

DOI:10.19651/j.cnki.emt.2416859

## 改进 YOLOv8 的城市行车道路障碍物检测算法研究\*

向 雷<sup>1,2</sup> 蒋文波<sup>1,2</sup>

(1. 西华大学电气与电子信息学院 成都 610039; 2. 西华大学信号与信息处理四川省高校重点实验室 成都 610039)

**摘要:** 针对目前城市道路复杂环境下障碍物检测精度不足、检测速度慢、模型参数量大和小目标障碍物检测效果不佳的问题,提出一种改进的 YOLOv8n 轻量级城市行车道路障碍物检测算法。首先,制作 MRObstacle 城市道路障碍物目标检测数据集,扩展了障碍物检测种类与数量;其次,设计全新的 SPS\_C2f 改进主干网络,降低网络参数量与提升检测速度,添加 M\_ECA 注意力模块至网络的 Neck 部分,提升网络检测速度与特征表达能力;再次,融合 BiFPN 特征金字塔和添加小目标检测头,更好地捕捉小尺寸障碍物的特征;最后,使用可优化边界框宽度与高度值的损失函数 MPDIoU,提升网络边界框回归性能。相比于原 YOLOv8n 算法,该算法的 mAP0.5 指标提升 2.04%,达到 97.12%;FPS 值提升 12.08 fps,达到 107.45 fps;网络参数量减少 10%,降低至 2.73 MB。该算法在减少参数量的同时提高了检测精度和速度,可更好应用于城市行车道路障碍物检测任务。

**关键词:** 障碍物检测;YOLOv8;改进 C2f 模块;改进注意力机制;损失函数

**中图分类号:** TP391;TN0 **文献标识码:** A **国家标准学科分类代码:** 510.99

## Research on improved YOLOv8 urban driving road obstacle detection algorithm

Xiang Lei<sup>1,2</sup> Jiang Wenbo<sup>1,2</sup>

(1. School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, China;

2. Sichuan Provincial Key Laboratory of Signal &amp; Information Processing, Xihua University, Chengdu 610039, China)

**Abstract:** Aiming at the current problems of insufficient obstacle detection accuracy, slow detection speed, large number of model parameters and poor detection of small target obstacles in the complex environment of urban roads, an improved YOLOv8n lightweight urban driving road obstacle detection algorithm is proposed. Firstly, the MRObstacle urban road obstacle target detection dataset is produced to extend the types and numbers of obstacle detection; secondly, a new SPS\_C2f backbone network is designed to improve the backbone network, to reduce the number of network parameters and to improve the detection speed, and the M\_ECA attention module is added to the Neck portion of the network, to improve the network detection speed and the feature expression ability; thirdly, the BiFPN is integrated with a feature pyramid and a small target detection algorithm is added to the network, feature pyramid and adding a small target detection head to better capture the features of small-sized obstacles; finally, using the loss function MPDIoU that optimises the values of the bounding box width and height to improve the performance of the network bounding box regression. Compared with the original YOLOv8n algorithm, the mAP0.5 metric of this algorithm is improved by 2.04% to 97.12%, the FPS value is improved by 12.08 fps to 107.45 fps, and the volume of the network parameter is reduced by 10% to 2.73 MB. This algorithm improves the detection accuracy and speed while decreasing the number of parameters, and it can be better applied to the urban road obstacle detection task.

**Keywords:** obstacle detection; YOLOv8; improved C2f module; improved attention mechanisms; loss function

## 0 引言

近年来,随着人工智能技术的蓬勃发展,汽车行业正在

加快推动智能化建设<sup>[1]</sup>。然而,目前要实现自动驾驶还有许多技术难题需要攻克,其中环境感知在确保行车安全和提高驾驶体验方面发挥着至关重要的作用<sup>[2]</sup>。由于城市道

收稿日期:2024-09-10

\* 基金项目:四川省科技厅应用基础研究项目(2019YJ0455)、西华大学研究生科创竞赛项目(YK20240002)资助



权衡检测精度和速度后,本文选择在最轻量的 YOLOv8n 作为基础,对模型进行改进,以提高对于城市道路上的各种障碍物的检测性能。改进后的模型结构图如图 2 所示。主要有 4 个改进部分。在 Backbone 中引入 SPS\_C2f 结构降低网络参数量;为提高网络对障碍物目标的特征融合

能力,在 Neck 中引入 M\_ECA 注意力机制和融合 BiFPN 特征金字塔;新增一个 160×160 的小目标检测头 P2,提升复杂交通环境下小目标障碍物的检测精度;引入 MPDIoU 损失函数,增强网络收敛稳定性,更好地提升网络边界框回归性能。改进的结构如图 2 上方黑色实线框所示。

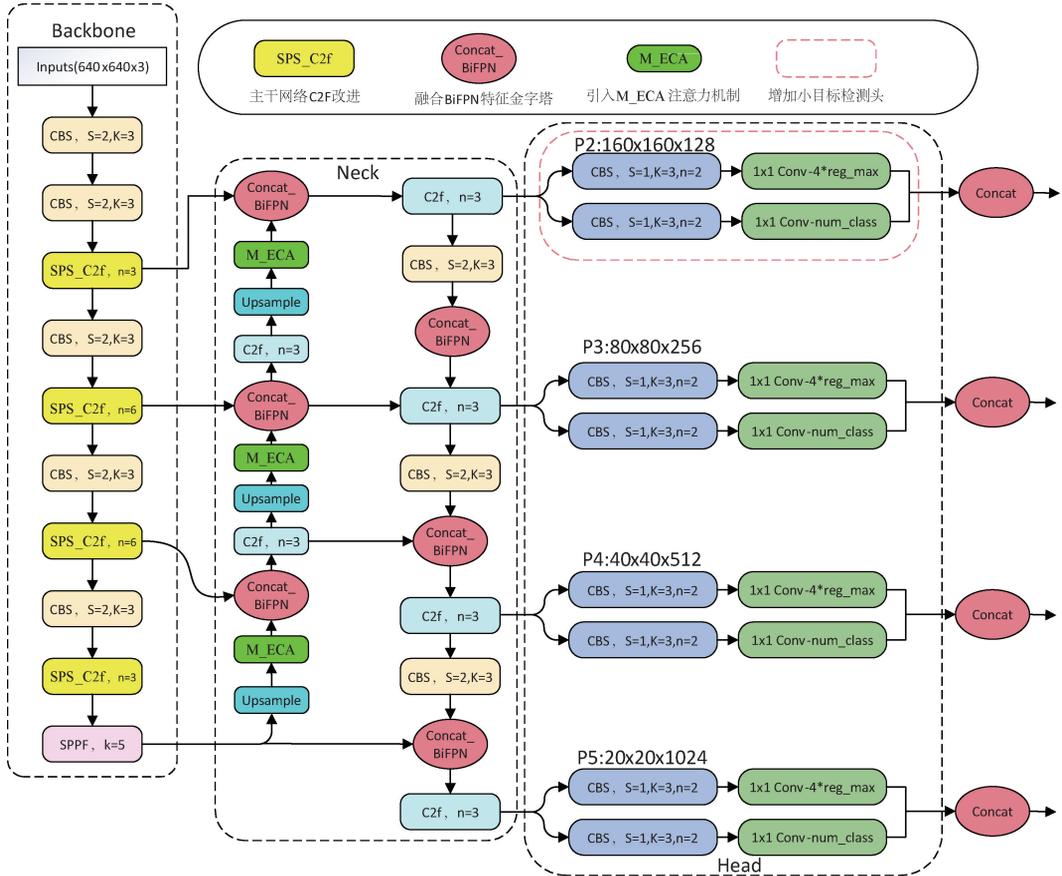


图 2 改进 YOLOv8n 网络结构

Fig. 2 Improvement of YOLOv8n network structure

1) 主干网络轻量化

YOLOv8 网络采用了多尺度特征融合的思想,有利于更充分地利用多个尺度特征。然而,由于通道数量的增加与拼接,可能导致多个通道携带相似或相同的信息,产生特征图冗余的问题。为了解决这个问题,设计 SPSCConv 来减轻模型的重量,同时仍能保持精度。如图 3 所示,使用 SPSCConv 构成的 SPSBlock 来替换主干网络 C2f 模块中的部分 Bottleneck,改进的 C2f 结构取名为 SPS\_C2f。如图 3(b)所示,改进后的主干网络取名为 SPSBackbone。

SPSBlock 结构如图 4(b)所示,SPSConv 中 SConv 表示在标准卷积前加上一个批处理归一化层,SPConv 表示在 PConv(部分卷积)前增加一个批处理归一化层,添加批处理归一化层可使中间层输出更加稳定。shuffle 将训练模型的数据集进行打乱,使数据排列拥有一定的随机性,避免过拟合以及提升模型预测表现。SPSConv 结构如图 4(a)所示。

PConv 可使网络减少冗余计算<sup>[10]</sup>,但可能使模型检测精度下降,因此采用一个普通卷积与其叠加来保持精度。PConv 原理如图 5 所示,它使输入通道的 1/4 进行卷积操作,让剩余通道保持不变,然后将经过卷积处理的 1/4 通道与未处理的通道拼接输出。PConv 的计算量 C 与内存访问量 M 为:

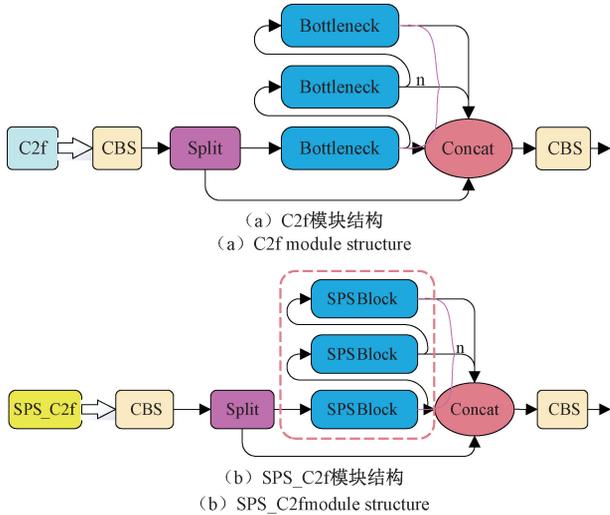
$$C = h \times w \times \omega \times k^2 \times c_p^2 \quad (1)$$

$$M = h \times w \times 2c_p + k^2 \times c_p^2 \quad (2)$$

其中,  $h$  和  $w$  分别代表特征图的高与宽,  $k$  为卷积核的大小,  $c_p$  为常规卷积的通道数。

2) 引入 M\_ECA 注意力机制

ECA 注意力模块通过快速一维卷积生成通道注意,因其无需降维的信道交互特性,相比传统的 SE 通道注意力模块更具轻量化<sup>[11]</sup>。ECA 模块通过平均池化(GAP)获得特征集合,然而城市道路图像背景占比较大并且障碍物数量众多,特征重要性不够突出会导致次要特征影响检测结

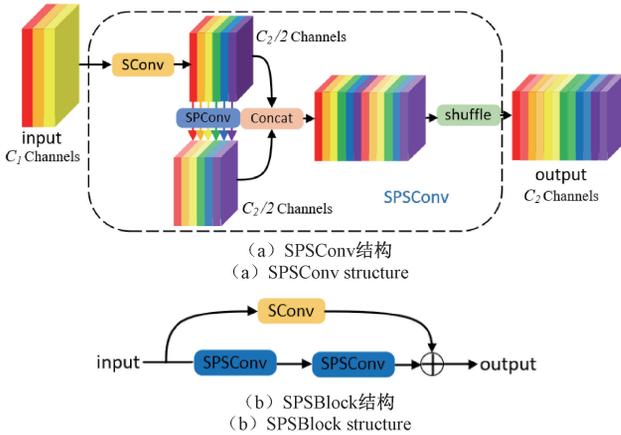


(a) C2f模块结构  
(a) C2f module structure

(b) SPS\_C2f模块结构  
(b) SPS\_C2f module structure

图 3 C2f 模块改进

Fig. 3 C2f module improvements



(a) SPSCConv结构  
(a) SPSCConv structure

(b) SPSBlock结构  
(b) SPSBlock structure

图 4 SPSCConv 与 SPSBlock 结构

Fig. 4 SPSCConv and SPSBlock structure

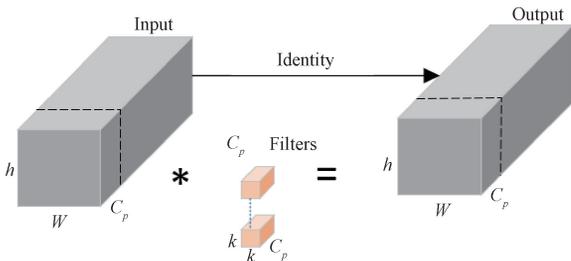


图 5 PCConv 原理

Fig. 5 PCConv principle

果,因此在 ECA 模块结构中增加最大池化(MAP)获得特征集合,通过 GAP 与 MAP 获取特征信息后结合处理,可以使网络避免重要信息丢失,改进后的结构取名为 M\_ECA,如图 6 所示。

图 6 中的  $K$  为自适应计算一维卷积核大小,公式如下:

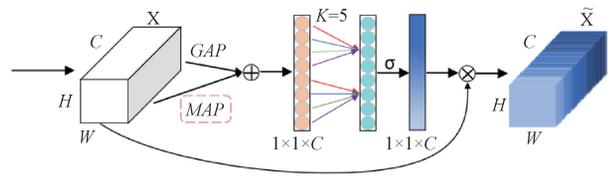


图 6 M\_ECA 注意力模块

Fig. 6 M\_ECA attention module

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (3)$$

其中,  $C$  为通道数,  $\gamma$  与  $b$  是超参数,  $|t|_{\text{odd}}$  表示取最接近  $t$  的奇数。

### 3) BiFPN 与小目标检测头

YOLOv8 网络架构中采用了 FPN+PAN(特征金字塔)结构,FPN 传递特征信息的特点是自上而下和横向连接的,如图 7(a)所示<sup>[12]</sup>。PAN 的底层输入完全依赖于 FPN 处理后的特征信息,可能会使从骨干网络中获取的原始特征信息部分丢失,从而对检测准确性产生影响。

BiFPN 在 FPN 的基础上删除了只有一条输入边的节点,采用双向跨尺度连接和加权特征融合,加强各个节点融合不同特征对特征网络的贡献,实现更好的准确性和效率平衡。如图 7(b)所示<sup>[13]</sup>。针对本研究中的城市行车道路障碍物检测任务,本文引入加权双向特征金字塔 BiFPN 结构,融合 BiFPN 的 YOLOv8 网络可以使特征金字塔各个层级更好的连接,有助于提供更全局的和更重要的特征信息,提高对于不同尺度的障碍物的检测能力。

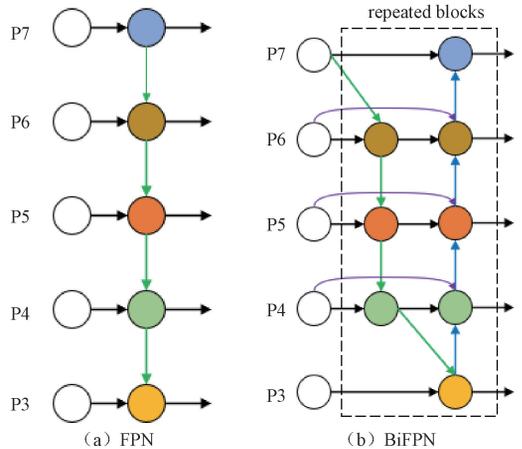


图 7 FPN 和 BiFPN 结构

Fig. 7 FPN and BiFPN structures

行车道路上有着众多的小目标障碍物,它们往往具有较少的特征信息和较低的分辨率,容易被漏检或错误分类。因此,本文在头部网络中新增加了一个小目标检测头 P2,其结构如图 2 红色虚线框中所示。改进后的网络共有 4 个检测头,分别对应颈部网络端特征图大小为  $160 \times 160$ 、 $80 \times 80$ 、 $40 \times 40$ 、 $20 \times 20$ 。长和宽为 160 的特征图具有原始图像中更多的细节信息,能够在城市复杂道路环境下有效

捕捉小目标障碍物的特征信息。

4)改进损失函数

YOLOv8网络使用CIoU作为边界框损失函数,通过考虑目标框与锚框的中心点距离、长宽比和重叠面积来计算损失,公式如式(4)~(7)所示。

$$\mathcal{L}_{CIoU} = 1 - IoU + \frac{\rho^2(B_{gt}, B_{pred})}{C^2} + \alpha\nu \quad (4)$$

$$IoU = \frac{B_{gt} \cap B_{pred}}{B_{gt} \cup B_{pred}} \quad (5)$$

$$\nu = \frac{4}{\pi^2} (\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_{pred}}{h_{pred}})^2 \quad (6)$$

$$\alpha = \frac{\nu}{(1 - IoU) + \nu} \quad (7)$$

其中, IoU(intersection over union)为真实框和预测框的交并比;  $B_{gt}$  代表真实框,  $B_{pred}$  代表预测框;  $\rho^2(B_{gt}, B_{pred})$  表示预测框和真实框中心点之间的欧氏距离;  $C$  表示包含预测框和真实框最小矩形的对角线长度;  $\alpha$  为平衡参数;  $\nu$  用来衡量长宽的相似性。

从上述可以看出,当预测的边界框和真实框长宽比一致而长度和宽度值不相等时,(如图8,虚线框为真实框,实线框为预测框),CIoU损失函数中的 $\nu = 0$ ,将会失去惩罚作用,损失函数无法被优化,而MPDIoU损失函数在这种预测情况下不会失效<sup>[14]</sup>。

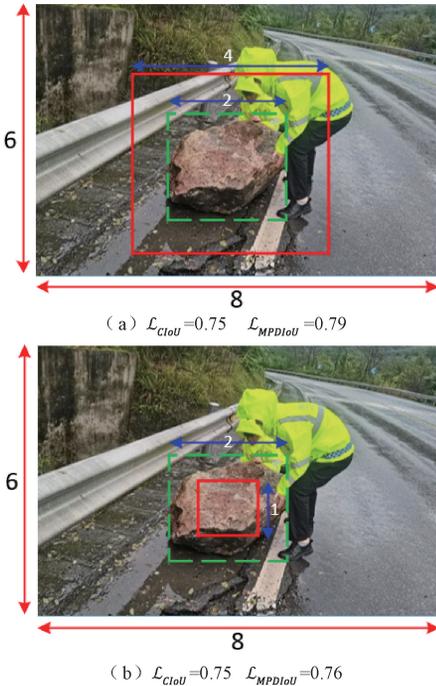


图8 两种情况下的边界框回归结果

Fig. 8 Boundary box regression results for both cases

可见,CIoU损失函数在不同的预测情况下具有相同的值,在道路障碍物检测过程中,降低了模型的收敛速度和精度。因此,本文采用MPDIoU损失函数替换CIoU对

网络的进行优化。其计算公式如式(8)~(10)所示。

$$\mathcal{L}_{MPDIoU} = 1 - IoU + \frac{d_1^2}{h^2 + w^2} + \frac{d_2^2}{h^2 + w^2} \quad (8)$$

$$d_1^2 = (x_1^{pred} - x_1^{gt})^2 + (y_1^{pred} - y_1^{gt})^2 \quad (9)$$

$$d_2^2 = (x_2^{pred} - x_2^{gt})^2 + (y_2^{pred} - y_2^{gt})^2 \quad (10)$$

其中,  $h$  和  $w$  分别代表输入图片的高和宽;  $(x_1^{pred}, y_1^{pred}), (x_2^{pred}, y_2^{pred})$  分别代表预测框左上角与右下角的坐标;  $(x_1^{gt}, y_1^{gt}), (x_2^{gt}, y_2^{gt})$  分别代表真实框左上角与右下角的坐标;  $d_1$  为预测框和真实框左上角的距离;  $d_2$  为预测框和真实框右下角的距离。

由公式可知,当预测的边界框和真实框长宽比一致而长度和宽度值不相等时,该损失函数的惩罚项不为0。相比CIoU,使用MPDIoU增强了网络收敛的稳定性,提高模型在城市道路障碍物检测中的准确率。

2 实验设计与结果分析

2.1 数据集与实验环境

针对现存道路障碍物数据集中障碍物种类较少,而实际道路上存在的障碍物种类较多的问题。本文采用从网络上收集的方式来获取所需图片数据,使用Labelimg工具进行标注,再对数据进行处理和增强,制作的数据集取名为MRObstacle。数据集中部分障碍物类别图例如图9所示。

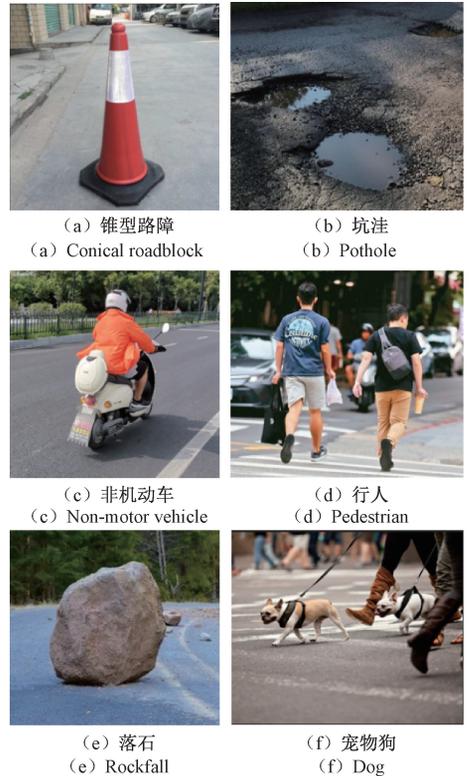


图9 MRObstacle数据集部分图例

Fig. 9 MRObstacle dataset partial legend

障碍物类别以及数据增强后的数量如表 1 所示。

表 1 障碍物类别及数量

Table 1 Type and number of obstacles

编号	障碍物类别	数据增强后的数量/张
0	汽车	315
1	行人	236
2	非机动车	550
3	球形路障	616
4	锥形路障	307
5	警示柱	259
6	宠物(狗为主)	769
7	落石	515
8	减速带	940
9	大裂缝	430
10	塌方	737
11	坑洼	1 192
12	车祸	1 059
	共计	7 925

障碍物检测所使用的标签种类分为以上 13 类,是城市道路上车辆正常行驶中较常出现的障碍物物体。汽车、行人等类别图片张数虽少,但这两类在其他类别图片上频繁出现,也进行了标注,以确保整个数据集的实例数量较为均衡。并按照 7:2:1 的比例随机划分为训练集、测试集和验证集,达到一个检测数据集应有的良好性能。

本文中所有实验均表 2 所示的主机环境下进行。

训练参数设置: 优化器选择为 SGD (stochastic gradient descent), 动量参数为 0.937, 初始学习率为 0.01, 最终学习率为 0.000 1, batchsize 为 32, 一整轮训练共 300 个 epochs。

## 2.2 实验评价指标

城市行车道路上障碍物检测是一个实际的且需要很强实时性的任务, 本文使用精确率(precision, P)、平均精度(mean average precision, mAP)、召回率(recall, R)、参数量

表 2 实验训练环境

Table 2 Experimental training environment

项目	型号
操作系统	Windows 11 专业版 64 位
CPU	Inter i9-12900
GPU	Geforce GTX3090 24 G
编译器	PyCharm
算法框架	Pytorch1.13+Cuda11.7
编程语言	Python3.9

(Params)、计算量(GFLOPs)和每秒帧率(frames per second, FPS)。计算公式如式(11)~(15)所示。

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$AP = \int_0^1 P(R) dR \quad (13)$$

$$mAP = \frac{\sum_{m=1}^M AP}{M} \quad (14)$$

$$FPS = \frac{N(p)}{T(p)} \quad (15)$$

其中,  $TP$  为实际是正类被预测为正类样本的数量;  $FP$  为实际是负类被预测成正类样本的数量;  $FN$  为实际是正类被预测成负类样本的数量。  $AP$  表示曲线与准确率  $P$  为纵轴和召回率  $R$  为横轴的坐标系下围成的面积;  $M$  表示数据集中类别数量。  $N(p)$  为处理图像的总数;  $T(p)$  为处理图像总花费时间。

## 2.3 实验结果与分析

### 1) 主干网络对比实验

本节在 YOLOv8n 的基础上对其特征提取主干网络进行改进, 选取部分近年来主流的主干网络进行替换, 主干网络对比实验结果如表 3 所示。

表 3 主干网络对比实验结果

Table 3 Backbone network comparison experiment

主干网络	P/%	R/%	mAP@0.5/%	Params/ $10^6$	GFLOPs/G	FPS/fps
CSPDarknet	92.95	91.12	95.08	3.09	8.1	92.37
Biformer <sup>[15]</sup>	<b>93.86</b>	91.25	<b>95.18</b>	11.41	35.5	77.54
mobileNetV3 <sup>[16]</sup>	87.63	85.13	89.87	<b>2.35</b>	<b>5.7</b>	102.33
Ghostnet <sup>[17]</sup>	90.82	87.26	91.82	3.35	6.9	93.62
VaillaBlock <sup>[18]</sup>	91.83	88.77	93.16	3.21	6.2	95.26
CovNext <sup>[19]</sup>	92.54	88.52	93.24	50.80	126.4	42.82
SPSBackbone	92.62	<b>91.85</b>	95.06	2.64	7.6	<b>112.47</b>

根据表 3 可以看出:

(1)SPSBackbone 在降低极少准确率的情况下, 拥有

更快的检测速度,其中参数量降低了约14.5%,召回率R也是比较网络中的最佳。

(2)Biformer网络模型虽获得更高的准确性,但在计算速度、检测速度和模型大小上损失过大,无法满足城市道路场景中障碍物检测的实时性要求。

(3)在原始模型的基础上替换了轻量化模型中的经典网络mobileNetV3,在Params与GLOPs方面表现优异,但在P、R及mAP0.5指标上有很大程度降低。

综合比较几个主流特征提取网络的性能,证明改进的SPSBackbone检测效果更好。

## 2)注意力机制对比实验

为了使模型的精确度提高,本文在YOLOv8n基础上添加近年来主流的注意力机制进行对比实验。为了直观地体现每种注意力在MRObstacle数据集上地改进效果,在数据集中随机选取一张图片进行可视化处理,并以热力图的形式进行体现,如图10所示。

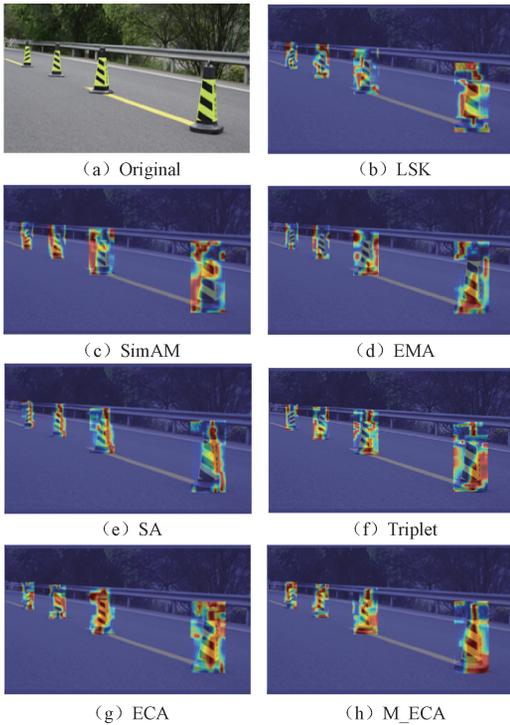


图10 注意力机制热力图效果对比

Fig.10 Attention mechanism heat map effect comparison

根据图10可知,ECA与M\_ECA比其他注意力机制对于道路障碍物的关注度更加明显,M\_ECA对于道路后排小障碍物关注度比ECA更好。

为了进一步验证在MRObstacle数据集上不同注意力机制的效果,在YOLOv8n原始模型上的相同位置引入不同注意力机制进行对比实验。具体实验结果如表4所示。

由表4数据可知:

(1)在精度P指标上,添加注意力机制均有不同程度的提升,其中添加EMA、ECA、M\_ECA与SA注意力机制

表4 注意力对比实验结果

Table 4 Results of the attention comparison experiment

注意力机制	P/%	R/%	mAP@0.5/%
—	92.95	91.12	95.08
LSK <sup>[20]</sup>	93.65	91.32	95.27
SimAM <sup>[21]</sup>	93.42	<b>91.88</b>	95.29
EMA <sup>[22]</sup>	94.12	90.52	95.23
SA <sup>[23]</sup>	94.08	90.85	95.46
Triplet <sup>[24]</sup>	93.32	91.26	95.45
ECA <sup>[25]</sup>	94.24	91.14	95.63
M_ECA	<b>94.29</b>	91.42	<b>95.85</b>

分别提升1.17%、1.29%、1.34%与1.13%。

(2)在mAP方面,添加ECA和M\_ECA注意力机制相对提升较明显,分别提升0.55%和0.77%。

(3)添加SimAM注意力机制使模型的召回率R最高,但P、mAP相比其他注意力机制涨幅并不明显。

为了满足城市道路复杂交通场景下障碍物检测的精确性要求,本文选择实验中检测精度最高的M\_ECA注意力机制作为改进模型的一部分。

## 3)消融实验

为进一步验证本文提出的各改进方法的有效性,对相关改进模块进行消融实验。实验结果如表5所示。

从表5消融实验结果可知:

(1)主干网络中C2f改进成SPS\_C2f后,模型的R和FPS分别提高了0.73%和17.1fps,网络的参数量较原模型下降了约14.5%;添加M\_ECA注意力机制模块使模型的P、mAP两项指标上分别获得1.34%、0.77%的提升;网络融合BiFPN与添加小目标检测头后,由于添加小目标检测层导致模型深度增加,因此模型的计算量增加4.1GFLOPs,同时FPS也略有下降,但模型对于小目标障碍物的检测表现有了明显提升;更换了MPDIoU损失函数后,模型在R、mAP分别提高了0.5%、0.26%。

(2)对比①和⑥实验可知,替换SPS\_C2f主干网络和添加M\_ECA注意力模块后的模型在P、mAP和FPS上分别获得0.77%、0.76%和16.15fps的提高。

(3)实验⑦在融合BiFPN与添加小目标检测头后,虽然增加了一些网络模型计算量,相较于实验⑥,在模型参数量降低约10%的情况下,P、R和mAP分别获得提升。

(4)融合4项改进模块的实验⑧,相较于基础模型,在仅增加少量网络模型的GFLOPs的情况下,P、R、mAP50和FPS分别提升1.61%、1.51%、2.04%和12.08fps,模型参数量降低约11.7%。

综合分析表5的消融实验结果,验证了本文提出的各个改进模块都具有效果。

## 4)不同网络检测性能对比

为了验证本文算法在城市行车道路障碍物检测方面的优越性,同时验证本文自制MRObstacle数据集具有良

表 5 消融实验结果

Table 5 Results of ablation experiments

序号	SPS_C2f	M_ECA	BiFPN+P2	MPDIou	P/%	R/%	mAP@0.5/%	Params/ $10^6$	GFLOPs/G	FPS/fps
①	—	—	—	—	92.95	91.12	95.08	3.09	8.1	95.37
②	√	—	—	—	92.62	91.85	95.06	2.64	7.6	112.47
③	—	√	—	—	94.29	91.42	95.85	3.23	8.3	93.24
④	—	—	√	—	94.63	90.81	95.71	2.92	12.2	90.65
⑤	—	—	—	√	92.97	91.62	95.34	3.09	8.1	96.52
⑥	√	√	—	—	93.72	91.23	95.84	3.01	8.5	111.52
⑦	√	√	√	—	93.96	91.71	96.75	2.73	12.4	108.98
⑧	√	√	√	√	94.56	92.63	97.12	2.73	12.4	107.45

好的泛化性,采用相同的实验环境,将本文改进算法与目前主流的目标检测算法进行检测对比,其中包括二阶段的检测算法 Faster-RCNN 和 SSD<sup>[26]</sup> 网络模型、目前最新的

轻量型 YOLO 系列算法 YOLOv7-tiny、YOLOv10n 和 YOLOv11n。同时与最新的改进 YOLOv8n 检测道路图像的算法进行对比,实验结果如表 6 所示。

表 6 不同网络检测性能对比

Table 6 Performance comparison of different network detection

Base model	Backbone	P/%	R/%	mAP@0.5/%	Params/ $10^6$	GFLOPs/G	FPS/fps
Faster R-CNN	ResNet50	85.92	83.16	85.08	28.09	69.1	32.10
SSD	VGG16	86.37	87.22	87.24	12.32	23.3	45.56
YOLOv7-tiny	CSPDarknet53	89.43	87.26	91.45	7.26	10.8	82.42
YOLOv8n	CSPDarknet	92.95	91.12	95.08	3.09	8.1	95.37
YOLOv9t	CSPDarknet	91.23	88.92	93.62	<b>2.00</b>	7.7	103.42
YOLOv10n	CSPDarknet	93.02	90.36	95.23	2.72	8.4	95.37
YOLOv11n	CSPDarknet	93.36	91.22	95.10	2.60	<b>6.5</b>	99.20
文献[27]	CSPDarknet	89.31	83.36	91.20	11.87	38.7	88.50
文献[28]	CSPDarknet	<b>95.50</b>	86.10	95.90	13.30	12.5	92.06
本文	SPSBackbone	94.56	<b>92.63</b>	<b>97.12</b>	2.73	12.4	<b>107.45</b>

根据表 6 可以看出:

(1)YOLOv9t 的网络参数量均比其他对比网络低的情况下,其 P、R 和 mAP 值相较于本文改进网络,落后 3.33%、3.71% 和 3.5%。

(2)本文提出的改进模型在道路障碍物识别任务中,P、R、mAP 与 FPS 上较最新 YOLOv11n 均具有一定优势。P 值领先 1.2%,R 值领先 1.41%,mAP 值领先 2.02%,FPS 领先 8.25 fps。本文提出的改进模型在各项指标上均优于 Faster R-CNN、SSD 二阶段检测模型。

(3)相较于文献[27]与文献[28]中改进 YOLOv8n 的算法,本文算法仅在 P 值上略微落后,R、mAP、计算复杂度和 FPS 值均有优势,本文算法改进主干网络更具轻量化,因此在参数量指标上具有明显优势。

由此可见,本文所提出算法具有更高的检测精度与速

度,综合性能优于目前主流的检测算法。

#### 5) 算法验证

为了更直观地展示本文所提网络的检测效果,本研究将 YOLOv8n 和改进后的网络的检测障碍物结果对比,进行了可视化分析。如图 11 所示。

(1)通过 A 组实验中图像对比,可以看出改进算法检测障碍物的置信度明显提升,本文改进算法具有更高的识别精度。

(2)根据 B 组实验结果对比,可以看出改进算法可以检测出远处 YOLOv8n 漏检的障碍物目标,证明本文算法可以改善对于远处较小目标的障碍物识别不佳的问题。

(3)通过 C 组实验结果对比,可以看出改进算法对于障碍物目标定位更准确,可得到本文算法可以改善障碍物目标定位不准确以及目标特征表达不充分的问题。

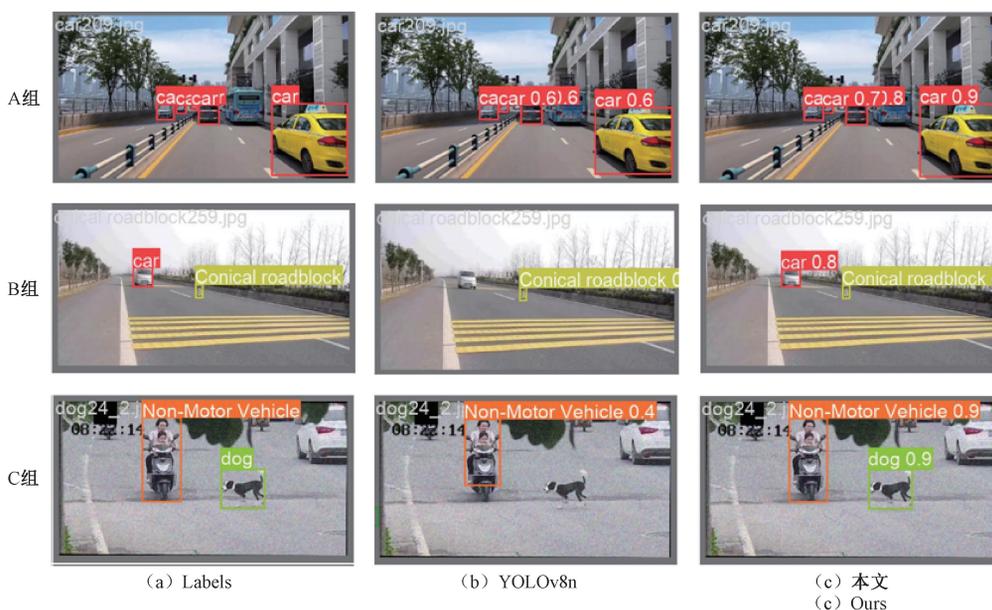


图 11 改进前后检测结果对比

Fig. 11 Comparison of test results before and after improvement

### 3 结 论

本文提出了一种改进的YOLOv8障碍物目标检测方法。设计SPS\_C2f改进主干网络,提高模型的检测速度;添加M\_ECA注意力机制模块、融合BiFPN和添加小目标检测分支,增强网络对于障碍物的特征提取效果;使用MPDIoU作为损失函数,提升网络的边界框回归性能。改进后的模型在自制MRObstacle数据集上实验结果表明,与YOLOv8n相比,本文提出的方法在精度、平均精度、召回率和FPS指标上分别提升1.61%、1.51%、2.04%和12.08fps,同时降低了模型的参数量,所提方法能够很好的胜任城市道路场景下障碍物目标检测任务。

未来将考虑到复杂道路场景下易出现遮挡型障碍物目标,存在像素重叠问题,以及障碍物目标出现旋转和同一障碍物目标出现缩放的问题,后续研究会针对该类问题做进一步改进。

### 参考文献

- [1] 张新钰,高洪波,赵建辉,等.基于深度学习的自动驾驶技术综述[J].清华大学学报(自然科学版),2018,58(4):438-444.  
ZHANG X Y, GAO H B, ZHAO J F, et al. A review of deep learning based autonomous driving technologies [J]. Journal of Tsinghua University (Natural Science Edition), 2018, 58(4): 438-444.
- [2] 冉险生,李锐,贺帅.基于改进YOLOv5s的道路障碍物检测算法[J].电子测量技术,2023,46(22):177-185.  
RAN X SH, LI R, HE SH. Road obstacle detection algorithm based on improved YOLOv5s [J].

Electronic Measurement Technology, 2024, 46(22): 177-185.

- [3] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015:1440-1448.
- [4] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with regionproposal networks[C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, Montréal, 2015:91-99.
- [5] HUU P N, THI Q P, QUYNH P T. Proposing lane and obstacle detection algorithm using YOLO to control self-driving cars on advanced networks[J]. Advances in Multimedia, 2022(5), DOI: 10.1155/2022/3425295.
- [6] HUANG P L, WANG S H, CHEN J Y, et al. Light weight model for pavement defect detection based on improved YOLOv7[J]. Sensors, 2023, 23(16): 7112.
- [7] JI B, XU J F, YANG L, et al. Improved YOLOv8 for small traffic sign detection under complex environmental conditions[J]. Franklin Open, 2024, 8: 100167.
- [8] YAO J L, XU S, HUANG F J, et al. Improved lightweight infrared road target detection method based on YOLOv8 [J]. Infrared Physics & Technology, 2024,141:105497.
- [9] AN R, ZHANG X CH, SUN M P, et al. GC-YOLOv9: Innovative smart city traffic monitoring solution[J]. Alexandria Engineering Journal, 2024, 106: 277-287.

- [10] CHEN J, KAO S H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2023, 23(7): 192023.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132-7141.
- [12] WU Y T, TANG S H M, ZHANG S H W, et al. An enhanced feature pyramid object detection network for autonomous driving [J]. Applied Sciences, 2019, 9(20):4363.
- [13] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020:10781-10790.
- [14] MA S L, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662,2023.
- [15] ZHU L, WANG X J, KE ZH H, et al. BiFormer: Vision transformer with bi-level routing attention[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323-10333.
- [16] 张奇,陈梦蝶,赵杰.融合特征金字塔和通道注意力的轻量车辆检测算法[J].国外电子测量技术,2023,42(1):41-48.  
ZHANG Q, CHEN M D, ZHAO J. Lightweight vehicle detection network fusing feature pyramid and channel attention[J]. Foreign Electric Measurement Technology, 2023,42(1):41-48.
- [17] 郝巨鸣,杨景玉,韩淑梅,等.引入Ghost模块和ECA的YOLOv4公路路面裂缝检测方法[J].计算机应用,2023,43(4):1284-1290.  
HAO J M, YANG J Y, HAN S H M, et al. YOLOv4 highway pavement crack detection method using Ghost module and ECA [J]. Journal of Computer Applications, 2023, 43(4): 1284-1290.
- [18] CHEN H T, WANG Y H, GUO J Y, et al. VanillaNet: The power of minimalism in deep learning[J]. Advances in Neural Information Processing Systems, 2024, 16(12):36-51.
- [19] LIU ZH, MAO H Z, WU CH Y, et al. A ConvNet for the 2020s[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022(6):11966-11976.
- [20] LI Y X, HOU Q B, ZHENG ZH H, et al. Large selective kernel network for remote sensing object detection[C]. 2023 IEEE/CVF International Conference on Computer Vision(ICCV), 2023:16748-16759.
- [21] YANG L X, ZHANG R Y, LI L D, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks[C]. 38th International Conference on Machine Learning(ICML), 2021:11863-11874.
- [22] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]. 48th IEEE International Conference on Acoustics, Speech and Signal Processing, 2023:1-5.
- [23] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks [C]. 2021 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP), 2021:2235-2239.
- [24] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: Convolutional triplet attention module [C]. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2020: 3138-3147.
- [25] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [26] BERG A C, FU CH Y, SZEGEDY C, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, 2016:21-37.
- [27] 周飞,郭杜杜,王洋,等.基于改进YOLOv8的交通监控车辆检测算法[J].计算机工程与应用,2024,60(6):110-120.  
ZHOU F, GUO D D, WANG Y, et al. Vehicle detection algorithm based on improved YOLOv8 in traffic surveillance [J]. Computer Engineering and Applications, 2024, 60(6): 110-120.
- [28] WANG B, LI Y Y, XU W J, et al. Vehicle-pedestrian detection method based on improved YOLOv8 [J]. Electronics, 2024, 13(11): 2149.

## 作者简介

向雷,硕士研究生,主要研究方向为基于深度学习的目标检测算法。

E-mail:xianglei@stu.xhu.edu.cn

蒋文波(通信作者),博士,教授,主要研究方向为深度学习算法与应用、光学信息处理方法与技术。

E-mail:jiangwenbo@mail.xhu.edu.cn