

基于 YOLO 算法的非机动车辆检测模型*

王树凤¹ 梁庆伟¹ 王宇航¹ 周倩²

(1. 山东科技大学, 青岛 266590; 2. 比亚迪汽车有限公司, 西安 710119)

【摘要】针对自动驾驶车辆目标检测过程中非机动车因体积小、易被遮挡而导致误检和漏检的问题,为提高非机动车的检测精度,对YOLOv4基础算法进行改进,利用跨阶段连接优化特征提取融合网络,在减少计算量的同时提高检测性能,并嵌入卷积块注意力模块(CBAM),通过通道和空间注意力权重分配来增大有效特征权重、提高检测精度,同时,利用自建的非机动车数据集,在锚框(Anchor)自适应匹配的基础上建立非机动车检测模型。最后,为验证模型的有效性,通过消融实验对比模型性能,结果表明,所提出的检测模型能够有效提高非机动车的检测和识别效果,较好地解决误检和漏检问题。

关键词: 非机动车检测 YOLOv4 算法 卷积块注意力模块 跨阶段连接 消融实验

中图分类号: U471.15 **文献标志码:** A **DOI:** 10.20104/j.cnki.1674-6546.20240223

Non-Motor Vehicle Detection Model Based on YOLO Algorithm

Wang Shufeng¹, Liang Qingwei¹, Wang Yuhang¹, Zhou Qian²

(1. Shandong University of Science and Technology, Qingdao 266590; 2. BYD Auto Co., Ltd., Xi'an 710119)

【Abstract】To address the issue of false and missed detection of non-motorized vehicles due to the small size and obstructed vision in autonomous vehicle target detection, this research refines YOLOv4 basic algorithm to bolster the accuracy of non-motorized vehicle detection. The optimized algorithm streamlines the feature extraction process through a cross-stage connection, concurrently diminishing computational overhead and bolstering detection efficiency. Additionally, Convolutional Block Attention Module (CBAM) is embedded to increase effective feature weights and improve detection accuracy through channel and spatial attention weights. A non-motorized vehicle detection model is established based on anchor adaptive matching using a self-built non-motorized vehicle dataset. To verify the effectiveness of the model, the performance of the model is compared through ablation experiments. The results show that the proposed detection model substantially improves the detection and recognition performance of non-motor vehicles, effectively solve the problems of missed and false detections.

Key words: Non-motor vehicle detection, YOLOv4 algorithm, Convolutional Block Attention Module (CBAM), Cross-stage connection, Ablation experiment

【引用格式】王树凤, 梁庆伟, 王宇航, 等. 基于YOLO算法的非机动车辆检测模型[J]. 汽车工程师, 2024(8): 8-14.

WANG S F, LIANG Q W, WANG Y H, et al. Non-Motor Vehicle Detection Model Based on YOLO Algorithm[J]. Automotive Engineer, 2024(8): 8-14.

1 前言

环境感知是自动驾驶汽车安全行驶的关键技术之一^[1],机动车辆、非机动车、行人是各类场景中的主要交通参与者,实时、准确识别各类目标是实现自动驾驶的前提。受交通场景的复杂性和可变性影响,非机动车辆易被遮挡,从而导致目标检测

算法易出现误检、漏检、鲁棒性较差等问题。

随着深度学习技术的发展,基于深度学习的目标检测算法逐渐代替传统检测方法。基于深度学习的目标检测方法主要分为一步法和两步法。两步法先提取候选区域的图像特征,再对目标进行分类,此类算法精度较高,但速度较慢,其发展经历了区域卷积神经网络(Region-based

*基金项目:山东省研究生优质专业学位教学案例库建设项目(SDYAL2023051)。

Convolutional Neural Network, R-CNN)、快速区域卷积神经网络(Fast R-CNN)、更快速区域卷积神经网络(Faster R-CNN)等阶段^[2]。一步法不需要预先生成候选框,直接在特征图上预测目标位置及其所属类别,检测速度大幅提升。其中,YOLO(You Only Look Once)算法^[3]在检测精度和速度方面较为均衡,成为目前目标检测领域中应用最广泛的模型之一。但是,单阶段算法不能很好地考虑不同层的特征,难以实现对小目标的实时准确检测。为解决这些问题,YOLO改进版本如YOLOv3和YOLOv4等相继提出,不断提升目标检测的性能和精度。袁小平等^[4]使用ResNeXt残差模块替换原有的残差模块,以提升网络特征提取能力,提出了改进YOLOv3算法;Seo等^[5]从重构损失函数的角度对算法进行改进,提高了对低分辨率目标图像的检测效果,降低了漏检率,但鲁棒性不足;Fu等^[6]将注意力机制嵌入目标检测模型,在舰船数据集中的小目标和多目标等的检测方面取得了较好的效果。

随着YOLO算法版本不断升级,在网络架构和性能上进一步优化改进,可提供更加快速和准确的目标检测效果,但结构愈发复杂,对设备的运算能力等要求逐步提高。本文以对设备要求不高的YOLOv4算法为基础,为解决非机动车实时检测时出现误检和漏检的问题,提出一种非机动车检测方法。采用跨阶段连接优化网络进行特征提取融合,提升网络学习能力,减轻计算压力;在特征融合部分嵌入卷积块注意力模块(Convolutional Block Attention Module, CBAM),降低误检率、提高特征识别效率;通过构造非机动车检测数据集,使用

K-均值(K-Means)算法计算合适的锚框(Anchor)参数,对模型进行训练,建立非机动车检测模型,最后利用消融实验验证模型的有效性。

2 YOLOv4基础算法及改进

为提高非机动车检测的精度,本文对YOLOv4算法进行优化,主要流程如图1所示。

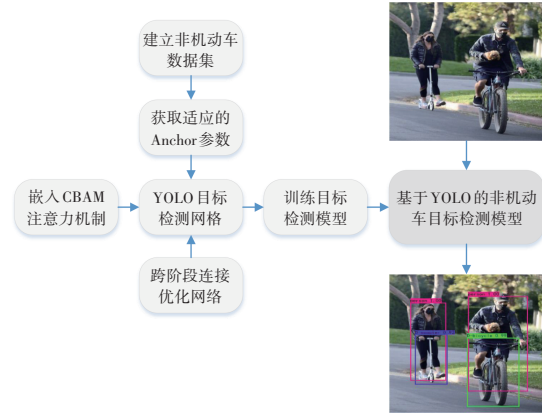


图1 目标检测流程

2.1 YOLOv4基础算法

YOLOv4采用基于卷积神经网络的端到端的架构,在YOLOv3基础上引入路径聚合网络(Path Aggregation Network, PANet)、空间金字塔池化(Spatial Pyramid Pooling, SPP)、马赛克数据增强(Mosaic Data Augmentation)等多种技术,其特点是检测模型高效、泛化能力强,可在普通图形处理单元(Graphics Processing Unit, GPU)上进行快速准确的训练。整体架构可以分为3个主要组件,即主干网络(Backbone)、颈部(Neck)和预测头(Head),如图2所示。

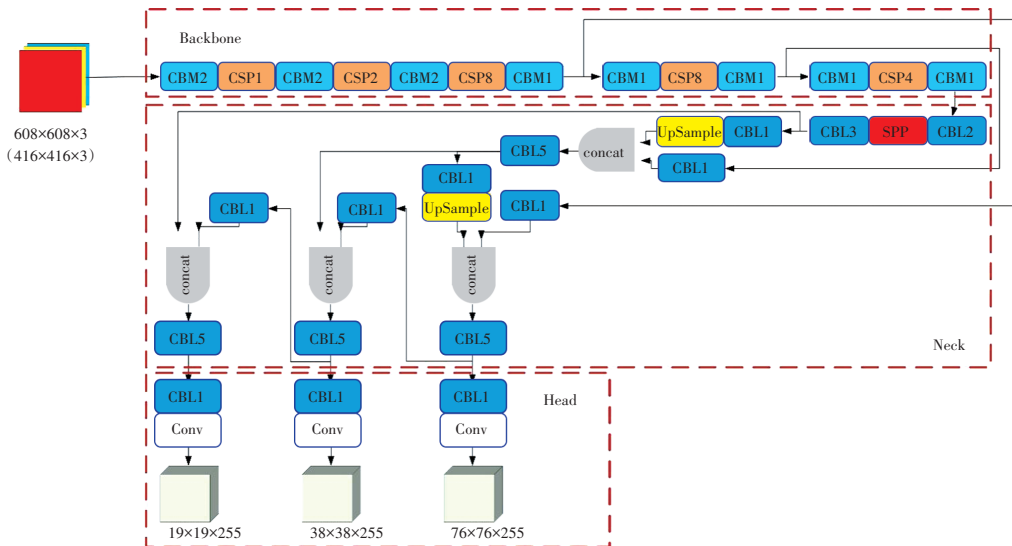


图2 YOLOv4网络结构

主干网络使用 CSPDarknet-53 进行图像的特征提取, CSPDarknet-53 是在 Darknet-53 的基础上,通过堆叠和组织卷积块模块(Convolutional Block Module, CBM),增加跨阶段局部网络(Cross Stage Partial Network, CSPNet)结构,实现图像特征提取。CSPNet 可有效缓解梯度消失或爆炸的问题,减少计算量,同时增强网络学习能力,保证准确率^[7]。

颈部使用结合 SPP 的 PANet 结构进行特征融合,提取多尺度的特征。PANet 通过上采样(UpSample)卷积块(CBL)的结合使用,可以在多个尺度上提取和融合特征,有助于网络捕获不同大小目标的细节,增强网络对小目标的检测能力,其中 CBL 是卷积神经网络中的基础模块,由卷积层、批归一化和 Leaky ReLU 激活函数 3 个部分组成。SPP 则通过在不同大小的池化窗口上应用池化操作来生成多尺度的特征图。两者结合更好地实现了浅层信息与更高层次语义特征的融合,缓解了小目标特征丢失问题,提高了检测的准确性。

预测头用于预测目标的位置和类别。该结构部分遵循 YOLOv3 的检测机制,根据颈部输出的特征图进行预测,通常包括多个卷积层,用于在不同尺度的特征图上生成边界框的预测,使用锚框来预测目标的位置,以及类别概率。

YOLOv4 的整体架构通过主干网络提取特征,利用 SPP 网络提取多尺度的特征,并使用预测头进行目标的位置和类别预测。端到端的设计使得 YOLOv4 在目标检测任务中具有较高的准确性和实时性能。

2.2 跨阶段连接优化网络

在网络推理过程中,特征的反复提取和融合涉及大量计算,为有效减少计算量,提高网络学习能力和检测性能,对 YOLOv4 的 PANet 前端进行拆分,复制特征层并采用跨阶段连接策略(Cross-Stage Connection Strategy)进行融合处理。与 CSPNet 类似,首先对主干网络特征提取部分的输出特征映射进行复制和拆分,本文的输入图像大小为 $416 \times 416 \times 3$,因此该部分的输出特征图为 $13 \times 13 \times 1024$,该特征图分为用于跨阶段连接的 $13 \times 13 \times 512$ 短边卷积特征图和用于维护原始特征融合路径的 $13 \times 13 \times 512$ 主卷积特征图。然后,将主干特征图经过 3×3 和 1×1 卷积运算后输入 SPP 模块,经池化运算,将主干特征图进行 1×1 和 3×3 卷积处理,最后

将输出特征图与短边卷积特征图融合,输出综合特征图。跨阶段连接的整体结构如图 3 所示。原始网络在更新不同层的权值时,有大量的梯度信息复用,而跨阶段连接优化网络通过分离梯度流,既保留原网络中特征重用的优势,又通过截断梯度流防止过多重复的梯度信息,减轻计算压力。

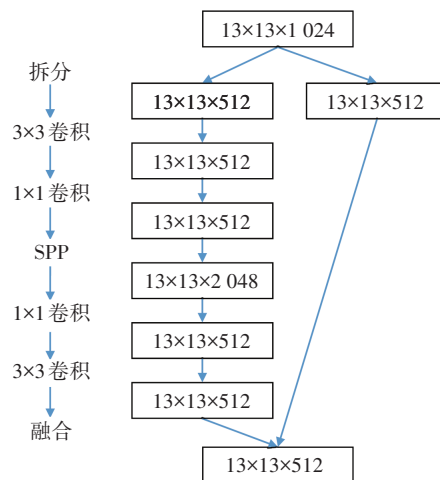


图3 跨阶段连接网络

2.3 嵌入卷积块注意力模块

CBAM 在通道和空间维度上对特征图进行加权,以增强网络对关键信息的捕捉能力,是一个轻量化、通用的模块,可即插即用。CBAM 由通道注意力模块和空间注意力模块串联组成,如图 4 所示。其中,通道注意力模块通过全局平均池化和全局最大池化操作,对输入特征层的每个通道生成权重,权重反映了通道的重要性。空间注意力模块在通道维度上进行操作,生成空间上的权重矩阵,用于强调或抑制特征图中的不同区域。CBAM 的核心是先后应用通道和空间注意力模块,集中提取这两个维度上的关键特征,然后进行两者的融合,可有效提升模型的性能和泛化能力^[8]。

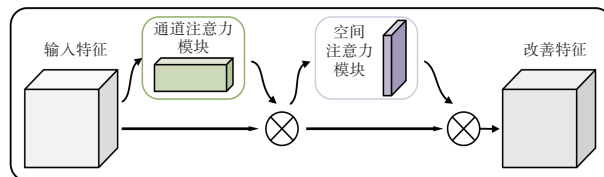


图4 CBAM 示意

为了提高非机动车等交通参与者的检测精度,在颈部与预测头之间嵌入 CBAM,如图 5 所示。针对不同尺度的特征层,通过计算并分配通道和空间注意力权值来增强有效特征的权重,抑制无效信息,从而提高检测性能。

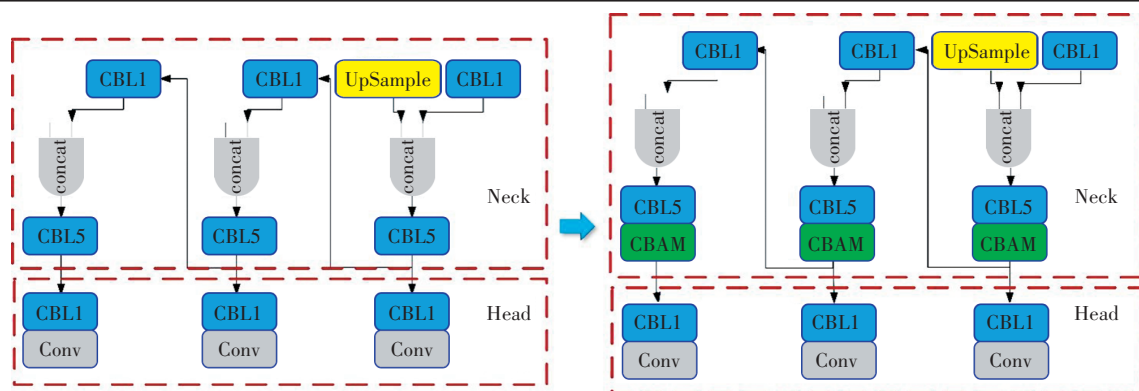


图5 CBAM的嵌入位置

3 非机动车检测模型有效性验证

为了验证改进的模型对非机动车的检测效果,首先在已有图像数据集的基础上,通过摄像头拍摄非机动车图像,建立相应的数据集,然后根据非机动车的特点,使用K-Means聚类算法对锚框参数进行匹配,根据评价指标,对改进模型进行消融实验,对比分析不同算法模型的效果。算法模型运行全部在配置英特尔至强 E5-2609 处理器、英伟达 Quadro M2000 和 64 位 Windows 10 操作系统的环境下,基于 TensorFlow 框架使用 Python 语言完成了模型训练与测试,除前文提到的图像大小等参数外,其他模型参数采用默认配置。

3.1 数据集的制作

为实现对非机动车的有效检测,本文利用摄像头等设备结合基于网络的图像资源,构建了一个以非机动车辆为主的新数据集,重点采集了交叉口、人行道等复杂的交通场景,该数据集共有 1 600 张图像,包含 7 000 余个检测对象。考虑到现有交

通工具的多样性,本文将平衡车添加到测试目标中。与公共交通数据集相比,自行车、摩托车、平衡车、三轮车等数量较多,形态差异大,对应骑行者姿态也有所不同,因此本文对数据集进行归类,如表 1 所示,采集的图像如图 6 所示。此外,非机动车驾驶员暴露在环境中时与行人危险性相似,均为弱势道路交通参与者,因此,本文将非机动驾驶员和行人标记为同一类别。然后使用 LabelImg 标注工具对各种类型的样本进行标注。

表 1 数据集样本类型和数量

类别	样本数量/个
单轮平衡车	65
两轮平衡车	106
摩托车	169
三轮车	270
自行车	756
行人	3 237
汽车	2 636



图6 数据集中的图像样本示例

3.2 基于K-Means聚类的锚框参数匹配

YOLO 模型针对目标检测的初始锚框参数由对包含 80 个类别的公共数据集聚类获得,而本文主要以非机动车为目标,其类别特征与公共数据集中目

标的大小、形状和位置不同,其默认锚框参数无法很好地适应非机动车数据集。为提升检测精度,本文利用 K-Means 聚类算法对目标进行自适应锚框参数计算,结果如表 2 所示。

表2 锚框参数调整结果

感受野	默认锚框参数	重新计算的锚框参数
大	(116,90)(156,198) (373,326)	(98,103)(121,245) (289,300)
中	(30,61)(62,45)(59,119)	(33,105)(53,52)(62,186)
小	(10,13)(16,30)(33,23)	(9,27)(18,61)(24,25)

从表2可以看出,重新计算的锚框参数与默认锚框参数差异较大,本文在新锚框参数的基础上,随机选取90%的图像作为训练集,其余10%作为测试集,用于训练和测试改进方案的检测效果。

3.3 评价指标

模型评价指标常用精确率(Precision)、召回率(Recall)、平均精度(Average Precision, AP)、平均精度均值(mean Average Precision, mAP)来评估检测模型的性能。

精确率 P 和召回率 R 的计算公式分别为:

$$P=N_{TP}/(N_{TP}+N_{FP}) \quad (1)$$

$$R=N_{TP}/(N_{TP}+N_{FN}) \quad (2)$$

式中: N_{TP} 、 N_{FP} 、 N_{FN} 分别为真正例、假正例、假负例的

数量。

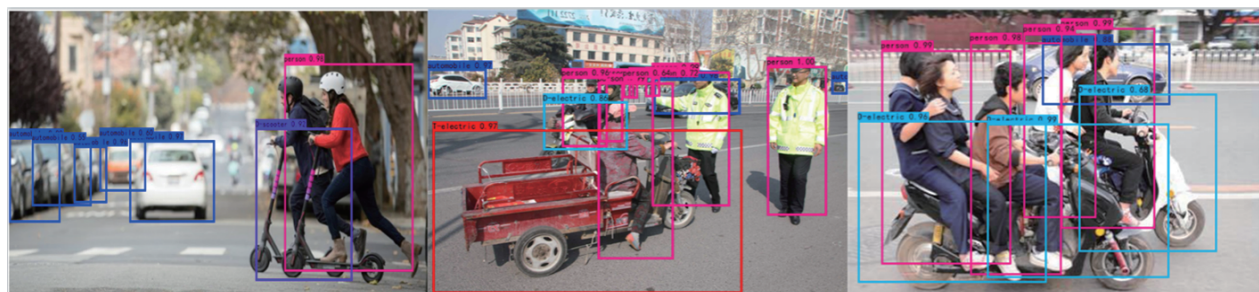
mAP是评价模型整体性能的重要指标,其表示检测模型中所有类别的AP的平均值,数值越大,表示识别准确率越高,计算公式为:

$$P_{mAP} = \frac{1}{m} \sum_{q \in m} P_A(q) \quad (3)$$

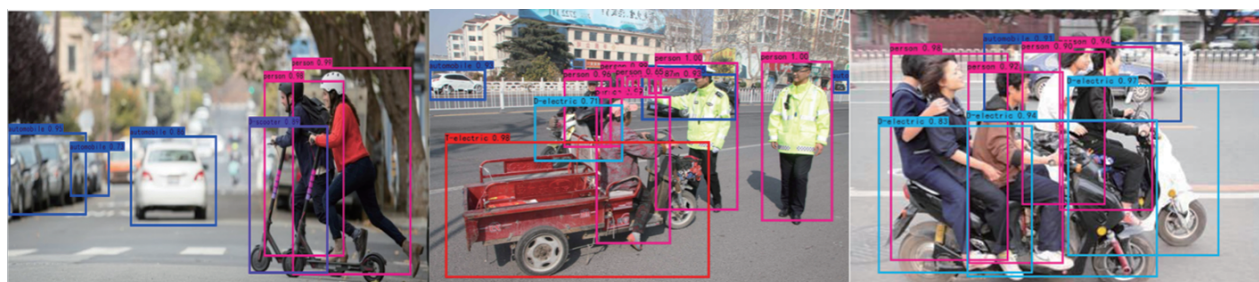
式中: m 为类别数量, q 为类别序号。

3.4 消融实验对比分析

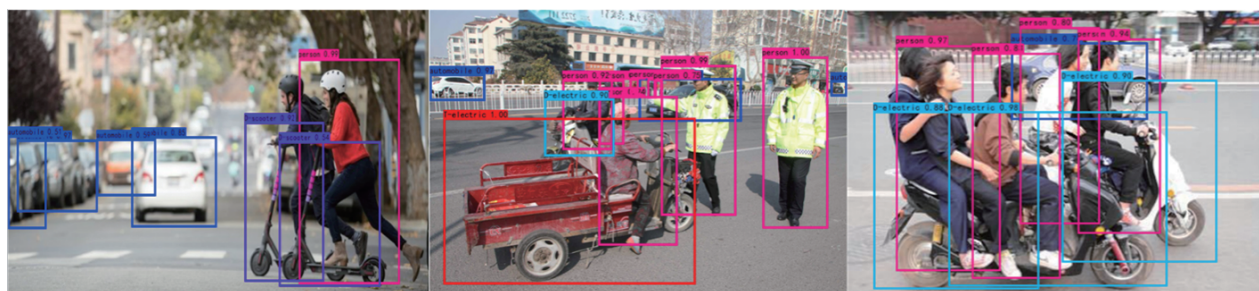
使用数据集训练目标检测模型,并在测试集上进行验证,采用消融实验对比模型性能。首先以YOLOv4原始模型为基线,在原始模型基础上增加跨阶段连接优化网络的模型记为YOLOv4-b,在原始模型基础上增加CBAM的模型记为YOLOv4-c,在原始模型基础上同时增加CBAM和跨阶段连接优化网络的模型记为YOLOv4-d。通过对比分析,考察跨阶段连接优化网络和CBAM对模型性能的影响。图7所示为目标检测效果对比,不同模型检测时平均精度及各类别平均的精确率、召回率、mAP的对比结果分别如表3、表4所示。



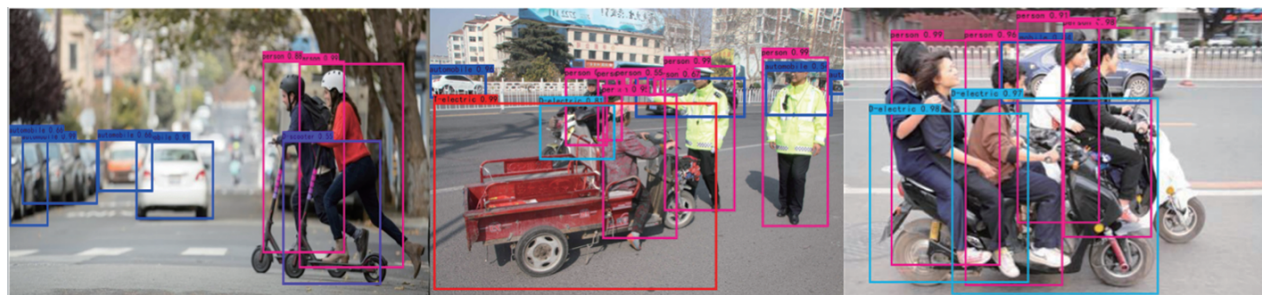
(a)YOLOv4



(b)YOLOv4-b



(c)YOLOv4-c



(d)YOLOv4-d

图7 不同模型的检测效果

表3 不同模型对应不同类别的平均精度

类别	平均精度			
	YOLOv4	YOLOv4-b	YOLOv4-c	YOLOv4-d
单轮平衡车	0.78	0.76	0.82	0.91
两轮平衡车	0.85	0.88	0.89	0.93
摩托车	0.89	0.91	0.93	0.94
三轮车	0.80	0.95	0.96	0.96
自行车	0.94	0.95	0.95	0.95
行人	0.91	0.93	0.93	0.94
汽车	0.90	0.88	0.89	0.89

表4 不同模型的精确率和召回率和mAP %

类别	YOLOv4	YOLOv4-b	YOLOv4-c	YOLOv4-d
精确率	82.67	90.22	90.79	93.39
召回率	81.38	82.07	81.97	82.02
mAP	87.05	91.27	91.31	93.26

由表3数据分析可知,跨阶段连接优化网络、CBAM均减少了模型对非机动的误检和漏检,提高了检测的精度,其中,CBAM的效果更好。具体来说,对于非机动车,平衡车、摩托车、三轮车的检测精度有较大幅度提升,但是对汽车的检测精度不及行人和自行车的检测精度,而且在改进模型中,其检测精度反而略有下降,主要原因在于自制的数据集中图片以交叉口、人行道等角度拍摄为主,汽车图像样本效果较差导致其检测精度不高(见图6)。另外,针对非机动车设置的锚框参数也不利于汽车的检测。由表4可知,跨阶段连接优化网络模型、CBAM模型以及组合模型均有效提高了模型检测的精确率和召回率,mAP分别提高了4.22百分点、4.26百分点和6.21百分点,改进模型的mAP可达93.26%。

跨阶段连接优化网络可以帮助模型更好地融合特征,既保留了骨干网中特征图的语义信息,又

融合了更深层次的特征信息,提高了检测的召回率;而CBAM可以更好地关注目标信息,提高了目标检测的精度;组合后的改进模型融合了以上两种模型的优势,能更准确地检测到目标。

4 结束语

本文以YOLOv4算法为基础,针对非机动车检测,增加跨阶段连接优化网络和CBAM对算法进行了改进,并建立自制的非机动车数据集,使用K-Means聚类算法获得合适的锚框参数,在此基础上建立非机动车识别模型,最后利用消融实验验证改进模型的有效性,结果表明,加入跨阶段连接优化网络和CBAM的非机动车检测模型的精确率和召回率均相应提升,mAP可达93.26%,验证了改进模型的有效性。

参考文献

- [1] ZHU J, LI H Y ZHANG T. Camera, LiDAR, and IMU Based Multi-Sensor Fusion SLAM: A Survey[J]. Tsinghua Science and Technology, 2024, 29(2): 415-429.
- [2] 金沙沙, 龙伟, 胡灵犀, 等. 多目标检测与跟踪算法在智能交通监控系统中的研究进展[J]. 控制与决策, 2023, 38(4): 890-901.
JIN S S, LONG W, HU L X, et al. Research Progress of Detection and Multi-Object Tracking Algorithm in Intelligent Traffic Monitoring System[J]. Control and Decision, 2023, 38(4): 890-901.
- [3] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的YOLO目标检测综述[J]. 电子与信息学报, 2022, 44(10): 3697-3708.
SHAO Y H, ZHANG D, CHU H Y, et al. YOLO Object Detection Based on Deep Learning[J]. Journal of Electronics and Information Technology, 2022, 44(10): 3697-3708.
- [4] 袁小平, 马绪起, 刘赛. 改进YOLOv3的行人车辆目标检测算法[J]. 科学技术与工程, 2021, 21(8): 3192-3198.
YUAN X P, MA X Q, LIN S. An Improved Algorithm of

- Pedestrian and Vehicle Detection Based on YOLOv3[J]. Science Technology and Engineering, 2021, 21(8): 3192-3198.
- [5] SEO J, PARK H. Object Recognition in Very Low Resolution Images Using Deep Collaborative Learning[J]. IEEE Access, 2019, 7: 134071-134082.
- [6] FU H X, SONG G Q, WANG Y C. Improved YOLOv4 Marine Target Detection Combined with CBAM[J]. Symmetry, 2021, 13(4): 623.
- [7] WANG C E, LIAO H Y M, WU Y H, et al. CSPNet: A New Backbone that Can Enhance Learning Capability of CNN [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, WA, USA: IEEE, 2020.
- [8] WANG Y M, ZHANG Z, FENG L W, et al. A New Attention-Based CNN Approach for Crop Mapping Using Time Series Sentinel-2 Images[J]. Computers and Electronics in Agriculture, 2021, 184.
- (责任编辑 斛 畔)
- 修改稿收到日期为2024年7月11日。

《汽车工程师》自动驾驶汽车测试技术专刊征稿通知

随着人工智能、无线通信、高性能计算等新一代信息技术的不断发展与加速应用,自动驾驶汽车已成为当前全球汽车与交通领域深度融合发展的重要载体,当前全球范围内均在加速推进自动驾驶关键技术研究及产业化工作。科学、完善的测试评价是保证自动驾驶汽车整车与零部件的安全性、可靠性、稳定性的关键环节,是自动驾驶技术规模化落地应用的必要前提。然而,由于自动驾驶已成为“人-车-环境-任务”强耦合的复杂系统,其测试评价的难度与复杂度显著增加,测试成本高、周期长、覆盖度低、可信性不足、缺乏完善工具链等已成为自动驾驶汽车测试领域面临的关键难题。

为此,《汽车工程师》特推出“自动驾驶汽车测试技术”专刊,并邀请长安大学王润民老师担任客座主编,欢迎全国高校、科研院所和企事业单位研究团队踊跃投稿。

一、征稿主题(包括但不限于以下主题)

1. 自动驾驶测试评价理论与方法
2. 自动驾驶虚拟仿真测试技术
3. 自动驾驶 HiL/ViL 测试技术
4. 自动驾驶场地及道路测试技术
5. 基于数字孪生的自动驾驶测试技术
6. 车联网与车路协同测评技术
7. 自动驾驶功能安全与预期功能安全
8. 自动驾驶测试装备

二、投稿须知

1. 请登录 www.tjqc.cbpt.cnki.net 在线投稿。
2. 论文具体要求见本刊主页“下载中心”栏目《作者指南》,投稿时请在系统内稿件基本信息的中文标题前增加“【自动驾驶汽车测试技术专刊】”字样。
3. 投稿截止日期为2024年12月15日。
4. 论文录用并完成编辑加工后即在中国知网进行网络首发,纸质期刊预计出版时间为2025年5月15日。

三、特邀客座主编



王润民 高级工程师

工学博士,硕士研究生导师,长安大学智能网联汽车测试实验室主任,《汽车工程师》青年编委会委员。

长期从事智能车路协同关键技术及应用、智能网联汽车测试评价技术等领域的研究工作。先后主持国家自然科学基金重点项目课题、陕西省“两链”融合新能源与智能汽车重点专项、浙江省重点研发计划项目等各类纵向、横向科研课题10项,参与国家重点研发计划项目、国家自然科学基金项目等各类国家级、省部级项目20余项。曾获得国家级教学成果奖一等奖1项,省部级科技奖励二等奖1项。

《汽车工程师》编辑部

2024年7月