A Distant Traffic Sign Detection Algorithm Based on YOLOv8s-REMN

XU Yingzhe, DU Qingzhi*, SHAO Yubin, DUO Lin

(College of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In the field of traffic sign detection, challenges arise due to the small area coverage of distant traffic signs in the scene and the diverse scales of the signs. To overcome the above challenges, this paper presents an improved YOLOv8s-based traffic sign detection algorithm, YOLOv8s-REMN. First, the method introduces the receptive field attention convolution (RFAConv) into the backbone network to enhance the receptive field and feature extraction capability of the network. Second, the efficient attention-guided feature module (EAGFM) module is added to the neck network to optimize multi-scale feature fusion. Then, the multi-scale detail enhancement fusion (MSDEF) module is incorporated into the detection head to increase the small object detection head, improving the detection of small targets. Finally, the normalized Wasserstein distance (NWD) loss function replaces the CIoU loss function to optimize the bounding box regression and improve the precision of small object localization. Experimental results show that YOLOv8s-REMN achieves significant performance improvements on the TT100K dataset. Compared to the original YOLOv8s, mAP@0.5 increases by 6.6% and mAP@0.5:0.95 increases by 5.1%. The effectiveness of the algorithm is also validated on the Chinese Traffic Sign Detection dataset, CCTSDB2021, where YOLOv8s-REMN outperforms YOLOv8s with a 2.9% increase in mAP@0.5 and a 2.9% increase in mAP@0.5:0.95.

Highlights:

1. To address the challenge of detecting small traffic signs in the long-range vision, an improved model YOLOv8s-REMN is proposed.
2. The backbone RFAConv expands the receptive field, the neck EAGFM optimizes multi-scale fusion, and the detection head MSDEF enhances the extraction of small target details.
3. The NWD loss is adopted instead of CIoU to solve the problem of gradient disappearance for small targets and improve the accuracy of box positioning.
4. Double dataset verification shows a significant improvement in accuracy, meeting real-time requirements, and the comprehensive performance is superior to the mainstream comparison algorithms.

Key words: traffic sign detection; YOLOv8s; receptive field attention convolution (RFAConv); efficient attention-guided feature module (EAGFM); multi-scale detail enhancement fusion (MSDEF); normalized Wasserstein distance (NWD) loss

---

Foundation item: Open Research Fund of Yunnan Key Laboratory of Media Convergence(No. 220245201).

Received: 2025-04-22; **Revised**: 2025-06-05

***Corresponding author**, **E-mail**: 57960748@qq.com.

ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 41,No. 4,Jul. 2026,pp. 1212－1225
DOI:10. 16337/j. 1004-9037. 2026. 04. 020

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

基于YOLOv8s-REMN的远景交通标志检测算法

徐英哲，杜庆治，邵玉斌，朵　琳

(昆明理工大学信息工程与自动化学院,昆明650500)

摘　要：在交通标志检测领域,由于远景交通标志在场景中占比面积很小、标志尺度多样等特点,给检测带来了挑战。为解决上述问题,在YOLOv8s基础上进行了改进,提出了一种新型交通标志检测算法YOLOv8s-REMN。该方法在骨干网络中引入RFAConv,增强网络的感受野和特征提取能力;在颈部网络中加入EAGFM模块,优化多尺度特征融合;然后,在检测头部分加入MSDEF模块,增加小目标检测头,提升小目标检测能力;采用NWD损失函数替换CIoU损失函数,优化边界框回归,提高小目标定位精度。实验结果表明,YOLOv8s-REMN在TT100K数据集上实现了较大性能提升,相较于基线YOLOv8s,mAP@0.5提高了6.6%,mAP@0.5:0.95提高了5.1%。该算法的有效性也在中国交通标志检测数据集CCTSDB2021上得到了验证,YOLOv8s-REMN相较于YOLOv8s,mAP@0.5提高了2.9%,mAP@0.5:0.95提高了2.9%。

关键词：交通标志检测;YOLOv8s;感受野注意力卷积;跨尺度细节增强融合;有效注意力引导特征模块;归一化Wasserstein距离损失

中图分类号：TP391.4　　**文献标志码：**A

引用格式：徐英哲,杜庆治,邵玉斌,等.基于YOLOv8s-REMN的远景交通标志检测算法[J].数据采集与处理,2026,41(4):1212-1225. XU Yingzhe, DU Qingzhi, SHAO Yubin, et al. A distant traffic sign detection algorithm based on YOLOv8s-REMN[J]. Journal of Data Acquisition and Processing, 2026, 41(4):1212-1225.

引　　言

随着智能交通系统的发展[1],交通标志检测在自动驾驶、驾驶辅助系统和道路安全中至关重要。准确快速识别交通标志有助于保障交通安全和提升通行效率,因而对交通标志的检测提出了快速且高精度的要求[2]。但复杂环境、远距离、小尺寸标志和尺度多样性给交通标志识别带来挑战。图1(a)为近距离交通标志,标志清晰;图1(b)为远距离场景,右上角放大了难以明显识别的小标志。远景照片中交通标志占比面积很小,很难像近景照片一样可被明显观察。

(a) 近距离交通标志示例
(a) Example of close-range traffic signs

(b) 远距离交通标志示例
(b) Example of long-distance traffic signs

图1　远近距离交通标志场景示例

Fig.1　Examples of traffic sign scenarios at near and far distances

交通标志检测的算法研究主要采取了基于特征工程的方法以及基于深度学习的目标检测方

基金项目：云南省媒体融合重点实验室开放项目(220245201)。
收稿日期：2025-04-22;**修订日期：**2025-06-05

法[3]。传统的交通标志检测方法多采用基于特征工程的方法,如HOG[4]、SVM[5]等。然而,这些方法在复杂场景下的鲁棒性较差,尤其在处理小目标和多尺度目标时,性能受到较大限制。近年来,得益于深度学习理论的发展,目标检测领域的相关方法有了显著的成果与提升。根据检测阶段的数量不同,分为两类。一类是二阶段算法,需要先生成候选区域,再去对候选区域进行识别。Fast R-CNN[6]、Faster R-CNN[7]、Mask RCNN[8]和Cascade R-CNN[9]等算法都是二阶段算法,二阶段目标检测算法需要生成一系列候选区域,尽管具有较高精度,但检测速度慢,难以实现交通标志实时检测任务。

另一类为一阶段算法,其代表算法包括SSD[10]系列算法以及YOLO系列算法。YOLO系列模型凭借端到端的检测架构与出色的实时性,成为目标检测领域的关键方法。YOLOv8经过广泛测试与应用,在各类实际场景中展现出卓越的稳定性与可靠性。通过优化网络结构,YOLOv8在提高检测精度的同时,兼顾了推理速度,使其成为广泛应用于目标检测任务的主流模型。然而,从YOLOv9开始的后续版本(如YOLOv10、YOLOv11)逐步向特定场景优化,针对特定环境的检测需求进行了定制化调整,不再像YOLOv8那样具备高度的通用性。因此,在交通标志检测这个需要兼顾小目标和大目标检测,并追求实时性与精度平衡的任务中,YOLOv8仍然是理想的选择。Han等[11]提出的EDN-YOLO创新算法,以EfficientVit替换骨干网络、采用高效解耦检测头,并结合CIoU与归一化Wasserstein距离(Normalized Wasserstein distance, NWD)损失构建优化损失函数,显著提升了复杂场景下多尺度交通标志的检测性能,然而其检测精确度略低。Zhang等[12]提出的CR-YOLOv8算法,在特征提取阶段引入注意力模块,强化通道与空间特征,助力小物体关键信息的学习;在特征融合阶段引入RFB模块,提升特征多样性与多尺度目标检测能力;同时改进损失函数,平衡训练时的多尺度目标,进而提升模型泛化能力。然而,在部分对精度要求极为严苛的应用场景下,该算法的检测精确度仍然不够。Zhang等[13]提出的YOLO-BS在YOLOv8框架中新增了小目标检测层以增强特征提取能力,同时引入双向特征金字塔网络,提升了对多尺度目标的建模能力。朱强军等[14]通过用FasterNetBlock替换BottleNeck、用小目标检测层替代大目标检测层,并采用Wise-IoU损失函数替换CIoU损失函数,在提升交通标志检测精度的同时,大幅度降低参数量与计算量,显著提高检测速度,但其可检测的交通标志类别只有27类。综上所述,这些改进方法在检测性能方面均取得了一定提升,但均未充分考量交通标志检测场景中小目标占比较大这一因素。因此,针对远距拍摄条件下受遮挡、光照不良及背景复杂干扰的交通标志,现有方法在检测精度方面仍存在一定局限,易发生误检和漏检。

针对上述难点,本文提出了一种改进YOLOv8s的交通标志检测模型YOLOv8s-REMN,通过对骨干网络、颈部网络以及检测头进行改进,并更换损失函数来提升检测精度。具体改进方法如下。

RFAConv模块(R):使用感受野注意力卷积(Receptive field attention convolution, RFAConv)替换YOLOv8s的骨干网络中的4个标准卷积,结合感受野增强与注意力机制,提升对远距离信息的感知能力,增强对小目标的检测效果。

EAGFM(E):在颈部网络部分引入高效注意力导向特征模块(Effective attention-guided feature modulation, EAGFM),添加到C2f模块之后,通过全局特征调制机制,提升多尺度特征融合效果,优化特征信息表达。

MSDEF模块(M):将跨尺度细节增强融合(Multi-scale detail enhancement fusion, MSDEF)模块添加到检测头中,使其增加一个小目标检测头,利用深度可分离卷积与特征上采样机制,提升小目标和细节特征的检测能力。

NWD损失函数(N):在边界框回归过程中,采用NWD损失函数,优化目标框的位置和形状估计,提高边界框的定位精度,特别适用于小目标检测任务。

1 改进YOLOv8s模型

1.1 YOLOv8s模型介绍与分析

YOLOv8网络主要由3个部分组成:特征提取网络(Backbone)、特征融合网络(Neck)与检测头(Head)。YOLOv8能兼顾大、中、小目标检测,主要是因为其特征融合能力较强、Anchor-Free设计、解耦头优化、梯度增强模块、动态正负样本分配以及高效损失函数,这些优点使其在不同目标尺度下均具备良好的检测性能,同时保持了较高的计算效率和推理速度。尽管YOLOv8适合通用目标检测任务,但在交通标志检测任务上仍存在小目标占比较大、恶劣环境适应性、相似标志区分、遮挡处理等问题,因此需要结合改进策略进一步优化,以提升实际应用效果。考虑到YOLOv8中YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l和YOLOv8x五个不同模型的训练时间与检测效果,YOLOv8s在其中能兼顾模型大小、精确度和实时性。因此,选取YOLOv8s进行改进,模型如图2所示。

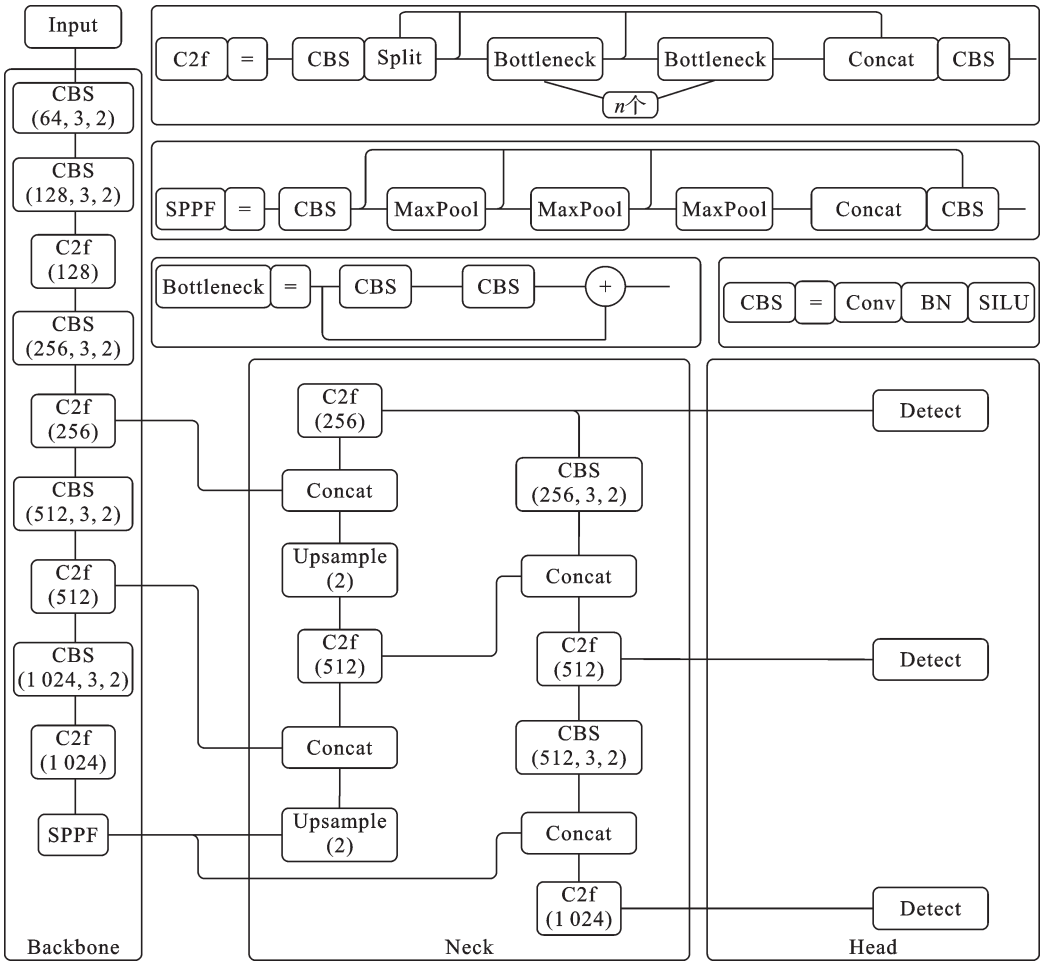

图2 YOLOv8s模型

Fig.2 YOLOv8s model

1.2 改进方法

1.2.1 RFCAConv模块

标准卷积在目标检测任务中有着重要的作用,但标准卷积存在以下局限性。

(1)局部感受野有限。卷积核只能聚焦于局部区域,原本YOLOv8中使用的1×1或者3×3卷积,只能覆盖1×1或者3×3的感受野,难以捕捉远距离的全局信息。想要获取更多的全局讯息,需要通过层级堆叠来扩大感受野,但这会增加计算量,且易出现梯度消失问题。

(2)缺乏全局信息建模能力。目标检测任务往往需要全局信息(如上下文关系),但标准卷积仅能依赖深度堆叠来获取更大感受野,导致信息传递不高效。

(3)对不同尺度的目标适应性较差。小目标容易被特征图下采样导致信息丢失,而大目标可能由于卷积核固定的大小无法充分利用全局信息。

为了解决标准卷积的局限性,Zhang等[15]提出了RFAConv。本文使用RFAConv来替换原本YOLOv8中的标准卷积。

RFAConv结合了感受野增强与注意力机制的优点,克服了标准卷积在目标检测任务中的局限性。核心思想是利用池化操作获取权重分布,并通过自适应加权特征生成机制提升感受野,从而增强网络的特征表达能力。相较于传统卷积,RFAConv具有以下优点:

(1)引入池化层获取全局权重信息。传统卷积仅能通过固定大小的卷积核(如1×1或3×3)提取局部特征,而RFAConv通过池化操作提取全局上下文信息,从而提升感受野,使不同尺度的特征均能被有效建模。这使得RFAConv能够在不增加过多计算成本的情况下增强远距离依赖关系,提升检测精度。

(2)自适应加权机制提升特征表达能力。RFAConv通过自适应权重计算,对不同位置的特征赋予不同的重要性,从而提高检测能力。相较于传统的固定权重卷积,RFAConv允许网络根据输入数据动态调整特征映射的权重,使得关键区域的特征得到更有效的强化,而非关键区域的噪声影响被削弱。RFAConv不仅注重通道间的注意力,还结合了空间注意力,提升特征选择能力。

(3)计算量适中,适用于实时目标检测任务。传统方法往往需要深度堆叠卷积层来扩展感受野,而RFAConv通过全局池化和自适应加权机制,在较浅层就可有效提升感受野,降低了模型深度需求,从而减少计算量。由于RFAConv不会引入额外的高昂计算成本,因此在实时目标检测具有较强的适应性,能够在保证检测精度的同时提升推理速度。

RFAConv的计算过程如下

$$\boldsymbol{X}_{\text{out}}=\text{Conv}_{1\times 1}\left(\sum_{i=1}^{N}\boldsymbol{W}_i\cdot\sigma\left(\text{BN}\left(\text{Conv}_{K_i}\left(\boldsymbol{X}\right)\right)\right)\right) \tag{1}$$

RFAConv使用多个卷积核计算不同尺度的特征图,即

$$\boldsymbol{F}_i=\sigma\left(\text{BN}\left(\text{Conv}_{K_i}\left(\boldsymbol{X}\right)\right)\right) \tag{2}$$

式中:$\boldsymbol{X}\in\mathbf{R}^{B\times C\times H\times W}$为输出特征图,$K_i$为第$i$组卷积核的大小(如3×3,5×5,7×7),$\text{Conv}_{K_i}$表示大小为$K_i\times K_i$的卷积操作,BN表示批归一化,$\sigma$表示激活函数,这里使用ReLU激活函数。通过计算得到多个不同感受野的特征图$\boldsymbol{F}_1,\boldsymbol{F}_2,\cdots,\boldsymbol{F}_N$,其中$N$为不同感受野的数量。

为了让网络自动选择最合适的感受野,使用式(3)计算每个感受野的权重。

$$\boldsymbol{W}=\text{Softmax}\left(\text{Conv}\left(\text{GAP}\left(\boldsymbol{X}\right)\right)\right) \tag{3}$$

GAP为全局平均池化,用于计算每个通道的全局均值。对于输入特征图$\boldsymbol{X}$的每个通道,GAP会计算该通道所有像素值的平均值。然后对结果进行一个卷积操作,并对卷积后的结果进行归一化处理,以此确保所有权重之和为1,从而计算出注意力权重$\boldsymbol{W}$,最终得到$\boldsymbol{W}_1,\boldsymbol{W}_2,\cdots,\boldsymbol{W}_N$作为不同感受野的权重。

有了不同感受野的特征$\boldsymbol{F}_i$和权重$\boldsymbol{W}_i$,再进行加权求和,即

$$\boldsymbol{X}_{\text{out}}=\sum_{i=1}^{N}\boldsymbol{W}_i\cdot\boldsymbol{F}_i \tag{4}$$

式中$\boldsymbol{W}_i\cdot\boldsymbol{F}_i$表示逐通道加权。

在YOLOv8s网络中，骨干网络和颈部网络均采用标准卷积。为了提升模型的感受野并增强对小目标的检测能力，在骨干网络中将其中4个标准卷积替换为RFAConv，以增强特征提取能力，使模型能够更有效地捕捉复杂场景中的关键特征。

1.2.2 EAGFM

在目标检测任务中，YOLOv8通过特征金字塔网络(Feature yramid network, FPN)融合自顶向下(Top-down)和自底向上(Bottom-up)的路径，以增强多尺度特征的表达能力。然而，传统FPN结构在特征传递过程中存在尺度不一致性问题，即高层特征的语义信息较强，而底层特征的空间信息较丰富，导致跨尺度特征融合不足。此外，不同尺度的特征图在检测相同目标时可能存在表征不一致的问题，从而影响最终的检测性能。尤其是在小目标检测场景下，FPN结构的局限性导致检测精度下降。

为解决上述问题，本文在跨视图交互模块(Cross-view interaction module，CVIM)[16]的基础上提出了EAGFM。CVIM是一种轻量化、通用的特征融合模块，主要用于提升跨视图信息共享和融合效率，特别适用于轻量化的图像超分辨率网络或其他计算机视觉任务。通过左右视图的特征交互提取视图间的互补信息，保留图像细节。优化输入特征的维度和去除冗余操作，使其适配轻量化网络。综合左右视图的特征交互，提升融合后图像质量。

EAGFM在CVIM架构基础上引入了像素注意力机制[17]以及动态加权策略，通过计算高、低视图特征之间的点积注意力来实现特征融合，其中注意力机制采用单头结构。这一改进不仅增强了跨视图特征交互的精准度，还通过合理的权重分配提升了对特征信息的深度挖掘能力，它能够依据不同尺度的特征进行动态加权，使得跨尺度特征融合更加平滑自然，保证了多尺度信息之间的一致性。其模块结构图如图3所示。图3中EAGFM模块的工作流程如下。

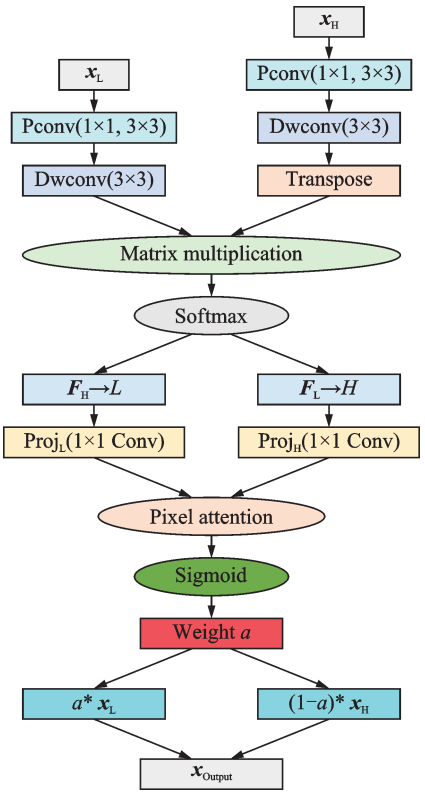

图3 EAGFM结构图

Fig.3 Structure of EAGFM

(1)根据输入低层级特征图$\boldsymbol{x}_L$和高层级特征图$\boldsymbol{x}_H$，如式(5,6)所示。通过不同的卷积层映射到查询$\boldsymbol{Q}$和$\boldsymbol{V}$空间。

$$\begin{cases}\boldsymbol{Q}_L = PConv_{1\times1}(PConv_{3\times3}(\boldsymbol{x}_L))\\ \boldsymbol{Q}_H = PConv_{1\times1}(PConv_{3\times3}(\boldsymbol{x}_H))\end{cases} \tag{5}$$

$$\begin{cases}\boldsymbol{V}_L = PConv_{3\times3}(\boldsymbol{x}_L)\\ \boldsymbol{V}_H = PConv_{3\times3}(\boldsymbol{x}_H)\end{cases} \tag{6}$$

(2)通过不同层级特征图的查询和转置的查询，计算跨视图的注意力矩阵，即

$$\boldsymbol{A} = \mathrm{Softmax}\left(\frac{\boldsymbol{Q}_L \cdot \boldsymbol{Q}_H^{\mathrm{T}}}{\sqrt{\boldsymbol{C}}}\right) \tag{7}$$

式中$\boldsymbol{C}$为特征维度，用于缩放点积注意力计算。

(3)根据注意力矩阵$\boldsymbol{A}$和$\boldsymbol{V}$，按式(8)，计算跨视图交互特征。

$$\begin{cases}\boldsymbol{F}_{H\to L} = \boldsymbol{A}\cdot\boldsymbol{V}_H\\ \boldsymbol{F}_{L\to H} = \boldsymbol{A}^{\mathrm{T}}\cdot\boldsymbol{V}_L\end{cases} \tag{8}$$

(4)将交互特征$\boldsymbol{F}_{H\to L}$和$\boldsymbol{F}_{L\to H}$通过像素注意力块生成动态权重a，即

$$a = \sigma(\mathrm{PixelAttention}(\boldsymbol{x}_L + \boldsymbol{x}_H, \boldsymbol{F}_{H\to L} + \boldsymbol{F}_{L\to H})) \tag{9}$$

式中：σ表示Sigmoid函数；PixelAttention表示像素注意力模块，用于生成融合权重。具体实施方式如图4所示。

(5)按照式(10),使用动态权重a对不同层级特征进行加权融合。

$$\boldsymbol{x}_{\text{out}}=a\cdot\boldsymbol{x}_{\text{L}}+(1-a)\cdot\boldsymbol{x}_{\text{H}} \tag{10}$$

1.2.3 MSDEF模块

在交通标志检测任务中,小目标占比较高,但高层特征图的分辨率较低,容易导致目标细节信息丢失,从而影响检测性能。为了解决该问题,基于多尺度细节增强(Multi-scale detail enhancement,MSDE)模块[18]提出了MSDEF模块。MSDE可以补充显著物体中缺乏的细节信息,解决在解码过程中上采样无法恢复关键细节的问题。包含主分支和辅助分支,主分支负责逐步上采样特征图,辅助分支则从输入图像中提取精细细节信息,并与主分支融合。

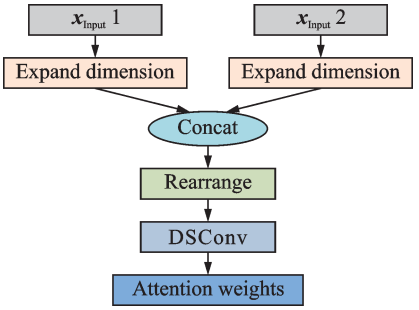

图4 像素注意力模块

Fig.4 Pixel attention module

MSDEF模块基于MSDE进行改进,嵌入特征上采样模块并优化特征融合方式。这样增强了模型对小目标细节的保留能力,精准捕捉细微特征。在特征融合时,MSDEF使用最近邻插值将低分辨率特征图上采样至最高分辨率,倍数依层次确定,优化融合策略,促进不同尺度特征信息的高效互补,让信息交互更加充分。同时,为符合交通标志检测实时性的要求,MSDEF使用深度可分离卷积[19],大幅减少计算量,实现性能与效率的平衡。MSDEF模块的如图5所示。

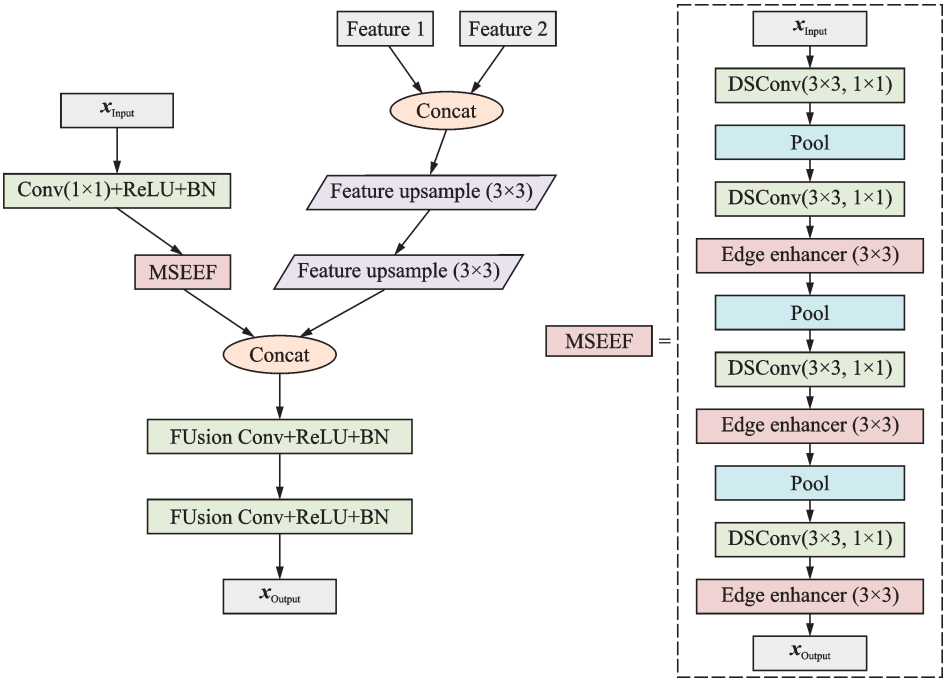

图5 MSDEF模块结构图

Fig.5 Structure of the MSDEF module

为了高效地将MSDEF模块到加入到YOLOv8中,对检测头结构进行了改进。在P2层(浅层特征层)引入MSDEF,并通过MSDEF进行特征融合,从而提升多尺度检测性能,增强算法对小目标的检测能力。最终,检测头Detect由[P2,P3,P4,P5]共同输入,确保充分利用不同尺度的信息,提高检测精度。

1.2.4 NWD损失函数

在YOLOv8s中,用NWD损失函数替换原有的CIoU损失函数。NWD是一种基于Wasserstein距

离的改进方法，专门用于衡量两个边界框之间的相似性。通过计算预测框与真实框在形状和位置上的差异来确定损失值。相较于传统的CIoU，NWD对边界框的形状变化更加敏感，能够更精准地处理不同形状的目标。在交通标志检测小目标占比较高的任务中，NWD能更准确地衡量预测框与真实框之间的差异，从而提升小目标的检测精度。此外，NWD在优化过程中提供了更平滑的梯度，有助于加速模型收敛，提高训练稳定性和检测性能。对于两个2D高斯分布P和Q，其Wasserstein距离w定义为

$$w(P,Q)=\sqrt{\|\boldsymbol{\mu}_P-\boldsymbol{\mu}_Q\|^2+\mathrm{tr}(\boldsymbol{\Sigma}_P+\boldsymbol{\Sigma}_Q-2(\boldsymbol{\Sigma}_P^{1/2}\boldsymbol{\Sigma}_Q\boldsymbol{\Sigma}_Q^{1/2})^{1/2})} \tag{11}$$

式中：$\boldsymbol{\mu}_P$和$\boldsymbol{\mu}_Q$为两个分布的均值（中心点），$\boldsymbol{\Sigma}_P$和$\boldsymbol{\Sigma}_Q$为两个分布的协方差矩阵。

为了归一化，NWD通常表示为

$$\mathrm{NWD}(P,Q)=\exp\left(-\frac{w(P,Q)}{C}\right) \tag{12}$$

式中C为一个归一化系数。据图像的尺寸设置为图像的对角线长度，这样可以将预测框中心的欧几里得距离归一化为[0,1]范围，从而确保损失值具有合适的尺度。如果输入图像尺寸是$W\times H$，则归一化系数C设置为

$$C=\sqrt{W^2+H^2} \tag{13}$$

1.3 YOLOv8s-REMN

为应对交通标志检测任务中的小目标检测问题，提出了YOLOv8s-REMN算法。首先将骨干网络中的4个标准卷积替换为RFAConv，在图5中，RBS模块是由RFAConv、批量归一化（Batch normalization，BN）以及激活函数SiLU（Sigmoid-weighted linear unit）依次组合而成的复合模块，其次在颈部网络的C2f之后添加EAGFM，优化特征融合。然后，设计了含有MSDEF模块的小目标检测头，并引入NWD损失函数。YOLOv8s-REMN的整体结构如图6所示。

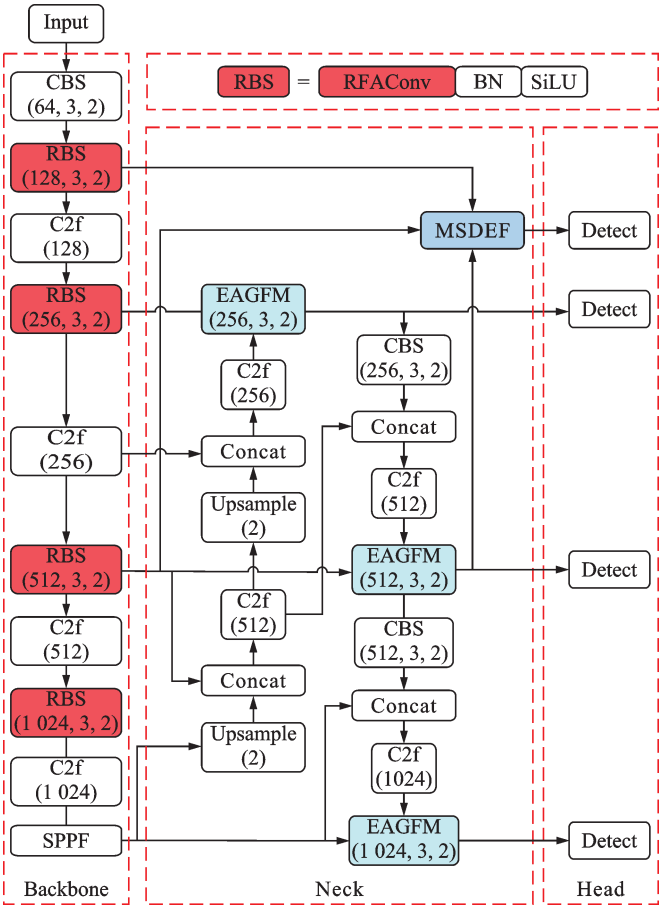

图6 YOLOv8s-REMN网络结构图

Fig.6 Network architecture of YOLOv8s-REMN

2 实验结果与分析

2.1 数据集介绍

TT100K数据集[20]由清华大学和腾讯的一个联合实验室编译并公开，其从中国5个不同城市的腾讯地图数据中心下载了10万张2 048像素×2 048像素的高分辨率街景图像，包括城市道路、高速公路和乡村道路，涵盖了中国多个城市的街景，以及广泛的照明和天气条件，以满足不同的应用需求。该数据集涵盖30 000个交通标志实例，共涉及221种不同类型的道路交通标

志。由于部分交通标志类别实例数量过少,难以保障模型训练的成功,在数据集预处理阶段,剔除了无标签和出现频率低的类别。最终,选取其中45个实例数量超过100个的交通标志类别,用于后续的深入分析。依据国际光学与光子学学会对小物体的定义,成像尺寸小于总像素数的0.12%的目标被认为是小物体。本次实验所使用的数据集中,84%的数据符合该定义标准。具体而言,训练集包含6 793张图片,验证集包含1 949张图片,测试集包含996张图片。

2.2 实验配置

在本实验期间,训练epoch设置为300,将每个训练周期(Epoch)的批次数量确定为8,同时将动量(Momentum)的值设定为0.937,输入图像尺寸为640像素×640像素。为了进一步提升模型的性能表现,采用了随机梯度下降(SGD)优化器,并将权重衰减率设置为0.000 5,以此来对模型进行训练。为优化学习策略,前290个epoch均使用Mosaic数据增强,最后10个epoch不使用。在训练过程中,初始学习率设置为0.01。为实现更稳定和高效的收敛,采用了余弦退火(Cosine annealing)策略来动态调整学习率,使其在训练过程中逐渐减小,最终趋近于0,从而有助于模型在训练后期更好地收敛到最优解。该策略可有效避免模型在训练中后期出现振荡或过早陷入局部最优的问题。本次实验所采用的环境具体配置情况如表1所示,关键参数汇总如表2所示。

表1 实验环境的具体配置

Table 1 Experimental environment configuration

参数	配置
操作系统	Windows11
CPU	12th Gen Intel(R) Core(TM) i5-12600KF
GPU	GPUNVIDIA RTX 4070ti (16 GB)
内存	32 GB
深度学习框架	Pytorch 2.5.1
编译器	PyCharm
CUDA版本	11.3

表2 训练关键参数

Table 2 Key training parameters

参数	设置值
Epoch	300
Batch size	8
Optimizer	SGD
Momentum	0.937
Weight decay	0.000 5
Input size	640×640
Mosaic augmentation	前290轮启用,后10轮禁用
Learning rate	0.01初始,Cosine annealing

2.3 评价指标

选取平均精度均值(Mean average precision, mAP)和每秒处理帧数(Frames per second, FPS)作为主要评估标准。其中,mAP用于反映模型在目标检测任务中的整体识别准确性,而FPS则用以评估图像处理的效率,数值越高表明模型运算速度越快。此外,为从多个维度评估算法性能,亦考虑引入精确率(Precision, P)与召回率(Recall, R)作为辅助指标。精确率用于衡量检测结果中正确识别目标的比例,其值越高,表示检测更为准确;召回率则关注模型发现目标的能力。

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{14}$$

召回率旨在衡量模型对目标样本的覆盖程度,用于评估模型是否能够全面地检索出所有应被识别的样本。

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{15}$$

P-R(Precision-recall)曲线用于描述精确率与召回率之间的关系。其中,精确率反映模型对检测结果的正确识别能力,而召回率则衡量其对全部真实目标的覆盖程度。该曲线下方所围成的面积即为平均精度(AP),而对所有类别AP值取平均后所得的mAP,则是评价目标检测模型性能的重要指标。

$$\begin{cases} \mathrm{AP}=\int_0^1 p(r)\mathrm{d}r \\ \mathrm{mAP}=\dfrac{1}{n}\sum_{i=1}^{n}\mathrm{AP}_i \end{cases} \tag{16}$$

本文采用的mAP指标为mAP@0.5与mAP@0.5:0.95。mAP@0.5指在IoU阈值设为0.5的条件下,分别计算每一类别的平均精度,并对其取平均值所得的结果。而mAP@0.5:0.95在目标检测领域是一个更为全面的性能评估标准,表示在多个不同IoU阈值(取值范围从0.5至0.95,步长为0.05)下所计算出的平均精度的均值。

为了在测试过程中获得更准确的FPS测量结果,测试时将批量大小设置为8,与模型训练阶段保持一致,并对实验数据进行分析对比。FPS所代表的含义是模型在每秒钟内能够处理的图像帧数数量。一般来说,FPS的数值越大,表明该模型对于图片的处理能力越出色、高效。在实际的应用场景当中,当模型的FPS达到30时,便足以满足实时监测的相关需求了。

2.4 消融实验

为了检验所提模型改进方法的实际效果,设计了一组消融实验。具体操作为:在YOLOv8s网络结构的基础上,每次仅引入1个改进模块,并进行单独训练。每次训练均采用相同的超参数设定,包括训练轮数、每轮的批次数量、动量系数以及其他优化器相关参数,确保实验结果的公平性,在TT100K数据集上开展相关的训练和测试工作,结果如表3所示。

表3 消融实验结果

Table 3 Ablation experimental results

模型	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量/10^6	FPS	GFLOPs
YOLOv8s	84.3	73.8	83.9	65.9	11.14	116.28	28.5
YOLOv8s+RFAConv	88.3	75.5	85.9	67.3	11.19	104.17	29.1
YOLOv8s+EAGFM	84.9	75.0	85.1	66.6	13.33	108.70	32.7
YOLOv8s+MSDEF	88.9	78.4	88.6	69.6	11.30	100.00	56.9
YOLOv8s+NWD	87.5	76.5	86.1	67.3	11.14	103.09	28.5
YOLOv8s-REMN	89.9	81.1	90.5	71.0	13.55	72.99	61.5

RFAConv:在YOLOv8s的骨干网络中使用RFAConv替换标准卷积后,模型检测性能得到明显提升。RFAConv引入了自适应感受野调节机制,能够根据不同尺度目标动态调整感受野,从而提升模型的特征提取能力。性能方面,精确率提升至88.3%,召回率提升至75.5%,mAP@0.5提升至85.9%,mAP@0.5:0.95提升至67.3%。与此同时,模型参数量略增至11.19×10^6,GFLOPs从28.5增至29.1,推理速度略降至104.17 FPS,显示出在较小计算代价下实现较大精度提升。

EAGFM:在YOLOv8s的颈部网络中加入EAGFM模块后,检测性能有所提升EAGFM利用全局特征调制机制与自适应权重学习策略增强目标区域特征表达能力。该模块将参数量提升至13.33×10^6,GFLOPs增加至32.7,推理速度降至108.7 FPS。尽管计算量有所增加,但检测性能依然提升,mAP@0.5达到85.1%,mAP@0.5:0.95提升至66.6%,适用于对精度要求较高的任务场景。

MSDEF:将MSDEF融合进YOLOv8s的检测头后,模型检测精度显著增强。该模块通过多尺度特征融合与跨尺度注意力机制,提升了不同层级特征之间的交互效率,尤其优化了小目标的检测效果。MSDEF模块使参数量略增至11.30×10^6,GFLOPs增加较多至56.9,推理速度下降至100 FPS。但其检测性能出色,mAP@0.5高达88.6%,mAP@0.5:0.95达69.6%,适合计算资源充足的高精度检测任务。

NWD:在YOLOv8s中用NWD替换原有的CIoU损失函数后,目标检测效果得到改善。NWD采用最优传输理论计算预测框与真实框之间的Wasserstein距离,更合理地刻画边界框分布,从而提升回

归精度。该方法在不增加模型参数量和GFLOPs的前提下,mAP@0.5提升至86.1%,mAP@0.5:0.95达到67.3%,但由于损失计算复杂度增加,推理速度降至103.09 FPS。

在YOLOv8s基础上引入RFAConv、EAGFM、MSDEF与NWD模块后,构建的YOLOv8s-REMN模型在目标检测性能方面取得显著提升。各模块分别针对特征提取、信息融合、尺度一致性与回归优化进行增强,形成结构间的互补与协同,共同提升了检测精度与推理效率。RFAConv提供可调节感受野,增强了骨干网络的特征提取能力,使模型能够更快聚焦于关键区域,从而减少无效计算。EAGFM尽管引入一定的计算开销,但其部署于颈部网络阶段,主要作用于特征整合环节,对整体推理路径的影响相对较小。MSDEF强调轻量化结构与高效信息交互,其引入的跨尺度特征融合方法进一步增强了特征融合效果。NWD通过优化目标框回归的分布一致性,使预测框更加贴近真实框,提高检测精度,尤其对小目标定位具有明显优势。尽管整体参数量上升至13.55×10^6,GFLOPs增加较多至61.5,但由于引入的模块在结构设计上对特征通道冗余进行了有效压缩与优化,模型推理速度仍保持在72.99 FPS,满足实时检测的需求,验证了该模块组合在提升检测性能的同时具备良好的实时性与实用价值。

图7用来说明YOLOv8s-REMN相较于原始算法的性能提升。图7(a)为数据集中原始图像,图像中的交通标志作放大处理。图7(b)为YOLOv8s算法的检测结果,第1、2和4幅图片中的交通标志均为被识别出来,第3幅图片中的交通标志虽然被识别出来,但被错检为限速40,实际为限速50。图7(c)为YOLOv8s-REMN算法的检测结果,均正确识别出交通标志。显然,YOLOv8s-REMN可以准确地检测出YOLOv8s错检或者遗漏的交通标志。

(a) 原始图像
(a) Original image

(b) YOLOv8s算法检测结果
(b) Detection results of the YOLOv8s algorithm

(c) 本文算法检测结果
(c) Detection results of the propsed algorithm

图7 部分检测结果展示

Fig.7 Presentation of partial detection results

2.5 特征融合方法对比实验

为验证所提多尺度特征融合方法MSDEF模块的有效性，选取了多种主流特征融合方法作为对比对象，并在TT100K数据集上进行了系统实验。在保持其他实验条件一致的前提下，比较分析了各方法在多尺度特征融合效果及整体检测性能方面的表现。表4总结了不同特征融合方法在mAP@0.5指标下的检测结果。

由表4可知，将MSDEF模块引入YOLOv8s模型后，其mAP@0.5提升至88.6%，在所有对比方法中表现最优。与原始YOLOv8s模型相比，提升了4.7%；相较于采用其他主流特征融合模块（如BiFPN、ODL、RFB）的模型，也表现出更明显的性能优势。这充分表明了MSDEF模块在多尺度特征融合方面的有效性与优越性。

表4 不同特征融合方法对比结果（mAP@0.5）

Table 4 Comparison of different feature fusion methods (mAP@0.5)

模型	mAP@0.5/%
YOLOv8s	83.9
YOLOv5s	73.6
YOLOv5s+BiFPN[21]	74.0
YOLOv8s+ODL[22]	84.5
YOLOv8s+RFB[11]	86.0
YOLOv10s	82.4
YOLOv8s+MSDEF	88.6

2.6 对比实验

为验证所提算法的有效性，选取了多个主流目标检测模型作为对比对象。表5显示了不同检测算法在TT100K数据集上，以mAP@0.5为指标的性能表现。结果表明，所提出的YOLOv8s-REMN算法在交通标志检测测试中优于所有其他算法。

表5 各算法性能对比实验（mAP@0.5）

Table 5 Performance comparison of various algorithms (mAP@0.5)

模型	mAP@0.5/%
SSD[9]	76.3
SSD+AlignedMatching[23]	84.7
Faster R-CNN[6]	69.5
TSP-RCNN[24]	81.2
Sparse R-CNN[25]	82.0
Deformable DETR[26]	77.1
AIE-YOLO[27]	84.8
YOLO-SG[28]	75.8
EDN-YOLO[10]	79.1
CR-YOLOv8[11]	86.9
YOLOv8s-DDA[29]	87.2
YOLO-BS[12]	90.1
所提算法	90.5

2.7 泛化实验

在公开可用的数据集中国交通标志检测数据集CCTSDB2021[30]上进行实验，以验证YOLOv8s-REMN算法在不同数据集上的泛化能力。CCTSDB2021是一个开源交通标志数据集，由长沙理工大学的专家团队创建，包含近20 000张交通标志图像，涉及近40 000个交通标志实例。这些图像来自中国的实际道路驾驶场景，其中大多数是车载视角拍摄的。训练集和测试集共包含17 856张图像及其标注信息，并被分为3类：禁止标志（Prohibition signs）、警告标志（Warning signs）和指示标志（Directional signs）。根据表6实验结果，相较于YOLOv8s，YOLOv8s-REMN的mAP@0.5提高了2.9%，mAP@0.5:0.95提高了8.7%。模型的参数量从11.12×10^6增加到13.54×10^6，仅提升了约21.77%，而对应的FPS从178.57降到了128.21，GFLOPs从28.4增加至61.3。

表6 CCTSDB2021数据集的实验结果

Table 6 Experimental results on the CCTSDB2021 dataset

模型	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量/10^6	FPS	GFLOPs
YOLOv8s	88.0	75.4	82.8	50.0	11.12	178.57	28.4
YOLOv8s-REMN	91.1	76.6	85.7	58.7	13.54	128.21	61.3

3 结束语

本文提出了一种基于YOLOv8s改进的交通标志检测算法YOLOv8s-REMN，并在TT100K和CCTSDB2021数据集上进行了实验验证。通过引入RFAConv、EAGFM、MSDEF模块以及NWD损失

函数,该方法有效提升了复杂交通场景下小目标检测的能力,同时保持了较高的推理速度,并在检测精度方面取得了显著提升。实验结果表明,相较于基线YOLOv8s,YOLOv8s-REMN在多个关键指标上均有明显提升。其中,mAP@0.5由83.9%提升至90.5%,增幅6.6%;mAP@0.5:0.95从65.9%增长至71.0%,提升5.1%;精确率由84.3%提升至89.9%,提高5.6%;召回率从73.8%增至81.1%,提升7.3%。此外,泛化实验结果表明,该模型在不同数据集上的适应性和鲁棒性较强,进一步验证了其在实际应用中的有效性。

在与当前主流目标检测算法的比较中,YOLOv8s-REMN在交通标志检测任务上的表现优于YOLOv5s、CR-YOLOv8及YOLOv8s等算法,展现出较强的竞争力。具体而言,RFAConv增强了感受野与全局信息建模能力,提高了小目标的检测精度;EAGFM模块优化了尺度特征融合策略,从而提升检测的稳定性;MSDEF模块强化了细节特征表达,使检测结果更加精确;NWD损失函数改进了目标框的回归能力,提升了模型对小目标的定位精度。

综上所述,YOLOv8s-REMN在智能交通系统中的交通标志检测任务中展现了良好的实用价值,可应用于自动驾驶、智能交通管理等多个领域。未来将进一步优化模型结构,以降低计算开销、提高推理效率,并探索更先进的特征融合机制,去进一步提升小目标检测能力。

参考文献:

[1] 张松涛.基于深度学习的单目图像深度估计方法研究[D].武汉:武汉大学,2020.
ZHANG Songtao. Research on monocular image depth estimation method based on deep learning[D]. Wuhan: Wuhan University, 2020.

[2] 贾子豪,王文青,刘光灿.改进YOLOv5的轻量化交通标志检测算法[J].数据采集与处理,2023, 38(6): 1434-1444.
JIA Zihao, WANG Wenqing, LIU Guangcan. An improved lightweight traffic sign detection algorithm based on YOLOv5[J]. Journal of Data Acquisition and Processing, 2023, 38(6): 1434-1444.

[3] 陈飞,刘云鹏,李思远.复杂环境下的交通标志检测与识别方法综述[J].计算机工程与应用, 2021, 57(16): 65-73.
CHEN Fei, LIU Yunpeng, LI Siyuan. Survey of traffic sign detection and recognition methods in complex environment[J]. Computer Engineering and Applications, 2021, 57(16): 65-73.

[4] ZAKLOUTA F, STANCIULESCU B. Segmentation masks for real-time traffic sign recognition using weighted HOG-based trees[C]//Proceedings of 2011 14th International IEEE Conference on Intelligent Transportation Systems. Washington DC: IEEE, 2011: 1954-1959.

[5] LE T T, TRAN S T, MITA S, et al. Real time traffic sign detection using color and shape-based features[C]//Proceedings of Intelligent Information and Database Systems: Second International Conference. Hue City: [s.n.], 2010: 268-278.

[6] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2015: 1440-1448.

[7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28.

[8] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2017: 2961-2969.

[9] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 6154-6162.

[10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Proceedings of 14th European Conference. Amsterdam: Springer, 2016: 21-37.

[11] HAN Y, WANG F, WANG W, et al. EDN-YOLO: Multi-scale traffic sign detection method in complex scenes[J]. Digital Signal Processing, 2024, 153: 104615.

[12] ZHANG L J, FANG J J, LIU Y X, et al. CR-YOLOv8: Multiscale object detection in traffic sign images[J]. IEEE Access, 2023, 12: 219-228.

[13] ZHANG H, LIANG M, WANG Y. YOLO-BS: A traffic sign detection algorithm based on YOLOv8[J]. Scientific Reports, 2025, 15(1): 7558.

[14] 朱强军,胡斌,汪慧兰,等.基于轻量化YOLOv8s交通标志的检测[J].图学学报, 2024, 45(3): 422-432.

ZHU Qiangjun, HU Bin, WANG Huilan, et al. Detection of traffic signs based on lightweight YOLOv8s[J]. Journal of Graphics, 2024, 45(3): 422-432.

[15] ZHANG X, LIU C, YANG D, et al. RFAConv: Innovating spatial attention and standard convolutional operation[J]. arxiv preprint arxiv: 2304.03198, 2023.

[16] LI Y, ZOU W, WEI Q, et al. Multi-level feature fusion network for lightweight stereo image super-resolution[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2024: 6489-6498.

[17] ZHAO H, KONG X, HE J, et al. Efficient image super-resolution using pixel attention[C]//Proceedings of Computer Vision—ECCV 2020 Workshops. Glasgow: Springer, 2020: 56-72.

[18] GAO S, ZHANG P, YAN T, et al. Multi-scale and detail-enhanced segment anything model for salient object detection[C]//Proceedings of the 32nd ACM International Conference on Multimedia. [S.l.]: ACM, 2024: 9894-9903.

[19] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 1251-1258.

[20] ZHU Z, LIANG D, ZHANG S, et al. Traffic-sign detection and classification in the wild[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 2110-2118.

[21] 傅融,逯洋,彭淼.YOLOv5s模型的改进及其在交通标志检测上的应用[J].遥感信息,2024, 39(6): 87-93.
FU Rong, LU Yang, PENG Miao. Improvements to the YOLOv5s model and its application in traffic sign detection[J]. Remote Sensing Information, 2024, 39(6): 87-93.

[22] 李冰,朱孝峰,管嘉俊,等.基于LiteTS-YOLO的交通标志检测[J].自动化与仪表,2025, 40(1): 82-89, 94.
LI Bing, ZHU Xiaofeng, GUAN Jiajun, et al. Traffic sign detection based on LiteTS-YOLO[J]. Automation and Instrumentation, 2025, 40(1): 82-89, 94.

[23] KANG S H, PARK J S. Aligned matching: Improving small object detection in SSD[J]. Sensors, 2023, 23(5): 2589.

[24] SUN Z, CAO S, YANG Y, et al. Rethinking transformer-based set prediction for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2021: 3611-3620.

[25] SUN P, ZHANG R, JIANG Y, et al. Sparse R-CNN: End-to-end object detection with learnable proposals[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 14454-14463.

[26] ZHU X, SU W, LU L, et al. Deformable DETR: Deformable transformers for end-to-end object detection[J]. arxiv preprint arxiv: 2010.04159, 2020.

[27] YAN B, LI J, YANG Z, et al. AIE-YOLO: Auxiliary information enhanced YOLO for small object detection[J]. Sensors, 2022, 22(21): 8221.

[28] HAN Y, WANG F, WANG W, et al. YOLO-SG: Small traffic signs detection method in complex scene[J]. The Journal of Supercomputing, 2024, 80(2): 2025-2046.

[29] NIU M, CHEN Y, LI J, et al. YOLOv8s-DDA: An improved small traffic sign detection algorithm based on YOLOv8s[J]. Electronics, 2024, 13(18): 3764.

[30] ZHANG J M, LV Y R, TAO J J, et al. A robust real-time anchor-free traffic sign detector with one-level feature[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2024, 8(2): 1437-1451.

作者简介:

徐英哲(2000-),男,硕士研究生,研究方向:目标检测、图像水印,E-mail: 2302005803@qq.com。

杜庆治(1977-),**通信作者**,男,高级实验师,硕士生导师,研究方向:通信与信息系统,E-mail: 57960748@qq.com。

邵玉斌(1970-),男,教授,硕士生导师,研究方向:无线通信技术。

朵琳(1974-),女,副教授,硕士生导师,研究方向:下一代网络、智能信息处理、智能通信。

(编辑:夏道家)