

基于改进 YOLOv8 的大件货物运输目标障碍物检测方法

谢兴国¹, 闵丹^{1,*}, 江宇星², 况雲赟², 邹佳鑫², 魏存鑫²

1. 四川省大件运输有限公司, 四川 成都, 610000

2. 西华大学, 汽车与交通学院, 四川 宜宾, 644000

摘要: 大件货物运输过程中障碍物类型多样、环境复杂, 传统目标检测算法在识别精度和实时性方面存在不足。为实现复杂运输场景下障碍物的快速、精准识别, 本文在 YOLOv8 模型基础上进行了结构改进。通过在 Backbone 与 Neck 层引入多残差模块 (MRM)、FasterC2f 模块 (含部分卷积 PConv) 和空间注意力机制 (SA), 分别提升特征提取能力、优化特征融合效率并强化显著区域响应。基于自建大件运输障碍物数据集开展系统的消融实验与对比实验。实验结果表明, 改进模型在 mAP@0.5 和 mAP@0.5:0.95 上分别提升 4.4% 和 7.5%, F1 值提高 4 个百分点; 与 YOLOv5、YOLOv7、Faster R-CNN 等主流算法相比, 改进 YOLOv8 的 mAP@0.5 达 0.911, mAP@0.5:0.95 为 0.712, 推理速度提升至 108 帧/秒。研究表明, 本文提出的改进 YOLOv8 模型在检测精度、实时性及鲁棒性方面均显著优于传统算法, 能够有效应对光照变化、遮挡及多尺度目标检测任务, 为大件运输障碍物的智能识别与安全监测提供了高效可靠的技术支撑。

关键词: 大件运输; YOLOv8; 多残差模块; 空间注意力机制; 障碍物检测

Large Cargo Transport Obstacle Detection Method Based on Improved YOLOv8

Xingguo Xie¹, Dan Min^{1,*}, Yuxing Jiang¹, Yunyun Kuang¹, Jiaxin Zou¹, Cunxin Wei¹

1. Sichuan Oversize Transport Company Limited, Chengdu, Sichuan, 610000, China

2. Department of Automobile and Transportation, Xihua University, Yibin, Sichuan, 644000, China

Abstract: During the transportation of oversized cargo, obstacle types are diverse and the environment is complex, presenting challenges for traditional object detection algorithms in terms of both recognition accuracy and real-time performance. To achieve rapid and accurate obstacle identification in complex transport scenarios, this paper introduces structural improvements based on the YOLOv8 model[1]. By integrating a Multi-Residual Module (MRM) in the Backbone, a FasterC2f module (incorporating Partial Convolution PConv) in the Neck, and a Spatial Attention (SA) mechanism, we enhance feature extraction capability, optimize feature fusion efficiency, and strengthen responses in salient regions, respectively. Systematic ablation and comparative experiments were conducted using a self-built dataset of obstacles in oversized cargo transportation. Experimental results demonstrate that the improved model achieves increases of 4.4% and 7.5% in mAP@0.5 and mAP@0.5:0.95, respectively, with a 4 percentage point improvement in the F1-score. Compared to mainstream algorithms such as YOLOv5, YOLOv7, and Faster R-CNN, the enhanced YOLOv8 model reaches a mAP@0.5 of 0.911 and a mAP@0.5:0.95 of 0.712, with an inference speed increased to 108 frames per second. The research indicates that the improved YOLOv8 model proposed in this paper significantly outperforms traditional algorithms in detection accuracy, real-time performance, and robustness. It effectively handles challenges such as illumination variations, occlusions, and multi-scale object detection tasks, providing efficient and reliable technical support for the intelligent identification and safety monitoring of obstacles in oversized cargo transport.

Keywords: Oversized Cargo Transport; YOLOv8; Multi-Residual Module; FasterC2f; Spatial Attention Mechanism; Obstacle Detection

大件运输是指车货从轮胎起高度超过 4 米,车货宽度超 2.55 米,车货长度超过 18.1 米的运输情况。车货规格只要超过以上情况的任一情况就属于大件运输。在运输过程面临许多未知因素(因车货超限且受环境影响成分很大),以及障碍物、减速带、路况等情况都可能对运输过程造成严重的财产损失及安全事故。目前,大多数大件运输在运输过程中采用多部门协调配合与雷达检测障碍物进行运输,雷达检测采用双回波进行检测障碍物^[1],但雷达检测在复杂天气(雾天、雨天)下可能漏检及忽略慢速或者静止障碍物。尽管雷达检测对动态障碍物的识别有巨大提升,但在对行人、车辆、障碍物、路况的图像检测仍有着很大的提升空间。人眼观察对人、车、路、环境任一因素未考虑到将会对安全、财产造成极大的损失^[2]。

随着人工智能的兴起,YOLO(You Only Look Once)目标检测在识别物体及信息捕捉方面有着极大的作用,它通过将图片或者视频动画处于检测中心对动画进行识别、卷积输出结果得到想要的目标信息。帮助驾驶员在驾驶途中提前获取道路信息。

从 2016 年 YOLOV1 问世以来通过不断的优化检测模型、卷积的输出为目标检测极大提升了检测准确性及应用适配性^[3]。随着深度学习的不断发展,众多学者将 YOLO 系列模型运用于交通领域,将其用于目标的检测、识别物体。马晓君等^[4]基于 YOLOv5 模型,对该模型进行改进及引入坐标注意力检测机制等对高速列车运行途中障碍物的检测。代少升等^[5]采用 Inner-CIoU 作为边框函数改进 YOLOv8n 的盲区目标检测算法,降低复杂场景下的漏检率与误检率。王巍等^[6]提出基于机器视觉的驾驶员视野盲区安全预警方法研究,通过对行人特征提取及行人特征融合对车辆盲区进行检测与预警。诸多作者将计算机视觉运用于交通领域,对交通领域做出极大贡献,但仅仅局限于某一种目标物检测或在某一种极端天气情况下进行检测,未提出更多场合的融合及各种障碍物的识别。识别算法较为单一,而实际交通场景及大件运输场景的情况是多种多样的,遇到障碍物及极端天气的情况亦是多种多样的。

因此,本研究基于改进 YOLOv8 模型,通过对数据集增强以对各种场景(雨天、雾天、雪天)进行检测及应对。引入多残差模块 Efficient Model(EMO),同时融合了 Transformer 的有效组件。通过结合,EMO 实现统一视角处理轻量级模型的设计,创新地将 CNN 和注意力机制相结合。引入多残差模块会加大模型的参数量和计算量,在此基础上引入空间注意力机制,空间注意力机制可提升模型对小目标、遮挡物的检测精度。将两者融合,可将被检测物体浅层和深层特征进行融合从而提高检测精度。在模型的 Neck 部分,将 c2f 模块替换为 fasterc2f, Fasterc2f 是一种轻量化的神经网络架构,其可以在提高计算速度而不降低检测准确率。通过对原始 YOLOv8 模型三个模块的改进,将三个部分进行结合处理,可以在提升目标检测速度的同时也提高了目标检测的精确率。

1 YOLOv8 算法的改进

1.1 网络结构概述

YOLOv8 模型属于典型的单阶段目标检测网络,整体结构由 Backbone、Neck 和 Detection Head 三个部分组成。

Backbone(主干网络)负责对输入图像进行多层次特征提取,逐级获取从低级纹理到高级语义的多尺度信息。该部分采用基于 C2f 模块的分层卷积结构,并通过 CBS(Conv-BN-SiLU)组合实现高效特征表达。Neck(颈部结构)采用路径聚合网络思想,实现特征的多尺度融合。Neck 连接 Backbone 与检测头,在自顶向下与自底向上路径中交替进行特征传递,使浅层的细节特征与深层的语义特征得以充分融合。

Detection Head（检测头）采用解耦式结构，将回归分支与分类分支分离处理，使用 Anchor-Free 机制实现端到端的预测，提高模型的检测精度与收敛稳定性。

在大件运输场景中，检测对象往往具备尺度跨度大、结构细长、部分遮挡、背景复杂等特征。传统 YOLOv8 的标准结构在使用过程中依赖建模、跨尺度融合以及对显著目标区域的响应上存在不足。针对这些问题，本文从 Backbone 与 Neck 两个层面进行优化，引入了三项关键改进：多残差模块（MRM）、FasterC2f 模块（含部分卷积 PConv）、空间注意力机制（SA）。改进后的网络结构图如图 1 所示。

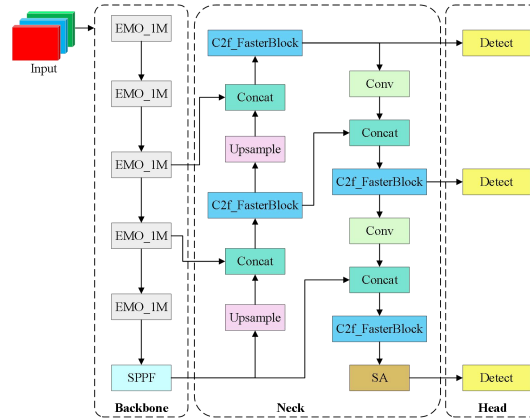


图 1 改进 YOLOv8 的核心模块结构框图

Fig. 1 Architecture of the core improved YOLOv8 module.

1.2 多残差模块

传统 YOLOv8 模型的骨干网络在浅层主要提取局部纹理特征，深层则侧重于语义信息整合，在大件运输场景中，对大型目标的整体结构与轮廓建模能力存在不足。为此，本文提出一种多残差模块（Multi-Residual Module, MRM），嵌入于 YOLOv8 的骨干网络中以增强多尺度特征关联与长程依赖建模能力。该模块采用双分支架构：主分支沿用反向残差块（IRB）的设计^[7]通过“ 1×1 卷积- 3×3 深度卷积- 1×1 卷积”实现特征压缩与重构，保障信息的高效流动；辅助分支则融合跨尺度可分离卷积与通道注意力机制，以捕获更丰富的空间上下文信息^[8]。两分支输出通过门控机制进行自适应融合，门控权重由轻量化子网络动态学习，从而在保留局部细节的同时强化结构感知与尺度适应能力。

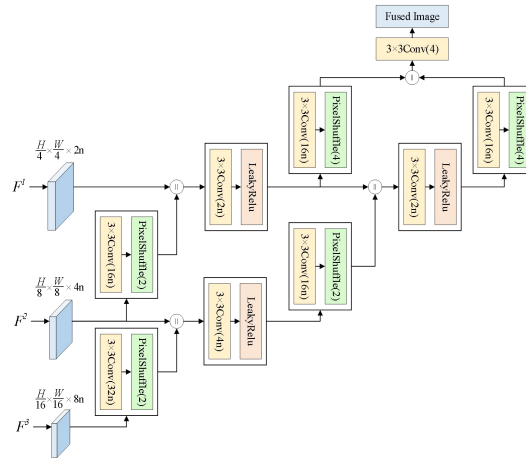


图 2 多残差模块（MRM）结构示意图

Fig. 2 Structure of the Multi-Residual Module (MRM)

具体实现中，如图 2 将原网络中高层 C2f 模块替换为 MRM，使模型在浅层保持纹理提取能力，而在

中高层提升结构特征的代表质量。主分支采用可扩展的瓶颈比控制参数规模,辅助分支利用多尺度卷积增强空间上下文建模。门控融合基于全局池化与 Sigmoid 函数生成归一化权重,确保特征融合的平衡性与训练稳定性。各卷积层均配备批归一化与 SiLU 激活函数,以维持数值稳定。该改进在增强模型整体表达能力的同时,保持了与原 YOLOv8 架构的兼容性。

1.3 Neck 改进: FasterC2f 模块(含部分卷积 PConv)

针对 YOLOv8 颈部网络中 C2f 模块在高分辨率特征融合时存在计算冗余的问题,本文提出一种高效的 FasterC2f 结构。该模块基于部分卷积(PConv)思想^[9]。

如图 3,基于部分卷积(PConv)的 FasterNet 模块采用结构化通道划分策略以优化计算效率。该模块将输入特征图沿通道维度划分为两部分:仅选取前 C_P 个通道(通常 $C_P = C/4$)作为主分支,执行 3×3 空间卷积以提取局部特征;其余通道则通过恒等映射直接传递。两部分特征在通道维度进行拼接后,依次经由 1×1 卷积、线性瓶颈结构与 ReLU 激活函数实现特征重构,最终通过残差连接与输入特征融合。该设计通过减少参与空间卷积的通道数量,显著降低了计算复杂度,同时保持了特征表达的完整性,实现了精度与效率的有效平衡。

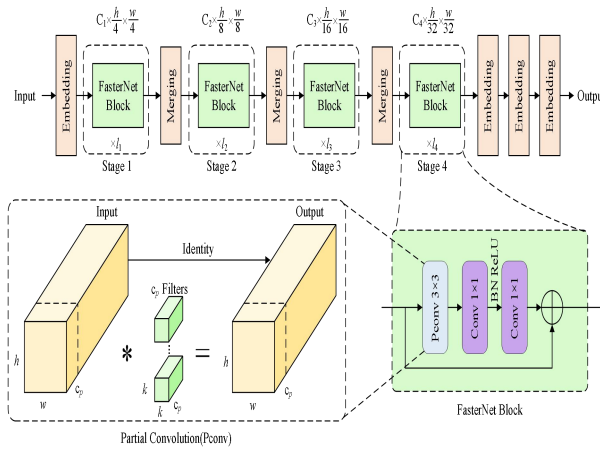


图 3 FasterC2f 模块(含部分卷积 PConv)结构示意图

Fig. 3 Structure of the proposed Faster C2f module (with PConv)

1.4 Neck 改进: 空间注意力机制 SA

在大件运输环境中,存在大量复杂背景与噪声干扰,例如路边的标识物、地面反光或空中线缆等。这些因素常导致小目标或部分遮挡目标的检测困难。为增强模型对目标区域的关注度,本文在 Neck 融合阶段引入了空间注意力机制(SA)^[10]。SA 的核心思想是通过学习空间维度上的权重分布,自适应地调整不同位置特征的重要性,从而让网络更加专注于潜在目标区域。

FasterC2f 模块在结构上继承了 C2f 的多分支特征聚合形式,但通过通道划分机制将输入特征图分为两部分:一部分执行标准空间卷积以提取局部空间信息,另一部分直接通过逐点卷积(1×1 卷积)或恒等映射快速传递全局语义特征。随后,两路特征在通道维度拼接,并通过残差连接与逐点卷积进行再融合。设参与空间卷积的通道比例为 p ,则网络的空间卷积计算量与显存访问量近似按 p 比例缩减,从而显著提升运算效率。该模块既保持了 YOLOv8 原有的跨层复用特性,又在高分辨率特征融合中引入了更优的计算分配策略。

空间注意力模块通过对输入特征图在通道维度上分别进行平均池化和最大池化,获得两张反映不同统计特性的二维图像。两者在通道维上拼接后,通过一个 7×7 卷积层捕获局部空间相关性,输出经 Sigmoid

激活函数映射为空间权重矩阵。最终，该权重矩阵与原特征图进行逐元素乘法，实现显著区域的增强与无关区域的抑制。公式如下：

$$F'(x, y, c) = A(x, y) \times F(x, y, c) \quad (1)$$

其中， $A(x, y)$ 为空间注意力权重，取值范围[0,1]。该机制在不增加大量参数的前提下，使模型在特征融合阶段具备空间选择性，从而提高了特征图的辨别力与目标响应强度。

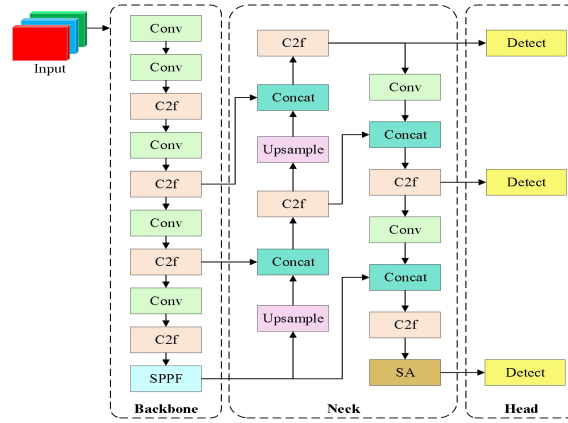


图 4 空间注意力机制 (SA) 结构示意图

Fig 4. Structure of the Spatial Attention (SA) mechanism

在网络实现中，如图 4 SA 模块被嵌入到 Neck 的多尺度融合节点中，具体位于上采样与横向连接后的拼接层之后、进入后续卷积前的位置。此时的特征图包含丰富的语义与细节信息，适合通过空间加权进一步强化目标区域。卷积核大小通常设为 7×7 ，当分辨率较高或资源受限时可缩减至 5×5 。模块输出无需额外归一化操作，以减少计算开销。该机制与 FasterC2f 模块协同作用，使模型在保持高效计算的同时，能够对重要空间位置进行重点关注，从而实现高鲁棒性检测。

2 实验结果与分析

2.1 数据集构建

针对大件运输场景中复杂障碍物识别的需求，本文构建了专用数据集。数据集共计 8,200 张图片，划分为训练集 (70%)、验证集 (20%)、测试集 (10%)。类别包括，风电叶片、运输车辆、工程机械、建筑构件、限高杆、交通锥、警示标志及路障等。图像分辨率统一为 1280×720 ，数据采集涵盖车载摄像头、无人机及固定监控等多视角场景。采集环境覆盖白天、夜晚、雨雪、雾霾等多种气候条件，以保证模型的泛化能力。

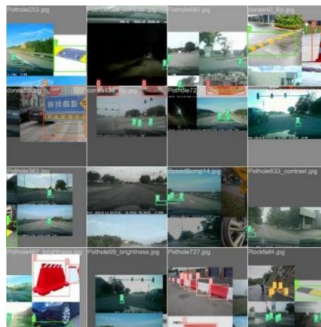


图 5 自建数据集部分样例图像

Fig 5. Sample images from the custom dataset

2.2 实验环境与参数设置

2.2.1 软件环境

深度学习框架使用 PyTorch 2.2+CUDA 12.1, Python 版本为 3.10。训练与可视化使用 TensorBoard、Matplotlib。

2.2.2 训练参数

输入图像统一调整为 640×640。批量大小设置为 16, 优化器采用 AdamW (权重衰减 1e-4), 初始学习率为 0.001%, 并使用余弦退火 (Cosine Annealing) 策略进行动态调整, warmup 3 epoch。

总训练轮次为 300, 启用 EMA (指数滑动平均) 以稳定模型收敛。

数据增强包括随机缩放、裁剪、亮度扰动、颜色抖动、仿雾化与仿雨雪增强等。损失函数采用 Focal Loss (分类)^[11]与 WIoU (边界框回归) 联合优化^[12]。

2.2.3 推理与评估设置

推理阶段使用 PyTorch 原生推理, 不开启 TensorRT 加速。NMS (非极大值抑制) 阈值设为 IoU=0.5, 置信度阈值为 0.25。所有速度指标均在单卡 (batch=1) FP16 精度下测得。

2.3 评估指标

模型性能评价采用以下指标:

(1) Precision (精确率):

$$P = \frac{TP}{TP + FP} \quad (2)$$

表示模型预测为正样本的目标中, 有多少是真实目标。

(2) Recall (召回率):

$$R = \frac{TP}{TP + FN} \quad (3)$$

表示真实目标中有多少被模型检测出。

(3) mAP@0.5:IoU=0.5 阈值下的平均精度, 反映模型总体检测能力。

mAP@0.5:0.95:在 IoU∈[0.5,0.95]区间内的平均精度, 更严格地衡量模型的鲁棒性。

F1-score: 综合考虑 P 和 R 的调和平均:

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (4)$$

FPS (帧率): 每秒处理帧数, 衡量模型的实时检测能力。

2.4 消融实验

为了验证本文所提出的三项结构改进, 设计了系统性的消融实验来验证 YOLOv8 模型性能提升的具体作用及其相互协同关系

通过在相同训练条件与数据集下, 逐步引入各个模块并比较性能指标的变化, 分析每个改进模块的独立贡献及其在联合应用时的叠加效果, 从而确定各结构在检测精度、鲁棒性与推理速度方面的影响。

2.4.1 多残差模块

该模块替换原 YOLOv8 主干网络中的标准 C2f 块,其核心设计思想是通过多路径残差结构增强特征的多层表达能力。MRM 模块包含主分支与辅助分支两条路径,主分支采用标准的卷积+BN+SiLU 激活组合,用于局部特征提取;辅助分支采用空洞卷积(dilated conv)与深度可分离卷积组合,以扩大感受野并捕获跨尺度特征;两分支输出通过残差连接与逐通道加权融合,从而提升特征层的表示密度与语义一致性。在推理阶段,该模块不会增加额外计算量,仅略微增加参数数量,整体计算复杂度保持轻量。

2.4.2 FasterC2f 模块

在 Neck 层的特征融合阶段,将原 YOLOv8 的 C2f 模块替换为 FasterC2f 模块。引入部分卷积(Partial Convolution, PConv)机制,仅对部分通道执行卷积操作,减少计算冗余并保持主要特征通路。设输入特征为 $X \in R^{C \times H \times W}$,将通道划分为两部分:

$$X = [X_a, X_b], X_a \in R^{\rho C \times H \times W}, X_b \in R^{(1-\rho)C \times H \times W} \quad (5)$$

其中, ρ 表示卷积比例(本文取 $\rho=0.5$)。仅对 X_a 进行卷积计算 X_b 直接连接或经逐点卷积再融合。最终输出为:

$$Y = \text{Concat}(\text{Conv}(X_a), X_b) \quad (6)$$

这种结构显著降低了计算量,同时保持跨尺度特征流动的连续性。

2.4.3 SA 模块

为增强模型在复杂背景和小目标场景下的判别能力,在每个 Neck 层融合节点后引入空间注意力机制。对输入特征 $F \in R^{C \times H \times W}$,分别进行通道维度的平均池化与最大池化,得到两个单通道特征图 F_{avg} 与 F_{max} ,随后将其拼接并经过 7×7 卷积与 Sigmoid 激活函数得到空间注意力权重图 $A(x, y)$ 。

$$A = \sigma \left(f^{7 \times 7} \left(\left[\text{AvgPool}(F); \text{MaxPool}(F) \right] \right) \right) \quad (7)$$

输出特征为:

$$F' = A \odot F \quad (8)$$

其中“ \odot ”表示逐元素乘法。通过该机制,模型能够自适应地为不同空间位置分配权重,突出目标显著区域,抑制无关背景干扰。

引入 MRM 模块后,模型结构感知增强,对大目标边缘轮廓保持更稳定。加入 FasterC2f(含 PConv)后,高分辨率层特征融合效率提高, mAP@0.5:0.95 提升 3.8%引入 SA 模块后,模型在小目标和遮挡物检测上明显改善, F1 提升 0.03%。

为验证三项改进模块的有效性,在 YOLOv8n 基线模型上逐步加入 MRM、FasterC2f(含 PConv)和 SA,其性能变化如表 1 所示。

表 1 消融实验结果

Table 1. Ablation study results.

模型结构	Precision	Recall	mAP@0.5	mAP@0.5:0.95	F1
YOLOv8n(基线)	0.897	0.877	0.867	0.637	0.83
+MRM(多残差模块)	0.893	0.798	0.881	0.657	0.84
+FasterC2f(PConv)	0.921	0.799	0.899	0.695	0.85
+SA(空间注意力)	0.910	0.823	0.900	0.693	0.86
+MRM+FasterC2f+SA	0.932	0.843	0.911	0.712	0.87

由表 1 可得出, MRM 模块使 $mAP@0.5$ 提升 1.4%, $mAP@0.5:0.95$ 提升 2.0%, 验证了其在特征提取与结构建模方面的优势。Faster C2f 模块带来 $mAP@0.5$ 提升 3.2%, $mAP@0.5:0.95$ 提升 5.8%, 在保持精度的同时显著提高了推理速度。SA 模块进一步提升小目标检测表现, F1 值上升至 0.86, 说明注意力机制有效增强了显著区域响应。

当三模块联合应用时, 模型在所有指标上达到最优, Precision=0.932、 $mAP@0.5:0.95=0.712$, 整体提升最显著。这表明三者功能上互为补充, MRM 提供结构表达基础, FasterC2f 优化特征传输效率, SA 聚焦显著目标。三者协同实现了结构增强、融合高效与空间聚焦的统一优化目标。

由图 6 可知, 引入 MRM 模块后, 模型结构感知增强, 对大目标边缘轮廓保持更稳定; 加入 FasterC2f (含 PConv) 后, 高分辨率层特征融合效率提高, $mAP@0.5:0.95$ 提升 3.8%; 引入 SA 模块后, 模型在小目标和遮挡物检测上明显改善, F1 提升 0.03%。

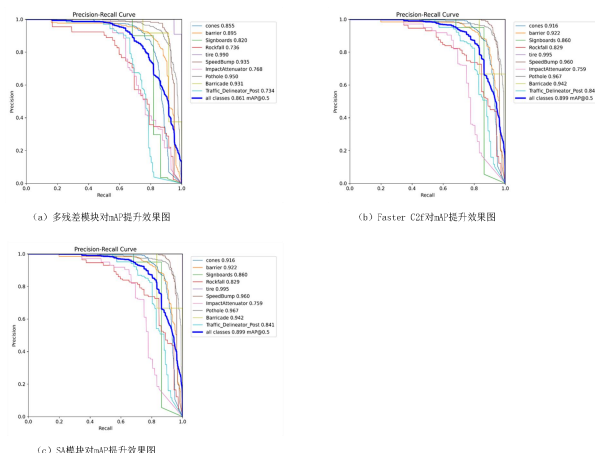


图 6 各模块对 mAP 值的提升趋势曲线图

Fig 6. mAP improvement trends of different modules

综上所述, 通过消融实验验证了多残差模块 MRM、FasterC2f (含 PConv) 与空间注意力机制 (SA) 的联合改进使 YOLOv8 在大件运输障碍物检测任务中实现“高精度、高效率、高鲁棒性”的统一, 为复杂工业环境下的实时目标检测提供了一种轻量化且高性能的可行方案。

2.5 对比实验

为全面验证所提出改进 YOLOv8 模型 (引入多残差模块 MRM、Neck 层 FasterC2f 含 PConv 与空间注意力机制 SA) 在大件运输障碍物识别任务中的性能优势, 本文在统一数据集与实验条件下, 与主流检测模型 YOLOv5s、YOLOv7-tiny、Faster R-CNN 进行了横向对比实验。

所有对比模型均在同一自建大件运输障碍物数据集上训练和测试, 采用相同的数据增强策略、输入尺寸 (640×640)、批量大小 (batch size=16) 和优化策略。训练轮次均为 300, 优化器采用 AdamW, 学习率初始值为 0.001%, 并使用余弦退火 (Cosine Annealing) 策略进行动态调整。推理阶段采用相同的非极大值抑制 (NMS) 阈值 (IoU=0.5) 与置信度阈值 (0.25)。硬件平台为 NVIDIA RTX 4090 GPU, 软件环境为 PyTorch 2.2+CUDA 12.1。在 YOLOv8n 基础上引入 MRM、FasterC2f (含 PConv) 与 SA 模块, 以同时增强特征表达力、融合效率与空间显著性。实验结果如下表 2:

表 2 不同检测模型在大件运输障碍物数据集上的性能对比结果

Table 2. Performance comparison of different detection models on the large-cargo transport obstacle dataset.

模型	Precision	Recall	mAP@0.5	mAP@0.5:0.95	FPS
Faster R-CNN	0.871	0.832	0.845	0.621	38
YOLOv5s	0.902	0.851	0.876	0.652	85
YOLOv7-tiny	0.911	0.868	0.892	0.667	90
改进 YOLOv8 (MRM+FasterC2f+SA)	0.932	0.843	0.911	0.712	108

从表中可以看出,改进 YOLOv8 在 $mAP@0.5=0.911$ 、 $mAP@0.5:0.95=0.712$ 两项关键指标上均达到最高值,相较于 YOLOv7-tiny 分别提升 2.1%与 4.5%,相较于 Faster R-CNN 提升幅度更达 6.6%与 9.1%。Precision 最高,为 0.932,说明模型在复杂背景下的误检率最低,目标识别结果最为稳定。尽管本文模型在 Recall 上略低于 YOLOv7 (0.843 vs 0.868),但由于 Precision 提升明显,其综合 F1 值更高;其 Recall 略低主要与阈值设置及部分复杂遮挡场景下的置信度分布有关,若在部署端适当降低置信度阈值,可进一步提高召回率。从 FPS 对比来看, Faster R-CNN 仅 38FPS,难以满足实时检测要求; YOLOv5s 与 YOLOv7-tiny 分别为 85FPS 与 90FPS,而改进 YOLOv8 达到 108FPS,提升约 20%。这得益于 FasterC2f 模块中部分卷积(PConv)的引入,使特征融合过程中仅对关键通道进行卷积计算,大幅减少冗余计算量,实现了轻量化与高效化的兼顾。

综上所述,该对比实验验证了本文提出的改进 YOLOv8 模型优于其他模型,尤其在 $mAP@0.5:0.95$ 与 FPS 上取得显著优势。

3 讨论

为实现大件货物运输场景下障碍物的快速精准识别,本文在 YOLOv8 模型基础上提出了一种改进的障碍物检测算法。通过在 Backbone 和 Neck 层中引入多残差模块(MRM)、FasterC2f 模块(含部分卷积 PConv)与空间注意力机制(SA),实现了对网络结构的轻量化优化和特征提取能力的增强。

在实验验证中,消融实验结果表明,三项改进模块均能提升模型的检测性能和鲁棒性,其中联合使用时模型表现最优, $mAP@0.5$ 提升 4.4%, $mAP@0.5:0.95$ 提升 7.5%, F1 值提升 4 个百分点。在对比实验中,改进 YOLOv8 模型的 $mAP@0.5$ 与 $mAP@0.5:0.95$ 分别达 0.911 和 0.712,推理速度(FPS)达 108 帧/秒,相较 YOLOv7-tiny 和 Faster R-CNN 等主流检测算法具有更高的检测精度与实时性能。

综合结果表明,改进 YOLOv8 模型在大件运输障碍物识别中实现了高精度、高效率和高鲁棒性的统一,能在复杂光照、遮挡及多尺度场景下保持稳定表现,具备良好的工程应用潜力。未来研究可进一步结合多模态传感、知识蒸馏与时序信息,拓展算法在智能运输与工业检测领域的应用。

参考文献

- [1] 代少升,周漫,余自安,等.基于改进 YOLOv8n 的货车盲区目标检测方法[J].半导体光电, 2025, 46(01): 172-179.
- [2] 李维东,郭蕊,张磊,等.基于 PCA 改进层次分析法的多目标大型货物运输通道选择[J].中国安全生产科学技术, 2021, 17(02): 135139.
- [3] 岑伟迪,江佳玲,黄勃,等. YOLO 十年演进:从实时检测先锋到多任务智能前沿[J/OL]. 武汉大学学报:理学版, 1-16[2025-10-17].
- [4] 马晓君,王栋,刘德胜,等.基于改进 YOLOv5 的降雪天气高速列车障碍物检测[J].计算机仿真, 2025, 42(01): 155-161+451.
- [5] 代少升,周漫,余自安,等.基于改进 YOLOv8n 的货车盲区目标检测方法[J].半导体光电, 2025, 46(01): 172-179.
- [6] 王巍,白龙,马欢畅,等.基于机器视觉的驾驶员视野盲区安全预警方法研究[J].计算机科学, 2023, 50(S1): 498-504.

- [7] Zhang J, Li X, Li J, et al. Rethinking mobile block for efficient attention-based models[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2023: 1389-1400.
- [8] Vo D M, Lee S W. Semantic image segmentation using fully convolutional neural networks with multi-scale images and multi-scale dilated convolutions[J]. Multimedia tools and applications, 2018, 77(14):18689-18707.
- [9] Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 12021-12031.
- [10] 王新, 王赛. 基于改进 YOLOv5 算法的交警手势识别[J]. 电子测量技术, 2022, 45(02): 129-134.
- [11] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.
- [12] Tong Z, Chen Y, Xu Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. arXiv preprint arXiv: 2301.10051, 2023.

¹ **第1作者简介:** 谢兴国 (1973-), 男, 工程师, 研究方向: 大件运输。E-mail: linlangang@scdj-trans.com。

*** 通讯作者简介:** 闵丹 (1988-), 女, 硕士, 工程师, 研究方向: 大件运输。E-mail: 0720240001@xhu.edu.cn。