

DOI:10.19651/j.cnki.emt.2417400

# 改进 YOLOv8s 的校园智能清扫车障碍物 检测与测距算法<sup>\*</sup>

郭志军<sup>1,2</sup> 叶世文<sup>1</sup> 庞明天<sup>1</sup> 王丁健<sup>1</sup> 杜林林<sup>1</sup>

(1. 河南科技大学车辆与交通工程学院 洛阳 471000; 2. 河南科技大学智能农业装备全国重点实验室 洛阳 471000)

**摘要:** 针对校园智能清扫车障碍物检测精确度低、检测速度慢以及模型复杂度高的问题,提出一种改进 YOLOv8s 的校园智能清扫车障碍物检测与测距算法 YOLOv8s-FDR。在 YOLOv8s 算法框架的基础上,将主干网络替换为参数量和内存访问量更小的 FasterNet 网络,以降低模型复杂度并提高检测速度;然后设计了 SPPF-DAM 模块,以残差方式引入可变形注意力机制,提高模型对多尺度目标特征的感知能力;其次,在特征融合网络中采用 Partial-RFEM 进行下采样,以捕获非局部上下文特征和局部目标特征,提高检测精确度;最后,添加了测距功能,降低硬件成本。实验结果表明,改进算法与原算法相比 mAP 提高了 3.6%,模型计算量和参数量相较于原模型分别降低了 19.72% 和 15.27%。实际环境测试显示,YOLOv8s-FDR 算法的检测速度达到 38.44 fps,远高于原算法的 17.12 fps,能够满足校园智能清扫车正常运行的性能要求。

**关键词:** 校园智能清扫车;YOLOv8s;可变形注意力机制;性能测试

**中图分类号:** TP391.4;TN914 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## Improve the obstacle detection and ranging algorithm of YOLOv8s campus intelligent sweeper

Guo Zhijun<sup>1,2</sup> Ye Shiwen<sup>1</sup> Pang Mingtian<sup>1</sup> Wang Dingjian<sup>1</sup> Du Linlin<sup>1</sup>

(1. College of Vehicle and Traffic Engineering, Henan University of Science and Technology, Luoyang 471000, China;

2. State Key Laboratory of Intelligent Agricultural Power Equipment, Henan University of  
Science and Technology, Luoyang 471000, China)

**Abstract:** In order to solve the problems of low accuracy, slow detection speed and high model complexity of obstacle detection of campus intelligent sweepers, a modified YOLOv8s obstacle detection and ranging algorithm for campus intelligent sweepers YOLOv8s-FDR was proposed. On the basis of the YOLOv8s algorithm framework, the backbone network is replaced by the FasterNet network with smaller parameters and memory access, so as to reduce the complexity of the model and improve the detection speed. Then, the SPPF-DAM module was designed to introduce the deformable attention mechanism (DAM) in the form of residuals to improve the model's perception of multi-scale target features. Secondly, Partial-RFEM was used for downsampling in the feature fusion network to capture the non-local context features and local target features to improve the detection accuracy. Finally, the ranging function is added to reduce hardware costs. Experimental results show that compared with the original algorithm, the mAP of the improved algorithm is increased by 3.6%, and the amount of model computation and parameters is reduced by 19.72% and 15.27%, respectively. The actual environment test shows that the detection speed of the YOLOv8s-FDR algorithm reaches 38.44 fps, which is much higher than the 17.12 fps of the original algorithm, which can meet the performance requirements of the normal operation of the campus intelligent sweeper.

**Keywords:** campus intelligent sweeper; YOLOv8s; deformable attention mechanisms; performance testing

## 0 引言

随着社会不断发展,各类智能装备相继出现,其中智能

清扫车作为一种新兴的智能装备,正逐渐成为清洁作业的理想选择<sup>[1]</sup>。然而,复杂的作业环境和不断变化的道路信息严重影响了障碍物识别的精度和效率。因此,在校园道

收稿日期:2024-11-19

<sup>\*</sup> 基金项目:国家自然科学基金(51675163)项目资助

路环境下,实现智能清扫车对障碍物的快速、精准识别成为视觉系统研究的关键<sup>[2-3]</sup>。

近年来,随着计算机硬件和机器视觉技术的飞速发展,基于深度学习的目标检测方法逐渐涌现<sup>[4]</sup>。基于深度学习的目标检测方法可分为基于区域候选框的方法和回归方法<sup>[5]</sup>。其中,基于区域候选框的方法通常采用两个阶段:第一阶段通过选择性搜索或区域建议网络生成候选区域,第二阶段则对候选区域进行目标分类和位置回归,这类算法主要包括 R-CNN<sup>[6]</sup>和 Faster R-CNN<sup>[7]</sup>等。基于回归的目标检测方法是单阶段的,此类算法在推理速度上有显著优势,但精度相对较低,代表性算法有 SSD(single shot multibox detector)<sup>[8]</sup>和 YOLO(you only look once)<sup>[9]</sup>等。相较于传统目标检测算法,YOLO 算法自动提取图像的低级和高级特征,具有高度自适应性。高德勇等<sup>[10]</sup>以 YOLOv8 为基础模型,构建了 C2F-DBB 模块,提出了 PA-FPN 的特征融合方式,大幅提高了多尺度道路目标检测精度和速度,但检测精度有待提高。Xiong 等<sup>[11]</sup>基于 YOLOv5 提出了一种具有自适应特征融合和改进注意力机制的小目标检测算法,在特征提取阶段融入 SoftPool,从而减少标准下采样技术固有的小目标关键边缘信息的丢失,但在检测速度方面存在不足。胡伟超等<sup>[12]</sup>基于 YOLOv8 增加小目标检测头,优化特征层之间的信息融合,提高了模型对复杂场景密集人群的检测精度和对不同尺度目标的适应性,但模型复杂度过高。刘罡等<sup>[13]</sup>基于 YOLOv8 对感受野注意力卷积进行轻量化改进,引入轻量化点采样算子,提高了对复杂道路场景的适应能力,但对动态障碍物多尺度目标的适应性不足,检测精度仍有提高空间。

通过对上述研究内容分析,以上方法主要集中研究模型检测精度,模型复杂程度和计算量不断增加。复杂模型难以在计算资源有限的设备上流畅运行,限制了实际应用。因此,论文针对行人、汽车和路障等常见校园道路障碍物为视觉检测对象,并针对校园智能清扫车障碍物检测精确度低、检测速度慢以及模型复杂度高的问题,提出了一种面向校园道路障碍物视觉检测的 YOLOv8s-FDR 算法,能够有效提升校园道路环境中障碍物的多尺度特征的提取,均衡检测速度和精确度。主要包括如下贡献:首先在校园内使用相机在不同时间段、天气条件以及光照条件下对校园内常见障碍物进行了采集并对采集到的图像进行了标注,主要包括行人、汽车和路障等;然后利用 FasterNet 替换网络主干,降低模型的复杂度和参数量,提升算法的检测速度,使其更适用于实时场景;其次设计 SPPF-DAM 模块减少环境背景噪声对目标检测的干扰,增强模型对动态障碍物目标特征的感知能力,提高复杂场景下的检测性能;接着利用 Partial-RFEM 模块通过多次特征通道的拼接,有效捕获非局部上下文特征和局部目标特征,从而解决目标局部语义缺失的问题,额外的特征提取层用于增强特征融合,提高障

碍物检测精度;最后在改进的算法中添加了测距功能,使得在无需借助激光雷达的情况下,仍然能够完成精确的测距任务。

## 1 校园道路障碍物视觉检测和测距方法

### 1.1 改进的 YOLOv8s 结构

本文以 YOLOv8s 作为基准模型,虽 YOLOv8s 模型是 YOLOv8 系列中的轻量化版本,有较少的参数量和计算量,能为校园道路智能清扫车在实际应用场景中带来更高的稳定性和可靠性;但在面临一些复杂环境时容易出现检测精度不高,检测速度有限的问题。本文对 YOLOv8s 模型进行优化,在速度和精度之间达到一个不错的平衡。为降低算法的参数量、计算量,同时兼顾模型精度,实现高精度和高速度检测模型在智能清扫车上的良好运行,本文从 3 个方面对 YOLOv8s 算法模型进行改进。改进后 YOLOv8s 结构如图 1 所示。

为了减少模型参数量,降低模型复杂度,首先将 YOLOv8s 主干网络替换为结构更简单、更轻量化的 FasterNet 网络。其次设计了 SPPF-DAM 模块,以残差方式引入可变形注意力机制(deformable attention module, DAM),提高模型对多尺度目标特征的感知能力。最后,在特征融合网络中引入了基于密集扩张连接卷积(receptive field expansion module, RFEM)模块改进的 Partial-RFEM,以捕获非局部上下文特征和局部目标特征,解决目标局部语义缺失的问题,提高检测精度。最后在改进的模型中添加了测距算法,通过摄像头来实现对障碍物目标距离的获取。

### 1.2 FasterNet 轻量化主干网络

YOLOv8s 的主干网络采用了包含大量 Bottleneck 结构的 C2f 模块,虽然该结构能够有效提取特征,但其模型复杂度较高,无法适用于像校园智能清扫车这类计算资源有限的边缘设备。现阶段的轻量化网络如 MobileNet<sup>[14]</sup>和 GhostNet<sup>[15]</sup>等,通过深度卷积或组卷积提取空间特征,虽然减少了计算量,但算子的内存访问量增加,导致计算效率低下。而 Chen 等<sup>[16]</sup>提出的 FasterNet 发现计算量的减少并不一定会使神经网络的运行速度在相同水平上提升,并证明低效的浮点运算主要是由于算子的频繁内存访问所致。图 2 展示了 PConv 的工作原理及 FasterNet 的结构,该结构包含多个阶段,每个阶段的 FasterNet 块均有 PConv 层 PWConv 层,充分利用了来自所有通道的信息。

假设  $h, w$  表示输入特征图的长和宽,  $c$  表示常规卷积输入通道数,  $c_p$  表示 PConv 参与卷积通道数,  $k$  为卷积核尺寸,  $r$  表示参与卷积率,则 PConv 的计算量  $FLOP_{PConv}$  和内存访问量  $MAC_{PConv}$  分别表示为:

$$FLOPs_{PConv} = h \times w \times c_{in} \times c_{out} \quad (1)$$

$$MAC_{PConv} = h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (2)$$

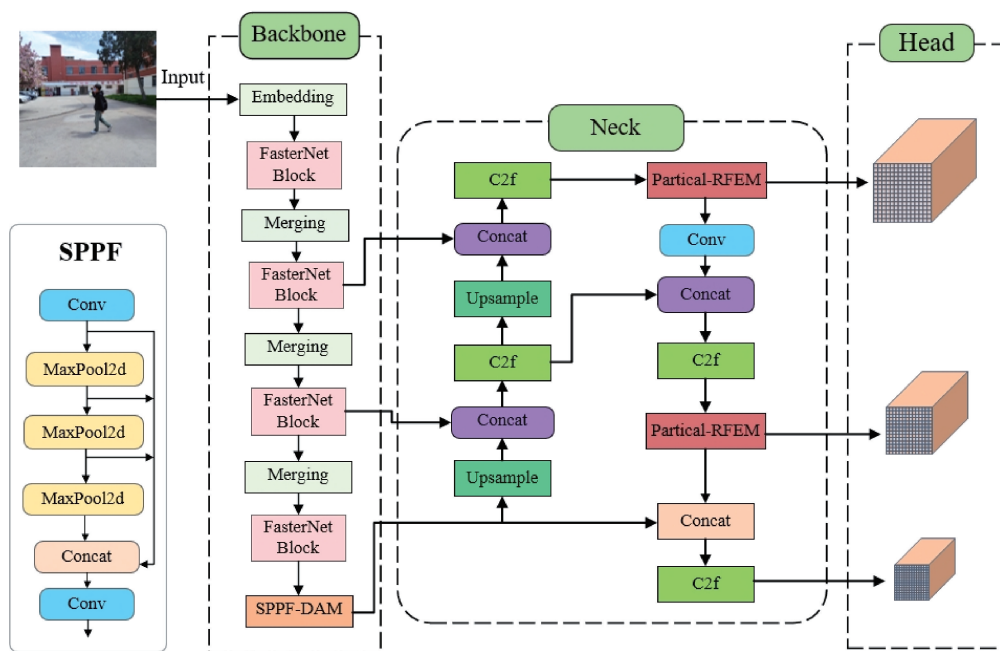


图1 改进YOLOv8s网络结构

Fig. 1 Improved YOLOv8s network structure

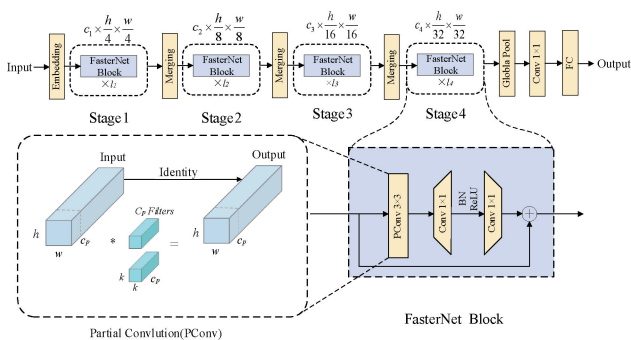


图2 FasterNet的结构图和PConv的工作原理

Fig. 2 Working principle of PConv and structure of FasterNet

设置 PConv 参与卷积通道数占常规卷积输入通道数的  $1/4$ , 则 PConv 的计算量仅为常规卷积操作的  $1/16$ , 内存访问量约为常规卷积的  $1/4$ 。由此可见, PConv 将大幅度减少计算量和内存访问量。

### 1.3 SPPF-DAM 模块

在 YOLOv8s 网络中, SPPF 模块的作用是在不同尺度上提取特征, 从而捕捉目标物体在不同尺度上的细节信息。然而, 在校园道路环境中, 障碍物的尺度信息时刻变化, 且骨干特征提取网络往往会产生大量负样本, 会导致样本分布不平衡及细节信息缺失, 从而影响模型的收敛性和特征学习。SPPF 模块的结构中包含多次最大池化操作, 这会引起局部特征丢失的问题, 而引入注意力机制可以帮助模型聚焦于有用的特征信息, 关注细节语义信息, 抑制不重要的特征。因此, 在 SPPF 模块中以残差形式引入了可变形注意力机制 DAM<sup>[17]</sup>, 可以动态调整关注区域,

适应不同大小的目标障碍物, 从而提高多尺度特征的提取能力, 进而提高模型性能。如图 3 所示, 借鉴残差连接的思想, 将 DAM 模块嵌入残差路径中, 形成了 SPPF-DAM 模块。以残差形式引入注意力机制不会影响原始特征的输出, 同时将获得的特征与原始特征拼接, 增强了输出特征的完整性, 实现了对模块的优化。

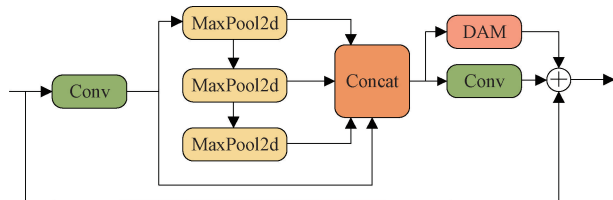


图3 SPPF-DAM 结构

Fig. 3 SPPF-DAM architecture

### 1.4 DAM 注意力机制

图 4 展示了 DAM 在 Vision Transformer<sup>[18]</sup> 中的实现过程, 包括以下步骤: 首先, 在输入特征图上生成一个统一的参考点网格, 为后续处理提供基础位置; 然后, 通过轻量级子网络计算查询令牌的偏移量, 并据此动态调整参考点的位置, 以突出关键图像区域; 接着, 在修订后的参考点位置上进行特征采样, 并应用双线性插值方法, 以精确采样周围的图像信息; 最后, 执行标准的多头注意力计算, 有效地捕捉特征之间的相互关系。

### 1.5 Patical-RFEM

YOLOv8s 算法在特征融合过程中, 由于小目标的局部语义信息不足, 导致小目标检测变得困难。因此, 需要

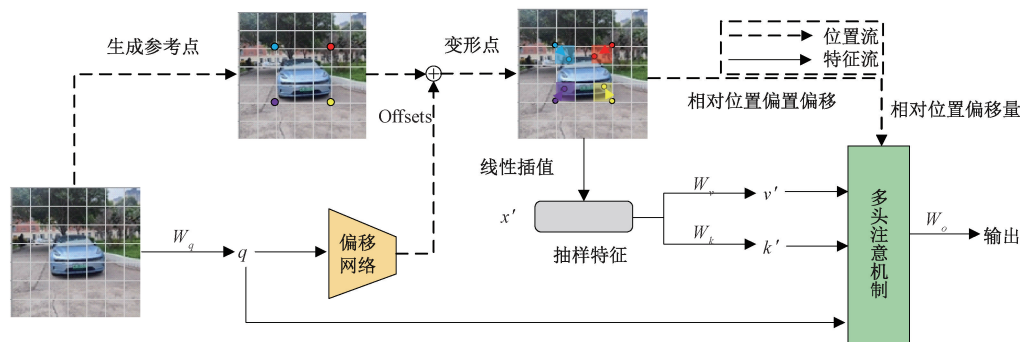


图 4 可变形注意力机制结构

Fig. 4 Deformable attention module architecture

提取具有大接收域的非局部上下文信息,以作为重要的补充和补偿。为此,研究人员提出了一种基于密集扩张卷积的特征融合模块 RFEM<sup>[19]</sup>,旨在平衡非局部上下文特征和局部目标特征的捕捉。RFEM 模块如图 5 所示, $F$  是输入的融合特征图,以及  $R_i (i=1,2,3)$  是中间变量。DConv $_{3\times 3}$  表示  $3\times 3\times C\times (C/2)$  扩张卷积,扩张率为 2,DConv $_{3\times 3}$  和 DConv $_{3\times 3}$  是  $3\times 3\times (C/4)\times (C/4)$  扩张卷积,扩张率分别为 4 和 8。

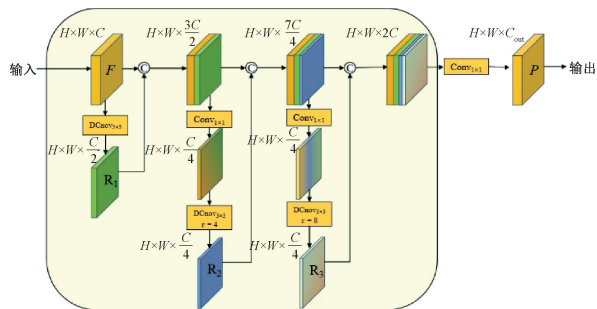


图 5 RFEM 结构

Fig. 5 Structure of RFE

RFEM 模块通过设计不同的扩张率解决了层与层之间缺乏相关性而导致的局部信息丢失问题,实现了在参数量较少的情况下对全局特征的有效获取。然而,过大的扩张率会导致内部数据结构和空间层级信息的丢失,并增加计算量,从而降低模型的运行速度。为解决上述问题,增加了一个支路以拓展模块的深度,从而获取更多的特征信息,提高对目标特征的提取和融合能力,减少过拟合的风险。改进后的结构如图 6 所示,命名为 Partial-RFEM 特征堆叠模块;将第 1 个和第 4 个并行支路的卷积形式替换为 PConv,这样在不增加参数量和计算量的情况下能更好地提取特征,另外降低了第 3 支路的扩张率,降低模块的计算量。改进后的结构不仅减少了因主干网络大量卷积提取特征所导致的信息丢失问题,还在保证检测速度的同时提高了对小目标障碍物的检测精度。

#### 1.6 添加测距功能

虽然 YOLOv8s-FDR 算法能够处理视觉传感器输入

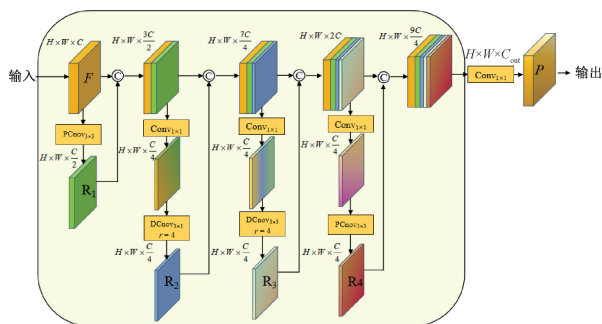


图 6 Partial-RFEM 模块结构

Fig. 6 Structure of the Partial-RFEM

的 RGB 图像并完成目标检测任务,但它缺乏与目标的距离信息。为了有效地完成避障感知任务,需要将目标检测与深度距离信息结合起来。

目前存在的测距方法有激光测距、结构光测距、双目视觉测距和单目视觉测距等,其中结构光测距的功耗小对光线的敏感度较低,更符合校园智能清扫车的工作需求,因此本研究采用结构光测距方法来解决这一问题。本文使用的摄像头为 Geminipro 双目相机,其有两个 IR 相机和一个 RGB 相机以及一个红外点阵投射器等,另外此相机可通过 OpenNI 工作包调用,测距的有效范围为 0.25~5 m,满足校园智能清扫车使用要求。其工作原理图如图 7 所示。

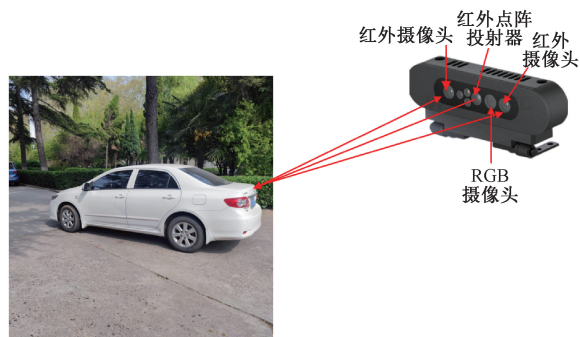


图 7 结构光主动测距原理

Fig. 7 Structured light active ranging principle



具体步骤如下:首先,利用 matlab 对 RGB-D 相机进行标定,以获得相机的焦距、畸变系数、旋转矩阵和平移向量等参数;然后,通过 OpenNI 工具包调用摄像头使投射器发射点阵图案;接着进行图案解码与匹配,在通过三角测量计算深度并输出深度图;最后,在 RGB 图像中进行目标检测,计算出预测框中心点的坐标位置,使获取的深度图和 RGB 图像进行匹配并将计算得到的深度信息返回并显示在 RGB 图像上。经过实际测量,通过结构光进行测距的结果与实际距离误差在 5 cm 左右,满足校园智能清扫车工作要求。距离测定效果如图 8 所示。



图 8 距离测定效果  
Fig. 8 Distance measurement renderings

2 数据集构建和模型训练

2.1 校园道路障碍物数据集

为了合理构建校园道路障碍物数据集,在河南科技大学校园内进行了图像采集,在不同时间段(早、中、晚)、不同天气条件(晴天、阴天、雾天)以及不同光照条件(顺光、逆光等)下进行采集,经过筛选后留存了 4 500 张图片作为初步数据集。使用 Labelimg 软件对采集的图片进行人工标注,障碍物类别包括:汽车、路障、摩托车、电动车、自行车、猫和狗,如图 9 所示,此外还引入了对应的约 5 000 张 VOC2012 公开数据集的部分图像扩充了数据集。

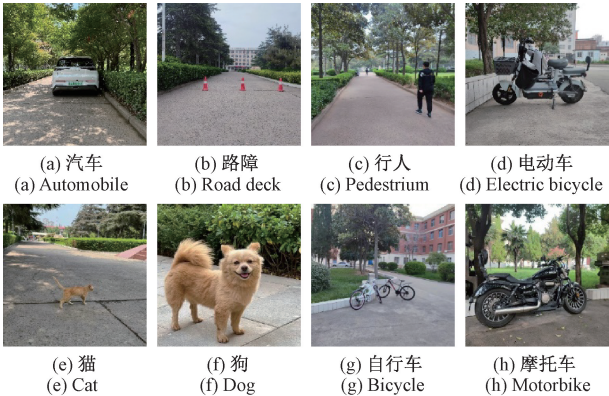


图 9 数据集部分样本  
Fig. 9 Partial sample plot of the dataset

视觉检测算法在训练过程中需要大量的图像数据进行监督学习,以便更好地捕捉数据的复杂性和多样性,并减少过拟合的风险。因此,采用了运动模糊、随机平移、随机缩放和边缘强化等数据预处理方法,将图片样本数量扩

充至 12 000 张。数据集随后按 8:1:1 的比例划分为训练集、验证集和测试集。

2.2 实验环境和评价指标

模型训练实验环境如表 1 所示。

表 1 实验环境信息表

Table 1 Experimental environment information table	
配置	型号
操作系统	Ubuntu 20.04
编程语言	Python 3.8
深度学习框架	Pytorch 2.0.0
CPU	Intel(R) Xeon(R) CPU E5-2680 v4
GPU	NVIDIA Tesla P40 24 GB
RAM	DELL DDR4 3 200 MHZ 32 GB

视觉检测性能测试平台为一个由电机驱动四轮差速转向的校园智能清扫车如图 10 所示,搭载了 Jetson Nano B01 上位机模块和视觉检测模块等,来模拟智能清扫车工作过程中对障碍物的视觉检测。



图 10 校园道路智能清扫车  
Fig. 10 Campus road smart sweeper

使用精确率(Precision)、召回率(Recall)和总类别平均精确度(mean average precision,mAP)等指标来评价视觉检测的精确度;使用模型参数量(Params)和浮点运算次数(FLOPs)来评估视觉检测模型的复杂度;使用每秒帧数(frames per second,FPS)来衡量视觉检测的速度。其公式分别为:

Precision = TP / (TP + FP) (3)

Recall = TP / (TP + FN) (4)

AP = ∫₀¹ P(r) dr (5)

mAP = 1/n ∑\_{i=1}^n AP\_i (6)

FPS = (n × batch) / ∑\_{i=1}^n (T\_end - T\_start) (7)

其中,  $TP$  为检测的目标正样本数量;  $FP$  为模型误检的目标数量;  $FN$  为模型误检及漏检的数量;  $n$  代表共检测的批次数;  $batch$  表示每批次检测的图片数量;  $(T_{end} - T_{start})$  表示每轮检测消耗的时间。

### 2.3 模型训练

模型训练的输入图片分辨率为  $640\text{ pixel} \times 640\text{ pixel}$ , 批量大小为 32, 迭代次数为 200 个 Epoch。梯度更新方面, 选择随机梯度下降策略, 动量为 0.937, 权重衰减为 0.0005, 学习率范围为 0.1~0.01, 采用余弦退火率更新策略。

## 3 实验结果与分析

### 3.1 主干网络模型复杂度对比实验

为验证 FasterNet 主干网络对降低模型复杂度的有效性, 与 YOLOv8s 算法原有的主干网络以及不同轻量化主干网络进行对比, 对比实验结果如表 2 所示。

表 2 不同轻量化主干网络的性能  
Table 2 Performance of different lightweight backbone networks

主干网络	FLOPs/ ( $G \cdot s^{-1}$ )	Params/ MB	$mAP_{0.5}$ / %	$mAP_{0.5+0.95}$ / %
YOLOv8s	28.4	11.13	85.63	56.81
MobileNetV3	21.2	9.94	76.21	50.11
ShuffleNetV2	16.4	6.38	73.85	46.93
FasterNet	21.6	8.62	80.35	52.18
GhostNetV2	21.0	9.40	78.03	49.02

由表 2 可知, 相较于 YOLOv8s 算法原有的主干网络, YOLOv8s-FDR 算法采用 FasterNet 作为轻量化主干网络, 使得 FLOPs 和 Params 分别下降了 23.59% 和 22.55%。这表明 FasterNet 有效降低了模型的复杂度, 使模型更加轻量化。在  $mAP_{0.5}$  和  $mAP_{0.5}$  两个指标上, FasterNet 仍然保持较高的性能, 分别达到了 80.35% 和

52.18%。这说明 FasterNet 在降低模型复杂度的同时, 依然具备较强的不同尺度特征提取能力。

### 3.2 注意力残差模块精确度对比实验

为了验证 DAM 注意力机制模块的优越性, 以 FasterNet 作为主干网络的基础上, 加入了 SPPF-DAM 模块。对比使用不同注意力机制的注意力残差模块对障碍物检测精确度的影响, 并设置不使用注意力残差模块进行对照, 实验结果如表 3 所示。

表 3 不同注意力机制的精确度  
Table 3 The accuracy of different attention mechanisms

注意力机制	AP <sub>0.5</sub>			mAP <sub>0.5</sub>	mAP <sub>0.5+0.9</sub>
	行人	汽车	路障		
无	83.21	88.22	82.65	80.35	52.18
CBAM	82.37	88.97	80.08	84.98	53.23
SA	84.92	91.86	81.53	86.82	54.74
SimAM	82.17	90.39	83.70	86.46	53.34
DAM	86.79	91.05	84.22	87.66	55.48

通过  $mAP_{0.5}$  和  $mAP_{0.5+0.9}$  这两个指标可以看出, 带有 DAM 注意力机制的注意力残差模块效果提升最显著。表明带有 DAM 注意力机制的模块通过动态调整关注区域, 聚焦于多尺度特征信息的提取, 有效抑制了道路背景信息的干扰, 从而显著提升了算法的检测精度。

### 3.3 视觉检测算法构建策略消融实验

为了更好地分析本文不同改进策略的有效性, 在 YOLOv8s 算法基础上, 依次添加本文改进策略进行消融实验。将本小节实验不同分组所涉及的算法依次命名为算法 1、2、3 和 4, 即: 算法 1: YOLOv8s; 算法 2: 算法 1 + FasterNet; 算法 3: 算法 2 + SPPF-DAM; 算法 4: 算法 3 + Partial-RFEM; 消融实验结果如表 4 所示。

表 4 消融实验结果  
Table 4 Ablation test results

分组	模型	FLOPs/( $G \cdot s^{-1}$ )	Params/MB	$mAP_{0.5}$ /%	$mAP_{0.5+0.9}$ /%	FPS/fps
算法 1	YOLOv8s	28.4	11.13	85.63	56.81	42.73
算法 2	算法 1+FasterNet	21.6	8.62	80.35	52.18	71.26
算法 3	算法 2+SPPF-DAM	21.7	8.68	87.66	57.48	66.53
算法 4	算法 3+Partial-RFEM	22.5	9.23	88.71	58.22	64.82

与原算法相比, 引入 FasterNet 轻量化主干网络改进 YOLOv8 算法后,  $mAP@0.5$  和  $mAP@0.5:0.95$  分别下降了 6.2% 和 8.1%, 表明轻量化主干网络的引入导致了算法特征提取能力的下降, 影响了障碍物视觉检测的精确度。然而, 模型复杂度显著降低, 具体表现为 FLOPs 从 28.4 G/s 减少至 21.6 G/s, 参数量从 11.13 MB 减少至

8.62 MB。同时, FPS 从 42.73 fps 提升至 71.26 fps, 说明引入轻量化主干网络后, 视觉检测速度得到了显著提升, 满足了实时性需求。在引入注意力残差模块后,  $mAP_{0.5}$  和  $mAP_{0.5+0.95}$  分别提升了 87.66% 和 57.48%。引入 Partial-RFEM 多分支卷积模块后, 模型复杂度并未显著增加, 同时  $mAP_{0.5+0.9}$  提升至 58.22%, 超过 YOLOv8s 算法的

56.81%。这表明 Partial-RFEM 模块能够在保持较小计算成本的情况下,加强小尺度特征信息的提取和融合能力,提高了检测矩形框的回归定位精度,进而提高了障碍物视觉检测的精确度。

3.4 不同视觉检测算法性能对比实验

为了分析 YOLOv8s-FDR 算法校园道路障碍物视觉检测性能,将 YOLOv8s-FDR 算法与其他主流视觉检测算法进行对比实验,并增加精确率和召回率评价指标,对比实验结果如表 5 所示。

表 5 模型对比实验

Table 5 Model comparison experimental

算法	Precision/ %	Recall/ %	FLOPs/ (G · s <sup>-1</sup> )	Params/ MB
Faster-RCNN	67.14	65.31	132.2	41.31
CenterNet	80.61	76.45	50.09	13.75
YOLOX-M	79.29	74.62	73.70	25.31
YOLOv5s	82.32	81.16	16.50	7.20
YOLOv7	81.55	79.33	104.70	36.92
YOLOv9s	84.71	85.44	26.70	7.20
YOLOv8s	85.63	84.05	28.40	11.13
YOLOv8s-FDR	88.71	90.26	22.50	9.23

实验结果表明改进的 YOLOv8s-FDR 在精度、召回率以及计算复杂度方面均优于其他算法。具体而言, YOLOv8s-FDR 的精度达到 88.71%,召回率为 90.26%,显著高于 YOLOv8s 的 85.63%和 84.05%。该算法在预测类别准确性方面表现优异,校园道路障碍物的误检概率较低。此外, YOLOv8s-FDR 的 FLOPs 为 22.5 G/s,较 YOLOv8s 的 28.4 G/s 显著减少且仅高于 YOLOv5s,显示出其有更高的计算效率。在模型参数量方面,改进算法为 9.23 MB,仅高于 YOLOv5s 和 YOLOv9s 这两种算法,而相对于 YOLOv8s 的 11.13 MB 进一步减少,表明该算法更加轻量化,但在精确度和召回率方面高于 YOLOv5s

和 YOLOv9s。通过对比实验可以得出,本文改进的模型在满足实时性要求的情况下具有较高的检测性能,适合在资源受限的环境中部署。

图 11 更加直观的展示了常见障碍物视觉检测算法的检测速度与精确度性能。相较于其他算法, YOLOv8s-FDR 算法的检测速度最快,精确度最高。因此, YOLOv8s-FDR 算法在保证较高检测精确度的同时,检测速度最高,综合性能最佳,适合应用于智能清扫车视觉检测任务。

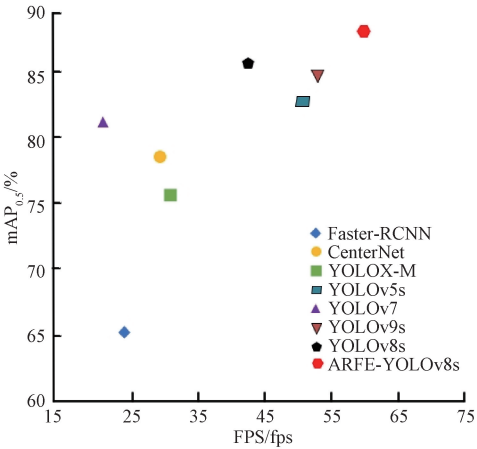


图 11 检测速度和精度对比

Fig. 11 Comparison of detection speed and precision

3.5 模型检测结果对比

为了能够更好的观察改进算法的检测效果,选取了多种环境下的数据集来进行测试,原算法和改进算法的测试结果分别如图 12 和 13 所示。在图 12 和 13 中不难看出在强光照以及昏暗的环境下,本文改进模型能够充分检测到所有目标,而原模型检测中出现了漏检,此外改进模型检测精度普遍高于原模型。另外从图中可以看出本文改进的模型能够准确识别远距离的目标,而原模型对于远处的目标检测精度大大降低。通过上述对比试验结果,充分验证了本文提出的 YOLOv8s-FDR 模型的检测性能以及在校园智能清扫车上应用的可行性。

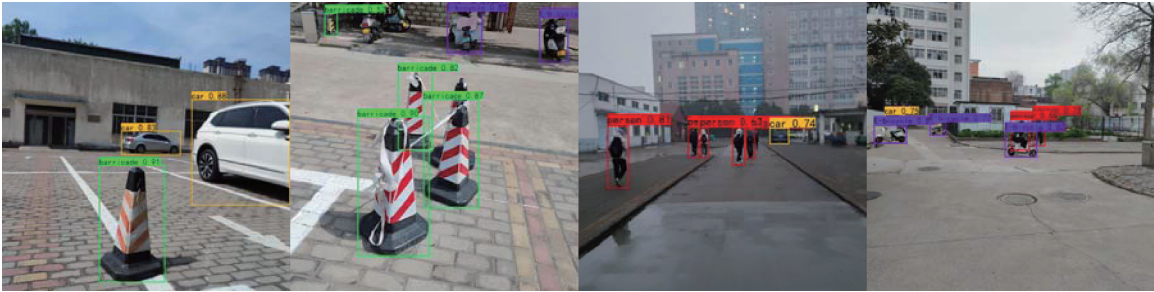


图 12 YOLOv8s 模型检测效果

Fig. 12 Detection effect of YOLOv8s model

3.6 校园智能清扫车视觉检测性能测试

在校园道路上来模拟智能清扫车的作业过程,在测试

之前,分别将 YOLOv8s 和 YOLOv8s-FDR 算法部署到智能清扫车中。当清扫车行驶至大致相同的作业位置时,分



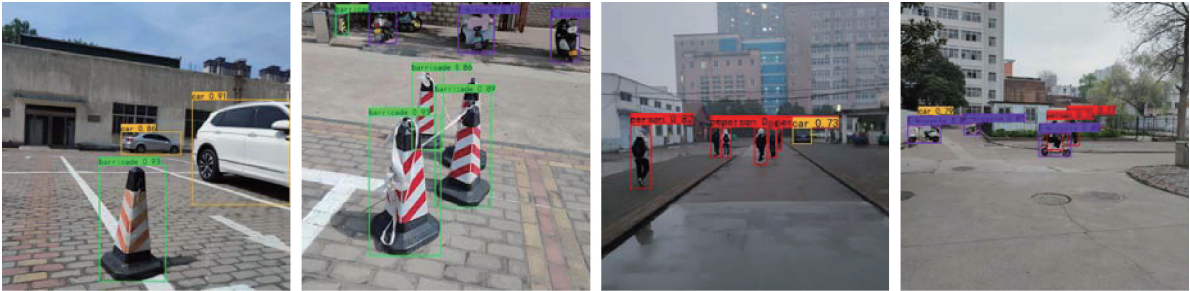


图 13 改进 YOLOv8s 模型检测效果

Fig. 13 Improved the detection effect of YOLOv8s model

别截取两种算法的实时检测画面,对比算法检测性能。YOLOv8s 算法的检测效果如图 14 所示, YOLOv8s-FDR 算法的检测效果如图 15 所示。由测试结果可知,改进的 YOLOv8s 算法检测出的检测精确度最低为 85% 优于原算法。由检测画面左上角实时显示的 FPS 指标可知,将 YOLOv8s 算法移植到智能清扫车系统中运行后, FPS 仅为 17.12 fps, 表明在进行障碍物视觉检测时,检测画面显示卡顿,因此 YOLOv8s 算法达不到检测实时性要求。而本文构建的 YOLOv8s-FDR 算法在添加了测距功能后, FPS 可达到 38.44 fps, 同时能够精准显示出目标的距离,满足本文搭建的智能清扫车视觉检测实时性作业要求。

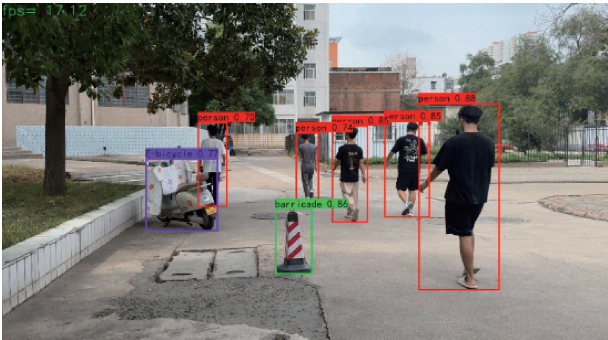


图 14 YOLOv8s 算法检测实时画面

Fig. 14 The YOLOv8s algorithm detects real-time images



图 15 YOLOv8s-FDR 算法检测实时画面

Fig. 15 The YOLOv8s-FDR algorithm detects real-time images

4 结 论

针对校园智能清扫车障碍物检测精确度低、检测速度慢以及模型复杂度高的问题,提出了一种改进 YOLOv8s 的校园智能清扫车障碍物检测与测距算法,以 YOLOv8s 为基础,将主干网络替换为 FasterNet 轻量化网络,降低模型参数量和复杂度并提升视觉检测速度;引入 SPPF-DAM 注意力残差模块减少背景噪声的干扰,提升了算法捕捉多尺度特征的能力;使用 Partial-RFEM 模块进行下采样,扩大感受野,捕捉更大范围内的上下文信息。构建校园道路障碍物数据集并验证算法优化策略的有效性;通过结构光测距,获取了障碍物的距离信息。相较于 YOLOv8s 算法,改进后的算法  $mAP_{0.5}$  提高了 3.6%,计算量和参数量分别降低了 19.72% 和 15.27%,检测速度提升了 51.7%。在实际校园道路环境中测试智能清扫车视觉检测的性能表现,测试结果,添加了测距功能的 YOLOv8s-FDR 算法 FPS 实际表现为 38.44 fps,达到视觉检测实时性要求。未来需要继续扩充数据集,提高模型的泛化性,实现在更复杂的环境中稳定工作。

参考文献

[1] 高瞩,丁栋,姜乾喆,等. 基于场景交互理论的智能道路清扫车安全设计研究[J]. 包装工程,2023,44(6): 410-419.  
GAO ZH, DING D, JIANG Q ZH, et al. Research on safety design of intelligent road sweeper based on scene interaction theory[J]. Packaging Engineering, 2023,44(6):410-419.  
[2] 吕睿. 基于双目视觉的智能清扫车道路目标检测[D]. 长春:吉林大学,2024.  
LYU R. Road target detection of intelligent sweeper based on binocular vision [D]. Changchun: Jilin University,2024.  
[3] 史立秋,程亮,江霖,等. 校园小型无人驾驶清扫技术研究[J]. 经济师,2020(12):220-221.  
SHI L Q, CHENG L, JIANG L, et al. Research on small unmanned cleaning technology in campus [J].



- Economist, 2020(12):220-221.
- [4] 牛为华,殷苗苗. 基于改进YOLOv5的道路小目标检测算法[J]. 传感技术学报, 2023, 36(1):36-44.  
NIU W H, YIN M M. Road small target detection algorithm based on improved YOLOv5[J]. Journal of Sensing Technology, 2023, 36(1):36-44.
- [5] KIRAN B R, SOBH I, TALPAERT V, et al. Deep reinforcement learning for autonomous driving: A survey [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(6): 4909-4926.
- [6] 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.
- [7] GIRSHICK R. Faster-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, 2016: 21-37.
- [9] 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.
- [10] 高德勇,陈泰达,缪兰. 改进YOLOv8n的道路目标检测算法[J]. 计算机工程与应用, 2024, 60(16): 186-197.  
GAO D Y, CHEN T D, MIAO L. Improved YOLOv8n road target detection algorithm[J]. Computer Engineering and Applications, 2024, 60(16):186-197.
- [11] XIONG X, HE M, LI T, et al. Adaptive feature fusion and improved attention mechanism based small object detection for UAV target tracking [J]. IEEE Internet of Things Journal, 2024, 11(12):21239-21249.
- [12] 胡伟超,皮建勇,胡倩,等. 面向复杂场景密集行人检测的YOLOv8改进模型[J]. 电子测量技术, 2024, 47(14):159-169.  
HU W CH, PI J Y, HU Q, et al. YOLOv8 improved model for dense pedestrian detection in complex scenes[J]. Electronic Measurement Technology, 2024, 47(14):159-169.
- [13] 刘罡,闫曙光,刘钰,等. 基于感受野增强的复杂道路场景目标检测研究[J]. 电子测量技术, 2024, 47(13):157-166.  
LIU G, YAN SH G, LIU Y, et al. Research on object detection in complex road scene based on receptive field enhancement [J]. Electronic Measurement Technology, 2024, 47(13):157-166.
- [14] HOWARD A G. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. ArXiv preprint arXiv:1704.04861, 2017.
- [15] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [16] CHEN J, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [17] XIA Z, PAN X, SONG S, et al. Vision transformer with deformable attention[C]. IEEE/CVF Conference on computer Vision and Pattern Recognition, 2022: 4794-4803.
- [18] KHAN S, NASEER M, HAYAT M, et al. Transformers in vision: A survey [J]. ACM Computing Surveys(CSUR), 2022, 54(10s): 1-41.
- [19] HAN Y, LIAO J, LU T, et al. KCPNet: Knowledge-driven context perception networks for ship detection in infrared imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 61: 1-19.

## 作者简介

郭志军(通信作者),博士,教授,主要研究方向为车辆系统动力学、车辆智能控制等。

E-mail:gzhij1970@163.com

叶世文,硕士研究生,主要研究方向为深度学习,目标检测。

庞明天,硕士研究生,主要研究方向为智能车辆路径跟踪。

王丁健,博士研究生,主要研究方向为智能车辆路径规划、燃料电池电解槽。

杜林林,硕士研究生,主要研究方向为智能车辆路径规划。