



基于改进的 YOLOv5 模型和射线法的车辆违停检测

摘要

车辆违法停车将会降低道路通行效率,引发交通拥堵和交通事故.传统的车辆违停检测方法参数量大且准确度低.为此,本文提出了一种使用改进的 YOLOv5 模型和射线法的车辆违停检测方法.首先设计了轻量化的特征提取模块,减少模型参数量;其次在模型中加入注意力机制,从通道维度和空间维度增强模型的特征提取能力,保证模型精度;接着使用混合数据增强丰富数据集样本,提升复杂背景下的检测效果;然后选用 EIoU 作为损失函数提高模型定位能力.实验结果表明,改进后的模型均值平均精度达到 91.35%,比原始 YOLOv5s 提升 1.01 个百分点,并且参数量减少 35.79%.最后将改进后模型与射线法结合,在 Jetson Xavier NX 嵌入式平台的检测速度可以达到约 28 帧/s,能够实现实时检测.

关键词

车辆违停检测;YOLOv5s 算法;Ghost 卷积;注意力机制;射线法

中图分类号 TP391.41

文献标志码 A

收稿日期 2023-04-02

资助项目 国家自然科学基金(62171228);江苏省研究生科研与实践创新计划(SJCX21_0354)

作者简介

庄建军,男,博士,教授,研究方向为人工智能、嵌入式系统. jjzhuang@nuist.edu.cn

徐子恒(通信作者),男,硕士生,研究方向为计算机视觉、目标检测. 2278976145@qq.com

0 引言

随着经济的快速发展,我国汽车产销量也在不断增加.根据公安部 2022 年 12 月的统计数据,全国机动车保有量达 4.15 亿辆,其中汽车保有量为 3.18 亿辆,汽车已成为人们主要的交通出行方式之一.在汽车数量高速增长的同时,城市停车难和交通拥堵问题日趋严重,违章占道停车现象屡见不鲜.违章停车不仅会扰乱道路秩序、降低通行效率,还会带来交通安全隐患.

传统的违章停车检测主要采用地磁技术^[1],通过车辆上的金属进行识别,这种方法简单稳定,但是传感器的检测范围很小,只适用于单个停车位,不满足道路、停车场等大范围检测需求.因此,基于视频流的违停检测开始慢慢普及.早期基于视频流的检测方法主要利用传统图形处理技术.例如:王殿海等^[2]利用混合高斯模型获取背景与运动目标以在场景相对复杂的城市主干道上检测车辆违停;Ye 等^[3]采用自适应阈值和二值化消除背景干扰,再采用形态学提取车辆目标;杨祖萁等^[4]通过混合高斯模型提取出运动目标,利用 Meanshift 算法跟踪目标,最后使用卷积神经网络判断静止目标是否为车辆,该方法与传统方法相比,准确率提升了 63%;唐洁^[5]对设定的禁停区域进行灰度直方图计算,当灰度发生变化且趋于稳定时,用 SSD 算法检测物体是否为车辆,实现指定区域内的违停检测.

近年来,基于深度学习的目标检测算法不断涌现,其中具有代表性的 R-CNN^[6]、YOLO^[7] 系列算法更是被广泛应用,越来越多的研究者开始使用目标检测方法识别车辆,来解决传统方法识别速度慢、易受环境干扰等问题.例如:刘寒迪等^[8]使用对小目标检测效果好的 SSD^[9] 网络检测施工车辆,采用 MobileNet 作为主干网络,在检测速度和精度上都有提高;Tang 等^[10]在 SSD 网络中引入空间注意力模块,提升了模型的检测精度,并通过前后帧间预测框的交并比实现车辆目标跟踪,判断车辆是否停车;丁冰等^[11]采用基于卷积神经网络的 YOLOv3 模型作为车辆检测网络,极大地提高了车辆检测的精度,并利用 DeepSORT 跟踪算法作为停车检测,实现了高速公路隧道中的停车检测任务;许璧麒等^[12]将通道空间注意力模块引入 YOLOv5 网络中,提升小目标检测精度,并通过 DeepSORT 实现车辆的逆行检测.上述方法普遍存在参数量大、实时性差的问题,在实际应用中不适合在性能受限的嵌入式设备上运行.因此,本文提出一种使用改进 YOLOv5 和

1 南京信息工程大学 电子与信息工程学院,南京,210044

射线法的车辆违停检测方法,能够在保证模型精度的同时减少算法复杂度和参数量.主要工作如下:

1) 利用通道混洗和通道复用卷积重构YOLOv5s的特征提取模块,减少模型的参数量;在模型中引入改进后的注意力机制,增强模型特征提取能力;优化损失函数提高模型回归精度,加速模型收敛.

2) 对UA-DETRAC数据集进行混合数据增强,提升样本的复杂度;使用k-means聚类算法在数据集上获取更适合的初始锚框.

3) 通过消融实验、不同模型对比实验,验证模型改进的有效性;将改进后的模型与射线法结合,部署在Jetson Xavier NX平台.

1 理论基础

1.1 YOLOv5 算法原理

YOLO系列是典型的一阶段目标检测网络,其中YOLOv3、YOLOv4、YOLOv5^[13]、YOLOX和YOLOv7应用较为广泛.考虑到网络需要在嵌入式设备下运行,体积超过50MB的YOLOv3、YOLOv4和YOLOv7并不适合本文的实验环境.YOLOv5和

YOLOX则根据模型深度和宽度的不同分为s、m、l、x 4个版本,其中s版本的模型体积最小,推理速度最快.YOLOv5s在COCO数据集上的精度比YOLOXs稍低但不超过1%,模型体积却小了20%,检测速度也更快,因此本文选择YOLOv5s作为基础网络.YOLOv5可分为输入端(input)、主干网络(backbone)、颈部网络(neck)和输出端(head)4个部分,模型网络结构如图1所示.

YOLOv5的输入端先对输入图像进行预处理,然后采用Mosaic数据增强^[14]、自适应锚框计算、自适应图像缩放.主干网络的作用是特征提取,首先使用Focus结构对进入主干网络前的图片进行切片,将输入图片的长宽缩减一半,通道则扩充成4倍.C3模块为跨阶段局部融合网络,能够有效获取特征图特征.SPP是空间金字塔池化,能将任意大小的特征图转变成固定大小的特征向量.Neck采用FPN^[15]和PAN^[16]融合的结构,FPN自顶向下传递高层语义特征,PAN自底向上传递定位特征,将语义特征和定位特征混合,可以进一步提升对密集目标的检测效果.YOLOv5的输出包括80×80、40×40、20×20 3个尺度的特征图,用于检测不同大小的目标.采用CIoU损

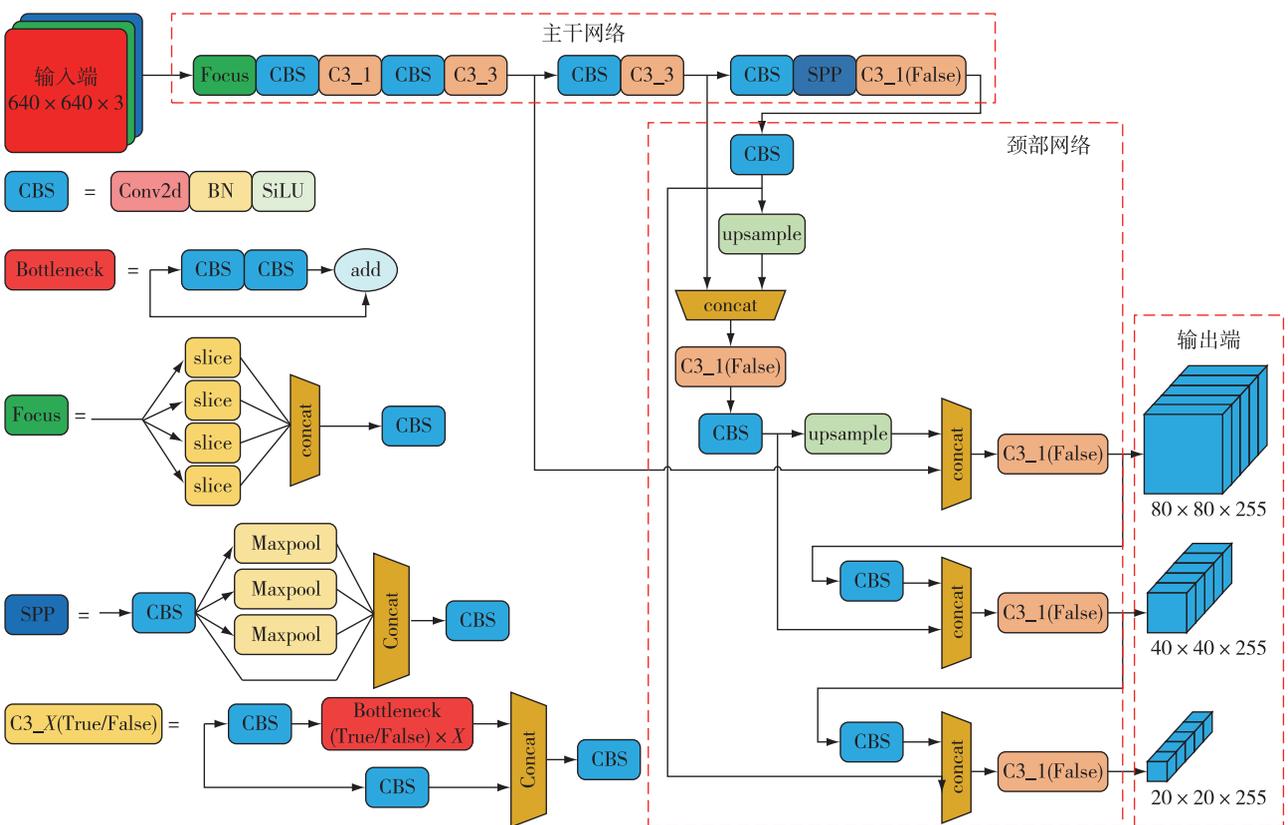


图1 YOLOv5网络结构

Fig. 1 YOLOv5 network structure

失函数来评价预测框和真实框之间的距离,解决了模型训练过程中预测框和真实框不重合时造成的收敛速度慢的问题.使用非极大值抑制(NMS)保留效果最好的预测框并去除冗余的预测框,增强模型对有遮挡目标的识别能力,从而完成目标检测过程.

1.2 基于射线法的区域入侵

考虑到在不同的地点禁停区域拥有不同的形状,以及不同距离产生的透视形变,摄像头图像中的区域通常由多边形构成,因此无法像矩形框通过几何运算快速获得交并比(IOU)一样,以此判断车辆是否在禁停区域内.本文以车辆外接矩形的中心点作为车辆的位置,根据中心点是否在多边形内判断车辆是否违停.此方法从车辆质心处向任意方向发射一条射线,根据射线穿过多边形边界的次数判断车辆质心是否在多边形内.射线法效果如图 2 所示,当射线穿过边界次数为奇数时说明质心在多边形内部,反之在多边形外部.

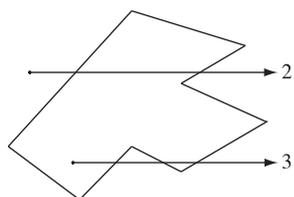


图 2 射线法示意

Fig. 2 Schematic of ray method

2 算法的改进

2.1 YOLOv5 改进

2.1.1 轻量化特征提取模块

C3 是 YOLOv5 中的特征提取模块,在网络中多次使用,但是在关键特征提取过程中容易产生很多冗余特征,使得网络模型整体参数多、计算量大.为了降低模型的计算复杂度、提升实时性,本文引入 Ghost 卷积模块^[17],其结构如图 3 所示.Ghost 卷积模块的输出分为两个部分,第一部分由输入特征图通过普通卷积生成,第二部分则是对第一部分的输出使用线性计算得出.将两部分在通道维度叠加组成整个模块的输出,通过线性计算生成特征图,避免重复特征被再次送入网络中进行卷积计算,减少了模型的计算量.

利用 Ghost 卷积模块代替 Bottleneck 中的普通卷积,搭建 Ghostunit 模块,其结构如图 4 所示,降低了模型的参数量和计算量.与原始网络中 Bottleneck

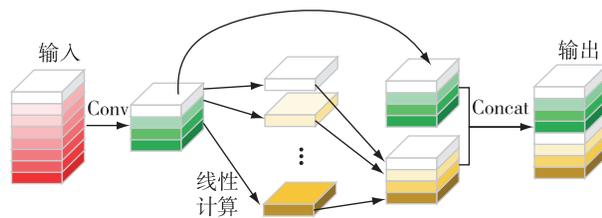


图 3 Ghost 卷积结构

Fig. 3 Ghost convolution structure

不同的是,第一次 Ghost 卷积模块不再减少输入的通道数,而是将通道数提升一倍,第二个 Ghost 卷积模块再将通道数恢复到输入数量,这是因为模块中使用了深度可分离卷积,在参数下降的同时提取的特征也会减少,因此采用先升维再降维的方式增强特征提取能力.

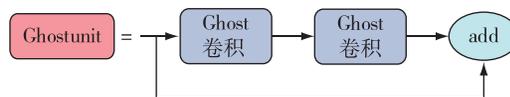


图 4 Ghostunit 结构

Fig. 4 Ghostunit structure

以 C3 结构为基础,设计了 SG-CSP (Shuffle Ghost- Cross Stage Partial) 轻量化特征提取模块,结构如图 5 所示,使用 Ghostunit 替代 Bottleneck,并用 Channel Split 和 concat 操作代替一维卷积实现通道的增减.Channel Split 是将特征图从通道维度一分为二,concat 是将两个分支的特征图在通道维度合并.由于输出是两个部分叠加而成,因此在 concat 后加入 Channel Shuffle,该模块将特征图的通道维度拆成二维,执行矩阵转置后再合并为一维,能够均匀打乱两个分支的输出,保证通道间的信息交流.

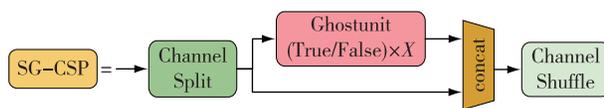


图 5 SG-CSP 结构

Fig. 5 SG-CSP structure

2.1.2 融合注意力机制

注意力机制是人们在机器学习模型中嵌入的一种特殊结构,用来学习和计算输入数据对输出数据的贡献大小.与人眼获取图像信息的方式类似,注意力机制通过对不同部分的特征图赋予权重,使网络更容易关注图像中重要信息并抑制无用信息,以选择出图片中对目标检测更有利的特征.

CBAM(Convolutional Block Attention Module)注意力机制^[18]是一种结合通道和空间的注意力模块,包含两个子模块:通道注意力模块(Channel Attention Module)和空间注意力模块(Spatial Attention Module),分别进行通道和空间上的注意力操作,这使得CBAM能够作为即插即用的模块快速集成到现有的网络架构中.CBAM模块的结构如图6所示.

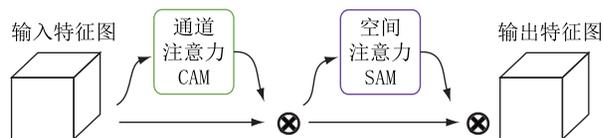


图6 CBAM结构
Fig. 6 CBAM structure

通道注意力模块的结构如图7所示.首先分别经过基于特征图长宽的全局最大池化(Global Max Pooling)和全局平均池化(Global Average Pooling),得到两个 $1 \times 1 \times C$ 的特征图.接着将它们送入两层的全连接网络中,第一层网络的神经元个数为 C/r (r 为减少率),可以减少模块计算量,第二层网络的神经元个数为 C ,两层之间使用ReLU激活函数激活.最后将全连接层输出的两个特征的对应元素加和,再经过sigmoid激活函数生成通道注意力权重.将权重和输入特征图对应元素相乘得到通道注意力特征图并作为空间注意力模块的输入.

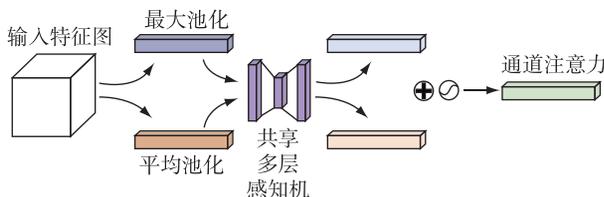


图7 通道注意力模块结构
Fig. 7 Channel attention module structure

空间注意力模块的结构如图8所示.首先做基于特征图通道的全局最大池化和全局平均池化,得到两个 $H \times W \times 1$ 的特征图.将两个特征图进行通道维度的叠加操作形成 $H \times W \times 2$ 的特征图,经过一个 7×7 的卷积降维成 $H \times W \times 1$.最后经过sigmoid生成空间注意力权重,将权重和输入特征图相乘,得到最终的特征图.

CBAM模块采用先通道后空间的形式,相比只采用通道注意力的SE更能够帮助网络获取有利的特征.但是模块采用了串行结构,使得通道信息和空

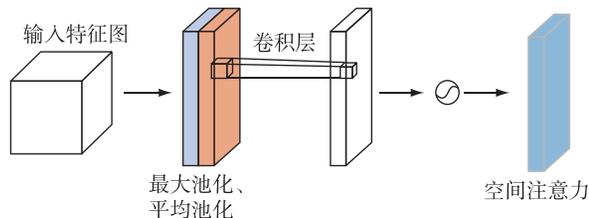


图8 空间注意力模块结构
Fig. 8 Spatial attention module structure

间信息起到的作用相同,而事实上,网络中不同层的特征图对通道和空间信息有不同的关注度.因此,本文改进了CBAM模块并将其命名为P-CBAM(Parallel-Convolutional Block Attention Module),模块结构如图9所示.

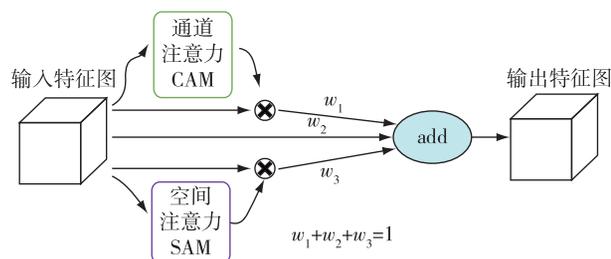


图9 P-CBAM结构
Fig. 9 P-CBAM structure

特征图输入模块后分别送入3个分支,前2个分支分别送入通道注意力模块和空间注意力模块,以获得通道特征图 and 空间特征图,最后一个分支保持不变.3个输出各自乘上一个归一化后的可训练权重系数后按照对应元素相加得到输出.

2.1.3 损失函数优化

YOLOv5的损失函数包含回归损失、分类损失和置信度损失,其中IOU损失用于衡量真实框和预测框之间的距离,是回归损失中的重要部分.YOLOv5采用的CIoU损失函数考虑了边界框回归的重叠面积、中心点距离和宽高比.CIoU损失函数的公式如下:

$$s_{\text{IOU}} = \frac{A \cap B}{A \cup B}, \quad (1)$$

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

$$\alpha = \frac{v}{(v - s_{\text{IOU}}) + v}, \quad (3)$$

$$L_{\text{CIoU}} = 1 - s_{\text{IOU}} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v. \quad (4)$$

其中: v 用来衡量预测框和真实框宽高之间比例的一

致性; w 和 h 分别表示预测框的宽和高; w^{gt} 和 h^{gt} 分别表示真实框的宽和高; α 是用于平衡比例的参数; b 和 b^{gt} 分别表示预测框和真实框的中心点坐标; $\rho(\cdot)$ 为计算欧式距离; c 表示预测框和真实框最小外接矩形的对角线距离. 从 CIUO 的损失函数公式可以看出, 它包含了预测框和真实框重叠区域面积的 IOU 损失、两框中心点之间的距离损失和两框的宽高比损失. CIUO 损失函数通过迭代可以将预测框不断向真实框移动, 还可以保证预测框和真实框的宽高比更为接近.

尽管 CIUO 克服了 IOU 的缺点, 但是在宽高比的描述上, CIUO 使用的是相对值, 难以反映两框之间的真实差异, 有时候会阻碍模型的有效优化. 因此, 本文将 YOLOv5 的损失函数更新为 EIOU. EIOU 在 CIUO 惩罚项的基础上将宽高比的影响因子拆开, 利用预测框和目标框的长和宽分别计算损失, 明确地衡量了重叠区域、中心点距离和边长这 3 个几何因素的差异. EIOU 损失的公式如下:

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - s_{IOU} + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2}. \quad (5)$$

其中: C_w 和 C_h 分别是预测框和目标框最小外接矩形的宽和高. 将宽高比的损失项拆分成宽和高进行独立计算, 能够提高模型回归精度.

2.2 射线法优化

射线法是根据射线穿过多边形边界的次数来判断质心与多边形的相对位置的, 而顶点本身就处于多条边上, 当质心是多边形的某个顶点时, 顶点本身所处边的数量将算作射线穿过边界的次数从而影响结果, 如图 10a 所示. 因此, 在使用射线法前应判断质心是否在顶点上, 若符合则可以说明车辆已经进入禁停区域. 同理, 质心在多边形的边上也会影响结果, 如图 10b 所示, 提前判断点在边上也可说明车辆已经进入禁停区域.

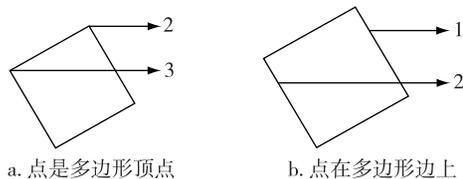


图 10 特殊情况 1

Fig. 10 Special case 1

当质心不在多边形的顶点和边上时, 才会使用射线法进行判别, 而此时又有两种奇异情况需要考虑. 第一种是射线与多边形的某一条边重叠, 如图 11a 所示, 射线经过重叠边时会额外增加一次边界穿越次数, 这影响了最终结果, 因此在使用射线法时, 应当忽略多边形中与射线平行的边. 第二种是射线会经过多边形的顶点, 如图 11b 所示. 因此, 根据边的折叠方式采取了两种处理方式, 当顶点所在的边位于射线两侧, 则增加一次边界穿越次数, 若位于射线同侧, 则不增加穿越次数.

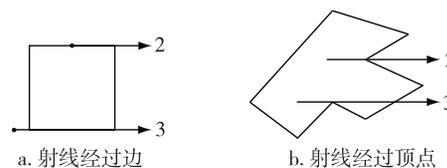


图 11 特殊情况 2

Fig. 11 Special case 2

改进后的射线法能够适用于凸多边形、凹多边形、不规则多边形等各类形状, 满足车辆违停判断的需求.

3 实验与分析

3.1 数据集的制作和前处理

本文实验的数据集是 UA-DETRAC 车辆数据集, 该数据集包含 24 个不同地点, 超过 14 万帧画面, 8 250 个手动标注的车辆, 总共 121 万目标对象外框, 满足车辆检测的要求, 因此选取其中 9 600 张样本进行实验. 数据集中部分样本如图 12 所示.

数据集制作好后, 将对其使用 Mosaic 加 Mixup 的混合数据增强. Mosaic 数据增强是把 4 张图片通过随机缩放、剪裁、排布、色域变化等方式进行拼接, 并按照 4 个方向进行图片和框的组合, 数据增强效果如图 13a 所示. Mixup 数据增强则将随机的 2 张样本按比例混合, 同时样本的标签也按比例混合, 效果如图 13b 所示, 通过 2 个随机数据进行正负样本融合成新的样本, 能够使数据集数量翻倍, 并且生成的样本中含有更多的目标物. 通过混合数据增强使图像远远脱离自然图片的真实分布, 能够提升模型在多目标和复杂背景下的检测效果.

3.2 锚框值设置

在目标检测任务中, 会提前生成一些先验信息来约束预测框的范围大小, 通过这些先验框信息大致确定预测框的范围. 本文的数据集目标尺寸偏小,

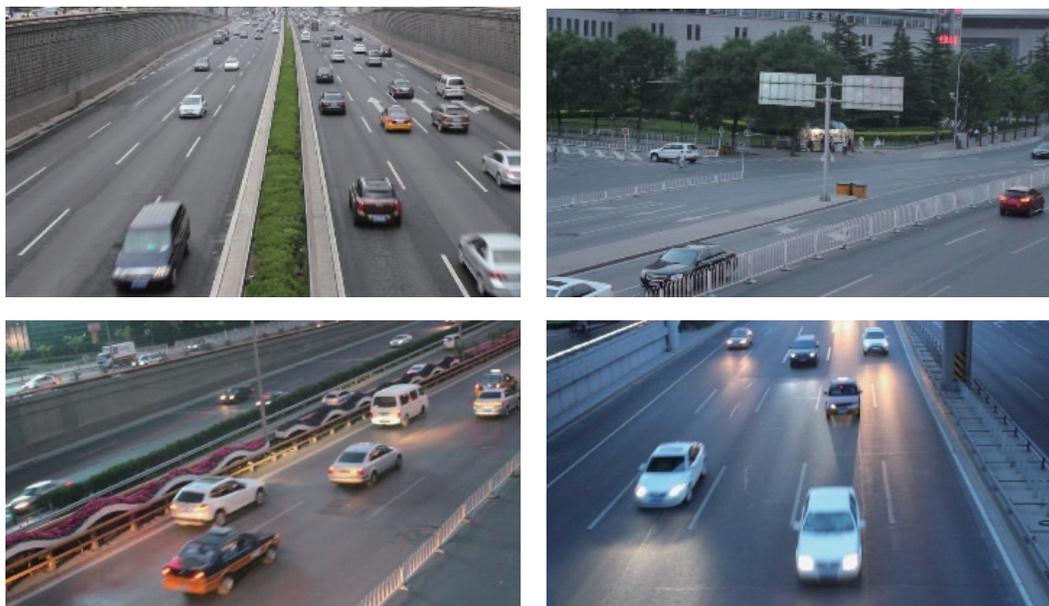
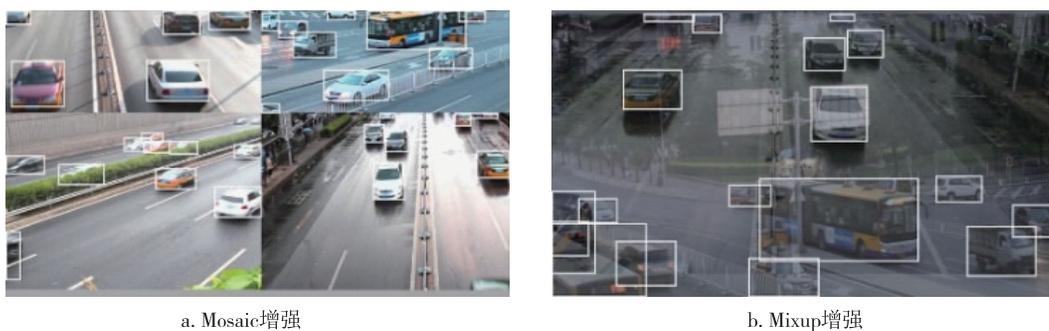


图 12 数据集样本
Fig. 12 Dataset samples



a. Mosaic增强

b. Mixup增强

图 13 数据增强效果
Fig. 13 Data enhancement effect

不适合原始网络在 COCO 数据集上形成的锚框,因此使用 k -means 聚类的方式在数据集上重获锚框值,锚框数据如表 1 所示.

表 1 锚框值

Table 1 Anchor frame values

特征层	锚框值
小目标	(17,26) ,(24,35) ,(34,46)
中目标	(42,66) ,(57,53) ,(59,97)
大目标	(86,74) ,(99,126) ,(144,190)

3.3 实验平台和评价指标

为了对比 YOLOv5s 在改进前后的性能差异,需要在硬件平台上训练和验证,此过程会产生巨大的计算量,对硬件的性能需求非常高,因此本文使用带有 NVIDIA 高性能显卡的主机作为模型训练和测试

的平台,并选择 Jetson Xavier NX 作为模型部署的嵌入式平台,主机端和嵌入式端的配置参数如表 2 所示.

表 2 主机端和嵌入式端配置参数

Table 2 Configuration parameters of host side and embedded side

软硬件配置	主机端	嵌入式端
CPU	i5-10400f	Carmel ARM
GPU	NVIDIA 3070ti	NVIDIA Volta
内存/GB	16	16
操作系统	Windows10	Ubuntu18.04
CUDA	11.5	10.2
深度学习框架	Pytorch1.11.0	Pytorch1.10.0
开发工具	Python3.7	Python3.6

在搭建好网络模型后,需要将不同的网络模型

在数据集上进行训练.为了保证所有网络模型训练结果的公平性,将统一设置所有模型的超参数,具体设置如表 3 所示.

表 3 模型超参数设置

Table 3 Model hyperparameter settings

参数类型	参数值
输入尺寸	640×640
优化器	Adam
初始学习率	0.001
学习率调整	余弦退火
batch_size	8
epoch	250

为了定量评价实验结果,本文使用精确度 (Precision, P)、召回率 (Recall, R) 和均值平均精度 (mAP, mean Average Precision) 来作为模型准确性能的评价指标.精确度是指分类器判定目标为正样本并且判断正确的部分占有所有分类器判定为正样本的比例;召回率是指模型预测为正样本并判断正确的部分占有所有正样本的比例;平均精度指的是以召回率为横轴,精确度为纵轴组成的曲线,而曲线所围成的面积就是平均精度;均值平均精度是数据集中所有类别的平均精度的均值,在本文实验中采用的 mAP 指标是指 PASVAL VOC 数据集中所使用的 $mAP_{0.5}$,即 IOU 阈值设定为 0.5 时所有类别 AP 的平均值.上述评价指标计算公式如下:

$$P = \frac{n_{TP}}{n_{TP} + n_{FP}}, \quad (6)$$

$$R = \frac{n_{TP}}{n_{TP} + n_{FN}}, \quad (7)$$

$$P_{AP} = \int_0^1 P(R) dR, \quad (8)$$

$$P_{mAP} = \frac{1}{N} \sum_{i=1}^N P_{AP,i}. \quad (9)$$

模型在轻量化方面,使用参数量和权重大小作为模型大小的评价指标,使用每秒可处理的图片数

量 (FPS) 作为模型检测速度的评价指标.

3.4 实验结果分析

在验证改进后的完整模型之前,可以通过消融实验逐步探究所提改进方法对原始模型的影响.消融实验将 SG-CSP 模块、P-CBAM 模块、Mixup 增强和 EIOU 损失函数逐步加入到 YOLOv5s 模型中,结果如表 4 所示.由表 4 可知,SG-CSP 模块替换 C3 模块后,由于参数量的减少,不利于模型在复杂环境下提取车辆特征,模型的 mAP 值下降了 2.03 个百分点.为了保持模型的特征提取能力,在模型的 backbone 和 neck 之间加 P-CBAM 注意力模块,使得模型能够自主选择空间信息和通道信息,与使用 SG-CSP 的模型相比 mAP 提升了 2.06 个百分点,也稍高于基准模型.训练阶段,在 Mosaic 数据增强基础上额外使用 Mixup 增强,并用 EIOU 代替 CIOU 计算交并比损失,与基准模型相比 mAP 分别提升 0.57 和 1.01 个百分点.最终模型的 mAP 值对比基准网络都有更好的表现,并且参数量减少 35.78%,模型大小下降 36.16%.

根据消融实验结果不难发现,使用注意力机制有效地弥补了轻量化带来的负面影响.为了验证 P-CBAM 注意力模块相较于其他注意力模块有更好的表现,以轻量化后的 YOLOv5s 模型作为基准网络,同样在 backbone 和 neck 之间的位置引入不同的注意力模块,使用 mAP、模型体积、参数量作为评价指标,对比结果如表 5 所示.由表 5 可知:使用 SE (Squeeze-and-Excitation) 模块对比基准模型 mAP 值几乎没有变化,模型大小和参数却上升最多;使用 ECA (Efficient Channel Attention) 后模型 mAP 值提升了 1.12 个百分点,对模型大小和参数量影响最小;CBAM 和 P-CBAM 都采用空间注意力模块和通道注意力模块,性能提升效果明显.由于 P-CBAM 增加了自适应权重使模型能够在通道特征和空间特征之间做选择,mAP 值对比使用 CBAM 提升了 0.35 个百分点,模型大小和参数量对比基准模型有些许增加,但是 mAP 值提升了 2.06 个百分点,对比使用

表 4 消融实验

Table 4 Comparison of ablation experiment

模型	mAP/%	模型大小/MB	参数量/ 10^6	FPS/(帧/s)
YOLOv5s	90.34	27.16	7.07	63
YOLOv5s+SG-CSP	88.31	17.17	4.46	68
YOLOv5s+SG-CSP+P-CBAM	90.37	17.34	4.54	66
YOLOv5s+SG-CSP+P-CBAM+Mixup	90.91	17.34	4.54	66
YOLOv5s+SG-CSP+P-CBAM+Mixup+EIOU	91.35	17.34	4.54	66

ECA还提升了0.94个百分点.实验结果表明,P-CBAM注意力模块能够有效提升轻量化模型的特征提取能力,弥补轻量化带来的负面影响.

表5 不同注意力模块对照实验

Table 5 Comparative experiment of different attention modules

基准模型	注意力	mAP/%	模型大小/MB	参数量/ 10^6
YOLOv5s +SG-CSP	None	88.31	17.17	4.46
	SE	88.32	17.84	4.64
	ECA	89.43	17.17	4.46
	CBAM	90.02	17.34	4.54
	P-CBAM	90.37	17.34	4.54

为了进一步验证改进后的目标检测网络在车辆检测方面有更好的优势,将改进后的网络模型与多个现阶段常用的YOLO系列目标检测网络进行对比,实验结果如表6所示.其中:Ghost-v5s和Mobile-v5s是在YOLOv5s的基础上,分别将主干网络替换为GhostNet和MobileNetV3网络;YOLOv5s、YOLOXs和YOLOv7为非轻量化模型;Ghost-v5s、Mobile-v5s和YOLOv4-tiny为轻量化模型.

首先在模型轻量化方面,本文采用的算法以YOLOv5网络模型为基础,融合通道分割和轻量化卷积,有效地减少了模型体积和参数量,由表6可以看出,本文算法的参数量和模型体积不仅小于非轻量化模型,还小于YOLOv4-tiny之类的轻量化模型.

其次在性能指标方面,本文算法在P-CBAM注意力机制、Mixup数据增强和EIOU损失的加持下,mAP值仅低于YOLOv7,但模型大小只有YOLOv7的12%,参数量只有YOLOv7的16%.在召回率方面,本文算法弱于YOLOXs和YOLOv7,但在精确度方面,本文算法优于所有对比算法.

最后在处理速度方面,YOLOv4-tiny以142帧/s的处理速度远超其他网络模型,但是召回率太低,导致mAP值明显低于其他模型.本文算法达到了66帧/s,比原版YOLOv5s的处理速度稍快,但本文算法

在其他方面对比YOLOv5s表现均更好.

为了验证本文算法在实际情况下检测速度,选择有遮挡、光线差和模糊的图像,同时对YOLOv5s和本文算法进行测试.置信度均设置为0.5,非极大值抑制的IOU阈值设置为0.3.图14a、b分别为YOLOv5s和本文算法的检测结果,可知本文算法的检测结果均正确,对于YOLOv5s漏检或错检的目标,本文算法依然可以正确检测.

3.5 违停检测算法部署与应用

基于实际应用性能考虑,本文选用的硬件平台为Jetson Xavier NX(以下简称为NX).NX是NVIDIA发布的一款用于AI智能计算的高性能终端,它包含非常丰富的硬件资源,适合目标检测模型的部署.本文采用640×640大小的图片作为输入,在NX上的检测速度可以达到约28帧/s,能够满足实时检测的需求.将车辆检测算法和射线法结合,对比人为设定好的禁停区域,能够快速并准确地进行违停检测,多个应用场景的效果展示如图15所示.

4 结论

为了解决车辆违停检测算法复杂、模型大的问题,基于YOLOv5s网络提出一种轻量化的车辆违停检测方法.通过构建SG-CSP模块减少模型的参数;引入P-CBAM注意力机制,增强主干网络的特征提取能力;训练阶段运用混合数据增强并使用EIOU损失代替CIOU损失,提升模型精度.实验结果表明,改进后的YOLOv5s在参数量减少35.78%的同时精确度提升1.01个百分点.将本文提出的轻量化模型与射线法结合并部署在嵌入式平台Jetson Xavier NX上,可以实现约28帧/s的检测速度,达到实时性的需求.

未来的研究中,将着力解决数据集样本不平衡时,因样本数量较少导致某一类别召回率偏低的问题,如在损失函数中使用权值均衡等方法进一步优化模型,提升模型精度.

表6 不同注意力模块对照实验

Table 6 Comparative experiment of different target detection models

模型	精确度/%	召回率/%	mAP/%	模型大小/MB	参数量/ 10^6	FPS/(帧/s)
YOLOv5s	88.03	87.51	90.34	27.16	7.07	63
YOLOXs	85.26	89.47	90.98	34.35	8.94	55
YOLOv7	87.50	89.78	91.94	142.30	27.21	44
Ghost-v5s	88.24	87.69	90.10	23.92	6.18	44
Mobile-v5s	88.37	85.97	88.00	24.17	6.48	51
YOLOv4-tiny	85.69	67.54	72.77	22.49	5.88	142
Our-Model	89.18	87.78	91.35	17.34	4.54	66

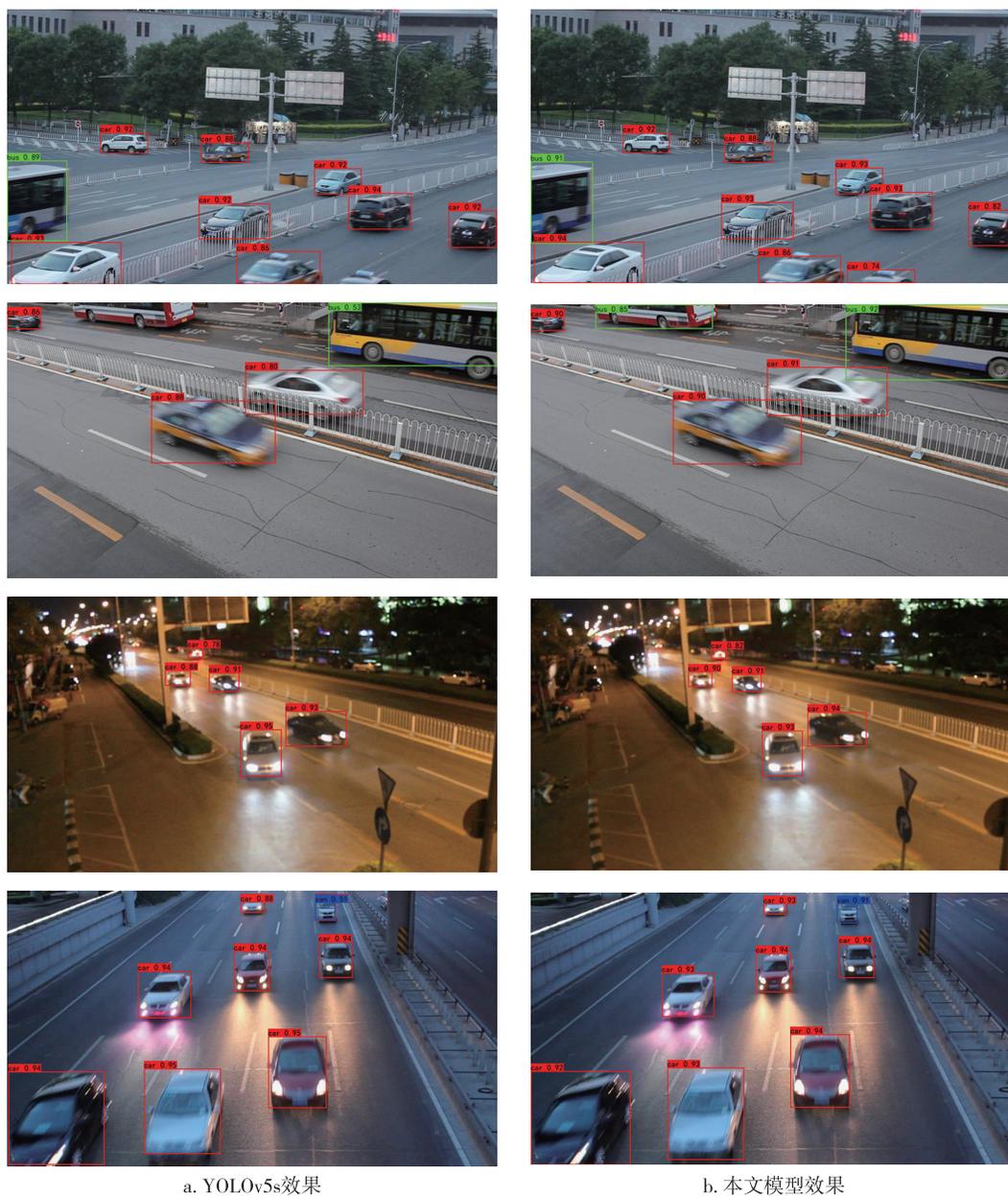


图 14 模型检测效果对比

Fig. 14 Comparison of model detection performance



图 15 嵌入式平台违停检测效果

Fig. 15 Rendering of illegal parking detection on embedded platform

参考文献

References

- [1] 张增超,李强,孙红雨,等.基于地磁传感器和UWB技术的停车位车辆检测方法与其实现[J].传感技术学报,2019,32(12):1917-1922
ZHANG Zengchao, LI Qiang, SUN Hongyu, et al. Parking vehicle detection method and implementation based on geomagnetic sensor and UWB technology [J]. Chinese Journal of Sensors and Actuators, 2019, 32 (12): 1917-1922
- [2] 王殿海,胡宏宇,李志慧,等.违章停车检测与识别算法[J].吉林大学学报(工学版),2010,40(1):42-46
WANG Dianhai, HU Hongyu, LI Zhihui, et al. Detection and recognition algorithm of illegal parking [J]. Journal of Jilin University (Engineering and Technology Edition), 2010, 40(1) : 42-46
- [3] Ye Q, Zhang Y M, Ma L. Video-based detection method of illegal parking vehicles [J]. Energy Procedia, 2011, 11 : 4949-4957
- [4] 杨祖蓂,丁洁,刘晋峰.一种新的结合卷积神经网络的隧道内停车检测方法[J].重庆大学学报,2021,44(6):49-59
YANG Zuliang, DING Jie, LIU Jinfeng. A new tunnel vehicle stopping detection methodology combined with convolutional neural network [J]. Journal of Chongqing University (Natural Science Edition), 2021, 44(6) : 49-59
- [5] 唐洁.基于改进的卷积神经网络的违章停车检测[D].合肥:安徽大学,2019
TANG Jie. Illegal parking detection based on improved convolution neural network [D]. Hefei: Anhui University, 2019
- [6] 梅舒欢,闵巍庆,刘林虎,等.基于Faster R-CNN的食品图像检索和分类[J].南京信息工程大学学报(自然科学版),2017,9(6):635-641
MEI Shuhuan, MIN Weiqing, LIU Linhu, et al. Faster R-CNN based food image retrieval and classification [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2017, 9 (6): 635-641
- [7] 温秀兰,焦良葆,李子康,等.复杂环境下小尺度烟火目标检测研究[J].南京信息工程大学学报(自然科学版),2023,15(6):676-683
WEN Xiulan, JIAO Liangbao, LI Zikang, et al. Small scale smoke & fire target detection in complex environment [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(6) : 676-683
- [8] 刘寒迪,赵德群,陈星辉,等.基于改进SSD的航拍施工车辆检测识别系统设计[J].国外电子测量技术,2020,39(7):127-132
LIU Handi, ZHAO Dequn, CHEN Xinghui, et al. Design of detection and recognition system of aerial photography construction vehicle based on improved SSD [J]. Foreign Electronic Measurement Technology, 2020, 39 (7): 127-132
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [C] // European Conference on Computer Vision. Cham: Springer, 2016: 21-37
- [10] Tang H R, Peng A M, Zhang D M, et al. SSD real-time illegal parking detection based on contextual information transmission [J]. Computers, Materials & Continua, 2020, 62(1) : 293-307
- [11] 丁冰,杨祖蓂,丁洁,等.基于改进YOLOv3的高速公路隧道内停车检测方法[J].计算机工程与应用,2021,57(23):234-239
DING Bing, YANG Zuliang, DING Jie, et al. Detection method of highway tunnel vehicle stopping based on improved YOLOv3 [J]. Computer Engineering and Applications, 2021, 57(23) : 234-239
- [12] 许璧麟,马志强,宝财吉拉呼,等.基于YOLOv5的高速公路小目标车辆逆行检测模型[J].国外电子测量技术,2022,41(11):146-153
XU Biqi, MA Zhiqiang, BAO Caijilahu, et al. Small target vehicle in highway retrograde detection model based on YOLOv5 [J]. Foreign Electronic Measurement Technology, 2022, 41(11) : 146-153
- [13] 庄建军,叶振兴.基于改进YOLOv5m的电动车骑行者头盔与车牌检测方法[J].南京信息工程大学学报,2024,16(1):1-10
ZHUANG Jianjun, YE Zhenxing. Helmet and license plate detection for electric bike rider based on improved YOLOv5m [J]. Journal of Nanjing University of Information Science & Technology, 2024, 16(1) : 1-10
- [14] 韩泽山,杨红柳,赵建光,等.基于改进YOLOv5s的口罩佩戴检测算法研究[J].信息技术与信息化,2023(2):22-25
HAN Zeshan, YANG Hongliu, ZHAO Jianguang, et al. Research on mask wearing detection algorithm based on improved YOLOv5s [J]. Information Technology and Informatization, 2023(2) : 22-25
- [15] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21 - 26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944
- [16] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18 - 23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 8759-8768
- [17] 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). June 13 - 19, 2020, Seattle, WA, USA. IEEE, 2020: 1577-1586
- [18] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [C] // European Conference on Computer Vision. Cham: Springer, 2018: 3-19

Illegal parking detection based on improved YOLOv5 model and ray method

ZHUANG Jianjun¹ XU Ziheng¹ ZHANG Ruoyu¹

¹ School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

Abstract Illegally parked vehicles reduce road traffic efficiency, and cause traffic congestion even traffic accidents. Traditional vehicle detection methods are perplexed by a large number of parameters and low accuracy. Here, we propose a method using the improved YOLOv5 model and ray method to detect illegally parked vehicles. First, a lightweight feature extraction module is designed to reduce the amount of model parameters. Second, the attention mechanism is added to the model to enhance its feature extraction ability from both channel dimension and spatial dimension to ensure the model's accuracy. Then, the mixed data is used to enhance and enrich the dataset samples thus improve the detection performance in complex backgrounds, and EIoU is selected as the loss function to improve the model's positioning performance. Experiments show that the mean accuracy of the improved YOLOv5 model reaches 91.35%, which is 1.01 percentage points higher than that of the original YOLOv5s, and the number of parameters is reduced by 35.79%. Finally, the improved YOLOv5 model is combined with the ray method, which can reach real-time inspection speed of 28 frames per second on the embedded platform of Jetson Xavier NX.

Key words illegal parking detection; YOLOv5; Ghost convolution; attention mechanism; ray method