

DOI:10.19651/j.cnki.emt.2417036

结合非相邻特征的轻量化道路多目标检测算法^{*}

黄艳国¹ 彭建¹ 方敏杰¹ 吴水清²

(1.江西理工大学电气工程与自动化学院 赣州 341000; 2.江西永安交通设施科技股份有限公司 吉安 343000)

摘要: 在复杂道路环境下进行多目标检测时,现有算法存在识别效果差、参数量大、计算复杂度高等问题,不适合在资源有限的移动端设备部署。对此,基于YOLOv7-tiny提出一种结合非相邻特征的轻量化道路多目标检测算法。首先,设计轻量级非相邻特征金字塔网络Tiny-AFPN对不同尺度的非相邻特征进行融合,减少因尺度差异造成的特征损失,实现更丰富的跨尺度信息交互,提升模型的泛化能力。其次,通过引入分布偏移卷积DSConv,重新设计了高效层聚合网络ELAN,并命名为ELAN-DS,在优化高效层聚合网络、降低模型复杂度的同时提升特征表达能力。最后,采用MPDIoU损失函数,增强网络对目标的检测能力,提高了边界框回归的准确性。在SODA10M的实验中,改进后的YOLOv7-tiny算法相较于原模型的准确率、mAP@0.5和召回率分别提升了1.4%、1.4%、5.9%,参数量和计算量分别减少了8.2%和41.5%,有效降低了模型的参数量和计算复杂度,大幅度提升了模型的检测速度,为在边缘设备上的部署提供了可能性。

关键词: 智能交通;车辆检测;轻量化;YOLOv7-tiny;深度学习

中图分类号: TP391.4; TN914 文献标识码: A 国家标准学科分类代码: 520.20

Lightweight road Multi-target detection algorithm combining asymptotic feature

Huang Yanguo¹ Peng Jian¹ Fang Minjie¹ Wu Shuiqing²(1. School of Electric Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China;
2. Jiangxi Yong'an Traffic Facilities Technology Co., Ltd., Ji'an 343000, China)

Abstract: In complex road environments, existing algorithms for road multi-target detection suffer from poor recognition performance, large number of parameters, and high computational complexity, making them unsuitable for deployment on resource-limited mobile devices. To address these issues, a lightweight road multi-target detection algorithm combining non-adjacent features is proposed based on YOLOv7-tiny. First, the design of the Tiny-AFPN combines non-adjacent features of different scales, reducing the loss of features caused by scale differences and achieving richer cross-scale information interaction. Secondly, with the introduction of DSConv, the ELAN was redesigned and named ELAN-DS, improving the expression of features while optimizing the efficient layer aggregation network and reducing the complexity of the model. Finally, the use of the MPDIoU loss function improves the accuracy of bounding box regression and enhances the network's target detection capabilities. In the experiments on SODA10M, compared with the original YOLOv7-tiny model, the improved algorithm increased accuracy, mAP@0.5, and recall by 1.4%, 1.4%, and 5.9%, respectively. It also reduced the number of parameters and computation by 8.2% and 41.5%, respectively. This effectively reduces the number of parameters and the computational complexity, substantially improves the detection speed of the model, and provides the possibility for deployment on edge devices.

Keywords: intelligent traffic; vehicle detection; lightweight; YOLOv7-tiny; deep learning

0 引言

随着自动驾驶技术的发展,目标检测技术在保障道路

交通安全方面的作用日益凸显。其核心任务是在道路图像中精准识别并定位各种交通标志、行人、车辆等各种道路使用者及其他潜在障碍物,从而为自动驾驶车辆及智能交通

收稿日期:2024-10-04

* 基金项目:国家自然科学基金(72061016)、多维智能感知与控制江西省重点实验室项目(2024SSY03161)、江西省研究生创新专项资金项目(YC2023-S654)、吉安市自然科学基金(20244-018595)项目资助

系统提供充足的路况数据,辅助其作出更为明智和安全的决策。因此,对道路目标的高效、准确识别,对于保证行车及行人安全具有重要意义^[1]。

近年来,深度学习技术的飞速发展为目标检测领域带来了革命性的进步。基于深度学习的目标检测方法主要可分为两阶段(two-stage)算法和单阶段(one-stage)算法。两阶段目标检测方法首先依赖于预先生成的候选区域(Region Proposal),然后使用卷积神经网络(CNN)对这些区域进行分类和位置定位。典型的两阶段算法包括RCNN^[2]、Fast R-CNN^[3] 和 Faster R-CNN^[4] 等,尽管这些算法精度较高,但由于需要进行两次特征提取,使得对图像的处理时间较长,因此在实时检测应用中受到限制。相比之下,以SSD^[5]、YOLO系列^[6-7]为代表的单阶段目标检测方法则更为高效。该类算法直接生成图像中物体的类别概率和位置坐标,只需提取一次特征即可实现目标检测。这种方法拥有较高的检测速度,但在精度等方面还存在不足,对此,许多研究人员结合算法的优缺点设计有效的模块,以获得更高的检测精度和更快的检测速度。

原睿等^[8]针对 YOLOv7 模型进行了改进,优化了特征融合与瓶颈注意力机制,减少了特征损失。但会导致模型参数量增大,检测帧率降低,需在性能与效率间权衡。黄全宇等^[9]在 YOLOv5s 算法的基础上引入 SE 模块,同时将主干网络中的部分高运算量模块替换为 Ghost 模块,提升目标在遮挡情况下的检测精度。然而,模型整体性能不足,检测精度有限,在遮挡目标和小目标检测上还有提升空间。Wang 等^[10] 在 YOLOv5 中采用深度可分离卷积,增加 Bneck 结构数量,在保证检测精度的同时使检测速度得到极大提升。贾玮迪等^[11] 将卷积替换为改进型跨阶段深度可分离卷积,并设计上下层特征融合模块,提出了一种基于 YOLOv5m 的轻量化车脸检测方法。但牺牲了大目标检测精度,并且检测速度较慢。许晓阳等^[12] 通过融合深度可分离卷积提出了 ELAN-DW 模块,显著提高了算法对不同尺度目标的检测效果。然而,该算法使用具有动态非单调 FM 的 WIoU v3 损失函数,增加了模型参数的复杂性和调优的难度,并且在训练过程中增加了计算资源的消耗。李琳等^[13] 提出了一种基于 Haar 小波下采样的 HWD_YOLOv9 算法用于车辆与行人目标检测,有效改善了原算法中的计算容量的冗余和小目标的漏检问题。但使用 YOLOv9 算法使得整体计算量以及参数量较大。闫世洋等^[14] 基于 YOLOv8 算法设计了全新的 C2f_SC 模块,增加 160×160 尺度的检测头来改进目标检测层,有效的提高了检测精度。但该方法仅针对交通标志检测中的误检、漏检问题,没有解决遮挡、重叠等交通标志的检测问题。

尽管上述算法在一些方面进行了优化,但未能权衡好性能与效率的关系,精度的提升往往以参数量的增加和检测速度的降低作为代价。此外,在检测模型较大、道路场景下目标互相遮挡、同类目标尺度较多且跨度较大以及存在

漏检和误检等问题上,还存在较大改进空间。因此,如何设计合理的目标检测算法,在复杂场景中兼顾检测性能与实时性能,是研究的重点和难点。针对上述问题,本研究以 YOLOv7-tiny 算法为基准,提出结合非相邻特征的轻量化目标检测算法,在保证检测精度的前提下,进一步解决复杂环境下目标重叠、遮挡以及目标尺度不均造成的误检、漏检现象,并使之能在低性能的边缘设备上部署。本研究主要工作为:

- 1) 设计轻量化非相邻特征金字塔网络 (tiny asymptotic feature pyramid network, Tiny-AFPN) 作为新的特征融合网络,加强非相邻层级的信息进行特征融合,减少特征在传播和交互过程中造成的丢失或退化,增强对小目标、遮挡目标的检测能力,提升模型的检测性能与泛化能力。
- 2) 使用分布偏移卷积(distribution shift convolution, DSConv)设计 ELAN-DS 模块,提升网络特征提取能力的同时降低模型的参数量和计算量,能够在资源及算力有限的设备上进行部署,提高模型的普适性。

- 3) 针对 CIoU 在真实框与预测框宽高比例相同、但宽高值不同时处理效果不佳的问题,采用 MPDIoU 损失函数代替 CIoU 损失函数,有效改善边界框的回归效果,加快模型的收敛速度。

1 YOLOv7-tiny 模型

YOLO 系列目标检测算法经过不断更新迭代目前已经发布到 YOLOv10^[15]。YOLOv7 目标检测算法是由 Bochkovskiy 等^[16]于 2022 年 7 月推出的单阶段目标检测网络。YOLOv7-tiny 则是在 YOLOv7 的基础对结构进行了简化,是 YOLOv7 系列中针对边缘设备或计算资源有限的场景而设计的轻量级模型。该算法的网络结构主要包括输入端(Input)、骨干网络(Backbone)、颈部网络(Neck)、和预测头(Head) 4 个部分,网络结构如图 1 所示。输入端的图像首先经过骨干网络生成 3 个不同尺度的特征图传入颈部网络,从而在特征图的基础上引入更多的上下文信息,进一步增强特征融合能力,最后经过预测头得到 3 个不同尺度的预测结果。虽然 YOLOv7-tiny 模型在轻量级模型中表现出了很好的检测精度和速度,但仍然存在以下不足:

- 1) YOLOv7-tiny 模型的参数量和计算量仍有改进的空间。
- 2) 检测小目标或密集目标时,YOLOv7-Tiny 在提供上下文信息和特征融合方的能力有限,导致检测效果不佳。
- 3) 在复杂多变的环境中,YOLOv7-tiny 容易出现漏检和误检的情况。

因此,这些不足限制了其在车辆检测中广泛使用。

2 改进 YOLOv7-tiny 模型

针对 YOLOv7-tiny 的缺点,提出了一种结合非相邻特

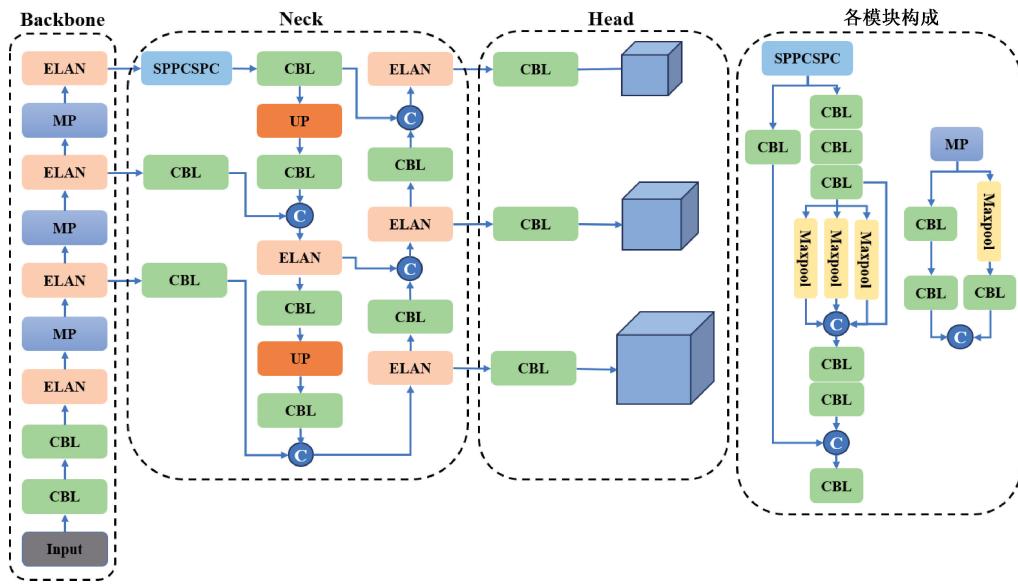


图 1 YOLOv7-tiny 网络结构

Fig. 1 YOLOv7-tiny network structure

征的轻量级 YOLOv7-tiny 模型,改进后的网络结构如图 2 所示。首先,设计 Tiny-AFPN 模块,增强非相邻层级的信息融合,提高模型的特征融合能力;将标准卷积替换为 DS 卷积,依照梯度路径设计方法设计 ELAN-DS 模块,从而提

高网络的特征提取能力;最后,通过采用 MPDIoU 损失函数来替代 CIoU 损失函数,加速模型的收敛过程,进一步提高检测精度。

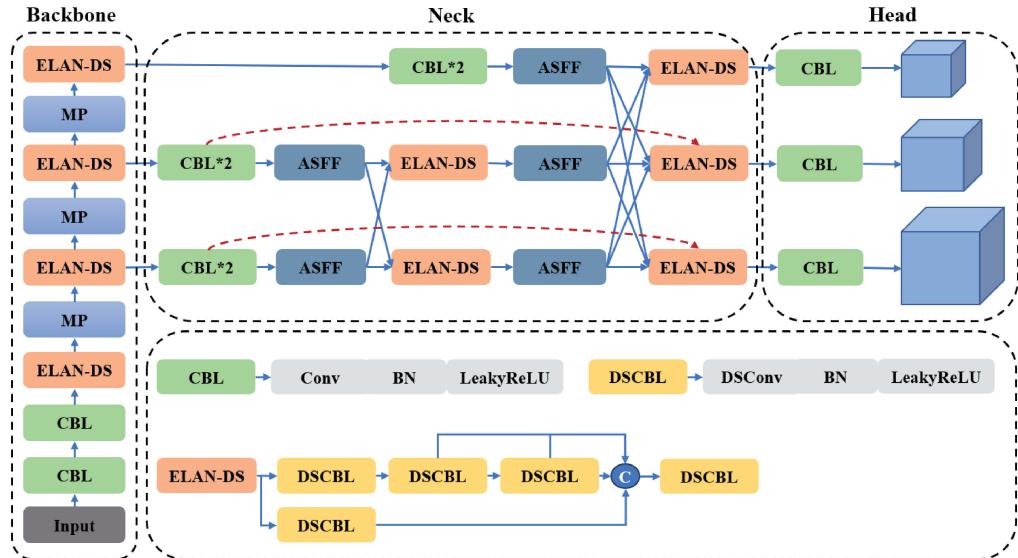


图 2 改进的 YOLOv7-tiny 网络结构

Fig. 2 Improvement of YOLOv7-tiny network structure

2.1 DSConv 模块和轻量级 ELAN-DS 模块设计

分布偏移卷积 DSConv^[17]因其出色的性能而被广泛用于模型轻量化研究中。在内存和计算复杂度方面,DSConv 将传统的卷积核分解为可变量化核(variable quantized kernel, VQK)和核分布偏移(kernel distribution shift, KDS)。如图 3 所示,VQK 是一种高效的卷积核压缩方法,在保持精度的前提下降低了参数量与模型大小,最终

近似达到使用全量化卷积核的效果。分布偏移机制则依赖于内核中的 KDS 以及通道中的分布偏移(channel distribution shift, CDS)来调整 VQK 的位置,使得模型能够适应输入数据的分布变化,从而更好捕捉全局信息。本文使用 DSConv 代替标准卷积来提取图像特征,降低模型的参数量和计算复杂度的同时,最大程度保证特征提取的准确性;同时在 YOLOv7-tiny 中引入 DSConv 对 ELAN

模块进行改进。

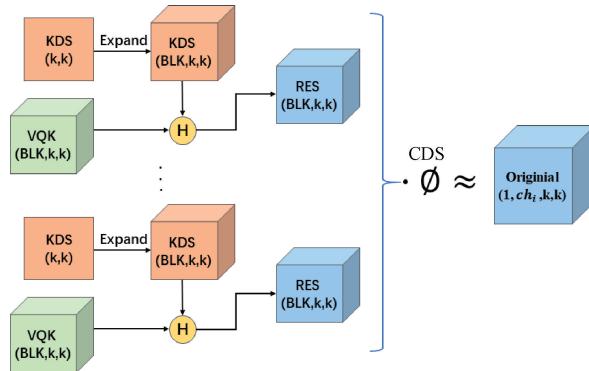


图 3 DSConv 工作方式

Fig. 3 How DSConv works

图 4 为不同高效层聚合网络 (efficient layer aggregation networks, ELAN) 的梯度路径设计策略, 其中图 4(a)为 ELAN 模块, 该模块包含两条路径, 第 1 条路径

经过一个 1×1 的卷积来改变特征图的通道数, 第 2 条路径则先经过一个 1×1 的卷积模块来调整通道数, 然后再经过 4 个 3×3 的卷积模块, 逐步提取特征信息; 图 4(b)则是 YOLOv7 在 ELAN 基础上改进的 E-ELAN 模块; 图 4(c)为 YOLOv7-tiny 中的 ELAN 模块, 在图 4(a)的第 2 条分支的基础上减少了两个 3×3 的卷积, 保持一定精度的同时降低了部分参数量与计算量; 通过对 ELAN 进行分析并参考梯度路径设计策略, 本研究设计了图 4(d)所示的 ELAN-DS 模块。该模块首先利用 1×1 卷积核用于通道数的调整, 接着利用两个步长为 1、卷积核大小为 3×3 的 DSConv 替换模块中的标准卷积进行特征提取, 使得模块在不同尺度下获取更大的感受野, 从而获得更多上下文信息。最后在模块的第 4 层使用 1×1 卷积核进行特征整合和调整。相较于原模型中的 ELAN 模块, ELAN-DS 模块更注重空间中的特征提取, 通过融合不同路径上的特征, 能够提取到更多不同层次的细节特征和语义信息进行融合, 进而提升模型的检测性能。

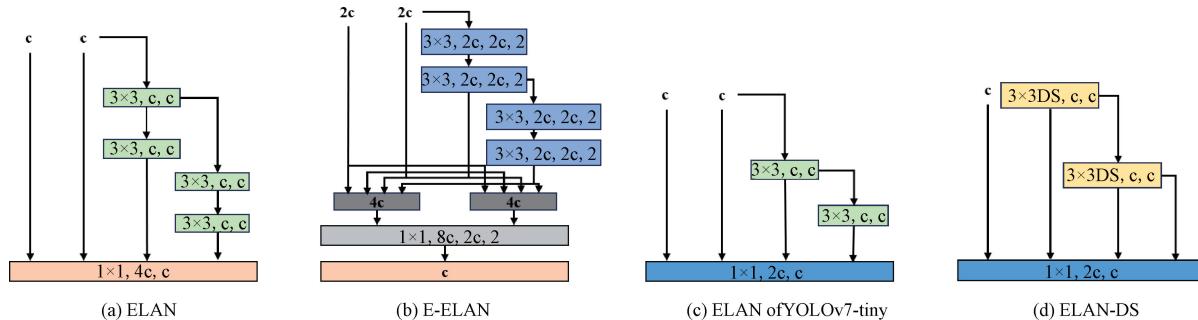


图 4 ELAN 梯度路径设计

Fig. 4 ELAN gradient path design

为评估本文提出的 ELAN-DS 模块的有效性, 对 YOLOv7-tiny 模型使用该设计, 评估本文提出的 ELAN-DS 模块的优劣性, 实验结果如表 1 所示。

表 1 改进后的模型参数

Table 1 Parameters of the improved model

网络模型	参数数量/M	GFLOPs
YOLOv7-tiny	6.1	13.5
ELAN-DS	6.3	7.9

由表 1 可知, 经过改进后的模型参数相较于原模型略有增加, 但浮点运算量远远低于原模型, 保持了更高的效率, 同时提升了网络的特征提取能力, 为之后的检测和分类提供了更为丰富的特征信息。

2.2 轻量化多尺度特征融合网络 Tiny-AFPN

目标检测模型通常引入特征金字塔结构来解决尺度变化的问题。在现有特征金字塔体系结构中, 特征在传播和交互过程容易造成信息丢失或退化。非相邻特征金字塔网络 (asymptotic feature pyramid network, AFPN)^[18] —

种利用低层、高层和顶层特征信息进行多尺度特征融合的网络。由于非相邻层次特征之间的语义差距较大, 在多级特征融合过程中, 该网络采用自适应空间融合机制 (adaptive spatial feature fusion, ASFF) 进行空间权重分配, 如图 5 所示, 使得不同层级的特征可以以适当的比例进行融合。

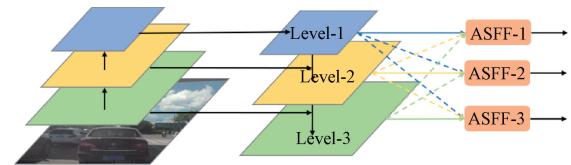


图 5 自适应空间融合

Fig. 5 Adaptive spatial fusion operation

为了满足实时检测车辆的需求并进一步解决特征丢失问题, 在 AFPN 的基础上设计一种轻量化非相邻特征融合网络 Tiny-AFPN, 3 种网络结构如图 6 所示。如图 6(a)所示, PAN 是一个双向金字塔结构, 能够有效融合自顶向下和自底向上的特征图, 使得模型能够更好的提取不同尺度的目标。但伴随着卷积的深入, PAN 结构易造成特征丢

失,于是引入更为复杂的非相邻特征金字塔网络 AFPN;如图6(b)所示,AFPN结构通过渐进连接来加强非相邻层之间的直接交互,有效地避免信息损失和冗余,提高特征的表达能力。

为满足轻量化要求并保证检测网络的特征提取能力,

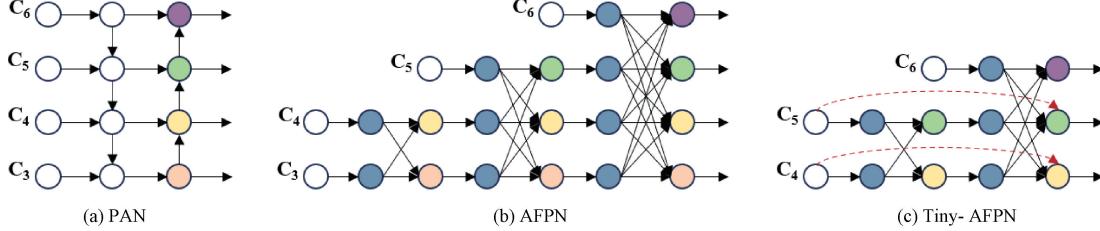


图6 PAN、AFPN、Tiny-AFPN 网络结构

Fig. 6 PAN, AFPN, Tiny-AFPN network structure

在多级特征融合过程中,由于底层和顶层特征差异大,直接将C4、C5和C6层融合不合理,因此采用特定的融合策略。首先采用ASFF对特征层C4和C5的权重进行分配,随后运用非相邻特征融合策略进行融合,接着加入C6层按相同方式与C4、C5融合,经过特征融合步骤,最终产生一组多尺度特征图{P4,P5,P6}用于道路目标预测。融合过程可以表示为对n个特征层提取的特征向量按分配的空间权重进行加权计算,其计算公式如式(1)所示。

$$y_{ij}^n = \alpha_{ij}^n \cdot x_{ij}^{1 \rightarrow n} + \beta_{ij}^n \cdot x_{ij}^{2 \rightarrow n} + \gamma_{ij}^n \cdot x_{ij}^{3 \rightarrow n} \quad (1)$$

式中: $x_{ij}^{k \rightarrow n}$ 表示从特征层k到n的位置(i,j)处的特征向量, y_{ij}^n 表示结果特征向量, α_{ij}^n 、 β_{ij}^n 、 γ_{ij}^n 表示层次特征在层次n的空间权重,受 $\alpha_{ij}^n + \beta_{ij}^n + \gamma_{ij}^n = 1$ 约束。

为了评估本研究提出的轻量化网络Tiny-AFPN的优劣性,对YOLOv7-tiny模型分别使用PAN、AFPN、Tiny-AFPN网络在SODA10 M数据集上进行对比分析。由表2可知,Tiny-AFPN在参数量和浮点运算量方面均有下降。具体而言,Tiny-AFPN的参数量仅为5.6 MB,相较于PAN的6.1 MB和AFPN的6.3 MB,分别减少了8.2%和11.1%。这种结构使得Tiny-AFPN部署在资源受限的设备时具有一定优势。同时,Tiny-AFPN的浮点运算量为12.6 G,低于PAN和AFPN,表示其在执行相同任务时所需的计算量更少,能带来更快的推理速度和更低的能耗。在mAP@0.5方面,Tiny-AFPN比PAN高1.4%,比AFPN仅减少0.2%,但在参数量和计算量方面,Tiny-AFPN网络更低,这表明Tiny-AFPN在保证精度损失不大的前提下更加轻量化。

表2 特征融合网络对比结果

Table 2 Feature fusion network comparison results

网络结构	mAP@0.5/%	参数量/M	GFLOPs
PAN	71.2	6.1	13.5
AFPN	72.8	6.3	14.8
Tiny-AFPN	72.6	5.6	12.6

本研究基于AFPN结构,保留C4-C6为特征输入层,减去C3层,使模型更轻量化。同时结合残差结构思想(如图6(c)虚线箭头),将输入特征层与处理后的输出层相加,保留更多细节信息,并将ELAN-DS模块替换ELAN模块以优化特征表达能力,减少计算复杂度。

2.3 损失函数

YOLOv7-tiny使用的是CIoU损失函数^[19],假定预测框为 $B = [x, y, w, h]$,真实框为 $B^{gt} = [x^{gt}, y^{gt}, w^{gt}, h^{gt}]$,CIoULoss的计算公式为:

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

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (3)$$

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

CIoU损失函数考虑了重叠面积、中心点距离和长宽比3种几何参数。然而在预测框与真实框具有相同的宽高比但宽度和高度值完全不一致的情况下,CIoU损失函数无法进行优化,降低了边界框回归的收敛速度和准确性。如图7所示,相比较于CIoULoss,MPDIoULoss^[20]使用一种基于最小点距离的边界框相似度比较度量法,直接最小化预测框与真实框之间的左上角和右下角点距离,并且MPDIoULoss包含了现有损失函数中考虑的所有相关因素,例如重叠或非重叠面积、中心点距离以及宽度和高度的偏差,解决了宽高比值相同但实际值完全不同的问题,使得边界框回归效果得到改善。

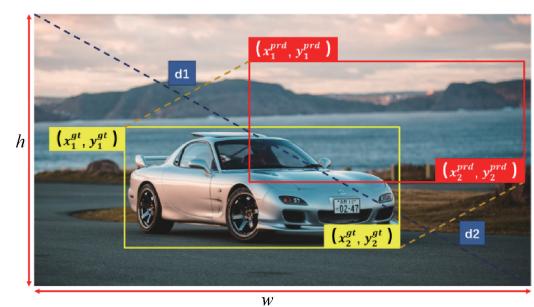


图7 MPDIoU 的影响因素

Fig. 7 Factors of MPDIoU

MPDIoULoss计算公式为:

$$MPDIoU = 1 - IoU + \frac{d_1^2}{h^2 + w^2} + \frac{d_2^2}{h^2 + w^2} \quad (5)$$

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (6)$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (7)$$

式中: (x_1^{gt}, y_1^{gt}) 、 (x_2^{gt}, y_2^{gt}) 分别表示真实框的左上角和右下角点坐标, (x_1^{prd}, y_1^{prd}) 、 (x_2^{prd}, y_2^{prd}) 分别表示预测框的左上角和右下角点坐标, w 和 h 为输入图像的宽和高。

为进一步验证 MPDIoU 损失函数的有效性, 本文在硬件设施及参数设置相同的情况下, 使用 YOLOv7-tiny 模型作为基准模型, 对 MPDIoU 与 CIoU 损失函数在 SODA10M 数据集上进行对比分析, 同时以训练前后的 Loss 曲线来衡量模型性能。

各类 Loss 值对比如图 8 所示, 分别为改进算法前后的 box_loss、obj_loss、cls_loss, 从图 8 可得, MPDIoU 损失函数在各项损失值上均表现出显著的优化, 改进后的算法比原算法的 Loss 值更小且算法收敛速度更快, 表明改进后的算法模型在训练数据上的拟合程度较好, 且具有更好的泛化能力。

3 实验结果分析

3.1 数据集与实验环境

本研究改进后的算法模型主要在 SODA10M 数据

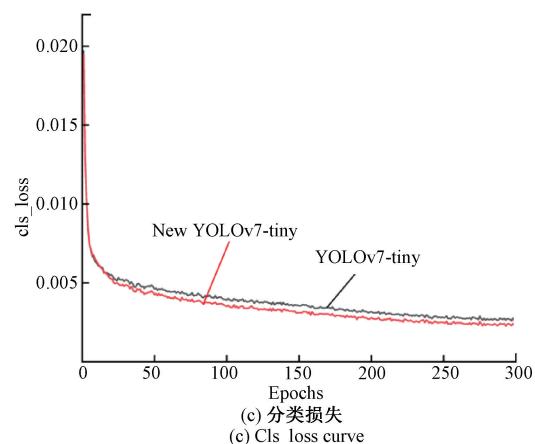
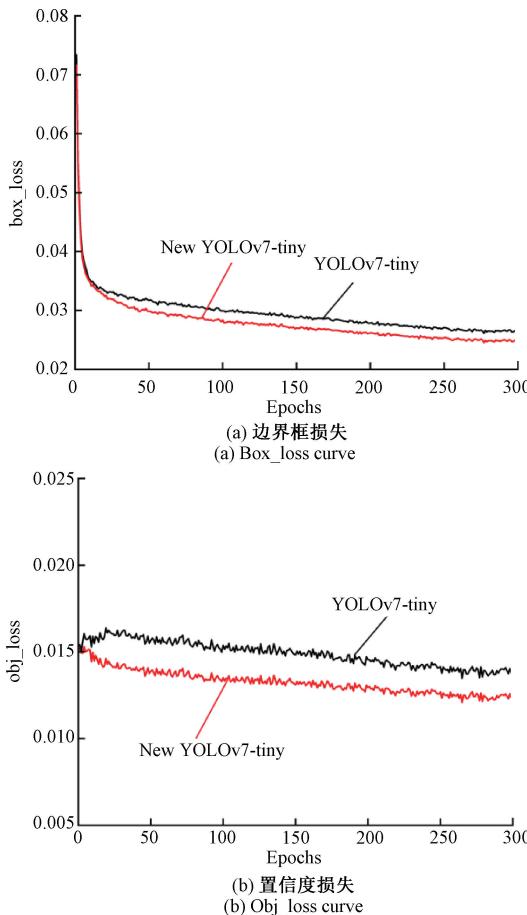


图 8 各类 Loss 值对比

Fig. 8 Comparison of loss values for each category

集^[21]进行研究分析, 同时在 KITTI 数据集^[22]和 Cityscapes 数据集^[23]上进行通用性实验验证, 以此验证本研究改进后的算法在道路目标检测方面的有效性和泛化性。

1) SODA10M 数据集

SODA10M 是由华为诺亚方舟实验室与中山大学联合发布的 2D 基准数据集, 该数据集包含了一千万张多样性丰富的无标签道路场景图片以及两万张带标签图片。本文选取其中 5 000 张标签为 Pedestrian、Cyclist、Car、Truck、Tram、Tricycle 的图像, 并按照 7:2:1 的比例划分训练集、验证集和测试集。

2) KITTI 数据集

KITTI 数据集是自动驾驶领域广泛使用的计算机视觉数据集, 主要用于评估目标检测与识别算法在车载环境中的性能。该数据集由车内第一视角拍摄, 共包含 7 481 张图像, 涵盖 Car、Cyclist、Tram、Pedestrian、Van、Person_sitting、Misc、Truck、DontCare 9 个类别, 训练集、验证集和测试集的划分标准同 SODA10M 数据集。

3) Cityscapes 数据集

Cityscapes 数据集是一个针对城市街景理解的 2D 基准数据集, 总共有 5 000 张覆盖全球多个城市的真实场景图像。与 SODA10M 数据集相比, 该数据集包含了大量高质量、精细标注的街景图像, 展现了极高的多样性和复杂性。训练集、验证集和测试集的划分标准同 SODA10M 数据集。

3.2 评价指标

为确保实验的稳定性和可重复性, 采用统一的实验环境与参数配置, 具体参数设置如表 3 所示。

在本次研究中, 采用模型参数量、浮点运算量 (FLOPs)、平均精度均值 (mAP)、精确率 (precision, P)、召回率 (recall, R) 以及每秒检测帧数 (FPS) 来评估目标检测模型的表现效果。如式(8)~(11)所示, TP、FP 和 FN 分别表示正样本被正确识别、误报的负样本以及漏报的正样本数量。

表3 实验环境设置

Table 3 Lab environment settings

实验环境	参数
操作系统	Window 10
处理器	12th Gen Intel(R) Core(TM) i7-12650H@2.30 GHz
GPU	Nvidia GeForce RTX4060
内存	16 GB
代码编程	python 3.11.4
深度学习框架	Pytorch 2.0.1
加速库	Cuda 11.7
epoch	300
batch_size	8

平均精度均值通过计算各个类别的平均精度(AP)的总和除以整体类别数,是衡量模型整体检测性能的关键指标,本文选取mAP@0.5和mAP@0.5:0.95作为评估模型精度的指标。

表4 消融实验
Table 4 Ablation experiment

模型	Tiny-AFPN	ELAN-DS	MPDIoU	参数量/M	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	GFLOPs
1				6.1	82.1	61.8	71.2	46.8	13.5
2	✓			5.6	83.2	62.3	72.2	48.3	12.6
3		✓		6.3	83.6	64.2	70.7	47.7	7.9
4			✓	6.1	83.3	62.9	71.8	47.3	13.5
5	✓	✓		5.8	83.7	62.6	72.3	48.6	8.3
6	✓		✓	5.6	82.9	63.5	72.5	48.4	12.6
7		✓	✓	5.9	82.4	62.8	72.3	47.6	7.6
8	✓	✓	✓	5.8	83.9	65.7	72.6	48.9	7.9

由表4可知,采用轻量级Tiny-AFPN特征融合网络相对于原网络,参数量下降了0.5 MB,准确率、mAP@0.5和mAP@0.5:0.95分别提升了1.1%、1.0%、1.5%。说明Tiny-AFPN特征融合网络在保证轻量化的同时,能有效地利用多尺度特征融合机制对不同层级特征进行融合,以提高目标检测的准确性和泛化能力。

此外,通过采用本研究设计的轻量级ELAN-DS模块替换原模型中的ELAN模块,虽然mAP@0.5下降了0.5%,模型的参数量略有提升,但浮点运算量降低了40%,远远低于原模型,且准确率较原模型提升了1.5%。这一结果表明,本文设计的轻量化ELAN-DS模块很好的降低了模型的计算复杂度,并且通过融合不同层级多个路径的语义信息与细节特征,有效地捕捉检测目标的上下文信息,在参数量与准确性之间取得了很好的平衡。

值得注意的是,本文采用的MPDIoU损失函数相较于CIOU损失函数,模型的准确率、mAP@0.5和mAP@0.5:

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \frac{\sum AP}{NC} \times 100\% \quad (9)$$

准确率表示预测结果中正样本与所有样本的比值。

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

召回率等于数据集中被正确识别出的正样本数量除以所有真实的正样本数量的值。

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

FPS反映模型在单位时间内能够处理的图片数量。

3.3 消融实验

为了验证本文所提各项改进的有效性,设计一组消融实验在SODA10M数据集上进行对比分析,了解不同模块对模型性能的影响。为确保实验的公平性和准确性,训练过程均采用相同的参数设置,实验结果如表4所示。其中“✓”表示消融实验中采用该模块,否则表示未使用。

0.95分别提高了1.2%、0.6%、0.5%,这一提升归功于MPDIoU损失函数充分利用了边界框回归的几何性质并综合考虑了现有损失函数中的所有相关因素,从而提高模型的泛化能力。

综上所述,通过对YOLOv7-tiny和不同的模块组合进行消融实验分析,本研究改进算法在模型参数量、计算复杂度与精度方面都表现出更优越的性能。

3.4 算法对比实验

为了验证改进后算法的轻量化程度、计算量与准确性等衡量指标,将本研究提出的模型与YOLO系列算法、SSD、Faster-RCNN以及其他文献改进算法进行对比实验,结果如表5所示。

从参数量的角度来看,改进算法的参数量仅为5.8 M,远低于YOLOv5s、YOLOv7-tiny、YOLOv8s等轻量级算法。这一显著优势使得改进算法在模型大小和计算资源消耗上具有更好的表现,特别是在资源受限的环境下,改

表 5 与其他算法的对比实验

Table 5 Comparative experiment with other algorithms

网络模型	参数量/M	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	GFLOPs	FPS
SSD	26.3	81.2	60.1	69.3	45.5	62.8	168.3
Faster-RCNN	136.8	82.7	61.6	68.2	45.1	369.8	37.6
YOLOv3	61.5	80.9	59.7	61.7	40.3	154.6	56.2
YOLOv5s	9.1	81.1	60.4	68.5	45.2	15.8	119
YOLOv7	37.2	85.7	66.3	74.2	49.7	105.2	86.2
YOLOv7-tiny	6.1	82.5	61.8	71.2	46.8	13.5	188.7
文献[24]	43.6	85.9	66.9	74.8	49.9	121.5	98
YOLOv8s ^[25]	11.2	84.8	66.1	73.5	49.2	28.6	146.7
YOLOv9 ^[26]	70.1	85.4	66.5	74.3	49.9	313.4	43.6
改进算法	5.8	83.9	65.7	72.6	48.9	7.9	213.5

进算法能够更高效地运行,同时保持较好的检测性能。

其次,在精确率和召回率方面,改进算法的精确率达到了 83.9%,召回率达到了 65.7%,在同类算法中保持着较高的检测水平。尽管其精确度和召回率可能不是所有算法中最高的,但考虑到其低参数量、低计算量的特点,这一表现已经相当可观。

此外,在 mAP@0.5/% 和 mAP@0.5:0.95/% 两个评估指标上,改进算法分别为 72.6%、48.9%,相较于 SSD(69.3% 和 45.5%)、Faster-RCNN(68.2% 和 45.1%)以及 YOLOv3(61.7% 和 40.3%)而言,改进算法在 mAP@0.5/% 上略有优势,在 mAP@0.5:0.95/% 上则展现出更强的竞争力。这表明改进算法在复杂场景和存在多种不同尺度目标的检测任务中具有较好的性能。

在计算量方面,改进算法的计算量仅为 7.9 G,这一数值远低于 SSD 的 62.8 G、Faster-RCNN 的 369.8 G 以及 YOLOv3 的 154.6 G。同时,与一些轻量级算法相比,如 YOLOv5s 的 15.8 G、YOLOv7-tiny 的 13.5 G 和 YOLOv8s 的 28.6 G,改进算法的计算量也相对较低,在计算资源要求较高的场景下具有更好的表现。

最后,从 FPS 来看,改进算法的处理速度高达 213.5 fps,远高于 YOLOv7 的 105.2、文献[24]的 98 fps 以及 YOLOv9 的 43.6 fps。这表明改进算法在实际应用中仍然具有较好的实时性,能够满足大多数实时检测任务的需求。

通过算法对比实验可以得出,本文改进的 YOLOv7-tiny 算法在保持较高精度的同时实现了轻量化。此外该算法还展示出了较低的计算量以及较快的实时检测速度,优于其他对比算法的性能,在这为模型在资源有限的边缘终端设备上部署提供了可行性。

3.5 通用性对比试验

为了进一步验证改进后算法的泛化能力,将改进后的算法、YOLOv7-tiny 和 YOLOv7 算法在 KITTI 数据集和 Cityscapes 数据集上进行泛化实验,训练前后的结果如

表 6、表 7 所示。

表 6 KITTI 数据集泛化实验

Table 6 KITTI dataset generalization experiment

网络模型	P/%	R/%	mAP@0.5/%	GFLOPs
YOLOv7-tiny	83.8	73.5	80.9	13.5
YOLOv7	86.4	76.9	84.1	105.2
改进算法	85.5	75.8	82.4	7.9

表 7 Cityscapes 数据集泛化实验

Table 7 Cityscapes dataset generalization experiment

网络模型	P/%	R/%	mAP@0.5/%	GFLOPs
YOLOv7-tiny	85.3	75.1	81.8	13.5
YOLOv7	86.8	77.4	85.3	105.2
改进算法	86.1	76.7	83.7	7.9

由表 6 可知,相较于 YOLOv7-tiny 模型,改进后的算法在 KITTI 数据集上的准确率、召回率、mAP@0.5 分别提升了 1.7%、2.3%、1.5%;同时计算量减少了 5.6 G,显著降低了计算成本;相较于 YOLOv7 算法,改进后的算法很好的平衡了模型的参数量、计算量与检测精度的关系,能够很好的部署在低性能设备上。

由表 7 可知,相较于原模型,改进后的算法在 Cityscapes 数据集上的各项参数均获得了不同程度的提升,其中准确率、召回率和 mAP@0.5 分别提升了 0.8%、1.6%、1.9%;相较于 YOLOv7 模型,改进后的算法在保持较高检测精度的同时显著降低了计算成本,体现了模型的实际部署价值。

实验结果表明,改进后的算法能够保留更多特征信息,提高检测性能,在复杂道路场景下有着更好的鲁棒性,进一步验证了改进算法的有效性和泛化能力。

3.6 检测效果分析

为了更加直观的展示改进后算法的实际效果,图 9 展

示了原模型与改进 YOLOv7-tiny 算法在 4 个场景下的检测对比结果,包含小目标密集、背光且小目标多、光线不均与目标被遮挡以及光照不足的场景,白框表示有明显检测改善的区域。从图 9 可以看出,原模型识别目标的置信度相对较低,且误检、漏检较多,本研究算法的类别检测精度比原模型有一定的提高。图 9(a)所示的场景中,原模型将绿衣服行人右侧物体误检测为汽车,且未能检测到远距离的车辆,存在漏检现象。改进后的算法能够检测到更多出现在车辆视角下的多个目标,有效降低了漏检率。如图 9(b)所示,在背光且背景复杂的环境下,原模型很难检测到图中右侧的模糊目标,检测精度较低。如图 9(c)所示,在光线不均且大多数目标被遮挡的情况下,原模型错误的将

图中左侧被遮挡的不同车辆识别为同一辆车,识别效果较差。如图 9(d)所示,原模型在光照不足的场景下没有识别到右侧目标,且识别目标的置信度较低。相比之下,嵌有 ELAN-DS 模块的轻量级 Tiny-AFPN 网络能够扩大网络的感知范围,加强非相邻信息的融合,从而捕捉到更多的特征信息。因此,从对比图中可以看出,改进后的算法能够在光照不均以及背景复杂的环境下,降低光线以及环境带来的干扰,提高对小目标、遮挡目标以及模糊目标的检测率,减少了漏检和误检的情况,同时对同一目标的检测置信度也有所提高,在复杂道路场景中有着更好的鲁棒性,更卓越的检测性能,满足了车辆在行驶过程中轻量化、高精度检测的需求。

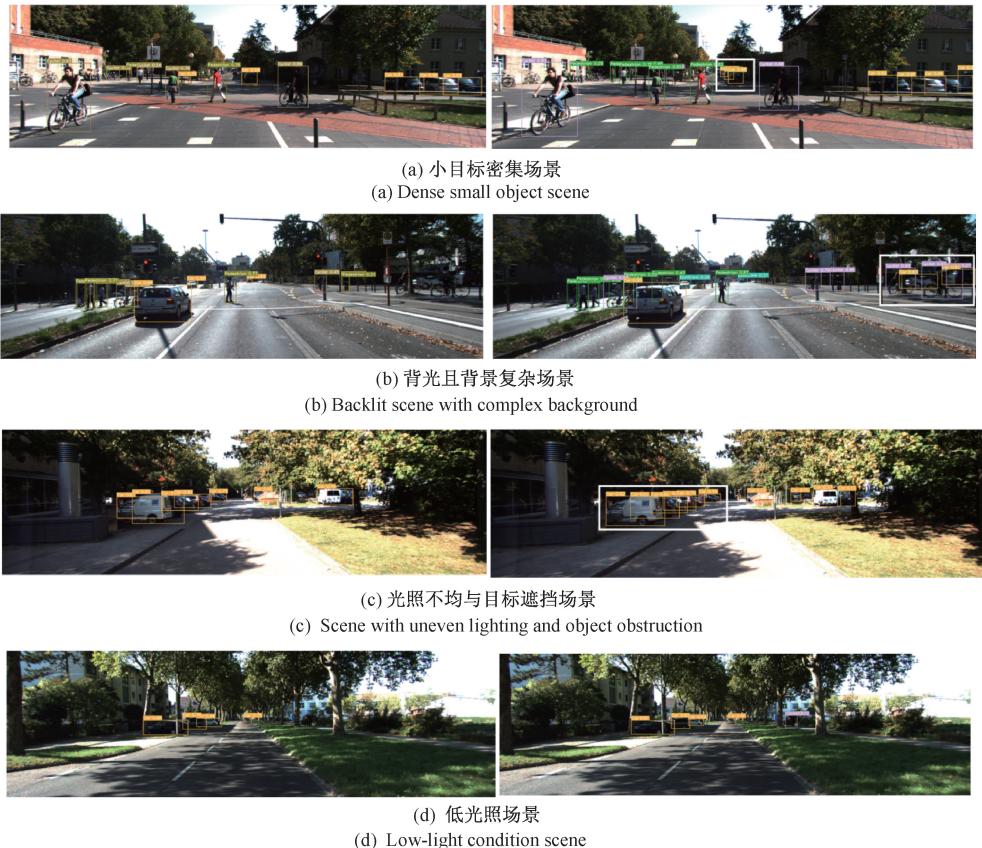


图 9 不同场景下的检测效果对比(白框表示有明显检测改善的区域)

Fig. 9 Comparison of detection effects in different scenes (white boxes indicate areas of significant detection improvement)

4 结 论

为了进一步降低道路多目标检测算法的参数量和计算量,同时提高待检测目标的精度,本文提出了一种结合非相邻特征的轻量级 YOLOv7-tiny 道路多目标检测算法。首先,采用 Tiny-AFPN 多尺度特征融合网络,对非相邻级别的特征进行融合,加强不同尺度特征间的传播,减少不同尺度特征因融合时造成的特征损失,提升模型的泛化能力。其次,利用 DSConv 的优势设计了 ELAN-DS 模块,在

模型进行特征聚合时,增强特征信息的传递和表达,使得模型在复杂道路场景下能够更好地捕捉和识别目标。最后,在损失函数方面,将 CIoU 损失函数更换为 MPDIoU 函数,改进预测框与真实框具有相同的宽高比,但宽度和高度值完全不同时无法优化的情况,使得模型在边界框回归时效果更佳,收敛速度也得到了提升。实验结果表明,本文提出的改进算法在准确率、召回率、平均准确率均值、FPS 等关键指标上均高于原算法模型,证明了本文算法能够部署在低算力环境下的辅助驾驶系统中,并在复杂道路

场景下完成目标检测任务。在后续的研究中,将考虑进一步提高模型的检测精度并且继续对模型进行压缩,提高模型的实用性。

参考文献

- [1] 侯学良,单腾飞,薛靖国.深度学习的目标检测典型算法及其应用现状分析[J].国外电子测量技术,2022,41(6):165-174.
- [2] HOU X L, SHAN T F, XUE J G. Analysis of typical deep learning-based object detection algorithms and their application status [J]. Foreign Electronic Measurement Technology, 2022,41(6):165-174.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [4] GIRSHICK R. Fast region-based convolutional network method[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [5] REN SH 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, 2015, 39(6): 1137-1149.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, 2016: 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLOv3: An incremental improvement [C]. 31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 2018.
- [9] 原蕾,王科俊.基于注意力机制与特征融合的改进YOLOv7车辆检测方法[J].国外电子测量技术,2023,42(9):49-57.
- [10] YUAN L, WANG K J. Improved YOLOv7 vehicle detection method based on attention mechanism and feature fusion [J]. Foreign Electronic Measurement Technology, 2023, 42(9): 49-57.
- [11] 黄全宇,胡斌杰,朱婷婷.面向道路交通场景的轻量级目标检测方法[J].现代电子技术,2022,45(3): 88-95.
- [12] HUANG T Y, HU B J, ZHU T T. Lightweight object detection method for road traffic scene [J]. Modern Electronics Technique, 2022,45(3): 88-95.
- [13] WANG CH P, WANG H Q, YU F J, et al. A high-precision fast smoky vehicle detection method based on improved Yolov5 network [C]. 2021 IEEE International Conference on Artificial Intelligence and Industrial Design(AIID), 2021: 255-259.
- [14] 贾玮迪,余鹏飞,余国豪,等.基于改进 YOLOv5m 的轻量化车脸检测方法[J].电子测量技术,2023,46(12):125-133.
- [15] JIA W D, YU P F, YU G H, et al. Lightweight car front detection method based on improved YOLOv5m [J]. Electronic Measurement technology, 2023, 46(12): 125-133.
- [16] 许晓阳,高重阳.改进 YOLOv7-tiny 的轻量级红外车辆目标检测算法[J].计算机工程与应用,2024,60(1):74-83.
- [17] XU X Y, GAO C H Y, Improved YOLOv7-tiny lightweight infrared vehicle target detection algorithm[J]. Computer Engineering and Applications, 2024,60(1): 74-83.
- [18] 李琳,靳志鑫,俞晓磊,等. Haar 小波下采样优化 YOLOv9 的道路车辆和行人检测[J].计算机工程与应用,2024,60(20):207-214.
- [19] LI L, JIN ZH X, YU X L, et al. Road vehicle and pedestrian detection based on YOLOv9 for Haar wavelet downsampling[J]. Computer Engineering and Applications, 2024,60(20):207-214.
- [20] 闫世洋,罗素云.基于 SC-YOLOv8 的交通标志检测算法研究[J].电子测量技术,2024,47(15):117-124.
- [21] YAN SH Y, LUO S Y. Research on traffic sign detection algorithm based on SC-YOLOv8 [J]. Electronic Measurement Technology, 2024, 47 (15): 117-124.
- [22] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458,2024.
- [23] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [24] NASCIMENTO M G, FAWCETT R, PRISACARIU V A. Dsconv: Efficient convolution operator [C]. IEEE/CVF International Conference on Computer Vision, 2019: 5148-5157.
- [25] YANG G Y, LEI J, ZHU ZH K, et al. AFPN: Asymptotic feature pyramid network for object detection[J]. ArXiv preprint arXiv:2306.15988, 2023.
- [26] 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.]
- [20] MA S L, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662, 2023.
- [21] HAN J H, LIANG X W, XU H, et al. SODA10M: A large-scale 2D self/semi-supervised object detection dataset for autonomous driving [J]. ArXiv preprint arXiv: 2106.11118, 2021.
- [22] UHRIG J, SCHNEIDER N, SCHNEIDER L, et al. Sparsity invariant CNNs [C]. 2017 International Conference on 3D Vision(3DV), IEEE, 2017: 11-20.
- [23] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes dataset for semantic urban scene understanding [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 2016:3213-3223.
- [24] ZHAO L L, ZHU M L. MS-YOLOv7: YOLOv7 based on multi-scale for object detection on UAV aerial photography[J]. Drones, 2023, 7(3): 188.
- [25] 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), IEEE, 2024: 1-6.
- [26] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [J]. ArXiv preprint arXiv: 2402.13616, 2024.

作者简介

黄艳国,博士,教授,主要研究方向为智能交通系统、智能控制。

E-mail:1915598484@qq.com

彭建(通信作者),硕士研究生,主要研究方向为交通目标检测、图像处理。

E-mail:Acctive1999@163.com

方敏杰,硕士研究生,主要研究方向为图像处理。

E-mail:6120220325@mail.jxust.edu.cn

吴水清,教授级高级工程师,主要研究方向为智能交通系统、智能控制。

E-mail:1599639769@qq.com