DOI:10.19651/j.cnki.emt.2519458

改进 YOLOv11 的小目标交通标志检测算法*

刘 丽[1,2] 张初夏[1] 张 硕[1] 李宇健[1] 王 强[1,3]
(1.华北电力大学计算机系 保定 071000; 2.复杂能源系统智能计算教育部工程研究中心 保定 071000;
3.河北省能源电力知识计算重点实验室 保定 071000)

摘 要: 针对小目标交通标志检测中存在的特征表达能力不足、复杂背景干扰强、定位精度低等问题,提出了一种改进 YOLOv11 的多尺度特征增强算法 TSD-YOLOs。通过异构部分卷积和跨尺度特征融合结构设计 CSP-MSFPF 模块,有效增强小目标的特征表达能力;在网络颈部构建前景增强特征金字塔网络 FE-FPN,通过上下文关联与前景增强机制突出关键目标,抑制复杂背景干扰;采用 Haar 小波下采样保留高频细节并降低计算开销,缓解小目标特征在下采样过程中的损失;引入 WIoU v3 损失函数,通过动态调节梯度权重优化正负样本分布,提升模型的定位精度;使用基于 LAMP 分数的通道剪枝策略对改进后的模型进行压缩,在保证检测性能的同时显著降低模型复杂度与计算量。实验结果表明,相较于基线模型,TSD-YOLOs 在 TT100k 数据集上 mAP50、mAP50:95 分别提升 2.6%、2.1%,检测速率达 242 fps;在 CCTSDB 数据集上 mAP50、mAP50:95 分别提升 1.6%、3.9%;且 TSD-YOLOs 计算量降低 27.6%、模型大小压缩 58.7%、参数量减少 59.6%。实验结果充分验证了 TSD-YOLOs 在保证实时性的同时,显著提升了复杂场景下小目标交通标志的检测精度与鲁棒性。

关键词: 目标检测;交通标志;YOLOv11;跨尺度特征融合;前景增强
中图分类号: TP391;TN919.5 **文献标识码**: A **国家标准学科分类代码**: 520.6040

Improved small target traffic sign detection algorithm of YOLOv11

Liu Li[1,2] Zhang Chuxia[1] Zhang Shuo[1] Li Yujian[1] Wang Qiang[1,3]
(1. Department of Computer Science, North China Electric Power University, Baoding 071000, China;
2. Engineering Research Center of Intelligent Computing for Complex Energy Systems, Ministry of Education, Baoding 071000, China;
3. Hebei Key Laboratory of Knowledge Computing for Energy Power, Baoding 071000, China)

Abstract: To address the challenges of insufficient feature representation, strong background interference and low localization accuracy in small traffic sign detection, an improved multi-scale feature enhancement algorithm TSD-YOLOs based on YOLOv11 is proposed. Firstly, a CSP-MSFPF module is designed using heterogeneous partial convolutions and cross-scale feature fusion, effectively enhances the feature expression of small targets. Secondly, a foreground enhancement feature pyramid network is constructed in the neck of the network to improve target focus and suppress background noise through contextual correlation and foreground enhancement. In addition, the Haar wavelet downsampling is adopted to retain high-frequency details while reducing computational overhead, alleviating information loss for small targets. Meanwhile, the WIoU v3 loss function is utilized to dynamically adjust gradient allocation and optimize sample distribution, improving localization accuracy. Finally, a channel pruning strategy based on the LAMP score is applied to compress the improved model, which significantly reduces model complexity and computation while maintaining detection performance. Experimental results demonstrate that compared to the baseline model, TSD-YOLOs achieves 2.6% and 2.1% improvements in mAP50 and mAP50:95 respectively on the TT100k dataset, with a detection rate of 242 fps; On the CCTSDB dataset, mAP50 and mAP50:95 improved by 1.6% and 3.9%, respectively. Furthermore, the computational cost is reduced by 27.6%, the model size is compressed by 58.7%, and the number of parameters is reduced by 59.6%. The experimental results adequately verify that TSD-YOLOs significantly improves the detection accuracy and robustness of small-target traffic signs in complex scenes while ensuring real-time performance.

Keywords: object detection;traffic sign;YOLOv11;cross-scale feature fusion;foreground enhancement

0 引 言

随着科学技术的不断发展以及计算机硬件算力的持续提升,目标检测技术得到了广泛应用,极大地推动了自动驾驶与智能辅助驾驶等智慧交通技术的快速发展。其中,交通标志检测作为智慧交通系统的重要组成部分,在辅助驾

收稿日期:2025-07-26
* 基金项目:北京市自然科学基金(4254105)项目资助

驶员安全行驶、预警交通风险以及提升道路交通管理效率等方面发挥着关键作用。

在实际道路场景中,车载摄像头捕捉的图像通常包含尺寸较小的交通标志。虽然现有目标检测算法在通用任务上表现优异,但在应对此类小目标检测时仍面临诸多挑战。主要问题在于:一方面,交通标志本身包含的特征信息有限,难以被模型充分提取和表达;另一方面,常规的下采样操作极易丢失小目标的关键细节,进一步影响检测精度。此外,复杂背景、光照变化以及目标遮挡等现实因素,也会显著增加误检和漏检的风险。因此,优化小目标交通标志检测技术已成为智能交通领域的研究热点。攻克这一难题,对于提升智能驾驶系统的安全性与可靠性具有重要的理论价值和广阔的应用前景。

目前,小目标交通标志检测方法的研究主要分为基于传统手工特征的方法与基于深度学习的方法两大类。早期研究主要依赖人工提取的特征,包括形状特征、颜色特征及其融合方式。杨斐等[1]通过改进分割算法提高检测效率与完整性。尽管基于手工特征的方法在一定程度上具有较好的实时响应与鲁棒性,但在应对尺寸较小的交通标志时存在明显不足。由于小目标所包含的特征信息有限,传统方法往往难以精准刻画其细节特征,从而导致检测准确率下降,难以满足实际道路环境中对精度和可靠性的高要求。

深度学习领域中的目标检测模型主要分为两大类。第 1 类是以 Faster R-CNN[2]、Mask R-CNN[3] 和 Cascade R-CNN[4] 为代表的双阶段(two-stage)目标检测模型,这类模型通常具有较高的检测精度,适用于对精度要求较高的场景,但由于其在推理过程中需要生成并处理大量候选框,计算复杂度较高,难以满足实时性要求。第 2 类是以 YOLO[5]、SSD[6]、DETR[7] 等为代表的单阶段(one-stage)目标检测模型。该类模型有更快的检测速度,适用于对实时性要求较高的应用场景。然而,由于缺乏候选区域筛选机制,单阶段模型在检测精度方面相对较弱,尤其在小目标检测任务中仍存在一定的性能瓶颈。考虑到交通标志检测主要应用于车辆行驶过程,具有对实时性与准确性要求并重的特点,因此相关研究多集中于单阶段目标检测算法在小目标交通标志检测任务中的应用与改进。

张莹等[8]提出了基于 YOLOv7 的多通道特征融合学习网络,通过引入跨通道信息连接、浅层特征融合模块以及高分辨率检测头,有效增强了小目标缺陷的检测能力。然而,该方法对复杂背景下多尺度语义信息的深度交互建模能力有限,且网络结构定制性较强,在跨场景泛化与实时部署的灵活性方面仍存在一定不足。艾强等[9]基于 YOLOv8 模型提出了融合噪声抑制与深层语义增强的特征融合模块(noise suppression and semantic enhancement,NSSE),并引入多尺度通道注意力机制(multi-scale convolutional attention,MSCA),在提升模型对不同尺度目标的感知能力及几何形变鲁棒性方面取得了一定成效。然而,该方法主要侧重于通道维度的特征加权,缺乏对多感受野特征建模及跨尺度空间信息深度交互的有效设计,对小目标的细粒度细节表达和复杂背景下的稳定检测能力仍存在一定局限。俞林森等[10]基于 YOLOX 模型设计了轻量化骨干网络并引入前景注意力感知模块(foreground attention perception module,FAPM),在噪声抑制和前景增强方面取得了一定提升,但该方法仅在通道与空间维度进行加权,缺乏跨尺度信息交互机制,限制了模型对多尺度小目标的适应能力。Tang 等[11]基于 YOLOv11 模型设计 HybridFocus 模块与双向动态多尺度特征金字塔网络(bidirectional dynamic multi-scale feature pyramid network,BiDMS-FPN),在增强多尺度特征融合和抑制背景干扰方面取得了良好效果。然而,HybridFocus 模块主要依赖大核卷积与多尺度池化聚合上下文信息,导致对小尺度目标的边缘细节捕获能力不足。彭杰等[12]基于 YOLOv11 模型提出 C3k2_GG 与 SPPFAPSA 模块,并结合 GhostConv、全局注意力和 SimAM 机制以增强小目标特征表达。然而,多种注意力机制叠加易造成信息冗余与响应区域重叠,使模型过度聚焦显著区域,削弱了对细粒度特征的捕捉能力。Chu 等[13]基于 YOLOv11 模型采用 GhostNetV3 改进主干网络,并引入深度可分离卷积模块,在保持较高精度的同时有效降低了计算复杂度。但在结构轻量化的过程中牺牲了部分特征表达能力,对小目标的细节特征提取不足,进而影响复杂场景下的检测稳定性。Wang 等[14]基于实时检测 Transformer(real-time detection Transformer,RT-DETR)模型通过多尺度序列融合模块与级联分组注意力机制增强多尺度特征建模能力,在提升小交通标志检测精度与实时性能方面取得了良好效果。然而,该方法在高分辨率场景中仍存在计算开销较高的问题,且对小目标细粒度空间结构与局部细节特征的显式增强能力仍有一定局限。

综上所述,当前改进型 YOLO 模型在多尺度特征融合、轻量化结构设计和注意力机制集成方面虽取得显著进展,但仍存在小目标特征表达不足、上下文信息交互不充分以及特征冗余与噪声干扰等问题。针对上述挑战,本文提出一种基于多尺度特征增强的交通标志小目标检测算法 TSD-YOLOs,综合考虑特征多尺度融合、前景特征强化与细节信息保留等关键因素,兼顾检测精度与实时性。主要贡献如下。

1)针对单一感受野在上下文建模与交通标志细粒度特征捕捉方面的不足,设计了多尺度特征部分融合模块(multi-scale feature partial fusion,MSFPF),并构建了跨阶段部分连接多尺度特征部分融合模块(cross stage partial multi-scale feature partial fusion,CSP-MSFPF)模块以替代原始网络中的 C3k2 结构。该模块通过对部分通道进行多种感受野的异构卷积,有效提取并融合不同尺度的语义信息,增强多尺度特征间的信息交互与小目标的感知能力。

2)为优化小目标在复杂背景下的检测能力,在特征融合阶段构建了前景增强金字塔网络(foreground

enhancement feature pyramid network，FE-FPN)。该结构融合金字塔上下文信息，通过矩形自校准模块(rectangular self-calibration module，RCM)增强前景区域表达，并设计了跨尺度特征融合模块(cross-scale feature fusion，CSF)与语义特征融合模块(semantic feature fusion，SFF)实现高效特征融合的同时增强关键特征表示，实现模型在复杂场景中对交通标志关键特征的显著增强与对背景干扰的有效抑制，达到协同优化的目标。

3)为减少下采样过程中关键信息的损失，使用 Haar 小波池化(Haar wavelet downsampling，HWD)替代传统的步进卷积，在降低计算复杂度的同时保留更多边缘、纹理等细节信息。

4)引入智能交并比(wise intersection over union，WIoU) v3 作为新的回归损失函数，采用动态非单调机制来设计更合理的梯度增益分配策略，有效地降低部分标注样本质量低的不利梯度增益，提升模型对交通标志的定位性能和泛化能力。

5)使用基于层自适应幅值的剪枝算法(layer-adaptive magnitude-based pruning，LAMP)分数的通道剪枝策略对改进后的模型进行压缩，在保证检测性能的同时显著降低模型复杂度与计算量。

1　TSD-YOLOs

YOLOv11 是 Ultralytics 于 2024 年发布的目标检测模型，在 YOLOv8 的基础上进行了多项结构优化。主要改进包括使用 C3k2 模块替代原有的 C2f 模块，在 SPPF 模块后新增具备注意力机制的 C2PSA 模块，并在检测头中引入两个深度可分离卷积(DWConv)；此外，模型在深度与宽度参数上进行了大幅调整，以实现更优的检测精度与推理速度之间的平衡。YOLOv11 根据网络深度与宽度的不同，划分为 n、s、m、l、x 5 种模型规模。其中，YOLOv11-s 模型在计算量和参数规模上小于 m、l、x 模型，在检测精度上优于 n 模型，能够较好地兼顾精度与计算开销。因此，本文选择 YOLOv11-s 作为小目标交通标志检测的基线模型。

针对交通标志检测中存在的有效特征获取受限、背景干扰显著以及小目标定位精度低等问题，本文提出了一个融合多尺度特征增强与前景感知机制的模型 TSD-YOLOs，如图 1 所示。首先，对原有网络中的 C3k2 模块进行改进，设计了 CSP-MSFPF，利用多个异构核部分卷积和跨尺度特征连接，强化多尺度特征之间的信息交互，并增强对小目标的感知能力。其次，构建前景增强金字塔 FE-FPN，整合上下文信息并引入 RCM 模块，突出前景区域特征表达；同时构建 CSF 与 SFF，在实现高效多尺度融合的基础上，提升网络对关键区域的识别能力，有效抑制复杂背景干扰。此外，采用 HWD 替代传统下采样方式，兼顾效率与精度，在减少信息损失的同时保留高频细节特征。引入 WIoU v3 回归损失函数优化边界框回归，通过动态调节梯度优化样本标注质量不均的不利影响，增强小目标的定位准确性。最后，通过 LAMP 分数通道剪枝实现模型压缩，有效减少计算量并保持检测性能稳定。

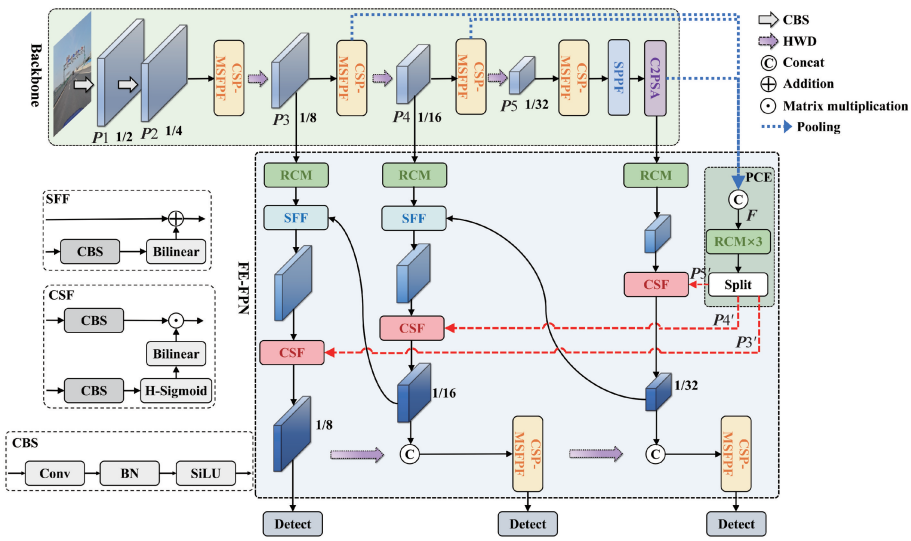

图 1　TSD-YOLOs 网络结构

Fig. 1　TSD-YOLOs network structure

1.1　CSP-MSFPF 模块

在交通标志检测任务中，小目标检测始终是一个重要且极具挑战性的研究方向。由于小目标具有尺度较小、特征模糊、易被背景干扰等特性，现有目标检测方法在此类任务中常面临检测精度不足的问题。YOLOv11 网络中的 C3k2 模块在一定程度上实现了多尺度特征融合，提升了目标检测性能。然而，该模块采用固定大小的卷积核限制了感受野的扩展能力；同时，全通道处理方式使得浅层特征的细粒度信息在后续卷积过程中逐渐被稀释或丢失。在面对遮挡严重或目标密集的场景时，模型难以充分捕获

小目标的上下文语义,容易出现漏检和误检等问题。

为此,本文借鉴 FasterNet[15] 中关于高效特征建模的卷积思想,设计了 MSFPF 模块,通过对部分特征通道施加多种感受野的异构卷积操作,实现对不同尺度语义信息的高效提取与融合。并在此基础上构建了 CSP-MSFPF 模块,如图 2 所示,用于替代 YOLOv11 主干网络中的 C3k2 模块,旨在提升其对小目标的表征能力与检测准确性。MSFPF 采用 3×3、5×5 和 7×7 三种不同尺寸的卷积核构建多尺度分支,并只在部分通道进行 5×5 和 7×7 卷积,其余通道则跳过复杂操作,直接参与特征融合。通过分组处理的方式减少冗余计算,同时保留每个尺度的原始特征信息,使得模型能够高效地学习丰富的特征表达。相比于传统全通道卷积,该方法显著降低了计算成本,同时仍能有效提取不同尺度的信息,确保小目标的关键特征得以保留。最后,利用 1×1 卷积进行通道信息交互,并通过残差连接与原始输入融合,以进一步增强对尺度变化的鲁棒性,提升小目标的特征表达能力和检测精度。MSFPF 模块表示为:

$$\boldsymbol{f} = \mathrm{Conv}_{1\times1}(\mathrm{Concat}(\varphi_{7\times7}(\boldsymbol{x}_{2a}),\boldsymbol{x}_{2b},\boldsymbol{x}_{1b})) + \boldsymbol{x} \quad (1)$$

$$\boldsymbol{x}_{1a},\boldsymbol{x}_{1b} = \mathrm{Split}(\varphi_{3\times3}(\boldsymbol{x})) \quad (2)$$

$$\boldsymbol{x}_{2a},\boldsymbol{x}_{2b} = \mathrm{Split}(\varphi_{5\times5}(\boldsymbol{x}_{1a})) \quad (3)$$

其中,$\boldsymbol{x}$ 表示输入的特征;$\varphi_{k\times k}$ 表示卷积核大小为 k 的卷积 Conv;$\boldsymbol{x}_{ia}$、$\boldsymbol{x}_{ib}$ 表示卷积后的特征拆分的结果;$\boldsymbol{f}$ 表示 MSFPF 模块输出结果。

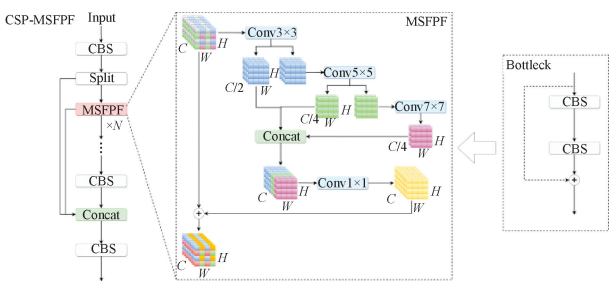

图 2　CSP-MSFPF 模块

Fig. 2　CSP-MSFPF module

1.2　FE-FPN 网络

在交通标志检测任务中,由于部分远距离交通标志尺寸较小或处于复杂背景下,其可辨性显著下降,易发生误检和漏检现象。从视觉感知的角度出发,该问题可通过增强模型对交通标志前景的关注能力、提升其与背景区域的区分能力来有效缓解。为此,本文设计了 FE-FPN 网络,如图 1 所示,该结构融合金字塔上下文信息,通过 RCM 模块[16]增强前景区域表达,并设计了 CSF 模块与 SFF 模块实现高效特征融合的同时增强关键特征表示,有效抑制复杂背景干扰。

1)金字塔上下文提取模块

金字塔上下文提取(pyramid context extraction,PCE)PCE 模块由 3 个 RCM 模块组成,负责从多尺度特征中提取上下文信息进行前景聚焦融合,即对骨干网络生成的不同尺度的、分辨率分别为 $[H/2\times W/2, H/4\times W/4, H/8\times W/8, H/16\times W/16, H/32\times W/32]$ 的特征 $\boldsymbol{P}1$、$\boldsymbol{P}2$、$\boldsymbol{P}3$、$\boldsymbol{P}4$、$\boldsymbol{P}5$ 进行处理:为保证网络计算的高效性,舍弃较大尺度的特征 $\boldsymbol{P}1$、$\boldsymbol{P}2$,对较小尺度的特征 $\boldsymbol{P}3$、$\boldsymbol{P}4$ 和 $\boldsymbol{P}5$ 进行平均池化(average pooling,AP)将其统一下采样至 $H/64\times W/64$,随后拼接生成金字塔特征 $\boldsymbol{F}$。

将特征 F 输入至多个堆叠的 RCM 中进行特征交互,并提取尺度感知语义特征。最后,将 PCE 模块获得的特征拆分为 $\boldsymbol{P}'3$、$\boldsymbol{P}'4$ 和 $\boldsymbol{P}'5$,用于与各阶段的特征进行融合。该过程可表示为:

$$\boldsymbol{F} = \mathrm{Concat}(\mathrm{AP}(\boldsymbol{P}3,8),\mathrm{AP}(\boldsymbol{P}4,4),\mathrm{AP}(\boldsymbol{P}5,2)) \quad (4)$$

$$\boldsymbol{P}'3,\boldsymbol{P}'4,\boldsymbol{P}'5 = \mathrm{Split}(\mathrm{PCE}(\boldsymbol{F})) \quad (5)$$

其中,$\mathrm{AP}(\boldsymbol{P},x)$ 表示将特征 $\boldsymbol{P}$ 以因子 x 进行下采样的平均池化操作,而 $\boldsymbol{P}'3$、$\boldsymbol{P}'4$ 和 $\boldsymbol{P}'5$ 代表具有金字塔上下文信息的特征表征。

2)RCM 模块

在 FE-FPN 中,RCM 模块使模型能够更多地关注前景。RCM 的结构如图 3 所示,由矩形自校准注意(rectangular self-calibration attention,RCA)、批量归一化(batch normalization,BN)和多层感知机(multilayer perceptron,MLP)组成。RCA 通过水平和垂直方向的池化提取轴向全局上下文信息,生成两个方向的特征向量并融合,实现对感兴趣矩形区域的建模,从而增强模型对目标区域的感知能力。该过程可以用数学表示为:

$$\boldsymbol{y} = \mathrm{AvgPool}_{\mathrm{C}}^{\mathrm{H}}(\boldsymbol{x}) \oplus \mathrm{AvgPool}_{\mathrm{C}}^{\mathrm{V}}(\boldsymbol{x}) \quad (6)$$

其中,$\boldsymbol{x}$ 表示输入的特征,$\mathrm{AvgPool}_{\mathrm{C}}^{\mathrm{H}}(\boldsymbol{x})$ 表示水平池化,$\mathrm{AvgPool}_{\mathrm{C}}^{\mathrm{V}}(\boldsymbol{x})$ 表示垂直池化,$\oplus$ 表示广播向量加法,$\boldsymbol{y}$ 表示水平轴向特征向量和垂直轴向特征向量融合的结果。

随后通过形状自校准函数 ϕ_{C} 对注意力区域进行调整,使其更加契合前景目标。该函数采用两个大核条带卷积进行解耦校准,分别处理水平方向和垂直方向的注意力图。首先,利用水平条带卷积调整水平方向的形状特征,增强每一行特征对前景目标的表达能力。随后,应用批量归一化并引入 ReLU 激活函数,以提升模型的非线性表示能力。接着,通过垂直条带卷积进一步增强垂直方向的形状感知能力。RCM 模块的核心优势在于,通过将水平与垂直方向的卷积解耦,使模型能够更灵活地适应不同形态的目标,从而在复杂场景中显著提升目标检测的精度与鲁棒性。形状自校准函数 ϕ_{C} 为:

$$\phi_{\mathrm{C}}(\boldsymbol{y}) = \delta(\mathrm{SConv}_{k\times1}(\mathrm{relu}(\mathrm{BN}(\mathrm{SConv}_{1\times k}(\boldsymbol{y}))))) \quad (7)$$

其中,SConv 表示大核条带卷积,k 表示条带卷积的核大小,BN 表示批处理归一化,$\mathrm{relu}(\boldsymbol{x})$ 表示 ReLU 函数,δ 表示 Sigmoid 函数。

在此基础上将原始输入 $\boldsymbol{x}$ 与形状自校准函数处理的输出 $\phi_{\mathrm{C}}(\boldsymbol{y})$ 进行融合,以提升前景特征在图像中的表达权重。该过程首先采用 3×3 深度可分离卷积进一步提取输

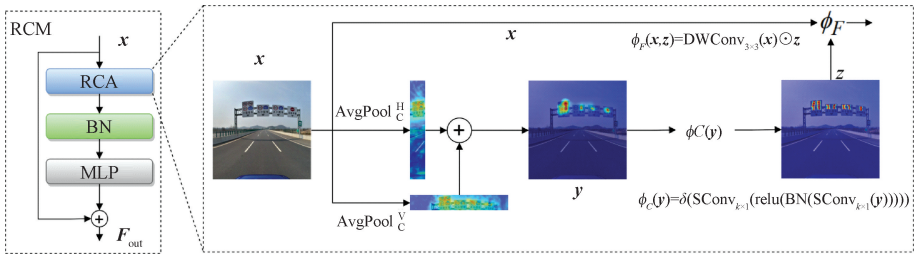

图 3　RCM 结构

Fig. 3　RCM structure

入特征的局部细节信息，继而通过 Hadamard 乘积实现对前景区域的加权强化。融合函数 $\phi_F(\boldsymbol{x},\boldsymbol{z})$ 可表示为：

$$\phi_F(\boldsymbol{x},\boldsymbol{z}) = \mathrm{DWConv}_{3\times3}(\boldsymbol{x})\odot\boldsymbol{z} \tag{8}$$

其中，$\mathrm{DWConv}_{3\times3}$ 表示具有内核大小为 3×3 的深度卷积，$\boldsymbol{x}$ 表示原始输入，$\boldsymbol{z}$ 表示上一步 $\phi_C(\boldsymbol{y})$ 得到的注意力特征，⊙ 表示 Hadamard 乘积。

最后，在 RCA 之后加入 BN 和 MLP 来细化特征，并采用残差连接进一步增强特征复用。该过程可表示为：

$$\boldsymbol{F}_{\mathrm{out}} = \mathrm{MLP}(\mathrm{BN}(\phi_F(\boldsymbol{x},\phi_C(\boldsymbol{y})))) + \boldsymbol{x} \tag{9}$$

其中，$\boldsymbol{F}_{\mathrm{out}}$ 表示 RCM 模块输出的特征。

3)CSF 模块和 SFF 模块

CSF 模块旨在实现不同尺度特征的高效对齐与深度融合。该模块负责将 RCM 模块获得的低层次特征 $\boldsymbol{x}_{\mathrm{h}}$ 与 PCE 模块产生的高层次特征 $\boldsymbol{P}'_i$ 进行融合。两个特征分别经过卷积操作后，使用 H-sigmoid 激活函数对特征图 $\boldsymbol{P}'_i$ 调整权重，并使用双线性插值将其上采样到与 $\boldsymbol{x}_{\mathrm{h}}$ 相同的空间尺寸。随后，将两组特征进行融合，以实现低层与高层语义信息的有效整合。CSF 模块可表示为：

$$\mathrm{CSF}(\boldsymbol{x}) = \mathrm{Conv}(\boldsymbol{x}_{\mathrm{h}})\odot\rho(\sigma(\mathrm{Conv}(\boldsymbol{P}'_i))) \tag{10}$$

其中，Conv 为无激活函数的卷积操作；ρ 表示双线性插值；σ 为 H-sigmoid 函数；⊙ 表示逐元素乘法；CSF($\boldsymbol{x}$) 表示融合后的输出特征图。

SFF 模块主要负责不同语义层级前景特征的有效整合。该模块首先采用双线性插值对来自不同阶段的特征图进行空间尺度对齐，确保各通道间的信息能够在统一尺度下充分交互。随后，利用 1×1 卷积进一步优化通道维度，使多尺度信息得以充分融合。最终，通过逐元素融合操作，实现不同语义层级信息的有效融合，进一步提升模型对前景区域的判别能力。可表示为：

$$\mathrm{SFF}(\boldsymbol{x}) = \boldsymbol{x}_{\mathrm{h}} + \mathrm{Conv}_{1\times1}(\rho(\boldsymbol{x}_{\mathrm{l}})) \tag{11}$$

其中，$\boldsymbol{x}_{\mathrm{h}}$ 表示 RCM 模块获得的特征；$\boldsymbol{x}_{\mathrm{l}}$ 表示 CSF 模块获得的特征；$\mathrm{Conv}_{1\times1}$ 表示 1×1 的卷积操作；ρ 表示使用双线性插值将特征图上采样到与 $\boldsymbol{x}_{\mathrm{h}}$ 相同的空间尺寸；SFF($\boldsymbol{x}$) 表示融合后的输出特征图。

FE-FPN 网络通过引入 PCE 模块与 RCM 模块，构建了兼具语义丰富性与空间自适应性的特征表达基础。在此基础上，进一步设计 CSF 模块与 SFF 模块，分别实现不同语义层级特征的深度融合与空间尺度的一致对齐，有效强化前景区域的特征表达能力。FE-FPN 整体提升了模型对关键目标的聚焦能力，在复杂背景干扰与远距离小目标检测任务中展现出更强的鲁棒性与识别性能。

1.3　HWD 改进下采样

传统的下采样虽然能够扩大感受野并降低数据维度，但由于下采样过程中的信息丢失导致边界信息模糊、相似交通标志难以区分。为此，本文引入了 Haar 小波下采样[17]，以替换原有的下采样卷积层。HWD 通过频域变换的方式提取不同尺度的特征，简化计算的同时保留更多的边缘与纹理信息，有效减少了下采样过程中的信息损失，从而提升了小目标检测的鲁棒性和检测精度。

HWD 模块由无损特征编码模块和特征表示学习模块组成，如图 4 所示。其中，无损特征编码模块主要通过 Haar 小波变换对输入特征进行转换，同时降低其空间分辨率。Haar 小波变换在保留完整信息的基础上实现特征降维，有助于降低计算复杂度，同时保留关键的结构特征。H_0 和 H_1 分别代表低通和高通分解滤波器，分别用于从图像中提取近似信息和高频信息；符号 ↓2 表示下采样操作。特征表示学习模块由标准卷积层、BN 层和 ReLU 激活函数组成，旨在进一步提取判别性特征并抑制冗余信息，增强模型对关键信息的表达能力。对于二维图像的 Haar 分解，其本质可视为分别对所有列和所有行执行一维 Haar 分解。其中，Haar 小波变换中第 1 级一维 Haar 变换所使用的小波基函数与尺度函数定义为：

$$\begin{cases}\phi_1(x) = 1/\sqrt{2}\,\phi_{1,0}(x) + 1/\sqrt{2}\,\phi_{1,1}(x)\\ \psi_1(x) = 1/\sqrt{2}\,\phi_{1,0}(x) - 1/\sqrt{2}\,\phi_{1,1}(x)\end{cases} \tag{12}$$

$$\phi_{j,k} = \sqrt{2^j}\,\phi(2^j x - k), k = 0,1,\cdots,2^j - 1 \tag{13}$$

其中，参数 j 和 k 分别表示为 Haar 基函数的阶段(或图像处理域中的尺度)和顺序(或 2D 图像的方向)。此外，$\phi_{0,0}(x)$ 可定义为：

$$\phi_{0,0}(x) = \phi_0(x) = \begin{cases}0, & x < 0\\ 1, & 0 \leqslant x < 1\\ 0, & x \geqslant 1\end{cases} \tag{14}$$

因此，可以使用 0 阶段 Haar 基函数来表示 1 阶段的 Haar 变换：

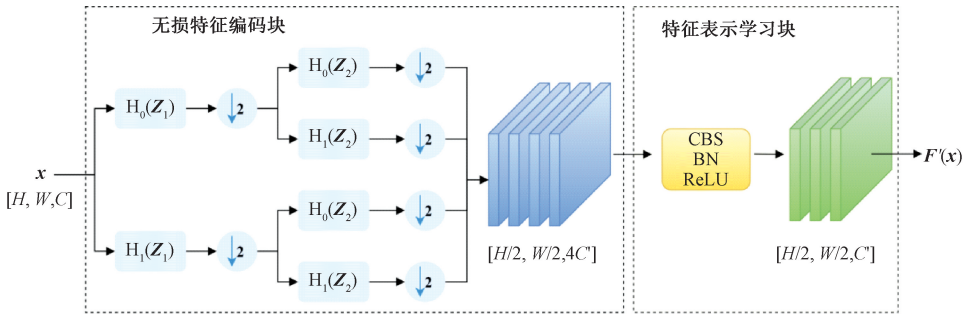

图 4　HWD 模块

Fig. 4　HWD module

$$\begin{cases}\phi_1(x)=\phi_0(2x)+\phi_0(2x-1)\\ \psi_1(x)=\phi_0(2x)-\phi_0(2x-1)\end{cases} \tag{15}$$

如图 5 所示,Haar 小波变换将原始图像有效地分解为 4 个分量,在空间分辨率减半的同时,分别捕获图像的低频信息和高频细节(包括水平方向、垂直方向和对角线方向)。该变换不仅能够保留更多的图像关键信息,还通过增加特征映射的通道数来维持信息的完整性,在降低分辨率的同时为后续网络层提供更加丰富且有效的特征输入。

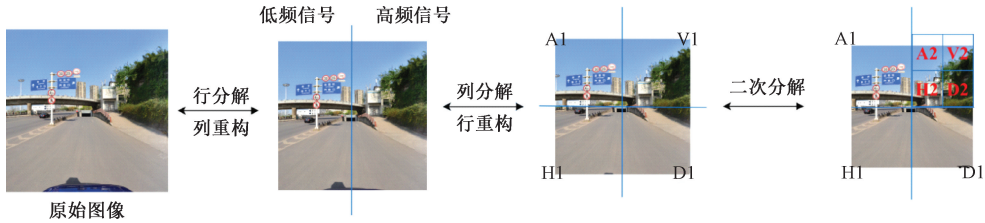

图 5　Haar 小波的运算过程

Fig. 5　The operation process of the Haar wavelet

HWD 能够在下采样过程中保留更多信息,增强了模型对多样化交通标志样式的感知能力,从而显著提高了交通标志识别的准确性和在复杂实际场景下的适应性。

1.4　优化损失函数

YOLOv11 使用完整交并比损失(complete intersection over union,CIoU)[18]作为损失函数,对训练数据的质量要求较高。然而,在实际交通标志检测任务中,数据集中各类目标尺寸大小不一、分布不均衡,存在一定比例的低质量标注,会对特征学习产生不利影响,导致模型收敛速度变慢,鲁棒性降低。为提升模型的定位能力并加快训练收敛速度,本文引入 WIoU[19] v3 作为边界框回归损失函数。一方面,当训练数据存在低质量标注时,WIoU v3 借助动态非单调聚焦机制,通过离群度评估锚框质量,有效抑制几何因素(如位置偏差、长宽比)对模型产生过大的负面影响;另一方面,对于预测框与真实框重合度较高的样本,该损失函数适当减弱几何因素的惩罚,引导模型以较少的训练干预实现更强的泛化能力。WIoU 示意图如图 6 所示,WIoU v3 的定义为:

$$L_{\mathrm{WIoUv3}}=r\times R_{\mathrm{WIoU}}\times L_{\mathrm{IoU}} \tag{16}$$

$$r=\frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{17}$$

$$R_{\mathrm{WIoU}}=\exp\left(\frac{(x-x_{\mathrm{gt}})^2+(y-y_{\mathrm{gt}})^2}{(w_{\mathrm{R}}^2+h_{\mathrm{R}}^2)^*}\right) \tag{18}$$

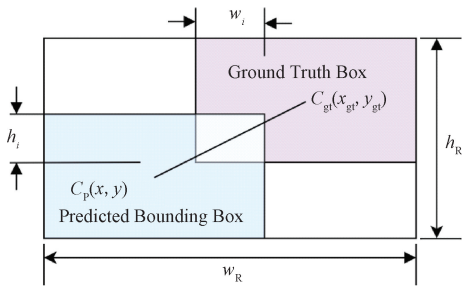

图 6　WIoU 示意图

Fig. 6　WIoU diagram

$$L_{\mathrm{IoU}}=1-\mathrm{IoU},\ L_{\mathrm{IoU}}\in[0,1] \tag{19}$$

其中,r 为非单调聚焦系数,δ、α 为可调节的超参数;(x,y)、$(x_{\mathrm{gt}},y_{\mathrm{gt}})$ 分别表示预测框和真实框的中心点坐标;w_{R}、h_{R} 表示最小包围框的宽高;R_{WIoU} 表示预测框和真实框的匹配程度;L_{IoU} 表示预测框与真实框的重叠程度;IoU 表示预测框与真实框的交并比;* 表示 w_{R} 和 h_{R} 从计算图中分离出来的操作,防止 R_{WIoU} 产生阻碍收敛的梯度。

离群度 β 用于衡量锚框质量,定义为:

$$\beta=\frac{L_{\mathrm{IoU}}^*}{\overline{L_{\mathrm{IoU}}}}\in[0,+\infty) \tag{20}$$

其中,L_{IoU}^* 为单调聚焦系数的梯度增益,与 L_{IoU}^* 定义相同,但其在训练过程中会依据目标情况的变化而动态计算

调整；$\overline{L_{IoU}}$ 为动量 m 的滑动平均值，对 L_{IoU}^{*} 进行动态归一化处理，维持合理的梯度更新强度，提升训练效率。离群度越小表示锚框质量越高，对于这类高质量锚框，分配较小的梯度增益，有助于边界框回归将关注点集中在普通质量的锚框上，提升整体检测性能。同时，对于离群度较大的锚框即低质量样本，同样分配较小的梯度增益，可有效抑制其产生较大的有害梯度，避免对模型训练造成干扰。

WIoU v3 采用合理的梯度增益分配策略动态优化高质量和低质量锚框在损失中的权重，使得模型专注于平均质量样本，提高了模型的整体性能。

1.5　基于 LAMP 分数的通道剪枝

模型剪枝技术是一项关键的模型压缩与加速方法，能够显著提高计算效率并增强模型在资源受限环境中的部署能力。该技术通过识别并移除神经网络中的冗余参数或结构组件，在尽量保持模型精度的前提下，大幅减少参数数量、降低计算开销，并提升推理速度。

采用 LAMP 算法[20]实施剪枝操作。与其他剪枝方法不同，LAMP 属于全局剪枝方法，能有效避免层崩溃问题。LAMP 在评估权重上使用了基于计算目标连接的权重幅值的平方的方法，LAMP 分数定义为：

$$\text{score}(u;W)=\frac{(W[u])^2}{\sum_{v\geqslant u}(W[v])^2} \tag{21}$$

其中，$W[u]$、$W[v]$ 为网络中第 u、v 个连接的权重。$(W[u])^2$ 表示目标连接的权重幅值的平方；$\sum_{v\geqslant u}(W[v])^2$ 表示同一层中从目前连接开始，该层所有具有更大权重幅度连接的平方和。

2　实验结果与分析

2.1　实验环境与参数设置

本实验所用操作系统为 Ubuntu 20.04，硬件配置：CPU 为 Intel(R) Xeon(R) Platinum 8462Y＋ CPU@2.80 GHz，GPU 为 NVIDIA RTX 4090，显存 24 GB。深度学习框架版本为 Pytorch2.2.2＋cu121。实验训练参数如表 1 所示。

表 1　实验参数设置

Table 1　Experimental parameter setting

参数	数值
输入图像	640×640
批量大小	16
迭代次数	500
动量因子	0.937
权重衰减	0.000 5
初始学习率	0.01
优化器	SGD
线程数	4
是否预训练	否

2.2　实验数据集

TT100K 数据集[21]图像来源于腾讯地图街景数据，包含 10 万张分辨率为 2 048×2 048 pixels 的图像和约 3 万个真实交通标志实例，目标尺寸范围为 16×20 pixels～160×160 pixels，覆盖多种光照与天气条件，具备较高场景复杂性与挑战性。数据集中共包含 232 类交通标志，其中 201 类具有标注，且类别分布极不均衡。为提高训练稳定性与模型泛化能力，本文在预处理阶段剔除实例数不足 100 的类别，并按 7∶2∶1划分为训练集(6 598 张)、验证集(1 889 张)和测试集(970 张)，兼顾图像数量与类别分布的均衡性。

CTSDB2021 数据集[22]包含 17 856 张分辨率范围在 600×900 pixels～1 024×768 pixels 的图像，目标尺寸分布在 20×20 pixels～573×557 pixels。标志类型主要划分为 3 类，分别为警告类、禁止类与强制类。

2.3　评价指标

本文实验采用准确率(precision，P)、召回率(recall，R)、平均准确度均值(mean average precision，mAP)作为算法检测性能的评估指标。计算公式为：

$$P=\frac{\text{TP}}{\text{TP}+\text{FP}} \tag{22}$$

$$R=\frac{\text{TP}}{\text{TP}+\text{FN}} \tag{23}$$

$$\text{AP}=\int_0^1 P(R)\,\text{d}R \tag{24}$$

$$\text{mAP}=\frac{\sum_{i=1}^{n}\text{AP}_i}{n} \tag{25}$$

其中，TP 表示正确预测的正样本数量，FP 表示错误预测的负样本数量，FN 表示错误预测的正样本数量。AP 表示模型检测单一类别的准确度，mAP 表示对所有类别求平均 AP 值。

此外，使用参数量(Params)、浮点运算量(FLOPs)、内存占用(Model size)、推理延时(Latency)和每秒帧数(FPS)等指标评估模型复杂度和计算效率。其中 FLOPs 和 Params 用于评判模型的计算复杂度；Latency 即模型完成一次前向推理的时间，不包括图像预处理和非极大值抑制(non-maximum suppression，NMS)过程，用于评估网络的计算效率；FPS 代表模型的检测速度。

2.4　消融实验

为验证所提方法在交通标志检测任务中的有效性，设计了 4 组在相同环境与参数下的消融实验，结果如表 2 所示。实验 A 将原 C3k2 模块替换为 CSP-MSFPF 模块，mAP50 提升 1.1%，表明其在多尺度信息融合与小目标表征方面更具优势。实验 B 在此基础上引入 FE-FPN，mAP50 进一步提升 0.7%，有效增强前景感知与背景抑制能力。实验 C 加入 Haar 小波下采样，在 mAP50 提升 0.6%的同时降低了参数量与计算量，缓解了小目标信息在下采样过程中的损失。实验 D 进一步引入 WIoU v3 损

表 2　消融实验

Table 2　Ablation experiment

消融设置	CSP-MSFPF	FE-FPN	HWD	WIoU v3	LAMP	P/%	R/%	mAP50/%	mAP50:95/%	Params/($\times10^6$)	GFLOPs	Model size/MB	Latency/ms
Baseline	×	×	×	×	×	85.6	74.3	84.7	67.5	9.4	21.4	18.4	**3.40**
A	√	×	×	×	×	83.7	78.1	85.8	68.3	9.5	26.5	18.7	3.42
B	√	√	×	×	×	**87.4**	77.9	86.5	69.0	12.2	34.3	23.7	4.45
C	√	√	√	×	×	86.3	78.7	87.1	69.4	10.5	31.3	20.6	4.78
D	√	√	√	√	×	86.8	78.6	**87.8**	**70.0**	10.5	31.3	20.6	4.59
E	√	√	√	√	√	85.6	**79.0**	87.3	69.6	**3.8**	**15.5**	**7.6**	4.12

注:×为未添加模块;√为添加模块。加粗数据为最优值。

失函数,检测精度持续提升。实验 E 通过引入基于 LAMP 分数的通道剪枝策略,实现对改进模型的参数量与模型规模的有效压缩。综合结果表明,与基线模型相比,本文提出的改进算法在召回率提升 4.7%、mAP50 提升 2.6%、mAP50:95 提升 2.1%的同时,计算量降低 27.6%、参数量减少 59.6%、模型大小缩减 58.7%,各个方面均显著提升了模型的检测性能。

2.5　对比实验

1)不同模型对比实验

为进一步验证本文所提出算法在交通标志识别与检测任务中的优越性能,选取当前主流的 YOLO 系列模型进行对比实验。所有模型在相同的数据集上进行训练,训练参数与本文模型保持一致,差异性超参数均采用默认设置。各模型的对比实验结果如表 3 所示。TSD-YOLOs 在检测精度和模型效率之间实现了最优平衡。在仅 3.77×10^6 参数量、15.5 GFLOPs 计算量的前提下,TSD-YOLOs 实现了 87.3%的 mAP50 与 69.6%的 mAP50:95,显著优于 YOLOv5m、YOLOv8s 等主流模型。同时,TSD-YOLOs 模型的推理速度达到 242 fps,展现出较强的实时检测能力,尤其适用于对实时性要求高的设备。

表 3　不同模型的对比实验

Table 3　Comparison experiment of different models

算法	mAP50/%	mAP50:95/%	Params/($\times10^6$)	GFLOPs	Model size/MB	FPS/fps
YOLOv5m	82.7	64.6	21.01	48.4	40.6	329
YOLOv7	68.1	51.0	37.41	105.8	71.8	240
YOLOv8s	84.4	67.1	11.14	28.5	21.6	**400**
YOLOv9s	84.5	67.2	9.63	38.9	19.4	107
YOLOv10m	85.2	67.8	16.50	63.7	32.1	150
YOLOv11s	84.7	67.5	9.43	21.4	18.4	293
YOLOv12s	83.0	66.2	9.28	21.3	18.1	172
YOLOv13s	81.4	64.2	9.01	20.1	17.8	333
RT-DETR[23]	83.0	60.3	20.14	61.3	—	132
D-FINE[24]	87.1	67.2	19.22	56.5	—	—
DEIM[25]	**87.3**	67.4	19.22	56.5	—	—
RT-DETRmg[14]	83.1	—	16.70	51.4	—	169
文献[11]	81.7	—	7.74	21.3	—	45
文献[12]	84.4	64.6	**2.67**	—	—	—
Ghost-YOLO[13]	84.8	63.1	9.10	21.0	—	62
TSD-TOLOs	87.3	**69.6**	3.77	**15.5**	**7.6**	242

注:加粗数据为最优值。

在与 Transformer 架构检测器的对比中,TSD-YOLOs 同样表现出领先优势。相较于 RT-DETR[23],本文方法在 mAP50 和 mAP50:95 分别提升 4.3%和 9.3%;与 D-FINE[24]相比,本文方法分别提升 0.2%和 2.4%;相较于 DEIM[25],TSD-YOLOs 在 mAP50:95 提升 2.2%。此外,TSD-YOLOs 在计算复杂度方面显著低于 RT-DETR、D-FINE 和 DEIM,展现出更优的效率与实用性。

为体现 TSD-YOLOs 在交通标志检测领域的应用优

势，本文还将 TSD-YOLOs 与其他学者在交通标志检测领域的优秀算法进行对比。RT-DETRmg[14]是对 RT-DETR 的轻量化改进算法，而 TSD-YOLOs 在 mAP50 精度方面显著优于 RT-DETRmg，同时在模型参数量和计算开销上也表现出更高的效率。对于与基于 YOLOv11 改进的算法相比，TSD-YOLOs 的检测性能优于文献[11]和 Ghost-YOLO[13]；与文献[12]相比，虽然 TSD-YOLOs 的参数量略大，但 TSD-YOLOs 在 mAP50 和 mAP50:95 上分别领先 2.9%和 5.0%。

综合精度、计算开销、模型复杂度与检测效率等多个指标，TSD-YOLOs 展现出更优的综合性能，进一步验证了本文所提改进策略的有效性与先进性，为车载端交通标志实时检测提供了兼顾精度与效率的解决方案，具有较高的实际应用价值。

2)不同损失函数对比实验

为验证 WIoU 的有效性及其参数设置的合理性，将其与 GIoU[26]、DIoU[18]、EIoU[27]、SIoU[28]、ShapeIoU[29]、PIoU[30]进行对比，如表 4 所示。使用 WIoUv3($\alpha=1.4$ $\delta=5.0$)后模型的 mAP50、mAP50:95 分别提高 0.7%、0.6%。

表 4　损失函数的结果比较

Table 4　Comparison of loss functions results　(%)

损失函数	P	R	mAP50	mAP50:95
CIoU	86.3	78.7	87.1	69.4
DIoU	85.1	80.0	86.6	69.2
EIoU	86.3	77.2	86.1	69.2
SIoU	87.4	77.5	86.8	69.2
ShapeIoU	85.2	79.3	86.6	69.3
PIoU	84.2	79.7	87.0	69.6
GIoU	86.9	78.6	87.0	69.5
WIoU-v1	86.7	78.3	86.9	69.3
WIoU-v2	86.6	78.6	86.7	69.4
WIoUv3　$\alpha=1.7$　$\delta=2.7$	85.1	**80.4**	87.3	69.6
WIoUv3　$\alpha=1.9$　$\delta=3.0$	87.3	78.0	87.1	69.7
WIoUv3　$\alpha=1.6$　$\delta=4.0$	86.1	79.3	87.2	69.8
WIoUv3　$\alpha=1.4$　$\delta=5.0$	86.8	78.6	**87.8**	**70.0**
WIoUv3　$\alpha=1.2$　$\delta=6.0$	**87.9**	78.4	87.5	69.2

注：加粗数据为最优值。

3)不同剪枝率对比实验

不同剪枝率下模型参数量与精度的对比结果如表 5 所示。其中，剪枝策略后的数值(如 1.5)表示剪枝前后计算量的比值。以 LAMP2.0 为例，该值表示在采用 LAMP 策略的情况下移除了 50%的参数。实验结果表明，随着剪枝率的增加，模型参数量和计算量均呈下降趋势，但并不必然导致性能下降。例如，LAMP1.5 模型的精度未出现降低，说明同一层卷积核可能学习到大量相似特征，存在参数和计算冗余。在 LAMP2.0(50%剪枝率)下，模型参数量和体积显著减少，而检测性能基本保持不变，实现了轻量化与效率的平衡。综合考虑性能与模型压缩效果，本研究选择 50%的剪枝率，以在保证精度的同时显著降低模型体积。

2.6　泛化实验

为验证 TSD-YOLOs 算法的有效性与泛化能力，本文在 CCTSDB2021 交通标志检测数据集上与基线模型进行了对比实验。由表 6 可知，TSD-YOLOs 在各项性能指标上均优于 YOLOv11s，其中精度提升 0.6%、召回率提升 2.8%、mAP50 提升 1.6%、mAP50:95 提升 3.9%，充分体现了本文算法在复杂实际场景中的鲁棒性与良好的迁移能力。

2.7　可视化分析

1)特征可视化分析

本文在多种典型场景下对交通标志图像进行了可视化对比实验，涵盖远距离小目标、单一目标、多目标及复杂背景环境下的检测情况，并采用 Grad-CAM[31]生成热力图，如图 7 所示。图像第 1 列是原始输入图像，第 2 列为基线模型的热力图，第 3 列为本文改进模型 TSD-YOLOs 的热力图。

由于基线模型在小目标交通标志检测中存在特征提取不充分、下采样过程易导致关键信息丢失等问题，其注意力区域往往集中在交通标志的边缘，难以全面覆盖目标整体特征。尤其在第 2 行所示的复杂背景场景下，模型对小目标的感知能力明显不足，容易造成目标信息的遗漏。相比之下，TSD-YOLOs 能更准确地聚焦于目标的中心区域，且在远距离、多目标和复杂背景等多种场景下均展现

表 5　不同剪枝率的模型性能

Table 5　Model performance at different pruning rates

剪枝策略	mAP50/%	mAP50:95/%	Params/($\times10^6$)	GFLOPs	Model size/MB	FPS/fps
Base	**87.8**	70.0	10.59	31.3	20.6	217
LAMP1.5	87.7	**70.1**	5.16	20.8	10.3	222
LAMP1.7	87.7	70.0	4.43	18.4	8.9	230
LAMP1.9	87.6	69.8	3.92	16.2	7.9	236
LAMP2.0	87.3	69.6	3.77	15.5	7.6	242
LAMP2.1	86.1	68.8	3.48	14.0	7.1	251
LAMP2.3	85.7	68.2	**3.24**	**12.8**	**6.6**	**260**

表 6　算法改进前后在 CCTSDB 上的性能对比

Table 6　Performance comparison on CCTSDB before and after algorithm improvement　(%)

算法	*P*	*R*	mAP50	mAP50:95
YOLOv11s	96.0	92.5	96.6	75.8
TSD-YOLOs	**96.6**	**95.3**	**98.2**	**79.7**

注:加粗数据为最优值。

(a) 原始图像　(b) 基线模型　(c) TSD-YOLOs

(a) Original figure　(b) Baseline

图 7　改进模型与基线模型的热力图对比

Fig. 7　Comparison of the heatmaps of the improved model with the baseline model

出更强的感知能力和特征覆盖范围。这一优势主要得益于 CSP-MSFPF 模块对细粒度特征的增强、FE-FPN 对前景目标的有效聚焦以及 HWD 策略对高频细节的保留,有效提升了交通标志的检测精度和鲁棒性。

2)检测结果分析

为直观验证 TSD-YOLOs 在交通标志检测任务中的有效性,本文选取测试集中 4 张典型图像进行目标检测,并与基线模型进行可视化对比,如图 8 所示。

(a) 原始图像　(b) 基线模型　(c) TSD-YOLOs

(a) Original figure　(b) Baseline

图 8　TSD-YOLOs 与基线模型检测结果对比

Fig. 8　Comparison of the detection results between TSD-YOLOs and the baseline model

(1)基线模型存在类别误判问题:第 1 行图像中,尽管两种模型均检测出两个交通标志,但基线模型将“pm20”误识为“pl40”,而 TSD-YOLOs 借助 Haar 小波下采样在压缩特征同时保留关键细节,有效提升了对相似类别的辨识能力。

(2)基线模型漏检小尺度目标:第 2 行图像中,TSD-YOLOs 成功检测出两个不同尺度的目标,而基线模型漏检了小尺寸的“io”标志,说明 CSP-MSFPF 模块强化了多尺度特征的提取与融合,提升了小目标的检测效果。

(3)基线模型易受复杂背景干扰:TSD-YOLOs 在复杂场景中表现更稳健。第 3 行图像中,TSD-YOLOs 准确识别出背景干扰中的"p10",而基线模型未能检测;第 4 行图像中,得益于 FE-FPN 对前景特征的强化与背景抑制能力,TSD-YOLOs 识别出 8 个远距离交通标志,而基线模型仅检测出 6 个。

综上所述,TSD-YOLOs 在类别相似目标识别、复杂背景干扰抑制以及远距离小目标检测方面均显著优于基线模型,验证了本文提出方法在交通标志检测任务中的有效性与实用价值。

3　结　　论

针对小目标交通标志检测中存在的特征表示不足、背景干扰强以及下采样信息丢失等问题,本文从特征提取、特征融合、下采样策略以及损失函数等方面进行优化,提出了 TSD-YOLOs 算法。首先,提出 CSP-MSFPF 模块增强特征提取与跨尺度融合能力,提升模型对小目标的辨识效果;其次,设计 FE-FPN 结构,结合上下文引导与前景增强策略,强化关键区域特征表达,实现在复杂场景中前景增强与干扰抑制的协同优化;再次,采用 Haar 小波下采样替代传统卷积操作,在降低计算负担的同时更好地保留边缘细节,有效缓解小目标在下采样过程中的信息损失问题;同时,引入 WIoU v3 作为回归损失函数,通过动态非单调采样机制自适应调整正负样本的梯度分布,缓解样本标注质量不均衡问题,进一步提升了模型的定位精度;最后,在模型优化阶段引入基于 LAMP 分数的通道剪枝机制,以减少冗余参数并控制计算量,从而实现性能与效率的双重优化。实验结果表明,TSD-YOLOs 在 TT100K 数据集上相较基线模型的 mAP50 提升 2.6%,mAP50:95 提升 2.1%,召回率提升 4.7%,误检率和漏检率均显著下降;检测速度达到 242 fps,满足实时检测需求。在 CCTSDB 数据集上,TSD-YOLOs 在准确率和召回率上分别提升 0.6% 和 2.8%,mAP50 提升 1.6%,mAP50:95 提升 3.9%,进一步验证了该算法良好的检测性能与泛化能力。且 TSD-YOLOs 计算量降低 27.6%、模型大小压缩 58.7%、参数量减少 59.6%。后续研究将聚焦于提升模型在类别不均衡以及雾霾、低可见度等极端条件的检测能力,以满足实际应用中对不同环境的需求。

参考文献

[1] 杨斐,秦勃,李伟. 基于多帧视频图像的交通标志实时检测[J]. 计算机科学与应用,2017,7(5):463-472.
YANG F, QIN B, LI W. Real-time traffic sign detection based on multi-frame video images [J]. Computer Science and Application, 2017, 7(5): 463-472.

[2] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[3] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]//IEEE International Conference on Computer Vision, 2017: 2961-2969.

[4] CAI ZH W, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 6154-61629.

[5] SAPKOTA R, FLORES-CALERO M, QURESHI R, et al. YOLO advances to its genesis: A decadal and comprehensive review of the you only look once (YOLO) series[J]. Artificial Intelligence Review, 2025, 58(9): 274.

[6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//IEEE Conference on Computer Vision, 2016.

[7] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]// European Conference on Computer Vision, 2020: 213-229.

[8] 张莹,邓华宣,王耀南,等. 基于多通道特征融合学习的印制电路板小目标缺陷检测[J]. 仪器仪表学报,2024,45(5):10-19.
ZHANG Y, DENG H X, WANG Y N, et al. Small defects detection of PCB based on multi-channel feature fusion learning[J]. Chinese Journal of Scientific Instrument,2024,45(5):10-19.

[9] 艾强,冯永安,王灵超,等. 融合通道剪枝的轻量化 YOLOv8 交通标志检测算法[J/OL]. 电子测量技术,1-14 [2025-12-15]. https://link.cnki.net/urlid/11.2175.tn.20251124.1021.008.
AI Q, FENG Y AN, WANG L CH, et al. Lightweight YOLOv8 traffic sign detection algorithm incorporating channel pruning [J/OL]. Electronic Measurement Technology,1-14[2025-12-15]. https://link.cnki.net/urlid/11.2175.tn.20251124.1021.008.

[10] 俞林森,陈志国. 融合前景注意力的轻量级交通标志检测网络[J]. 电子测量与仪器学报,2023,37(1):21-31.
YU L S, CHEN ZH G. Lightweight traffic sign detection network with fused foreground attention[J]. Journal of Electronic Measurement and Instrumentation, 2023,37(1):21-31.

[11] TANG J Y, XU B, LI J, et al. Ghost-YOLO-GBH: A lightweight framework for robust small traffic sign detection via GhostNet and bidirectional multi-scale feature fusion[J]. Eng, 2025, 6(8): 196.

[12] 彭杰,于惠钧,夏梦,等.改进 YOLO11n 的小目标交通标志检测方法[J/OL].重庆理工大学学报(自然科学),1-11[2025-10-23]. https://link.cnki.net/urlid/50.1205.T.20250922.0935.004.
PENG J, YU H J, XIA M, et al. Improved YOLO11n traffic sign detection for small targets[J/OL]. Journal of Chongqing University of Technology (Natural Science), 1-11[2025-10-23]. https://link.cnki.net/urlid/50.1205.T.20250922.0935.004.

[13] CHU X S, ZHOU Z P, LIAO W P, et al. Ghost-YOLO: A lightweight traffic sign detection framework via GhostNetV3[C]//2025 4th Conference on Fully Actuated System Theory and Applications (FASTA), 2025: 1557-1562.

[14] WANG Y Q, CHEN J L, YANG B, et al. RT-DETRmg: A lightweight real-time detection model for small traffic signs[J]. The Journal of Supercomputing, 2025, 81(1): 307.

[15] CHEN J R, KAO S H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.

[16] NI Z L, CHEN X H, ZHAI Y J, et al. Context-guided spatial feature reconstruction for efficient semantic segmentation[C]//European Conference on Computer Vision, 2024: 239-255.

[17] XU G P, LIAO W T, ZHANG X, et al. Haar wavelet downsampling: A simple but effective downsampling module for semantic segmentation[J]. Pattern Recognition, 2023, 143: 109819.

[18] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]//AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

[19] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[J]. ArXiv preprint arXiv:2301.10051, 2023.

[20] LEE J, PARK S, MO S, et al. Layer-adaptive sparsity for the magnitude-based pruning[C]//International Conference on Learning Representations, 2020.

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

[22] ZHANG J M, ZOU X, KUANG L D, et al. CCTSDB 2021: A more comprehensive traffic sign detection benchmark[J]. Human-centric Computing and Information Sciences, 2022, 12, DOI:10.22967/HCIS.2022.12.023.

[23] ZHAO Y A, LYU W Y, XU S L, et al. Detrs beat yolos on real-time object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

[24] PENG Y S, LI H B, WU P X, et al. D-FINE: Redefine regression task in DETRs as fine-grained distribution refinement[J]. ArXiv preprint arXiv: 2410.13842, 2024.

[25] HUANG S H, LU ZH C, CUN X D, et al. Deim: Detr with improved matching for fast convergence[C]//Computer Vision and Pattern Recognition Conference, 2025: 15162-15171.

[26] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.

[27] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.

[28] GEVORGYAN Z. SIoU loss: More powerful learning for bounding box regression[J]. ArXiv preprint arXiv:2205.12740, 2022.

[29] ZHANG H, ZHANG SH J. Shape-IoU: More accurate metric considering bounding box shape and scale[J]. ArXiv preprint arXiv:2312.17663, 2023.

[30] CHEN Z M, CHEN K AN, LIN W Y, et al. PIoU loss: Towards accurate oriented object detection in complex environments[C]//Computer Vision-ECCV 2020: 16th European Conference. Springer International Publishing, 2020: 195-211.

[31] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//IEEE International Conference on Computer Vision, 2017: 618-626.

作者简介

刘丽,博士,副教授,主要研究方向为人工智能和计算机视觉。

E-mail:liuli@ncepu.edu.cn

张初夏(通信作者),硕士研究生,主要研究方向为目标检测。

E-mail:zhang_chu_xia@163.com

张硕,硕士研究生,主要研究方向为遥感目标检测。

E-mail:1442860914@qq.com

李宇健,硕士研究生,主要研究方向为目标检测。

E-mail:lyjyjs12138@ncepu.edu.cn

王强,博士,讲师,主要研究方向为计算机视觉、深度学习、持续学习。

E-mail:52152569@ncepu.edu.cn