A Lightweight Improved Traffic Sign Detection Algorithm Based on YOLOv8

Bozhang Liu,Ye Tao*,Wenhua Cui

Abstract—**With the rapid development of artificial intelligence technology, traffic sign recognition has become a critical component of the environmental perception system in intelligent driving, attracting increasing attention. However, the complexity of real world conditions—such as variable weather, changing lighting, and background interference—continues to pose significant challenges to the accuracy and robustness of recognition systems. To address these issues, this paper proposes an innovative, ultra lightweight traffic sign detection model based on YOLOv8n, aiming to achieve more efficient and accurate detection in complex environments. The model incorporates several structural optimizations: an Efficient Multi Scale Attention (EMA) module is introduced at the end of the backbone to significantly enhance key feature extraction; a lightweight GSConv convolution is adopted to greatly reduce computational complexity and memory consumption; the pooling structure is improved to boost detection speed; and an additional small object detection head is added to improve multi scale feature fusion and enhance the detection accuracy for small targets. The proposed model is evaluated on two representative datasets, TT100K and CCTSDB 2021. Experimental results show that, compared with the original YOLOv8n, the model reduces parameters by 0.95M and model size by 0.72M, while achieving mAP improvements of 7.2% and 1.8% on the two datasets, respectively. This model strikes an excellent balance between accuracy and real time performance, outperforming the original YOLOv8n and demonstrating strong practical value. It offers a more reliable and efficient solution for traffic sign detection in intelligent transportation systems.**

Index Terms—**Traffic sign detection, Lightweight, YOLOv8n, GSConv, Attention mechanism**

I. Introduction

THE technology of traffic sign detection and recognition is a crucial component of autonomous driving systems. Its core objective is to accurately extract key information such as road speed limits, driving directions, and construction zones to ensure the safe and reliable operation of autonomous vehicles. However, in practical applications, the accuracy and reliability of traffic sign detection are often affected by varying weather conditions, complex road environments, and various potential interference factors. Additionally, the high complexity of deep learning models results in low operational efficiency in resource constrained environments such as in vehicle systems and mobile terminals. Therefore, designing an efficient traffic sign detection algorithm that is adaptable to various complex scenarios has become a key research focus.

Currently, traffic sign detection methods can be broadly categorized into two types: traditional feature based methods and deep learning based methods. Traditional methods primarily rely on manually defined features such as color and shape for traffic sign recognition. While these methods have lower computational costs, they are susceptible to variations in lighting, weather conditions, and occlusions, leading to weaker generalization capabilities. In contrast, deep learning based detection methods can automatically extract deep features of traffic signs, offering greater adaptability and detection accuracy, making them the mainstream approach. Since Redmon [1] first introduced the YOLO (You Only Look Once) detection algorithm, it has undergone rapid development in recent years due to its significant advantages in both detection accuracy and speed. This evolution has led to the emergence of various versions, with YOLOv8 currently being widely adopted for its accurate and fast detection performance. YOLOv8 consists of five models of different sizes, determined by scaling factors. Among them, YOLOv8n is a smaller version with fewer network layers and parameters. Although its detection accuracy is lower than that of larger models, it offers higher detection speed. However, in complex traffic scenarios, factors such as multi scale variations of traffic signs, background interference, and sample imbalance can lead to false detections and inconsistencies in recognition performance. To address these issues, this paper proposes an enhanced traffic sign detection method based on the YOLOv8n model.

Since Redmon [1] first introduced the YOLO (You Only Look Once) detection algorithm, it has undergone rapid development in recent years due to its significant advantages in both detection accuracy and speed. This evolution has led to the emergence of various versions, with YOLOv8 currently being widely adopted for its accurate and fast detection performance. YOLOv8 consists of five models of different sizes, determined by scaling factors. Among them, YOLOv8s is a smaller version with fewer network layers and parameters. Although its detection accuracy is lower than that of larger models, it offers higher detection speed. However, in complex traffic scenarios, factors such as multi scale

Manuscript received March 28, 2025; revised June 28,2025.

This work was supported by National Natural Science Foundation of China (62272093),Social Science Federation project of Liaoning Province (2025lslybwzzkt-100),and Postgraduate education and teaching reform research project of Liaoning Province (LNYJG2024092).

Bozhang Liu is a postgraduate student majoring in computer technology at School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (e-mail: liubz_1998@163.com).

Ye Tao* is an Associate Professor at the School of Computer Science and Software Engineering, University of Science and Technology LiaoNing, Anshan 114051, China (Corresponding author to provide e-mail: taibeijack@163.com).

Wenhua Cui is a professor of School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan,114051, China (e-mail: taibeijack@126.com).

variations of traffic signs, background interference, and sample imbalance can lead to false detections and inconsistencies in recognition performance. To address these issues, this paper proposes an enhanced traffic sign detection method based on the YOLOv8n model.

II. Related Work

In recent years, researchers have proposed various deep learning based traffic sign detection methods to improve detection accuracy and efficiency. For example, Rishabh [2] et al. introduced a multi branch, multi task convolutional neural network (CNN) architecture that enables simultaneous traffic sign detection and classification, accelerating detection speed while maintaining accuracy. Li et al. optimized Faster R-CNN by integrating the ResNet50 network with an attention-guided contextual feature pyramid network, enhancing the detection capability for small traffic signs in complex backgrounds.

Xiang Xinjian et al. applied preprocessing techniques such as image space transformation, illumination compensation, and brightness adjustment to enhance traffic sign image quality, combining these with Mask R-CNN to improve detection accuracy. Additionally, Jing Fangke et al. proposed a dual head detection structure tailored for small objects, effectively capturing small target features while reducing model parameters. Saxena [3] et al. optimized the PANet structure of the YOLOv4 neck network and introduced feature scales specifically for small object detection, enabling the algorithm to better adapt to complex road environments. Mahadshetti [4] et al. incorporated YOLOv7, the SE module, and attention mechanisms to enhance the model's ability to capture the salient features of traffic signs. Wu Mengmeng et al. proposed an adaptive feature enhancement object detection network called YOLO-AFENet, which significantly improved the detection accuracy of small objects.

In optimizing the YOLO series models, Luo Yutao [5] et al. introduced a channel attention calibration module and combined it with a dual path enhancement structure to optimize the prediction branch. They also employed the K means++ clustering algorithm to improve YOLOv5s detection capability for small objects. Zhao et al. proposed a lightweight feature extraction backbone network and optimized the ESGBlock structure using an attention mechanism to reduce computational complexity. Zhu Ningke et al. improved the inverted residual structure in MobileNetv3 and applied it to the YOLOv5 backbone network, enhancing the model's lightweight efficiency. Cao et al. proposed a lightweight backbone network based on GhostNet, effectively reducing the parameter count and model size of YOLOv5s. Peng [6] et al. introduced a lightweight context aware traffic sign detection network into YOLOv8n to reduce network computational costs.

Additionally, Liu Fei et al. replaced the original YOLOv5s backbone network with BoTNet (Bottleneck Transformer Network) and designed a lightweight network called C3GBneckv2. Integrating the Ghostv2 Bottleneck and channel attention mechanisms enhanced the model's feature extraction capability. Lan Hong [7] et al. proposed an efficient multi scale feature pyramid fusion network, MPANet, which improves computational speed by downscaling shallow feature maps and incorporating lightweight modules such as BBot and C2fGhost.However, despite the significant progress in deep learning based traffic

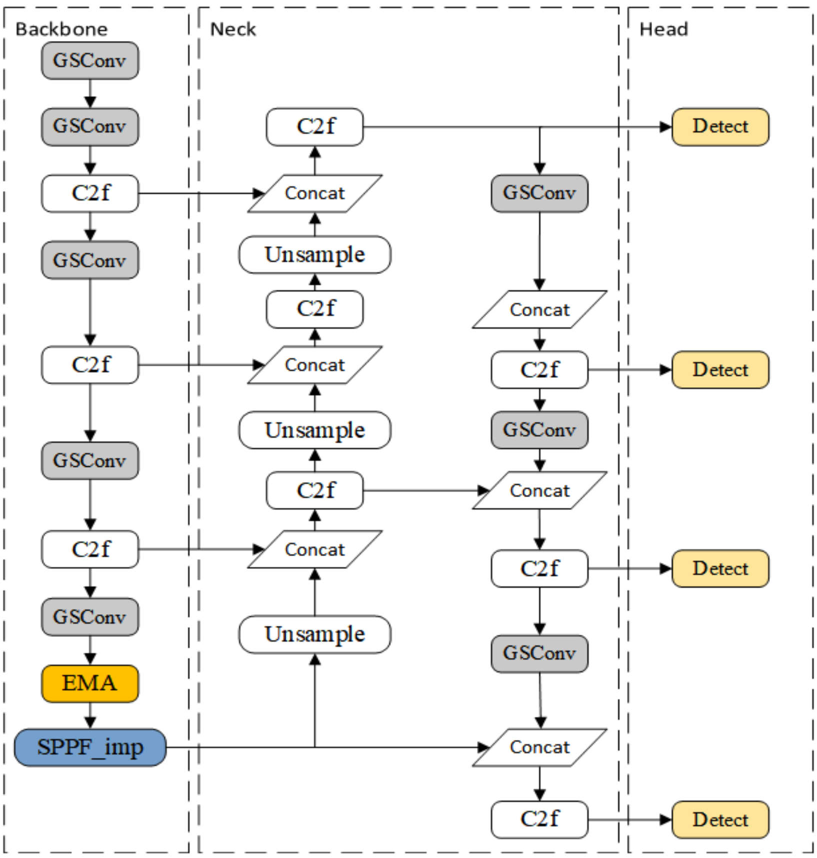


Fig. 1. Improved YOLOv8 network structure diagram

sign detection, several challenges remain in complex scenarios. These include susceptibility to occlusions, background interference, and high computational costs, necessitating further optimization and improvement. Deep learning based traffic sign detection eliminates the need for manual feature extraction. By training on large amounts of labeled data, deep learning models learn nonlinear functions that transform images into a feature space where a linear classifier can quickly distinguish categories, enabling accurate traffic sign recognition.

The proposed algorithm is an improved version of YOLOv8n. As a groundbreaking advancement in object detection, YOLOv8 has achieved significant progress in balancing real time detection accuracy and computational efficiency through systematic architectural reconstruction and multi dimensional algorithmic innovations. It has made breakthroughs in optimizing the tradeoff between computational efficiency and detection accuracy.

By reconstructing the backbone network and feature fusion mechanisms, YOLOv8 introduces the C2f (Cross Stage Partial Context Fusion) structure, which enhances gradient propagation efficiency and multi scale feature representation. At the training paradigm level, YOLOv8 innovatively employs a dynamic sample weighting mechanism, which adaptively adjusts the ratio of positive and negative samples and the confidence threshold, mitigating the common issue of class imbalance in object detection. Architecturally, this model integrates object detection, instance segmentation, and keypoint detection into a unified framework. Its hierarchical design allows for the generation of sub models with varying complexity (Nano/Small/Medium/Large/XLarge) by adjusting depth and width factors, achieving a linear tradeoff between parameter size and inference speed. In the MS COCO benchmark, the YOLOv8 X version attained an average precision (AP) of 53.9%, an 8.2% improvement over previous models, while maintaining an end to end processing speed of under 30ms, demonstrating its applicability in resource constrained environments. The training process incorporates multi scale hybrid augmentation strategies, a self optimizing anchor mechanism, and task oriented loss function design, exhibiting strong robustness against occlusions and small object detection challenges. In terms of open source ecosystem support, YOLOv8 provides a comprehensive model compression toolchain and multi platform deployment solutions. Its technological capabilities have been successfully applied in cutting edge fields such as intelligent medical image analysis, autonomous drone navigation, and precision industrial inspection, offering innovative solutions for the engineering deployment of lightweight visual perception systems.

The main contributions of this paper are as follows:

Integration of the EMA [8] attention mechanism into the YOLOv8 backbone. By reshaping certain channels into the batch dimension and grouping the channel dimension into multiple sub features, this approach preserves channel information while reducing computational overhead. The EMA module recalibrates channel weights by encoding global information and captures pixel level relationships through cross dimensional interactions, effectively addressing challenges caused by complex backgrounds in traffic scenarios.

• Replacing standard convolution (Conv) with GSConv [9]. This lightweight convolution module enhances the receptive field[10], while reducing the number of parameters, improving the efficiency of large kernel convolutions and overall detection performance.

• Modifying the pooling layer structure. The redesigned pooling mechanism expands the model's receptive field, enhances robustness, and improves the integration of multi scale features.

• Adding a small object detection head. This enhancement strengthens the model's ability to detect small objects, improving detection performance in complex traffic environments.

This study conducted comprehensive experimental evaluations on the improved YOLOv8n lightweight traffic sign detection algorithm using two widely recognized datasets: TT100K and CCTSDB2021.[11] The results demonstrate that the proposed method not only significantly reduces computational complexity and model size but also maintains, and in some cases improves, detection accuracy across diverse scenarios. This balance between performance and efficiency makes the algorithm particularly well suited for deployment in resource constrained computing environments, such as embedded systems or edge devices in intelligent transportation systems.

III. LISTLIGHT AND IMPROVED YOLOV8N ALGORITHM

Figure 1 shows the architecture of the new generation object detection model proposed in this study, which realizes the comprehensive performance exceeding YOLOv8n[12]

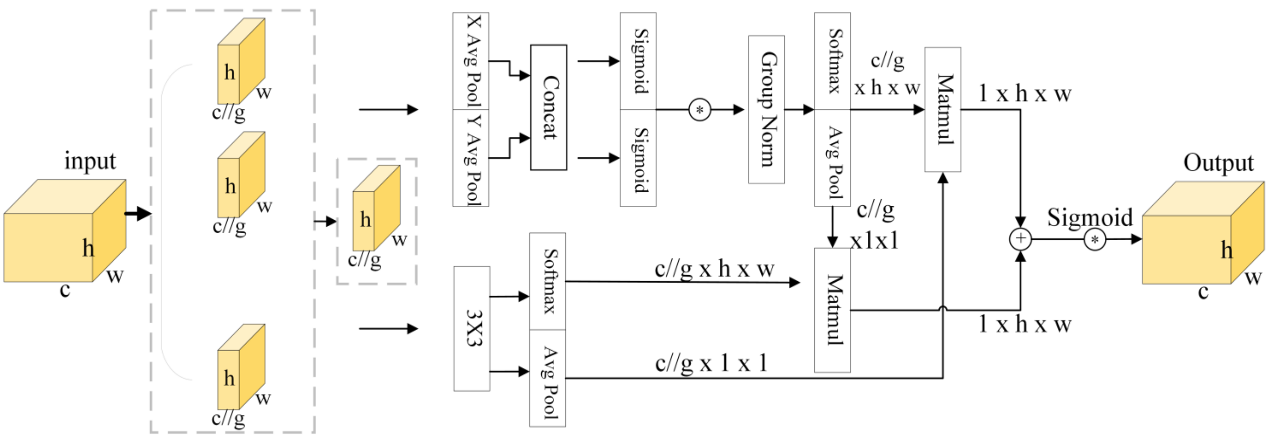


Fig. 2. EMA attention network structure diagram

through multi dimensional technological innovation. In terms of architecture design, the model innovatively integrates four core modules: EMA attention module based on two dimensional dynamic channel space calibration, which enables the network to adaptively enhance the contrast of occlocclted objects and background by constructing feature importance weight graphs[13],[14],[15] and significantly improve the feature focusing capability in complex lighting and dense scenes; The improved SPPF_imp module adopts the multi level gated void pooling strategy and introduces a learnable feature screening mechanism based on the traditional spatial Pyramid structure not only preserves the fusion efficiency of cross scale context information but also optimizes the feature representation ability of different scale targets through dynamic receptive field regulation. [16]The lightweight GSConv operator designed for edge computing needs, through the hybrid architecture of packet convolution and channel recombination, reduces the computing load by 55% while maintaining the semantic integrity of deep features by using dynamic kernel parameter sharing technology, providing an efficient reasoning basis for mobile terminal deployment. The cascade detector system designed for small objects constructs a complete detection link from feature retention to precise location through high resolution shallow feature guidance, cross scale feature pyramid two-way interaction, and prediction frame optimization algorithm based on probability density estimation. All modules form closed loop optimization through a cooperative working mechanism: EMA module preferentially selects key semantic features, SPPF_imp extends multi scale sensing domain and eliminates feature redundancy, GSConv implements dynamic computational resource allocation during feature transfer, and finally, small target detection head completes fine-grained feature decoding.[17],[18],[19] This architecture design not only realizes the strong adaptability of the model in complex scenarios but also significantly improves the processing ability of challenging scenarios such as large scale differences, high target density, and strong background interference in detection tasks under the premise of maintaining real time reasoning speed through intelligent optimization of the calculation path. It provides a new generation of solutions with high efficiency and robustness for automatic driving environment perception, industrial precision parts inspection, and remote sensing image analysis.

A. Ema attention mechanism

The Efficient Multi Scale Attention (EMA) module is a novel Attention mechanism based on improved Coordinate Attention, which aims to enhance feature representation through collaborative modeling of channel dimensions and spatial dimensions. Unlike traditional attention mechanisms, EMA is capable of capturing feature dependencies at both a local and global scale in an image, providing a more refined representation of objects in complex scenes. Especially in the traffic sign detection task, the dataset usually contains a large number of targets with a small scale and is easily affected by background interference, which makes the model susceptible to the impact of background noise during detection[20], resulting in the reduction of recognition accuracy. In order to solve this problem, the EMA module is introduced into the YOLOv8n backbone network to enhance the feature expression ability of the model for small targets. Specifically, by building long and short range feature dependencies[21], EMA enables the model to more accurately capture the subtle features of traffic signs, effectively separating targets from interference information in complex contexts. The introduction of this module not only improves the detection accuracy of the model but also improves the adaptability of the model to different scale targets while maintaining the computational efficiency and ensuring the robustness and reliability of the traffic sign detection task. The specific structure of EMA is shown in the figure.

B. Improvement of the pooling layer

In object detection tasks, the spatial pyramid pool fast module (SPPF) is often used to extract multi scale features and enhance the model's ability to perceive objects of different sizes. However, the original SPPF of YOLOv8 mainly relies on the 5×5 maximum pooling operation of a fixed size, and only obtains the characteristics of different receptive fields through multiple pooling, which lacks the full utilization of global information. Therefore, we propose an improved version of the SPPF module (SPPF_imp), which retains the original SPPF structure and introduces an adaptive pooling mechanism to enhance the global feature extraction capability. Specifically[22], SPPF_imp first reduces dimension by 1×1 convolution to reduce computational effort and improve computational efficiency. Then, when performing multiple 5×5 maximum Pooling to extract local features of different scales, Adaptive Max Pooling and Adaptive Avg Pooling are introduced to obtain the most significant features and global average information, respectively. These features from different sources are spliced in the channel dimension, and after 1×1 convolution fusion, the enhanced feature expression is finally generated. Compared with the original SPPF, SPPF_imp can integrate local and global information more comprehensively, improve the robustness of the model to the shape and scale changes of the target, and thus improve the detection accuracy and generalization ability. Figure 3 shows the structure of the improved sppf.

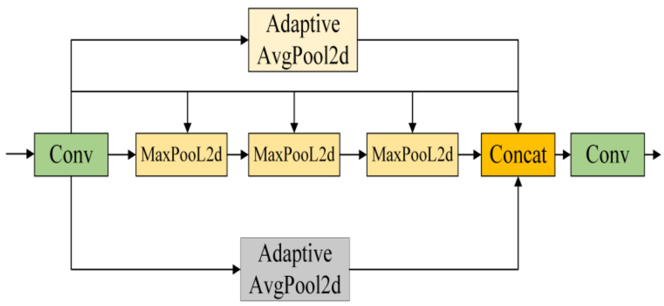


Fig. 3. SPFF Improvements

C. GSConv

GSConv[9] convolution is an improved convolution structure consisting of standard convolution, depth separable convolution, and shuffle operation. It is designed to improve model performance and detection accuracy while maintaining computational efficiency and accelerating convergence during training. Its overall structure is shown in the figure 3.

Assume that the number of channels for the input feature is C_{in}, and the number of channels for the output feature is C_{out}. First, the input features are processed by the standard convolution layer, and the number of channels is reduced to $C_{out}/2$ while the basic features are extracted in order to reduce the computation amount and improve the computation efficiency. Then, the features after dimensionality reduction are sent into the deep separable convolution layer, where channel by channel convolution is used to extract local features of each channel, while point by point convolution is responsible for integrating information between channels to further enrich the feature expression ability. Then, the features extracted by depth separable convolution are fused with the features directly output by the standard convolution layer to enhance the complementarity of information and retain more spatial structure information and semantic information.

After the merged features are shuffled, the channels are rearranged to improve cross channel information exchange. Finally, the output features with the number of channels as Cout are generated. This design fully combines the powerful capability of standard convolution in feature extraction with the advantage of deep separable convolution in computational efficiency so that GSConv can not only improve the expressibility of the model under limited computational resources but also achieve a reasonable balance between the number of parameters and the amount of computation, and finally improve the overall performance and applicability of the model.

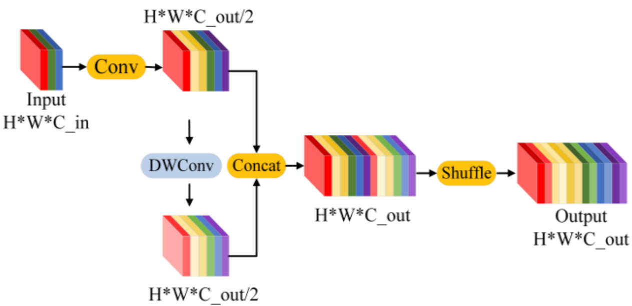


Fig. 4. GSConv structure diagram

D. Small target detection head

Aiming at the detection challenge of small traffic signs (10×10 pixels) in vehicle scenes, this study carried out multi scale perception enhancement optimization of YOLOv8n architecture. The original model adopts 640×640 input resolution and generates three groups of 80×80,40×40 and 20×20 feature maps through five stages of downsampling. The minimum detection head corresponds to an 8×8 receptive field. In the traffic sign data obtained by a wide angle vehicle lens, the physical size of the tiny target is often smaller than the network base receptor field, resulting in the attenuation of high frequency details during deep feature extraction. Traditional multi scale detection architecture makes it difficult to achieve subpixel level feature analysis. Therefore, this study proposes a 160×160 super resolution detection mechanism based on hierarchical feature reconstruction. The innovative implementation path includes: first, the 160×160 scale feature map of the third layer of the backbone network is retained, and local feature enhancement is carried out through the C2f module. Then, a cross stage feature fusion channel is constructed, and the deep 80×80 feature map is spatially aligned with the shallow high resolution features by up sampling. Finally, the two branch feature refining architecture is designed to integrate shallow texture details and deep semantic information, and finally, the fine grained target location is realized by adding detection heads. By establishing a cooperative mechanism of "deep semantic guidance and shallow feature compensation", the scheme significantly enhances the representation ability of the network for small size targets while maintaining the lightweight characteristics of the model.

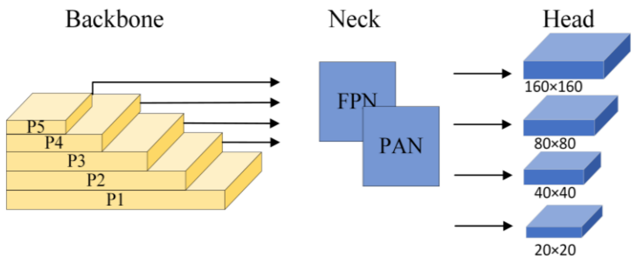


Fig. 5. Small target detection head

IV. Experimental Results and Analysis

A. Datasets used in the experiment

This study utilizes two publicly available traffic sign detection datasets CCTSDB and TT100K to comprehensively evaluate the proposed model's performance across various environmental conditions, weather scenarios, and object scales. The China Traffic Sign Detection Benchmark (CCTSDB) is designed to simulate real world driving environments and emphasizes complex and adverse road conditions. It contains a total of 13,828 images, including 11,062 for training and 2,766 for testing. The dataset features a wide range of challenging scenarios, such as fog, rain, glare from strong light, background occlusion, image blurring, and diverse camera angles. Traffic signs in CCTSDB are categorized into three functional types: mandatory, warning, and prohibited. This variability enables the model to learn more robust and generalizable features, improving its detection accuracy and reliability in degraded or dynamic conditions typical of real traffic scenes.

In contrast, the TT100K dataset, jointly developed by Tsinghua University and Tencent, emphasizes high resolution imagery and small target detection. It consists of 10,000 images at a resolution of 2048×2048 pixels, with over 30,000 traffic signs collected from real urban and suburban roads. A distinctive feature of TT100K is the large proportion of small scale traffic signs, some as small as 32×32 pixels, which are often difficult to detect due to limited pixel information and interference from complex backgrounds.

To reduce class imbalance and improve training efficiency, categories with fewer than 100 instances were excluded, resulting in a filtered dataset containing 45 sign categories. The final training and testing sets include 6,105 and 3,065 images, respectively. This preprocessing not only balances the dataset but also strengthens the model's stability and adaptability in small object detection tasks.

In summary, the CCTSDB dataset is primarily used to test the model's robustness under harsh weather and complex visual conditions, while TT100K provides a focused evaluation of its ability to detect small scale targets in high resolution scenes. By combining these two datasets, the study offers a comprehensive and realistic assessment of the model's performance in diverse traffic scenarios, ensuring its effectiveness and reliability in practical traffic sign detection applications.

Fig. 6. CCTSDB2021 dataset label type

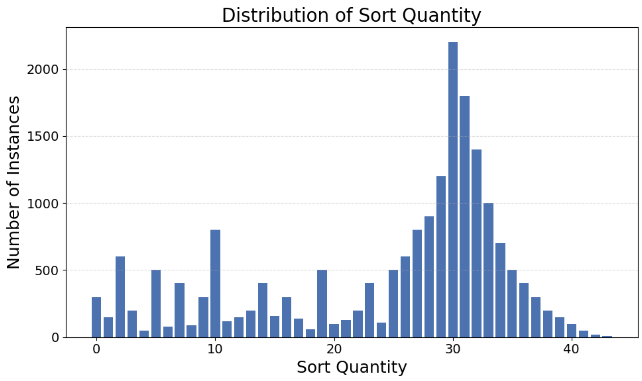


Fig. 7. TT100K dataset label type

B. *Experimental environment and parameter setting*

This experiment was conducted on a Windows 10 operating system, using Python as the programming language and PyTorch (version 2.4.0) as the deep learning framework, with CUDA version 12.1. The hardware configuration included an NVIDIA GeForce RTX 3070 GPU with 8GB of video memory and an Intel(R) Core(TM) i7-10700F processor, which provides sufficient performance for large scale deep learning model training. To standardize the input data format, the image size was adjusted to 640×640 during training. The model was trained for 150 epochs on the TT100K dataset with a batch size of 4. For the CCTSDB2021 dataset, the model was trained for 100 epochs, with the momentum and weight decay parameters set to 0.937 and 0.0005, respectively, to enhance model stability and generalization. The learning rate was set to 0.01 and adjusted using a cosine annealing scheduling algorithm to optimize the training process. Additionally, Mosaic data augmentation was applied during the final 10 training epochs to further improve the model's adaptability to diverse scenarios.

C. *Experimental environment and parameter setting*

The experimental results of this paper use the commonly used evaluation indexes of object detection: accuracy (P), recall rate (R), and average detection accuracy (mAP@0.5) as the main indexes of model evaluation. mAP@0.5 measures the average accuracy of the model for all classes of targets at 0.5 overlaps (IoU) between the predicted border and the actual border of the target. Specifically, mAP is calculated by calculating the accuracy and recall rate of the model under different detection thresholds and then represented by the area value under the P-R curve. The vertical axis of the P-R curve is Precision, which represents the proportion of correctly predicted positive samples in all predicted positive samples. The horizontal axis is Recall, which represents the proportion of samples correctly predicted to be positive to all actual positive samples. In this way, mAP can comprehensively evaluate the overall performance of the model under different detection tasks, especially the accuracy and recall ability of the model when dealing with target detection.

These indicators are calculated by the following formula:

$$\text{Precision}=\frac{TP}{TP+FP} \tag{1}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{2}$$

$$AP=\frac{\sum_{i=1}^{N} P_i}{N} \tag{3}$$

$$mAP=\frac{\sum_{j=1}^{N} AP_j}{M} \tag{4}$$

D. *Experimental results of the TT100K dataset*

To fully verify the effectiveness of the proposed algorithm in traffic sign detection tasks, this study selects classic models such as YOLOv3, YOLOv4, YOLOv5s, and YOLOv7-tiny as comparison models. The performance comparison results of different algorithms are shown in Table 1, where "ours" represents the proposed algorithm.

Table 1 summarizes the performance comparison of various models on the TT100K dataset. As shown, the proposed model (Ours) achieves the best overall performance, with a Precision (P) of 72.2%, Recall (R) of 68.2%, and mAP@0.5 of 73.8%, while maintaining a lightweight architecture with only 2.21M parameters. Compared with other representative models, such as YOLOv5 (P: 72.1%, R: 70.7%, mAP@0.5: 70.1%) and YOLOv11 (P: 71.6%, R: 65.2%, mAP@0.5: 72.8%), the proposed approach demonstrates superior detection accuracy and better balance between precision and recall. Moreover,it significantly

TABLE II
EXPERIMENTAL RESULTS OF DIFFERENT MODELS ON THE TT100K DATASET

Model	EMA	SPPF_imp	GSConv	Head	mAP@0.5	Parameters(M)	ModelSize（M）
Yolov8n					67.0	3.01	5.1
	√				68.0	2.59	4.9
		√			67.9	3.12	6.11
			√		67.1	2.83	5.63
				√	73.8	3.11	6.20
	√	√			68.9	2.79	4.60
	√	√	√		72.1	2.38	4.54
	√	√	√	√	74.2	2.06	4.38

TABLE I
EXPERIMENTAL RESULTS OF DIFFERENT MODELS ON THE TT100K DATASET

Model	P%	R%	mAP@0.5	Parameters(M)
YOLOv5s	72.1	70.7	70.1	7.02
YOLOv7-tiny	56.5	47.1	46.4	4.82
ImproveYOLOv8n[23]	70.0	67.9	71.8	4.37
YOLOv10	65.8	61.5	66.5	5.5
YOLOv11	71.6	65.2	72.8	5.37
Ours	72.2	68.2	73.8	2.21

reduces model complexity compared to models like YOLOv5 (7.02M parameters) and YOLOv10 (5.5M parameters), which is particularly advantageous for deployment in resource constrained environments. These results validate the effectiveness of the proposed improvements in enhancing both accuracy and model efficiency for traffic sign detection tasks.

As shown in Table 2, introducing the EMA attention module alone improves mAP@0.5 by 1% and significantly increases precision. Additionally, Parameters and Model Size decrease by 0.42M and 0.2M, respectively, indicating that the EMA module can reduce false detection rates while decreasing model size, verifying its ability to represent high level semantic features. When introducing the SPPF improvement module alone, the improved SPPF module enhances feature representation capabilities, expands the receptive field, improves object perception, reduces information loss, enhances detection accuracy, and increases computational efficiency. These improvements are confirmed in the ablation experiments. When introducing the convolution module alone, mAP@0.5 increases by 0.1%. Although the accuracy improvement is minor, the Parameter count is significantly reduced, demonstrating that GSConv effectively reduces the model's parameter size. When introducing the small object detection head alone, mAP@0.5 increases by 6.8%, a significant improvement compared to the original YOLOv8n model. This validates the effectiveness of the small object detection head, though it comes with an increase in parameter count and model size. When both the EMA attention module and the improved SPPF module are introduced into the YOLOv8 network, mAP@0.5 increases by 1.9%, while both Parameters and Model Size decrease, indicating that the EMA attention module and the improved SPPF module complement each other, enhancing accuracy while reducing model size and complexity.When the EMA attention module, SPPF improvement module, and GSConv module are introduced together, mAP@0.5 increases by 5.1%, demonstrating the effectiveness of the three modules. When all four modules (EMA attention module, SPPF improvement module, GSConv module, and small object detection head) are incorporated, mAP@0.5 increases by 7.2%, achieving the highest detection accuracy among all improvements, the model's parameter count is reduced by 0.95M, and model size decreases by 0.72M, proving that these four enhancements work synergistically to improve accuracy while making the model more lightweight, further optimizing detection performance.

Figure 8 shows the detection results of the proposed algorithm and YOLOv8n algorithm. In the first picture, it can be found that the traffic signs on the way are closely arranged, and it can be found that the improved algorithm proposed in this paper can accurately locate the six traffic signs in the figure, and no error detection or omission occurs. In contrast, YOLOv8n only detected four traffic signs and missed two traffic signs that appeared repeatedly. Compared with the improved algorithm, the YOLOv8 algorithm has the situation of missing detection. The second test picture is of a road situation in a dark light scene. The original YOLOv8 algorithm only detected four traffic signs, and there was one wrong detection. The smaller "i2r" sign is identified as "i2" class, and the blocked "i4r" sign is not recognized. By contrast, the improved algorithm in this paper can not only accurately identify the blocked traffic sign but also has a better detection accuracy under dark conditions.

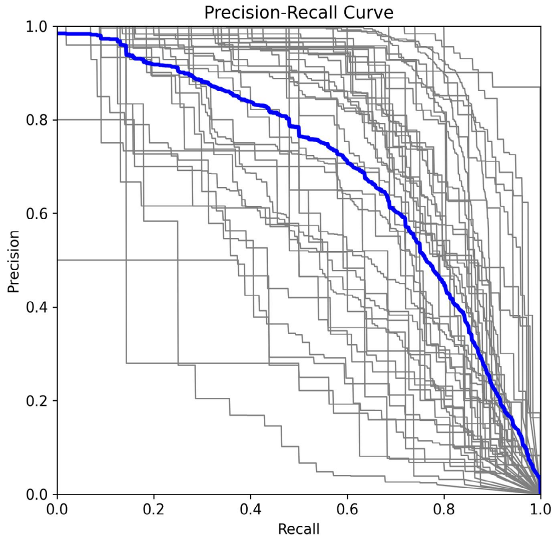
Figure 9 compares the precision-recall (P-R) curves before and after algorithm improvement. In object detection tasks, the P-R curve serves as a key metric for evaluating model performance, illustrating the trade off between precision and recall across varying confidence thresholds. As shown in the figure, the P-R curve of the improved YOLOv8n (Fig. (a)) is compared with that of the original YOLOv8n (Fig. (b)).Overall, the improved model’s curve is closer to the upper left corner, indicating higher precision while maintaining a high recall rate. Moreover, the improved YOLOv8n exhibits smaller fluctuations across different classes or experimental conditions, suggesting enhanced generalization ability and more stable detection results. In contrast, the original YOLOv8n shows a more scattered curve with a sharper drop in precision, implying a higher risk of false positives or missed detections in traffic sign scenarios.In summary, by optimizing the network architecture, improving the loss function, and enhancing data processing strategies, the improved YOLOv8n achieves better detection accuracy and stability, leading to superior performance in object detection tasks.

TABLE III
EXPERIMENTAL RESULTS OF DIFFERENT MODELS ON THE CCTSDB DATASET

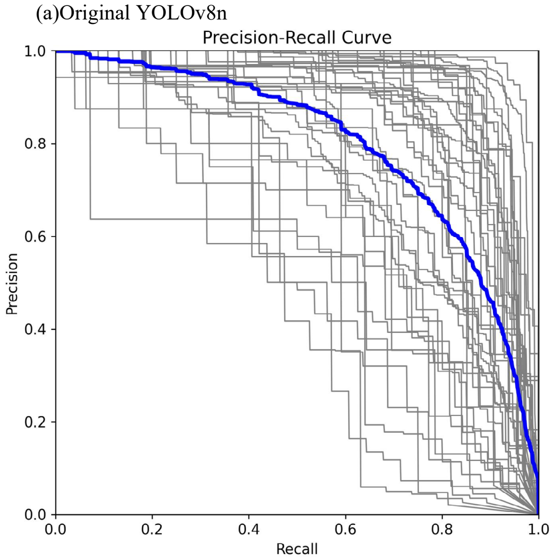
Model	P%	R%	mAP @0.5	Parameters(M)	ModelSize (M)
SSD	79	62.4	70.5	26.28	91.2
YOLOv5s	84.4	72.7	79.8	7.02	14.7
YOLOv7-tiny	81.3	70.1	73.2	4.82	10
YOLOv8n	94.8	91.1	95.2	3.01	6.3
ImproveYOLOv8n[23]	86.0	66.8	74.7	2.11	-
Ours	95.6	93.7	97.5	1.96	4.07

TABLE IV
EXPERIMENTAL RESULTS OF DIFFERENT MODELS ON THE CCTSDB DATASET

Model	EMA	SPPF_imp	GSConv	Head	mAP@0.5	Parameters(M)	ModelSize（M）
Yolov8n					95.2	3.01	5.1
	√				68.0	2.79	4.9
	√	√			67.9	2.59	4.56
	√	√	√		96.6	2.21	4.38
	√	√	√	√	97.5	1.96	4.07



(a)Original YOLOv8n



(b) Improved YOLOv8n

Fig. 9. Comparison of P-R curves between the improved YOLOv8n algorithm and the original YOLOv8n algorithm

E. Experimental results of the CCTSDB2021 dataset

In order to further verify the generalization ability of the proposed algorithm, a comparison experiment was conducted with other mainstream algorithms on the CCTSDB2021 dataset. The experimental results are shown in Table 3.

As can be seen from Table 3, the improved algorithm proposed in this paper has reached the highest values in P, R, and mAP@0.5, which is significantly improved compared with other mainstream classical algorithms and superior to the four classical algorithms. Moreover, the improved algorithm proposed in this paper is more reflected in lightweight. Compared with the improved YOLOv8 algorithm, the Parameter value of the improved algorithm proposed in this paper is reduced by 0.15M, and it is more than one third lower than that of YOLOv8n. The parameter number and model size of the improved algorithm proposed in this paper are only 1.96M and 4.07M. Much smaller than other comparison models.

Table 4 presents the results of ablation experiments conducted on the CCTSDB2021 dataset to evaluate the effects of different module combinations on model performance. In the experiment, we successively introduced EMA, SPPF_imp, GSConv, and Head modules to observe their effects on the Model mAP@0.5, Parameters, and Model Size. The results show that when the EMA module is used alone, the mAP@0.5 of the model reaches 95.2, the parameter number is 3.01M, and the model size is 5.1M. After adding the SPPF_imp module, mAP@0.5 decreases significantly to 68.0, parameter quantity to 2.79M, and model size to 4.9M. With the further introduction of the GSConv module, mAP@0.5 is increased to 67.9, the parameter number is reduced to 2.59M, and the model size is 4.56M. Finally, after adding the Head module, mAP@0.5 reaches the highest value of 97.5 with 1.96M parameters and 4.07M model size. These results show that by combining different modules reasonably, the number and size of parameters of the model can be significantly reduced while maintaining high performance, and the model can be lightweight.

Figure 10 presents a comparison of different methods for traffic sign detection. Subfigures (a) and (b) show two distinct traffic sign scenarios. Each subfigure consists of three panels: the original image, the heatmap generated by YOLOv8n, and the heatmap produced by our method. The original images display the actual traffic signs along with surrounding environmental details. In the heatmaps, regions in darker shades (closer to red) indicate a higher probability of traffic sign presence as determined by the model. From the figure, we can observe differences in the heatmap

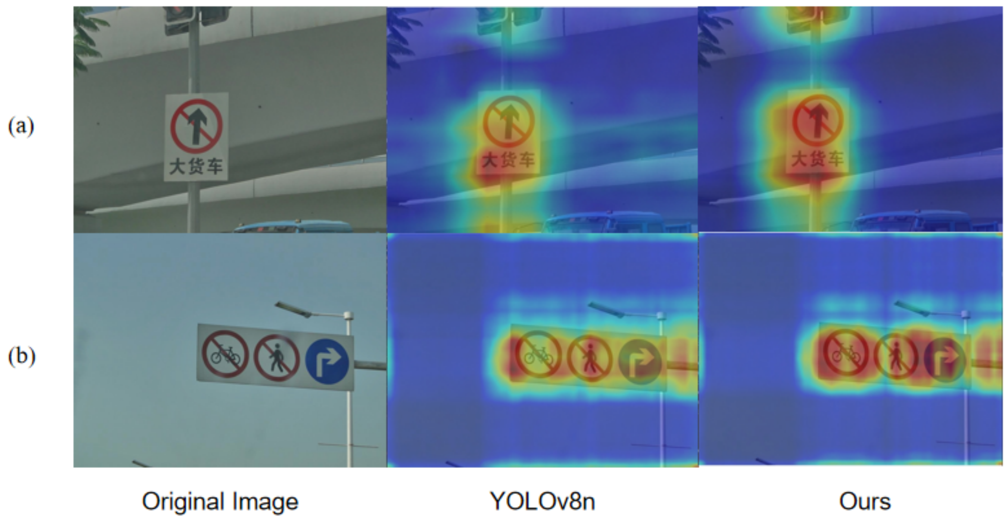


Fig. 10. Comparison of heat maps of different methods in the traffic sign detection task

distributions between YOLOv8n and our method across both scenarios. Our method's heatmaps are more concentrated and accurately cover the actual signs, highlighting its superior Accuracy and robustness in traffic sign detection. In contrast, YOLOv8n's heatmaps show some over diffusion and less accurate detection in certain areas, which further underscores the advantage of our method in this task.

V. CONCLUSION

In this paper, a lightweight traffic sign detection method based on an improved YOLOv8n architecture is proposed to address the challenges of false detections, missed detections, and high computational complexity in complex traffic scenarios. The optimized Spatial Pyramid Pooling Fast (SPPF) module enhances multi scale feature extraction and improves detection of small targets. The incorporation of the Efficient Multi scale Attention (EMA) mechanism strengthens feature representation and robustness against background interference. A dedicated small object detection head further improves detection accuracy for distant, low resolution, or densely arranged traffic signs. Additionally, the integration of Ghost Spatial Convolution (GSConv) effectively reduces redundant computation, minimizes model size, and accelerates convergence, enabling real time performance with lower resource consumption. Experimental results demonstrate that the proposed method achieves superior detection accuracy while significantly reducing missed and false detections, meeting the stringent demands of autonomous driving systems. Future work will focus on enhancing detection robustness under adverse weather and lighting conditions to improve adaptability for real world autonomous driving applications..

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 779–788, 2016.

[2] R. Uikey, H. R. Lone, and A. Agarwal, “Indian traffic sign detection and classification through a unified framework,” IEEE Trans. Intell. Transp. Syst., vol. 25, no. 10, pp. 14866–14875, 2024.

[3] S. Saxena, S. Dey, M. Shah, and A. Author, “Traffic sign detection in unconstrained environment using improved YOLOv4,” Expert Syst. Appl., vol. 238, pp. 121836, 2024.

[4] R. Mahadshetti, J. Kim, and T. W. Um, “Sign-YOLO: Traffic sign detection using attention-based YOLOv7,” IEEE Access, vol. 12, pp. 132689–132700, 2024.

[5] Y. T. Luo and Q. Gao, “Traffic sign detection based on channel attention and feature enhancement,” J. South China Univ. Technol. (Nat. Sci. Ed.), vol. 51, no. 12, pp. 64–72, 2023.

[6] Y. Peng, K. Chen, Y. Chen, and A. Author, “A lightweight context-aware traffic sign detection algorithm based on YOLOv8-n,” in Proc. 2024 IEEE 2nd Int. Conf. Control, Electron. Comput. Technol. (ICCECT), pp. 1578–1583, 2024.

[7] H. Lan and H. Z. Wang, “An improved algorithm combining lightweight and multi-scale fusion for traffic sign detection,” Comput. Eng., vol. 50, no. 10, pp. 1–13, 2024.

[8] D. Ouyang, S. He, G. Zhang, and A. Author, “Efficient multi-scale attention module with cross-spatial learning,” in Proc. ICASSP 2023 – 2023 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), pp. 1–5, 2023.

[9] H. Li, J. Li, and H. Wei, “Slim-neck by GSConv: A lightweight design for real-time detector architectures,” J. Real-Time Image Process., vol. 21, no. 3, p. 62, 2024.

[10] A. Jain, A. Mishra, A. Shukla, and A. Author, “A novel genetically optimized convolutional neural network for traffic sign recognition: A new benchmark on Belgium and Chinese traffic sign datasets,” Neural Process. Lett., vol. 50, no. 3, pp. 3019–3043, 2019.

[11] J. Zhang, X. Zou,and L. D. Kuang,“CCTSDB 2021: A more comprehensive traffic sign detection benchmark,” Hum.-Centric Comput. Inf. Sci., vol. 12, 2022.

[12] Donghao Hou, Yujun Zhang, and Jia Ren, "A Lightweight Object Detection Algorithm for Remote Sensing Images," *Engineering Letters*, vol. 33, no. 3, pp704-711, 2025

[13] Wutao Du, Xinyu Ouyang, Nannan Zhao, and Yifan Ouyang, "BCS-YOLOv8s: A Detecting Method for Dense Small Targets in Remote Sensing Images Based on Improved YOLOv8s," IAENG International Journal of Computer Science, vol. 52, no. 2, pp417-426, 2025

[14] Jiacheng Xu, Yonghui Yang, and Maoxiang Chu, "MPGA-YOLOv8: A Leaf Disease Detection Network Combining Multi-Scale Progressive Feature Network and Global Attention," IAENG International Journal of Computer Science, vol. 52, no. 2, pp436-446, 2025

[15] Liuai Wu, Nannan Lu, Xiaotong Yao, and Yong Yang, "Enhanced Helmet Wearing Detection Using Improved YOLO Algorithm," IAENG International Journal of Computer Science, vol. 51, no. 10, pp1426-1439, 2024

[16] Chao Chen, Bin Wu, and Yongguo Shi, "An Improved YOLOv5 Algorithm for Detecting Target," IAENG International Journal of Computer Science, vol. 51, no. 10, pp1454-1461, 2024

[17] Lanxue Dang, Gang Liu, Yan-e Hou, and Hongyu Han, "YOLO-FNC: An Improved Method for Small Object Detection in Remote Sensing Images Based on YOLOv7," IAENG International Journal of Computer Science, vol. 51, no. 9, pp1281-1290, 2024

[18] Mengjia Yan, Jianqiang Wang, Wenlong Zhao, Hongju Chen, and Yongzhi Chen, "An Economical Snow-melting Approach with Snow and Ice Detection Based on Autonomous Driving Scenarios," *Engineering Letters*, vol. 32, no. 7, pp1255-1265, 2024

[19] Junshang Fu, and Ying Tian, "Improved YOLOv7 Underwater Object Detection Based on Attention Mechanism," *Engineering Letters*, vol. 32, no. 7, pp1377-1384, 2024

[20] Xiaoming Zhang, and Ying Tian, "Improved YOLOv5s Traffic Sign Detection," *Engineering Letters*, vol. 31, no.4, pp1883-1893, 2023

[21] S. Xie, M. Zhou, C. Wang, and S. Huang, “Csppartial-yolo: A lightweight yolo-based method for typical objects detection in remote sensing images,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023.

[22] J. Lin, Y. Zhao, S. Wang, and Y. Tang, “Yolo-da: An efficient yolo based detector for remote sensing object detection,” IEEE Geoscience and Remote Sensing Letters, 2023.

[23] E. J. Xiong, R. F. Zhang, and Y. H. Liu, “Ghost YOLOv8 detection algorithm for traffic signs,” Comput. Eng. Appl., vol. 59, no. 20, pp. 200–207, 2023.