

基于改进 YOLOv8-DeepSort 景区人流量监测算法研究

宋江民¹
SONG Jiangmin

摘要

针对景区人流量监测中精度不足和实时性差的问题,文章提出了一种基于改进 YOLOv8-DeepSort 的监测算法。首先,引入基于注意力的多尺度特征融合(AMSF),增强对小目标和密集人群的检测精度;其次,通过引入 Ghost Conv 模块进行轻量化设计,有效降低了计算量和内存消耗,从而显著提升了算法的实时性。实验表明,改进后的 YOLOv8-DeepSort 算法在复杂场景中表现出更高的鲁棒性和实时性。该算法为景区人流量监测提供了一种高效、准确的解决方案。

关键词

景区人流量监测; YOLOv8-DeepSort; 注意力机制; Ghost Conv

doi: 10.3969/j.issn.1672-9528.2025.08.014

0 引言

近几年来,随着国民经济的发展,人们的生活水平也随之不断提高,相较于物质追求,人们开始趋向于精神追求,因此旅游业逐渐成为全球经济的重要组成部分,特别是景区在旅游业中扮演着至关重要的“角色”,吸引着大量游客的光临^[1]。景区人流量的有效监测不仅对于景区管理者优化游客流动、提升游客体验具有重要意义,而且对于保障游客安全、应对突发事件、合理规划景区资源的分配也具有至关重要的作用。尤其在大规模游客集中的景区中,如何高效、精确地监测人流密度和流动趋势,已成为亟待解决的关键问题^[2]。

传统的人流量监测方法大多依赖于人工统计、安保人员巡逻或静态的传感器设备,这些方法在实际应用中面临着实时性差、监测范围有限以及数据分析处理能力不足等问题。随着人工智能技术的发展,计算机视觉在智能监控领域中的应用逐渐成为解决这一问题的有效途径。基于深度学习的目标检测和跟踪算法,特别是 YOLO 一系列算法,其凭借高效的目标检测性能和出色的实时性优势,已经成为人流量监测任务中的主流技术之一^[3]。然而,尽管 YOLO 算法在物体检测方面取得了显著进展,但应该怎样进一步提升其在复杂景区环境中的表现,尤其是在景区这种高密度人群环境下的检测精度,依旧是一个重大挑战。

DeepSort 算法作为一种基于卡尔曼滤波的目标跟踪方法,能够在 YOLO 目标检测的基础上,实现对人流的精准跟踪和高效计数。但现有的 YOLOv8 和 DeepSort 算法在复杂场景下的结合,尤其是在景区这种动态、多变的环境中,仍然面临着一些问题,如遮挡、目标之间的相互干扰以及环境光照的变化等,限制了其在实际应用中的准确性和鲁棒性^[4]。

为此,本文提出了一种基于改进的 YOLOv8-DeepSort 的景区人流量监测算法。针对 YOLOv8 的检测精度和实时性问题,结合多尺度特征融合、注意力机制等技术以及 Ghost Conv^[5] 模块对其进行优化,在有效地减少了模型的计算量和参数的同时还提升了模型在复杂场景中的检测效果。通过实验验证,改进后的算法能够在景区人流量监测任务中实现更高的检测准确率和更好的跟踪效果,大力支持了景区的智能化管理的同时,也为游客的安全提供了有效保障。

本文的研究不仅为人流量监测技术提供了新的解决思路,也为基于深度学习的智能视频分析技术在实际场景中的应用提供了宝贵经验。

1 YOLOv8-DeepSort 算法

YOLOv8-DeepSort 算法融合了 YOLOv8 目标检测与 DeepSort 多目标跟踪两种算法的技术优势,在人流量统计、视频监控等领域得到了广泛应用。YOLOv8 作为目标检测的核心模块,在检测精度和运行效率上相比传统方法都有显著提升,能够快速识别图像中的多个目标并准确标出其位置。DeepSort 作为多目标跟踪模块,在 SORT 算法的基础上,结合卡尔曼滤波和匈牙利算法,有效解决了目标遮挡、轨迹重叠等复杂场景下的跟踪问题,显著提升了模型的跟踪和抗干扰能力^[6]。在本文的实验中,将 YOLOv8 和 DeepSort 这两种算法结合起来使用,YOLOv8 负责实时检测目标并提供位置信息,DeepSort 则利用这些信息对目标进行持续跟踪,确保目标身份的连续性和运动轨迹的准确性。这种组合不仅能够满足实时视频分析对检测速度的高要求,还能在复杂场景下提供稳定且可靠的跟踪效果,因此,本文所使用的算法在景区智能管理、安防监控和智能交通等复杂场景中具有广阔的应用前景。

1. 西安石油大学 陕西西安 710065

2 YOLOv8-DeepSort 目标检测网络改进

2.1 引入基于注意力的多尺度特征融合技术

在目标检测任务中，尤其是在景区这种人员密集且环境复杂的场景下，传统的 YOLOv8 算法在处理不同尺度的目标时存在着一定的局限性，特别是对于小目标的检测，YOLOv8 在这方面的精度严重不足。为了解决这一问题，本文实验引入了基于注意力的多尺度特征融合技术（AMSF），旨在提高 YOLOv8 对各种尺度目标的检测能力，尤其是在景区这类密集人群场景下对小目标进行检测的表现。多尺度特征融合的关键在于将不同尺度的特征图进行整合，从而使得模型能够在各个层次上充分捕捉图像中的细节信息，YOLOv8 的低层特征能够更好地捕捉细节和小目标，而高层特征则侧重于较大的目标及其语义信息。通过将这些信息进行有效融合，可以使得模型在保持大目标高精度检测的同时，提升对小目标的检测能力，从而来解决小目标漏检错检的问题。

如图 1 所示，其 Neck 部分使用注意机制进行多尺度特征融合，而不是传统检测器中的简单级联操作。

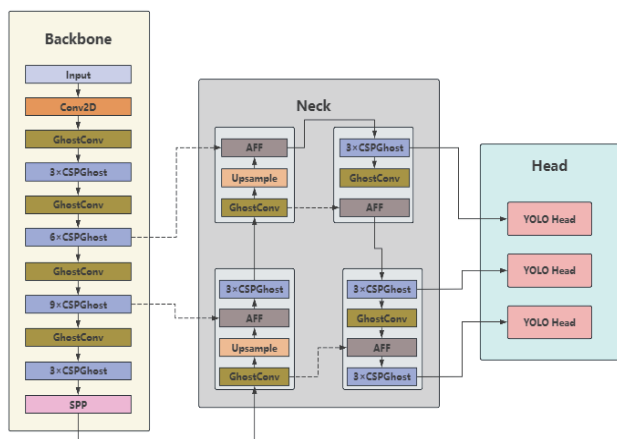


图 1 YOLOv8 网络拓扑结构

引入基于注意力的多尺度特征融合后，YOLOv8^[7] 充分发挥了不同层次特征的优势，大幅提升了在密集环境下的检测效果。通过融合低层和高层的特征信息，算法不仅能捕捉到小目标，而且还能通过高层的语义信息判断目标的类别和位置，从而在更加复杂的场景中提供更准确的检测结果。尤其在人员密集的景区环境中，这一改进使得 YOLOv8 能够精确地区分出相互接近的各个目标，而且还减少由于遮挡或重叠而导致的误判和漏检问题。改进后的模型能够更好地处理景区人群中的小目标，在保证大目标检测精度的同时，提升了算法对小目标的检测能力，使得整体检测结果更加的稳定和准确。

2.2 引入 Ghost Conv

在景区人流量监测任务中，所选择的目标检测算法不仅需要具备高精度，还必须能够在实时性和计算资源有限的

情况下高效运行。为此，本文实验采用了 Ghost Convolution（Ghost 卷积）作为一种优化卷积操作的方法，以此来提升 YOLOv8 模型的性能，特别是在计算量和内存消耗方面的优化。Ghost Conv 的引入使得 YOLOv8 算法能够在保持检测精度的同时，显著提高处理效率，满足景区这种高动态、复杂环境下的实时监控需求^[8]。

Ghost Conv 的基本思想是通过生成“Ghost 特征”来替代传统卷积操作中部分冗余的计算，从而大幅度减少计算量并降低内存消耗。传统卷积通常通过对输入的每个像素执行卷积运算，生成多个特征图，这一过程在网络层数较多时往往会带来极大的计算压力。与此不同的是，Ghost Conv 通过生成一小部分主特征，并利用简单的线性操作生成“Ghost 特征”，这些 Ghost 特征能够有效补充主特征，保留足够的图像信息，而且计算成本远低于传统卷积的计算成本^[9]。因此，在此项目中，Ghost Conv 帮助 YOLOv8 在进行目标检测时大幅减少了计算负担和内存占用，提升了处理速度，同时保证了精度。

如图 2 所示，（a）为传统卷积，（b）为 Ghost 卷积。与传统卷积相比，Ghost Conv 的优势主要体现在计算复杂度、内存消耗和实时处理能力方面。传统卷积需要执行大量的计算，特别是当卷积核数量增加时，计算量呈指数级增长^[10]。即使在一些较小的网络架构中，积层的计算量和内存需求也会随着图像分辨率的提高而急剧增加，尤其是在景区这样的大范围监控中，计算成本会成为瓶颈。Ghost Conv 通过仅生成少量的主特征和 Ghost 特征，显著降低了计算量，使得 YOLOv8 在处理高分辨率图像时能够保持较低的计算负担，从而提高了整体运行速度。

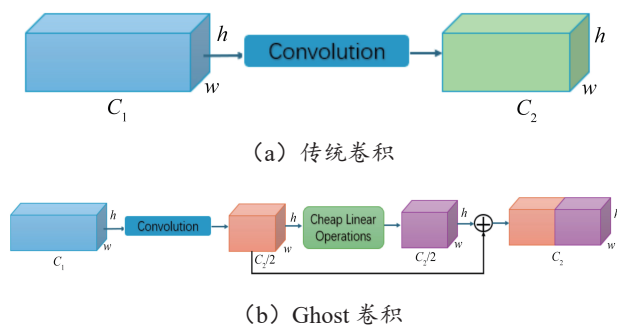


图 2 传统卷积与 Ghost 卷积对比图

在卷积神经网络中，标准卷积操作与 Ghost 卷积在计算复杂度和参数规模上呈现出显著区别。以输入张量维度为例，假设输入特征图的通道数为 c ，空间维度为 $h \times w$ ，经过初始卷积变换后，输出特征图的维度将发生变化。假设卷积核尺寸为 d ，经过 s 次特征变换后，对传统卷积运算与 Ghost 卷积运算的计算量进行对比分析，用公式表示为：

$$r_s = \frac{n \times h' \times w' \times c \times k \times k}{\frac{n}{s} \times h' \times w' \times c \times k \times k + (s-1) \times \frac{n}{s} \times h' \times w' \times d \times d} \approx s \quad (1)$$

传统卷积和 Ghost Conv 的计算量对比:

$$r_c = \frac{n \times c \times k \times k}{\frac{n}{s} \times c \times k \times k + (s-1) \times \frac{n}{s} \times d \times d} \approx s \quad (2)$$

3 实验及结果分析

3.1 构建景区人流检测数据集

为全面构建一个高质量的景区人流检测数据集, 本文从多种数据源中采集并整合了一个大约 3 000 张的包含丰富场景信息的数据集, 涵盖了景区入口、热门景点、狭窄通道以及开阔广场等关键区域。数据集中包含了不同时间段(如工作日、周末、节假日)以及多种天气条件(如晴天、雨天、雾天)下的场景, 使得数据更具有多样性。为确保数据质量, 使用专业的图像标注工具对每张图片或视频帧中的人流进行了精准标注, 包括边界框(Bounding Box)、密度图(Density Map)以及移动轨迹信息。在数据划分上, 按照 7:2:1 的比例将数据集划分为训练集、验证集和测试集, 训练集用于模型的训练, 验证集用于参数调优, 而测试集则用于最后的性能评估^[11]。这一划分方式既保证了模型训练的充分性, 又增强了实验结果的科学性和可靠性, 为评估人流检测算法的实际效果奠定了坚实基础。除此之外, 本实验数据集还通过数据增强技术(如旋转、缩放、翻转等)对数据进一步扩充, 以此来提高模型的泛化能力。

3.2 实验环境

在本文实验中, 所采用的操作系统为 Windows10 专业版。使用 PyTorch 深度学习框架, 版本号为 2.5.1, 且为 GPU 版本。实验的编译语言为 Python 3.9.21。CUDA 版本是 12.4。CPU 为 12 代英特尔酷睿 i5-12450H, 主频 2.00 GHz。运行内存为 16 GB。GPU 是 NVIDIA GeForce RTX 4060, 显存为 16 GB。

除所需的实验环境外, 本实验还设置了大量的超参数。具体而言, 实验中将输入模型的图片尺寸像素统一调整为 640 px×640 px, 以确保数据的一致性。此外, 每次实验的训练周期(Epoch)设定为 200 次, 批量大小(batch-size)设置为 32。在训练的初始阶段, 学习率(lr0)被设定为 0.01, 这一参数用于控制权重更新的步长, 从而确保模型训练的稳定性。在训练中后期, 避免在最优解附近出现震荡, 本实验采用了学习速率衰减策略, 从而提高模型的训练效率和最终性能。

3.3 评价指标

在本文实验中, 为了更好地评估改进 YOLOv8-DeepSort 的目标检测性能, 因此采用了 F_1 分数作为主要的评价指标。 F_1 分数是目标检测领域中一种广泛使用的评估指标, 综合平衡了模型的精确率(Precision)和召回率(Recall)。其计算

过程主要包括以下几个步骤:

(1) 精确度和召回率^[12] 计算: 精确度是指模型正确检测到的目标占所有检测到的目标的比例, 而召回率是指模型正确检测到的目标占所有真实目标的比例。

(2) F_1 分数计算: 通过调和平均精确度和召回率, 得到 F_1 分数。 F_1 分数的计算公式为:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

该分数能够平衡精确度和召回率, 避免因其中一项过高而掩盖另一项的不足。

(3) 平均 F_1 分数^[13] 计算: 计算出所有类别的 F_1 分数后, 再对所有类别的 F_1 分数取均值, 得到平均 F_1 分数(mF_1)。 mF_1 值越高, 表示模型在所有类别上的整体检测性能越优。本文实验只考虑 people 这一种类别, 因此使用 F_1 分数作为平均 F_1 分数(mF_1)。

通过使用 F_1 分数作为评价指标, 本实验能够更直观地反映模型在目标检测任务中的综合表现, 同时兼顾了精确度和召回率的平衡。

3.4 消融实验结果与分析

本文以 YOLOv8-DeepSort 为基准, 依次添加本文所提出的改进措施, 进行一系列消融实验, 添加各个措施后具体的实验结果如表 1 所示。

表 1 消融实验结果

方法	AM SF	Ghost Conv	P/%	R/%	F_1 /%	Params /10 ⁶	FLOPs /10 ⁹
base			56.4	77.3	65.2	6.1	13.2
A	√		62.4	82.2	70.9	6.5	13.7
B		√	58.7	77.9	66.95	5.3	9.9
C	√	√	62.4	82.8	71.17	5.5	10.4

在对 YOLOv8-DeepSort 模型进行改进时, 方法 A 添加了 AMSF 模块, 显著提升了模型的性能, Precision、Recall 和 F_1 指标分别提升了 6%、4.9% 和 5.7%, 虽然添加 AMSF 后的模型的参数量和计算量有所增加, 但模型的整体性能得到了显著提升。方法 B 通过引入 Ghost Conv 模块, 优化了模型性能的同时计算量有所减少, Precision、Recall 和 F_1 指标分别提升了 2.3%、0.6% 和 1.75%。方法 C 同时添加了 AMSF 和 Ghost Conv 模块, 进一步提升了模型精度, Precision、Recall 和 F_1 指标分别提升了 6%、5.5% 和 5.97%。虽然参数和计算量有所增加, 但该方法在精度和效率上达到了最佳平衡, 整体性能表现优异。

3.5 对比实验分析

为验证 YOLOv8-DeepSort 算法相比于当前一些主流算法的优越性, 将 YOLOv8-DeepSort 算法与当前的一些主流算法进行比较实验结果如表 2 所示。

表 2 对比实验结果

方法	Precision /%	Recall /%	F_1 /%	Params /10 ⁶	FLOPs /10 ⁹
YOLOv3-tiny	56.8	78.6	66.1	8.7	5.9
YOLOv4s	57.5	79.7	66.7	6.3	14.5
YOLOv5s	58.8	80.7	68.1	6.7	16.5
YOLOv6s	61.2	81.5	70.1	5.6	27.7
US	62.4	82.3	71.0	5.4	11.3

从表 2 的实验对比结果可以看出,本文提出的改进后的 YOLOv8-DeepSort 算法在检测精度和计算效率两个方面均展现出显著优势。在性能指标方面,本文实验的算法在 Precision、Recall 和 F_1 分数在 3 个关键指标上分别达到了 62.4%、82.3% 和 71.0%,相较于当前主流的 YOLOv5s (58.8%、80.7%、68.1%) 和 YOLOv6s (61.2%、81.5%、70.1%) 均有不同程度的提升。特别是在 Precision 和 F_1 分数上,改进后的 YOLOv8-DeepSort 算法分别比 YOLOv5s 高出 3.6 个百分点和 2.9 个百分点,比 YOLOv6s 高 1.2 个百分点和 0.9 个百分点,充分体现了其在目标检测任务中的高精度优势。在计算效率方面,本文算法的参数量仅为 5.4×10^6 ,明显低于 YOLOv5s (6.7×10^6) 和 YOLOv6s (5.6×10^6),同时其 FLOPs 为 11.3×10^9 ,相较于 YOLOv5s (16.5×10^9) 和 YOLOv6s (27.7×10^9) 分别减少了 31.5% 和 59.2%。改进后的 YOLOv8-DeepSort 算法在参数量和计算量上都展现出了显著的优势,进一步凸显了其轻量化设计的优越性。

总而言之,改进后的 YOLOv8-DeepSort 算法不仅在对目标的检测精度上取得了大幅度的提升,而且还在计算效率上取得了显著突破,实现了高精度与高效率的统一^[14]。

4 检测效果

图 3、图 4 展示了本文提出的算法在实际应用场景中的检测结果,图 3 为原始图片,图 4 则是改进后的 YOLOv8-DeepSort 算法在该图片上的检测结果。从图 4 中可以看出,改进后的 YOLOv8-DeepSort 算法在实际应用上不仅能够准确地定位大目标,还显著提高了对小目标的识别准确率。这一改进使得模型在实际应用中的可靠性得到了有效提升,尤其是在复杂场景下的目标检测任务中,改进后的模型表现尤为突出。



图 3 原始图片

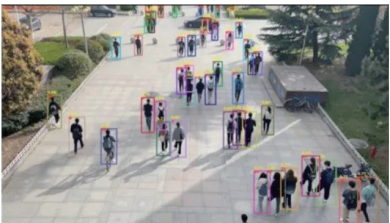


图 4 改进后 YOLOv8-DeepSort

5 结语

景区人流量监测是现代智慧旅游管理和安全保障的重要组成部分,对于提升游客体验、保障景区安全以及实现可持续发展具有重要意义。随着旅游行业的蓬勃发展,景区内的人流量也随之持续增长,尤其是在节假日和旅游旺季,景区内的人流量往往会急剧增加,这就导致了景区的管理和服务面临着巨大的压力。传统的检测方法效率较低,且难以提供实时数据支持,无法满足现代景区管理的需求。在继承现有技术的基础上,本文对数据采集与处理模块进行了优化升级,并引入先进的计算机视觉技术与深度学习算法,从而大幅提高了人流检测的精度和实时性。

尽管改进后的人流检测系统在准确性和实时性方面取得了显著进展,但在复杂场景下的检测精度和鲁棒性仍有进一步提升的空间。未来,将继续优化该系统,探索多模态数据融合技术,整合视频监控、移动传感器等多种数据源,以更全面地反映人流动态,提升复杂场景下的检测精度。同时,也将持续关注深度学习领域的最新进展,引入更高效的网络架构和算法,进一步提高系统的处理速度和资源利用效率。

参考文献:

[1] 毛坤,朱学军,赖惠鸽,等.SGW-YOLO:多尺度融合优化的PCB缺陷检测算法[J/OL].激光与光电子学进展,1-21[2025-05-30].<http://kns.cnki.net/kcms/detail/31.1690.tn.20250212.1723.004.html>.

[2] 张轩宇,周思航,黄健,等.基于高阶空间特征提取的无人机航拍小目标检测算法[J].计算机工程与应用,2025,61(12):210-221.

[3] 艾峰,邓耀华.用于缺陷检测的YOLOv8轻量化设计方法[J].电子测量技术,2025,48(4):181-190.

[4] 许景科,索祥龙,周磊.改进YOLOv8的无人机航拍小目标检测算法[J].计算机工程与应用,2025,61(11):119-131.

[5] 盛伟,刘明剑,刘殿臣.面向人员密集与遮挡的实时目标检测方法[J/OL].计算机工程与科学,1-10[2025-01-30].<http://kns.cnki.net/kcms/detail/43.1258.tp.20250103.0853.002.html>.

[6] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]// Proceedings of the IEEE international conference on computer vision. Piscataway: IEEE, 2017: 2961-2969.

基于 CNN-LSTM 的欠驱动采摘机械手抓取力控制方法

胡朝晖¹ 董艳珂¹
HU Zhaohui DONG Yanke

摘要

现有研究因未充分考虑目标特性与动态交互作用,易导致抓取失衡、滑脱和果实损伤等问题。为此,文章提出一种基于 CNN-LSTM 的抓取力控制方法。利用 CNN 提取目标图像特征,解决传统方法对水果特性建模不足的问题;通过 LSTM 捕捉接触力、力矩与扭矩的长期动态平衡关系,优化静力学模型预测能力;结合非线性反馈与逆动力学前馈补偿,提升欠驱动系统的控制鲁棒性。实验结果表明,所提方法在稳定性、角偏差、滑落次数及果实损伤方面均显著优于对比方法,验证了其在高动态采摘场景中的有效性。表明 CNN-LSTM 模型通过多模态数据融合与动态关系学习,为欠驱动机械手的精准抓取控制提供了新思路,兼具理论价值与工程应用潜力。

关键词

CNN-LSTM; 抓取力控制; 长期依赖特征; 非线性补偿项; 欠驱动

doi: 10.3969/j.issn.1672-9528.2025.08.015

0 引言

采摘机械手作为自动化设备的核心,在农业等领域广泛应用,可提升效率、降低成本并减少人工误差。控制力是其性能的关键因素,优化抓取力控制方法能显著提高采摘精度和效率,降低生产成本,改善农产品质量^[1]。不仅可以降低农业生产的成本,提高农作物的质量和产量,还可以为其他

领域的自动化设备提供一种新的控制方法和技术支持,这对于提高我国在机器人技术领域的国际竞争力具有重要意义。

梁栋等人^[2]利用闭环矢量方程建立机械手速度映射模型,并基于拉格朗日方程构建逆刚体动力学模型,设计 PD 补偿力矩控制方法,但未充分考虑复杂形状物体的交互作用,导致抓取稳定性不足。杨明博等人^[3]基于拉格朗日方程建立运动模型,设计自适应滑膜函数调整连杆参数,并通过李雅普诺夫稳定性算法优化抓取力控制,但仍难以完全适应复杂

1. 郑州工商学院信息工程学院 河南郑州 451400

- [7] LI C Y, LI L L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07)[2024-06-12]. <https://doi.org/10.48550/arXiv.2209.02976>.
- [8] REN S Q, HE K M, SUN J, 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.
- [9] 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]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway: IEEE, 2023: 7464-7475.
- [10] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2025-01-22]. <https://doi.org/10.48550/arXiv.1804.02767>.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4:

optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2024-11-23]. <https://doi.org/10.48550/arXiv.2004.10934>.

- [12] ZHANG Y, GUO Z Y, WU J Q, et al. Real-time vehicle detection based on improved YOLOv5[J]. Sustainability, 2022, 14(19): 12274.
- [13] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1580-1589.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision—ECCV 2018. Berlin: Springer, 2018: 3-19.

【作者简介】

宋江民(2001—),男,陕西商洛人,硕士研究生,研究方向:目标检测。

(收稿日期:2025-03-19 修回日期:2025-07-31)