

doi:10.3969/j.issn.1003-3106.2025.11.002

引用格式:李明芳. DMP-YOLO:面向自动驾驶的多尺度目标检测算法[J]. 无线电工程, 2025, 55(11): 2142-2152. [LI Mingfang. DMP-YOLO: A Multi-scale Object Detection Algorithm for Autonomous Driving[J]. Radio Engineering, 2025, 55(11): 2142-2152.]

## DMP-YOLO: 面向自动驾驶的多尺度目标检测算法

李明芳

(贵州大学 大数据与信息工程学院, 贵州 贵阳 550025)

**摘要:** 自动驾驶场景下的目标检测面临复杂环境干扰、多尺度目标分布及目标遮挡等挑战, 现有算法在特征融合能力、细节表征精度和定位回归性能方面仍存在不足。为此, 提出了一种改进的 YOLOv8 检测算法 DMP-YOLO。使用多分支辅助特征金字塔网络 (Multi-Branch Auxiliary Feature Pyramid Network, MAFPN) 优化原颈部结构, 增强复杂交通场景中的多尺度特征融合能力; 在骨干网络中提出 C2f\_DEConv 模块, 将标准卷积替换为细节增强卷积 (Detail-Enhanced Convolution, DEConv), 通过高频特征保持与局部纹理强化, 显著提升对小尺度车辆及遮挡目标的细节捕捉能力; 引入 PIoUv2 (Powerful Intersection over Union version 2) 损失函数优化改进边界框损失, 通过动态尺度敏感因子与几何约束优化, 改善目标边界框的回归精度。在 KITTI 数据集上的实验表明, DMP-YOLO 的各项性能指标都有显著提升, 其中 mAP@0.5 达到 89.0%, 较基线 YOLOv8 提升了 2.6 个百分点, mAP@0.5:0.95 提升了 2.9 个百分点, 为自动驾驶场景下的高精度实时检测提供了有效解决方案。

**关键词:** 自动驾驶; YOLOv8; 多分支辅助特征金字塔网络; PIoUv2 损失函数

中图分类号: TP183; TP391.41

文献标志码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1003-3106(2025)11-2142-11



## DMP-YOLO: A Multi-scale Object Detection Algorithm for Autonomous Driving

LI Mingfang

(School of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

**Abstract:** Target detection in autonomous driving scenarios faces challenges such as complex environmental interference, multi-scale target distribution and target occlusion, and existing algorithms are still deficient in feature fusion capability, detail characterization accuracy and localization regression performance. To this end, an improved YOLOv8 detection algorithm, DMP-YOLO, is proposed. The original neck structure is optimized using Multi-Branch Auxiliary Feature Pyramid Network (MAFPN) to enhance the multi-scale feature fusion capability in complex traffic scenarios; C2f\_DEConv is proposed in backbone network module, which replaces the standard convolution with Detail-Enhanced Convolution (DEConv) to significantly improve the detail capturing ability of small-scale vehicles and occluded targets through high-frequency feature preservation and local texture enhancement; the Powerful Intersection over Union version 2 (PIoUv2) loss function is introduced to optimize the improved bounding-box loss, which improves the regression accuracy of the target bounding-box through the optimization of dynamic scale-sensitive factors and geometric constraints. Experiments on the KITTI dataset demonstrate that DMP-YOLO achieves significant improvements across all key performance metrics, with mAP@0.5 reaching 89.0% (2.6% improvement compared with the baseline YOLOv8) as well as 2.9% improvement for mAP@0.5:0.95, which provides an effective solution for high-precision real-time detection in autonomous driving scenarios.

**Keywords:** autonomous driving; YOLOv8; MAFPN; PIoUv2 loss function

### 0 引言

随着计算机视觉和人工智能技术的快速发展, 基于视觉感知的交通场景理解在环境感知系统中扮

演着日益重要的角色<sup>[1]</sup>。其中, 车辆和行人检测作为核心任务之一, 直接影响着自动驾驶决策的准确性和道路安全监控的可靠性。实际道路环境中, 检测算法需要应对多变的光照条件、复杂的背景干扰、

收稿日期: 2025-07-10

基金项目: 国家自然科学基金(62166006, 52168011)

Foundation Item: National Natural Science Foundation of China (62166006, 52168011)

目标遮挡以及尺度变化等挑战,这对检测模型的鲁棒性和适应性提出了更高要求。提升交通目标的检测性能不仅有助于增强自动驾驶系统的环境感知能力,还能为智慧城市建设中的交通管理和安全预警提供关键技术支撑,具有重要的研究价值和应用前景<sup>[2]</sup>。

在深度学习的推动下,目标检测领域形成了2种主流技术路径:两阶段检测算法与单阶段检测算法。两阶段算法以R-CNN系列算法<sup>[3]</sup>为典型代表,通常包含区域建议生成和区域分类回归2个主要步骤。这种分步处理方式虽然保证了较高的检测精度,但计算过程相对冗长,导致检测速度受限。单阶段算法,以YOLO<sup>[4]</sup>和SSD<sup>[5]</sup>为代表,则将检测视为一个统一的回归任务,直接预测图像中目标的类别和位置。这种端到端的设计实现了极高的检测速度,非常适合于实时性要求高的场景。尽管在理论精度上可能略逊于两阶段算法,但单阶段算法凭借其效率优势,在诸多实际应用中展现出强大的竞争力。

近年来,针对交通场景目标检测的改进研究取得了显著进展,众多学者从不同角度对现有算法进行了优化。刘宇等<sup>[6]</sup>通过构建CTNet主干网络和引入Vit-Block,有效提升了密集目标和小尺度物体的检测性能;Jiao等<sup>[7]</sup>设计的RS-YOLO算法创新性地提出了多尺度路径聚合特征金字塔网络(Multi-Scale Path Aggregation Feature Pyramid Network, MSPAFPN),并引入加权混洗融合算法,显著提升了层间特征交互能力。杨磊等<sup>[8]</sup>通过特征感受野融合模块(Enhanced Feature Field of View Module, EFFVM)和浅层特征检测头,增强了模型对遮挡目标和小目标的检测能力;杨博等<sup>[9]</sup>在YOLOv8n中嵌入可变形卷积(Deformable Convolution Network, DCN)和全局注意力机制(Global-attention Mechanism, GAM),进一步提高了复杂背景下的特征提取能力;李康等<sup>[10]</sup>通过在颈部网络引入全维动态卷积与GAM,并结合MPDIoU损失函数优化定位精度,显著提升了模型的检测性能与速度。针对无人机视角的特殊性,Fan等<sup>[11]</sup>设计了一种新型多尺度特征融合模式,通过在特征金字塔网络(Feature Pyramid Network, FPN)中引入上采样操作,有效解决了特征传播退化问题;马荣贵等<sup>[12]</sup>通过引入Transformer与注意力机制对网络结构进行增强,显著提升了模型对小目标和遮挡目标的检测性能。此外,Wang等<sup>[13]</sup>提出的模糊感知注意力网络(Blur-aware Attention

Network, BANet)通过多通道注意力模块(Multichannel Attention, MCA)和特征融合策略提升了小目标检测效率。Khalili等<sup>[14]</sup>通过广义特征金字塔网络(Generalized Feature Pyramid Networks, GFPNs)和高效多尺度注意力(Efficient Attention Module, EMA)进一步优化了小目标检测性能。Chen等<sup>[15]</sup>设计的VisioSignNet通过双交互式神经网络架构,创新性地解决了交通标志检测中的长距离像素依赖问题。在轻量化方面,Mahaur等<sup>[16]</sup>提出基于泰勒准则排序的通用网络剪枝策略(Versatile Network Pruning, VNP),在保持检测精度的同时大幅减小了模型复杂度。古佳欣等<sup>[17]</sup>利用动态分组卷积和Transformer重构主干网络,显著降低了计算复杂度。Bie等<sup>[18]</sup>则通过深度可分离卷积和C3Ghost模块对YOLOv5进行轻量化改进,显著降低了模型参数量。Cai等<sup>[19]</sup>通过引入GhostNet轻量化结构和混合局部通道注意力机制,在保持检测精度的同时显著降低了模型复杂度,有效提升了复杂交通场景下的检测效率。

然而,现有方法在复杂交通场景中仍存在多尺度目标检测中的特征融合不充分、细节特征提取能力较弱及遮挡场景下的目标定位精度不足等问题,检测精度还有待进一步提高。为此,本文提出改进算法DMP-YOLO,在保证实时性的同时提升对遮挡目标和小尺度目标的检测鲁棒性,主要包括以下3点改进:

① 将MAFPN网络结构引入YOLOv8的颈部设计,构建了具有跨尺度动态交互能力的特征金字塔,显著提升了模型对多尺度交通目标的特征融合效果;

② 采用DEConv模块重构C2f特征提取结构,通过高频特征保持机制强化了对小尺度目标和细节纹理的特征表达能力;

③ 通过引入PIoUv2损失函数改进边界框回归过程,利用其动态尺度感知特性有效提升了遮挡目标的定位准确率。

## 1 YOLOv8网络结构

YOLOv8作为YOLO系列目标检测算法的经典版本,在继承YOLOv5优秀架构的基础上进行了多项创新性改进。该算法采用模块化设计思想,整体架构延续了YOLO家族经典的“输入-骨干-颈部-头部”四段式设计,如图1所示。

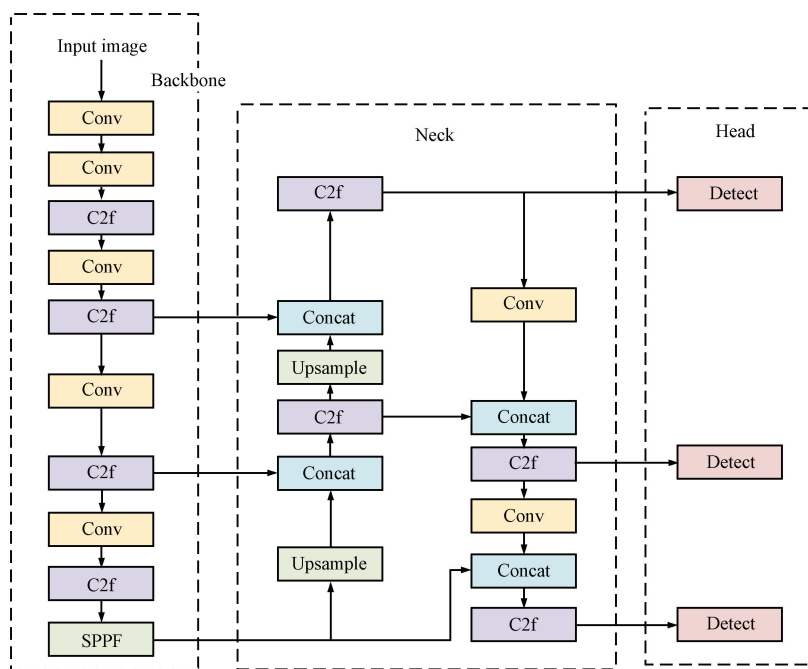


图 1 YOLOv8 网络结构  
Fig. 1 YOLOv8 network structure

在输入预处理阶段, YOLOv8 采用自适应图像缩放技术动态调整输入尺寸, 结合 Mosaic 数据增强策略有效提升了模型对不同尺度目标的适应能力, 并在训练后期自动关闭 Mosaic 增强以避免训练不稳定问题。骨干网络 (Backbone) 是特征提取的核心, 主要由 CBS 模块、C2f 模块和快速空间金字塔池化 (Spatial Pyramid Pooling Fast, SPPF) 模块 3 个核心组件构成。其中 CBS 模块由标准卷积、批归一化 (Batch Normalization, BN) 和 SiLU 激活函数组成, 负责基础特征提取和下采样操作; C2f 模块通过梯度分流和残差连接设计, 在保持轻量化的同时实现了更丰富的梯度信息流动; SPPF 模块则通过级联的最大池化操作实现多尺度感受野的特征融合。这种模块化设计使得骨干网络在特征提取能力和计算效率之间取得了良好平衡。特征融合模块采用创新的双向特征金字塔结构, 将 FPN 自上而下的语义信息传递与路径聚合网络 (Path Aggregation Network, PAN) 自下而上的空间细节补充有机结合。该部分移除了传统上采样操作中的冗余卷积层, 并使用 C2f 模块替代原有 C3 模块, 显著提升了特征融合效率。检测头部分采用解耦式设计, 将分类任务和回归任务分离为独立分支, 不仅降低了模型复杂度, 还解决了

传统耦合检测头在复杂场景中的定位不准确问题。同时, YOLOv8 摒弃了传统的基于锚框的检测方式, 采用无锚点机制直接预测目标位置和尺寸, 大幅简化了参数调优过程。

基于不同应用需求, YOLOv8 提供 5 种尺度模型, 其中 YOLOv8n 作为基准模型在保持较高检测精度的同时具有最小参数量, 特别适合部署在计算资源受限的边缘设备上。这种灵活的架构设计使 YOLOv8 能够适应从轻量级到高精度检测的多样化应用场景。

## 2 DMP-YOLO 算法

针对现有 YOLOv8 模型在复杂场景下多尺度特征融合不充分、细节特征提取能力不足以及边界框回归精度受限等问题, 本文提出一种改进的 YOLOv8 网络模型。在颈部网络部分采用 MAFPN 替代原 FPN 结构, 通过引入跨尺度特征交互机制和自适应权重分配策略, 增强模型对不同尺度目标的特征融合能力, 提升密集场景下的语义信息关联性。其次, 在骨干网络优化方面, 提出基于细节增强卷积的 C2f\_DEConv 模块, 通过将标准卷积替换为具有高频特征保持特性的细节增强卷积 DEConv, 重构原

有 C2f 模块的结构设计,使网络能够在保持轻量化的同时更有效地捕捉目标的边缘细节和纹理特征,显著提升对小尺寸目标的特征提取能力。最后,在损失函数优化上采用 PIoUv2 替换 CIoU 损失函数,通

过引入动态尺度敏感因子和几何约束优化项,在保持交并比(Intersection over Union, IoU)度量优势的基础上进一步优化模型对遮挡目标和非常规比例目标的边界框回归精度。DMP-YOLO 网络结构如图 2 所示。

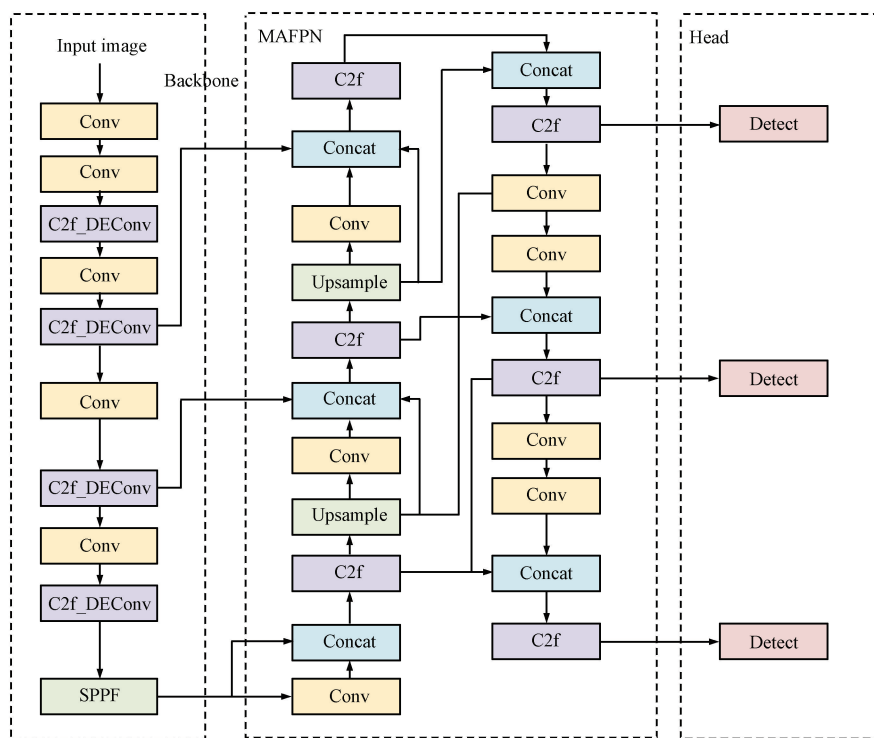


图 2 DMP-YOLO 网络结构

Fig. 2 DMP-YOLO network structure

### 2.1 MAFPN 模块

在自动驾驶的实际应用中,道路场景中的目标具有显著的尺度多样性,包括近处的大型车辆、远处的行人以及各种尺寸的交通标志等。传统 FPN 在特征融合过程中存在浅层空间信息丢失和跨尺度特征交互不足的问题,难以有效应对道路目标的多尺度检测挑战。因此,引入 MAFPN<sup>[20]</sup> 优化改进颈部网络,其核心在于通过浅层辅助融合(Superficial Assisted Fusion, SAF)模块强化高分辨率特征的保留与利用,确保小尺度目标的细节信息不被丢失;同时采用高级辅助融合(Advanced Assisted Fusion, AAF)模块构建动态非线性特征交互机制,显著提升多尺度特征的融合效率。这种双路径增强设计有效解决了传统方法中特征融合不充分的本质问题,使网络对不同尺度目标都具备均衡的检测能力。

SAF 模块通过双向连接机制将骨干网络的浅层特征与深层特征进行融合,有效保留了小目标检测所需的细节信息,其结构如图 3 所示。具体实现中,SAF 采用 1×1 卷积控制浅层特征的通道数,并通过 Concat 操作将下采样的高层特征、同层特征和上采样的低层特征进行融合,其输出特征可表示为:

$$P'_n = \text{Concat}(\delta(C(\text{Down}(P_{n-1}))), P_n, U(P'_{n+1})). \quad (1)$$

AAF 模块在深层网络中构建了更密集的多尺度特征交互路径,通过整合高分辨率浅层特征、低分辨率浅层特征以及相邻层特征,显著提升了中等尺度目标的检测性能,结构如图 4 所示。AAF 采用均衡的通道控制策略,确保各层特征对输出的贡献权重一致,其输出特征表示为:

$$P''_n = \text{Concat}(\delta(C(\text{Down}(P'_{n-1}))), \delta(C(\text{Down}(P''_{n-1})), P'_n, C(U(P'_{n+1}))). \quad (2)$$

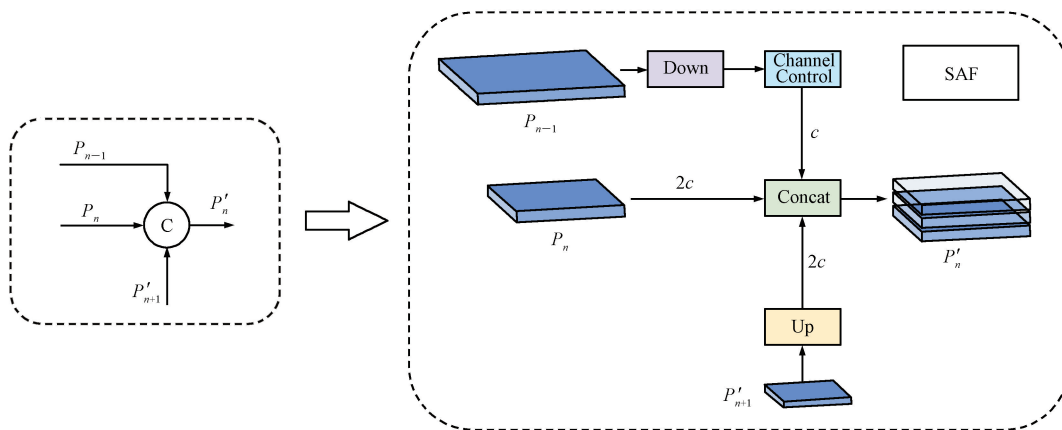


图 3 SAF 模块结构

Fig. 3 Architecture of SAF module

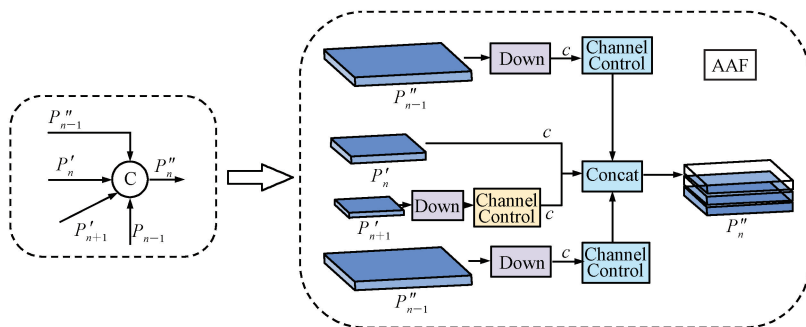


图 4 AAF 模块结构

Fig. 4 Architecture of AAF module

2.2 C2f-DEConv 模块

针对复杂交通场景中目标尺度变化大、细节特征易丢失的问题,传统 C2f 模块虽然通过梯度分流和多分支结构实现了高效的特征传递,但其标准卷积操作在处理交通场景中的小目标和边缘特征时存在明显的局限性。特别是在远距离行人、自行车等小尺度目标的检测任务中,原始模块容易丢失关键

的纹理和轮廓信息,导致定位精度下降。本文提出采用细节增强卷积( DEConv)<sup>[21]</sup>改进 YOLOv8 的 C2f 模块,构建新型 C2f\_DEConv 结构,如图 5 所示。DEConv 采用可重新参数化的卷积核设计,通过计算输入特征图与不同卷积核处理后特征的差异来增强细节感知能力,在保持特征提取能力的同时有效控制了参数增长。

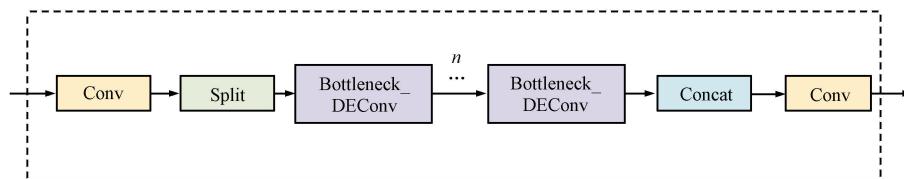


图 5 C2f\_DEConv 模块

Fig. 5 C2f\_DEConv module

在模块设计上,首先构建了 Bottleneck\_DEConv 基础单元,通过 2 个级联的 DEConv 操作实现特征变换,如图 6 所示。与传统卷积不同,DEConv 通过计算输入特征图与其经过不同卷积核处理后的特征

图之间的差异,有效地增强了网络对边缘和纹理等细节特征的敏感性。这种差异特征提取机制特别适用于交通场景中的小目标检测,如行人和自行车等。基于此,进一步构建了 C2f\_DEConv 模块,该模块保

留了原始 C2f 的多分支结构特性,但通过引入可重新参数化技术,在保持强大特征提取能力的同时显著降低了模型的复杂度。在训练阶段,模块采用多分支结构学习丰富的特征表示;而在推理阶段,这些分支可以通过参数融合技术合并为单一通路,既保证了性能又不增加推理时的计算负担。

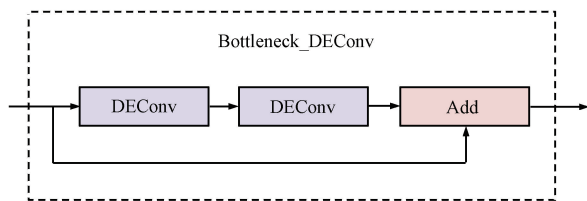


图 6 Bottleneck\_DEConv 模块  
Fig. 6 Bottleneck\_DEConv module

### 2.3 损失函数改进

在目标检测任务中,边界框回归的精度直接影响检测性能。YOLOv8 原本采用的 CIoU 损失函数通过引入中心点距离惩罚和长宽比约束,相比传统的 IoU 损失有了显著改进。其数学表达式为:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b_{gt})}{c^2} + \alpha\nu, \quad (3)$$

$$\alpha = \frac{\nu}{(1 - IoU) + \nu}, \quad (4)$$

$$\nu = \frac{4}{\pi^2} \left( \arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2, \quad (5)$$

式中:IoU 表示预测框与真实框的交并比,  $\rho^2(b, b_{gt})$  表示二者中心点的欧氏距离平方,  $c$  表示最小外接矩形的对角线长度,  $\alpha$  表示权重系数,  $\nu$  用于衡量长宽比的一致性,  $w_{gt}$  和  $h_{gt}$  表示真实框的宽度和高度,  $w$  和  $h$  表示预测框的宽度和高度。尽管 CIoU 在一般场景表现良好,但在复杂交通场景下存在明显不足。首先,其固定的几何惩罚项会导致锚框在回归过程中不必要地扩大尺寸,这种现象在车辆密集场景下尤为明显,容易造成误检。其次,对远处小目标的检测效果欠佳,因为小目标的中心点偏差和长宽比变化对损失函数的贡献相对较小,导致梯度信号较弱。此外,在目标部分遮挡情况下,CIoU 的刚性几何约束会使回归过程不稳定。

针对动态遮挡场景下目标定位精度不足的问题,本文提出采用 PIoUv2 损失函数进行改进,其通过 2 个关键创新有效解决了传统 CIoU 在遮挡目标边界框回归中的定位偏差问题。首先是引入了尺寸自适应惩罚机制,该机制根据目标框的实际尺寸动态调整惩罚权重,其公式表示为:

$$\lambda = \lambda_0 \cdot (w_{gt} \cdot h_{gt})^\gamma, \quad (6)$$

式中:  $\gamma$  为尺度敏感因子,经实验验证取 0.5 时效果最佳。这种设计使得模型能够自适应地处理不同尺度的目标,特别是避免了小目标检测中锚框过度扩大的问题。第二个创新是非单调注意力机制,该机制通过 Sigmoid 函数构建注意力权重,其表达式为:

$$Attention(q) = \frac{1}{1 + \exp(-\alpha(q - \beta))}, \quad (7)$$

式中:  $q$  表示当前锚框的质量(用 IoU 值衡量),超参数  $\alpha$  和  $\beta$  分别控制注意力曲线的陡峭程度和中心位置。这种设计特别强化了对中等质量锚框 ( $0.3 < IoU < 0.7$ ) 的关注,而这部分样本在交通场景中往往对应着部分遮挡或形变较大的目标。完整的 PIoUv2 损失函数计算公式可表示为:

$$L_{PIoUv2} = (1 - IoU + \lambda \cdot Penalty) \cdot Attention(q). \quad (8)$$

## 3 实验与结果分析

### 3.1 数据集

选取广泛应用的 KITTI<sup>[22]</sup> 自动驾驶基准数据集进行模型性能验证。该数据集包含城市道路、高速公路及乡村环境等多种复杂交通场景下的 7 481 张标注图像。针对交通目标检测任务特点,对原始标注类别进行了合理整合:将 Van、Truck 和 Tram 统一归为 Car 类别, Person\_sitting 并入 Pedestrian 类别,最终保留 Car、Pedestrian 和 Cyclist 三个典型交通目标类别。数据集按照 8 : 1 : 1 的比例划分为训练集(5 984 张)、验证集(748 张)和测试集(749 张)。

### 3.2 实验环境及参数配置

本研究构建了完整的深度学习实验平台,硬件方面采用 NVIDIA GeForce RTX 4060 显卡进行模型训练与推理加速,搭配 CUDA 12.6 并行计算架构以提升运算效率。软件环境基于 PyTorch2.5.0 深度学习框架搭建,开发语言使用 Python 3.10.16 版本。训练参数如表 1 所示。

表 1 实验参数

Tab. 1 Experiment parameters

参数名	参数值
图像尺寸/pixel	640×640
学习率	0.01
优化器	SGD
批量大小	16
迭代轮次	200
动量	0.937

3.3 评价指标

在检测精度方面,选取精确率(Precision)、召回率(Recall)和平均精度均值(mean Average Precision, mAP)作为核心指标,其中 mAP 包含 mAP@0.5 和 mAP@0.5:0.95 两个评价标准,分别反映算法在宽松和严格 IoU 阈值下的检测性能,mAP@0.5 采用 0.5 的 IoU 阈值,而 mAP@0.5:0.95 则在 0.5~0.95 的 IoU 内以 0.05 为步长计算平均值。模型复杂度通过参数量(Parameters)和计算量(FLOPs)进行量化评估,二者共同决定了模型的计算效率和资源需求。各指标计算公式如下:

$$Precision = \frac{TP}{TP+FP}, \quad (9)$$

$$Recall = \frac{TP}{TP+FN}, \quad (10)$$

$$AP = \int_0^1 Precision(Recall) d(Recall) \quad (11)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (12)$$

式中:TP 表示正确检测的目标数,FP 表示误检数, FN 表示漏检数,AP<sub>i</sub> 表示第 i 个类别的平均精度,N 表示类别总数。

3.4 实验结果分析

3.4.1 损失函数对比

为了全面评估 PIoUv2 损失函数的有效性,本研究在相同实验环境下与 CIoU、EIoU、SIoU、ShapeIoU 和 PIoU 等主流边界框回归损失函数进行了对比实验。如表 2 所示,PIoUv2 在多个关键指标上展现出显著优势。在检测精度方面,PIoUv2 取得了 89.2% 的 mAP@0.5 值,相比基准 CIoU (86.4%) 提升 2.8 个百分点,同时以 62.4% 的 mAP@0.5:0.95 显著超越其他对比方法。特别值得注意的是,PIoUv2 在召回率指标上达到 80.8%,较 CIoU (76.9%) 提升 3.9 个百分点,这表明其在减少漏检方面具有突出优势。虽然 SIoU 在精确率 (90.4%) 上略优于 PIoUv2 (90.3%),但其召回率和 mAP 指标均明显落后,显示出 PIoUv2 在精度-召回平衡方面的优越性。

表 2 不同损失函数结果对比

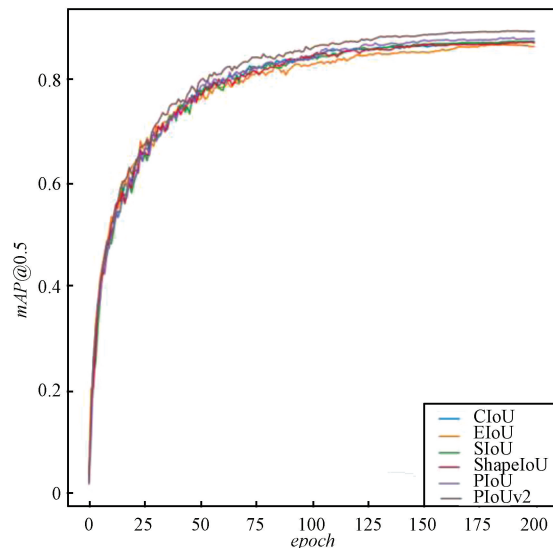
Tab. 2 Comparison of different loss function results

Metrics	Precision/%	Recall/%	mAP@0.5	mAP@0.5:0.95
CIoU	89.9	76.9	86.4	59.8

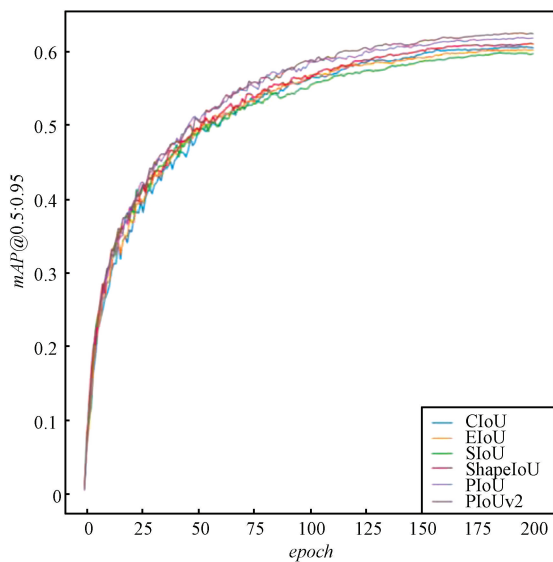
续表

Metrics	Precision/%	Recall/%	mAP@0.5	mAP@0.5:0.95
EIoU	87.7	76.2	85.5	59.7
SIoU	<b>90.4</b>	77.6	86.2	59.6
ShapeIoU	88.3	77.8	85.7	60.1
PIoU	90.1	77.3	87.3	61.3
PIoUv2	90.3	<b>80.8</b>	<b>89.2</b>	<b>62.4</b>

mAP@0.5 和 mAP@0.5:0.95 的训练过程曲线如图 7 所示,可以看出 PIoUv2 损失函数在整个训练过程中展现出明显的优势:就 mAP@0.5 指标而言,其训练曲线始终维持在最高位较其他损失函数保持领先优势;而在更严格的 mAP@0.5:0.95 指标上,PIoUv2 不仅收敛速度更快,且在训练后期 (epoch>100) 的稳定表现显著优于对比方法。特别值得注意的是,虽然 SIoU 等传统损失函数在训练初期 (epoch<50) 表现出相近的上升趋势,但随着训练深入,PIoUv2 的曲线逐渐与其他方法拉开差距,这种后期分化现象印证了其动态尺度敏感因子和几何约束项在模型优化后期的持续有效性。曲线波动分析显示,PIoUv2 的振幅较小,表明其对复杂场景中的尺度变化具有更强的适应性。这些训练动态特征从优化过程角度佐证了 PIoUv2 在自动驾驶目标检测任务中的优越性。



(a) mAP@0.5



(b) mAP@0.5:0.95

图7 训练过程曲线

Fig. 7 Training progress curves

### 3.4.2 消融实验

为系统评估各改进模块的有效性,本研究以 YOLOv8n 为基线模型,通过渐进式增加改进模块的方式进行消融实验,实验结果如表 3 所示。结果表明,当

仅将原始颈部结构替换为 MAFPN 时,模型在保持参数数量不变(3.0 M)的情况下,mAP@0.5 从 86.4% 提升至 87.7%,mAP@0.5:0.95 从 59.8% 提升至 62.3%,FLOPs 从 8.1 G 增加到 8.7 G,验证了 MAFPN 在可接受的计算成本增加下对多尺度特征融合能力的提升效果。在此基础上引入 PIoUv2 损失函数后,模型性能得到进一步改善,mAP@0.5 提升至 89.1%,mAP@0.5:0.95 达到 63.2%,同时召回率从 76.9% 显著提升至 82.3%,表明 PIoUv2 有效优化了目标定位精度。最终,当进一步将 C2f 模块替换为 C2f-DEConv 后,模型精度提升至 91.6%,表明该模块显著改善了模型对小尺度目标的细节捕捉能力。虽然 mAP@0.5 较前一阶段略有下降至 89.0%,但整体性能更为均衡,且 FLOPs 仅轻微增至 8.8 G,参数数量基本保持不变。综合分析表明,MAFPN 主要增强了多尺度特征融合能力,PIoUv2 显著提升了目标定位精度,而 C2f-DEConv 则优化了模型的细节特征表达能力,三者协同作用使模型实现了综合性能的显著提升,为实际应用提供了有效的精度-效率平衡方案。

表 3 消融实验结果

Tab. 3 Results of ablation experiment

Baseline	MAFPN	PIoUv2	C2f-DEConv	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%	FLOPs/G	Parameters/M
YOLOv8n				89.9	76.9	86.4	59.8	8.1	3.006
	✓			90.7	78.3	87.7	62.3	8.7	2.986
	✓	✓		88.1	82.3	89.1	63.2	8.7	2.986
	✓	✓	✓	91.6	79.9	89.0	62.7	8.8	2.987

### 3.4.3 主流算法对比

为全面评估本研究所提 DMP-YOLO 模型的性能优势,选取了当前主流的轻量级目标检测模型进行对比实验,包括 YOLOv3-tiny、YOLOv5n、YOLOv8n 及其多个改进版本(YOLOv8-EfficientViT、Hyper-YOLO 等),以及最新发布的 YOLOv10、YOLO11 等基准模型。如表 4 所示,MLG-YOLO 在多项关键指标上展现出显著优势:在检测精度方面,mAP@0.5 达到 89.0%,分别较 YOLOv8n、YOLOv5n 提升 2.6、3.2 个百分点;在模型效率方面,仅需 8.8 G FLOPs 和 3.0 M 参数,计算复杂度显著低于 YOLOv3-tiny (18.9 G

FLOPs) 和 YOLOv8-EfficientViT (9.4 G FLOPs)。特别值得注意的是,MLG-YOLO 以 91.6% 的精度和 79.9% 的召回率实现了当前最优的精度-召回平衡,这主要得益于 MAFPN 结构与 DEConv 模块的协同作用。与同参数规模(3.0 M)的 YOLOv8n 相比,MLG-YOLO 在保持相同参数数量的情况下,mAP@0.5:0.95 提升 2.9 个百分点,验证了所提方法在特征融合和细节提取方面的创新价值。实验结果充分表明,MLG-YOLO 在精度与效率的平衡上达到了当前轻量级检测模型的先进水平。

表 4 不同算法性能对比结果

Tab. 4 Performance comparison of different algorithms

Models	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%	FLOPs/G	Parameters/M
YOLOv3-tiny	87.8	68.5	77.9	52.3	18.9	12.1
YOLOv5n	87.7	76.2	85.8	58.2	7.1	2.5
YOLOv8n	89.9	76.9	86.4	59.8	8.1	3.0
YOLOv8-EfficientViT	89.1	76.8	86.2	58.7	9.4	4.0
Hyper-YOLO	87.6	79.9	86.9	61.7	10.8	3.9
YOLOv9t	88.2	76.4	85.6	58.7	7.6	2.0
YOLOv10n	84.1	74.2	84.0	58.1	6.5	2.2
YOLO11n	88.2	75.8	85.4	58.8	6.3	2.5
YOLO12n	86.3	78.0	85.7	58.2	6.3	2.5
MLG-YOLO	91.6	79.9	89.0	62.7	8.8	3.0

3.4.4 可视化分析

为直观展示改进模型的检测性能,图 8 对比了 DMP-YOLO 与基线 YOLOv8n 在不同场景下的检测效果。第一行可视化图展示了强光背景下的检测结果,得益于 PLoUv2 损失函数的几何约束优化,DMP-YOLO(图 8(c))较 YOLOv8n(图 8(b))表现出更精确的目标定位能力,有效抑制了过曝区域存在的检测精度过低问题。第二行可视化图显示,在复杂背景干扰下,YOLOv8n 出现明显的误

检(如将背景物体识别为目标),而 DMP-YOLO 通过 C2f-DEConv 模块增强的特征选择机制避免了此类错误。第三行可视化图进一步验证了改进模型对小目标和遮挡场景的鲁棒性:YOLOv8n 存在漏检(如未识别部分遮挡车辆),而 DMP-YOLO 凭借细节增强卷积和 MAFPN 模块的多尺度特征增强能力,实现了更完整的目标覆盖。这些可视化结果与定量分析一致,证实了 DMP-YOLO 在复杂环境下的检测优势。

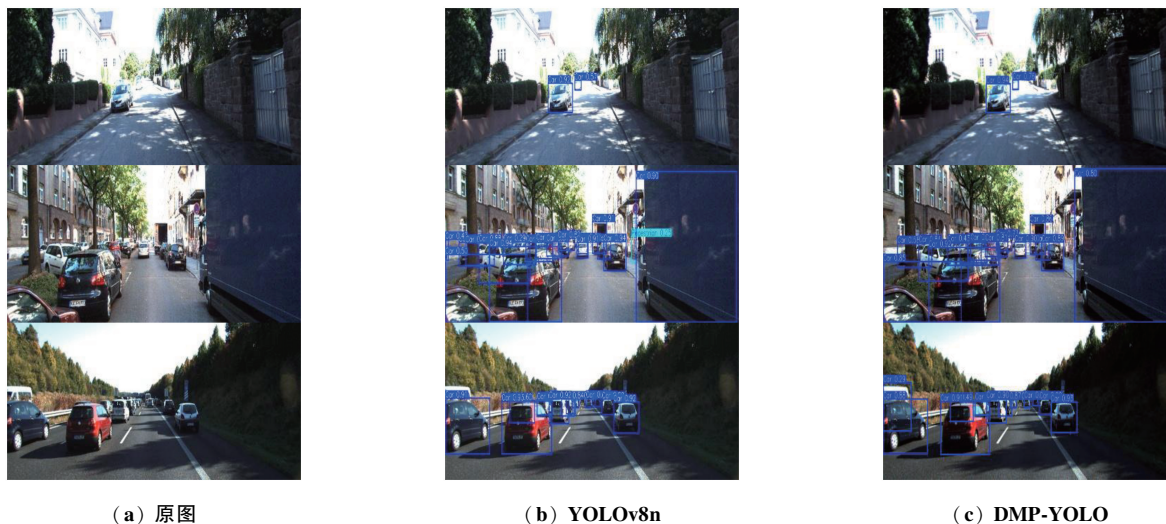


图 8 模型改进前后检测效果对比

Fig. 8 Comparison of detection results before and after model improvement

为了更加充分、直观地分析改进模型 DMP-YOLO 的有效性,本研究通过 Grad-CAM 技术对模型特征关注区域进行可视化分析如图 9 所示。从图 9(c)可以看出,相比基线模型 YOLOv8n,DMP-

YOLO 的热力图呈现出更精准的特征聚焦特性;在远距离小目标检测场景中,改进模型的热力分布能够完整覆盖目标区域,而基线模型则表现出明显的注意力分散现象;面对复杂背景干扰时,DMP-YOLO

的热力集中分布于车辆边缘等判别性特征区域,有效抑制了背景误激活;特别在遮挡场景下,改进模型的热力峰值稳定定位在目标几何中心,展现出更强的鲁棒性。这些可视化证据与定量实验结果一致,

共同证实了 MAFPN 多尺度特征融合、C2f-DEConv 细节增强以及 PIoUv2 损失函数在提升模型特征选择能力方面的协同作用,降低了对外部环境干扰的敏感性,同时增强了对目标本质特征的捕捉能力。

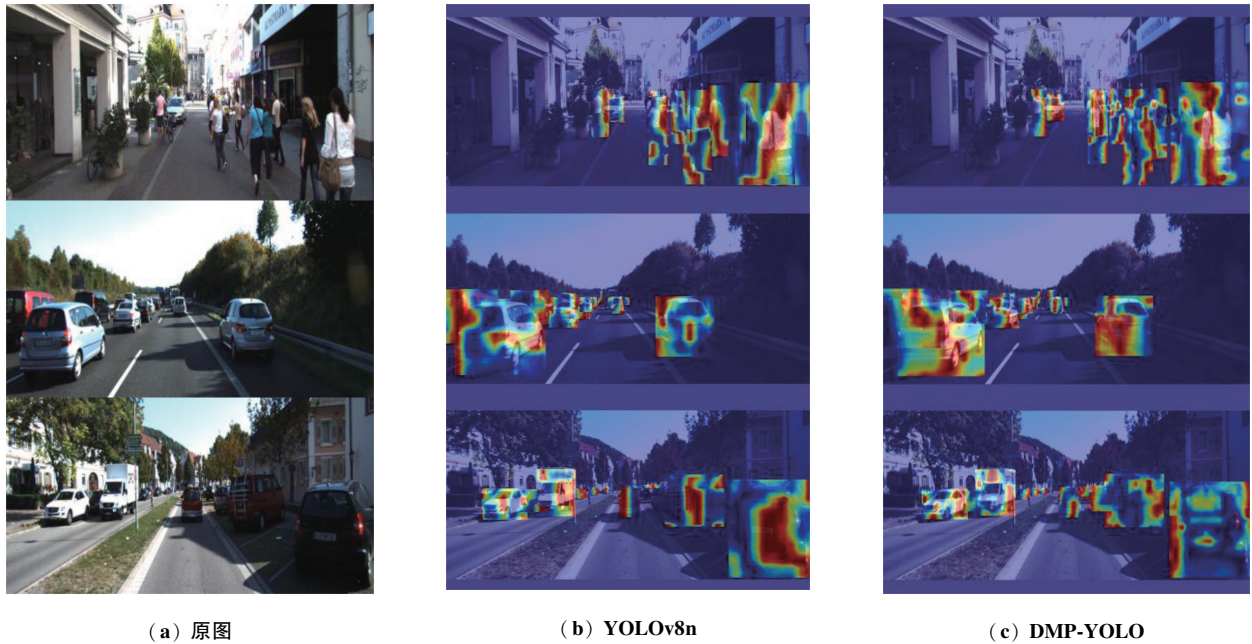


图9 模型改进前后热力图对比

Fig.9 Comparison of heatmaps before and after model improvement

#### 4 结束语

针对复杂交通场景中多尺度目标检测和复杂环境适应性等关键问题,本文提出了一种面向自动驾驶的目标检测算法 DMP-YOLO。以 YOLOv8n 为基础,MAFPN 通过浅层辅助融合和高级辅助融合双路径设计改进颈部网络,实现了跨尺度特征的高效交互与融合;C2f-DEConv 采用独特的卷积结构增强了对小目标和细节特征的提取能力;PIoUv2 则通过动态尺度敏感因子和几何约束优化,显著提升了遮挡场景下的定位精度。实验结果表明,DMP-YOLO 在保持参数量和计算量基本不变的情况下,Precision、Recall、mAP@50、mAP@50:95 分别提升了 1.7、3、2.6、2.9 个百分点,验证了本文改进算法的有效性。此外,还与其他先进算法进行了比较,检测性能明显优于其他算法,展现了其在自动驾驶实际应用中的潜力。虽然,DMP-YOLO 在检测精度上有了很大的提升,但在轻量化方面还需要对模型进一步优化,未来研究将着重探索模型在面向实际车载系统的轻量

化部署方案,以进一步提升自动驾驶环境感知系统的实用性和可靠性。

✦

#### 参考文献

- [1] ZHAO Z Q, ZHENG P, XU S T, et al. Object Detection with Deep Learning: A Review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212-3232.
- [2] BOUKERCHE A, HOU Z J. Object Detection Using Deep Learning Methods in Traffic Scenarios[J]. ACM Computing Surveys (CSUR), 2021, 54(2): 1-35.
- [3] 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, 2016, 39(6): 1137-1149.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-time Object Detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779-788.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector [C] // Computer Vision-ECCV 2016: 14th European Conference. Amsterdam: Springer, 2016: 21-37.

- [6] 刘宇,高尚兵,张秦涛,等. 面向复杂交通场景的目标检测模型 YOLO-T[J]. 山东科学, 2024, 37(6): 104-115.
- [7] JIAO B W, WANG Y L, WANG P, et al. RS-YOLO: An Efficient Object Detection Algorithm for Road Scenes [J]. Digital Signal Processing, 2025, 157: 104889.
- [8] 杨磊,陈艳菲,李海鸣,等. 基于改进 YOLOv8 的自动驾驶场景目标检测算法[J]. 计算机工程与应用, 2025, 61(1): 131-141.
- [9] 杨博,胡珍珍. 基于 YOLOv8n 改进算法的自动驾驶目标检测 [J/OL]. 控制工程, 1-7 [2025-07-10]. <https://doi.org/10.14107/j.cnki.kzgc.20240494>.
- [10] 李康,宋文广. 改进 YOLOv5 的道路车辆目标检测方法 [J]. 无线电工程, 2025, 55(3): 493-499.
- [11] FAN Q, LI Y T, DEVECI M, et al. LUD-YOLO: A Novel Lightweight Object Detection Network for Unmanned Aerial Vehicle [J]. Information Sciences, 2025, 686: 121366.
- [12] 马荣贵,张翼,董世浩. 基于无人机影像的改进 YOLOv5 道路目标检测 [J]. 无线电工程, 2025, 55(1): 1-10.
- [13] WANG S Y, QU Z, LI C J, et al. BANet: Small and Multi-object Detection with a Bidirectional Attention Network for Traffic Scenes [J]. Engineering Applications of Artificial Intelligence, 2023, 117: 105504.
- [14] KHALILI B, SMYTH A W. SOD-YOLOv8—Enhancing YOLOv8 for Small Object Detection in Traffic Scenes [J]. Sensors, 2024, 24(19): 6209.
- [15] CHEN Y, LUO H L. VisioSignNet: A Dual-interactive Neural Network for Enhanced Traffic Sign Detection [J]. Expert Systems with Applications, 2024, 255: 124688.
- [16] MAHAUR B, MISHRA K K, KUMAR A. An Improved Lightweight Small Object Detection Framework Applied to Real-time Autonomous Driving [J]. Expert Systems with Applications, 2023, 234: 121036.
- [17] 古佳欣,陈高华,张春美. YOLOv8-DEL: 基于改进 YOLOv8n 的实时车辆检测算法研究 [J]. 计算机工程与应用, 2025, 61(1): 142-152.
- [18] BIE M L, LIU Y Y, LI G N, et al. Real-time Vehicle Detection Algorithm Based on a Lightweight You-Only-Look-Once (YOLOv5n-L) Approach [J]. Expert Systems with Applications, 2023, 213: 119108.
- [19] CAI F, QU Z, XIA S F, et al. A Method of Object Detection with Attention Mechanism and C2f\_DCNv2 for Complex Traffic Scenes [J]. Expert Systems with Applications, 2025, 267: 126141.
- [20] YANG Z Q, GUAN Q, ZHAO K, et al. Multi-branch Auxiliary Fusion YOLO with Re-parameterization Heterogeneous Convolutional for Accurate Object Detection [C] // Chinese Conference on Pattern Recognition and Computer Vision (PRCV). Urumqi: Springer, 2024: 492-505.
- [21] CHEN Z X, HE Z W, LU Z M. DEA-Net: Single Image Dehazing Based on Detail-enhanced Convolution and Content-guided Attention [J]. IEEE Transactions on Image Processing, 2024, 33: 1002-1015.
- [22] GEIGER A, LENZ P, URTASUN R. Are We Ready for Autonomous Driving the KITTI Vision Benchmark Suite [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 3354-3361.

作者简介

李明芳 女, (2000—), 硕士研究生。主要研究方向: 机器学习与人工智能、目标检测。