

DSA-YOLO: 一种改进YOLOv11n的铁轨表面损伤检测算法

陈猛^{1,2,3}, 邹丽^{1,2,3}, 韩彦强^{1,2,3}

(1. 大连交通大学 智能轨道工程学院; 2. 大连交通大学 辽宁省轨道交通装备焊接与可靠性重点实验室;
3. 大连交通大学 大连市区块链技术与应用重点实验室, 辽宁 大连 116028)

摘要: 铁路轨道是保障列车行驶的基本设施, 随着运营里程不断增加, 铁轨表面会出现不可逆的损伤, 严重时会影响列车运行的风险。为此, 设计了一种改进YOLOv11n的铁轨表面损伤检测算法。首先, 利用imageio库增加雪雾渲染数据, 提升模型泛化能力、训练效率; 其次, 在骨干网络中集成轻量化动态卷积, 显著加强铁轨损伤的特征提取能力; 再次, 构建SCSA机制实现骨干部分特征融合, 充分利用模型中的多层次语义信息, 优化模型检测任务的表现; 最后, 利用ADown降采样模块中细粒度的深层结构, 改善空间特征的捕获效能, 减少额外计算负担。在铁轨损伤数据集上的实验表明, DSA-YOLO算法的mAP50达到98.1%, 参数量减少到 2.29×10^6 , 优于比较算法。在公共钢铁数据集上的泛化实验表明, DSA-YOLO算法相较于比较模型在mAP50指标上提升了5%, 证明了所提模型在不同数据集上的泛化性较强。

关键词: 铁轨表面损伤检测; YOLOv11n; 数据增强; 空间-通道协同; 动态卷积

DOI: 10.11907/rjdx.251230

扫描二维码阅读全文:



中图分类号: TP317.4

文献标识码: A

文章编号: 1672-7800(2026)007-0152-09

DSA-YOLO: An Improved YOLOv11n Algorithm for Rail Surface Damage Detection

CHEN Meng^{1,2,3}, ZOU Li^{1,2,3}, HAN Yanqiang^{1,2,3}

(1. School of Railway Intelligent Engineering, Dalian Jiaotong University;

2. Liaoning Key Laboratory of Welding and Reliability of Rail Transportation Equipment, Dalian Jiaotong University;

3. Dalian Key Laboratory of Blockchain Technology and Application, Dalian Jiaotong University, Dalian 116028, China)

Abstract: Railway tracks are the basic infrastructure that ensures the operation of trains. As the operating mileage continues to increase, irreversible damage may occur on the surface of the tracks, which can increase the risk of train operation in severe cases. For this purpose, an improved YOLOv11n railway track surface damage detection algorithm was designed. Firstly, using the Imageio library to increase snow and fog rendering data can improve the model's generalization ability and training efficiency; Secondly, integrating lightweight dynamic convolution into the backbone network significantly enhances the feature extraction capability of rail damage; Once again, establish the SCSA mechanism to achieve the fusion of backbone features, fully utilize the multi-level semantic information in the model, and optimize the performance of model detection tasks; Finally, utilizing the fine-grained deep structure in the ADown downsampling module improves the capture efficiency of spatial features and reduces additional computational burden. Experiments on the rail damage dataset show that the mAP50 of the DSA-YOLO algorithm reaches 98.1%, and the parameter count is reduced to 2.29×10^6 , which is superior to the comparative algorithms. The generalization experiment on the public steel dataset showed that the DSA-YOLO algorithm improved the mAP50 index by 5% compared to the comparison model, proving that the proposed model has strong generalization on different datasets.

Key Words: rail surface damage detection; YOLOv11n; data augmentation; spatial-channel coordination; dynamic convolution

收稿日期: 2025-04-25

基金项目: 辽宁省应用基础研究计划项目(2023JH2/101300236)

作者简介: 陈猛(1999-), 男, CCF 学生会员, 大连交通大学轨道智能工程学院硕士研究生, 研究方向为智能算法及其应用; 邹丽(1980-), 女, 博士, CCF 会员, 大连交通大学轨道智能工程学院教授、硕士生导师, 研究方向为智能计算、计算机视觉、大数据处理; 韩彦强(1997-), 男, 大连交通大学轨道智能工程学院硕士研究生, 研究方向为智能算法及其应用。本文通讯作者: 邹丽。

0 引言

铁轨在列车行驶过程中承载着支撑作用,是铁路设施维护的核心部分。随着长期摩擦、震动和运营里程的增加,铁轨表面可能出现永久性损伤,例如裂纹、塌陷、剥落和锈蚀^[1]。这些损伤一旦出现,往往会迅速扩展并恶化,可能引发铁轨破裂,造成人员伤亡,因此定期排查并及时修复缺陷对铁轨维护至关重要。

目前,铁轨表面损伤的检测方法主要有:目视检查、涡流探伤、激光检测和红外热像检测等^[2]。目视检查简便,但容易遗漏微小损伤且存在主观性。涡流检测是一种无损技术,通过检测铁轨表面涡流变化来识别裂纹、坍塌和锈蚀等损伤,通常用于表面复杂的材料。激光检测通过反射和评估激光精确识别铁轨损伤,但设备成本高昂且对金属光滑度的要求较高。红外热像检测通过捕捉金属温度差异,判断铁轨表面损伤,但该方法易受温度、湿度影响,且操作成本较高。综上,单独使用这些传统检测方法,难以满足现代铁路对高效性和精确性的要求。

随着深度学习迅速发展,卷积神经网络因其卓越的特征提取能力而被广泛应用,已成为轨道表面损伤检测的主要技术手段。深度学习的检测方法可分为单阶段和双阶段检测算法。双阶段方法先生成潜在的目标区域,再对这些区域进行精确分类和位置调整。Faster R-CNN、Mask R-CNN和R-CNN是双阶段方法的典型代表,尽管在检测精度上表现优异,但计算开销较大、速度较慢^[3-5]。Xu等^[6]探讨了Faster R-CNN、Mask R-CNN在道路裂缝及铁轨损伤智能检测中的应用,发现联合训练策略能有效提升识别精度。Faghih-Roohi^[7]等设计了深度卷积神经网络(Deep Convolutional Neural Network, DCNN)自动检测铁轨损伤的方法,通过比较不同的激活函数特征和网络结构验证了该方案的有效性。

单阶段检测算法将目标定位与分类整合在单个网络中进行处理,直接预测边界框和类别概率,实现了高效的实时目标检测。YOLO和SSD是单阶段检测算法的典型代表,以高速检测和实时性强著称^[8,9]。Lu等^[10]提出改进YOLOv5s的GC-YOLO模型,在骨干网络中加入Ghost卷积模块、坐标注意力(Coordinate Attention, CA)机制,以提升绝缘子损伤检测的精度和推理速度。Wen等^[11]优化YOLOv8n模型以提升车轮胎面缺陷检测精度。首先,改进检测层来减少水渍干扰并增强小目标检测;其次,引入SiOU损失加快收敛,使算法的平均精确度达到96.95%,优于原始模型。总体而言,单阶段目标检测算法的核心优势在于检测高效、结构简洁。

虽然,上述算法已广泛应用于铁轨表面损伤检测的工程领域,但面对小型损伤尺寸和背景相似的场景时检测十分困难^[12]。此外,光照条件、雪雾天气等自然因素也可能

影响检测效果,模型容易出现漏检、误判和检测延迟问题。

为此,本文提出基于YOLOv11的改进模型DSA-YOLO算法,在保证检测速率的同时,提升了模型检测精度。同时,使用imageio库增强原始数据集中的雪雾场景,有助于提升模型在不同环境下的泛化能力和鲁棒性。具体地,从3个方面对YOLOv11模型进行优化:①为了降低误检问题,引入动态卷积(Dynamic Convolution, DyConv),通过动态调整网络层的深度和宽度来提升模型灵活性、检测性能、计算效率和准确性;②针对模型冗余问题,在骨干网络的末层嵌入设计的空间一通道协同模块(Selective Channel Spatial Attention, SCSA),以充分利用多语义信息的协同效应,提升模型的特征表达能力;③提出ADown下采样方法替代传统的卷积下采样,以降低模型复杂度,提升铁路缺陷检测效率。

1 YOLOv11模型

YOLOv11包括YOLOv11n、YOLOv11s、YOLOv11m、YOLOv11l和YOLOv11x,与YOLOv8具有相同规模的尺寸大小。YOLOv11包含输入层、骨干网络(backbone)、颈部结构(neck)、头部结构(output)和锚框机制。目前,YOLOv11是单阶段检测中效果最优的SOTA模型,通过改进骨干网络和颈部结构,显著增强了模型的特征提取和处理能力,相较于其他YOLO版本,检测灵活性和稳定性更优秀。

YOLOv11的开发源于YOLO系列的持续改进与优化,网络结构如图1所示。YOLOv11的创新之处:①引入C3k2卷积模块取代YOLOv8中C2f模块,C3k2模块优化了跨层级环节(Cross-Stage Partial, CSP)的计算性能,使用两个较小的卷积代替一个大卷积,在提升运算速率的同时保持了均衡性^[13];②在主干网络部分引入Cross Stage Partial with Spatial Attention(C2PSA)模块,还保留了原始的空间金字塔池一快速块(Spatial Pyramid Pooling-Fast, SPPF),使模型能高效地关注重要部分;③通过头部结构接收颈部传递的特征信息,以生成目标框和类别结果。

2 DSA-YOLO模型

为了解决YOLOv11算法存在的中长距离依赖、难以迅速获取底层语义信息与漏检问题,提出了DSA-YOLO模型,如图2所示。首先,改进骨干网络,借助动态卷积模块取代原始Conv层以强化局部信息捕捉能力;其次,在骨干网络最底层嵌入SCSA机制,有助于不同通道和空间位置之间的信息流动与交互,以强化模型聚焦关键区域,增强网络协同效应,减少数据信息冗余;最后,在颈部结构融入ADown降采样模块,通过注意力机制增强特征选择的精准性和自适应调整能力,以提升模型运算效率。

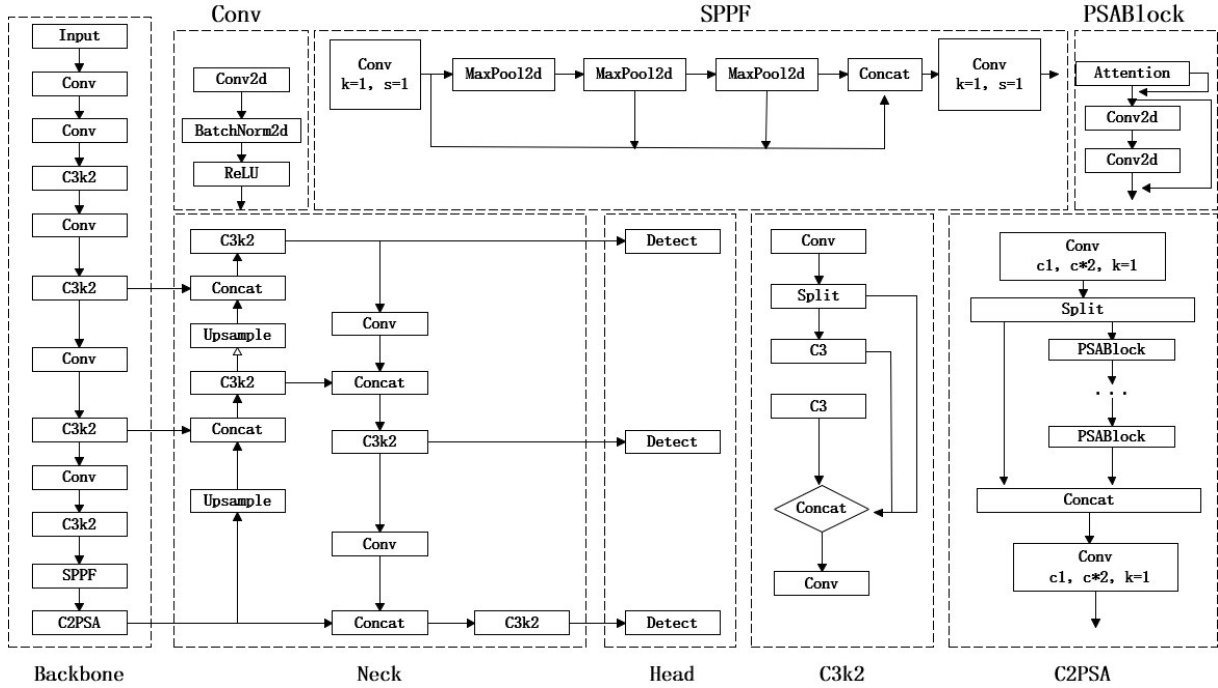


Fig. 1 YOLOv11 network architecture

图1 YOLOv11网络结构

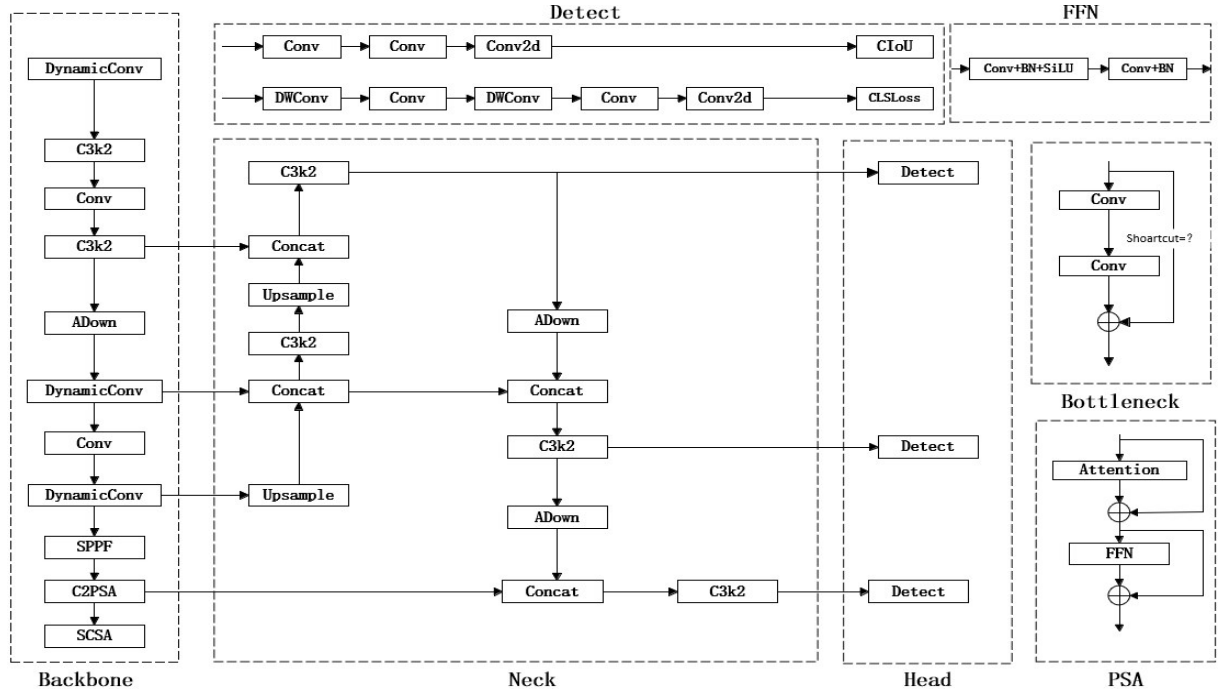


Fig. 2 DSA-YOLO model structure

图2 DSA-YOLO模型结构

2.1 Dynamic Convolution

尽管在基准模型中, C3k2模块能有效提取图像的高阶特征并实现参数共享,但堆叠多个卷积层可能会出现梯度爆炸或梯度消失问题。针对铁轨表面损伤检测中因训练不稳定易出现的误识别问题, Han等^[14]提出一种提升模型性能的动态卷积网络,在保证预训练模型参数量可控的前提下,取得了更优的性能和计算效率。

为此,本文将动态卷积神经网络应用于YOLOv11中,如

图3所示。动态卷积通过动态聚合多个卷积核来提取特征,并依据输入样本特征实时调整卷积核参数,不仅提升了卷积的灵活性,还增强了模型对多样化输入的自适应能力。

动态卷积类似于动态感知机制,基本思想是根据输入数据动态调整卷积参数,既增强了模型表达能力,又能实时适应输入。本文使用 $y = g(W^T x + b)$ 表示静态感知器, W 为权重, b 输入函数的偏置, g 为激活函数。通过聚合 k 个线性函数 $\{\bar{W}_k^T x + \bar{b}_k\}$ 来定义感知器。

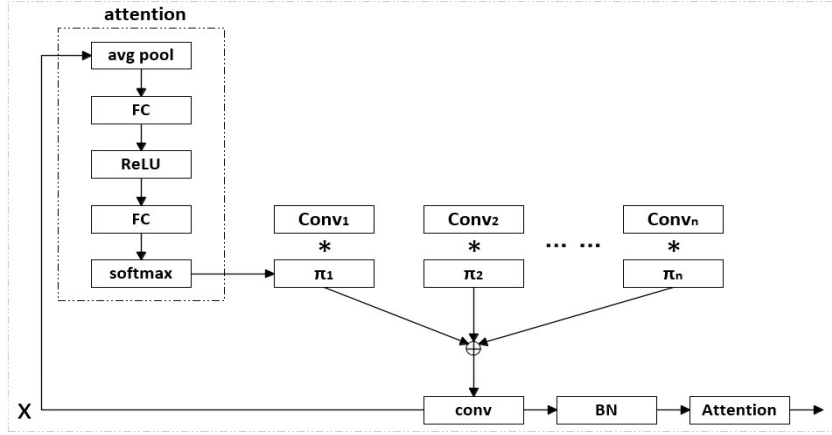


Fig. 3 Dynamic convolution architecture

图 3 Dynamic Convolution 结构

$$y = g(\bar{W}^T(x)x + \bar{b}(x)) \quad (1)$$

$$\bar{W}(x) = \sum_{k=1}^K \pi_k(x) \bar{W}_k, \quad \bar{b}(x) = \sum_{k=1}^K \pi_k(x) \bar{b}_k \quad (2)$$

$$s.t. \quad 0 \leq \pi_k(x) \leq 1, \quad \sum_{k=1}^K \pi_k(x) = 1 \quad (3)$$

在聚合函数中,动态权重 $\bar{W}(x)$ 和偏置 $\bar{b}(x)$ 经过线性变换可生成网络输出 y 。需要注意的是,注意力权重 $\{\pi_k(x)\}$ 依据输入 x 动态变化,满足 $0 \leq \pi_k(x) \leq 1$ 且总和为 1。在这种情况下,聚合模型 $\bar{W}^T(x)x + \bar{b}(x)$ 是一个非线性函数, π_k 表示第 k 个线性函数 $\{\bar{W}_k^T x + \bar{b}_k\}$ 的注意力权重,网络结构会根据不同的输入动态调整权重和偏置,以提升模型灵活性和适应性。虽然,动态卷积在提升模型性能方面优势明显,但引入了过多的自由度,可能导致模型参数发生过拟合现象,特别是在数据集特征较为简单的情况下。

2.2 SCSA

传统 YOLOv11 模型虽能准确捕获图像中的复杂细节,但存在资源消耗过多和计算效率低下的问题,无法充分利用多语义信息进行协同工作,并且不同特征之间存在语义差异。为此,本文提出一种由共享多语义空间注意力 (Multi-Semantic Spatial Attention, SMSA) 和渐进式通道自注意力 (Progressive Channel Self-Attention, PCSA) 组成的注意力机制 SCSA。

共享多语义空间注意力通过多尺寸 1D 卷积捕获关键特征进行空间特征增强。在神经网络中,该技术会显著减少模型的计算成本和复杂度。首先,对输入特征图沿宽度和高度维度进行分解;其次,对每个维度应用平均池化得到两个单向的 1D 结构 $X_H \in R^{B \times C \times W}$ 和 $X_W \in R^{B \times C \times H}$;最后,将特征集分为 k 个大小相等的独立特征 X_H, X_W , 以捕获上下文结构关系,并进一步分解得到子特征,这些子特征彼此独立,有助于高效地提取多语义空间信息

$$X_H^i = X_H[(i-1) \times \frac{C}{k}, i \times \frac{C}{k}] \quad (4)$$

$$X_W^i = X_W[(i-1) \times \frac{C}{k}, i \times \frac{C}{k}] \quad (5)$$

式中: i, k 分别表示当前索引和特征图的区域数目; C 代表特征图的通道数; X_H 为输入特征图在高度维度上的数

据; X_H^i 为第 i 个划分的高度区域, $i \in [1, k]$ 。

渐进式通道自注意力通过计算输入特征的相似度和贡献度进行通道增强,以缓解 SMSA 中不同特征的语义差异。需要注意的是,SCSA 依赖空间特征,适用于半监督学习任务。本文将 SCSA 集成到 YOLOv11 骨干网络的末层,以增强模型对不同尺寸铁轨损伤的检测能力,从而提升模型性能和泛化能力。

图 4 为 SCSA 的结构。首先,通过 SMSA 对原始特征图进行平均池化,计算高宽方向的均值以得到两个全新的特征图 x_h, x_w ; 其次,在不同尺度下分割特征图,应用 1D 卷积从关键区域提取信息并将其拼接起来进行归一化,有助于增强不同语义信息的表达;再次,空间注意力权重通过门控机制获得,PCSA 可设计下采样方法缩放特征图,防止过拟合现象并提升模型运行效率;接下来,进行归一化处理,利用 1×1 深度卷积生成查询 (Q)、键 (K) 和值 (V),以计算通道间的关系;最后,对注意力权重和特征值进行矩阵乘法,以生成加权特征图,并将其作为通道增强的输出。

尽管,SCSA 机制结合了空间和通道注意力,显著提升了特征表达能力,但也增加了模型复杂度。在噪声较多的情况下,分配不平衡的注意力可能会影响检测效果,此外过度依赖空间或通道特征也容易引发过拟合现象。

2.3 ADown

YOLOv11 采用传统卷积进行降采样,易导致微小损伤特征丢失,影响铁轨表面损伤检测性能。为了解决这一问题,引入 ADown 降采样替代传统卷积,在保留更多关键特征信息的基础上,既能有效降低特征图的空间维度,又能提升深层语义特征的建模能力^[15]。

图 5 为 ADown 的结构。首先,输入特征映射 X 通过平均池化在通道维度上可将其分为两个子特征映射 X_1, X_2 ; 其次, X_1 通过卷积核大小为 3 的传统卷积降低空间维度,减少参数复杂度, X_2 通过最大池化捕获更多特征信息,再经卷积核大小为 1 的 CBS 进行降维;最后,将两个子特征映射拼接、整合,通过多模态特征融合得到 ADown 模块的输出特征映射 Y 。

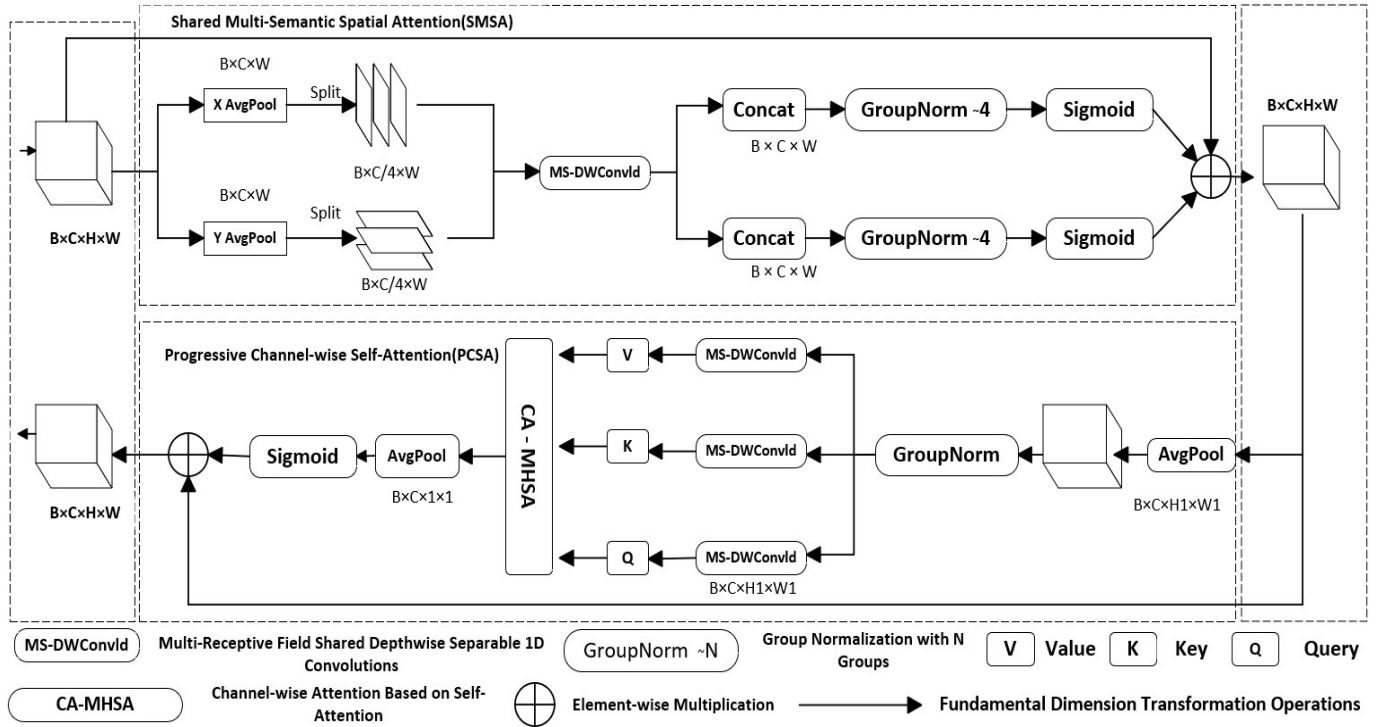


Fig. 4 SCSA architecture

图4 SCSA 结构

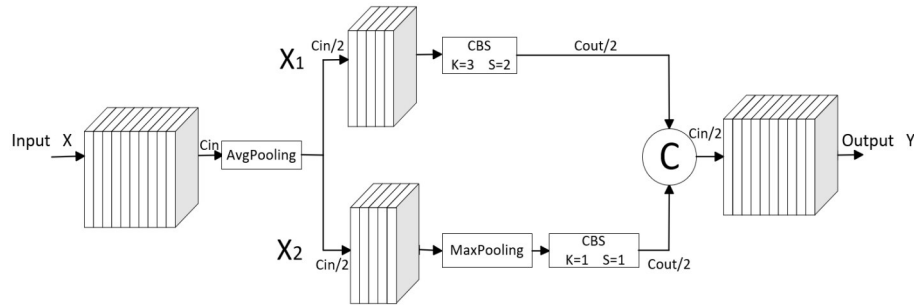


Fig. 5 ADown architecture

图5 ADown 结构

然而,由于ADown降采样引入了注意力机制、自适应门控和额外的矩阵计算,结构更复杂,增加了单次前向推理的延迟。此外,在多尺度特征融合过程中,ADown生成的特征图可能在空间分布上与传统模块存在差异,从而影响了后续的特征对齐和融合效果。

3 实验结果与分析

3.1 数据集

本文数据集由大疆 Air3 无人机(配备 4 800 万有效像素,1/1.3 英寸广角相机和 3 倍中长焦相机)在沈大铁路上拍摄所得。无人机的镜头焦距为 24 mm 的广角,高空视角可覆盖更广阔的铁轨区域,有助于捕捉不同角度的轨道损伤。

目前,该数据集已进行了严格的筛选和细致的重标记工作,原始数据集约 900 张,分辨率均为 2 560×1 920,但很

容易受到外界污渍及各种复杂环境的干扰,例如光照变化和恶劣天气(如雾、雪、雨)。因此,本文采用 imageio 库方法模拟雪雾场景来增强原始数据集,提升模型在不同应用场景下的检测能力、泛化性和鲁棒性。扩增后,本文数据集共有 3 290 张(包含轨道锈蚀、轨道剥落、轨道塌陷 3 类),按照 8:1:1 的比例划分得到验证集、测试集各 329 张。

轨道锈蚀是指易遭遇外界环境干扰引发的铁轨表面腐蚀。轨道剥落是指受到列车不断冲击,导致铁轨材质部分脱落的现象。轨道塌陷是指铁轨经过长时间的磨损导致铁轨表面出现凹坑等变形。模型使用的铁轨表面损伤数据集包括初始数据集和增强数据集,如图 6、图 7 所示。其中,图 7(a)—图 7(c)为雪景增强图像;图 7(d)—图 7(f)为雾化处理图像。

3.2 实验环境

本文操作环境基于 AutoDL 服务器站点。硬件设备为: Intel Xeon Platinum 8255C 处理器、NVIDIA GeForce

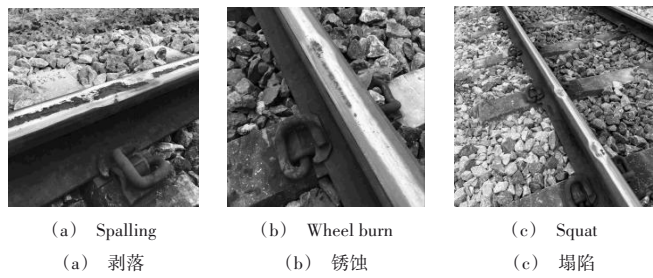


Fig. 6 Original dataset

图 6 原始数据集

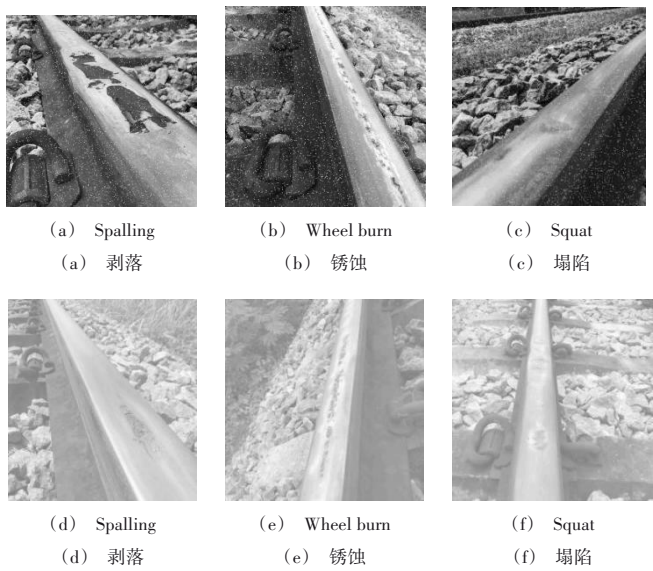


Fig. 7 Enhanced dataset

图 7 增强后的数据集

RTX 3080 显卡、10 GB 内存。软件环境为:服务器配置为 Ubuntu 16.04、CUDA11.3、Python 3.8 和 Pytorch 2.2.1 版本。模型的训练参数配置如表 1 所示。

3.3 评价指标

在铁轨损伤检测算法的性能指标中,核心指标包括模型复杂度和检测精度。其中,模型复杂度涉及计算量和存储需求,通常情况下参数数量越多,模型存储开销越大,提升模型精度往往会增加计算和存储成本。

目前,检测精度的常用评估指标包括:精确率(P)、召

Table 1 Model training parameter configuration

表 1 模型的训练参数配置

训练参数	参数值
Batch Size	32
Epoch	200
Optimiser	Adam
Momentum	0.937
Learning Rate	0.01
Weight Decay	0.000 5

回率(R)和平均精度均值(mean Average Precision, mAP)。P用来衡量模型预测结果有多少个真正的正类,R用来衡量模型能找出多少个真正的正类。

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$R = \frac{TP}{TP + FN} \tag{7}$$

式中:TP表示准确预测正类样本的数量;FP表示将实际负类样本错误预测为正类;FN表示将实际正类样本错误预测为负类的数量。

mAP是评估模型在所有类别上整体表现的综合指标。

$$AP = \int_0^1 P(R) dR \tag{8}$$

$$mAP = \frac{1}{n} \int_{i=1}^n AP_i \tag{9}$$

式中:n表示类别数量;AP表示特定类别在不同置信度阈值下精度—召回率曲线下的面积。

3.4 消融实验

本文基于 YOLOv11n 基准模型进行消融实验,以评估动态卷积、SCSA 机制和 ADown 下采样 3 种优化方案的有效性,并验证表 2 中列出的 7 种排列组合方案的性能。为了确保训练结果有效,需统一系统硬件和参数,通过模型参数量、精确率(P)、召回率(R)、平均精度均值(mAP@50)及每秒帧数(Frames Per Second, FPS)评价模型性能,实验结果如表 2 所示。

由表 2 可知,Yolov11n 算法的性能,铁轨表面损伤检测任务的平均准确度为 95.8%。实验表明,在基准模型中分别融合 DyConv、SCSA 和 Adown 等 3 种优化策略,均能有效提升模型识别准确度。其中,使用 DyConv 替代骨干网络

Table 2 Ablation experiment

表 2 消融实验

Model	Parameter/M	P/%	R/%	mAP@50/%	FPS/Frames/s	GFLOPs
YOLOv11n	2.58	95.9	92.5	95.8	143.64	6.3
YOLOv11n+DyConv	2.50	98.1	94.7	97.9	166.51	5.5
YOLOv11n+SCSA	2.62	97.9	91.6	96.3	160.09	6.4
YOLOv11n+ADown	2.34	98.5	93.7	97.1	158.49	5.8
YOLOv11n+SCSA+ADown	2.38	97.4	91.7	96.8	156.52	5.9
YOLOv11n+DyConv+SCSA	2.53	98.0	94.4	97.4	164.77	5.6
YOLOv11n+DyConv+ADown	2.26	98.0	94.2	97.4	153.06	5.0
DSA-YOLO	2.29	98.7	94.4	98.1	169.33	5.1

中的原始卷积,使模型的 mAP 提升至 97.9%,但参数规模未显著增加。由此可知,DyConv 机制能在不显著增加参数量或计算复杂度的前提下,提升模型检测性能。

基于 DyConv 模块进行 SCSA 扩展后,模型的 mAP 提升至 97.4%,证明了 SCSA 通过更有效地结合全局上下文信息,提升了模型的目标识别和全面覆盖能力。采用轻量化机制 ADown 优化颈部网络后,显著降低了计算成本,并将平均准确度提升至 98.1%。综上,改进后的 YOLOv11n 算法在检测性能、FPS 方面相较于改进前提升显著,证明了所提模型的可行性。

图 8 为 YOLOv11n 与 DSA-YOLO 算法的 mAP 和召回率曲线。由图 8(a)可知,经过 100 轮训练后,改进模型的 mAP 明显高于初始模型,并在 150 轮后趋于稳定,最终收敛。由图 8(b)显示,召回率在后期稳定,模型达到了最优状态,优化后的模型 mAP 达到 98.1%。

3.5 比较实验

本文在相同实验条件下,将 DSA-YOLO 算法与其他检测方法进行比较,实验结果如表 3 所示。由此可知,改进后的模型 mAP 达到 98.1%,远超现有算法。尤其在复杂环境中,DSA-YOLO 相较于现有算法能更有效地应对这些挑战。具体地,YOLOv9t 在小目标和恶劣场景下的检测性能不及其他算法。DSA-YOLO 算法相较于 YOLOv8n、YOLOv10n,在 mAP 上分别提升 2.6%、3.1%,参数量分别减少 0.72 M、0.41 M。

DSA-YOLO 相较于 YOLOv5s 而言,在召回率、mAP 上提升显著,GFLOPs 值相较于 YOLOv5s 增加了 0.7,导致模型在铁轨表面检测任务中需要更多计算资源。YOLOv9t 的速度、精度虽然较高,但消耗了较高的计算资源和内存。一方面,DSA-YOLO 算法中集成了 DyConv 模块、SCSA 和 ADown 下采样,实现了精度与运算效率的最佳平衡;另一方面,Faster R-CNN 通过多层卷积和金字塔模型(如 FPN)在一定程度上解决了尺度变化问题,但在后续检测阶段仍存在计算冗余。

此外,DSA-YOLO 相较于 EfficientNet,各项指标均有显著提升。RT-DETR 是 DETR(Facebook 团队提出)算法的改进版,尽管 mAP 达到了 97.1%,但推理速度明显低于

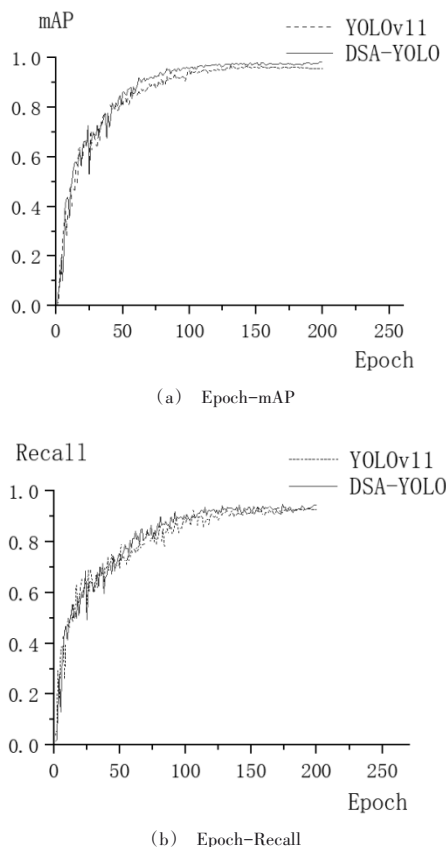


Fig. 8 mAP and recall curves of YOLOv11n and DSA-YOLO algorithm

图 8 YOLOv11n、DSA-YOLO 算法的 mAP 与召回率曲线

DSA-YOLO,影响了目标检测系统的实时响应能力。CE-YOLO 算法更轻量化,但 mAP、FPS 低于 DSA-YOLO。综上,改进后的算法在各项评价指标上优势明显,证明了 DSA-YOLO 模型相较于现有方法具有很强的竞争力。

3.6 算法验证

为了更直观地验证 DSA-YOLO 算法的检测性能,将 YOLOv10n、YOLOv11n 及 DSA-YOLO 算法在铁轨表面损伤识别中的效果进行可视化,实验结果如图 9 所示。其中,从上到下依次为 YOLOv10n、YOLOv11n 及 DSA-YOLO 算法的时间结果;图 9(b)为雾化场景,图 9(c)、图 9(d)为雪

Table 3 Comparative experiment

表 3 比较实验

Model	Parameters/M	Precision/%	Recall/%	mAP50/%	FPS/Frames/s	GFLOPs
Faster R-CNN ^[16]	136.70	50.6	61.7	51.44	42.65	200.8
SSD ^[17]	11.90	53.2	44.8	52.38	37.90	29.6
YOLOv5s	2.18	98.7	93.4	96.50	198.66	5.8
YOLOv8n	3.01	96.9	92.9	95.50	133.10	8.1
YOLOv9t	61.10	97.9	93.5	96.40	115.68	9.6
YOLOv10n	2.70	93.0	89.1	95.00	145.78	8.2
YOLOv11n	2.58	95.9	92.5	95.80	143.64	6.3
EfficientNet ^[18]	8.00	52.6	48.3	51.40	18.60	10.3
RT-DETR ^[19]	19.80	98.2	94.6	97.10	77.20	56.9
CE-YOLO ^[20]	2.18	98.7	93.4	96.50	143.35	5.8
DSA-YOLO	2.29	98.7	94.4	98.10	169.33	5.1

景增强效果。

由图 9(a)、图 9(b)可见,DSA-YOLO 算法具有更高的置信度,目标特征更明显;由图 9(c)、图 9(d)可见,原始算法无法完全检测到铁轨表面损伤,DSA-YOLO 算法却能准确定位轨道上的所有损伤,在不同复杂条件下的鲁棒性、迁移能力更强。综上,DSA-YOLO 算法能有效解决其他基准模型存在的小目标漏检问题。

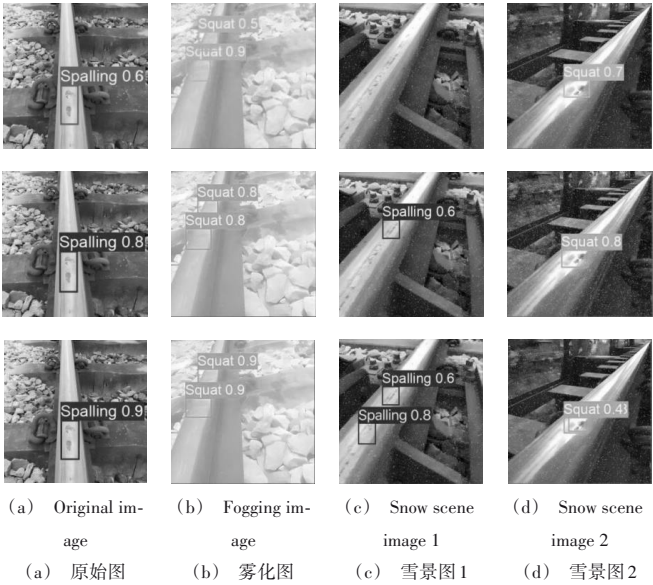


图 9 可视化展示

3.7 SCSA 热力图可视化分析

深度学习中的热力图(红色通常代表较高的数值和强度)可快速识别数据中具有特定趋势的区域,便于深入理解模型的内部机制,可提升模型对输入数据的敏感性。

为此,本文利用 Grad-CAM 方法生成 YOLOv11n 和改进 SCSA(空间-通道协同注意力)机制的热力图,如图 10 所示。图 10(a)为 YOLOv11n 的热力图,可以看到剥落缺陷并未完全被热力覆盖,出现了遗漏部分。图 10(b)为改进 SCSA 机制的热力图,剥落缺陷区域的热力覆盖范围更广。图 10(c)、图 10(d)分别为经过数据增强的雾图和雪图,通过注意力机制优化后,模型对缺陷区域的关注度更高。综上,改进的 YOLOv11n 相较于基准模型的目标检测热力值更高,检测效果更好。

3.8 泛化能力验证

为了验证本文算法的泛化性能,选用 Roboflow 平台公开的钢铁数据集(<https://roboflow.com/>)进行实验。该数据集包含 1 800 张图片,被分为正常和损坏两种类别。数据集的增强方式、划分方法、参数设置、评价指标与前文实验保持一致,实验结果如表 4 所示。由此可知,DSA-YOLO 算法的 mAP50(%) 值达到 86.6%,相较于基准模型提升 5%,参数量显著减少,表明了 DSA-YOLO 算法在目标检测任务中性能更优,在不同数据集上的泛化性能更好(见图 11)。

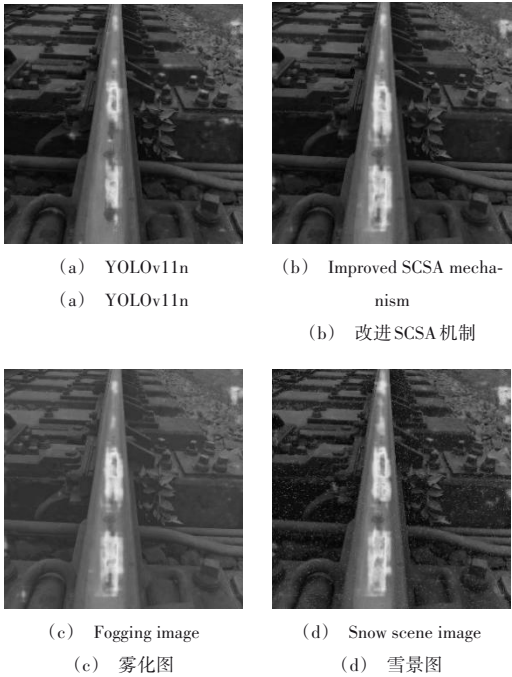


图 10 热力图

由图 11 可见,改进后的模型相较于原 YOLOv11n 模型,不仅能检测到更多目标,还具备更高的置信度,在铁轨损伤检测任务中的优势明显。

表 4 钢铁数据集的检测结果比较

Model	Parameters/M	Precision/%	Recall/%	mAP50/%	mAP50-95/%
YOLOv11n	2.58	77.3	98.5	81.6	61.9
DSA-YOLO	2.29	78.1	92.6	86.6	68.1

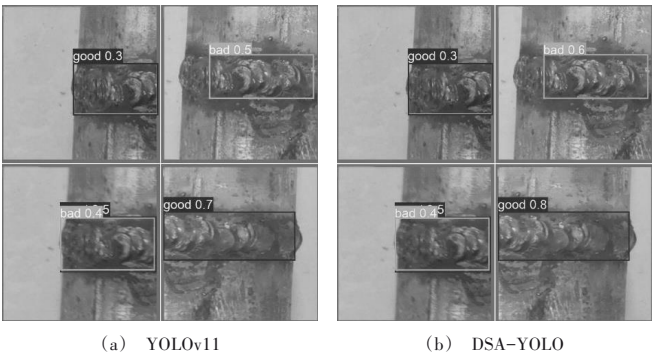


图 11 泛化性能比较

4 结语

本文提出一种基于 YOLOv11n 改进的模型 DSA-YOLO 算法,以解决传统模型在复杂场景中出现的漏检、误检和数据冗余问题。首先,采用 imageio 库增强雪雾化的数据来模拟真实场景,加强模型容错性;其次,提出一种轻量化

动态卷积替代骨干网络中的原始卷积,在不增加卷积层堆叠或通道数的情况下,提升模型表达能力;再次,设计空间—通道协同模块,充分利用多语义信息的协同效用,缓解语义差异并降低数据冗余;最后,在颈部网络中融入ADown降采样机制,强化不同特征层之间的信息交互与融合效果,有效降低了模型计算量。

实验表明,DSA-YOLO算法在铁轨表面损伤检测中的mAP50达到98.1%,参数规模降低至 2.29×10^6 ,证明了所提模型在提升精度的同时,显著减少了计算资源消耗,能在实际的铁轨表面损伤检测任务中得到有效应用。

未来,将持续优化模型性能,以更好地满足铁轨损伤检测的多样化需求,并进一步探索其在移动设备上的应用与部署。

参考文献:

- [1] WANG D, WU X, JIANG L. Automatic defect description of railway track line image based on dense captioning[J]. *Sensors*, 2022, 22(17): 6419.
- [2] LI X, GU J, JIN W, et al. Detection of rail defects using NDT methods[J]. *Sensors*, 2023, 23(10): 4627.
- [3] HUANG W X, JIANG N. Defect detection of fasteners based on improved faster R-CNN algorithm[J]. *Software Guide*, 2023, 22(5): 190–197.
黄午祥,江南. 基于改进Faster R-CNN算法的扣件缺陷检测[J]. *软件导刊*, 2023, 22(5): 190–197.
- [4] ZHANG B, ZHOU J, WANG F, et al. Touchscreen glass defect detection and recognition based on Mask R-CNN[J]. *Software Guide*, 2019, 18(2): 64–67, 71.
张博,周军,王芳,等. 基于Mask R-CNN的触摸屏玻璃疵病检测与识别[J]. *软件导刊*, 2019, 18(2): 64–67, 71.
- [5] NASUTION M D, AL-KHOWARIZMI, RAHMAT R F, et al. The role of faster R-CNN algorithm in the Internet of Things to detect mask wearing: the endemic preparations[J]. *International Journal of Electronics & Telecommunications*, 2023, 69(4): 172–176.
- [6] XU X Y, ZHAO M, SHI P X, et al. Crack detection and comparison study based on faster R-CNN and Mask R-CNN[J]. *Sensors*, 2022, 22: 1215.
- [7] FAGHIH-ROOHI S, HAJIZADEH S, NUNEZ A, et al. Deep convolutional neural networks for detection of rail surface defects[C]// *Proceedings of the International Joint Conference on Neural Networks*, 2016: 2584–2589.
- [8] ZHAO C G, XU D, ZHANG L F, et al. Rail surface defect detection based on image enhancement and improved YOLOX[J]. *Electronics*, 2023, 12(12): 2672.
- [9] ZHAO X H, LIU Q H. Improved SSD model for pavement disease image detection system[J]. *Software Guide*, 2020, 19(11): 217–220.
赵雪寒,刘庆华. 改进SSD模型的路面病害图像检测系统[J]. *软件导刊*, 2020, 19(11): 217–220.
- [10] LU D, RAO Z Q, DING B, et al. Research on defect detection method of railway transmission line insulators based on GC-YOLO[J]. *IEEE Access*, 2023, 11: 102635–102642.
- [11] WEN Y, GAO X R, LUO L, et al. Improved YOLOv8-based target precision detection algorithm for train wheel tread defects[J]. *Sensors*, 2024, 24(11): 3477.
- [12] DUAN Y X, QIU S, JIN W Q, et al. High-speed rail tunnel panoramic inspection image recognition technology based on improved YOLOv5[J]. *Sensors*, 2023, 23(13): 5986.
- [13] ZHANG D, CAO K, HAN K, et al. PAL-YOLOv8: a lightweight algorithm for insulator defect detection[J]. *Electronics*, 2024, 13(17): 3500.
- [14] HAN K, WANG Y H, GUO J Y, et al. Parameters are all you need for large-scale visual pretraining of mobile networks[C]// *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 17–21.
- [15] WANG C Y, YE H I, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[C]// *European Conference on Computer Vision*, 2025: 1–21.
- [16] CHEN J Y, HUANG J M. Deep learning (fast R-CNN)-based evaluation of rail surface defects[J]. *Applied Sciences*, 2024, 14(5): 1874.
- [17] MIN Y Z, WANG Z W, LIU Y, et al. FS-RSDD: few-shot rail surface defect detection with prototype learning[J]. *Sensors*, 2023, 23(18): 7894.
- [18] TAN M X, LE Q C. EfficientNet: rethinking model scaling for convolutional neural networks[C]// *International Conference on Machine Learning*, 2019: 6105–6114.
- [19] ZHAO Y, LYU W, XU S, et al. DETRs beat YOLOs on real-time object detection[C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2024: 16965–16974.
- [20] LI Y, YAN H, LI D, et al. Robust miner detection in challenging underground environments: an improved YOLOv11 approach[J]. *Applied Sciences*, 2024, 14(24): 11700.

(责任编辑:刘嘉文)