

路侧违停车辆移动影像检测方法

唐康^{1,2,3}, 孙玉^{1,2,3}, 仲道阳^{1,2,3}, 高嘉良^{1,2,3}, 陈崇成^{1,2,3}

1. 福州大学空间数据挖掘与信息共享教育部重点实验室, 福建 福州 350108;
2. 福州大学地理空间信息技术国家地方联合工程研究中心, 福建 福州 350108;
3. 数字中国研究院(福建), 福建 福州 350108

收稿日期: 2025-01-23; 修回日期: 2025-09-01

中图分类号: P237

文献标识码: A

文章编号: 1001-1595(2025)09-1712-15

基金项目: 国家自然科学基金(42171426); 福建省经济联盟科技创新团队项目; 地理信息科学与技术全国重点实验室开放基金(2024KFJJ025)

第一作者简介: 唐康(2000—), 男, 硕士生, 研究方向为数字地球与智慧社会。 E-mail: 1635138766@qq.com

通信作者: 孙玉 E-mail: jade.yusun@outlook.com

摘要:路侧停车区在缓解城市停车压力方面发挥了重要作用,但随着城市机动车保有量的持续增长,城市路侧停车位供需缺口持续扩大,由此导致的违停现象严重影响了交通效率和安全。现有基于定点摄像头或传感器的违停监控系统存在成本高、监控范围受限等问题。为此,本文提出了一种基于移动摄像头的路侧停车区疑似违停车辆检测方法。该方法利用嵌入式设备结合改进的目标检测算法(mAP@50提升3.3%)进行开发,依托自制数据集完成模型训练,实现了疑似违停车辆的实时检测,并有效支持大范围区域的疑似违停监控任务。通过与同时段、同路段的无人机跟踪正射影像对比分析,本文方法的违停检测平均查准率达0.87、查全率达0.88,检测速度高达53.96 fps,充分验证了本文方法的可行性和有效性。

关键词:嵌入式设备; 目标检测; 疑似违停车辆

随着城市化进程的加快,城市道路的停车管理问题日益凸显,城市中违停现象的泛滥不仅造成了交通堵塞,还在一定程度上危及驾驶人员的生命安全^[1]。因此,高效、精准的违停监测技术对提升交通管理智能化水平、优化城市治理具有重要意义。目前主流的违停监测方式包括人工巡逻取证、地磁与感应线圈技术,以及基于计算机视觉的智能监测。人工巡查取证的方案主要依靠交警部门工作人员现场对违停车辆进行取证。地磁与感应线圈技术主要利用地磁传感器、超声波传感器或者红外线传感器等设备,检测非停车区域是否有车辆停放。这些传感器可以通过测量车辆与地面之间的距离或者车辆停留的时间来判断是否存在违停行为。基

于计算机视觉的智能监测方案主要利用基于深度学习的图像处理技术对监控设备所获取的视频帧进行一系列处理,最终获取违停车辆信息。在上述3种方案中,人工巡查取证的方式可能由于工作人员的主观性导致误判,并且该方案效率低下;基于地磁与感应线圈技术的方案需要大量传感器,考虑到传感器的购入以及维护成本较高,无法做到大范围布控;基于计算机视觉的智能监测方案以其相对廉价的成本以及较强的自动化检测能力,成为目前主要的研究热点。文献[2]通过目标检测模型获取车辆目标,然后使用射线法和车牌识别进行车辆违停判定,最后在嵌入式设备上对监控视频中的违停车辆进行检测。文献[3]将优化后的SSD(single

引文格式: 唐康, 孙玉, 仲道阳, 等. 路侧违停车辆移动影像检测方法[J]. 测绘学报, 2025, 54(9): 1712-1726. DOI: 10.11947/j. AGCS. 2025. 20250038.

TANG Kang, SUN Yu, ZHONG Xiaoyang, et al. Moving video-based detection for roadside illegally parking vehicles[J]. Acta Geodaetica et Cartographica Sinica, 2025, 54(9): 1712-1726. DOI: 10.11947/j. AGCS. 2025. 20250038.

shot multi box detector)目标检测算法与运动跟踪算法相结合,对兴趣区内的违章车辆进行判断,该方法不仅克服了光照、天气的影响,还使得车辆的检测速度及精度都大幅度提高。文献[4]使用YOLOv8目标检测算法及目标跟踪算法以固定视角在确保较高准确率的前提下对违停车辆进行检测。文献[5]在YOLOv8中通过构建膨胀空间金字塔池化(DSPP)模块和基于分支注意力机制的膨胀空间金字塔池化(DSPPA)模块,有效提高了视角倾斜、车辆重叠遮挡等情况下违规占用停车位的查全率和查准率。

随着深度学习技术的应用,违停检测的成本已得到显著降低。然而,上述研究中的违停检测方案仍依赖定点摄像头。传统基于定点摄像头的监测方案虽然能够在一定范围内检测违停车辆,但其监控区域通常有限,难以满足大范围的停车监测的需求。为了实现灵活高效的实时违停车辆检测,许多学者对基于移动摄像平台的违停监测任务展开了研究。文献[6]利用安装在仪表板上的智能手机的后置摄像头,并执行视觉场景和位置分析,以识别潜在的停车违规行为。文献[7]在车载相机上使用SSD算法对车辆的最小违停单元进行检测,并将违停现象划分为6类。文献[8]利用目标跟踪算法结合基于GPS的前车运动判别算法对道路上的违停车辆进行检测。文献[9]采用无人机航拍策略对违停车辆进行识别。该方案容易受到停车区周边绿化设施的遮挡影响。

尽管基于移动摄像平台的违停监测技术在常规场景中已展现出良好的应用效果,但在复杂城市场景下面临挑战。实际道路环境中存在3类主要制约因素:其一,城市停车位存在垂直式、平行式、斜列式等多样化布局形态;其二,违停车辆呈现多角度停放、部分越界等复杂空间姿态;其三,混合交通场景中存在遮挡、光照变化等干扰因素。这些现实条件导致基于深度学习的检测方法面临双重困境:技术层面需要构建细粒度的违停分类体系,数据层面需要大量数据集的支撑,若将违停检测任务聚焦于特定类型的停车区,则能够消除大量由目标检测算法带来的误差。为此,本文聚焦于城市道路两侧平行式停车区中四轮机动车的压线、停于泊位外和占用双车位的违停行为,提出基于移动嵌入式设备的路侧停车区疑似违停车辆检测方法,研究内容包括:①针对YOLOv5^[10]在自制数据集下出现的误分类和定位模糊现象,在YOLOv5颈部网络引入CBAM(convolutional block attention module)注意

力机制^[11]和MPDIoU^[12]边框损失函数,旨在通过双重通道注意力模块自动学习并强调对大型和小型目标检测最为关键的通道特征和空间特征,同时优化CIoU^[13]的不足,使网络在反向传播过程中,各参数获得更合理的传播梯度;②路侧停车区疑似违停车辆检测机制的设计与实现。

1 YOLOV5目标检测算法改进

YOLOv5相较于其他YOLO系列算法具有快检测速度和高检测精度的优势,特别适合交通场景中基于嵌入式设备的实时目标检测任务,其网络结构和高效的训练机制,能够在不牺牲精度的情况下实现快速推理,在一定程度上优化了嵌入式设备推理速度慢的缺点。因此选其作为本文的基准目标检测算法。YOLOv5网络结构主要由主干网络(backbone)、颈部网络(neck)和头部网络(head)3部分组成。主干网络负责从输入图像中提取特征,颈部网络对不同尺度的特征图进行融合,以增强网络的整体表达能力,而头部网络层则用于最终的回归预测。本文在Neck层中的 20×20 和 80×80 特征图融合前引入了CBAM注意力机制旨在提高对大目标和小目标的检测能力;为了更合理地优化损失值,本文使用MPDIoU替代了YOLOv5中的CIoU。最终,本文方法的网络结构如图1所示。

1.1 CBAM注意力机制

颈部网络网络在YOLOv5中起到关键作用,其采用路径聚合网络(path aggregation network, PANet)^[14]在通道维度上叠加融合不同尺度的特征图(20×20 、 40×40 、 80×80)实现不同层级之间的语义特征和定位特征的传递。特征图越大,对应的感受野越小,越适合进行小型物体的检测。在本文中,大型目标和小目标所占比重较大,且对此二类目标的检测效果直接影响后续疑似违停车辆判别的精度。因此,图1中红色虚线框所标注的位置引入CBAM注意力机制,主要为了在进行 20×20 和 80×80 特征图融合之前抑制无关通道信息和增强目标所在位置空间信息以增强融合后特征图的表征能力,进而可以有效提高目标检测模型的检测能力。CBAM注意力模块由通道注意力子模块与空间注意力子模块串联构成,其结构如图2所示。在通道注意力机制中,首先对输入特征图(维度为 $W \times H \times C$)沿空间维度进行全局最大池化(global max pooling)和全局平均池化(global average pooling),分别生成两个 $1 \times 1 \times C$ 的通道特征描述量。随后

将这两个特征描述符分别输入共享参数的两层全连接神经网络(MLP),通过非线性变换提取通道间相关性。将两个 MLP 输出特征进行逐元素相加后,经 Sigmoid 函数归一化处理,最终输出多维通道权重向量。空间注意力机制则采用互补的维度处理策略:首先在通道维度上对特征图分别执行全局最大池化与全局平均池化,得到两个 $H \times W \times 1$ 的

空间特征映射。将这两个空间特征沿通道维度拼接为 $H \times W \times 2$ 的复合特征后,采用 7×7 卷积核进行跨通道信息融合,将通道维度压缩至 1,最终通过 Sigmoid 函数生成 $H \times W$ 空间维度的注意力权重矩阵。通道与空间注意力机制通过级联方式协同工作,实现对特征图"通道-空间"双维度的自适应校准。

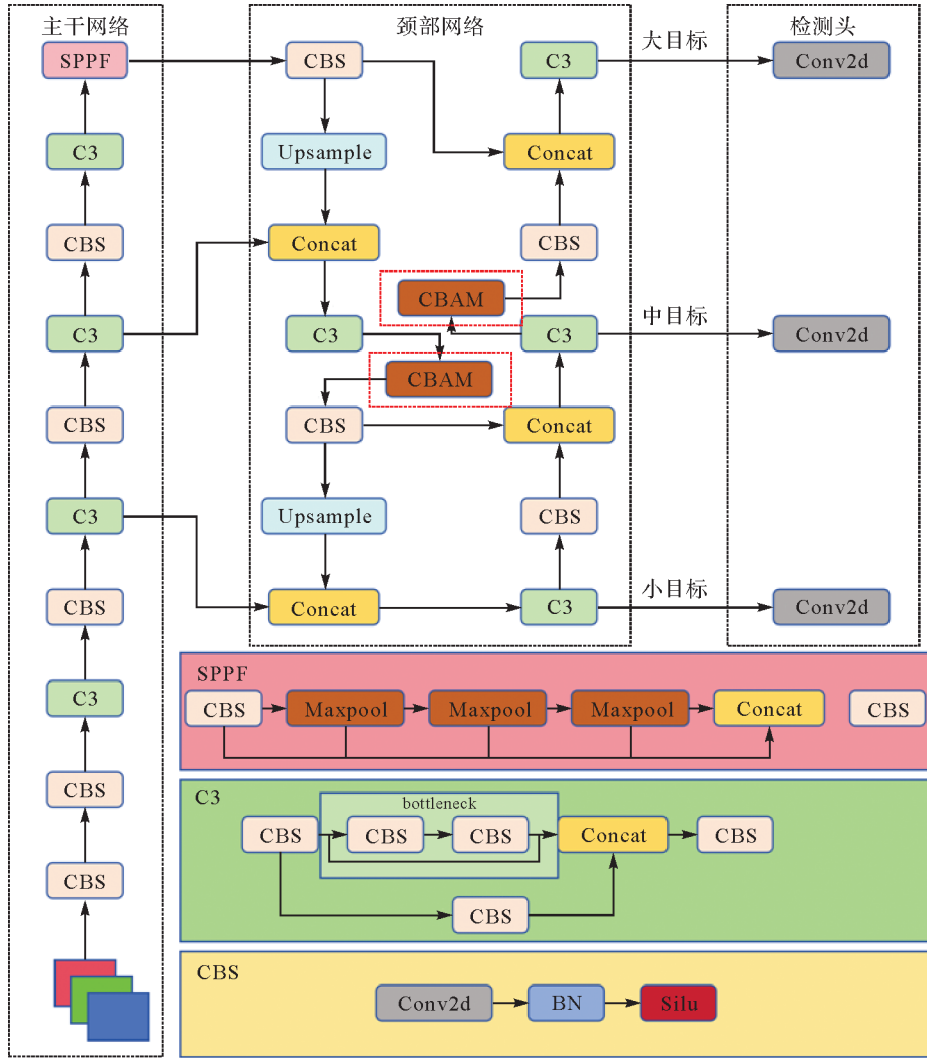


图 1 网络结构

Fig. 1 Network structure diagram

1.2 MPDIoU 边框损失函数

在模拟行车记录仪视角所采集的街景图像中,可能出现目标遮挡的情况,受重叠因素影响,会导致被挡目标部分信息缺失,从而无法准确描述重叠目标边界框长宽值的变化。YOLOv5 的边框损失函数采用 CIoU, CIoU 相较以往边框损失函数,不仅关注重叠区域,还考虑了预测框与标注框的中心点距离及长宽比的匹配。这使得模型在回归边框时能够更精确地调整位置和形状,从而提升检测精度

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

$$\alpha = \frac{v}{1 - IoU + v} \quad (2)$$

$$v = \frac{4}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

式中, IoU 代表预测框与标注框的交并比; $\rho^2(b, b^{gt})$ 代表预测框与标注框中心点的欧氏距离; c 代表能够同时包含预测框与标注框的最小矩形的对角

线距离; w^{gt} 、 h^{gt} 、 w 、 h 分别代表标注框的宽高和预测框的宽高。

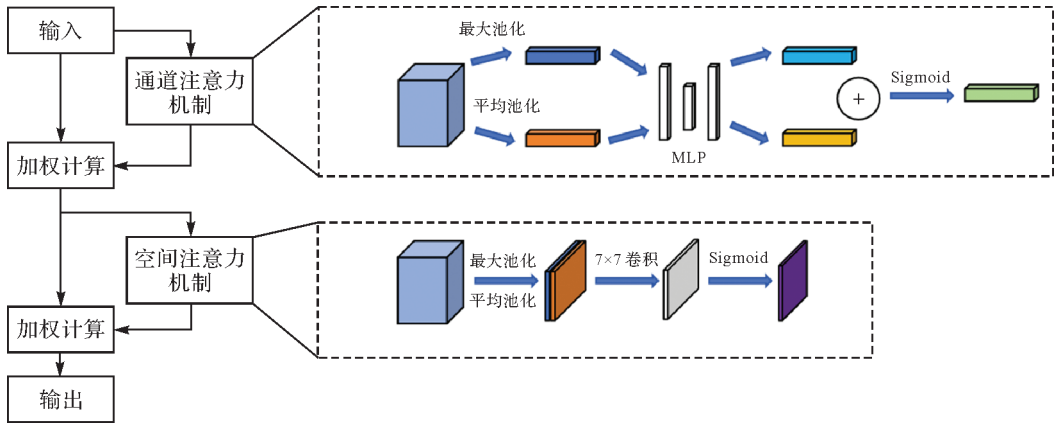


图 2 CBAM 结构
Fig. 2 CBAM structure

虽然 CIoU 拥有较强的梯度传递能力,但是在预测框与标注框之间拥有相同宽高比但数值不同的情况下(图 3),CIoU 会使两种情况损失值相同,使得网络在反向传播过程中所传递的梯度信息不合理,从而导致 YOLOv5 在处理重叠目标时易出现预测框偏移,甚至漏检的情况。因此,本文引入 MPDIoU 加以改善,MPDIoU 通过最小化预测边界框和真实边界框之间的左上角和右下角点之间的距离结合图像整体宽度与高度对边框损失值进行定义,以此更好地训练目标检测、字符级场景文本检测和实例分割的深度学习模型

$$LOSS_{MPDIoU} = 1 - IoU + \frac{d_1^2}{w^2 + h^2} + \frac{d_2^2}{w^2 + h^2} \quad (4)$$

式中, d_1 、 d_2 分别代表预测边界框和真实边界框之

间左上角和右下角点之间的距离; w 、 h 分别代表图像宽度与高度。

由图 3 可知,若使用 CIoU,则图 3(a)、(b)两种情况下的边框损失值均为 $1 - k_1^2$,这是由于预测框与标注框之间具有相同的宽高比和相同中心点,使得 CIoU 中的惩罚项失去了作用,只以 IoU 对损失值进行定义,进而无法对有效区分图 3(a)、(b)所展现的不同情况;若使用 MPDIoU 作为边框损失函数,则图 3(a)、(b)情况下的边框损失值分别为 $1 - k_1^2 + (1 - k_1)^2 (w^{gt} + h^{gt}) / 2(W^2 + H^2)$ 、 $1 - k_1^2 + (1 - k_1)^2 (w^{gt} + h^{gt}) / 2k_1^2(W^2 + H^2)$,由此可以看出 MPDIoU 的引入可以有效解决预测框与标注框之间拥有相同宽高比但数值不同的情况下损失值相同的问题,使得网络训练过程朝着最优方向进行。

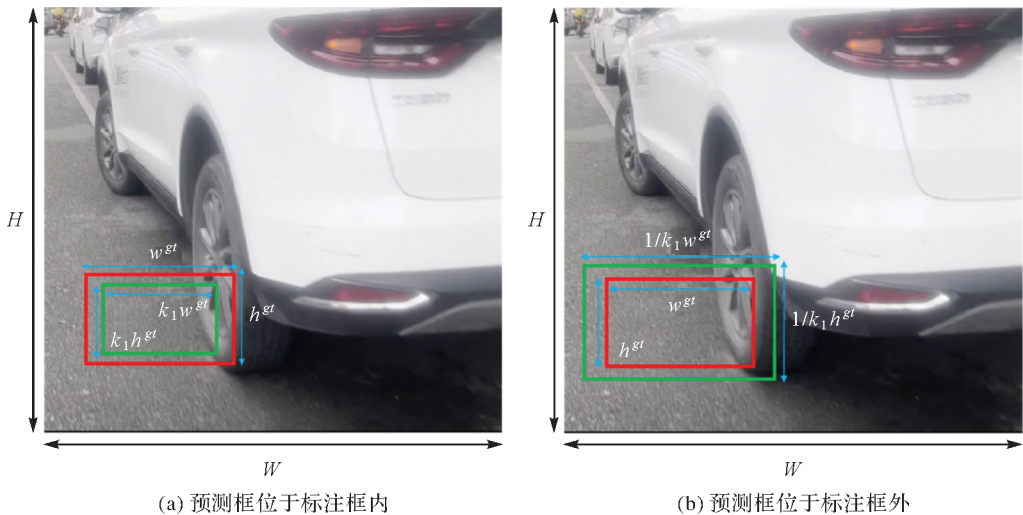


图 3 不同情况预测框
Fig. 3 Prediction boxes for different situations

Fig. 3 Prediction boxes for different situations

2 路侧停车区疑似违停车辆检测方案

疑似违停车辆检测方案的建立作为作为本文

主要的研究目标,主要涉及到目标检测与跟踪、基于虚拟线的启动机制、轮胎类别组合判别算法,整体方案流程如图 4 所示。

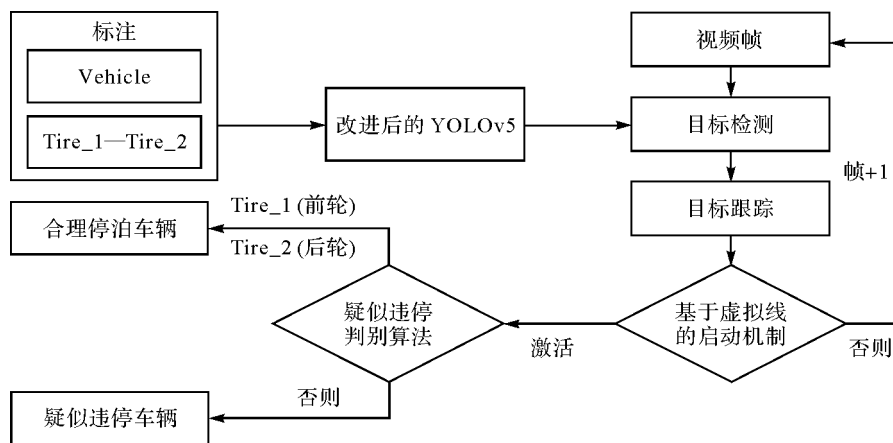


图 4 疑似违停判别方案流程

Fig. 4 Flowchart of suspected illegal parking discrimination strategy

2.1 目标检测与跟踪

目标检测使用优化后的 YOLOv5 目标检测算法将街景图像中的路侧停车区车辆划分为 3.2 小节所述的 5 种类别。由于本文研究的主体是车辆,而目标检测算法破坏了不同视频帧中同一车辆之间的客观联系,因此需要采用目标跟踪算法对不同视频帧中相同车辆进行独有 ID 赋值。目前主流的多目标跟踪算法主要有 DeepSORT 算法^[15]和 SORT (simple online and realtime tracking) 算法^[16],其中 DeepSORT 算法在 SORT 算法的基础上增加了一个特征提取网络,并且在匹配策略上采用同时基于运动信息和目标外观特征的方法。考虑到本文基于嵌入式设备,综合算力相对于 PC 端较弱,DeepSORT 的特征提取策略较为耗时,因此采用 SORT 算法对车辆进行目标跟踪。在本文中 SORT 算法的最大丢失帧数(max_age)设置为 5,生成 tracker 的最小持续检测帧(sort_min_hits)设置为 2, IoU 阈值(sort_iou_thresh)设为 0.2。

2.2 基于虚拟线的启动机制

在本文中,对车辆是否处于疑似违停状态只需进行单次判断,由于视频帧是连续的,而同一车辆会出现于相邻视频帧中,为规避连续视频帧中针对同一车辆的重复判别问题,本文采用一种基于虚拟线的方式^[17]作为疑似违停判别算法的启动器。该方法通过车辆运动轨迹与预设虚拟线的空间交互关系确定关键判别帧,确保每辆车仅触发一次违停判定。具体地,考虑到本文模拟行车记录仪视角,

路侧停车区域车辆呈现相对于摄像头的逆向运动特征,基于这个特点,笔者在屏幕合适位置设定一条垂直虚拟线 L ,连接相邻帧中同一车辆跟踪框中心点作为 L_c ,当 L 与 L_c 相交时,则触发并启动疑似违停车辆判别算法,如图 5 所示。为了确定虚拟线的最佳位置,本文在 $0.6 \times \text{width}$ 、 $0.7 \times \text{width}$ 、 $0.8 \times \text{width}$ 处分别设置虚拟线,基于所录制的街景视频数据进行试验测试。试验结果见表 1,其中触发数代表成功启动规范性判别算法的车辆数,触发比则是触发数与车辆总数的比值,可以看出, $0.7 \times \text{width}$ 位置的纵向虚拟线表现出最优试验结果,有 58 辆车启动了规范性判别算法,触发比达到了 0.98。

表 1 不同位置虚拟线触发情况

Tab. 1 Triggering conditions of virtual lines at different locations

虚拟线位置	总车辆数	触发数	触发比
$0.6 \times \text{width}$	61	58	0.95
$0.7 \times \text{width}$	61	60	0.98
$0.8 \times \text{width}$	61	34	0.56

2.3 疑似违停判别算法

在路侧停车区中,车辆的违停与否客观上取决于车辆与路侧停车位之间的相对位置关系。在 3.2 小节中,笔者对数据集进行了介绍,其中 Tire_1 至 Tire_4 很好地表现了车辆轮胎与停车位之间的相对位置关系。通过轮胎与停车位相对位置关系的组合,可以进一步表明车辆与停车位的相对位置关系,进而对疑似违停车辆进行判别。

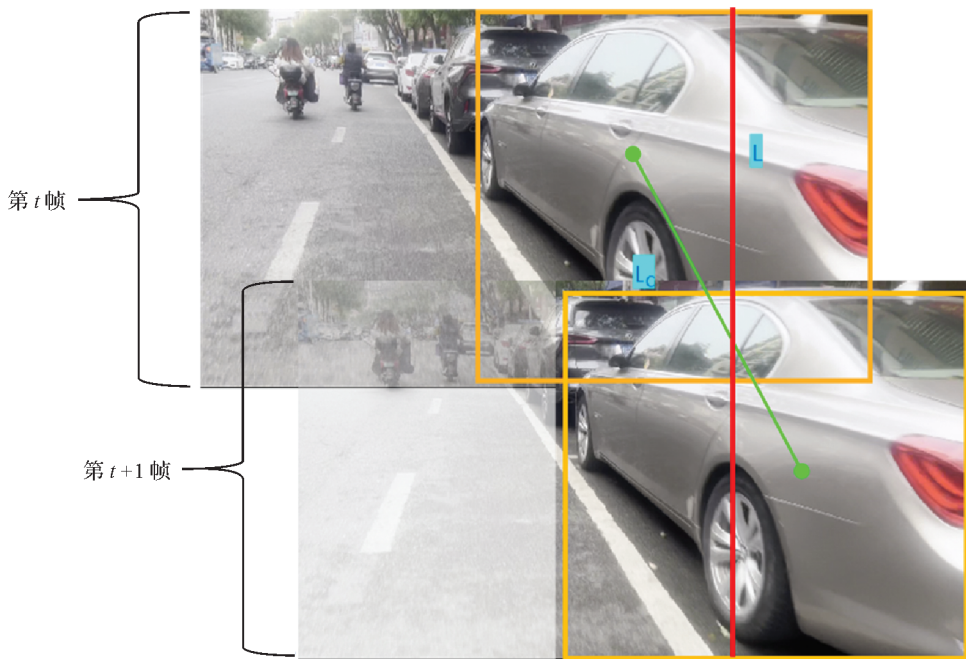


图 5 基于虚拟线的启动机制

Fig. 5 Virtual line-based startup mechanism

本文采用如图 6 所示的疑似违停判别算法对路侧停车区域的目标车辆进行系统性判别检测。在算法实施过程中,首先需要解决的核心技术难点在于建立轮胎与所属车辆之间的空间归属关系。由于常规目标检测算法仅能实现车辆与轮胎的独立识别,而目标跟踪算法亦难以直接构建二者之间的空间关联。为了解决这个问题,本文提出基于 IoU

的匹配机制。具体而言,当某一车辆启动违停判别算法时,算法将针对目标车辆进行多维度分析:首先根据目标跟踪算法所获取到的跟踪边界框信息,然后同步计算其与当前帧所有轮胎检测框的 IoU 值,最后通过数值筛选获得最大 IoU 数值且非零的两个轮胎检测目标作为匹配对象。

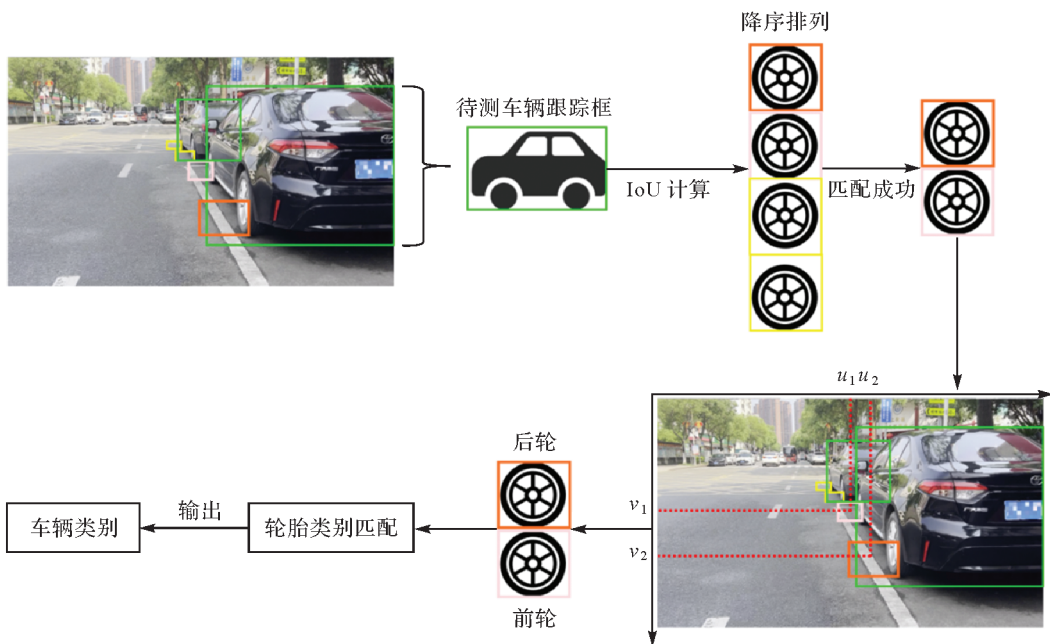


图 6 疑似违停判别算法

Fig. 6 The suspected illegal parking judgment algorithm

在完成初步匹配后,本文将进行精细化空间关系判定。算法自动提取与该车辆所匹配的两个轮胎检测框的中心点像素坐标 (u_1, v_1) 和 (u_2, v_2) ,基于像素坐标系横向数值和纵向数值的对比关系进行前后轮的区分。根据道路场景中车辆常规停放方向的特征,本文将横向与纵向坐标值较小的轮胎判定为前轮,相对较大者则对应后轮位置。在此基础上,算法通过分析前后轮类型组合模式与预设合规停车标准的比对,实现对车辆停放状态的智能判别,相关分类细节和规则见表2。由表2可知,当车辆前轮类别为Tire_1,后轮类别为Tire_2时,车辆为合理停泊车辆,其余情况皆为违停车辆。

表2 路侧停车行为分类规则

Tab. 2 Classification rules for roadside parking

Tire_1	Tire_2	Tire_3	Tire_4	疑似违停	合理停泊
前	后				✓
前		后		✓	
	前 & 后			✓	
前			后	✓	
	前	后		✓	
	后	前		✓	
后	前			✓	
后			前	✓	
		后	前	✓	
			前 & 后	✓	

3 试验结果

3.1 试验设置

在本文中,由于后续的路侧停车区疑似违停车辆检测方案基于嵌入式设备所运行,Jetson Orin NX 嵌入式设备拥有强大的图像计算能力。为了保证后续测试的稳定运行,本文在PC端对试验组模型进行训练后,需在Jetson Orin NX 嵌入式设备上对试验组模型进行推理时间测试。所有试验组均在相同的硬件设备、环境配置、参数设置、数据集下进行。本文试验所采用的硬件设备信息见表3。训练过程采用SGD优化器,学习率为0.001, batch size为2,动量因子为0.937,权值衰减为0.0005,训练轮次为100。

3.2 数据集介绍

本文针对行车记录仪视角下的路侧停车区车辆检测问题展开数据构建工作。尽管KITTI数据集^[18]作为行车记录仪视角的经典开源数据集具有视角适配性,但其在路侧停车场景下的车辆样本规模及类别划分难以满足实际需求。为此,笔者在福建省福州市闽侯县上街镇开展数据采集工作。

表3 试验硬件配置

Tab. 3 Experimental hardware configuration

配置	Jetson Orin NX	PC
OS	Ubuntu 20.04 LTS	Windows 10
CPU	6-core NVIDIA Arm Cortex A78AE v8.2 64 bit CPU 2MB L2+4MB L3	Intel(R) Core(TM) i7-11700
GPU	1024-core NVIDIA Ampere GPU with 32 Tensor Cores	GTX1660Super
Memory	16 GB 128 bit LPDDR5 102.4 GB/S	32GB (2933 MHz/ 3200 MHz)

研究针对路侧停车区街景影像采集需求,提出基于移动平台的动态数据采集方案,即采用iPhone14 智能手机搭载电动自行车构建移动采集系统,以1920×1080分辨率、30 fps帧率对包含路侧停车区域的城市场景进行多时段动态视频采集。为优化数据冗余并保证时序特征完整性,设计帧间采样间隔为6帧的视频关键帧提取策略,经数据清洗后最终构建涵盖不同光照条件、不同地面纹理,共1939张有效样本的高质量数据集。部分影像样本如图7所示。

依据道路交通标志标线规范(GB5768—2009)对停车区域的界定标准,结合违停车辆与停车位边线的空间位置关系,本文创新性地构建了如图8所示的五元分类体系。其中,Vehicle代表纵向停放车辆,Tire_1代表同时包括纵向停车线和横向停车线的轮胎区域,Tire_2代表仅包含纵向停车线的轮胎区域,Tire_3代表轮胎压住纵向停车线区域,Tire_4代表既不包含纵向停车线也不包含横向停车线的轮胎区域。数据标注方面,采用labelme工具进行精细化标注,确保边界框与车辆实际轮廓的精准匹配,最终形成包含6254个有效标注样本的数据集(表4)。在数据划分策略上,采用分层随机抽样方法将数据集按7:3比例划分为训练集(1357张)和验证集(582张),有效保证了两大子集在场景复杂度、光照条件及车辆分布上的均衡性。

表4 数据集标签分布

Tab. 4 Distribution of dataset labels

标签类别	标签数量	占比/(%)
Vehicle	1980	32
Tire_1	1267	20
Tire_2	1017	16
Tire_3	710	11
Tire_4	1280	20

3.3 评价指标

本文使用Flops(floating point operations)、查

准率 (Precision)、召回率 (Recall)、mAP (mean average precision) 作为试验的评价指标,其中 Flops 用于衡量模型在一次前向传播中所需执行的总浮点运算次数,查准率、召回率、mAP,公式为

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{mAP} = \frac{\sum_{i=1}^N \text{AP}_i}{N} \quad (7)$$

式中,TP 表示预测为正例且实际为正例的样本数;FP 表示预测为正例但实际为负例的样本数;FN 表示预测为负例但实际为正例的样本数;AP 代表由每一轮训练结果所构成的 PR 曲线的面积;mAP 表示一定 IoU 阈值下 N 个类别 AP 的平均值,本文 IoU 阈值设为 0.5。



(a) 沥青地面采集的图像



(b) 水泥地面采集的图像



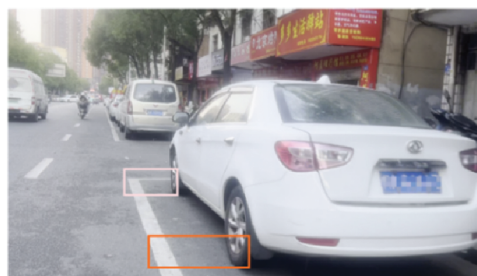
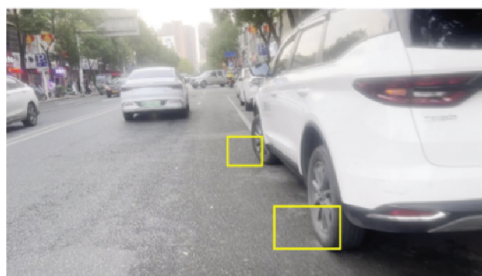
(c) 阴雨采集的图像



(d) 傍晚采集的图像

图 7 采集的图像

Fig. 7 Collected image



Vehicle Tire_1 Tire_2 Tire_3 Tire_4

图 8 自制数据集分类情况

Fig. 8 Classification of the custom dataset

3.4 YOLOv5 改进结果分析

为评估边框回归损失函数在目标检测任务中的性能差异,本文基于统一网络架构(YOLOv5)开展对照试验,分别采用 MPDIoU、Wise-IoU^[19] 及 Inner-IoU^[20] 3 种损失函数进行训练验证。由表 5 可知,MPDIoU 在查全率与 mAP@50 指标上显著优于其他损失函数,其查准率虽略低于 Wise-IoU,但综合性能表现最优,是本文数据集的最佳边框损失函数。MPDIoU 在边框回归任务中展现出的优越性源于其创新性的几何建模方式。相较于 CIoU、Wise-IoU 等传统损失函数,MPDIoU 通过最小化预测框与真实框顶点间的最大投影距离,实现了对边界框位置关系的更全面建模。这种设计突破性地解决了对比项存在的 3 个关键问题:首先,它无须像 CIoU 那样引入复杂的长宽比惩罚项,避免了因参数敏感导致的优化偏差;其次,相比 Wise-IoU 的动态权重策略,其几何特性具有更好的尺度不变性,能更稳定地处理不同尺寸目标;最后,通过检测框与标注框之间顶点投影距离进行损失函数定义,相较于 Inner-IoU 的单纯内部区域重叠度量更能精准控制边界框的绝对位置。

表 5 各损失函数试验结果

Tab. 5 Experimental results of various loss functions

边框损失函数	查准率	召回率	mAP@50
CIoU	0.911	0.908	0.941
MPDIoU	0.923	0.922	0.951
Wise-IoU	0.928	0.898	0.942
Inner-IoU	0.902	0.904	0.939

注:加粗表示最优结果。

在验证 MPDIoU 的有效性后,本文进一步将原始的 CIoU 替换为 MPDIoU,探讨不同注意力机制的引入对模型性能的影响。具体而言,在 Neck 层的相同位置分别添加了 CBAM、CA^[21]、EMA^[22] 和 MLCA^[23] 注意力机制,并对模型性能进行详细对比。结果表明(表 6),加入 CBAM 注意力机制后,模型的查准率、召回率及 mAP@50 值均显著高于其他注意力机制。这表明,CBAM 注意力机制能够更有效地优化特征提取与表达,在本文试验数据集下具有显著的优越性。CBAM 相较于 CA、EMA、MLCA 等注意力机制的优越性体现在双维度动态校准与高效特征聚焦,不同于 CA 和 MLCA 仅优化通道权重(忽略空间关系)或 EMA 侧重多尺度融合(缺乏显式空间定位),CBAM 通过通道注意力与空间注意力的串联互补机制,在低计算成本下更精准平衡特征通道的语义表达和空间结构的上下文感知。

表 6 各注意力机制试验结果

Tab. 6 Experimental results of various attention mechanisms

注意力机制	查准率	召回率	mAP@50
无	0.923	0.922	0.951
CA	0.925	0.935	0.965
EMA	0.935	0.934	0.968
MLCA	0.936	0.931	0.963
CBAM	0.940	0.939	0.972

注:加粗表示最优结果。

为了探究本文模型与目前主流单阶段目标检测模型的性能差别,选取 YOLOv3-tiny^[24]、YOLOv8、YOLOv9、YOLOv10^[25]、YOLOv11 单阶段目标检测算法与本文优化后的模型进行横向对比,通过查全率、查准率、mAP@50、模型大小及推理用时进行衡量,其中查全率、查准率、mAP@50、模型大小为 PC 端测试结果,推理用时为在 Jetson Orin NX 嵌入式设备上的测试情况,其结果见表 7。本文方法在性能指标上表现突出。具体而言,本文方法的查准率达到 0.940,查全率为 0.939,其 mAP@50 更是达到 0.972,优于所有对比模型。这表明本文方法在检测精度和召回率上具备显著优势。此外,本文方法的推理用时为 22 ms,尽管相比 YOLOv3-tiny 和 YOLOv5 略高,但远低于 YOLOv9 和 YOLOv10,显示出良好的实时性。同时,本文方法的大小为 14.05 MB,在保持较小模型规模的同时实现了优异的性能,兼具轻量化和高效性。相比之下,其他模型在性能上存在一定不足。例如,YOLOv10 的 mAP@50 仅为 0.898,推理用时达 33 ms,性能与效率较为平衡但总体表现不及本文方法;YOLOv9 虽然 mAP@50 达到 0.949,但推理用时高达 52.2 ms,难以满足高实时性需求;YOLOv8 在 mAP@50 上表现较优,为 0.951,但模型大小为 21.46 MB,推理用时为 24.5 ms,相较本文方法在轻量化方面存在一定劣势。

表 7 对比试验分析结果

Tab. 7 Comparative experimental analysis results

模型	查准率	召回率	mAP@50	模型大小 /MB	推理用时 /ms
YOLOv5	0.911	0.908	0.941	13.69	20.9
YOLOv8	0.912	0.930	0.951	21.46	24.5
YOLOv9	0.890	0.902	0.949	14.52	52.2
YOLOv10	0.815	0.864	0.898	15.75	33
YOLOv11	0.902	0.878	0.938	18.28	28.4
YOLOv3-tiny	0.902	0.863	0.927	23.23	18
本文模型	0.940	0.939	0.972	14.05	22

注:加粗表示最优结果。

各单阶段目标检测模型和本文方法的检测效果如图 9 所示。



图 9 对比试验结果

Fig. 9 Comparison of experimental results

通过对比结果发现 YOLOv5 在图 9(b)中出现定位模糊现象,针对此类情况,本文采用 MPDIoU 损失函数来充分挖掘水平矩形的几何特征,与其他损失函数不同,MPDIoU 损失函数基于最小点距离

的原理,将目标框的宽高偏差纳入考量,使得目标即便在重叠的情况下,本文方法也能够很好地预测出准确的边界框,从而提高本文方法对目标的定位性能。由图 9(a)可知,由于车位较宽,车辆轮胎与

停车位边界线的距离相对较远,众多算法无法准确地对目标进行分类;同时,由于 YOLOv10s 和 YOLOv11s 对该图像特征提取的不全面性,使得两种模型在该场景中还存在对车辆的漏检现象。由图 9(b)可知,由于曝光的影响,弱化了 YOLOv10 和 YOLOv3-Tiny 对相距较远车辆特征的提取效果,因此导致了漏检现象。图 9(c)中阴影的影响,使得除了 YOLOv8 以外的其他算法都出现误检现象。鉴于图 9(a)和图 9(c)中的漏检和误检现象,本文以 YOLOv5 为基准在 Neck 层引入 CBAM 注意力机制,通过通道注意力模块强化关键通道特征响应,利用空间注意力动态调整区域权重分布,以达到提高模型整体的检测效果。

为了评估改进后模型与原模型在性能上的差异,笔者在 PC 端上使用 PyTorch 深度学习框架,以 FP32 精度进行了消融试验,并将消融试验模型移植于嵌入式设备 Jetson Orin NX 进行推理时间测试,试验结果见表 8。同时引入 CBAM 注意力机制与 MPDIoU 边框损失函数时,本文方法在检测精度上显著优于原方法,mAP@50 提升 3.3%,仅增加 0.6×10^9 FLOPs 和 1.1 ms 推理耗时,验证了本文方法在有限计算资源下实现精度提升的有效性。当删除 CBAM 模块时,模型 mAP@50 值下降了 2 个百分点,FLOPs 下降 0.6×10^9 ,说明在 Neck 层引入 CBAM 有效地聚焦空间和通道层面的重点特征,加强了 20×20 、 80×80 特征图的表征能力,是模型复杂度增加的主要来源;当删除 MPDIoU 时,mAP@50 值下降 1 个百分点,推理时间和 FLOPs 值均未改变,说明 MPDIoU 对目标能够精确定位,充分挖掘了目标边界框的几何信息,提高了边界框回归的准确性,同时说明该改进项仅影响训练过程中的反向传播机制,不增加推理阶段的算力消耗。

表 8 消融试验分析结果

Tab. 8 Results of the ablation experiment analysis

CBAM	MPDIoU	FLOPs	mAP@50	Time/ms
		15.8×10^9	0.941	20.9
✓		16.4×10^9	0.962	22
	✓	15.8×10^9	0.951	20.9
✓	✓	16.4×10^9	0.972	22

3.5 疑似违停车辆检测方案测试

为了验证本文方法的有效性,笔者在福建省福州市闽侯县上街镇部分具有路侧停车区的道路上开展了测试试验。该测试路段具有一定的车辆停放密度,且光线环境充分。考虑到相机自身移动速度对目标检测算法结果影响较大,进而影响疑似违

停车车辆的判别,本文将硬件设备搭载于电动自行车以 10 km/h 的速度进行测试。在实际测试过程中,识别并记录了如图 10 所示的疑似违停车辆情况。根据不同轮胎类别组合特征,对测试中的车辆进行了 3 类违停行为分类:前后轮均为 Tire_4 类型,表示车辆停放位置偏离路侧停车位外;前后轮中包含 Tire_3 类型,表明车辆采用压线停放方式;前轮为 Tire_2 类型,后轮为 Tire_1 类型,说明车辆占用了双车位区域。从测试结果可以看出,本文方法在部分曝光严重的路段仍能成功识别潜在违规停车行为,进一步验证了本文方法在强光或高照明条件下也具备较高的稳健性和适应性。

为了对本文方法的准确性及实时性进行测试,在进行疑似违停车辆检测的同时采用大疆 Mavic 3M 无人机进行伴飞,以 4.84 cm 地面分辨率(ground sample distance, GSD),180 m 航高,80% 旁边重叠率,80% 纵向重叠率生成无人机正射影像,并采用目视解译的方法将无人机正射影像与本文方法所判别的目标车辆进行对比分析,同时评估本文方法的查全率、查准率及实时性。由于大多数路侧停车区周围树木遮挡较为严重,无法清晰地看出车辆与路侧停车位的相对位置关系,为了保证验证的可靠性,本文选取 4 条无人机可见度较高的路段进行验证。图 11 为采用无人机验证结果,可见在 4 个路段中均有疑似违停车辆,并且伴随小部分由目标检测算法误差和镜头摆动所导致的漏检或误检现象,而对于合理停泊车辆的检测,4 个路段中均未出现误检样本,只有路段 D 出现 3 个漏检样本,其主要原因在于漏检的合理停泊车辆外侧受到其余疑似违停车辆遮挡的影响,导致在基于虚拟线的启动机制中,虚拟线只与外侧疑似违停车辆的相邻帧跟踪框中心点的连线相交,因此忽略了对内侧合理停泊车辆的判别。

由于本文的主要对象是疑似违停车辆,因此笔者对图 11 中关于疑似违停车辆各类样本进行了详细统计,并结合查准率、召回率和检测速度等关键指标,对本文方法的表现进行了全面评估,评估结果见表 9。在图像分辨率为 940×544 的情况下,通过对所选取 4 条道路的测试结果表明,本文的疑似违停判别机制在这些道路上的平均查准率为 0.87,召回率为 0.88,平均检测速度达到 53.96 fps。这一结果不仅表明了本文方法在检测疑似违停车辆时具有较高的准确性和全面性,而且也反映出其在实时检测过程中具备较快的响应速度,能够满足实际应用中的实时性要求。因此,本文方法具有较

强的检测能力及较快的检测速度,能够有效应对复杂的道路环境。

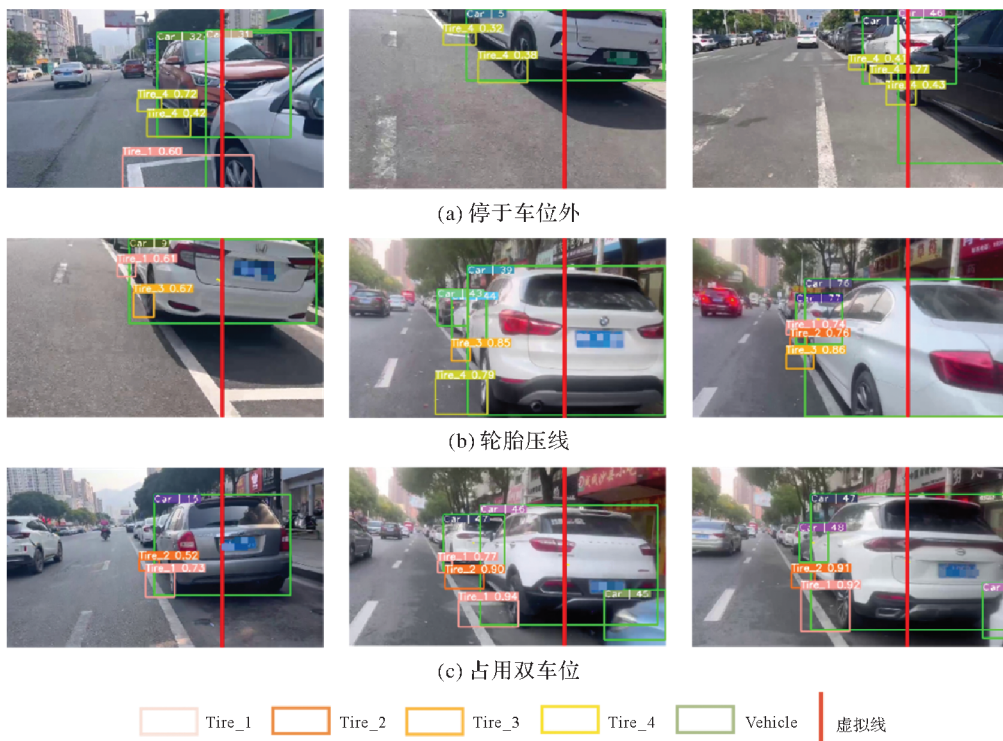


图 10 疑似违停车辆检测结果

Fig. 10 Detection results of suspected illegally parked vehicles

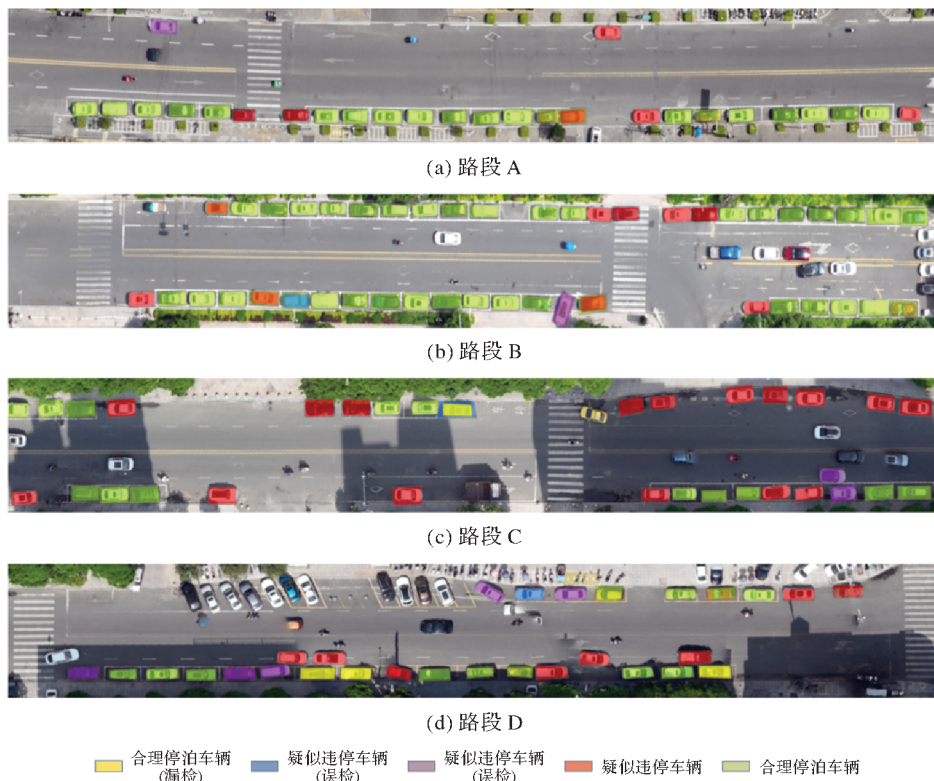


图 11 基于无人机正射影像的视觉解译结果

Fig. 11 Visual interpretation results based on UAV orthophoto

4 结论和展望

针对现有定点摄像头检测范围受限,以及基于移动摄像平台检测方案面临的停车位形态多样、停车空间姿态复杂、遮挡与光照影响等问题,本文提出了一种基于移动嵌入式设备的路侧停车区疑似违停车辆检测方法。该方法联合改进后的 YOLOv5 目标检测算法和 SORT 多目标跟踪算法,通过目标车辆前后轮类别组合,精准判別了复杂城市道路场景中平行式路侧停车区内四轮机动车的 3 种违停行为,同时实现了大范围区域性检测。基于无人机影像的验证结果表明,本文方法具有较高的查准率、查全率和检测速度。在未来的工作中,笔者将围绕基于动态背景的目标运动判断算法展开研究,重点引入激光雷达点云时序分析技术,通过多帧点云配准与目标轨迹追踪实现运动状态判別。具体而言,将采用改进的 NDT(正态分布变换)算法对连续帧点云进行非刚性配准,结合 IMU/GNSS 数据补偿自

车运动后,构建疑似违停车辆的三维运动矢量场;同时设计双模态校验机制,将 YOLOv5 检测框与点云投影进行空间对齐,利用激光雷达毫米波级测距精度验证目标位移真实性,有效消除视觉透视误差导致的伪运动。当检测到目标车辆在连续帧内位移量低于固定阈值,则判定为静止状态。将其与本文方法结合,进而实现行车记录仪视角下疑似违停车辆的运动性判断,若处于静止状态,则可将疑似违停车辆转变为违停车辆。

表 9 基于无人机正射影像的检测精度分析

Tab. 9 Detection accuracy analysis based on UAV orthophoto

ID	查准率	召回率	检测速度/fps	图像分辨率/像素
A	1.00	0.85	55.28	
B	0.9	0.9	53.17	
C	0.94	0.89	52.27	940×544
D	0.62	0.89	55.13	
平均	0.87	0.88	53.96	

参考文献

- [1] HUANG Y H, HSIEH C H. A decision support system for available parking slots on the roadsides in urban areas[J]. *Expert Systems with Applications*, 2022, 205: 117668.
- [2] 庄建军, 徐子恒, 张若愚. 基于改进的 YOLOv5 模型和射线法的车辆违停检测[J]. *南京信息工程大学学报*, 2024, 16(3): 341-351. ZHUANG Jianjun, XU Ziheng, ZHANG Ruoyu. Illegal parking detection based on improved YOLOv5 model and ray method[J]. *Journal of Nanjing University of Information Science & Technology*, 2024, 16(3): 341-351.
- [3] XIE Xuemei, WANG Chenye, CHEN Shu, et al. Real-time illegal parking detection system based on deep learning[C]//*Proceedings of 2017 International Conference on Deep Learning Technologies*. Chengdu: ACM Press, 2017: 23-27.
- [4] SHARMA N, BARAL S, PAING M P, et al. Parking time violation tracking using YOLOv8 and tracking algorithms[J]. *Sensors*, 2023, 23(13): 5843.
- [5] 陈伟, 王晓龙, 张晏玮, 等. 基于改进 YOLOv8 的高速公路服务区车辆违停检测[J]. *计算机工程*, 2024, 50(4): 11-19. CHEN Wei, WANG Xiaolong, ZHANG Yanwei, et al. Vehicle violation detection based on improved YOLOv8 in highway service areas [J]. *Computer Engineering*, 2024, 50(4): 11-19.
- [6] RANJAN A, MISRA P, VASAN A, et al. City scale monitoring of on-street parking violations with StreetHAWK[C]//*Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. New York: ACM Press, 2019: 31-40.
- [7] PENG Xinggan, SONG Rongzihan, CAO Qi, et al. Real-time illegal parking detection algorithm in urban environments[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(11): 20572-20587.
- [8] 程梁华, 黄瑞雪, 沈鑫. 适于高动态视频场景下的城市道路违停检测算法[J]. *计算机科学*, 2024, 51(12): 190-198. CHENG Lianghua, HUANG Ruixue, SHENG Xin. Urban illegal on-road parking detection algorithm for high dynamic video scenarios [J]. *Computer Science*, 2024, 51(12): 190-198.
- [9] 田爱军, 蔡旭阳, 陈玮, 等. 无人机端路面车辆违停检测及取证系统[J]. *测控技术*, 2021, 40(5): 67-74. TIAN Aijun, CAI Xuyang, CHEN Wei, et al. Vehicle illegal parking detection and evidence collection system on UAV[J]. *Measurement & Control Technology*, 2021, 40(5): 67-74.
- [10] Ultralytics. YOLOv5[EB/OL]. [2024-11-01]. <http://github.com/ultralytics/yolov5>.
- [11] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//*Proceedings of 2018 Computer Vision-ECCV*. Cham: Springer, 2018: 3-19.
- [12] MA Siliang, XU Yong, Mpdjou, a loss for efficient and accurate bounding box regression[EB/OL]. [2024-11-01]. <https://arxiv.org/>

abs/2307.07662.

- [13] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of 2020 AAAI conference on artificial intelligence. [S. l.]: PKP, 2020, 34(7): 12993-13000.
- [14] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake city:IEEE, 2018: 8759-8768.
- [15] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//Proceedings of 2017 IEEE International Conference on Image Processing. Beijing:IEEE, 2017: 3645-3649.
- [16] BEWLEY A, GE Zongyuan, OTT L, et al. Simple online and realtime tracking[C]//Proceedings of 2016 IEEE International Conference on Image Processing. Phoenix:IEEE, 2016: 3464-3468.
- [17] LIU Fei, ZENG Zhiyuan, JIANG Rong. A video-based real-time adaptive vehicle-counting system for urban roads[J]. PLoS One, 2017, 12(11): e0186098.
- [18] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: the KITTI dataset[J]. The International Journal of Robotics Research, 2013, 32(11): 1231-1237.
- [19] TONG Zanjia, CHEN Yuhang, XU Zzewei, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. [2024-11-01]. <https://arxiv.org/abs/2301.10051>.
- [20] ZHANG Hao, XU Cong, ZHANG Shuaijie. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[EB/OL]. [2024-11-01]. <https://arxiv.org/abs/2311.02877>.
- [21] HOU Qibin, ZHOU Daquan, FENG Jiashi. Coordinate attention for efficient mobile network design[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville:IEEE, 2021: 13708-13717.
- [22] OUYANG Daliang, HE Su, ZHANG Guozhong, et al. Efficient multi-scale attention module with cross-spatial learning[C]//Proceedings of 2023 IEEE International Conference on Acoustics, Speech and Signal Processing. Rhodes Island: IEEE, 2023: 1-5.
- [23] WAN Dahang, LU Rongsheng, SHEN Siyuan, et al. Mixed local channel attention for object detection[J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106442.
- [24] REDMON Joseph, FARHADI Ali. Yolov3: an incremental improvement[EB/OL]. [2024-11-01]. <https://arxiv.org/abs/1804.02767>.
- [25] WANG Ao, CHEN Hui, LIU Lihao, et al. Yolov10: real-time end-to-end object detection[EB/OL]. [2024-11-01]. <https://arxiv.org/abs/2405.14458>.

(责任编辑:张艳玲,张 琳)

Moving video-based detection for roadside illegally parking vehicles

TANG Kang^{1,2,3}, SUN Yu^{1,2,3}, ZHONG Xiaoyang^{1,2,3}, GAO Jialiang^{1,2,3}, CHEN Chongcheng^{1,2,3}

1. *Key Lab of Spatial Data Mining and Information Sharing of Ministry of Education, Fuzhou University, Fuzhou 350108, China;*

2. *National Engineering Research Center of Geospatial Information Technology, Fuzhou University, Fuzhou 350108, China;*

3. *Academy of Digital China (Fujian), Fuzhou 350108, China*

Abstract: Roadside parking zones play a significant role in alleviating urban parking pressure. However, with the continuous growth in urban motor vehicle ownership, the supply-demand gap for roadside parking spaces continues to widen, leading to severe illegally parking that significantly impacts traffic efficiency and safety. Existing illegally parking monitoring systems based on fixed-point cameras or sensors suffer from high costs and limited coverage. To address this, this paper proposes a solution for detecting suspected illegally parked vehicles in roadside parking zones using mobile cameras. The solution is developed using embedded devices combined with an improved object detection algorithm (achieving a 3.3% increase in mAP@50). Trained on a custom-built dataset, it enables real-time detection of suspected illegally parked vehicles and effectively supports large-area monitoring tasks for suspected illegal parking. Comparative analysis with simultaneous drone-tracked orthophoto imagery from the same road sections showed that the solution achieves an average precision of 0.87, a recall of 0.88, and a detection speed of 53.96 frames per second (fps), fully validating the feasibility and effectiveness of the proposed method.

Key words: embedded devices; object detection; suspected illegally parked vehicles

Foundation support: The National Natural Science Foundation of China (No. 42171426); Fujian Cultural and Tourism Economy Alliance Science and Technology Innovation Team-Digital Culture and Smart Tourism Fujian Provincial Science and Technology Innovation Team Project; Collaborative Knowledge Graph and Large Language Models for Maritime Shipping Risk Perception and Reasoning-Based Decision-Making (No. 2024KFJJ025)

First author: TANG Kang (2000—), male, postgraduate, majors in digital Earth and smart society.

E-mail: 1635138766@qq.com

Corresponding author: SUN Yu

E-mail: jade.yusun@outlook.com