure, and average mean Precision ( $mAP@0.5$  and  $mAP@0.5:0.95$ ).  $mAP@0.5$  represents the mean average precision in the situation that the Intersection over Union (IoU) threshold gets set to 0.5. On another hand,  $mAP@0.5:0.95$  represents the average value of  $mAP$  values which are calculated under different IoU threshold conditions. These threshold values lie from 0.5 to 0.95, with an increase of 0.05 between every step. This group of indicators can overall evaluate the model's abilities on the aspect of locating and classifying.

### 4.3 Ablation Experiments and Result Analysis

Taking the baseline YOLO model as the benchmark, this section conducts basic performance

verification and error analysis on the CCTSDB dataset, so as to provide a quantitative comparison standard for the ablation experiments of each subsequent improved module. The core result analysis is presented as follows.

As shown in the indicator curves in Figure 2, the optimal overall F1-score of the baseline model for all categories reaches 0.94 at the confidence threshold of 0.676. Within the confidence range from 0.2 to 0.8, the F1-scores of prohibition, indication and warning traffic signs all remain stably at a high level, and decline rapidly only when the confidence exceeds 0.8. The model precision rises continuously with the increase of confidence, and the overall precision reaches 1.00 at the confidence of 0.990, which proves that the prediction results with high confidence have extremely high reliability and an extremely low false detection risk.

The results of the Precision-Recall (PR) curve indicate that the overall mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.5 for the model attains 0.979. Specifically, the mAP for prohibition signs is 0.984, for indication signs is 0.977, and for warning signs is 0.976. The PR curve approaches the top - right corner of the coordinate system, which reflects a good equilibrium between high precision and high recall. The recall rate reaches 1.00 at the confidence of 0.000 and decreases slowly with the rising confidence. Its sharp drop in the high-confidence interval also reflects the missed detection risk of the model under strict threshold conditions.

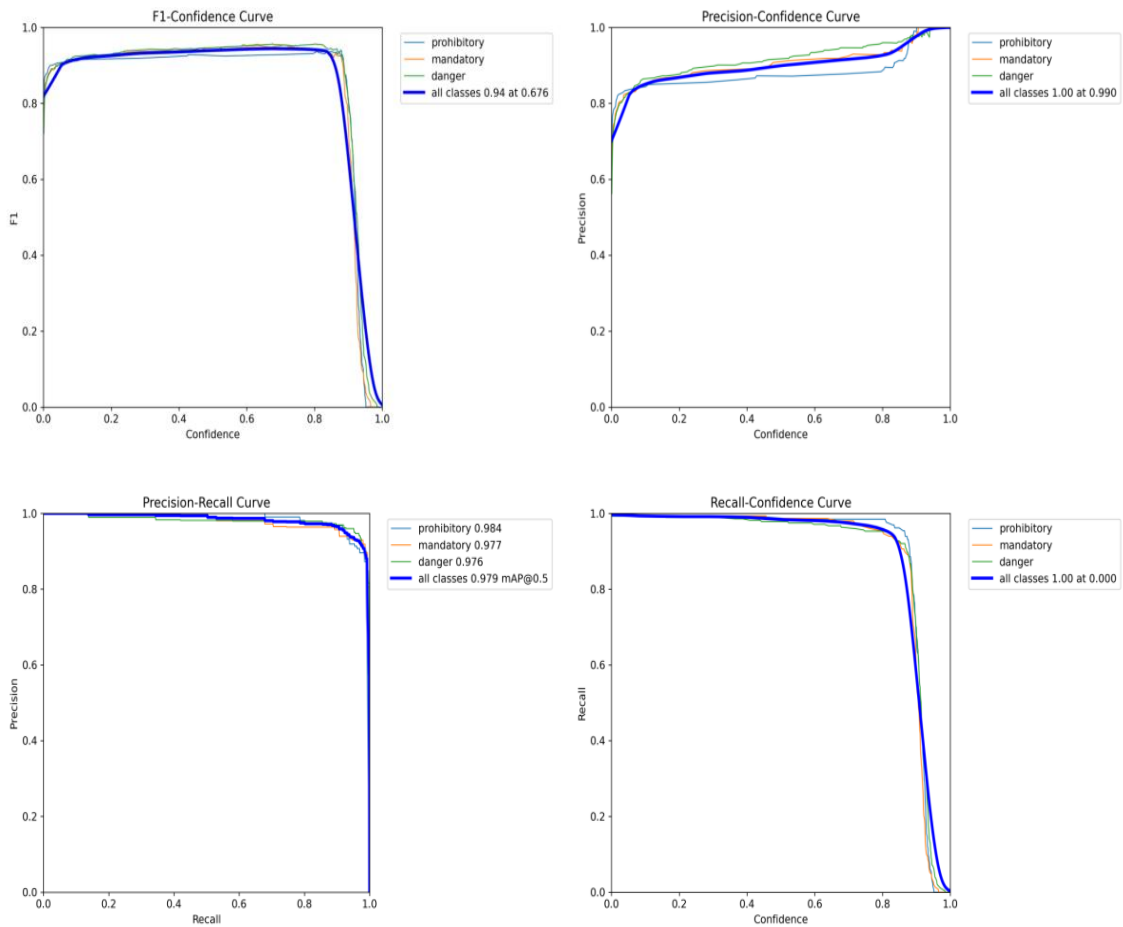


Figure 2: Confidence curve of the baseline YOLO traffic sign detection model

The confusion matrix results in Figure 3 demonstrate that the classification accuracy of the three types of traffic signs all reaches 99%, with no cross misdetection between categories.

Only a small number of targets are misjudged as the background, among which warning signs have the highest background misdetection rate of 0.44 and indication signs have the lowest rate of 0.26. Few background samples are incorrectly classified as targets, indicating that the model has clear category classification boundaries, and its core error mainly stems from target-background misjudgment in complex scenarios. The statistical distribution of dataset samples shows that the dataset is dominated by medium and small-scale targets, with the target aspect ratio concentrated in the range of 0 to 0.3. The number of warning sign samples is significantly larger than that of prohibition signs. The unbalanced sample distribution and the high proportion of small targets are the core restrictive factors affecting the model performance.

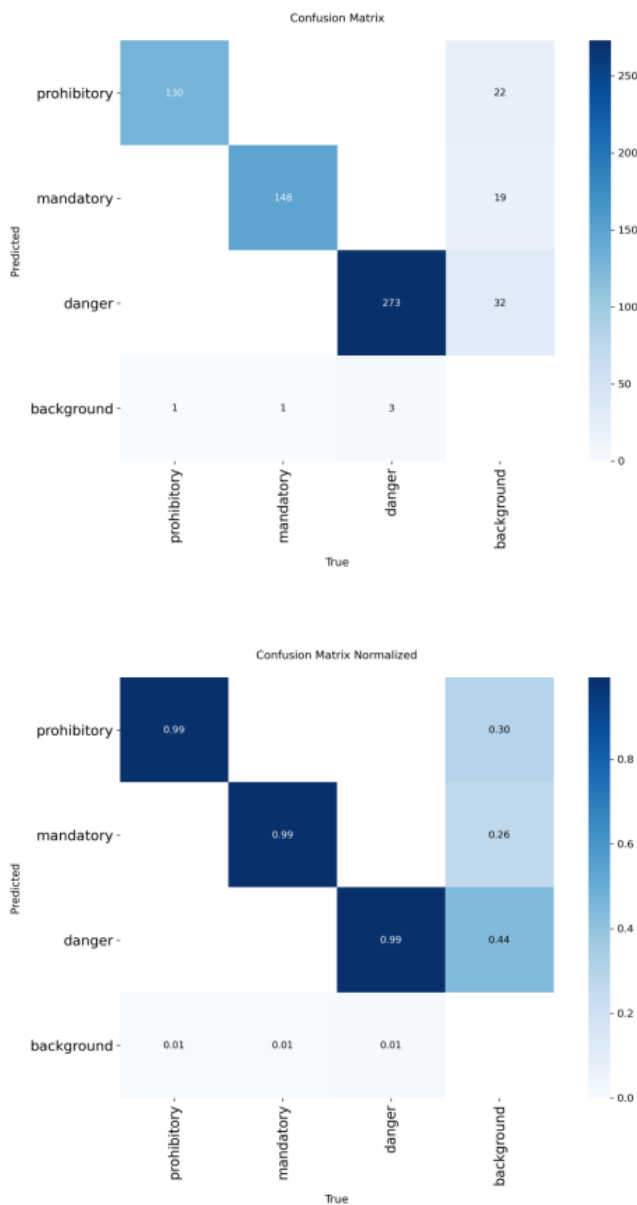


Figure 3: Confusion matrix of detection results for the baseline YOLO traffic sign detection model

The training convergence curves presented in Figure 4 indicate that the three distinct categories of loss on both the training dataset and the validation dataset steadily decrease

throughout the model training process. After 40 epochs, these losses start to reach a stable state, and there are no apparent signs of overfitting. Meanwhile, the mAP indicators rise steadily. The ultimate mAP@0.5 reaches a convergence value of 0.979, and the mAP@0.5:0.95 converges to 0.872. The model demonstrates excellent convergence and generalization capabilities. Moreover, there remains a significant scope for enhancing its accuracy. This situation offers a well - defined path for the subsequent optimization of the module.

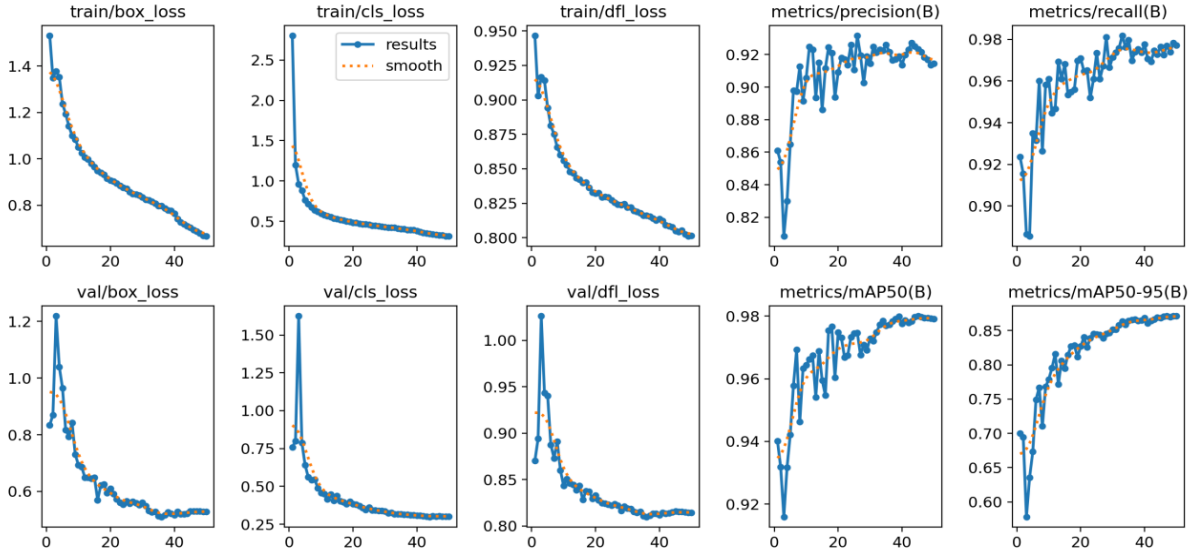


Figure 4: Training convergence curves of the baseline YOLO traffic sign detection model

The fundamental model validates the essential performance of the YOLO algorithm within the traffic sign detection task. In addition, it points out with exactitude its weak points, including the absence of enough ability to extract features in complicated situations, and the limited accuracy of detection for objects with multiple scales. To quantify the performance gain of each improved module, this section conducts ablation experiments under the single-variable control method to verify the actual effect of each optimization strategy.

The overall detection performance of the improved model is shown in Figure 5. Compared with the baseline model, the optimal overall F1-score of all categories rises from 0.94 to 0.95, and the corresponding confidence threshold increases from 0.676 to 0.815. It is demonstrated that the enhanced model is capable of sustaining outstanding comprehensive detection ability even under more stringent confidence constraints. Moreover, the dependability of the high - confidence prediction outcomes is significantly boosted. The model precision reaches 1.00 at the confidence of 0.998, which is further optimized compared with the baseline threshold of 0.990, and the suppression ability against false detection samples is effectively improved. The outcomes of the Precision - Recall (PR) curve indicate that the overall mean Average Precision at an Intersection over Union (IoU) threshold of 0.5 (mAP@0.5) for the enhanced model attains 0.982. This represents an increment of 0.003 when compared to the baseline model. Moreover, the Average Precision (AP) values for indication signs and warning signs both climb to 0.981. This method has the efficacy to solve the problem that the baseline model has insufficient performance on these two special categories. At a confidence level of 0.000, the retrieval rate attains 0.99. Moreover, the rate of decline in the high - confidence range is notably diminished. This effectively mitigates the issue of undetected objects under intricate lighting and occlusion circumstances.

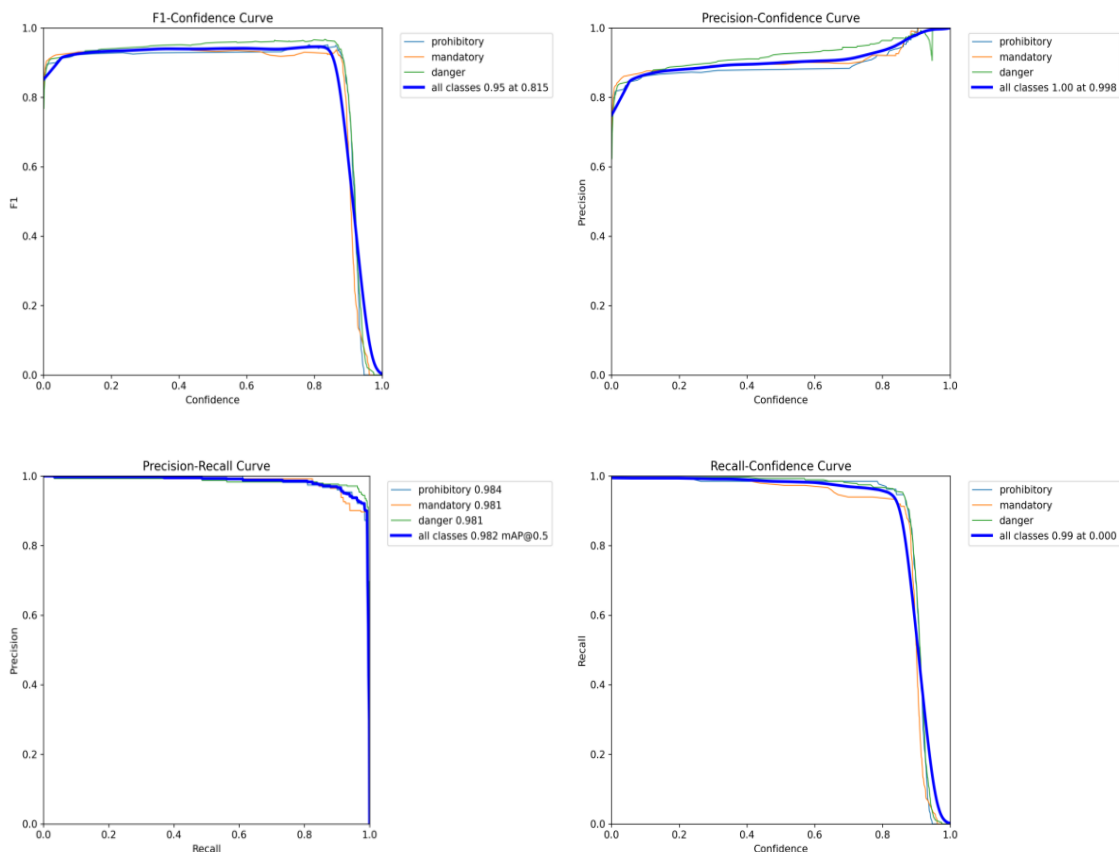


Figure 5: Confidence curves of the improved YOLO traffic sign detection model

The outcomes of the confusion matrix presented in Figure 6 indicate that the enhanced YOLO model sustains a classification precision of over 98% for all three kinds of traffic signs. Specifically, the classification precision of warning signs attains 100%, and there are no misdetections across different categories. The background misdetection rate of prohibition and indication signs is obviously reduced, and the number of background samples misclassified as targets is further decreased. The model achieves a prominent improvement in the discrimination accuracy between foreground targets and complex backgrounds. The improved strategies in this paper are designed according to the distribution characteristics of the dataset. By means of attention optimization and multi - scale feature integration, the suggested approach efficiently mitigates the challenges of category disparity and small - target recognition.

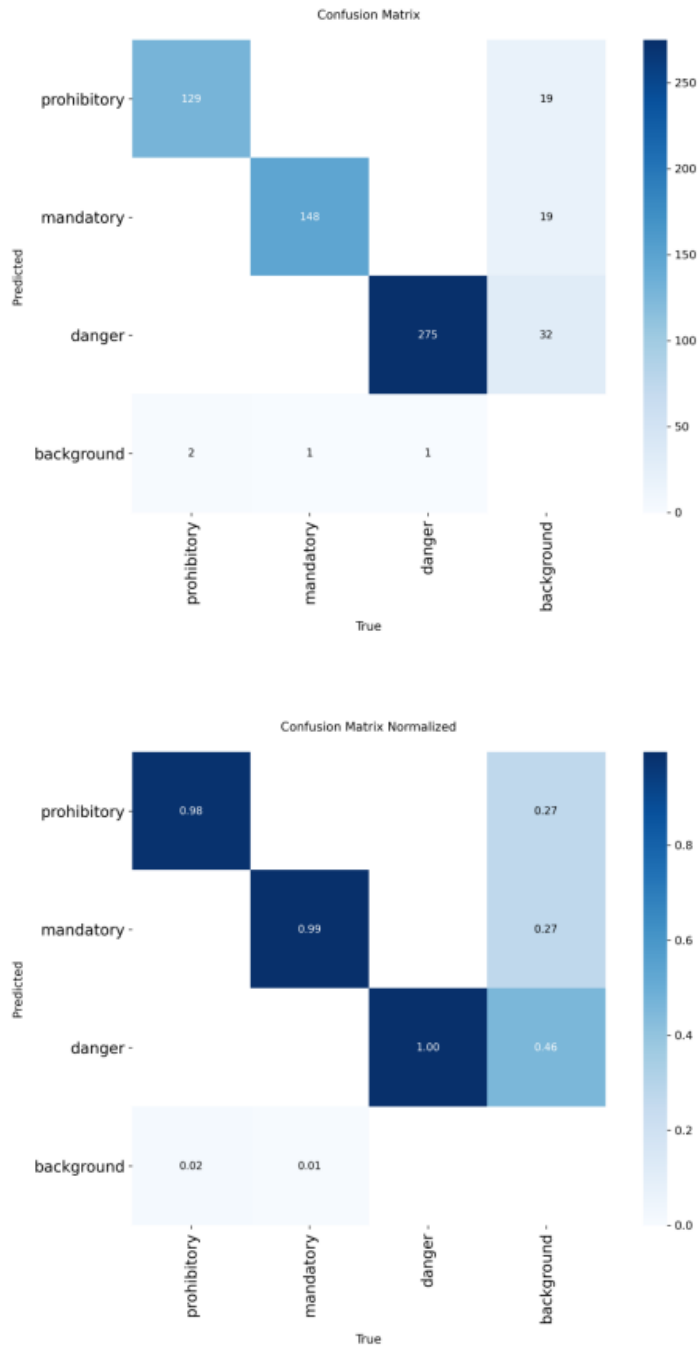


Figure 6: Confusion matrix of detection results for the improved YOLO traffic sign detection model

As can be observed from the training convergence curves presented in Figure 7, the enhanced model attains a more rapid rate of loss reduction and converges stably following 40 training epochs. There is no obvious oscillation or overfitting for the validation set loss. Meanwhile, the model acquires a higher final mAP value with a more stable convergence process, which fully verifies that the optimized modules effectively improve the model's convergence efficiency and generalization ability.

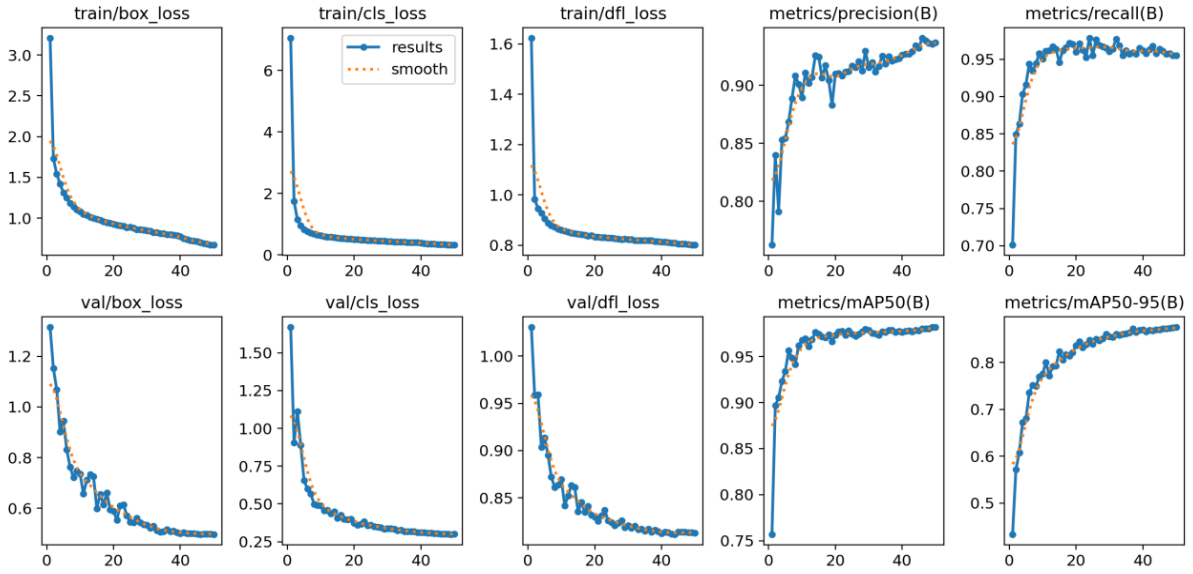


Figure 7: Training convergence curves of the improved YOLO traffic sign detection model

To quantify each module’s contribution, ablation experiments were performed under single-variable control. The effectiveness of each optimization was validated on CCTSDB and TT100K, with key results shown in Tables 2 and 3.

Ablation results show that on the CCTSDB dataset, integrating the Star attention module increases mAP@0.5 from 0.979 to 0.981, verifying its effectiveness in refining low-quality backbone features. The EMA strategy further lifts mAP@0.5 to 0.982 and boosts generalization. Adding the ASFF module raises mAP@0.5:0.95 to 0.876, a 0.004 improvement over the baseline. The fully enhanced model surpasses traditional Faster R-CNN.

On the category-rich and imbalanced TT100K dataset, the baseline yields an mAP@0.5 of 0.775. Incorporating the CA module, EMA strategy, and four detection heads improves mAP@0.5 to 0.808 and mAP@0.5:0.95 to 0.679, approaching Faster R-CNN. This demonstrates the proposed strategies’ adaptability and robustness for complex urban traffic.

Table 2: Detection Performance Results of Ablation Experiments on CCTSDB Dataset

	map@50	map@50:59
baseline	0.979	0.872
+star	0.981	0.874
+star+ema	0.982	0.872
+star+ema+asff	0.982	0.876
faster-rcnn	0.649	0.454

Table 3: Detection Performance Results of Ablation Experiments on TT100K Dataset

	map@50	map@50:59
baseline	0.775	0.642
+ca	0.778	0.644
+ca+ema	0.785	0.662
+ca+ema+4head	0.808	0.679
faster-rcnn	0.815	0.682

#### 4.4 Comparative Experiments and Result Analysis of Multiple Models

To verify the generalization and practical value of the improved YOLO model for complex multi-class traffic sign detection, multi-model comparisons are conducted on the TT100K dataset to evaluate its performance in real urban scenarios. The experimental results are presented in Figures 17–24.

According to the core curves in Figure 8, the improved YOLO model achieves an optimal overall F1-score of 0.75 on the TT100K dataset at a confidence threshold of 0.413. The thin gray lines denote subclass F1 curves, showing minor category differences while maintaining a high overall performance. The precision-confidence curve shows the precision reaches 1.00 at a confidence of 1.000, indicating no false detections for high-confidence predictions and strong false-alarm suppression.

The precision-recall curve yields an overall mAP@0.5 of 0.808, outperforming the baseline YOLO (0.775) and approaching Faster R-CNN (0.815). The recall-confidence curve shows a recall of 0.95 at a confidence of 0.000, reflecting a low miss rate at low thresholds. Recall declines steadily without sharp drops as confidence increases, proving stable performance across thresholds and suitability for practical deployment.

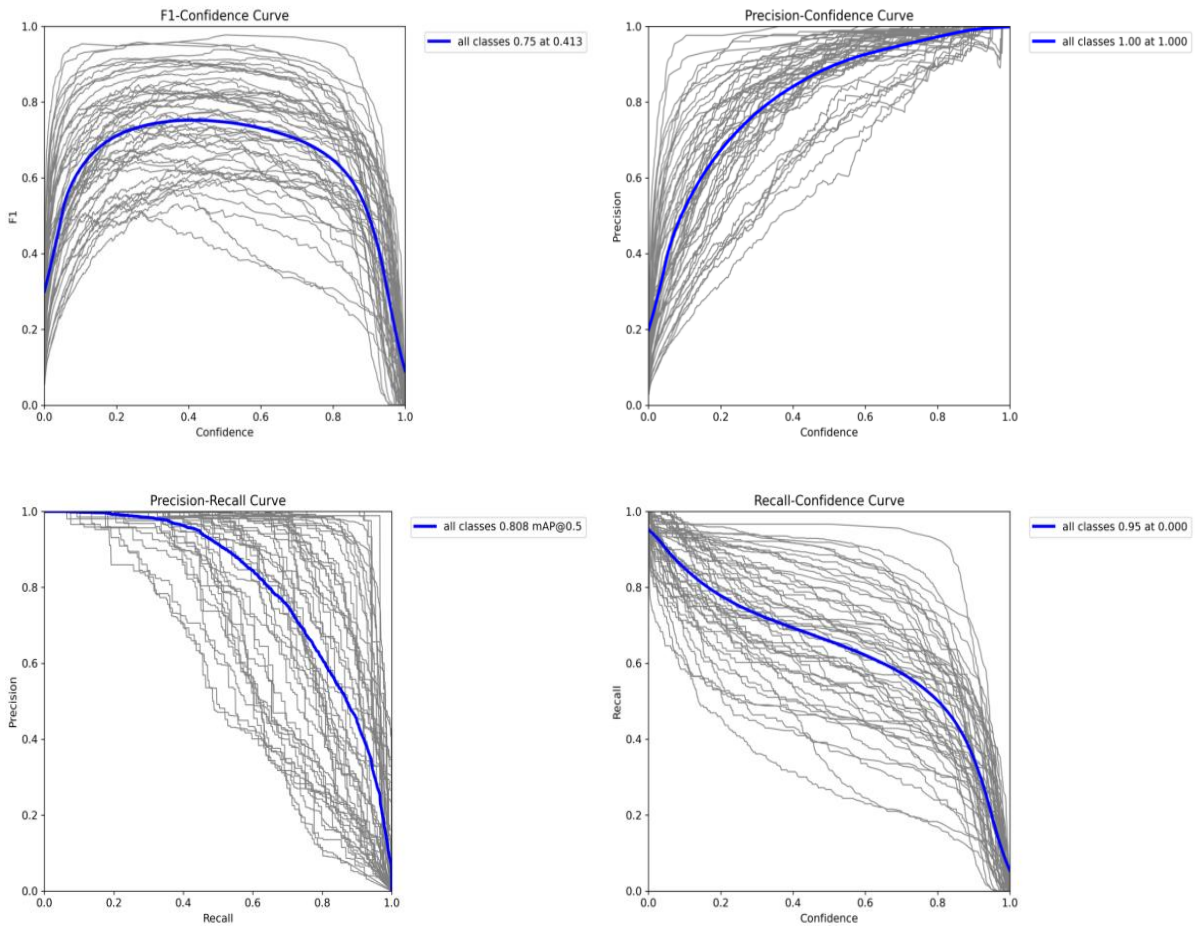


Figure 8: Confidence curves of the improved YOLO traffic sign detection model on TT100K dataset

The confusion matrix outcomes in Figure 9 can intuitively give people the model's classification performance that exists in multi-class situations. The TT100K dataset contains hundreds of fine-grained traffic sign categories, which suffers from high inter-class similarity





To sum up, the multi - model comparison experiments conducted on the TT100K dataset comprehensively confirm the efficacy and preeminence of the enhanced YOLO algorithm presented in this paper. When dealing with intricate scenarios characterized by multiple categories, long - tail sample distribution, and densely packed small targets, the model attains an optimal F1 - score of 0.75. Moreover, its mAP@0.5 shows a 3.3 - percentage - point increase compared to the baseline model, and its overall detection performance is nearly on par with that of Faster R - CNN. The confusion matrix proves its high classification accuracy for multiple categories, with only minor errors existing in a few long-tail categories. The training convergence curves show that the model maintains a stable training process without overfitting, and obtains greatly improved convergence efficiency and generalization ability. The improved strategies can well adapt to the detection challenges of complex traffic scenarios. It greatly boosts the detection accuracy while guaranteeing real-time inference speed, which can provide high-performance technical support for the vehicle-mounted traffic sign detection system.

## 5 Conclusion

To tackle the issues in urban traffic sign detection, such as significant scale changes, a large number of small targets in high density, and category imbalance, this paper presents an enhanced feature - boosted detection framework founded on YOLOv8n. Dataset - specific optimizations are implemented. For the CCTSDB dataset, the Star attention mechanism enhances the features of occluded and small targets. Instead of the traditional splicing method, the Adaptive Spatial Feature Fusion (ASFF) is used for adaptive multi - scale feature fusion. Regarding the TT100K dataset, the coordinate attention mechanism captures accurate location details. Moreover, the number of detection heads is increased from three to four to enhance the perception of ultra - small targets. The Exponential Moving Average (EMA) strategy is additionally employed to reduce training fluctuations. These promotion methods effectively promote the model's ability to pick up characteristics and find objects in complicated circumstances.

The proposed method is comprehensively verified through ablation and comparative experiments on CCTSDB and TT100K. On the CCTSDB dataset, the enhanced model attains a mean average precision (mAP) of 0.982 at an intersection over union (IoU) threshold of 0.5 and 0.876 at an IoU range from 0.5 to 0.95. This represents an increase of 0.003 and 0.004 respectively compared to the baseline model. At a relatively low confidence threshold, the model achieves a precision of 1.00, accompanied by a significantly reduced decline in recall. When applied to the more difficult TT100K dataset, the model achieves an optimal F1 - score of 0.75. The mAP at an IoU threshold of 0.5 improves from 0.775 to 0.808, showing a 3.3% increase, and the mAP at an IoU range from 0.5 to 0.95 reaches 0.679, with a 3.7% increase. The performance of this model is on par with that of Faster R - CNN, while still retaining the real - time processing advantage of YOLO. Confusion matrix analysis shows that the classification accuracy of the model remains above 98% in multi-class scenarios, the background false detection rate is significantly reduced, and the confusion between fine-grained categories is effectively alleviated.

The advanced frame that is proposed by this paper greatly promotes the detection accuracy of traffic signs in complex traffic conditions, and thus keeps the real-time detection work efficiency unchanged. This thus provides an efficacious and reliable technical scheme for the auxiliary drive systems of vehicles. But, the robustness of this model under extreme weather conditions which include rain, fog, and weak light in night time still requires further promotion. Furthermore, the detection precision of long-tail classes in the TT100K dataset

also needs to be further promoted. In the future, the research work will therefore investigate lighter attention mechanisms and data augmentation methods. Semi-supervised learning and knowledge distillation technique are introduced by us to relieve the problem of class unbalance, therefore to push the development of traffic sign detection technique toward lower calculation power consumption and higher environment adaptation ability.

## Funding

In 2024, this research was backed by the Science and Technology Innovation Project for Higher Education Institutions in Shanxi Province, specifically the project titled "Investigation into Traffic Sign Detection Leveraging Computer Vision" (Project No. 2024L540).

## Author's Profile

Yuhu Shi was born in Jinzhong City, Shanxi Province, China, in 1995. He holds a Master's degree from Taiyuan University of Technology and is currently affiliated with the Information Technology Institute at Shanxi Finance & Taxation College. His primary research interests lie in the fields of machine learning and image processing. syheducn@163.com

Shufang Lv was born in Luliang City, Shanxi Province, China, in 1988. She holds a Master's degree from Taiyuan University of Technology and is currently affiliated with the Information Technology Institute at Shanxi Finance & Taxation College. Her primary research interests focus on the applications of Internet of Things (IoT) technology and blockchain technology. 760944527@qq.com

Ting Yan was born in Xiaoyi City, Shanxi Province, China, in 1990. She holds a Master's degree from Taiyuan University of Technology and is currently affiliated with the Information Technology Institute at Shanxi Finance & Taxation College. Her primary research interest centers on data mining. 327335860@qq.com

Wangxin Li was born in Taiyuan, Shanxi Province, China in 1994. He holds a Master's degree from Xiangtan University and is currently affiliated with the Information Technology Institute at Shanxi Finance & Taxation College. His main research interest is in intelligent information processing. lwx360708@163.com

## References

- [1] Yang L X, Cui S, Zhang H R, Ma Y J. Traffic Sign Detection Algorithm Based on HRM-YOLO[J]. *Modern Electronic Technology*, 2026, 49(07): 190-198.
- [2] Weiguo Pan, Songjie Du, Bingxin Xu, Bin Zhang, Hongzhe Liu. ES-YOLO: Edge and Shape Fusion-Based YOLO for Traffic Sign Detection[J]. *Computers, Materials & Continua*, 2026, 87 (1):
- [3] Jiang H J, Ding J F, Li W. Small Traffic Sign Detection in Complex Environment Based on YOLOv8n[J]. *Journal of Dalian Minzu University*, 2026, 28(01): 27-33.
- [4] Li P F, Xiong Z X, Wang G B. YOLOv8-NTS: A Target Recognition Method for Traffic Sign Detection[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(12): 1868-1880.

- [5] Wang Q S, Qiao Y P, Li J E. Lightweight Traffic Sign Detection Model Based on Improved YOLOv8s[J]. Journal of Liaoning University (Natural Science Edition), 2025, 52(04): 329-337.
- [6] Zhiwei Zheng, Jin Cheng, Fanghua Jin. TFP-YOLO: Obstacle and Traffic Sign Detection for Assisting Visually Impaired Pedestrians[J]. Sensors (Basel, Switzerland), 2025, 25(18): 5879-5879.
- [7] Xiao Yu, Xin Zhao. YOLO-TCS: an enhanced multi-scale network for traffic sign detection integrating multi-level feature fusion and attention[J]. Multimedia Systems, 2026, 32(2): 110-110.
- [8] Wei Yi Dong, Ya Xuan Chen, Guang Yu Zou. Autotrinet YOLO triple attention framework for robust traffic sign detection[J]. Scientific reports, 2025, 15(1): 42247-42247.
- [9] Yili Xu, Xuanxuan Xiao. PST-YOLO: A lightweight, robust approach for traffic sign detection via polynomial-based and progressive scale fusions[J]. Digital Signal Processing, 2026, 169: 105745-105745.
- [10] Ruishi Liang, Minjie Jiang, Shuaibing Li. YOLO-DPDG: A Dual-Pooling Dynamic Grouping Network for Small and Long-Distance Traffic Sign Detection[J]. Applied Sciences, 2025, 15(20): 10921-10921.
- [11] Hongge Ren, Hairui Song, Haiqiang Liu, Anni Fan, Yingying Tan. HCF-YOLO: a high-performance traffic sign detection model with hybrid channel fusion and auxiliary box regression[J]. Journal of Real-Time Image Processing, 2025, 22(5): 191-191.
- [12] Ziyin Zhang, Bingyin Feng, Hao Liu. TS-YOLO: a small traffic sign detection algorithm for various harsh driving conditions in bad weather[J]. Scientific reports, 2026, 16(1): 4444-4444.
- [13] Qing Guo, Juwei Zhang, Bingyi Ren. YOLO-SDW: Traffic Sign Detection Algorithm Based on YOLOv8s Skip Connection and Dynamic Convolution[J]. Computers, Materials & Continua, 2025, 86(1): 1-20.
- [14] Peizhi Wang, Raihani Mohamed, Norwati Mustapha, Noridayu Manshor. YOLO-AEF: Traffic sign detection on challenging traffic scenes via adaptive enhancement and fusion[J]. Neurocomputing, 2025, 655: 131430-131430.
- [15] Shengsheng Tu, Mingming Kong, Chao Zhang. TSS-YOLO: A timely swift small traffic sign detection model[J]. Signal, Image and Video Processing, 2025, 19(12): 1029-1029.
- [16] Yuanyuan Dang, Hemin Wang, Yifeng Fan. WAC-YOLO: a real-time traffic sign detection model for complex weather with feature degradation suppression and occlusion awareness[J]. Journal of Real-Time Image Processing, 2025, 22 (4): 155-155.
- [17] Jun Li, QinWen Deng, WenXin Gao, Bing Yang, Lan Jia, Ju Zhou, HaiBo Pu. DSF-YOLO for robust multiscale traffic sign detection under adverse weather conditions.[J]. Scientific reports, 2025, 15 (1): 24550.

- [18] Ji Qiu, Wenbo Zhang, Siyuan Xu, Huiyu Zhou. DP-YOLO: A lightweight traffic sign detection model for small object detection[J]. *Digital Signal Processing*, 2025, 165: 105311-105311.
- [19] Yan Deng, Lidong Huang, Xinsu Gan, Youfang Lu, Shunxiang Shi. A heterogeneous attention YOLO model for traffic sign detection[J]. *The Journal of Supercomputing*, 2025, 81 (6): 765-765.
- [20] Luo X L, Lv W X, Shi Z Y, Liu R C. Lightweight Traffic Sign Detection Algorithm Based on Improved YOLOv8n[J]. *Laser & Optoelectronics Progress*, 2025, 62(12): 422-433.
- [21] Wang B, Xu H H, Sun D C, Yu Y F. Small-target Traffic Sign Detection Based on Improved YOLOv8 Algorithm[J]. *Journal of Computer Applications*, 2024, 44(S2): 274-279.
- [22] Huang Z Y, Fang Qiu, Guo X H. YOLOv8-TS Traffic Sign Detection Network Fused with Attention Mechanism[J]. *Modern Electronic Technology*, 2025, 48(01): 179-186.
- [23] Song J J, Zhang Z L. Small-target Traffic Sign Detection Algorithm Based on YOLOv8n[J]. *Science Technology and Engineering*, 2024, 24(36): 15548-15557.