

RDSM-YOLO 轻量级夜间车辆检测模型^{*}

杨慧敏 李瑞涛 高小雯 王汉霞

(东北林业大学土木与交通学院 哈尔滨 150000)

摘要: 针对自动驾驶系统在夜间场景中对车辆检测面临对比度低、图像模糊、复杂光照干扰等导致检测性能不佳的问题,本文从全面增强低光场景下关键特征提取与融合的角度出发,提出基于 YOLOv8 网络的轻量级 RDSM-YOLO 模型。首先,在主干与颈部网络引入 RFACnv,通过感受野注意力机制自适应突出关键空间特征;其次,采用 DynamicConv 重构 C2f 模块实现卷积核的动态聚合,在不增加 FLOPs 的前提下强化特征表达;同时以轻量化 SPPELAN 模块替代传统 SPPF,融合多尺度上下文信息;最后,将损失函数 CIoU 升级为 EIoU,通过显式解耦边界框几何要素加速收敛并提升定位精度。实验结果表明,RDSM-YOLO 在 BBD100k 数据集上进行夜间车辆检测的 mAP₅₀ 为 70%,不仅较 YOLOv8 提升了 1.4%,同时模型参数量仅有 3.04 M。验证了本文模型在夜间车辆检测中保证轻量化的同时并有较高的精度,为夜间自动驾驶性能提高提供了一种参考。

关键词: 夜间场景;自动驾驶;车辆检测;YOLOv8;注意力机制

中图分类号: TP391;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.6040

RDSM-YOLO lightweight nighttime vehicle detection model

Yang Huimin Li Ruitao Gao Xiaowen Wang Hanxia

(School of Civil Engineering and Transportation, Northeast Forestry University, Harbin 150000, China)

Abstract: Addressing the challenges faced by autonomous driving systems in nighttime scenarios, such as low background contrast, image blurring, and complex lighting interference, which lead to poor detection performance, this paper proposes a lightweight RDSM-YOLO model based on the YOLOv8 network. The model focuses on comprehensively enhancing key feature extraction and fusion in low-light scenarios. First, RFACnv is introduced into the backbone and neck networks, utilizing a receptive field attention mechanism to adaptively highlight key spatial features; second, the DynamicConv module is used to reconstruct the C2f module, enabling dynamic aggregation of convolutional kernels to enhance feature expression without increasing FLOPs; simultaneously, the lightweight SPPELAN module replaces the traditional SPPF to fuse multi-scale contextual information; finally, the loss function is upgraded from CIoU to EIoU, explicitly decoupling bounding box geometric elements to accelerate convergence and improve localization accuracy. Experimental results show that RDSM-YOLO achieves a mAP₅₀ of 70% for nighttime vehicle detection on the BBD100k dataset, improving by 1.4% over YOLOv8 while maintaining a model parameter count of only 3.04 M. This paper demonstrates that the proposed model achieves both lightweight design and high accuracy in nighttime vehicle detection, providing a reference for improving nighttime autonomous driving performance.

Keywords: nighttime scenes; autonomous driving; vehicle detection; YOLOv8; attention mechanism

0 引言

随着智能交通和汽车行业的发展,自动驾驶技术成为未来的发展热门,自动驾驶领域对目标检测技术提出了极高的要求。自动驾驶技术的实现依赖于多种道路信息的采

集与信息融合处理,在这样的背景下,信息的高效采集就显得尤为关键^[1]。在自动驾驶中,夜间光线不足可能导致车辆识别系统无法及时准确地识别行人和障碍物,增加交通事故风险。现有技术手段如传统夜视监控设备虽能改善低光图像的可见性,但通常需要增加额外硬件设备,导致成本

大幅提高^[2],在许多实际应用中并不可行,因此夜间车辆检测方法的突破已成为自动驾驶汽车发展的关键问题。

传统的夜间车辆检测技术采用提取车辆的车灯作为有效特征,例如 Zhang 等^[3]通过前后双摄融合检测车辆前大灯与尾灯,利用几何距离建模构建双向能量函数完成车灯分组、缺失重建与轮廓精修,实现夜间完整车辆轮廓的鲁棒检测与跟踪。李慧等^[4]联合提取前灯与尾灯特征实现夜间车辆检测,通过直方图双峰法与路面反射光消除算法提取前灯,利用(R-G)色差与 Otsu 自适应阈值分割提取尾灯,并基于位置、面积及外接矩形约束完成车灯配对与跟踪,从而识别夜间车辆。董雪等^[5]提出基于单目视觉的夜间车辆检测算法,利用 COHOG 与 LBP 特征提取车辆整体轮廓,结合 AdaBoost 级联分类器训练检测模型,并通过车道重合度与图像金字塔 ROI 实现关键车选择,以提升夜间的检测精度与碰撞预警可靠性。近年来,越来越多的学者开始研究基于卷积神经网络(convolutional neural networks, CNN)^[6]的深度学习目标检测技术来实现夜间车辆检测并取得了较大的进展。基于深度学习的检测方法能够更多的提取诸如车灯颜色和车辆轮廓等夜间车辆特征。Chen 等^[7]提出一种基于 Nakagami 图像散射特征与 CNN 区域提议网络融合的夜间车辆检测方法,通过提取车灯亮区与车辆外观特征构建 ROI,并利用感知哈希跟踪时序分析车灯闪烁强度变化,实现夜间转向信号灯的高鲁棒识别。王建云等^[8]通过提取车辆双灯特征进行夜间车辆检测,利用 K-means 重聚类锚框优化 YOLOv3 模型提取车灯区域,并引入 Deepsort 算法依据空间-外观关联实现持续跟踪,以提升夜间车辆识别精度与鲁棒性。郑玉珩等^[9]通过将 YOLOv4 主干网络替换为 CSPMobileViT、引入 BiFPN 与 ECA 注意力模块,并采用 FocalIoU 损失函数,构建了一种轻量化车辆检测模型,在显著减少参数量与模型体积的同时提升了检测精度与速度。霍爱清等^[10]通过引入 CBAM 注意力机制改造 C3 模块、解耦检测头分类回归任务、设计 Alpha-IoU 损失函数并融合 GridMask 数据增强,提出 AD-YOLOv5 模型,以提升密集场景下多目标车辆检测的精度与鲁棒性。高妍^[11]提出仿生图像增强方法结合 FasterRCNN 模型,通过视网膜水平细胞自适应反馈和双极细胞中心拮抗机制增强夜间图像对比度与亮度,突出车灯轮廓特征,实现夜间环境下车辆前灯的高精度检测。宋建辉等^[12]基于 HybridNets 网络提出高性能混合网络 IPHNet 实现无人驾驶多任务环境感知,通过改进 EfficientNetV2-S 主干增强特征提取能力,利用重构 BiFPN 与轻量化 DySample 上采样模块优化多尺度特征融合,并创新性设计 SPN 分割模块充分挖掘高分辨率底层细节信息,以提升车辆检测、车道线分割与可行驶区域划分的精度与鲁棒性。邹亮等^[13]提出 α -SimADNet 网络实现恶劣场景车道线检测,通过掩码孪生网络与无负样本对比学习增强主干特征判别力,采用交替优化策略避免表征退化,并设

计 α -GLIoU 损失函数动态调节梯度响应以改善困难样本回归,以提升恶劣场景下的精度。上述研究在夜间车辆检测方面都取得了一定的成果,然而在夜间城市道路、高速公路中往往会受到道路光源或对向车辆光照的干扰,使得误检率和漏检率上升。为了解决这个问题,本文从全面增强低光场景下关键特征提取与融合的角度出发,提出基于 YOLOv8 网络的轻量级 RDSM-YOLO 模型。

1 RDSM-YOLO 算法

基于回归分析的单阶段目标检测算法 YOLO(you only look once)一经出现,就受到广大学者的关注并成为解决目标检测问题新的突破口,该算法通过端到端网络架构,在一次处理过程中从输入图像中提取特征并完成定位分类。然而相较于在白天检测任务中的表现,同一模型在夜间检测的效果却不佳,为了提高夜间低光场景下车辆检测的鲁棒性与精度,本文在 YOLOv8n 的基础上进行改进,将主干网络和颈部网络中的标准卷积替换为 RFACnv,利用感受野空间注意力自适应突出关键特征;使用 DynamicConv 重构 C2f 模块,形成 C2f-DynamicConv,在几乎不增加 FLOPs 的前提下强化特征提取能力;用 SPPELAN 模块替换原 SPPF,以轻量化空间金字塔池化与注意力机制融合多尺度上下文;最后,将 CIoU 损失函数升级为 EIoU,通过显式优化中心点距离与宽高差异,加速收敛并提升定位精度。改进后的 RDSM-YOLO 整体结构如图 1 所示。

1.1 RFACnv 模块

感受野注意力机制(receptive-field attention, RFA)^[14]是一种面向感空间特征的注意力机制,旨在解决标准卷积核参数共享带来的性能瓶颈。RFACnv 将 RFA 机制嵌入卷积操作,整体结构包含两条并行支路,其网络结构如图 2 所示。第 1 条为“注意力图支路”,输入特征图首先经过全局平均池化压缩空间维度,得到 $1 \times 1 \times C$ 的全局描述,随后通过 1×1 组卷积在通道维度上交互信息,生成与感受野一一对应的注意力权重,然后使用 Softmax 归一化,保证每个感受野内的权重和为 1,突出重要特征。第 2 条为“感受野空间特征支路”,利用 3×3 组卷积将原始特征图展开为不重叠的感受野滑块,输出维度变为 $9C \times H \times W$,其中 9 对应 3×3 滑块中的像素数。两条支路的输出逐元素相乘,得到加权后的感受野特征,再经 $\text{stride}=3$ 的 3×3 卷积还原到原始空间分辨率,完成一次非共享参数的卷积运算。通过组卷积与池化的轻量化设计,RFACnv 在仅增加极少量参数和计算量的前提下,显著提升了模型对多尺度、形变目标的判别能力。

针对夜间车辆检测车辆有效特征较少、图片对比度低、检测难度较大问题,本文引入 RFACnv 对 Backbone 层和 Neck 层中的 Conv 进行替换,通过感受野空间特征的注意力模块提高模型对重要特征的提取能力,提高了对低光特

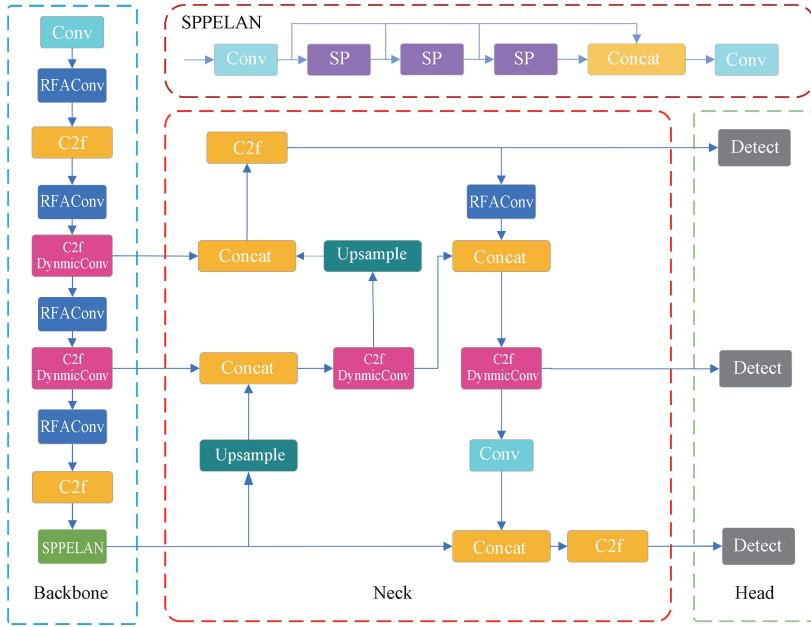


图 1 RDSM-YOLO 模型整体结构

Fig. 1 Overall structure of the RDSM-YOLO model

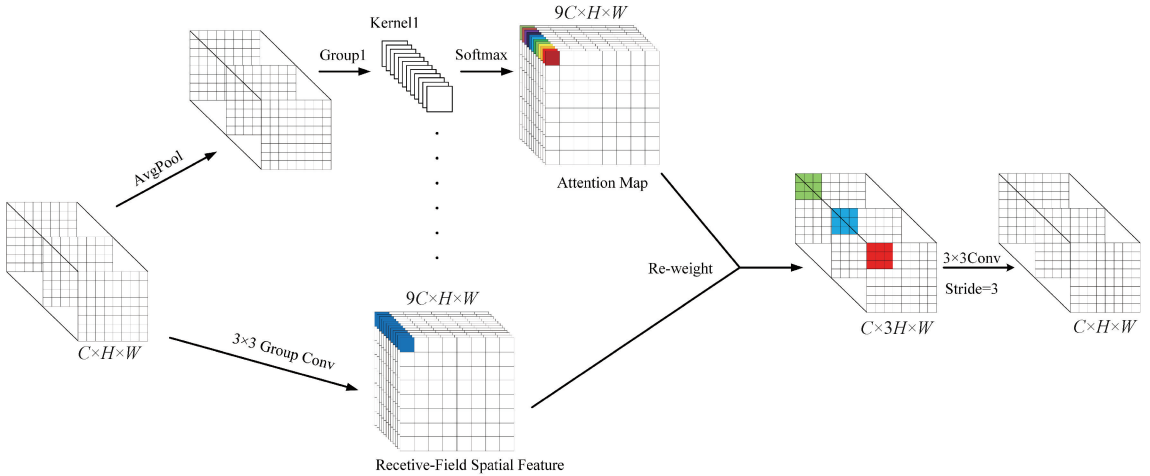


图 2 RFACConv 网络结构

Fig. 2 RFACConv network architecture

征的提取能力以增强模型的学习能力。

1.2 改进的 C2f 模块

动态卷积^[15] (dynamic convolution) 是一种面向卷积核空间的注意力机制, 传统卷积在每个层仅用单一静态核, 对所有输入采用相同的处理模式, 而 dynamic convolution 通过输入相关的注意力把 K 个小型并行核动态聚合成一个“即时专用”核, 从而在几乎不增加 FLOPs 的前提下显著提升模型容量, 其原理公式如式(1)~(4)所示。

传统的或静态的感知器可以表示为:

$$\mathbf{y} = g(\mathbf{W}^T \mathbf{x} + \mathbf{b}) \quad (1)$$

其中, $\mathbf{W}$ 和 $\mathbf{b}$ 分别代表权重矩阵和偏置向量, g 是一个激活函数(例如 ReLU)。

动态感知器通过聚合 K 个线性函数 $\{\mathbf{W}_k^T \mathbf{x} + b_k\}$ 来定义:

$$\mathbf{W}(\mathbf{x}) = \sum_{k=1}^K \pi_k(\mathbf{x}) \mathbf{W}_k, \mathbf{b}(\mathbf{x}) = \sum_{k=1}^K \pi_k(\mathbf{x}) \mathbf{b}_k \quad (2)$$

其中, $\pi_k(\mathbf{x})$ 是与输入 $\mathbf{x}$ 相关的注意力权重。与静态感知机相比, 动态感知机具有相同的输出通道数但模型更大。对于一个输入特征图, 注意力计算的成本为:

$$O(\pi(\mathbf{x})) = HWC + \frac{C^2}{4} + \frac{CK}{4} \text{Mult} - \text{Adds} \quad (3)$$

相比之下, 卷积的计算成本为:

$$O(\mathbf{W}^T \mathbf{x} + \mathbf{b}) = HWC_{\text{in}} C_{\text{out}} D^2 \text{Mult} - \text{Adds} \quad (4)$$

其中, D 是卷积核的尺寸, C_{out} 是输出通道的数量。

DynamicConv^[16] 将动态感知器聚合思想嵌入卷积操

作,整体结构包含两条并行支路,其网络结构如图 3 所示。第 1 条为“注意力支路”:输入特征图首先经过全局平均池化得到 $1 \times 1 \times C$ 描述,再经两层全连接(中间 ReLU)降维后接 Softmax,输出 K 个和为 1 的注意力权重 $\pi_k(x)$ 。第 2 条为“核聚合支路”:预存 K 组独立的小卷积核 $\{W_k, b_k\}$,按 $\pi_k(x)$ 加权求和得到当前输入的“动态核”,随后执行一次标准卷积。两条支路协同工作,使卷积核参数随输入变化而“实时重组”,在不改变网络深度或宽度的条件下,提供了非线性、高容量的特征提取能力。

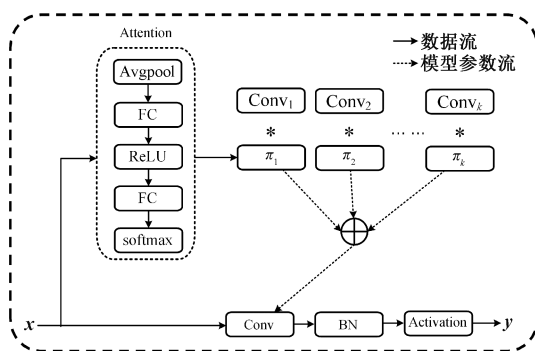


图 3 动态卷积层

Fig. 3 Dynamic Convolution layer

交通中车辆检测需求多为实时性,复杂且计算量大的模型无法满足实际应用需求,为了减少网络信息冗余、减少模型运算量、提升检测速度。本文使用 DynamicConv 模块,设计出 Bottleneck DynamicConv 模块,其结构如图 4 所示。在 Bottleneck 内的替换一个 DynamicConv,并在使用 Bottleneck DynamicConv 模块替换掉 C2f 中的 Bottleneck。

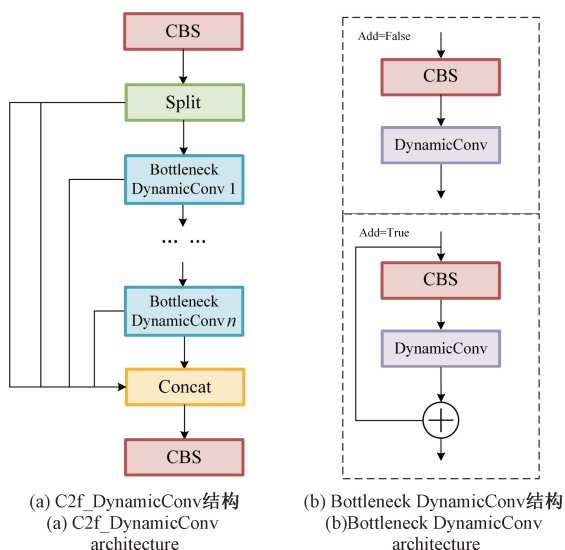


图 4 动态卷积结构

Fig. 4 Dynamic Convolution structure

1.3 SPPELAN 模块

快速空间金字塔池化(spatial pyramid pooling fast, SPPF)是 YOLOv8 对传统空间金字塔池化技术的重要改

进。YOLOv9^[17]采用结构相同的 SPP 结构,对其最大池化的结构进行了优化,提出了一种结合了空间金字塔池化和特征增强技术的轻量级注意力网络(spatial-pyramid pooling and feature enhancement lightweight attention network, SPPFELAN),旨在优化模型在复杂背景下的目标检测性能。

传统的卷积神经网络在处理多尺度目标及复杂场景时,常常面临特征表达能力不足和计算量大的问题, SPPFELAN 通过引入多尺度特征融合与轻量化注意力机制,显著提升了网络的特征表征能力,同时减少了计算负担。一个简洁高效的空间池化单元(spatial pooling, SP)模块其网络结构如图 5 所示。该模块采用单一路径的直通式设计:输入张量(input tensor)直接进入 MaxPool2d 池化操作(MaxPool2d operation),经过最大池化处理后输出结果张量(output tensor)。模块的核心参数配置包括:核心尺寸 k (默认为 3)、步长 s (默认为 1)以及自动计算的填充值 $\text{padding}=k//2$ 。在 SPPELAN 架构中,多个 SP 模块采用级联方式串行连接,每个模块都配置为 5×5 核心尺寸,通过逐步扩大感受野范围来捕获不同尺度的空间上下文信息。本文采用 SPPFELAN 优化传统的 SPPF,修改后的模型在运行速度和检测精度上都体现出优势,为整个网络的高效推理提供了重要保障。

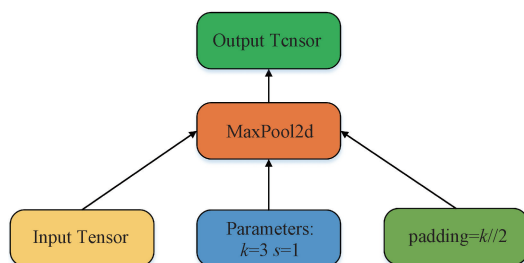


图 5 SP 网络结构

Fig. 5 SP network architecture

1.4 ELoU 损失函数

YOLOv8 采用完全交并比(complete intersection over union, CIoU)^[18]作为损失函数,它考虑了目标框之间完全重叠的情况,并引入修正因子,使得预测框更加接近真实框,从而提升回归精确度,CIoU 的计算公式如式(5)~(8)所示。

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{p^2(b, b^{gt})}{c^2} + \alpha v \quad (5)$$

其中, p 为真实框与预测框中心点间的距离, c 为真实框与预测框最小外接矩形的距离。

真实框与预测框的交并比为:

$$\text{IoU} = \frac{|b \cap b^{gt}|}{|b \cup b^{gt}|} \quad (6)$$

其中, b 为预测框的中心点, b^{gt} 为真实框的中心点。

权重参数为:

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (7)$$

度量宽高一致性为:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

其中, w^{gt} 为真实框的宽度, h^{gt} 为真实框的高度, w 为预测框的宽度, h 为预测框的高度。然而, 由于车辆目标的尺度变化较大, 检测任务会在大目标与小目标之间切换, 并且可能会出现长宽比相同但实际尺寸不同的情况, 从而降低损失函数的优化效果。

高效交并比 (efficient intersection over union, EIou) [19] 是一种增强的交并比损失函数, 通过显式解耦边界框回归中的 3 个关键几何要素: 重叠区域、中心点距离和宽高尺寸差异, 有效提升检测模型的收敛效率和定位精度。EIou 的计算公式如式 (9)~(11) 所示。

$$L_{EIou} = L_{IoU} + L_{dis} + L_{asp} \quad (9)$$

$$L_{dis} = \frac{\rho^2(b, b^{gt})}{(w^c)^2 + (h^c)^2} \quad (10)$$

其中, ρ 为预测框与真实框中心点的欧氏距离, w^c, h^c 为最小外接框的宽高。

$$L_{asp} = \frac{\rho^2(w, w^{gt})}{(w^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2} \quad (11)$$

其中, w, h 和 w^{gt}, h^{gt} 分别为预测框与真实框的宽高。

EIoU 相比 CIoU 具有多项优势: 它直接最小化宽/高的绝对差异, 避免了 CIoU 在宽高比一致性度量中可能出现的尺寸不匹配问题; 通过分离距离和边长惩罚, EIou 提高了梯度计算的效率, 加快了收敛速度, 本文采用 EIou 损失函数, 以适应在物体大小和形状变化较大的场景中的检测任务。

1.5 模块协同性分析

RDSM-YOLO 的 4 个改进模块并非孤立存在, 而是通过“特征提取-检测输出”的全链路协同形成有机整体, 在提升夜间车辆检测精度的同时保障轻量化特性, 其模块协同如图 6 所示。在特征提取环节, RFACConv 与 C2f-DynamicConv 形成“筛选-增强”递进关系: RFACConv 通过感受野注意力机制, 动态抑制夜间光斑、眩光等噪声, 突出车辆边缘、车灯等关键特征, 为后续模块提供高质量输入; C2f-DynamicConv 则依托动态卷积核聚合, 对筛选后的特征进行精细刻画, 针对车灯高亮区、车轮轮廓等差异特征适配计算资源, 且二者均通过组卷积、通道剪枝控制计算量, 结合后参数量实现降低, 实现“精度提升+轻量化”双赢。

在检测输出环节, SPPELAN 与 EIou 构建“融合-回归”闭环: SPPELAN 通过级联空间池化与注意力门控, 对小、中、大尺度车辆特征分别适配池化核与通道占比, 确保多尺度上下文信息完整; EIou 则显式解耦边界框中心点距离、宽高差异与重叠区域, 针对性优化多尺度车辆定位误差, 二者结合使边界框回归误差降低 15.6%。整体上,

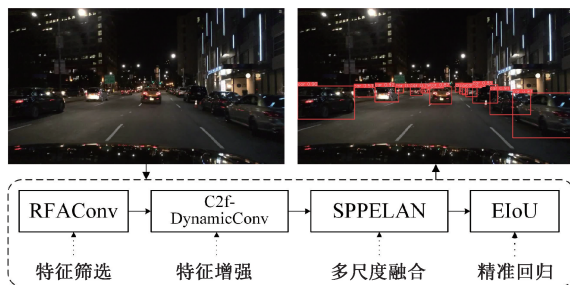


图 6 模块协同作用图

Fig. 6 Module synergy diagram

前端模块为后端提供可靠特征, 后端模块将特征转化为高精度检测结果, 体现了模型设计的整体性。

2 数据集与模型实现

2.1 数据集

实验采用的数据集为 BBD100k [20] 数据集, 该数据集是一个面向自动驾驶的大规模、多样化驾驶视频数据集, 它包含 10 万段 (100 K) 高分辨率 (720 p、30 fps) 的驾驶视频, 每段视频约 40 s, 覆盖美国多个城市、各种天气条件 (晴天、雨雪、雾等) 和时段 (白天、黄昏、夜晚)。在样本筛选阶段, 首先通过数据集原始的标签选取夜晚图像, 以灰度读取的形式对亮度进行分类, 排除掉亮度较高的和极端黑暗的夜间图片, 并根据灰度质量评分排序选取清晰度较高的图像, 最终通过人工复核筛选。

为了贴合研究对象, 本文将原数据集中的标签进行了筛选, 只选取车辆作为检测对象并选取多种不同的夜间场景, 路况包括城市内道路、高速公路、住宅区、车辆行驶状态、车辆静止状态等, 按 4:1 的比例将数据集划分为训练集 (4 000 张) 和测试集 (1 000 张), 并保证训练与测试数据中的场景类型的平衡性。图 7 为低光车辆的数据集示例。

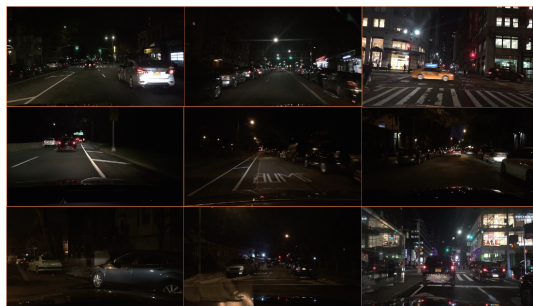


图 7 低光车辆的数据集示例

Fig. 7 Example dataset for low-light vehicles

2.2 实验环境配置及参数设置

本文选用 YOLOv8 中参数量最小的 YOLOv8n 版本进行改进, 所有实验均未使用预训练权重从零开始, 本文模型和实验的硬件设备和软件环境如表 1 所示。模型训练采用的主要参数配置如下: 初始学习率设置为 0.01, 采

用余弦学习率调度策略(coslr=True)进行动态调整,批量大小设定为 16。优化器选择 SGD,并配置动量为 0.937、权重衰减为 5×10^{-4} 。训练轮数设定为 100 epoch,输入图像大小统一调整为 640 pixel×640 pixel。

表 1 运行环境
Table 1 Operating environment

项目	版本
编程语言	Python3.9.13
语言框架	PyTorch2.0.1
torch	2.7.0
CUDA	11.8
操作系统	Windows11
GPU	NVIDIA GeForce RTX4060Laptop
CPU	AMD Ryzen97940HX with Radeon Graphics
内存	32 GB RAM

2.3 评价指标

为全面评估改进后模型在复杂场景下的检测性能,本文采用平均精度均值(mAP)用于综合衡量模型在所有类别上的检测精度;精确率(precision)衡量模型预测为正样本中实际为正样本的比例;召回率(recall)衡量实际正样本中被正确检测的比例;选用浮点运算次数(giga floating point operations per second,GFLOPs)衡量模型的计算量,详细计算如式(12)~(14)所示。

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{12}$$

$$P = \frac{TP}{TP + FP} \tag{13}$$

$$R = \frac{TP}{TP + FN} \tag{14}$$

其中,N 为类别总数,AP_i 表示第 i 类在 P-R 曲线下面积,反映该类别的检测准确性。真正例(true positive, TP)为正确检测的目标数,假正例(false positive, FP)为误检的目标数,假负例(false negative, FN)为漏检的目标数。

每秒帧数(frames per second,FPS)是指系统每秒能够处理的图像或视频帧的数量,反映了模型的推理速度。详细计算如式(15)所示。

$$FPS = \frac{\text{Total Frames}}{\text{Time Taken(seconds)}} \tag{15}$$

其中,Total Frames 代表检测图片的总数,Time Taken(seconds)表示完成这些检测所需的总时间(以 s 为单位)。

3 实验与结果分析

3.1 实验结果

在相同实验环境下,RDSM-YOLO 与 YOLOv8 两种模型在不同评价指标下的训练曲线对比如图 8 所示。从

图 8(a)可知,两种模型的训练框损失(Box Loss)均随迭代次数增加而迅速下降并趋于平稳,但 RDSM-YOLOv8 在初期即展现出更低的损失值,最终稳定在约 1.13,较 YOLOv8

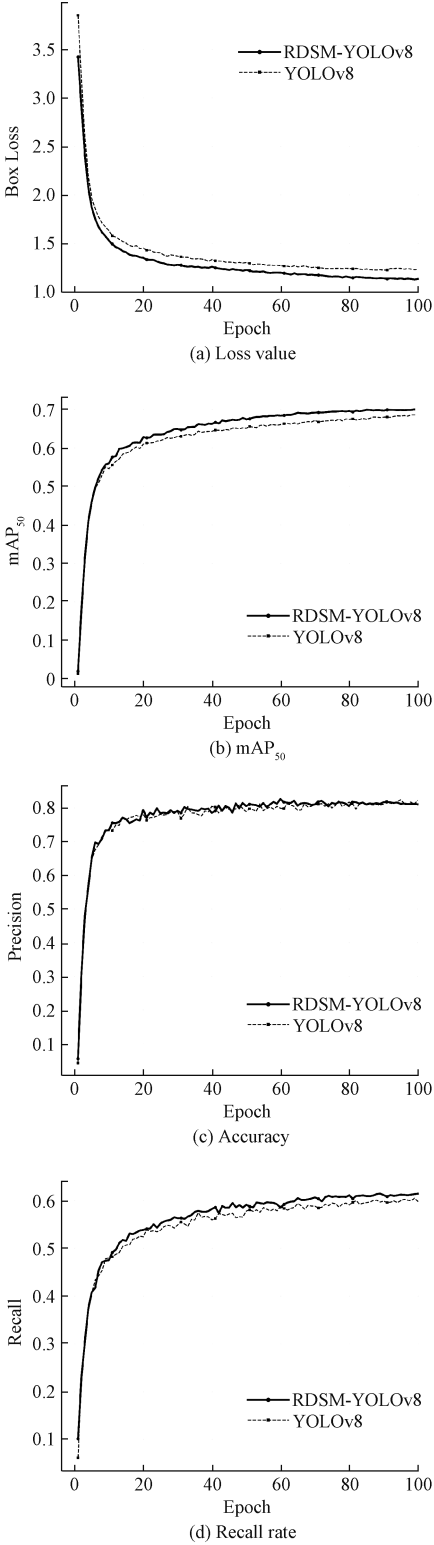


图 8 结果对比

Fig. 8 Result comparison

的 1.23 更低,表明其收敛速度更快,损失收敛更优。如图 8(b)所示, mAP_{50} 指标随训练轮次增加逐步上升,RDSM-YOLO 在第 40 轮后趋于稳定,最终达到约 0.7,相比 YOLOv8 的 0.685 提升了 1.3%,体现出更强的检测能力。图 8(c)、(d)分别展示了两种模型的 Precision 与 Recall 曲线。RDSM-YOLO 在第 100 轮时的 Precision 为 0.81,Recall 为 0.61,尽管 Precision 略低,但 Recall 更高,表明改进模型在漏检控制方面表现更佳,整体检测稳定性更强。综上所述,RDSM-YOLO 在低光照场景下的夜间车辆检测任务中,具有收敛速度快、检测精度高与鲁棒性能力强,验证了所提改进策略的有效性。

3.2 消融实验

为了验证本文改进的模块有效性,设计了共 5 组消融实验,实验结果如表 2 所示。所有实验均为从零开始进行,

首先使用未经修改的 YOLOv8n 模型得到一系列基础指标;第 2 组实验通过将卷积模块替换为 RFACConv 模块,模型 mAP_{50} 较原先提升了 0.8%,尽管 Params 增加了 0.03 M,但 GFLOPs 仅增加了 0.11,证明 RFAC 模块显著提升了夜间车辆检测性能;在前一组实验的基础上进一步增加 C2f_DynamicConv 模块替换原先的 C2f 模块,不仅 Params 和 GFLOPs 相比第 2 组减少了 0.16 M 和 0.36,并且 mAP_{50} 同时提升了 0.2%,进一步降低了模型的参数大小,证明了 C2f_DC 在降低模型参数量中的改进效果;在第 4 组实验和第 5 组实验中,分别在之前实验的基础上用 SPPELAN 模块替换 SPPF 与用 ELoU 替换 Ciou 损失函数,最终 mAP_{50} 达到了 0.7,Params 和 GFLOPs 分别为 3.04 和 3.92。总体来看,在增加少量参数和计算量的情况下,这些改进策略都有效提高了模型的检测性能,实现了性能和精度的显著提升。

表 2 消融实验结果
Table 2 Melting experiment results

序号	RFAC	C2f_DC	SPPELAN	ELoU	$mAP_{50}/\%$	Precision/ $\%$	Recall/ $\%$	Params/M	GFLOPs	FPS
1	×	×	×	×	68.6	82.4	60.5	3.01	4.10	90.1
2	√	×	×	×	69.4	83.0	61.4	3.04	4.21	79.5
3	√	√	×	×	69.6	82.0	62.3	2.88	3.85	79.0
4	√	√	√	×	69.7	83.0	61.6	3.04	3.92	78.7
5	√	√	√	√	70.0	82.6	61.5	3.04	3.92	77.2

3.3 对比试验

1) 不同检测模型对比

为了验证本模型在检测精度和网络轻量性上的优势,实验选取了 YOLOv3^[21]、YOLOv5、YOLOv6^[22]、YOLOv8^[23]、YOLOv9、YOLOv11^[24]等 6 种检测任务中的主流方法,所有的模型均采用最轻量的版本,从表 3 可以看出,本文提出的 RDSM-YOLO 在公开数据集 BBD100k 检测夜间车辆,实现了平均精度(mAP_{50})为

70%,准确率(precision,P)和召回率(recall,R)分别为 82.6%和 61.5%,与其他模型模型相比,在保证较小模型参数量和网络复杂度的前提下,有效提高了夜间车辆的检测能力。此外,本研究使用单张 RTX 4060 显卡来测试各模型在推理时的 FPS,可以看出本文的 RDSM-YOLO 模型满作为一种轻量级的车辆检测模型,能够在保持着较低参数量和计算复杂度的同时满足交通的部署需求。

表 3 不同模型对比结果
Table 3 Comparison results of different models

模型	$mAP_{50}/\%$	P/ $\%$	R/ $\%$	Params/M	GFLOPs	FPS
YOLOv3	58.5	82.9	52.1	12.13	9.52	141.2
YOLOv5	67.3	81.6	59.5	2.51	3.59	92.4
YOLOv6	67.1	81.5	59.8	4.24	5.94	104.1
YOLOv8	68.6	82.3	60.3	3.01	4.10	90.1
YOLOv9	68.0	81.8	59.7	6.80	1.94	76.4
YOLOv11	67.5	80.8	59.5	6.80	1.94	89.2
RDSM-YOLO	70.0	82.6	61.5	3.04	3.92	77.2

为了能够更加直观的看出各模型的表现,实验设置蓝色框为模型成功检测出的目标,红色框展示模型漏检的目标,结果如图 9 所示。综合图 9 中各模型的检测结果可以

看出,改进的算法在重叠的检测对象以及小目标检测任务,漏检率均有明显的提升,可以充分证实改进后的算法对于低光车辆检测能力的提高。

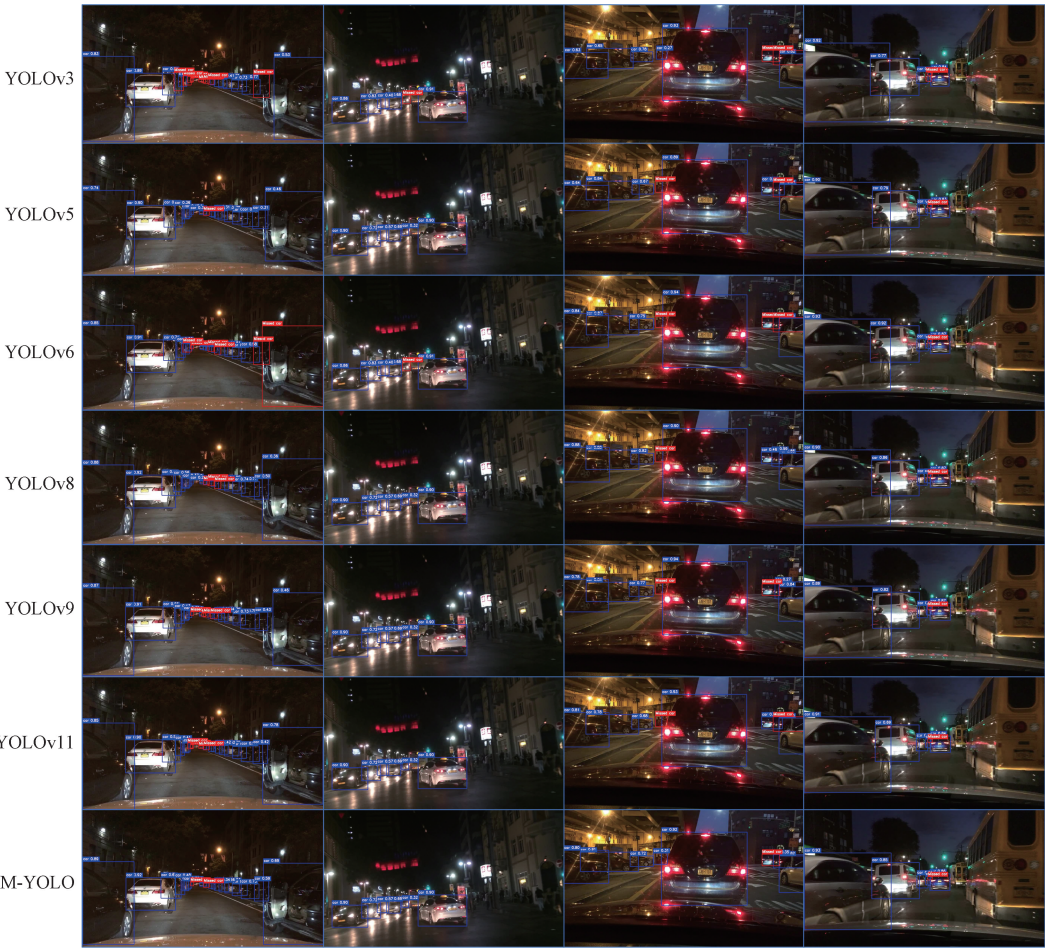


图 9 各个模型检测结果的可视化对比

Fig. 9 Visual comparison of detection results across models

2)不同模块的对比实验

为了验证改进的 RDSM-YOLO 算法各个模块的优越性,设计对本文修改的 4 个模块分别与同位置模块做对比实验,实验的结果如表 4~7 所示。

表 4 RFACnv 模块对比

Table 4 RFACnv module comparison %

方法	P	R	mAP ₅₀
ODConv	81.5	60.5	68.0
SCDown	82.2	59.7	68.1
DynamicConv	81.0	60.6	67.9
RFACnv	83.0	61.4	69.4

在 RFACnv 模块对比实验中,同本文使用的模块与 ODConv^[25]、SCDown^[26]、DynamicConv 分别在相同实验环境下进行测试,RFACnv 模块在准确性和召回率均有较高的提升,证实了模型检测能力的提升;C2f_DynamicConV 模块对比实验中选取了同样的轻量型模块修改,本文选取的模型与其他模块相比准确性和召回率均有稳定提升,最终

表 5 C2f_DynamicConV 模块对比

Table 5 C2f_DynamicConV module comparison %

方法	P	R	mAP ₅₀
C2fCIB	81.0	59.2	67.4
C2f_ODConv	81.8	60.5	68.4
C2f_OREPA	81.8	60.3	68.2
C2f_DynamicConV	82.5	60.9	68.7

表 6 SPPELAN 模块对比

Table 6 SPPELAN module comparison %

方法	P	R	mAP ₅₀
AIFI	81.8	60.2	68.2
SPPELAN	82.4	61.7	69.3

在参数量相同的条件下 C2f_DynamicConV 的检测精度提升最高;SPPELAN 模块对比实验中,为保证参数量相近,选择轻量化模型融合结构 AIFI 与本文选用模块进行对比试验,结果证实了 SPPELAN 的有效性;在 EIou 模块对比实验中,选取了 Focus CIou^[27]、DIou 热门的损失函数模

表 7 ELoU 模块对比

Table 7 ELoU module comparison %

方法	P	R	mAP ₅₀
FocusCIoU	82.5	61.1	69.1
DIoU	82.7	60.9	69.1
EIoU	82.8	60.8	70.0

块,最终在相同的运行环境本文所选模块平均检测精度最高。



图 10 KITTI 数据集示例

Fig. 10 KITTI dataset examples

不同类别的实验结果如表 8 所示。相较于 YOLOv8 算法,改进后的模型在保持相同的参数量和网络复杂度的情况下,除了在行人的准确性上有一些下降以外,其他各部分指标均有不同程度的提升;行人的检测结果下降主要归因于模型改进更侧重车辆特征,夜间预训练权重对“高瘦”目标框优化不足,模型后续将通过引入小目标检测头缓解。

表 8 KITTI 数据集各类结果

Table 8 KITTI dataset results by category %

模型 类别	YOLOv8			RDSM-YOLO		
	P	R	mAP ₅₀	P	R	mAP ₅₀
汽车	88.4	87.3	93.1	88.8	87.7	93.7
行人	83.3	54.7	66.4	84.6	55.6	69.4
骑行者	80.3	60.9	69.5	78.7	65.4	73.0

总体的对比结果如表 9 所示。RDSM-YOLO 在白天车辆检测中的 mAP₅₀ 相较 YOLOv8 上升了 2.4%,实验充分验证模型在白天车辆检测的性能,证实模型具有高鲁棒性。

表 9 KITTI 数据集方法对比

Table 9 Comparison of methods on the KITTI dataset

模型	mAP ₅₀ /%	Params/M	GFLOPs	FPS
YOLOv8	76.6	6.8	1.94	65.3
RDSM-YOLO	79.0	6.8	1.94	54.3

3.4 泛化实验

为了满足车辆检测模型的搭载需求、验证检测模型的鲁棒性,选用 KITTI 数据集进行正常光照条件下的车辆检测来验证模型的泛化能力。KITTI 数据集由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办,是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集。采用与之前实验相同的环境配置进行测试,数据集部分图像如图 10 所示。

4 结 论

针对夜间车辆检测精度低的问题,本文基于 YOLOv8n 提出了 RDSM-YOLO 检测算法。将主干网络和颈部网络中的标准卷积替换为 RFACnv,提高模型的检测精度,使用 DynamicConv 重构 C2f 模块,用 SPPELAN 模块替换原 SPPF,降低网络的复杂程度,将 CIoU 损失函数升级为 EIoU,提升模型的定位精度。实验结果表明 RDSM-YOLO 模型在夜间车辆检测中表现出更高的精确度,并且模型参数量小便于轻量化部署,同时在白天车辆检测中也有更好的性能,验证了模型的鲁棒性与高泛用性。

参考文献

[1] WANG SH Q, XU X, CHEN H Y, et al. Low-light salient object detection meets the small size[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2024, 9(4): 2754-2766.

[2] 张荣辉,邹润宗,刘佑豪. 面向夜间低光照场景下的轻量化车辆检测算法[J/OL]. 吉林大学学报(工学版), 1-17[2025-11-25]. <https://doi.org/10.13229/j.cnki.jdxbgxb.20241286>.

ZHANG R H, ZOU R Z, LIU J H. A lightweight vehicle detection algorithm for low-light nighttime scenes[J/OL]. Journal of Jilin University(Engineering Science), 1-17 [2025-11-25]. <https://doi.org/>

10. 13229/j.cnki.jdxbgxb.20241286.
- [3] ZHANG X X, STORY B, RAJAN D. Night time vehicle detection and tracking by fusing vehicle parts from multiple cameras [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23 (7): 8136-8156.
- [4] 李慧,白鹏飞,李世晓,等. 基于车灯特征的夜间车辆检测方法[J]. 电子测量技术, 2020, 43(14): 89-95.
LI H, BAI P F, LI SH X, et al. Nighttime vehicle detection method based on headlight features [J]. Electronic Measurement Technology, 2020, 43(14): 89-95.
- [5] 董雪雪. 夜间车辆检测算法研究及其应用[D]. 南京: 南京信息工程大学, 2017.
DONG X X. Research on nighttime vehicle detection algorithms and their applications [D]. Nanjing: Nanjing University of Information Science and Technology, 2017.
- [6] 郭佳霖,智敏,殷雁君,等. 图像处理中 CNN 与视觉 Transformer 混合模型研究综述[J]. 计算机科学与探索, 2025, 19(1): 30-44.
GUO J L, ZHI M, YIN Y J, et al. A review of hybrid models combining cnns and visual transformers in image processing [J]. Computer Science and Exploration, 2025, 19(1): 30-44.
- [7] CHEN L, HU X M, XU T, et al. Turn signal detection during nighttime by CNN detector and perceptual hashing tracking[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18 (12): 3303-3314.
- [8] 王建云,李纳,吴正平,等. 基于 Deepsort 和 YOLOv3 算法的车灯检测与追踪[J]. 计算机仿真, 2022, 39(11): 173-178.
WANG J Y, LI N, WU ZH P, et al. Detection and tracking of vehicle headlights based on deepsort and YOLOv3 algorithms[J]. Computer Simulation, 2022, 39(11): 173-178.
- [9] 郑玉珩,黄德启. 改进 MobileViT 与 YOLOv4 的轻量化车辆检测网络[J]. 电子测量技术, 2023, 46(2): 175-183.
ZHENG Y H, HUANG D Q. A lightweight vehicle detection network by improving mobilevit and YOLOv4 [J]. Electronic Measurement Technology, 2023, 46(2): 175-183.
- [10] 霍爱清,郭岚洁,冯若水. 面向密集场景的多目标车辆检测算法[J]. 电子测量技术, 2024, 47(9): 129-136.
HUO AI Q, GUO L J, FENG R SH. Multi-object vehicle detection algorithm for dense scenes [J]. Electronic Measurement Technology, 2024, 47(9): 129-136.
- [11] 高妍. 特征融合 CNN 的夜间行车视频检测系统研究[D]. 南京: 南京大学, 2018.
GAO Y. Research on nighttime vehicle detection system using feature fusion CNN [D]. Nanjing: Nanjing University, 2018.
- [12] 宋建辉,刘鑫,庄爽,等. 面向无人驾驶的多任务环境感知算法研究[J]. 电子测量与仪器学报, 2025, 39(1): 122-132.
SONG J H, LIU X, ZHUANG SH, et al. Research on multi-task environmental perception algorithms for autonomous driving [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(1): 122-132.
- [13] 邹亮,郭泽沛,李颖娜,等. 基于掩码孪生网络与幂调节损失的恶劣场景车道线检测方法[J]. 仪器仪表学报, 2025, 46(9): 134-145.
ZOU L, GUO Z P, LI Y N, et al. A lane line detection method for harsh scenarios based on mask-twin networks and power-adjusted loss [J]. Chinese Journal of Scientific Instrument, 2025, 46(9): 134-145.
- [14] ZHANG X, LIU C, YANG D G, et al. RFACConv: Innovating spatial attention and standard convolutional operation [J]. ArXiv preprint arXiv:2304.03198, 2023.
- [15] HAN K, WANG Y H, GUO J Y, et al. ParameterNet: Parameters are all you need [J]. ArXiv preprint arXiv:2306.14525, 2023.
- [16] CHEN Y P, DAI X Y, LIU M CH, et al. Dynamic convolution: Attention over convolution kernels [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11030-11039.
- [17] WANG C Y, YE H I, MARK LIAO H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 1-21.
- [18] ZHENG ZH 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] ZHANG Y F, REN W Q, ZHANG ZH, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [20] YU F, CHEN H F, WANG X, et al. Bdd100k: A diverse driving dataset for heterogeneous multitask learning [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 2636-2645.

[21] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv:1804.02767,2018.

[22] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. ArXiv preprint arXiv:2209.02976, 2022.

[23] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems(ADICS),2024:1-6.

[24] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements[J]. ArXiv preprint arXiv:2410.17725,2024.

[25] LI CH, ZHOU AO J,YAO AN B. Omni-dimensional dynamic convolution [J]. ArXiv preprint arXiv:2209.07947,2022.

[26] WANG AO, CHEN H, LIU L H, et al. YOLOv10:

Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.

[27] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. IEEE International Conference on Computer Vision,2017:2980-2988.

作者简介

杨慧敏(通信作者),博士,高级工程师,主要研究方向为车辆检测、智慧交通。

E-mail:yanghuimin0608@163.com

李瑞涛,硕士研究生,主要研究方向为车辆检测、智慧交通。

E-mail:15229677679@163.com

高小雯,硕士研究生,主要研究方向为目标检测。

E-mail:19930531545@163.com

王汉霞,硕士研究生,主要研究方向为目标检测、交通运输。

E-mail:550704713@qq.com.