文章编号:1673-0291(2026)03-0164-11 **DOI**:10.11860/j.issn.1673-0291.20250167

复杂环境下铁路轨道异物入侵轻量化检测研究

侯 涛,席 超,牛宏侠
(兰州交通大学 自动化与电气工程学院,兰州 730070)

摘 要:针对现有铁路轨道异物入侵检测算法易受到复杂环境的影响,导致检测精度低、计算量较大的问题,提出一种改进YOLOv12n的复杂环境下铁路轨道异物入侵检测方法(Railway Lightweighting-YOLO, RL-YOLO),实现轻量级部署.首先,引入高性能骨干网络(High Performance GPU NetV2, HGnetV2)替换原有的网络结构,并使用DWconv和Ghostconv对其进行组合改进,实现空间卷积-通道融合解耦、核心特征-廉价幽灵特征生成的双维度协同优化,在大幅降低模型参数量与浮点运算量的同时提升特征表征完整性.其次,采用轻量级非对称多级压缩检测头(Lightweight Asymmetric Dual-Head, LADH),进而对模型进行轻量化优化.再次,添加自适应细粒度通道注意力机制(Adaptive Fine-Grained Channel Attention, AFGCA),提高模型对复杂场景的理解能力.最后,采取Wise Focaler ShapeIoU(WF-ShapeIoU)损失函数,使模型更多地关注于复杂样本,提升模型的定位能力.研究结果表明:该方法在复杂环境下的平均检测精度$mAP_{0.5}$达89.1%,提升2.6%;处理速度达159帧/s,提升21帧/s;模型参数量减少41.7%,GFLOPS减少34.5%,满足实时性与轻量化需求.

关键词:铁路轨道;目标检测;复杂环境;异物入侵;轻量化
中图分类号:TP391.4;U298 **文献标志码**:A

Research on lightweight detection of foreign object intrusion on railway tracks in complex environments

HOU Tao, XI Chao, NIU Hongxia
(School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: To address the susceptibility of existing railway track foreign object intrusion detection algorithms to complex environments, which leads to low detection accuracy and high computational overhead, this paper proposes a lightweight detection method named Railway Lightweighting-YOLO (RL-YOLO). Based on an improved YOLOv12n architecture, this method is designed to facilitate lightweight deployment in complex environments. First, the High Performance GPU NetV2(HGNetV2) is introduced to replace the original backbone network, which is further optimized by integrating DWConv and GhostConv. This integration achieves dual-dimensional collaborative optimization, specifically, the decoupling of spatial convolution from channel fusion, and the generation of cost-effective

收稿日期:2025-10-16;**修回日期**:2026-03-03
基金项目:甘肃省重点研发计划-工业类(23YFGA0049);兰州市人才创新创业项目(2023-RC-13)
Foundation items: Gansu Province Key Research and Development Program under Industry(23YFGA0049); Talent Innovation and Entrepreneurship Project Lanzhou Ctiy(2023-RC-13)
第一作者:侯涛(1975—),男,四川中江人,教授,博士.研究方向为图像检测与智能控制. **email**:11230389@atu.lzjtu.edu.cn.

引用格式:侯涛,席超,牛宏侠.复杂环境下铁路轨道异物入侵轻量化检测研究[J].北京交通大学学报,2026,50(3):164-174.
HOU Tao, XI Chao, NIU Hongxia. Research on lightweight detection of foreign object intrusion on railway tracks in complex environments[J]. Journal of Beijing Jiaotong University, 2026, 50(3): 164-174.(in Chinese)

"ghost" features alongside core features. This approach significantly reduces the number of model parameters and floating-point operations while enhancing the integrity of feature representation. Second, Lightweight Asymmetric Dual-Head (LADH) is adopted to further streamline the model for lightweight deployment. Third, an Adaptive Fine-Grained Channel Attention (AFGCA) mechanism is incorporated to improve the model's capacity to interpret complex scenes. Finally, the Wise Focaler ShapeIoU (WF-ShapeIoU) loss function is employed to compel the model to focus more effectively on complex samples, thereby improving its localization capability. Experimental results demonstrate that the proposed method achieves a mean average precision (mAP@0.5) of 89.1% in complex environments, representing a 2.6% improvement over the baseline. The processing speed reaches 159 frames per second, an increase of 21 frames per second. Furthermore, the number of model parameters is reduced by 41.7%, GFLOPs is decreased by 34.5%. These results verify that the method successfully meets the demands for both real-time performance and lightweight deployment.

Keywords: railway tracks; object detection; complex environment; foreign object intrusion; lightweight deployment

在铁路运营中确保铁路安全及保障旅客安全是重中之重,若铁路轨道之中存在异物,可能导致列车车轮或轨道损坏从而造成行车延误,影响乘客的出行计划和货物运输,甚至于导致列车脱轨造成人员伤亡和财产损失,铁路轨道异物入侵已成为威胁铁路运输安全的核心风险之一.《交通强国建设纲要》[1]中明确要求推进铁路安全监测智能化,实现异物入侵实时预警.而深度学习在图像处理领域的成熟技术,为铁路入侵检测提供了高精度、实时的解决方案.

目前,基于深度学习的铁路异物入侵检测已被广泛研究.为解决铁路场景中异物目标尺度差异大、对比度低等检测难点,众多研究者通过特征增强、网络结构改进及损失函数优化等方式针对性提升检测准确率.文献[2]设计一种改进的加权双向特征金字塔网络结构,加强多尺度特征融合能力.文献[3]针对铁路异物检测痛点,构建全局语义与局部特征融合方法,有效降低漏检误检率,兼具良好实时性与鲁棒性.文献[4]提出基于特征降冗余的铁路异物检测算法,其利用改进局部卷积减少特征冗余,并引入SE注意力机制增强对关键目标的特征表达能力.文献[5]基于多尺度特征增强,通过扩散融合强化多尺度特征,融入可变尺寸步长模块实现自适应多尺度特征提取,引入滑动损失函数适配多尺度检测任务.文献[6]通过优化Mask-RCNN(Mask Region-based Convolutional Neural Network),提高低对比度目标的识别能力和对小尺寸异物的检测精度,但该模型在处理数据集时的计算效率较低.上述研究在精度优化上成效显著,但部分算法因其计算量大、复杂度高、对复杂环境泛化能力不足等问题,难以实现精度与实时性的同步提升.

为满足列车运行过程中异物检测的实时性要求,轻量化网络设计与计算复杂度优化成为另一研究方向.文献[7]设计基于浅层特征融合的轻量级算法(YOLO Lightweight and Shallow-feature Fusion, YOLO-LSF),提升检测精度速度的同时压缩模型体量,有效降低复杂环境下的漏检与误检概率.文献[8]构建一种轻量级高效铁路异物入侵检测模型.通过多通道增强特征注意力机制提升多尺度检测效果,设计轻量化共享检测头减少参数量,并结合自适应剪枝与知识蒸馏压缩模型.文献[9]针对每个像素赋予不同的膨胀率,以此降低模型的复杂度及计算成本.文献[10]基于YOLOv7-CGGA模型结合GhostConv卷积减少50%参数量,并采用Alpha_GIoU损失函数提高检测精度.尽管现有轻量化研究在参数缩减与检测提升方面表现优异,但在复杂场景下的泛化性及速度与精度的平衡上仍存在优化空间.

此外,部分学者从场景分割、目标关系建模等技术维度拓展了铁路场景感知能力的新方向.文献[11]设计的RailEINet网络利用特征对齐技术实现了列车运行场景的精准分割,通过强化语义信息提取完善铁路感知体系;文献[12]提出的Railformer框架则引入了RFEM模块与CCTF结构,RFEM用于增强小目标与场景结构的信息提取,CCTF则通过交叉耦合机制实现目标特征与关系特征的协同优化.此类研究着眼于场景的整体感知,为复杂环境下铁路轨道异物入侵检测提供了新的视角.

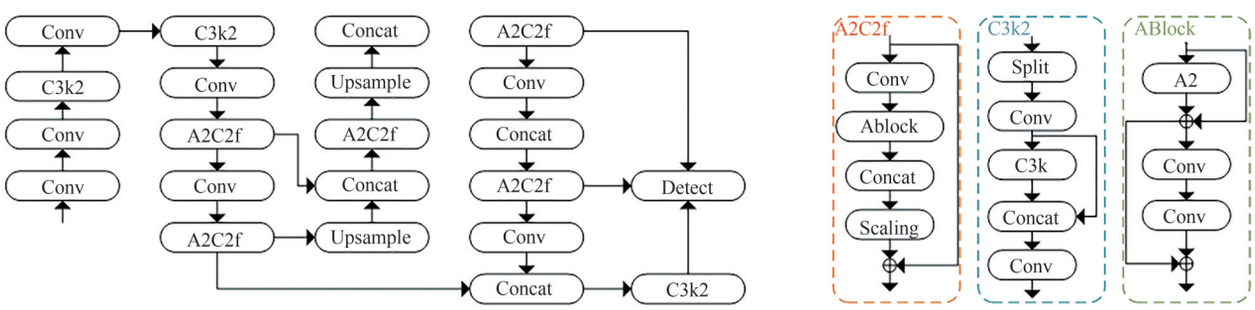
综上所述,现有的铁路轨道异物检测算法存

在难以平衡检测精度和速度的问题，鉴于YOLOv12n[13]能够更好地从复杂场景中提取关键特征并捕捉不同尺度的对象，故本文以YOLOv12n为基础，提出一种轻量化的铁路轨道异物入侵检测算法RL-YOLO，旨在适配铁路轨道复杂场景的轻量化与高精度检测需求，RL-YOLO算法构建全链路轻量化协同优化架构，通过多模块层级化协同实现功能互补与性能增益．算法以基于HGnetV2的轻量化特征提取网络DG-HGnet为核心支撑，通过嵌入轻量化卷积算子在保障多尺度特征表征能力的同时降低计算冗余．其次，引入自适应细粒度通道注意力机制AFGCA，其通过多粒度动态权重分配，以较低计算量实现全局与局部信息融合，从而输出高质量的特征．再次，LADH借助非对称压缩与双头协同机制，在精简模型计算复杂度的同时提升推理效率，与前端的轻量化特征提取网络形成高度适配的联动架构．而在特征输出环节则通过A2C2f替换C3k2，再度削减参数量与计算量，与DG-HGnet特征提取网络形成前后协同，完善轻量化链路．最后，设计WF-ShapeIoU损失函数，其通过差异化回归样本聚焦训练，在无额外推理开销前提下改善漏误检问题，为前端轻量化模块性能发挥提供后端保障．综上，完成各模块深度协同的全链路轻量化设计，使RL-YOLO高效达成轻量化与高精度检测的协同优化目标，适配复杂环境下铁路轨道异物入侵场景部署需求．

1 RL-YOLO算法整体结构

1.1 YOLOv12n算法

YOLOv12n系列算法主要由主干网络（Backbone）、颈部（Neck）和头部（Head）组成．Backbone负责从输入图像中提取特征，Neck负责连接主干网络和头部，对特征进行融合，Head负责预测．选用尺度最小的YOLOv12n作为基准模型，其通过注意力机制与混合架构的突破，在保持实时性的同时显著提升检测精度，尤其在复杂场景和动态环境中表现突出．YOLOv12n算法结构见图1．



（a）整体结构　　（b）模块细节

图1 YOLOv12n算法结构图

Fig.1 Structure diagram of YOLOv12n algorithm

1.2 RL-YOLO算法

YOLOv12n算法具有较好的检测效果，但由于铁路设备的计算资源有限，难以针对铁路轨道异物目标多样，图像背景环境复杂的情况做出有效检测[14]．故本文在YOLOv12n算法的基础上，提出一种铁路轨道异物入侵检测算法RL-YOLO，旨在保持高精度的同时减少计算资源的消耗．其结构见图2．

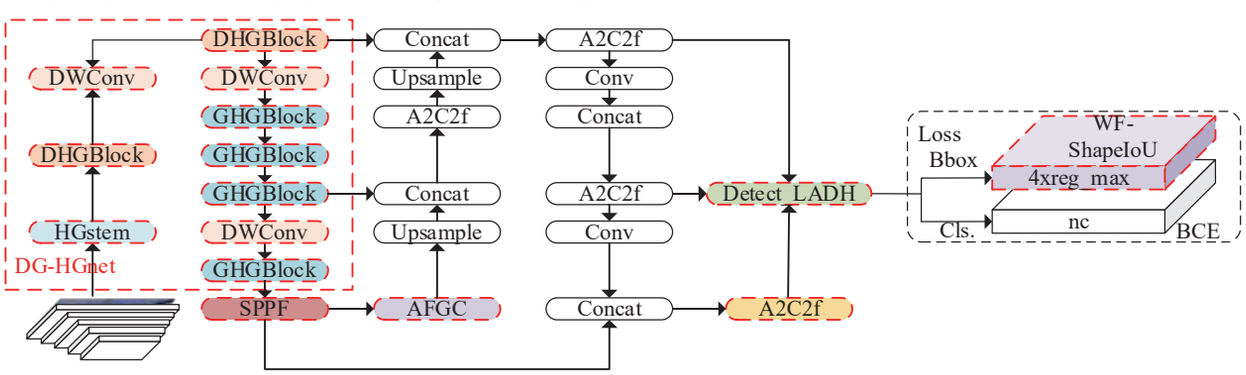


图2 RL-YOLO算法结构图

Fig.2 Structure diagram of RL-YOLO algorithm

2　RL-YOLO 算法过程

2.1　设计 DG-HGnetV2 特征提取网络

2.1.1　DWConv

DWConv[15]是深度可分离卷积的第一阶段,即通过逐通道独立卷积,以显著降低计算量和参数量,实现轻量化.

假设 D_F 为输入特征矩阵的宽高,D_K 为卷积核的大小,M 为输入特征矩阵的深度,N 代表输出特征矩阵的深度;在传统卷积中,参数量为 $D_K \times D_K \times M \times N \times D_F \times D_F$;而 DWConv 每个卷积核仅处理单通道,参数量为 $D_K \times D_K \times M \times D_F \times D_F$,仅为传统卷积的 $1/N$.其结构见图 3.

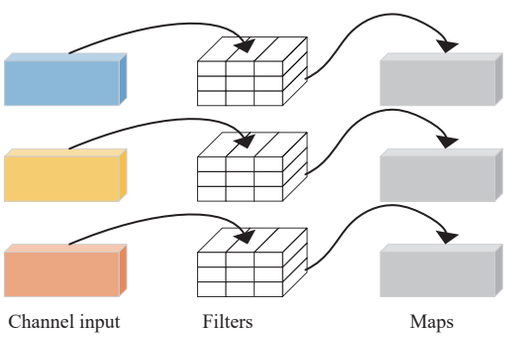


图 3　DWConv 结构图

Fig.3　Structure diagram of DWConv

2.1.2　GhostConv

GhostConv[16]旨在通过生成冗余特征图减少计算量,同时保持模型性能.其使用少量卷积核提取输入特征图的核心特征生成初始特征图,再对初始特征图进行线性变换,生成与内在特征图数量相近的冗余特征,最终拼接形成完整输出.其结构见图 4.

输入特征图 $\boldsymbol{X} \in \mathbb{R}^{H \times W \times C_{in}}$ 通过 1×1 卷积生成内在特征 $\boldsymbol{X}_{pri} \in \mathbb{R}^{H \times W \times C_{mid}}$,其表达式为

$$\boldsymbol{X}_{pri} = \mathrm{Conv}_{1\times1}(\boldsymbol{X}) \tag{1}$$

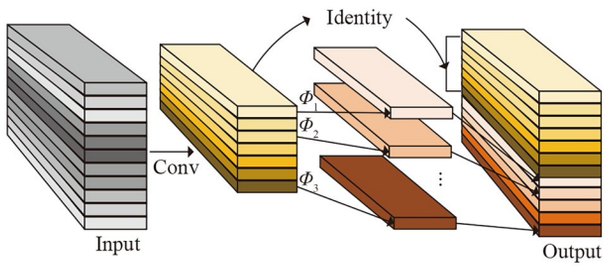


图 4　GhostConv 结构图

Fig.4　Structure diagram of GhostConv

$$C_{mid} = [\, C_{out}/s \,] \tag{2}$$

式中:C_{in} 为输入通道数;C_{mid} 为中间通道数;C_{out} 为输出通道数;H 与 W 为特征图尺寸;s 为冗余倍数.

再对 $\boldsymbol{X}_{pri}$ 进行深度可分离卷积,生成冗余特征 $\boldsymbol{X}_{cheap} \in \mathbb{R}^{H \times W \times C_{mid} \times (s-1)}$,其表达式为

$$\boldsymbol{X}_{cheap} = \mathrm{DWConv}(\boldsymbol{X}_{pri}) \tag{3}$$

式中:DWConv 的卷积核为 $K \times K \times C_{mid} \times 1$;$K \times K$ 为卷积核的维度.

最终输出表达式 $\boldsymbol{Y} \in \mathbb{R}^{H \times W \times C_{mid}}$ 为

$$\boldsymbol{Y} = \mathrm{Concat}(\boldsymbol{X}_{pri}, \boldsymbol{X}_{cheap}) \tag{4}$$

2.1.3　DG-HGnet

HGBlock 作为 HGnetV2[17]的核心特征提取单元,具备分层特征融合的能力,但受限于标准卷积的固有结构导致其计算复杂度偏高.由于卷积神经网络的浅层与深层特征存在显著的异质性:浅层负责提取边缘、纹理等低级局部特征,而深层聚焦于语义、类别等高级全局特征.故针对浅层特征空间关联性强、通道独立性高的特点,使用深度可分离卷积 DWConv 的逐通道卷积范式,在降低参数量的同时保留局部细节;针对深层特征的语义信息强、空间分辨率低的特点,采取 GhostConv 的核心特征与线性变换机制,以低成本扩展特征维度,保证复杂异物目标的识别性能.综上,本文将 DWConv、GhostConv 分别与 HGBlock 融合,提出 D-HGBlock 与 G-HGBlock,实现网络不同层级的精准轻量化优化,其结构见图 5.

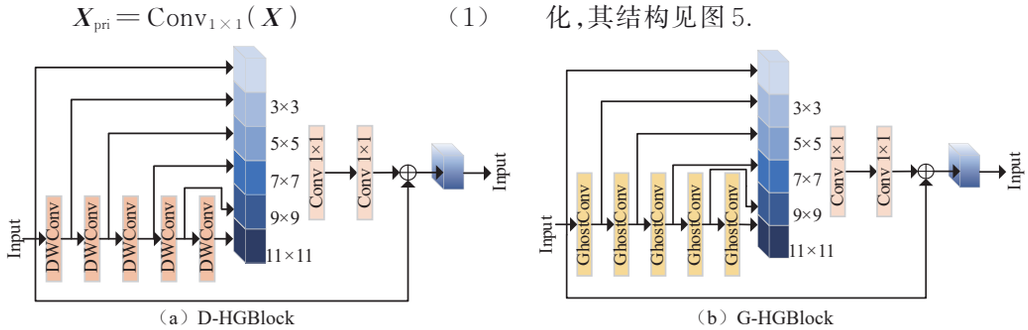


图 5　HGBlock 结构图

Fig.5　Structure diagram of HGBlock

HGnetV2采用分层特征提取机制，实现了全局语义信息与局部细节特征的有效平衡；通过渐进式下采样策略逐步降低特征图分辨率，在显著降低模型计算复杂度的前提下，扩大特征感受野，从而达成轻量化设计目标．为优化特征提取能力与模型效率，DG-HGnet以HGnetV2为基础框架进行改进：在网络前两层引入D-HGBlock模块，该模块通过逐通道独立卷积操作精简参数规模，实现轻量化的初步优化；在网络后四层则采用G-HGBlock模块，通过生成多尺度冗余特征表示，在保障模型检测精度不受损失的前提下，提升轻量化水平．DG-HGnet的网络结构如图6所示．

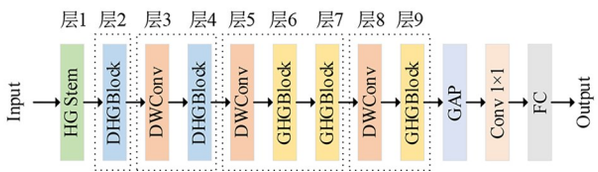


图6 DG-HGnet结构图

Fig.6 Structure diagram of DG-HGnet

2.2 设计WF-ShapeIoU损失函数

针对铁路轨道异物检测复杂场景下的漏检、误检及定位精度不足等问题，单一的目标检测损失函数难以适配该场景的特殊性与复杂性，导致检测性能受限．为此，首先，以ShapeIoU[18]的形状约束机制为基础解决目标形状匹配偏差问题．其次，引入Focaler[19]方法中困难样本挖掘策略改善样本权重分配不均问题．再次，结合WIoU[20]的动态权重调整机制实现对困难样本及小目标的精准聚焦．最后，形成从样本筛选到形状匹配及动态加权的三方协同优化，构建WF-ShapeIoU损失函数替代CIoU，以提升改进算法在铁路轨道异物入侵检测中的定位精度．目标定位示意图见图7．

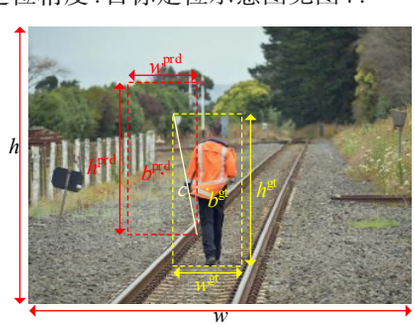


图7 目标定位示意图

Fig.7 Target positioning diagram

YOLOv12n采用CIoU的损失函数，CIoU表达式为

$$\mathrm{IoU}=\frac{|B\cap B^{gt}|}{|B\cup B^{gt}|} \tag{5}$$

$$L_{\mathrm{IoU}}=1-\mathrm{IoU} \tag{6}$$

$$L_{\mathrm{CIoU}}=1-\mathrm{IoU}+\frac{\rho^2(b^{prd},b^{gt})}{c^2}+\alpha v \tag{7}$$

$$v=\frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}}-\arctan\frac{w^{prd}}{h^{prd}}\right)^2 \tag{8}$$

$$\alpha=\frac{v}{(1-\mathrm{IoU})+v} \tag{9}$$

式中：IoU为交并比；B为预测框；B^{gt}为真实框；L_{IoU}为IoU损失；L_{CIoU}为CIoU损失；b^{prd}和b^{gt}分别为预测框和真实框的中心点；w^{gt}和w^{prd}分别为目标框和预测框的宽；h^{gt}和h^{prd}分别为目标框和预测框的高；ρ为计算两个中心点之间距离的函数；c为预测框和真实框最小闭包框的对角线长度；v为衡量长宽比一致性的项；α是权重系数．

而ShapeIoU通过引入形状约束和尺度权重，更精准地评估预测框的质量．它通过引入一个最大精度项来优化定位精度，帮助模型更好地学习目标的定位．ShapeIoU表达式为

$$L_{\mathrm{ShapeIoU}}=1-\mathrm{IoU}+S+0.5\times\Omega \tag{10}$$

$$S=\frac{h_{\mathrm{q}}\times(x_{\mathrm{c}}^{prd}-x_{\mathrm{c}}^{gt})^2}{c^2}+\frac{w_{\mathrm{q}}\times(y_{\mathrm{c}}^{prd}-y_{\mathrm{c}}^{gt})^2}{c^2} \tag{11}$$

$$\Omega=\sum_{t=w,h}(1-\mathrm{e}^{-w_t})^4 \tag{12}$$

$$w_{\mathrm{q}}=\frac{2\times(w^{gt})^{\mathrm{scale}}}{(w^{gt})^{\mathrm{scale}}+(h^{gt})^{\mathrm{scale}}} \tag{13}$$

$$h_{\mathrm{q}}=\frac{2\times(h^{gt})^{\mathrm{scale}}}{(w^{gt})^{\mathrm{scale}}+(h^{gt})^{\mathrm{scale}}} \tag{14}$$

$$\begin{cases} w_w=h_{\mathrm{q}}\times\dfrac{|w^{prd}-w^{gt}|}{\max(w^{prd},w^{gt})} \\ w_h=w_{\mathrm{q}}\times\dfrac{|h^{prd}-h^{gt}|}{\max(h^{prd},h^{gt})} \end{cases} \tag{15}$$

式中：$(x_{\mathrm{c}}^{prd},y_{\mathrm{c}}^{prd})$为预测框中心坐标；$(x_{\mathrm{c}}^{gt},y_{\mathrm{c}}^{gt})$为真实框中心坐标；scale是缩放因子；$w_{\mathrm{q}}$和$h_{\mathrm{q}}$分别表示水平方向和垂直方向上的权重系数；$S$为中心点距离；$w_{\mathrm{w}}$与$w_{\mathrm{h}}$分别为宽和高的相对误差权重；$\Omega$为非线性形状惩罚．

使用Focaler-IoU替换ShapeIoU中的IoU，Focaler-IoU关注于不同的回归样本，能够减少简单样本在损失中的权重，从而让模型更多地关注那些复杂样本．Focaler-IoU表达式为

$$\mathrm{IoU}^{\mathrm{Focaler}}=\begin{cases} 0 & \mathrm{IoU}<d \\ \dfrac{(\mathrm{IoU}-d)}{u-d} & d\leqslant\mathrm{IoU}\leqslant u \\ 1 & \mathrm{IoU}>d \end{cases} \tag{16}$$

式中：$[d,u]\in[0,1]$．F-ShapeIoU损失函数表达式为

$$L_{\text{F-ShapeIoU}}=1-\text{IoU}^{\text{Focaler}}+S+0.5\times\Omega \quad (17)$$

而 WIoU 通过动态调整样本权重,使模型在训练过程中更关注困难样本和小目标,解决了传统 IoU 损失在复杂场景下的局限性.故在 F-ShapeIoU 中引入 WIoU,WIoU 损失函数表达式为

$$L_{\text{WIoU}}=\exp\left(\frac{(x_c^{\text{prd}}-x_c^{\text{gt}})^2+(y_c^{\text{prd}}-y_c^{\text{gt}})^2}{w_g^2+h_g^2}\right)L_{\text{IoU}} \quad (18)$$

式中:w_g与h_g分别为真实框与预测框的最小外接矩形的宽与高.故 WF-ShapeIoU 损失函数表达式为

$$L_{\text{WF-ShapeIoU}}=\exp\left(\frac{(x_c^{\text{prd}}-x_c^{\text{gt}})^2+(y_c^{\text{prd}}-y_c^{\text{gt}})^2}{w_g^2+h_g^2}\right)L_{\text{F-ShapeIoU}} \quad (19)$$

2.3　引入 LADH 检测头

LADH[21]作为非对称解耦检测模块,旨在解决耦合检测头中分类与回归任务冲突、常规解耦检测头参数量冗余、推理效率偏低的问题.LADH 核心架构由双头并行解耦 Dual-Head 构成.Dual-Head 通过构建分类、回归两个完全并行独立的分支实现任务解耦,两分支虽共享预处理特征但参数完全独立,既规避任务优化目标冲突,又借助并行计算提升推理效率.LADH 的结构如图 8 所示.

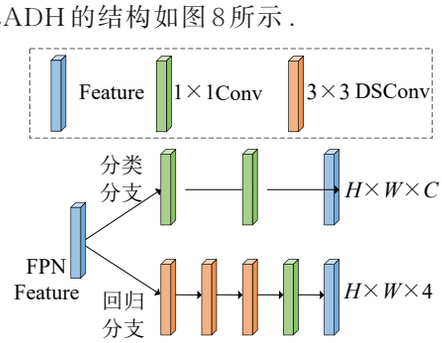


图 8　LADH 结构图

Fig.8　Structure diagram of LADH

输入特征图为$\boldsymbol{X}\in\mathbb{R}^{B\times C_{\text{in}}\times H\times W}$,$B$为批次大小.输入特征图经 2 个 1×1 卷积提取特征后,通过卷积输出类别得分,得到分类分支 classification(cls),其表达式为

$$\boldsymbol{X}_{\text{cls_feat}}=\text{Conv}_{1\times1}\times(\text{Conv}_{1\times1}(\boldsymbol{X};\ \boldsymbol{W}_{\text{cls1}},\boldsymbol{b}_{\text{cls1}});\ \boldsymbol{W}_{\text{cls2}},\boldsymbol{b}_{\text{cls2}}) \quad (20)$$

$$\boldsymbol{Y}_{\text{cls}}=\text{Conv}_{1\times1}(\boldsymbol{X}_{\text{cls_feat}};\ \boldsymbol{W}_{\text{cls_out}},\boldsymbol{b}_{\text{cls_out}}) \quad (21)$$

式中:$\boldsymbol{X}_{\text{cls_feat}}$为分类分支输入特征图;$\boldsymbol{Y}_{\text{cls}}$为分类分支输出特征图;$\boldsymbol{W}_{\text{cls1}}$、$\boldsymbol{W}_{\text{cls2}}$分别为两个 1×1 卷积的权重张量;$\boldsymbol{b}_{\text{cls1}}$、$\boldsymbol{b}_{\text{cls2}}$分别为两个 1×1 卷积的偏置向量;$\boldsymbol{W}_{\text{cls_out}}$、$\boldsymbol{b}_{\text{cls_out}}$分别为分类输出的卷积权重与偏置.

输入特征图经过 3 个 3×3 的深度可分离卷积 DSConv,1 个 1×1 的卷积降维后,输出边界框偏移量得到回归分支 regression(reg),其表达式为

$$\boldsymbol{X}_{\text{reg_feat}}=\text{DSConv}_{3\times3}(\text{DSConv}_{3\times3}(\text{DSConv}_{3\times3}\times(\boldsymbol{X};\ \boldsymbol{W}_{\text{reg1}},\boldsymbol{b}_{\text{reg1}});\ \boldsymbol{W}_{\text{reg2}},\boldsymbol{b}_{\text{reg2}});\ \boldsymbol{W}_{\text{reg3}},\boldsymbol{b}_{\text{reg3}}) \quad (22)$$

$$\boldsymbol{X}_{\text{reg_mid}}=\text{Conv}_{1\times1}(\boldsymbol{X}_{\text{reg_feat}};\ \boldsymbol{W}_{\text{reg_stem}},\boldsymbol{b}_{\text{reg_stem}}) \quad (23)$$

$$\boldsymbol{Y}_{\text{reg}}=\text{Conv}_{1\times1}(\boldsymbol{X}_{\text{reg_mid}};\ \boldsymbol{W}_{\text{reg_out}},\boldsymbol{b}_{\text{reg_out}}) \quad (24)$$

式中:$\boldsymbol{X}_{\text{reg_feat}}$为回归分支输入特征图;$\boldsymbol{X}_{\text{reg_mid}}$为回归中间特征图;$\boldsymbol{Y}_{\text{reg}}$为回归输出特征图;$\boldsymbol{W}_{\text{reg1}}$、$\boldsymbol{W}_{\text{reg2}}$及$\boldsymbol{W}_{\text{reg3}}$分别为 3 个 3×3 深度可分离卷积的权重;$\boldsymbol{b}_{\text{reg1}}$、$\boldsymbol{b}_{\text{reg2}}$及$\boldsymbol{b}_{\text{reg3}}$分别为 3 个 3×3 深度可分离卷积的偏置;$\boldsymbol{W}_{\text{reg_stem}}$、$\boldsymbol{b}_{\text{reg_stem}}$分别为经过 1 个 1×1 的卷积降维后的卷积权重与偏置;$\boldsymbol{W}_{\text{reg_out}}$、$\boldsymbol{b}_{\text{reg_out}}$分别为回归输出的卷积权重与偏置.

2.4　采用 AFGCA 注意力机制

AFGCA[22]模块核心目标在于突破通道注意力过度侧重全局特征而忽视局部细节表征的固有局限,通过融合全局与局部特征实现高效的通道校准,提升复杂场景的识别性能.一方面,借助全局池化语义表征能力与通道一致性假设实现全局平均池化,有效剔除特征冗余信息、聚合关键语义特征,同时抑制模型的过拟合;另一方面,基于局部感受野与稀疏关联机制,通过一维卷积或自注意力补充局部信息,适配通道局部特性.AFGCA 整体以注意力特征选择理论为核心,融合多理论并借鉴视觉双加工规律,形成兼顾全局与局部的校准方案.其结构如图 9 所示.

首先,对于输入特征图的第n个通道$\boldsymbol{F}_n$,使用全局平均池化(GAP)对所有空间像素点的灰度值取均值,得到该通道的全局响应U_n,其表达式为

$$U_n=\text{GAP}(\boldsymbol{F}_n)=\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}\boldsymbol{F}_n(i,j)\quad n\in[1,C] \quad (25)$$

式中:i与j分别为特征图高与宽的索引;C为通道数.

其次,引入局部交互矩阵,采用滑动窗口策略对相邻通道进行加权聚合,以此捕获通道间的局部依赖关系.局部信息向量$\boldsymbol{U}_{\text{lc}}$的第$a$个元素$U_{\text{lc},a}$表达式为

$$U_{\text{lc},a}=\sum_{z=1}^{K}\boldsymbol{G}_{a,z}\cdot U_{a+z-\lfloor K/2\rfloor}\quad a\in\{1,2,\cdots,C\} \quad (26)$$

式中:$\boldsymbol{G}_{a,z}$为学习权重矩阵;z为滑动窗口内的索引.

再次,全局信息向量$\boldsymbol{U}_{\text{gc}}$由对角矩阵$\boldsymbol{D}$与全局通道描述符$\boldsymbol{U}$的矩阵乘法得到;相关性矩阵$\boldsymbol{M}$由全局信息向量$\boldsymbol{U}_{\text{gc}}$与局部信息向量$\boldsymbol{U}_{\text{lc}}$的转置外积得到,然后对相关性矩阵$\boldsymbol{M}$行求和得到全局权重向量$\boldsymbol{U}_{\text{gc}}^{\text{w}}$,列求和得到局部权重向量$\boldsymbol{U}_{\text{lc}}^{\text{w}}$.为动态调节全局

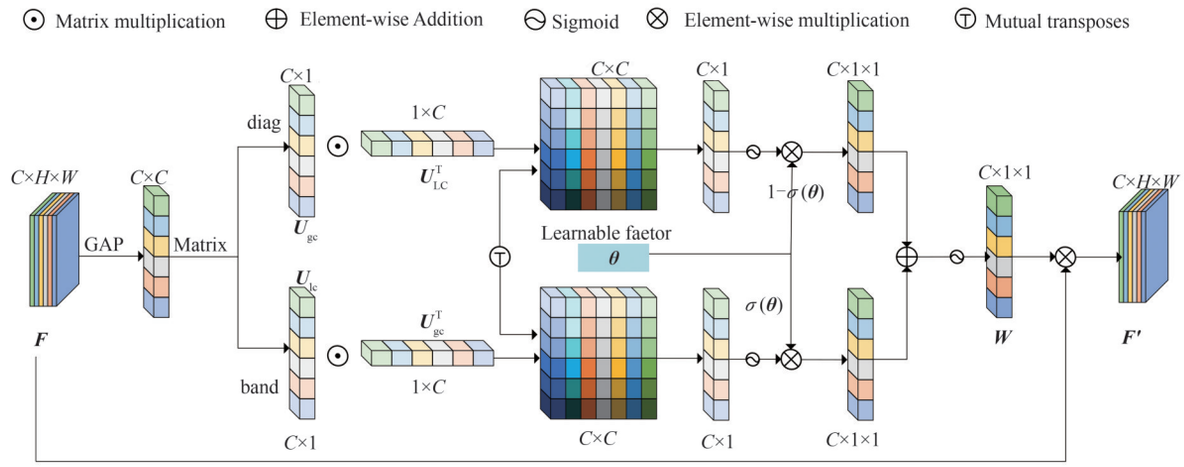


图 9 AFGCA 结构图

Fig.9 Structure diagram of AFGCA

权重与局部权重的贡献占比，引入可学习平衡因子 θ，通过 Sigmoid 激活将其映射至[0,1]区间，得到动态调节全局与局部权重的占比 $\boldsymbol{W}_{\text{fusion}}$，对融合后的权重向量再次执行 Sigmoid 激活，得到最后的通道权重，其表达式为

$$\text{Sigmoid}=\sigma(x)=\frac{1}{1+\mathrm{e}^{-x}} \tag{27}$$

$$\boldsymbol{W}=\sigma(\boldsymbol{W}_{\text{fusion}})=\sigma(\sigma(\theta)\cdot\sigma(\boldsymbol{U}_{\text{gc}}^{\text{w}})+(1-\sigma(\theta))\cdot\sigma(\boldsymbol{U}_{\text{lc}}^{\text{w}})) \tag{28}$$

最后，将生成的通道权重 $\boldsymbol{W}$ 作用于原始输入特征图 $\boldsymbol{F}$，实现强化关键通道、抑制冗余通道的特征校准，输出优化后的特征图 $\boldsymbol{F}'$ 表达式为

$$F_{\text{d}}'(i,j)=W_{\text{d}}\cdot F_{\text{d}}(i,j) \tag{29}$$

式中：W_{d} 为向量 $\boldsymbol{W}$ 的第 d 个元素.

3 实验与结果分析

3.1 实验环境和参数配置

模型训练实验环境见表 1.

表 1 模型训练实验环境

Tab.1 Model training experimental environment

配置名称	版本/参数	配置名称	版本/参数
操作系统	Ubuntu 20.04	内存/GB	12
CPU	Intel(R) Xeon(R) E5-v4	Python	11.8
GPU	NVIDIA RTX 3060	深度学习框架	Pytorch

在模型训练过程中，采用随机梯度下降的优化器以跳出局部最小值，用于随机梯度下降的动量因子设置为 0.937，而算法初始学习率设置为 0.01，批次大小 batch-size 设为 16，输入图像大小为 640×640，实验共训练 200 轮.

3.2 实验数据集

复杂环境下的图像特征是指在多样化的、非理想的环境条件下，图像中出现的各种特点和挑战.包括但不限于：强光、阴影、低光照等光照变化；背景可能包含多种物体、纹理和颜色；雨雾、夜间、沙尘等不同天气；部分目标可能被其他物体遮挡、图像中可能包含不同尺寸的物体且不同物体可能具有相似的纹理或颜色.

针对铁路轨道天气环境复杂多变的问题，本文通过网络开源数据集及图像，自制铁路轨道异物检测数据集 RD(Railway-Datasats)，将复杂多姿态的行人(person)和动物(animal)作为目标进行检测，同时挑选阴天、雪天、沙尘、能见度低场景及夜间等复杂环境构建数据集.数据集共 9 400 张图片，测试集 400 张，训练集 8 000 张，验证集 1 000 张.部分数据集样例见图 10.

3.3 评价指标

采用平均精确度 P、均值平均精度 mAP、召回率 R、FPS、参数量 Parameters 和计算量 GFLOPs 作为评价指标.其表达式为

$$P=\frac{T_{\text{P}}}{T_{\text{P}}+F_{\text{P}}} \tag{30}$$

$$R=\frac{T_{\text{P}}}{T_{\text{P}}+F_{\text{N}}} \tag{31}$$

$$\text{mAP}=\frac{1}{m}\sum_{i=1}^{n}P_{\text{A}}=\int_{0}^{1}P(R)\mathrm{d}R \tag{32}$$

式中：T_{P} 为阳性样本被正确分类(真阳性)；F_{P} 为阴性样本被错误分类(假阳性)；F_{N} 为阳性样本被错误分类(假阴性)；m 为类别数量.

3.4 模型训练

将 RL-YOLO 模型的 IoU 阈值设置为 0.5，训练 200 轮，训练结果见图 11.

由图 11 可知：RL-YOLO 和 YOLOv12n 在 0～



（a）雪天　（b）雨天　（c）沙尘　（d）夜间

（e）遮挡　（f）多目标　（g）模拟低光照　（h）模拟噪点

图 10　数据集样例

Fig.10　Sample dataset

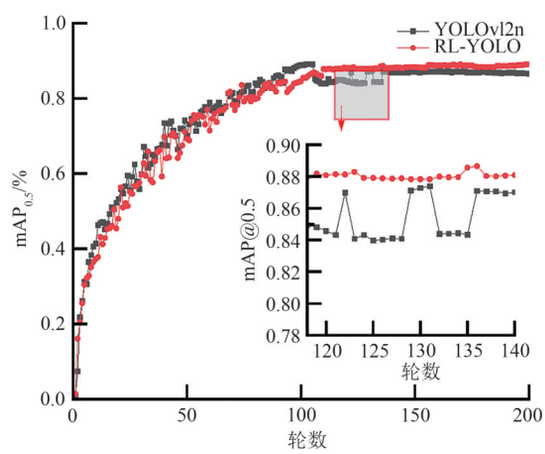


图 11　训练结果图

Fig.11　Training results chart

100 轮的训练前期均实现快速收敛，mAP_{50}同步快速上升，收敛效率基本持平；进入 100 轮后的收敛平台期后，RL-YOLO 的 $mAP_{0.5}$持续稳步抬升并最终趋于稳定，而 YOLOv12n 精度长期停滞在 0.87～0.88 区间，且出现了显著的锯齿状振荡，精度最低跌至 0.84，充分证明改进后的 RL-YOLO 相较 YOLOv12n，在保持同等训练收敛效率的同时，既实现了目标检测精度的显著提升，也大幅优化了模型训练的稳定性与鲁棒性.

3.5　消融实验

为验证 RL-YOLO 算法的各改进模块有效性及模块之间的协同作用，设计了一系列消融实验.实验以未加入任何改进模块的组别 0 作为基线模型，在此基础上，逐步加入 DG-HGnet、SPPF、AFGCA、LADH、A2C2f 和 WF-ShapeIoU 共 6 个改进模块，重点分析各个模块对模型检测性能、计算效率及推理速度的具体影响，实验结果见表 2.

表 2　消融实验

Tab.2　Ablation experiment

组别	DG-HGnet	SPPF	AFGCA	LADH	A2C2f	WF-ShapeIoU	$mAP_{0.5}$/%	GFLOPs	Param	FPS
0	×	×	×	×	×	×	86.5	5.8	2508734	138.3
1	√	×	×	×	×	×	84.2	4.8	1616494	157
2	√	√	×	×	×	×	89.1	5.0	1780718	150
3	√	√	√	×	×	×	86.5	5.0	1846516	157
4	√	√	√	√	×	×	88.4	3.8	1545716	161
5	√	√	√	√	√	×	87	3.8	1463668	159
6	√	√	√	√	√	√	89.1	3.8	1463668	159
7	√	√	×	√	√	√	86.3	3.8	1397870	159

注：“√”表示使用该模块，“×”表示未使用该模块.

由表 2 可知：将改进方法分别设计了 7 个实验，第 0 组为原模型；第 1 组在原模型的基础上加入轻量化特征提取网络 DG-HGnet，参数量减少 35.6%，GFLOPs 减少 17.2%，FPS 提升 18.7 帧/s，使其在降低一定精度的条件下，轻量化模型；第 2 组在第 1 组的基础上加入了 SPPF，检测精度提高 4.9%，实现较高精度检测；第 3 组在第 2 组的基础上加入了 AFGCA 模块，提升检测速度；第 4 组在第 3 组的基础上加入了 LADH 模块，通过对比 AFGCA 模块及 DG-HGnet 模块的融合改进，精度提升，在原模型的

基础上精度提升1.9%,参数量减少38.3%,计算量减少34.5%;第5组在第4组的基础上,改变了A2C2f模块,参数量进而降低;第6组在第5组的基础上加入WF-ShapeIoU,在保持模型参数量和计算量不变的前提下,使得精度提升2.1%;而第7组在第6组基础上减去AFGCA模块,验证了该模块在融合改进时的有效性.

3.6 改进算法泛化性验证实验

为验证改进算法的泛化能力与跨场景适配性能,通过选取公开数据集COCO进行验证实验.为保证验证实验的可靠性,在同一条件下,选取与RD数据集中相同的检测对象,验证跨场景性能.验证实验结果如表3所示.

表3 改进算法泛化性验证实验

Tab.3 Improvement of algorithm generalization validation experiment

检测模型	$mAP_{0.5}$/%	GFLOPs	Param	FPS
YOLOv12n	75.8	5.8	2508734	165
RL-YOLO	76.8	3.8	1463668	189

由表3可知:在COCO数据集的跨场景验证中,相较于基线模型YOLOv12n,本文所提RL-YOLO模型的$mAP_{0.5}$提升1.0%,同时GFLOPs降低34.48%,参数量压缩41.66%,推理FPS提升14.55%.实验结果表明,所提改进算法在脱离RD数据集的专属场景后,仍能实现检测精度与推理效率的同步提升,具备优异的泛化能力与跨场景适配性能,可满足不同场景下的目标检测任务需求.

3.7 对比实验

3.7.1 特征提取网络改进对比实验

为验证特征提取网络改进的合理性,通过对比YOLOv12n原有的backbone、HGnetV2及改进后的DG-HGnetV2网络进行对比实验,实验结果见表4.

表4 特征提取网络改进对比实验

Tab.4 Comparison experiment of feature extraction network before and after improvement

特征提取网络	$mAP_{0.5}$/%	GFLOPs	Param	FPS
backbone	86.5	5.8	2508734	138.3
HGnetV2	80.6	5.1	1677970	160
DG-HGnetV2	84.2	4.8	1616494	158

由表4可知:HGnetV2在YOLOv12原有的特征提取网络基础上,参数量减少33.1%,计算量减少12%,检测速度提升21.7帧/s,初步满足轻量化需求;基于HGnetV2改进的DG-HGnetV2与原有的特征提取网络相比,精度下降2.3%,但参数量减少35.6%,计算量减少17.2%,检测速度提升19.7帧/s,在保持一定检测精度的前提下满足轻量化.

3.7.2 DG-HGnetV2模块改进对比试验

为验证DG-HGnetV2改进模块的合理性,以组别0为基线,通过对比3个模块HGBlock(0)、D-HGBlock(1)及G-HGBlock(2)分别在图6中层1、层3、层5、层6、层7和层9的不同位置组合进行实验,实验结果由与基线对比的百分比表示,实验结果见表5.

表5 不同模块位置组合对比实验

Tab.5 Comparison experiment of different module position combinations

组别	层1	层3	层5	层6	层7	层9	$mAP_{0.5}$/%	GFLOPs/%	Param/%	FPS/%
0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	−1.4	−0.2	−2.5	−11
2	2	2	2	2	2	2	−3	−0.5	−3.7	−4
3	2	2	0	0	0	0	−2.2	−0.4	−2.6	−4
4	1	1	0	0	0	0	+2.1	−0.2	−2.6	−4
5	1	1	0	0	0	1	+1.3	−0.2	−3.7	−11
6	1	1	2	2	2	0	−2.2	−0.3	−3.2	−7
7	1	1	2	2	2	2	+4.2	−0.3	−3.7	−2

由表5可知:组别0为HGnetV2网络实验结果,在此基础上,组别7在精度、计算量及参数量组合最优,构成的最终的DG-HGnetV2特征提取网络.

3.7.3 不同算法对比实验

通过选取Faster-RCNN、SDD、YOLOv5、YOLOv7-tiny、YOLOv8s、YOLOv10n、YOLOv11和YOLOv12n等先进算法,与所提出RL-YOLO算法进行对比,旨在验证所提算法的有效性,实验结果见表6.

同时,为了验证改进算法在轻量化研究中的优越性,针对参数量、计算量及检测精度3方面,采取不同算法进行可视化对比,对比结果见图12.

由表6和图12可知:RL-YOLO与现有先进算法相比,检测精度和检测速度更高,计算量和参数量更小,满足铁路轨道异物入侵轻量化检测研究的要求.

同时,针对雪天、阴天,低照度等复杂环境利用不同算法进行可视化对比实验,实验结果见图13.

表 6　不同算法对比实验

Tab.6　Comparison of different algorithms

改进算法	$mAP_{0.5}$/%	FPS	GFLOPs	Param
Faster-RCNN	84.5	7.1	370.2	137 099 000
SDD	84.2	10.1	62.7	26 285 000
YOLOv5	84.4	115	7.1	2 508 854
YOLOv7-tiny	82.1	138	13.2	6 016 735
YOLOv8n	85.8	152	8.1	3 006 038
YOLOv10n	79.3	129	6.5	2 265 558
YOLOv11	87.6	111.9	6.3	2 582 542
YOLOv12n	86.5	138.3	5.8	2 508 734
RL-YOLO	89.1	159	3.8	1 463 668

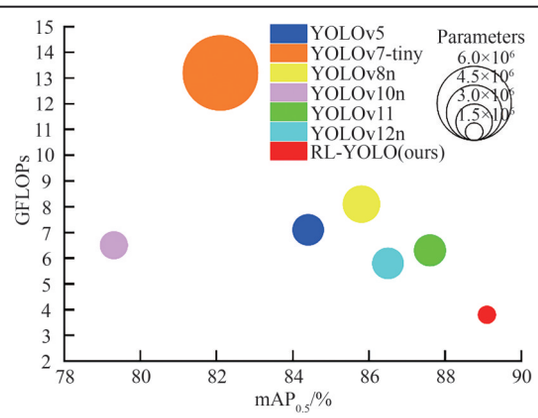


图 12　不同算法参数量,计算量及检测精度对比

Fig.12　Comparison chart of different algorithm parameters, computing power, and detection accuracy

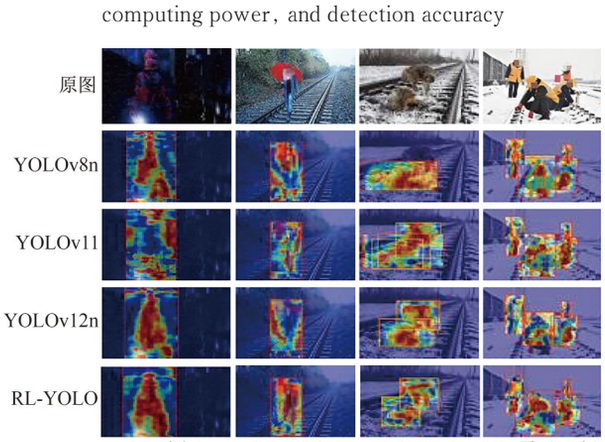


图 13　复杂环境不同算法对比实验

Fig.13　Comparison of different algorithms in complex environments

由图 13 可知:热力图突出显示了模型激活的区域,反映目标的感知能力.在复杂环境下,相较于其他先进算法,RL-YOLO 算法生成了更清晰的目标轮廓和更精确的前景激活,感知能力有着明显改善;检测时检测框的框选更为准确,减少了漏检误检的存在,在多目标情况和遮挡情况下也能精准检测;同时,热力图的颜色深浅能够更好地描述改进算法对检测目标的聚焦,体现其在复杂环境下检测的优越性.其感受野大于卷积网络,因此能够更好地捕捉整体上下文,进而提升前景激活的精度.

由表 5、表 6 及图 13 可知:相较于其他算法,RL-YOLO 算法在保持轻量化的前提下,精确检测到异物且保持了较高的置信度.证明 RL-YOLO 算法对于复杂环境下不同尺度复杂目标有着更加显著的检测优势、更为优异的感知能力及更好的泛化效果.

4　结论

1) DG-HGnet 网络在仅损失 2.3% 检测精度的前提下,实现 35.6% 的参数量压缩、17.2% 的计算量降低与 19.7 帧每秒的检测速度提升,为铁路异物入侵检测算法的轻量化部署提供核心支撑.

2) RL-YOLO 算法相较于 YOLOv12n,精度提升 2.6% 至 89.1%,实现参数量压缩 41.7%、计算量降低 34.5% 及检测速度提升 15%,有效改善了检测精度与轻量化难以兼顾的矛盾.

3) RL-YOLO 算法跨数据集精度衰减率仅为 13.8%,在复杂环境下拥有更强的目标感知能力,可有效降低遮挡、多目标场景下的漏检与误检,满足铁路轨道异物入侵实时轻量化检测的工程应用需求.

本算法在极端天气工况下或检测小尺度异物目标时,异物与背景的特征区分度会被大幅削弱,这直接导致算法表征能力及特征聚焦性能出现退化,最终造成检测精度的下降.后续研究将从以下两个方向展开深化探索:其一,构建动态特征增强模块,通过自适应特征权重校准,提升极端天气条件下目标与背景的特征差异度;其二,设计面向小目标的分层级检测架构,融合多尺度上采样与上下文语义信息聚合策略,强化小尺寸异物的特征表征能力.

参考文献(References):

[1] 新华社. 中共中央 国务院印发《交通强国建设纲要》[EB/OL]. (2019-09-19)[2025-12-13]. https://www.gov.cn/zhengce/2019-09/19/content_5431432.htm.
Xinhua News Agency. the Communist Party of China Central Committee and the State Council issued *the outline for building China's strength in transport* [EB/OL]. (2019-09-19)[2025-12-13]. https://www.gov.cn/gongbao/content/2019/content_5437132.htm. (in Chinese)

[2] 王楠, 侯涛, 牛宏侠. 多尺度特征融合的铁轨异物入侵检测研究[J]. 西安交通大学学报, 2024, 58(9): 139－153.
WANG Nan, HOU Tao, NIU Hongxia. Research on railway track foreign object intrusion detection based on multi-scale feature fusion[J]. Journal of Xi'an Jiaotong

University, 2024, 58(9): 139－153. (in Chinese)

[3] 袁花明, 薛云龙, 许剑, 等. 基于全局语义与局部特征融合的铁路异物侵限检测[J]. 计算机测量与控制, 2026, 34(1): 33－41.
YUAN Huaming, XUE Yunlong, XU Jian, et al. Railway foreign object detection based on the fusion of global semantics and local features[J]. Computer Measurement & Control, 2026, 34(1): 33－41. (in Chinese)

[4] 杜开华, 许贵阳, 白堂博. 基于特征降冗余的Vanilla-YOLOv8铁路异物侵限检测方法[J]. 北京交通大学学报, 2024, 48(5): 49－58.
DU Kaihua, XU Guiyang, BAI Tangbo. Vanilla-YOLOv8 railway foreign object intrusion detection method based on feature redundancy reduction[J]. Journal of Beijing Jiaotong University, 2024, 48(5): 49－58. (in Chinese)

[5] GAO J L, ZHU L Q, GUO B Q, et al. Multi-scale feature enhanced detection of foreign object intrusions on railways[J]. The Journal of Supercomputing, 2025, 81(6): 779.

[6] GAO S. A dense attention railway foreign object detection algorithm based on mask R-CNN[J]. IEEE Access, 2024, 12: 85761－85772.

[7] 侯涛, 李俊昶, 牛宏侠. 基于浅层特征融合的轻量级铁轨异物侵限检测研究[J]. 北京交通大学学报, 2025, 49(5): 209－220.
HOU Tao, LI Junchang, NIU Hongxia. Research on lightweight rail foreign object intrusion detection based on shallow feature fusion[J]. Journal of Beijing Jiaotong University, 2025, 49(5): 209－220. (in Chinese)

[8] GAO J L, ZHU L Q, GUO B Q, et al. Lightweight and efficient railway foreign intrusion detection with feature enhancement and optimization[J]. Measurement Science and Technology, 2026, 37(5): 056104.

[9] 曾璐, 彭东良, 江子璇, 等. 基于改进YOLOv8n的磁浮列车异物入侵检测算法[J]. 计算机工程与应用, 2025, 61(13): 165－175.
ZENG Lu, PENG Dongliang, JIANG Zixuan, et al. Magnetic levitation train foreign object intrusion DetectionAlgorithm based on improved YOLOv8n[J]. Computer Engineering and Applications, 2025, 61(13): 165－175. (in Chinese)

[10] 晏朋, 李遇鑫, 李治林, 等. 基于改进YOLOv7的高铁异物入侵检测算法[J]. 无线电工程, 2024, 54(5): 1099－1109.
YAN Peng, LI Yuxin, LI Zhilin, et al. High-speed rail foreign object intrusion detection algorithm based on improved YOLOv7[J]. Radio Engineering, 2024, 54(5): 1099－1109. (in Chinese)

[11] SUN T, GUO B Q, RUAN T, et al. RailEINet: A novel scene segmentation network for automatic train operation based on feature alignment[J]. Engineering Applications of Artificial Intelligence, 2024, 138: 109295.

[12] BAI D Y, GUO B Q, RUAN T, et al. Both objects and relationships matter: cross coupled transformer for railway scene perception [J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(7): 9812－9823.

[13] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors[EB/OL]. (2025-02-25)[2026-03-03].https://arxiv.org/abs/2502.12524.

[14] 姜香菊, 冯海照, 李涛. 基于FasterNet和YOLOv8s改进的铁路异物入侵快速检测方法[J]. 北京交通大学学报, 2024, 48(5): 39－48.
JIANG Xiangju, FENG Haizhao, LI Tao. Improved rapid detection method for foreign object intrusion on railroads based on FasterNet and YOLOv8s[J]. Journal of Beijing Jiaotong University, 2024, 48(5): 39－48. (in Chinese)

[15] WANG Z Y, DU X. YOLO-rail: an improved YOLO model for obstacle detection on railway tracks[J]. IEEE Sensors Journal, 2026, 26(5): 7528－7538.

[16] 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). Seattle, 2020: 1577－1586.

[17] ZHAO Y A, LV W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2024: 16965－16974.

[18] WANG G G, ZHAO X, DANG D H, et al. Enhancing object detection with shape-IoU and scale-space-task collaborative lightweight path aggregation [J]. Applied Sciences, 2025, 15(22): 11976－11976.

[19] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146－157.

[20] ZHENG Z H, WANG P, REN D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2022, 52(8): 8574－8586.

[21] QIU H, YANG D Y, CAO J H, et al. A lightweight track feature detection algorithm based on element multiplication and extended path aggregation networks [J]. Sensors, 2025, 25(18): 5753.

[22] SUN H, WEN Y, FENG H J, et al. Unsupervised bidirectional contrastive reconstruction and adaptive fine-grained channel attention networks for image dehazing[J]. Neural Networks, 2024, 176: 106314.