

文章编号:1007-2780(2026)09-1297-18

面向复杂道路场景的轻量化路面异物检测算法

张传伟^{1,2}, 尚召斌^{1*}, 张佳乐¹, 许 婧¹, 姚 豪¹

(1. 西安科技大学 机械工程学院, 陕西 西安 710054;

2. 陕西交通职业技术学院, 陕西 西安 710018)

摘要:针对车载前视复杂道路场景中部分路面异物成像尺度较小、边缘特征弱及背景干扰强的问题,提出基于YOLO11n的轻量化路面异物检测算法YOLO11-MSDF。设计小波频率解耦下采样(WFD-Down)模块,以高低频解耦保留边缘与纹理细节;构建方向空间感知金字塔快速池化(DSA-SPPF)模块,利用条形池化抑制结构化背景干扰;引入形态自适应C3(MAK-C3)模块,通过自适应采样增强不规则边界特征;设计层级非对称动态融合(HADF)模块,实现浅层细节与深层语义融合。实验结果表明,模型参数量为2.9M、计算量为7.5 GFLOPs,精确率(Precision)、召回率(Recall)和mAP@0.5分别为85.8%、73.9%和83.7%,较YOLO11n算法提高7.1、2.0和4.0个百分点,检测速度为90.5 FPS;外部数据集上3项指标分别提高5.2、1.3和1.9个百分点。合成雾和运动模糊实验表明,模型在多数退化场景下优于基线模型;车辆运行约束分析说明了实际车载应用的主要影响因素。

关键词:深度学习;路面异物检测;轻量化模型;特征融合;目标检测

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3724/CJLCD.2026-0129 **CSTR:**32172.14.CJLCD.2026-0129

Lightweight detection of foreign objects on road surfaces in challenging driving environments

ZHANG Chuanwei^{1,2}, SHANG Zhaobin^{1*}, ZHANG Jiale¹, XU Jing¹, YAO Hao¹

(1. School of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;

2. Shaanxi College of Communications Technology, Xi'an 710018, China)

Abstract: To address the issues of small apparent size, weak edge features, and strong background interference of some road-surface foreign objects in complex road scenes captured by forward-facing vehicle-mounted cameras, a lightweight YOLO11n-based detection algorithm, YOLO11-MSDF, is proposed. The WFD-Down module decouples high- and low-frequency features to preserve edge and texture details. The DSA-SPPF module employs strip pooling to suppress structured background interference. The MAK-C3 module introduces adaptive sampling to enhance irregular-boundary features. The HADF module fuses shallow-layer details with deep-layer semantic information. Experimental results

收稿日期:2026-07-06;**修订日期:**2026-08-17.

基金项目:国家自然科学基金(No. 51974229);中国博士后科学基金(No. 2025MD784118);陕西省教育厅专项科研计划(No. 25JK0549)

Supported by National Natural Science Foundation of China (No. 51974229); China Postdoctoral Science Foundation (No. 2025MD784118); Special Scientific Research Program of Shaanxi Provincial Department of Education (No. 25JK0549)

*通信联系人, E-mail: 2360601882@qq.com

show that the model contains 2.9 million parameters and requires 7.5 GFLOPs, achieving a precision of 85.8%, a recall of 73.9%, and an mAP@0.5 of 83.7%, which are 7.1, 2.0, and 4.0 percentage points higher than those of YOLO11n, respectively. The detection speed reaches 90.5 FPS on the experimental platform. On an external dataset, the three metrics improve by 5.2, 1.3, and 1.9 percentage points, respectively. Experiments under synthetic fog and motion blur show that the proposed model outperforms the baseline under most degraded conditions, while the analysis of vehicle operating constraints identifies the main factors affecting practical vehicle-mounted applications.

Key words: deep learning; road-surface foreign object detection; lightweight model; feature fusion; object detection

1 引言

在公路交通运行过程中,路面异物易引发车辆紧急制动、转向避让甚至失控碰撞,严重影响道路通行安全与运行效率。《交通强国建设纲要》提出推进公路安全监测智能化,提高交通安全预警能力。因此开展复杂道路场景下路面异物检测具有重要意义。

近年来,深度学习目标检测技术的发展为路面异物检测提供了有效解决方案^[1]。相比R-CNN^[2]、FasterR-CNN^[3]等两阶段检测算法,YOLO^[4]、SSD^[5]等单阶段检测算法具有检测速度快、实时性强和部署方便等优势,更适用于道路异物检测场景。其中,YOLO系列算法因兼顾检测精度与推理速度,已广泛应用于交通目标检测任务。

围绕道路障碍物与散落异物检测,国内外学者开展了大量研究。Zuo等^[6]通过优化YOLOv4先验框与CIoU损失函数,提高了道路障碍物检测精度,但模型计算复杂度较高。冉险生等^[7]在YOLOv5s中引入注意力机制与改进下采样结构,增强了目标特征提取能力,但模型推理速度下降明显。王龙春等^[8]针对自动驾驶场景中的小目标漏检问题,在YOLOv8s中引入RepConv、多尺度注意力和P2检测层,在提高小目标检测性能的同时降低了模型参数量,其主要面向一般道路障碍物检测,对不规则散落异物的适用性研究仍较少。李玉娟等^[9]提出ADMH-YOLOv8算法,通过动态上采样与多尺度融合提升了复杂道路环境下的检测性能,但对形态不规则散落异物的适应能力仍有限。Pang等^[10]提出RGE-YOLO算法,通过轻量化卷积与注意力机制提升了检测速度与精度,但其研究对象主要为单一包装袋目

标,针对多类别散落异物场景中的研究仍较有限。纪文杰等^[11]提出PATD-YOLO算法,通过双主干融合与小目标增强结构提升了复杂背景下的检测效果,但多类别异物检测稳定性仍有提升空间。

综上,复杂道路场景中的路面异物检测仍面临小尺度目标细节易丢失、背景干扰强及模型计算开销较大等问题。本文所述小尺度目标是指在模型输入图像中占据较小像素区域的目标。面向车载前视视觉感知,本文提出基于YOLO11n的轻量化路面异物检测算法,旨在持续处理道路前视图像,为后续风险评估与驾驶辅助预警提供目标类别、图像位置及置信度信息。自动检测的应用价值在于持续、自动地提供道路前方目标信息,作为驾驶员视觉观察的辅助感知信息及后续风险评估的输入,而非直接替代驾驶员判断;实际预警效果仍需结合目标距离、车辆运行状态及完整响应过程综合判断。主要改进如下:

(1)设计WFD-Down下采样模块,通过小波变换实现高低频特征解耦与融合,增强浅层网络对小尺度目标边缘与纹理细节的保留能力。

(2)构建DSA-SPPF空间池化模块,利用条形池化捕获长距离方向性上下文信息,增强模型对道路结构化背景的感知能力,降低标线、裂缝和阴影等背景干扰。

(3)引入MAK-C3特征提取模块,通过自适应采样机制动态调整卷积位置,提升模型对不规则异物形态的特征表达能力。

(4)设计HADP多尺度融合模块,通过跨层特征交互与自适应权重分配,实现浅层细节信息与深层语义信息的高效融合,提升多尺度异物检测性能。

2 相关理论

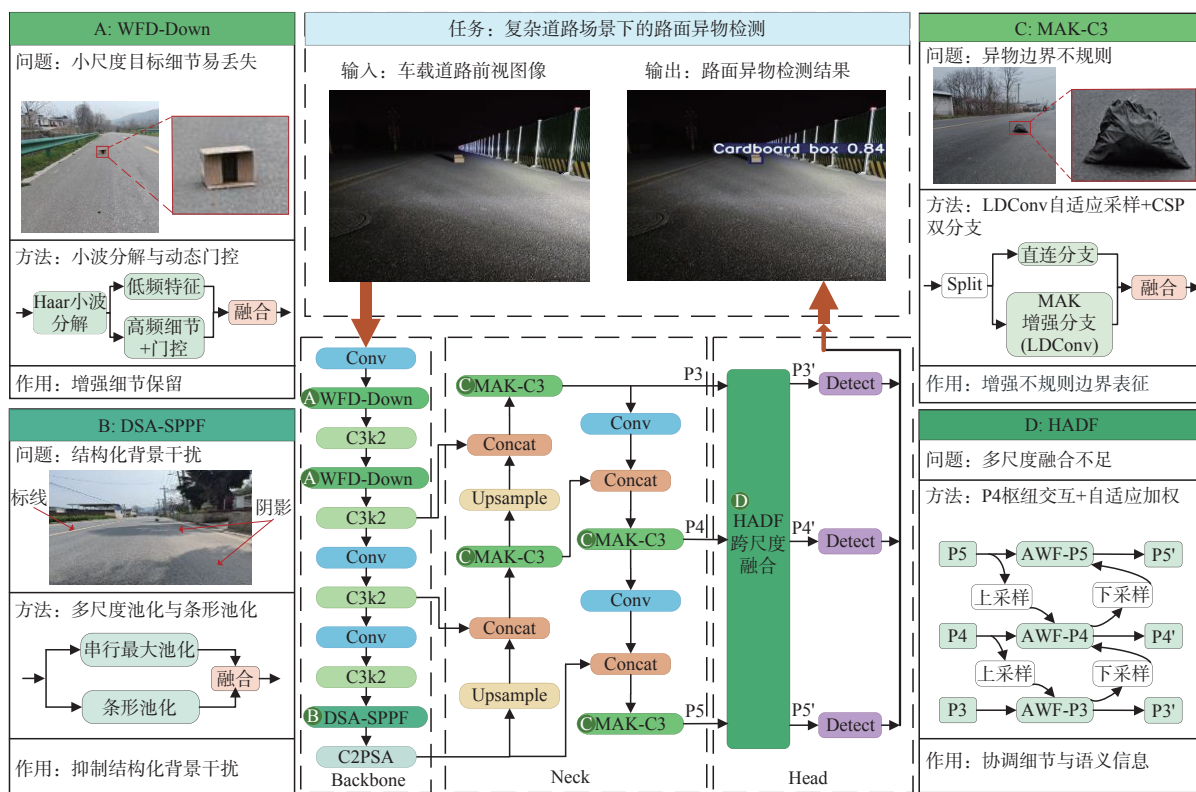
本文选用YOLO11n作为基线模型。该模型具有参数量小、计算开销较低等特点,适合作为轻量化路面异物检测研究的基础。

YOLO11n主要由主干网络(Backbone)、颈部网络(Neck)和头部网络(Head)三部分组成。Backbone采用C3k2模块进行特征提取,并结合C2PSA注意力机制增强对关键目标区域的特征表达能力。Neck通过多尺度特征融合与跨层连接,实现浅层细节信息与深层语义信息的有效聚合,提高模型对不同尺度目标的感知能力。Head采用解耦检测头结构和无锚框(Anchor-Free)策略^[12],分别完成目标分类与

边界框回归任务,从而提升检测性能。在损失函数方面,模型联合采用BCE损失、CIoU^[13]损失和DFL^[14]损失进行优化,以兼顾分类准确性与定位精度。

3 YOLO11-MSDF 模型设计

本文基于YOLO11n构建路面异物检测模型YOLO11-MSDF,以车载前视图图像为输入,输出目标类别、边界框及置信度,整体结构如图1所示。针对小尺度目标细节易丢失、道路背景干扰、异物边界不规则及多尺度特征融合不足等问题,分别引入WFD-Down、DSA-SPPF、MAK-C3和HADF模块。图1中A~D给出了对应检测问题、关键改进机制及模块在网络中的位置。



注: P3、P4、P5为输入HADF的颈部特征, P3'、P4'、P5'为融合后的特征; A~D对应4个改进模块, 框内仅示核心机制, 详细结构见正文。

图1 YOLO11-MSDF 整体网络结构及关键改进机制

Fig. 1 Overall architecture and key improvements of YOLO11-MSDF

3.1 WFD-Down 模块设计

受拍摄距离、环境条件等因素影响,部分路面异物在车载图像中占据的像素区域较小,其边缘与纹理等判别信息较弱。已有研究表明,远距离小目标易出现特征表达不足^[8],而连续下采样

还可能造成空间细节与边缘信息损失^[15]。为缓解这一问题,本文设计了小波频率解耦下采样(Wavelet Frequency Decoupling Downsampling, WFD-Down)模块,其结构如图2所示。

设输入特征为 X , WFD-Down首先引入二维

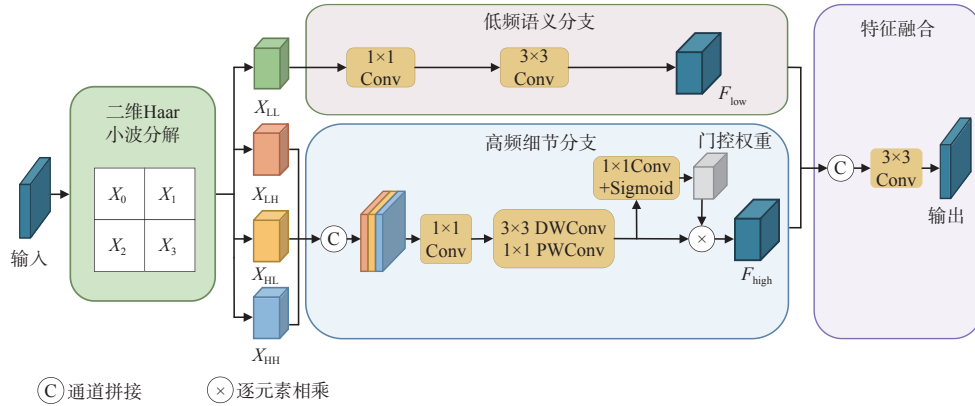


图2 WFD-Down模块结构

Fig. 2 Structure of WFD-Down module

Haar小波分解^[16-17]。通过高低通滤波将输入特征分解为低频与高频分量,如式(1)所示:

$$\{X_{LL}, X_{HL}, X_{LH}, X_{HH}\} = \text{DWT}_{\text{Haar}}(X), \quad (1)$$

式中: X_{LL} 表示低频分量,用于保留目标整体轮廓和场景结构; X_{HL} 、 X_{LH} 、 X_{HH} 表示不同方向的高频分量,用于描述目标边缘与纹理细节。为减少高低频特征之间的相互干扰,本文采用双分支结构对低频语义信息与高频细节信息进行建模。

在低频分支中,采用 1×1 卷积进行通道映射,并结合 3×3 卷积提取局部空间特征,得到低频结构特征 F_{low} ,如式(2)所示:

$$F_{\text{low}} = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(X_{LL})). \quad (2)$$

在高频分支中,首先对3类高频分量进行通道拼接,再利用 1×1 卷积、深度可分离卷积(DWConv)及逐点卷积(PWConv)对高频特征进行细化,以增强边缘与纹理细节表达,如式(3)所示。

$$\begin{cases} F_c = \text{Conv}_{1 \times 1}(\text{Concat}(X_{HL}, X_{LH}, X_{HH})) \\ \hat{F}_{\text{high}} = \text{Conv}_{1 \times 1}(\text{DWConv}_{3 \times 3}(F_c)) \end{cases}. \quad (3)$$

小波高频分量在保留目标边缘与纹理细节的同时,也会引入部分背景高频干扰信息。为此,借鉴门控卷积网络(Gated Convolutional Networks, GCNN)^[18]中的门控思想,利用Sigmoid函数生成自适应权重,并与高频特征逐元素相乘,以抑制杂散噪声响应并突出有效边缘细节。如式(4)所示:

$$F_{\text{high}} = \text{Sigmoid}(\text{Conv}_{1 \times 1}(\hat{F}_{\text{high}})) \otimes \hat{F}_{\text{high}}. \quad (4)$$

最后,将低频特征 F_{low} 与高频特征 F_{high} 进行通道拼接,并通过 3×3 卷积完成多频信息融合,得到输出特征 Y ,如式(5)所示:

$$Y = \text{Conv}_{3 \times 3}(\text{Concat}(F_{\text{low}}, F_{\text{high}})). \quad (5)$$

综上,WFD-Down通过高低频分支建模与门控融合,在压缩特征图尺寸的同时保留边缘与纹理信息,以减轻小尺度目标的细节损失。

3.2 DSA-SPPF模块设计

复杂道路场景中的车道线、路缘等结构化背景具有明显的方向连续特征,容易在高层语义特征中与异物目标产生响应混淆。已有道路视觉检测研究表明,复杂路面背景可能削弱细小目标的有效特征表达,而融合多尺度上下文信息有助于减轻背景干扰^[19]。传统SPPF模块采用各向同性池化进行上下文聚合,难以有效建模道路场景中的方向结构信息。为此,本文提出方向空间感知金字塔快速池化(Directional Spatial-Aware SPPF, DSA-SPPF)模块,其结构如图3所示。

设输入特征为 X ,首先通过 1×1 卷积压缩通道,得到中间特征 X_c :

$$X_c = \text{Conv}_{1 \times 1}(X). \quad (6)$$

DSA-SPPF保留原始SPPF的逐级最大池化分支,以提取多尺度上下文信息。设 $M(\cdot)$ 表示最大池化操作,则多尺度池化特征可表示为:

$$Y_1 = M(X_c), Y_2 = M(Y_1), Y_3 = M(Y_2). \quad (7)$$

为增强网络对方向结构信息的建模能力,模块进一步引入条形池化(Strip Pooling)^[20]分支,分别沿水平与垂直方向进行特征聚合,以提取道路场景中的方向结构特征,如式(8)所示:

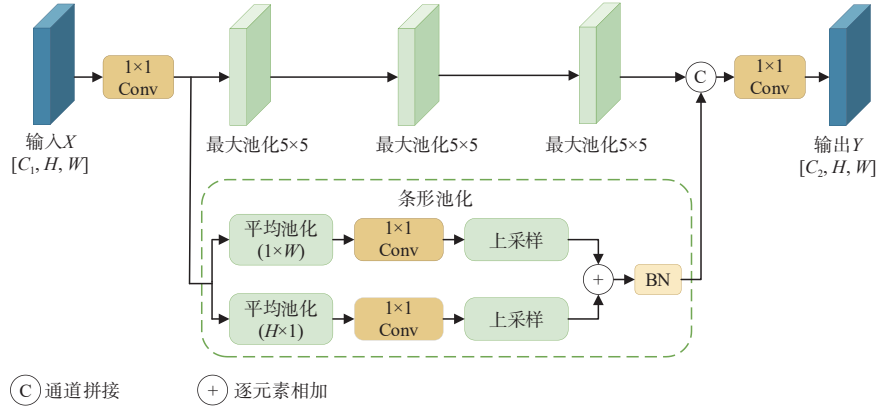


图3 DSA-SPPF 模块结构

Fig. 3 Structure of DSA-SPPF module

$$\begin{cases} S_h(c, i, 1) = \frac{1}{W} \sum_{j=1}^W X_c(c, i, j) \\ S_v(c, 1, j) = \frac{1}{H} \sum_{i=1}^H X_c(c, i, j) \end{cases}, \quad (8)$$

式中, S_h 与 S_v 分别表示水平和垂直方向聚合得到的条形上下文特征。随后,分别通过 1×1 卷积对两个方向的特征进行映射,并通过上采样恢复至中间特征(X_c)的空间尺寸,再将两路特征逐元素相加,并进行批归一化(BN),得到方向感知特征 F_d :

$$F_d = \text{BN}(\epsilon_w(\phi_h(S_h)) + \epsilon_h(\phi_v(S_v))), \quad (9)$$

式中, $\phi_h(\cdot)$ 与 $\phi_v(\cdot)$ 分别表示水平和垂直分支的 1×1 卷积映射, $\epsilon_w(\cdot)$ 与 $\epsilon_h(\cdot)$ 分别表示沿宽度和高度方向的特征扩展操作。最后,将中间特征 X_c 、多尺度空间特征 Y_1 、 Y_2 、 Y_3 以及方向感知特征 F_d 进行拼接,并通过 1×1 卷积完成特征融合,得到最终输出:

$$Y = \text{Conv}_{1 \times 1}(\text{Concat}(X_c, Y_1, Y_2, Y_3, F_d)). \quad (10)$$

综上,DSA-SPPF结合多尺度池化与方向条形池化,在保留上下文信息的同时,增强了网络对道路连续结构特征的感知能力,从而有效缓解复杂背景下的特征混淆问题。

3.3 MAK-C3模块设计

在真实道路场景中,路面异物通常具有边界不规则、局部形态变化明显等特点。传统卷积采用固定采样方式,采样位置受卷积核结构限制,难以根据目标局部形态动态调整,容易造成目标边缘特征提取不足,从而影响模型对复杂形态异物的识别与定位能力。

针对上述问题,本文引入线性可变形卷积(Linear Deformable Convolution, LDConv)^[21]的自适应采样机制,构建形态自适应C3(MAK-C3)模块。如图4所示,LDConv通过学习采样点偏移量,将传统固定采样扩展为可变形动态采样,使卷积核能够根据异物边缘与局部形态变化自适应调整采样位置,从而增强模型对不规则边界与复杂局部结构的建模能力,其计算过程如式(11)所示:

$$y(p) = \sum_{k=1}^K \omega_k \cdot x(p_0 + p_k + \Delta p_k), \quad (11)$$

式中: p_0 为当前位置, p_k 为初始采样位置, Δp_k 为可学习偏移量, ω_k 为卷积权重, K 为采样点数。

在此基础上,本文构建MAK-Bottleneck单元,如图5所示。该单元首先利用 3×3 卷积进行基础局部特征提取,再通过LDConv对异物边缘与局部形态特征进行自适应增强,并结合残差连接(Residual Shortcut)^[22]促进特征复用与梯度传播,从而提升网络训练稳定性。

进一步地,为兼顾复杂形态建模能力与计算效率,MAK-C3模块采用跨阶段局部网络(CSP)架构对输入特征进行双分支处理。如图6所示,直连分支用于保留原始特征信息,以减少网络信息损失;特征变换分支则通过串联堆叠 n 个MAK-Bottleneck单元,对异物边缘与局部形态特征进行进一步建模。最终,两路特征通过 1×1 卷积融合输出。

综上,MAK-C3模块通过结合自适应采样机制与CSP架构,增强了网络对复杂、不规则路面异物边界的特征建模能力,有效缓解了固定规则采样对局部形态表达受限的问题。

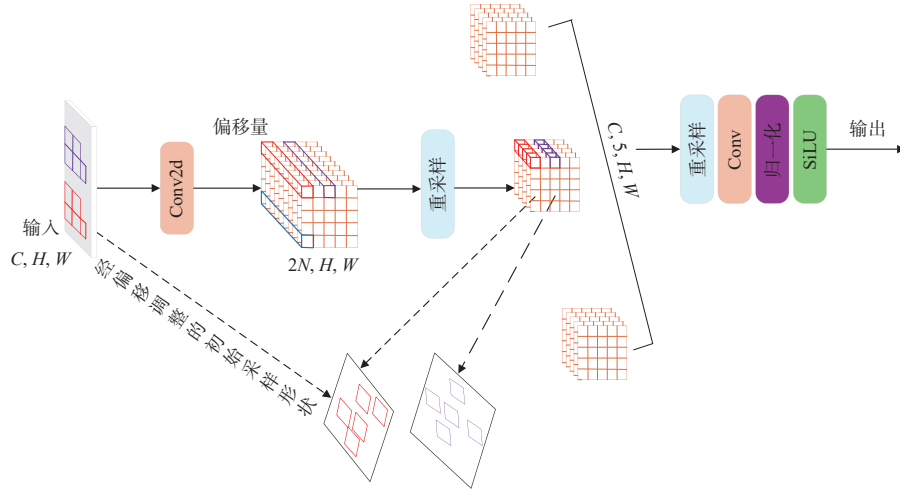


图4 LDConv 自适应采样机制

Fig. 4 Adaptive sampling mechanism of LDConv

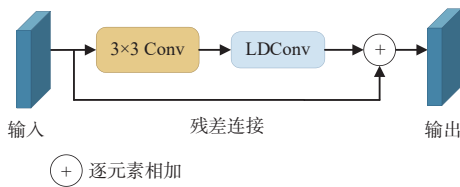


图5 MAK-Bottleneck模块结构

Fig. 5 Structure of MAK-Bottleneck module

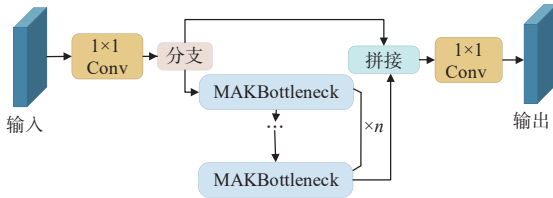


图6 MAK-C3模块结构

Fig. 6 Structure of MAK-C3 module

3.4 HADF 模块设计

在多尺度目标检测网络中,浅层高分辨率特征具有较丰富的位置与细节信息,而深层特征具有更强的语义表达能力,二者的有效融合有助于提升弱特征目标的识别能力^[23]。YOLO11n基线模型采用融合FPN^[24]与PANet^[25]的特征金字塔结构,主要依赖固定的信息传递路径与特征拼接方式。该结构虽然在通用场景下具有较好的适用性,但在复杂路面异物检测任务中,难以根据不同尺度目标与环境干扰动态调整特征融合权重,容易导致浅层细节信息减弱或深层语义表达不足。

为解决上述问题,本文设计层级非对称动态融合(Hierarchical Asymmetric Dynamic Fusion, HADF)模块,其结构如图7所示。该模块以中层特征P4为中心进行跨尺度信息交互:对于浅层检测层P3,引入上采样后的P4上下文信息,以增强语义表达并保留浅层细节;对于深层检测层P5,引入下采样后的P4特征,以增强高层特征的判别能力;而核心枢纽层P4则借鉴加权多尺度特征聚合思想^[26],同时融合来自P3的细节信息与P5的高层语义信息,从而实现不同尺度特征之间的协同融合,其融合路径可表示为:

$$\begin{cases} P'_3 = \text{AWF}(P_3, U(P_4)) \\ P'_4 = \text{AWF}(D(P_3), P_4, U(P_5)) \\ P'_5 = \text{AWF}(P_5, D(P_4)) \end{cases} \quad (12)$$

式中: $U(\cdot)$ 表示上采样操作, $D(\cdot)$ 表示下采样操作, $\text{AWF}(\cdot)$ 表示自适应加权融合单元。

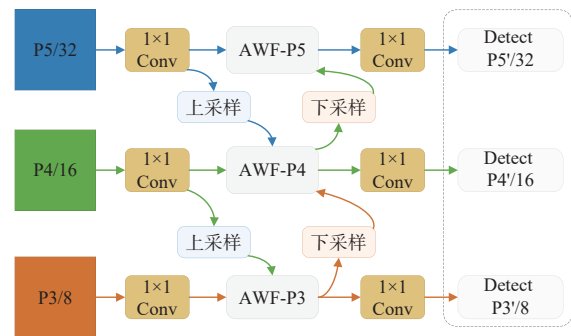


图7 HADF模块结构

Fig. 7 Structure of HADF module

为根据不同空间位置动态调整多源特征的融合权重,本文进一步设计自适应加权融合(Adaptive Weighted Fusion, AWF)单元,其结构如图8所示。AWF首先为各输入特征生成对应的空间权重,并利用Softmax函数进行归一化处理,随后对不同来源特征进行加权融合,如式(13)所示:

$$F_{out} = \text{Conv}_{3 \times 3} \left(\sum_{i=1}^n \alpha_i \odot F_i \right) + F_r, \quad (13)$$

式中: F_i 表示第*i*个输入特征, α_i 表示归一化后的空间权重, n 为输入特征数量, $\odot$ 表示逐元素相乘, F_r 表示残差特征。

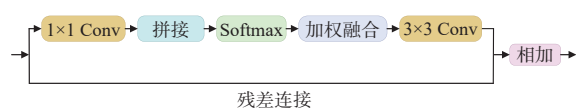


图8 AWF单元结构

Fig. 8 Structure of AWF unit

综上,HADF模块通过结合跨尺度信息交互与自适应加权融合机制,实现了浅层细节信息与深层语义信息的动态融合,从而缓解了传统固定融合方式下特征利用不充分的问题。

4 实验设计与结果分析

4.1 实验环境及训练参数设置

本文实验均在相同软硬件环境下完成。软件环境为Python3.12.12、PyTorch 2.7.1、CUDA12.8

和Ultralytics 8.4.6,硬件平台为NVIDIA GeForce RTX 5060 GPU。模型输入尺寸统一设置为640×640。测速时批量尺寸(batch size)设为1,经预热后统计多次单帧处理的平均耗时,计时范围包括前处理、模型推理和后处理,并据此计算FPS。主要训练参数见表1。

表1 模型训练参数设置

Tab. 1 Training parameter settings of the model

参数	数值
输入尺寸(Img_size)	640×640
学习率(Learning_rate)	0.01
迭代次数(Epoch)	200
批量尺寸(Batch_size)	16
线程数(Workers)	2
动量值(Momentum)	0.937
优化器(Optimizer)	SGD

4.2 数据集的构建

本文构建的RD(Road Debris)路面异物检测数据集来源于真实道路场景下的车载前视图图像。通过等间隔截取车载视频帧,减少相邻样本之间的冗余信息。数据集划分以视频序列为单位,以避免相邻或相似帧分布于不同子集。数据集覆盖晴天、雨天、低能见度及轻度沙尘等多种环境条件,可反映车载前视场景中的典型环境干扰。部分数据集样例见图9。

RD数据集包含砖块、交通锥、塑料袋、桶和



图9 RD数据集样例

Fig. 9 Samples from the RD dataset

纸箱等 5 类常见路面异物,共包含 2 170 张图像,其中训练集、验证集和测试集分别包含 1 446 张、362 张和 362 张图像。所有图像均采用 LabelImg 工具进行矩形框标注,并保存为 YOLO 格式。

本文以目标在模型输入图像中的标注框像素面积衡量其尺度,不以实际物理尺寸划分。将标注框按实际图像预处理方式映射至 640×640 输入坐标系,其面积计算如式(14)所示:

$$A=w\cdot h, \tag{14}$$

式中: w 、 h 分别为标注框的像素宽度和高度; A 为标注框面积,单位为 px^2 。本文依据目标在输入图像中的标注框面积进行尺度划分,将 $A<32^2$ 、 $32^2\leq A<96^2$ 和 $A\geq 96^2$ 的目标分别划分为小、中、大尺度目标^[27]。因此,交通锥等物理尺寸较大的物体,在远距离成像时也可能属于小尺度目标。

按照上述规则,RD 数据集 2 170 张原始图像共包含 2 368 个目标实例,其中小、中、大尺度目标分别为 1 454、764 和 150 个,占 61.4%、32.3% 和 6.3%,各类别分布见表 2。

表 2 RD 数据集各类别目标尺度分布
Tab. 2 Distribution of object scales by category in the RD dataset

类别	小尺度 目标数	中尺度 目标数	大尺度 目标数	目标 总数
砖块	337	146	22	505
交通锥	215	139	36	390
塑料袋	351	122	18	491
桶	308	147	33	488
纸箱	243	210	41	494
合计	1 454	764	150	2 368

考虑到实际道路场景中路面异物样本数量有限,为增强训练样本多样性,本文采用水平翻转与亮度增强两种方式对训练集进行数据扩充,以提高模型对目标位置变化及光照变化的鲁棒性。

增强后训练集扩充至 4 338 张,而验证集与测试集不参与数据增强,以保证实验评估的一致性。

4.3 评估指标

为评估模型性能,试验采用精确率(Precision, P)、召回率(Recall, R)、平均精度均值(Mean Average Precision, mAP)、参数量(Parameters)、浮点运算量(Floating Point Operations, FLOPs)和检测速度(Frames Per Second, FPS)6 个指标评估模型的综合性能。各指标计算公式如式(15)~(19)所示:

$$P=\frac{TP}{TP+FP}, \tag{15}$$

$$R=\frac{TP}{TP+FN}, \tag{16}$$

$$AP=\int_0^1 P dR, \tag{17}$$

$$mAP=\frac{1}{n}\sum_{i=1}^n AP_i, \tag{18}$$

$$FPS=\frac{1\,000}{T_{pre}+T_{inf}+T_{post}}, \tag{19}$$

式中: P 为精确率; R 为召回率; TP 为检测正确的目标数量; FP 为检测错误的目标数量; FN 为漏检数量; AP 为 P - R 曲线的线下面积; AP_i 为第 i 个类别的 AP 值; mAP 为所有类别的 AP 的平均值; n 为检测类别总数; FPS 为每秒帧数,时间为 1 000 ms; T_{pre} 、 T_{inf} 和 T_{post} 分别表示模型前处理、推理和后处理时间。

4.4 WFD-Down 模块性能验证

为了验证 WFD-Down 模块的性能和合理性,本文选取 SCDwn 和 ADown 两种代表性下采样结构进行对比实验,实验结果见表 3。

由表 3 可知,引入 WFD-Down 后,模型 mAP@0.5 和 mAP@0.5~0.95 分别达到 80.3% 和 57.2%,较基线提高 0.6 和 1.8 个百分点,且均高于 SCDwn 与 ADown 的对应结果。模型参数量基

表 3 不同下采样结构性能对比
Tab. 3 Performance comparison of different downsampling structures

模型	mAP@0.5/%	mAP@0.5~0.95/%	Params/M	FLOPs/G
YOLO11n	79.7	55.4	2.6	6.4
YOLO11n+SCDwn	78.9	54.8	2.5	6.1
YOLO11n+ADwn	78.8	54.4	2.5	6.2
YOLO11n+WFD-Down	80.3	57.2	2.6	7.2

注:粗体数字表示最优值;SCDwn 和 ADown 分别表示轻量化解耦下采样、空间域改进下采样。

本不变,计算量(浮点运算量)增加0.8 GFLOPs,表明该模块以较小的计算增量改善了检测性能。

4.5 DSA-SPPF 模块性能验证

为了验证 DSA-SPPF 模块的性能和合理性,本文将其与 SimSPPF、ASPP 进行对比,实验结果见表 4。

由表 4 可知,SimSPPF 虽保持了较低的计算复杂度,但其 mAP@0.5 却降至 79.3%,表明其多

尺度特征的表达能力仍存在一定不足;ASPP 虽扩大了感受野,但参数量和计算量分别增加至 2.9M 和 6.8 GFLOPs,其 mAP@0.5 降至 78.7%,表明其在该任务上的特征聚合效果有限。相比之下,引入 DSA-SPPF 后,在仅增加 0.2M 参数数量的条件下,mAP@0.5 及 mAP@0.5~0.95 较原始 SPPF 均提升约 1.0 个百分点,表明该结构在抑制背景干扰与空间上下文特征聚合方面具有一定优势。

表 4 不同空间上下文聚合结构性能对比

Tab. 4 Performance comparison of different spatial context aggregation structures

模型	mAP@0.5/%	mAP@0.5~0.95/%	Params/M	FLOPs/G
YOLO11n	79.7	55.4	2.6	6.4
YOLO11n+SimSPPF	79.3	55.7	2.6	6.4
YOLO11n+ASPP	78.7	55.1	2.9	6.8
YOLO11n+DSA-SPPF	80.7	56.4	2.8	6.5

注:粗体数字表示最优值;SimSPPF 和 ASPP 为常用空间上下文聚合结构。

4.6 MAK-C3 模块性能验证

为了验证 MAK-C3 模块的性能和合理性,本文将其与不同 C3k2 改进结构进行对比,实验结果见表 5。

由表 5 可知,引入 MAK-C3 后,模型参数量基

本不变,计算量增加 0.4 GFLOPs,mAP@0.5 和 mAP@0.5~0.95 分别达到 81.3% 和 57.6%,较基线提高 1.6 和 2.2 个百分点,且均高于其他对比结构。结果表明,MAK-C3 能够在较小计算开销增加的条件下改善路面异物检测性能。

表 5 不同 C3k2 改进结构性能对比

Tab. 5 Performance comparison of different improved C3k2 structures

模型	mAP@0.5/%	mAP@0.5~0.95/%	Params/M	FLOPs/G
YOLO11n	79.7	55.4	2.6	6.4
YOLO11n+C3k2-DWConv	79.4	55.9	2.5	6.5
YOLO11n+C3k2-STDConv	80.1	56.0	2.6	6.7
YOLO11n+C3k2-DCN	79.9	55.7	2.6	6.4
YOLO11n+MAK-C3	81.3	57.6	2.6	6.8

注:粗体数字表示最优值;DWConv、STDConv 和 DCN 分别表示深度可分离卷积、标准卷积和可变形卷积。

4.7 HADF 模块性能验证

为了验证 HADF 模块的性能和合理性,本文

将 HADF 与不含 AWF 的非对称融合策略 AF 进行对比,实验结果见表 6。

表 6 不同多尺度融合策略性能对比

Tab. 6 Performance comparison of different multi-scale fusion strategies

模型	mAP@0.5/%	mAP@0.5~0.95/%	Params/M	FLOPs/G
YOLO11n	79.7	55.4	2.6	6.4
YOLO11n+AF	79.4	55.3	2.8	7.3
YOLO11n+HADF	80.9	56.1	2.7	7.0

注:粗体数字表示最优值;AF 表示不含自适应加权融合单元 AWF 的非对称融合策略。

由表 6 可知,仅采用 AF 策略时,模型 mAP@0.5 降至 79.4%,说明非对称融合结构本身未带来性能改善。采用结合 AWF 的完整 HADF 结构后,mAP@0.5 和 mAP@0.5~0.95 分别达到 80.9% 和 56.1%,较基线提高 1.2 和 0.7 个百分点,参数量仅增加 0.1M。结果表明,引入自适应

权重分配有助于协调不同层级特征,在保持较小模型规模的同时改善检测性能。

4.8 消融实验

为验证各改进模块的有效性,本文以 YOLO11n 为基线模型,在相同数据集划分与训练配置下开展消融实验,实验结果见表 7。

表 7 YOLO11-MSDF 模型消融实验

Tab. 7 YOLO11-MSDF model ablation experiments

YOLO11n	WFD-Down	DSA-SPPF	MAK-C3	HADF	P/%	R/%	mAP@0.5/ %	mAP@0.5~ 0.95/%	Params/ M	FLOPs/ G	FPS
✓	—	—	—	—	78.7	71.9	79.7	55.4	2.6	6.4	120.5
✓	✓	—	—	—	73.9	77.2	80.3	57.2	2.6	7.2	106.3
✓	✓	✓	—	—	79.4	72.2	80.8	57.3	2.8	7.3	103.1
✓	✓	✓	✓	—	80.0	75.7	81.6	57.7	2.8	7.0	100.0
✓	✓	✓	✓	✓	85.8	73.9	83.7	58.5	2.9	7.5	90.5

注:粗体数字表示最优值;“✓”指使用了该模块,“—”指未使用该模块。

由表 7 可知,引入 WFD-Down 后,模型参数量基本不变,计算量增加 0.8 GFLOPs,mAP@0.5 提高 0.6 个百分点;进一步加入 DSA-SPPF 后,mAP@0.5 提高 0.5 个百分点,参数量与计算量分别增加 0.2M 和 0.1 GFLOPs。继续引入 MAK-C3 后,mAP@0.5 提高 0.8 个百分点,计算量降低 0.3 GFLOPs,表明该阶段在改善检测精度的同时减少了理论计算量。最后,加入 HADF 后,mAP@0.5 进一步提高 2.1 个百分点,达到 83.7%,表明其在已有改进基础上仍能带来性能增益。完整模型较基线模型的 mAP@0.5 提高 4.0 个百分点,参数

量和计算量分别增加 0.3M 和 1.1 GFLOPs,检测速度由 120.5 FPS 降至 90.5 FPS。总体而言,各模块的逐步引入改善了检测精度,完整模型在增加一定计算开销的同时保持了较小的模型规模,并在本文测试平台上表现出较好的实时处理能力。

为进一步分析性能提升的原因,图 10 展示了各模块逐步引入后的特征响应,其中 A、B、C 和 D 分别对应 WFD-Down、DSA-SPPF、MAK-C3 和 HADF。在所展示样本中,完整模型的高响应区域更加集中于目标附近,为表 7 中的检测性能改善提供了定性参考。

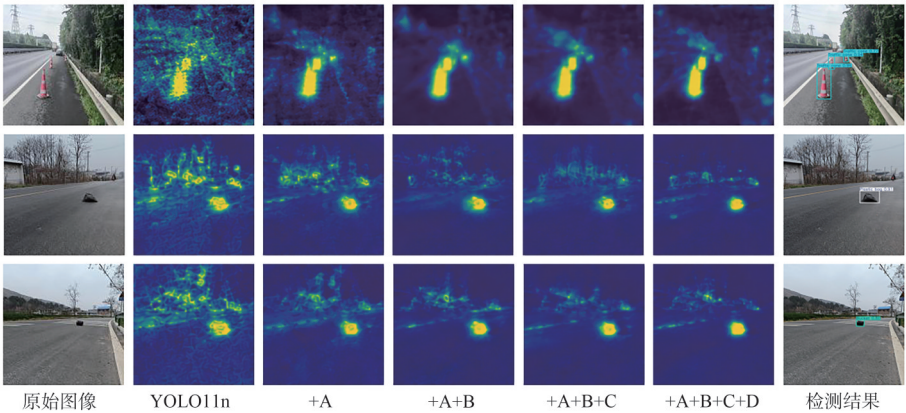


图 10 模块逐步引入后的特征响应图

Fig. 10 Feature response maps after progressive module introduction

4.9 与其他主流算法的对比实验

为评估 YOLO11-MSDF 的检测性能,本文将 其 与 RT-DETR、YOLOv6n、YOLOv8n、YOLOv8s、YOLOv10n、YOLOv10s、YOLO11n 和 YOLO11s 进行对比,结果见表 8。

由表 8 可知,YOLO11-MSDF 的 mAP@0.5 达到 83.7%,较 RT-DETR、YOLOv6n 和 YOLOv8n 分别提高 4.6、2.8 和 3.1 个百分点,参数量与计算

量均低于上述模型。与基线 YOLO11n 相比,其 mAP@0.5 提高 4.0 个百分点,参数量和计算量分别增加 0.3M 和 1.1 GFLOPs。与 YOLO11s 相比,其 mAP@0.5 提高 1.4 个百分点,参数量和计算量分别减少 6.5M 和 14.1 GFLOPs。此外,虽然其 mAP@0.5 低于 YOLOv8s,但参数量和计算量分别减少 8.2M 和 21.2 GFLOPs,体现了检测精度与模型规模之间取得了较好的平衡。

表 8 与主流目标检测算法的对比

Tab. 8 Comparison with mainstream detection algorithms

模型	$P/\%$	$R/\%$	mAP@0.5/%	mAP@0.5~0.95/%	Params/M	FLOPs/G	FPS
RT-DETR	83.9	76.6	79.1	57.4	92.3	103.4	52.9
YOLOv6n	79.8	71.8	80.9	55.7	4.2	11.9	151.5
YOLOv8n	79.6	71.7	80.6	56.5	3.0	8.2	149.2
YOLOv8s	84.8	77.7	86.3	62.7	11.1	28.7	142.8
YOLOv10n	70.7	67.9	73.3	52.3	2.7	8.4	138.8
YOLOv10s	75.2	72.3	79.5	57.5	8.0	24.8	116.2
YOLO11n	78.7	71.9	79.7	55.4	2.6	6.4	120.5
YOLO11s	82.7	73.8	82.3	58.6	9.4	21.6	105.2
YOLO11-MSDF	85.8	73.9	83.7	58.5	2.9	7.5	90.5

注:粗体数字表示最优值

在推理效率方面,YOLO11-MSDF 的检测速度由基线模型 YOLO11n 的 120.5 FPS 降至 90.5 FPS,表明检测精度的提升伴随一定的推理速度损失。总体而言,所提模型以 2.9M 参数量和 7.5 GFLOPs 计算量取得了 83.7% 的 mAP@0.5,并在本文测试平台上保持 90.5 FPS 的处理速度,在检测精度、模型规模与推理效率之间取得了较好的平衡,体现出较好的轻量化与实时检测能力。

4.10 可视化分析

为比较两种模型在复杂道路场景下的检测表现,本文选取 RD 测试集中的典型样本进行可视化分析,结果见图 11。在所展示的小尺度目标、夜间眩光及低照度等场景中,基线模型存在误检、漏检及部分目标检测置信度较低的情况。相比之下,YOLO11-MSDF 能够更有效地实现对目标的定位与识别,在复杂场景中的检测置信度相对更高。上述可视化结果展示了所提模型在典型复杂场景中的检测改善,为定量评测结果提供了直观补充。

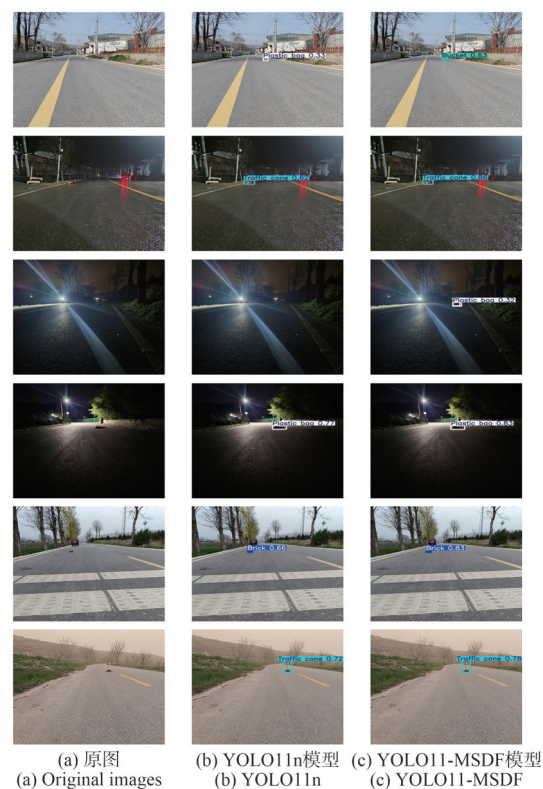


图 11 检测效果对比

Fig. 11 Comparison of detection results

4.11 不同合成雾强度下的检测性能分析

由于 RD 数据集中真实雾天样本较少,本文采用合成雾方法构建不同强度的雾化图像,以评估模型在雾化退化条件下的检测性能。首先将原始 RGB 图像归一化至 $[0,1]$,并根据图像纵向位置构造近似深度权重。将图像高度 40% 处设为近似地平线,其上方区域深度权重设为 1,下方区域经纵向归一化后采用指数 1.5 的幂函数由 1 逐渐衰减至 0。合成雾图像表示为:

$$I_t=I_0(1-\beta)+A\beta,$$
$$\beta=\text{clip}\big[SD_n(0.90+0.15N),0,0.75\big],$$

(20)

(21)

式中: I_0 和 I_t 分别为原始图像和合成雾图像; $A=230/255$ 为大气光强度; D_n 为近似深度权重; N 为经高斯平滑和归一化得到的随机扰动,高斯标准差为 25 px; S 为雾化强度控制系数;0.90 用于控制基础雾化程度,0.15 用于调节局部随机扰动幅度; β 为雾化权重,其上限设为 0.75,以限制图像过度泛白。

为保证实验可重复性,图像按文件名排序,第 i 幅图像采用随机种子 $(42+i)$ 生成扰动,同一图像在不同雾化强度下复用相同的深度权重和随机扰动,仅改变 S 。本文设置 $S=0.4,0.6,0.8$,分别生成低、中、高强度合成雾图像,如图 12 所示。

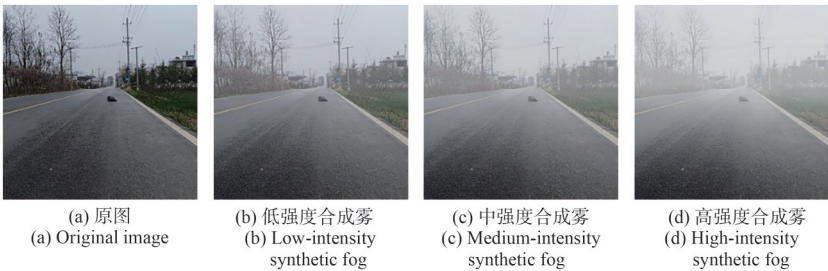


图 12 不同合成雾强度下的图像对比

Fig. 12 Comparison of images under different synthetic fog intensities

采用上述方法对 RD 测试集中的全部 362 张图像进行雾化处理,生成不同强度的合成雾图像,并保持图像尺寸、目标位置及原始标注不变。

生成图像仅用于测试,在相同评测配置下分别采用 YOLO11n 和 YOLO11-MSDF 进行检测,结果见表 9。

表 9 不同合成雾强度下模型检测性能对比

Tab. 9 Detection performance comparison under different synthetic fog intensities

场景	模型	$P/\%$	$R/\%$	mAP@0.5/%	mAP@0.5~0.95/%
原始图像	YOLO11n	78.7	71.9	79.7	55.4
	YOLO11-MSDF	85.8	73.9	83.7	58.5
低强度合成雾	YOLO11n	73.8	64.1	71.9	51.5
	YOLO11-MSDF	77.4	65.6	73.4	51.1
中强度合成雾	YOLO11n	61.1	53.7	59.0	41.0
	YOLO11-MSDF	68.8	57.7	61.5	42.2
高强度合成雾	YOLO11n	47.7	43.7	41.8	25.9
	YOLO11-MSDF	49.2	42.8	45.2	30.0

注:粗体数字表示最优值

由表 9 可知,随着合成雾强度增加,两种模型的检测性能均明显下降。低、中、高强度合成雾下,YOLO11-MSDF 的 mAP@0.5 分别较 YOLO11n 提高 1.5、2.5 和 3.4 个百分点,但低强度合成雾

下的 mAP@0.5~0.95 降低 0.4 个百分点,高强度合成雾下的召回率降低 0.9 个百分点,说明其优势并非体现在所有指标上。高强度合成雾下,YOLO11-MSDF 的召回率和 mAP@0.5 分别降

至 42.8% 和 45.2%,表明较强雾化退化仍会显著削弱模型的目标检出能力。上述实验反映了模型在可控雾化退化条件下的检测性能变化,但合成雾强度不直接对应实际气象能见度。因此,后续将结合真实雾天道路数据,进一步验证模型在大雾场景下的检测稳定性与适用性。

4.12 不同运动模糊强度下的检测性能分析

为评估模型对运动模糊的鲁棒性,本文在测试图像原有分辨率下,采用线性运动模糊核进行卷积处理,生成不同程度的运动模糊图像,其计算过程如式(22)所示:

$$I_b = I_0 * K(L, \theta), \tag{22}$$

式中, I_0 和 I_b 分别表示原始图像和运动模糊图像; $K(L, \theta)$ 为归一化的线性运动模糊核; L 为模糊长度,单位为像素(px); θ 为模糊方向。实验设置 $L=5$ 、10 和 15,分别生成轻、中、重 3 种程度的运动模糊图像,如图 13 所示。对于同一幅图像,在 $0^\circ\sim180^\circ$ 范围内随机生成一个模糊方向,并在 3 种模糊等级下保持方向一致,仅改变模糊长度。随机种子固定为 42,以保证实验可重复性。因此,各模糊等级的性能差异主要反映模糊长度变化对模型检测性能的影响。

实验选取 RD 测试集中的全部 362 张图像,按上述方法处理后保持图像尺寸不变,并沿用原始标注。在相同评测配置下,采用训练完成的 YOLO11n 和 YOLO11-MSDF 模型对各模糊等级的同一组图像进行测试,结果见表 10。

由表 10 可知,随着运动模糊程度增强,两种模型的检测性能均明显下降。轻、中度运动模糊下, YOLO11-MSDF 的各项检测指标均优于 YOLO11n;重度运动模糊下,其召回率、mAP@0.5 和 mAP@0.5~0.95 较 YOLO11n 分别提高 9.0、4.1 和 2.2 个百分点,但精确率降低 5.0 个百分点,表明目标检出能力的改善伴随着误检风险增加。为进一步分析模型在极端退化条件下的检测表现,本文选取高强度合成雾与重度运动模糊场景进行可视化分析,结果如图 14 所示。图 14(a)~(c)中,两种模型均将塑料袋误识别为纸箱,说明目标判别信息受损后仍可能出现类别混淆。图 14(d)~(f)中, YOLO11n 将塑料袋误识别为桶,置信度为 0.28,而 YOLO11-MSDF 能够正确识别为塑料袋,置信度为 0.56,说明改进模型在该样本上的类别判别有所改善。然而,重度运动模糊下, YOLO11-MSDF 的 mAP@0.5 仍降

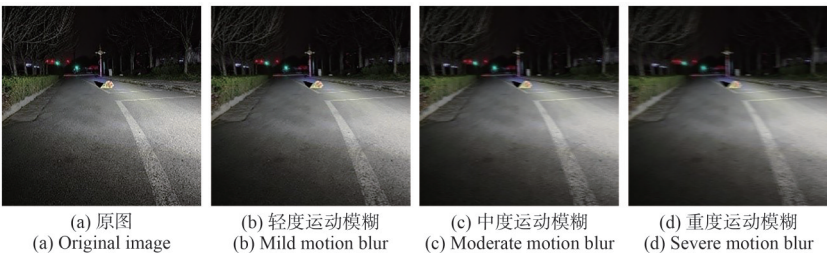


图 13 不同运动模糊强度下的图像对比

Fig. 13 Comparison of images under different motion blur intensities

表 10 不同运动模糊强度下模型检测性能对比

Tab. 10 Detection performance comparison under different motion blur intensities

场景	模型	<i>P</i> /%	<i>R</i> /%	mAP@0.5/%	mAP@0.5~0.95/%
原始图像	YOLO11n	78.7	71.9	79.7	55.4
	YOLO11-MSDF	85.8	73.9	83.7	58.5
轻度模糊 <i>L</i> =5	YOLO11n	78.7	68.1	76.9	52.4
	YOLO11-MSDF	83.6	69.6	79.2	53.8
中度模糊 <i>L</i> =10	YOLO11n	63.4	56.0	62.2	35.0
	YOLO11-MSDF	70.1	63.1	66.6	36.5
重度模糊 <i>L</i> =15	YOLO11n	55.9	40.6	44.0	21.6
	YOLO11-MSDF	50.9	49.6	48.1	23.8

注:粗体数字表示最优值



图 14 高强度合成雾与重度运动模糊下的检测结果对比

Fig. 14 Comparison of detection results under high-intensity synthetic fog and severe motion blur

至 48.1%, 说明模型对严重模糊的适应能力仍有限, 其性能下降可能与拖影导致的边缘与纹理信息损失有关。本文采用线性运动模糊构造可控退化条件, 模糊长度不直接对应实际车速。因此, 后续将利用真实车载动态视频验证模型的检测稳定性, 并探索融合多帧时序信息与运动补偿方法, 以改善严重运动模糊下的检测性能。

4.13 车辆运行参数与异物检测的应用约束

车载异物检测的实际应用效果受到目标距离、目标物理尺寸、车辆速度、制动能力及环境条件等因素的共同影响。为分析各因素与异物检测及车辆响应之间的制约关系, 本文分别从目标成像尺度和车辆停车距离两个方面进行理论分析。

在目标姿态近似不变、目标各部分深度差异相对较小的条件下, 设目标实际宽度和高度为 W_0 和 H_0 , 目标沿摄像头光轴方向的距离为 Z , 摄像头水平和垂直方向的等效焦距分别为 f_x 和 f_y , 则目标在图像中的成像尺寸可根据针孔模型近似表示为:

$$w_p = \frac{f_x W_0}{Z}, \quad h_p = \frac{f_y H_0}{Z}. \quad (23)$$

其成像面积为:

$$A_p = w_p h_p = \frac{f_x f_y W_0 H_0}{Z^2}, \quad (24)$$

式中: w_p 和 h_p 分别为目标的成像宽度和高度, A_p 为目标成像面积。由式(23)和式(24)可知, 在摄像头参数及目标尺寸保持不变的条件下, 当目标距

离增大至原来的 2 倍时, 其成像宽度和高度分别减小至原来的 1/2, 成像面积减小至原来的 1/4; 在目标距离不变时, 目标物理尺寸减小同样会导致成像尺度下降。由此可见, 目标距离与物理尺寸共同决定其在图像中的像素尺度, 远距离目标更易因成像尺度减小而增加检测难度。

为进一步说明上述关系, 本文以 RD 数据集中的砖块目标为例进行分析。该砖块经实测尺寸约为 $240 \text{ mm} \times 115 \text{ mm} \times 53 \text{ mm}$, 图像采集设备为 Intel RealSense D435i RGB 相机, 原始 RGB 图像分辨率为 1280×720 。模型输入前, 图像经等比例缩放并填充至 640×640 。通过 RealSense SDK 读取 RGB 相机内参, 并按图像缩放比例换算至模型输入尺度, 得到水平和垂直方向的等效像素焦距分别为 $f_x = 466 \text{ px}$ 和 $f_y = 469 \text{ px}$ 。考虑砖块主要可见平面, 在简化投影条件下取其宽度和高度分别为 0.240 m 和 0.115 m , 并选取 10 m 和 20 m 作为代表性距离。根据针孔成像模型计算, 两种距离下的成像尺寸分别约为 $11.2 \text{ px} \times 5.4 \text{ px}$ 和 $5.6 \text{ px} \times 2.7 \text{ px}$ 。其中, 10 m 和 20 m 仅作为代表性距离用于说明目标距离与成像尺度之间的关系, 并不代表模型实际的可靠检测距离。可见, 当砖块由较近位置向远处移动时, 其在图像中占据的像素区域明显减小, 边缘与纹理等有效判别信息随之减少, 从而增加远距离异物的检测难度。

在车辆运动方面, 按照第 4.1 节的测速口径,

YOLO11-MSDF以90.5 FPS运行时,平均单帧图像处理时间约为11.0 ms。为分析车速和制动能力的影响,在平直道路、静止异物和车辆匀减速制动的简化条件下,仅考虑单帧图像处理时间与制动过程,则理想化最低停车距离可表示为:

$$D_{\min} = \frac{v}{F} + \frac{v^2}{2a_b}, \quad (25)$$

式中: v 为车速; F 为检测速度; a_b 为有效制动减速度;式中第一项为单帧图像处理期间车辆行驶的距离,第二项为匀减速制动距离。参考AASHTO停车视距设计采用的 3.4 m/s^2 减速度^[28]。另一方面,相关研究表明,干燥铺装路面条件下,车速不低于48 km/h时轮胎路面摩擦系数约为0.45~0.70,低于48 km/h时约为0.55~0.80^[29]。为构造不同有效制动能力下的分析工况,本文取摩擦系数 $\mu=0.60$ 作为代表性分析值,对应理想减速度约为 5.9 m/s^2 。基于上述取值,本文选取 3.4 m/s^2 、 4.6 m/s^2 和 5.9 m/s^2 作为不同有效制动能力下的分析工况,其中 4.6 m/s^2 作为两者之间的中间工况。根据式(25)计算,当 $a_b=4.6 \text{ m/s}^2$ 时,车速为30 km/h、70 km/h和90 km/h对应的理想化最低停车距离分别约为7.64 m、41.31 m和68.21 m;当车速为70 km/h时,上述3种制动减速度对应的距离分别约为55.82 m、41.31 m和32.26 m。可见,车速升高或有效制动减速度降低均会增大理想化最低停车距离。

结合前述目标成像尺度分析可以看出,停车距离与目标成像尺度之间存在进一步的制约关系。当车辆所需停车距离增大时,为保证车辆有足够的制动余量,系统需要在更远位置可靠检出前方异物;而在目标物理尺寸和摄像头参数不变的条件下,检测距离增加又会使目标的成像宽度和高度近似按 $1/Z$ 减小,成像面积近似按 $1/Z^2$ 减小。因此,较高车速或较低制动能力不仅增加车辆停车距离,同时对远距离小尺度目标的可靠检测提出了更高要求。

进一步考虑感知、决策与执行过程中的总响应时间及安全裕量,在上述简化条件下,所需停车距离可近似估计为:

$$D_{\text{rep}} = vt_{\text{sys}} + \frac{v^2}{2a_b} + d_m, \quad (26)$$

式中: t_{sys} 为从感知开始至车辆产生有效制动的总响应时间,包括图像处理、决策和执行等环节的

延迟,驾驶辅助场景下还包括警报传递及驾驶员反应时间; d_m 为考虑道路状态、车辆运行状态及系统误差设置的安全裕量。以70 km/h为例,总响应时间每增加1 s,车辆将额外行驶约19.44 m。这里的可用距离指沿车辆行驶路径至异物的距离,与前述光轴方向距离并不完全相同;在摄像头朝向近似与行驶方向一致、目标位于车辆前方的简化条件下,可近似关联二者。只有在可靠检出异物时的可用距离大于所需停车距离时,车辆才具有一定的制动距离余量。因此,模型能够检出目标并不必然意味着能够及时完成制动。

环境条件还会同时影响视觉感知与车辆制动。雾、低照度及运动模糊可能削弱目标判别信息,雨雪、积水等路面状态则可能降低有效制动减速度,使感知距离与停车距离之间的矛盾加剧。本文的成像计算、退化实验与停车距离分析分别说明了上述约束,但尚未通过实车试验建立不同环境、车速与可靠检测距离之间的对应关系。因此,现有结果主要支持检测算法的性能评价与应用约束分析,实际预警效果仍需结合目标测距和动态车载试验进一步验证。

4.14 外部数据集对比实验

为进一步评估YOLO11-MSDF模型在外部公开数据集上的检测性能,本文在Roboflow平台公开发布的YOLOv7-Trash-03-22-2023数据集上,将YOLO11-MSDF与基线模型YOLO11n进行对比评测,实验结果见表11。

表11 外部数据集上的检测性能对比

Tab.11 Comparison of detection performance on the external dataset

模型	$P/\%$	$R/\%$	$mAP@0.5/\%$
YOLO11n	71.4	73.7	74.8
YOLO11-MSDF	76.6	75.0	76.7

注:粗体数字表示最优值

由表11可知,YOLO11-MSDF在外部数据集上的精确率、召回率和 $mAP@0.5$ 分别为76.6%、75.0%和76.7%,较YOLO11n提高5.2、1.3和1.9个百分点,表明所提模型在该数据集上同样取得了优于基线的检测结果。图15给出了典型样本的可视化对比。



图 15 外部数据集上的检测结果对比
Fig. 15 Comparison of detection results on the external dataset

5 结 论

(1) 针对复杂道路场景中小尺度目标细节易丢失、结构化背景干扰及异物边界不规则等问题,本文提出轻量化路面异物检测模型YOLO11-MSDF。模型在RD数据集上的mAP@0.5达到83.7%,较YOLO11n提高4.0个百分点,参数量、计算量和检测速度分别为2.9M、7.5 GFLOPs和90.5 FPS,并在外部公开数据集上取得优于基线模型的检测结果,表明所提模型在检测精度与计算开销之间具有较好的平衡。

(2) 合成雾与运动模糊实验表明,YOLO11-MSDF在多数实验场景下的整体检测表现优于

基线模型,但在高强度合成雾和重度运动模糊条件下性能仍明显下降。高强度合成雾下的漏检与类别混淆,以及重度运动模糊下精确率与召回率之间的权衡,表明当前方法对严重视觉信息损失的适应能力仍有限。

(3) 车载应用约束分析表明,目标距离与物理尺寸共同影响成像尺度,车辆速度、有效制动减速度及总响应时间共同决定所需停车距离。因此,检测性能的改善并不直接等同于实际预警效果的提升。后续将结合真实雾天数据、目标测距与动态车载试验,进一步研究不同环境和车辆运行条件下可靠检测距离与所需停车距离之间的关系。

参 考 文 献:

[1] 蒋弘毅,王永娟,康锦煌. 目标检测模型及其优化方法综述[J]. 自动化学报,2021,47(6):1232-1255.
JIANG H Y, WANG Y J, KANG J Y. A survey of object detection models and its optimization methods [J]. *Acta Automatica Sinica*, 2021, 47(6): 1232-1255. (in Chinese)

[2] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C]//*Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, 2014: 580-587.

[3] REN S Q, HE K M, GIRSHICK R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.

[4] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016: 779-788.

[5] LIU W, ANGELOV D, ERHAN D, *et al.* SSD: single shot MultiBox detector [C]. 14th European Conference on Computer Vision-ECCV 2016. Amsterdam, The Netherlands: Springer, 2016: 21-37.

- [6] ZUO X, YU J, XIAN T, *et al.* Road obstacle object detection based on improved YOLO V4 [J]. *International Journal of Advanced Network, Monitoring and Controls*, 2021, 6(3): 18-25.
- [7] 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 S, LI R, HE S. Road obstacle detection algorithm based on improved YOLOv5s [J]. *Electronic Measurement Technology*, 2023, 46(22): 177-185. (in Chinese)
- [8] 王龙春,方巍,张丽娟,等. 基于YOLOv8s改进的自动驾驶目标检测[J]. 液晶与显示, 2025, 40(5): 773-784.
WANG L C, FANG W, ZHANG L J, *et al.* Improved autonomous driving object detection based on YOLOv8s [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(5): 773-784. (in Chinese)
- [9] 李玉娟,葛动元,姚锡凡,等. 融合混合注意力与检测头的道路障碍物检测[J]. 科学技术与工程, 2026, 26(1): 253-261.
LI Y J, GE D Y, YAO X F, *et al.* Fusion of hybrid attention and detection head for road obstacle detection [J]. *Science Technology and Engineering*, 2026, 26(1): 253-261. (in Chinese)
- [10] PANG D F, GUAN Z W, LUO T, *et al.* RGE-YOLO enables lightweight road packaging bag detection for enhanced driving safety [J]. *Scientific Reports*, 2025, 15(1): 18306.
- [11] 纪文杰,宋廷伦,容小洛,等. PATD-YOLO:基于YOLOv11的道路障碍目标检测算法[J]. 计算机工程与应用, 2025, 61(21): 129-143.
JI W J, SONG T L, RONG X L, *et al.* PATD-YOLO: road obstacle object detection algorithm based on YOLOv11 [J]. *Computer Engineering and Applications*, 2025, 61(21): 129-143. (in Chinese)
- [12] HUANG J W, WANG K B, HOU Y, *et al.* LW-YOLO11: a lightweight arbitrary-oriented ship detection method based on improved YOLO11 [J]. *Sensors*, 2025, 25(1): 65.
- [13] ZHENG Z H, WANG P, LIU W, *et al.* Distance-IoU loss: faster and better learning for bounding box regression [C]// *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*. New York, NY, USA: AAAI, 2020: 12993-13000.
- [14] LI X, WANG W H, WU L J, *et al.* Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection [C]// *Proceedings of the 34th International Conference on Neural Information Processing Systems*. Vancouver, Canada: Curran Associates Inc., 2020: 1763.
- [15] 李智杰,惠爱婷,李昌华,等. 面向遥感图像道路提取的多尺度上下文感知网络[J]. 光学精密工程, 2025, 33(4): 610-623.
LI Z J, HUI A T, LI C H, *et al.* Multi-scale context-aware network for road extraction in remote sensing images [J]. *Optics and Precision Engineering*, 2025, 33(4): 610-623. (in Chinese)
- [16] LIU P J, ZHANG H Z, LIAN W, *et al.* Multi-level wavelet convolutional neural networks [J]. *IEEE Access*, 2019, 7: 74973-74985.
- [17] 刘旭,宋佩博,包芳勋,等. 融合小波卷积与频域注意力的小目标检测[J]. 中国图象图形学报, 2026, 31(4): 1125-1141.
LIU X, SONG P B, BAO F X, *et al.* Small-object detection model integrating wavelet convolution and frequency-domain attention [J]. *Journal of Image and Graphics*, 2026, 31(4): 1125-1141. (in Chinese)
- [18] DAUPHIN Y N, FAN A, AULI M, *et al.* Language modeling with gated convolutional networks [C]// *Proceedings of the 34th International Conference on Machine Learning*. Sydney, Australia: PMLR, 2017: 933-941.
- [19] 朱世松,高虹,芦碧波,等. 基于改进YOLOv8n的城市道路病害检测算法[J]. 液晶与显示, 2025, 40(9): 1356-1368.
ZHU S S, GAO H, LU B B, *et al.* Urban road defect detection algorithm based on improved YOLOv8n [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(9): 1356-1368. (in Chinese)
- [20] HOU Q B, ZHANG L, CHENG M M, *et al.* Strip pooling: rethinking spatial pooling for scene parsing [C]// *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, WA, USA: IEEE, 2020: 4002-4011.
- [21] ZHANG X, SONG Y Z, SONG T T, *et al.* LDConv: linear deformable convolution for improving convolutional neural networks [J]. *Image and Vision Computing*, 2024, 149: 105190.

- [22] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [23] 张洋, 孙海江, 张笑闻, 等. 基于局部上下文增强的快速 3D 弱特征目标检测方法[J]. *液晶与显示*, 2025, 40(11): 1675-1687.
- ZHANG Y, SUN H J, ZHANG X W, *et al.* Fast 3D weak feature object detection method based on local context enhancement [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(11): 1675-1687. (in Chinese)
- [24] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 936-944.
- [25] LIU S, QI L, QIN H F, *et al.* Path aggregation network for instance segmentation [C]//*Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA: IEEE, 2018: 8759-8768.
- [26] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection [C]//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, WA, USA: IEEE, 2020: 10778-10787.
- [27] PADILLA R, PASSOS W L, DIAS T L B, *et al.* A comparative analysis of object detection metrics with a companion open-source toolkit [J]. *Electronics*, 2021, 10(3): 279.
- [28] American Association of State Highway and Transportation Officials. *A Policy on Geometric Design of Highways and Streets* [M]. Washington, DC, USA: AASHTO, 2001.
- [29] JEONG Y, SON S, LEE B, *et al.* The braking-pressure and driving-direction determination system (BDDS) using road roughness and passenger conditions of surrounding vehicles [J]. *Sensors*, 2022, 22(12): 4414.

作者简介:



张传伟,男,博士,教授,2006年于西安交通大学获得博士学位,主要从事机电系统智能控制和智能网联汽车方面的研究。E-mail:zhangcw@xust.edu.cn



尚召斌,男,硕士研究生,2024年于陕西理工大学获得学士学位,主要从事目标检测和图像处理方面的研究。E-mail:2360601882@qq.com