

多尺度自适应的轻量化轨道病害检测方法

郝宜君¹, 郭新伟², 陈 帅², 黄贺祥³, 宋霄罡³

(1. 中铁一局集团有限公司, 陕西 西安 710054; 2. 中铁一局集团电务工程有限公司, 陕西 西安 710025;
3. 西安理工大学计算机科学与工程学院, 陕西 西安 710048)

摘要: 针对轨道病害检测中背景复杂、病害微小密集导致识别精度不足等问题, 提出多尺度自适应的轻量化轨道病害检测模型。首先, 设计动态特征增强模块, 通过分组卷积实现多尺度特征解耦融合, 并结合全局注意力机制强化关键特征, 提升模型对微小病害的细节捕捉能力; 其次, 引入大核选择卷积注意力模块, 利用可变形卷积构建空间选择机制与自适应感受野调节, 动态调整卷积核形状与范围, 有效抑制背景干扰, 增强病害定位精度; 最后, 将检测头回归损失函数替换为基于最小点距度量的MPDIoU损失, 通过优化边界框中心点及顶点距离计算并引入更强几何约束, 提升检测框拟合精度。实验表明, 相比原始YOLOv8n, 模型在轨道表面病害数据集上mAP50和mAP50:95分别提升2.3%和4.2%; 在轨道扣件数据集上分别提升3.3%和1.8%。结果表明, 该方法显著提升了轨道病害检测性能并具有工程实用价值。

关键词: 轨道病害检测; YOLOv8; 动态特征增强; 自适应感受野; 注意力机制

DOI: 10.11907/rjdk.251476

扫描二维码阅读全文: 

中图分类号: TP391.41; U216.3

文献标识码: A

文章编号: 1672-7800(XXXX)0XX-0001-09

Multi-scale Adaptive Lightweight Rail Defects Detection Method

XI Yijun¹, GUO Xinwei², CHEN Shuai², HUANG Hexiang³, SONG Xiaogang³

(1. China Railway First Bureau Group Co., Ltd., Xi'an 710054, China;
2. China Railway First Group Corporation Electrical Engineering Co., Ltd., Xi'an 710025, China;
3. School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: To address the issue of insufficient recognition accuracy caused by complex backgrounds and tiny, dense defects in rail defect detection, this paper proposes a multi-scale adaptive lightweight rail defect detection model. First, a Dynamic Feature Enhancement Module is designed to achieve multi-scale feature decoupling and fusion via group convolution, combined with a global attention mechanism to enhance key features and improve the model's ability to capture fine details of tiny defects. Second, a Large Kernel Selective Convolutional Attention Module is introduced, which constructs a spatial selection mechanism and adaptive receptive field adjustment using deformable convolution. This dynamically adjusts the shape and range of convolution kernels, effectively suppressing background interference and enhancing defect localization accuracy. Finally, the regression loss function of the detection head is replaced with the minimum point distance metric-based MPDIoU loss. By optimizing the calculation of bounding box center points and vertex distances and introducing stronger geometric constraints, the accuracy of detection box fitting is improved. Experiments show that compared to the original YOLOv8n, the model achieves improvements of 2.3% and 4.2% in mAP50 and mAP50:95, respectively, on the rail surface defect dataset, and 3.3% and 1.8% on the rail fastener dataset. The results demonstrate that this method significantly enhances rail defect detection performance and possesses practical engineering value.

Key Words: rail defect detection; YOLOv8; dynamic feature enhancement; adaptive receptive field; attention mechanism

收稿日期: 2025-07-01

基金项目: 国家自然科学基金面上项目(52372418); 陕西省创新能力支撑计划科技创新团队项目(2025RS-CXTD-006)

作者简介: 郝宜君(1973-), 男, 中铁一局集团有限公司正高级工程师, 研究方向为轨道维修、机电安装; 郭新伟(1968-), 男, 中铁一局集团电务工程有限公司高级工程师, 研究方向为智能交通; 陈帅(1990-), 男, 中铁一局集团电务工程有限公司高级工程师, 研究方向为轨道及地下工程; 黄贺祥(2000-), 男, 西安理工大学计算机科学与工程学院硕士研究生, 研究方向为计算机视觉; 宋霄罡(1987-), 男, 博士, 西安理工大学计算机科学与工程学院副教授、博士生导师, 研究方向为轨道交通智能运维、计算机视觉。本文通讯作者: 宋霄罡。

0 引言

铁路运输能否安全稳定运行,与轨道状态密切相关。轨道发生病害后,不仅会影响线路质量,还可能带来运行安全隐患,因而开展及时有效的病害检测十分必要。轨道病害可分为两类:表面病害和结构病害。表面病害主要包括轨道腐蚀、裂纹、脱落和凹坑等,结构病害主要体现为扣件螺栓状态异常、弹条丢失断裂与位置偏移等^[1-2]。随着人工智能特别是深度学习和计算机视觉技术的飞速发展,深度学习和计算机视觉技术发展较快,相关方法在工业检测中的应用不断深入,为铁路病害检测提供了新的技术路径。轨道和轨道扣件都是铁路线路的重要组成部分,其状态好坏直接影响列车运行的安全性和稳定性。传统的人工巡检方式不仅效率低、成本高,而且受制于人眼疲劳和环境因素,难以实现全天候、全覆盖、高精度的状态监控。因此,采用智能化方法对轨道及扣件进行自动化、实时化缺陷检测,已成为现代铁路运维体系升级的迫切需求和必然趋势。

在此背景下,深度学习驱动的计算机视觉方法逐渐成为轨道病害检测中的重要技术路径。与传统方式相比,这些方法能够直接从大量轨道图像中学习特征,并在复杂场景下完成目标识别,实现对轨道表面磨损、裂纹、剥落等缺陷和扣件缺失、松动、断裂等异常状态的检测,显著提升了检测效率和精度,为轨道设备的预防性维护提供了强有力的数据支撑。近年来,研究者们提出了多种基于深度学习的检测方法。例如,Guo 等^[3-4]分别基于 Transformer 和 Mask R-CNN 开展了轨道表面缺陷检测工作,取得了较好的检测效果。然而,Transformer 模型的庞大参数量以及 Mask R-CNN 固有的两阶段检测流程,都导致了高昂的计算复杂度,使得模型推理速度受限,难以满足实时检测需求,尤其在部署于资源受限的移动或嵌入式设备(如轨道巡检车、无人机)时面临严峻挑战。针对轨道扣件检测中普遍存在的背景复杂、干扰物多(如道砟、轨枕纹理)等难题,Chandran 等^[5]提出了一种创新思路,将深度学习模型与经典图像处理技术相结合,旨在提升模型在杂乱背景下的鲁棒性。与此同时,Gibert 等^[6]则采用了多任务学习框架,通过整合多个相关的检测子任务(如扣件定位、状态分类),期望利用任务间的相关性以提高神经网络在识别异常扣件时的整体表现。虽然这两种方法在特定场景下提升了性能,但前者采用的 ResNet-50 骨干网络计算量巨大,对硬件资源要求较高;后者复杂的多任务框架则显著增加了模型设计、训练过程和超参数调优的复杂性,其实际部署的便捷性和效率仍有待优化。

相比之下,以速度和效率见长的单阶段目标检测算法,尤其是 YOLO(You Only Look Once)系列模型^[7-9],因其“单次推理即可完成目标定位与分类”的简洁高效特性,在

铁路运行环境下的实时病害检测任务中表现尤为突出。本文选择 YOLOv8 作为基础模型,主要基于其优秀的轻量化特性及其在发布时所代表的速度—精度平衡优势。虽然更新的 YOLOv9/v10 等版本已发布,但本文提出的核心改进模块和损失函数旨在解决轨道病害特有的挑战,其设计思路 and 有效性同样适用于后续的 YOLO 版本或其他检测框架。选择 YOLOv8n 作为轻量高效的起点进行针对性改进,旨在最直接地解决实际工程需求。目前,研究者们已广泛尝试基于不同版本的 YOLO 框架进行针对性改进,以提升其在铁路场景下的病害检测性能。Feng 等^[10]对 YOLOv5s 模型进行了多项重要改进:引入全局注意力机制以聚焦关键区域、优化特征融合网络以增强多尺度特征表达能力、改进卷积结构以提升效率、采用 WIoU 损失函数以获得更优的边界框回归,以及优化锚框设计以适应轨道病害的特定形态。这些改进措施协同作用,显著提升了模型在轨道表面病害检测上的性能,准确率(Precision)和召回率(Recall)分别提高了 3.3% 和 3.9%,为实时高精度检测提供了有力方案。高嘉琳等^[11]则聚焦于改进 YOLOv4,通过优化网络结构和增强特征提取能力,在铁路扣件检测任务中实现了精度和效率的双重提升。然而,他们的方法在处理高度复杂的背景干扰时,对扣件的区分能力仍有提升空间。为了追求更极致的轻量化,Fu 等^[12]提出了基于 MobileNet 的 YOLOv4 算法,用轻量级网络替代原有的 CSP-Darknet53,以降低模型参数量和计算复杂度,同时增强对细微裂纹、锈蚀等特征的提取能力。但是,MobileNet 在轻量化的同时,也会带来一定的特征表达不足,特别是在深层语义建模和大感受野获取方面表现有限。当铁路场景背景较为复杂、目标本身细节较弱时,模型可能难以准确捕捉裂纹走向、分支等局部信息,从而影响检测效果。

综上所述,当前轨道病害自动化检测面临的核心挑战在于:如何在高背景复杂度(如道砟纹理、油污、光照不均)干扰下,鲁棒且精准地识别与定位微小(如毫米级裂纹)且密集分布的病害目标。尽管以 YOLO 系列为代表的高效检测器及其改进方案在速度和效率上取得了显著进展,但在应对上述核心挑战时仍存在明显局限。①细节捕捉不足:现有方法的特征融合机制或轻量化骨干在深层特征中极易丢失微小病害的关键细节信息,导致对小尺寸目标的漏检率较高;②背景抑制有限:现有注意力机制(如全局注意力)或结构设计难以动态区分高度相似背景噪声与真实病害,对复杂背景干扰的鲁棒性不足,易导致误检;③微小目标定位不准:传统 IoU 损失(如 CIoU)对微小目标框的中心点偏移和顶点位置偏差敏感度不足,难以满足病害精确测量的工程需求。

针对上述关键瓶颈,本文提出一种多尺度自适应的轻量化轨道病害检测模型,其核心创新点在于协同优化特征提取、背景抑制与定位精度。①动态特征增强模块(DFE):通过多分支异构卷积结构结合全局注意力

(GCA),实现多尺度特征解耦与关键细节强化,显著提升模型对微小密集病害的表征能力,弥补了现有方法在深层细节保留上的不足;②大核选择卷积注意力(LSK):引入可变形卷积构建的自适应空间选择机制,赋予模型动态调节感受野形状与范围的能力,聚焦病害区域同时强力抑制复杂背景干扰,克服了固定核注意力在背景噪声抑制上的局限性;③MPDIoU 损失函数:通过直接优化预测框与真值框顶点间的最小点距,强化边界框的几何约束,显著提升微小目标及不规则形状病害的定位精度,解决了传统 IoU 损失在此类场景下定位松散的问题。该模型在不显著增加计算负担的前提下,有效解决了复杂背景下微小密集轨道病害的高精度检测难题。具体而言,本文主要创新贡献体现在以下 3 个方面:

(1)为了增强模型对微小病害特征的识别和提取能力,对颈部网络进行了创新性改进。具体而言,在颈部网络中引入了动态特征增强(Dynamic Feature Enhancement, DFE)模块,用以替代原有的 C2f 模块。这一改进利用分组卷积和注意力选择机制,提升了网络对微小轨道病害的检测能力。

(2)为了解决轨道背景复杂导致的病害识别难题,对颈部网络进行了精心调整,引入了一种新颖的大核选择卷积(Large Selective Kernel, LSK)注意力模块。这一模块的加入,有效缓解了背景杂乱对特征学习的影响,提升了网络在复杂背景下对轨道病害的识别能力。

(3)为了提高轨道病害检测中边界框回归的准确性,对网络检测头的损失函数进行优化,引入了 MPDIoU 损失函数^[13]。这种损失函数通过减少预测边界框与实际标注边界框左上角和右下角点之间的距离,直接衡量两者的相似度。这种改进确保了在训练过程中,预测框能够更精确地对准目标框,从而提升了病害检测精度。

1 多尺度自适应的轻量化轨道病害检测模型

YOLOv8 在目标检测领域取得了显著成就,但在处理轨道病害检测这类复杂场景时,模型容易遇到漏检和误检问题。YOLOv8 在骨干网络、特征融合(Neck)和检测头(Head)设计上进行了多项优化,如更高效地跨阶段部分连接(CSP)结构、路径聚合网络(PAN)的增强版本,以及动态标签分配策略,使其在通用目标检测基准上表现优异。然而,将其直接应用于轨道病害检测,仍需克服三大关键挑战:①微小与密集病害识别难:轨道表面裂纹、细小凹坑等目标尺寸极小,有效特征信息稀少,极易被复杂背景淹没或在特征金字塔下采样过程中丢失;②复杂背景干扰强:道砟碎石的不规则纹理、钢轨表面的油污反光、光照条件的剧烈变化(如隧道内外)等,构成大量与病害特征相似的噪声,导致模型难以聚焦于真正的病害区域,产生漏检或误检;③边界框定位精度要求高:病害尺寸和位置直接影

响维修决策,传统的 IoU 系列损失函数在处理微小目标或长宽比异常的病害时,定位精度仍有提升空间。针对上述挑战,本文提出多尺度自适应的轻量化轨道病害检测模型,旨在在不显著增加模型复杂度的前提下,显著提升其在轨道病害检测任务中的性能。其网络结构如图 1 所示。

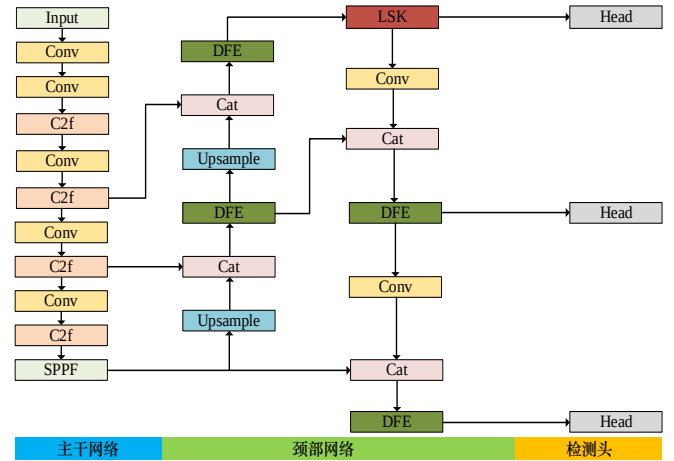


Fig. 1 Multi-scale adaptive lightweight rail defects detection model

图 1 多尺度自适应的轻量化轨道病害检测模型

1.1 动态特征增强模块

轨道病害检测面临的核心难题之一在于目标的尺寸极其微小(如细小的裂纹、剥落等),且这些病害常常密集分布在线路的局部区域。这种特性给基于卷积神经网络(CNN)的检测方法带来了显著挑战:一方面,在标准 CNN 的层级结构中,随着特征图通过池化或步长卷积进行下采样,微小目标本就稀少的特征信息极易被削弱甚至彻底丢失;另一方面,传统的单一尺寸卷积核因其感受野固定,难以有效适应轨道病害(尤其是形态各异的细长裂纹)多变的空问特征表达。为了克服上述难点,本文提出的特征增强模块(DFE)摒弃了原始 C2f 模块采用的单一特征提取路径,转而创新性地设计了一套多分支异构卷积结构。区别于 RFB 模块通过固定尺寸的膨胀卷积模拟人类视觉感受野^[13],或 ASF 模块采用静态加权融合多尺度特征^[14],本文提出的 DFE 模块创新性地融合了 3 个关键技术。①动态特征选择机制:通过分组卷积路径提取具有不同空间特性的特征(如捕捉纵向/横向裂纹的 $1 \times 3/3 \times 1$ 卷积、增强局部区域的 3×3 卷积),并结合注意力机制(GCA)根据输入图像内容动态地选择和强化对当前微小病害最有效的特征组合;②全局上下文感知增强:独立的 GCA 分支显式构建全局特征关联图,精准聚焦于稀疏的关键病害区域,有效抑制复杂背景干扰;③残差特征保护机制:保留原始输入特征的直连路径,确保微小目标的底层细节信息不被复杂变换覆盖丢失。这种特征解耦、动态选择与全局感知相结合的三重协同机制,是 DFE 模块的核心创新,使其能够更灵活、更精准地适应轨道病害微小、形态多样、密集分布且背景复杂的特点,从而显著提升模型对微小病害的细节捕捉能力和抗干扰能力。DFE 模块结构如图 2 所示。

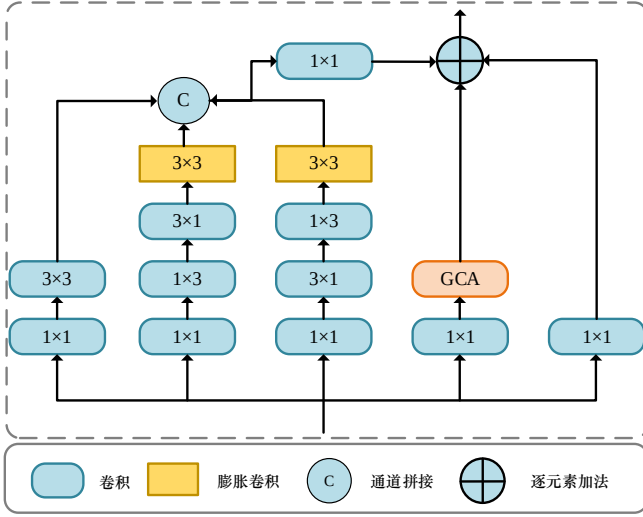


Fig. 2 DFM module structure

图2 DFM模块结构

在DFE模块中,取 F 为输入特征,每个分支首先采用 1×1 卷积优化输入特征图的通道数;接着,前3个分支通过 1×3 、 3×1 和 3×3 卷积核增强特征表达,并在第二、第三分支中加入膨胀卷积以提升特征感知;最后,这3个分支的特征图在通道维度上进行融合,如式(1)一式(4)所示。

$$F_1 = f_{conv}^{3 \times 3}(f_{conv}^{1 \times 1}(F)) \quad (1)$$

$$F_2 = f_{conv}^{3 \times 3}(f_{conv}^{3 \times 1}(f_{conv}^{1 \times 3}(f_{conv}^{1 \times 1}(F)))) \quad (2)$$

$$F_3 = f_{conv}^{3 \times 3}(f_{conv}^{1 \times 3}(f_{conv}^{3 \times 1}(f_{conv}^{1 \times 1}(F)))) \quad (3)$$

$$Y = C_{at}(F_1, F_2, F_3) \oplus f_{conv}^{1 \times 1}(F) \quad (4)$$

其中, F_1 、 F_2 和 F_3 表示第一、第二分和第三分支经过卷积后输出, $f_{conv}^{3 \times 3}(\cdot)$ 、 $f_{conv}^{1 \times 3}(\cdot)$ 、 $f_{conv}^{3 \times 1}(\cdot)$ 和 $f_{conv}^{1 \times 1}(\cdot)$ 分别表示卷积核大小为 3×3 、 1×3 、 3×1 和 1×1 的卷积操作, F 表示DFE模块输入特征表示, Y 表示3个分支进行特征融合的结果, $C_{at}(\cdot)$ 表示特征映射拼接, $\oplus$ 表示特征映射的元素相加。

第四分支采用全局注意力机制模块以提高模型对病害特征的敏感度,GCA模块网络结构如图3所示。输入特征首先通过3个独立的卷积层以增强网络的非线性能力,并对生成的 $H \times W \times C$ 的3个特征进行重塑,生成 Q 、 K 和 V 特征矩阵,如式(5)一式(7)所示。

$$Q = (R_{shape}(f_{conv}^{1 \times 1}(f_{conv}^{1 \times 1}(F))))^T \quad (5)$$

$$K = (R_{shape}(f_{conv}^{1 \times 1}(f_{conv}^{1 \times 1}(F)))) \quad (6)$$

$$V = (R_{shape}(f_{conv}^{1 \times 1}(f_{conv}^{1 \times 1}(F)))) \quad (7)$$

其中, Q 、 K 和 V 为输入特征经过卷积和特征重塑的结果, $R_{shape}(\cdot)$ 表示重塑运算, $f_{conv}^{1 \times 1}(\cdot)$ 表示卷积核大小为 1×1 的卷积操作, T 表示矩阵转置运算, F 表示DFE模块输入特征表示。

通过计算 Q 特征矩阵和 K 特征矩阵之间的相似度以获得特征注意力权重,这些权重表明了特征图中每个位置

的重要性。接着,将这些权重应用于特征矩阵 V ,从而得到全局注意力特征,如式(8)所示。

$$G = R_{shape}(V \otimes (S_{softmax}(Q \otimes K))) \quad (8)$$

其中, G 表示通过注意力权重生成的全局注意力特征, $\otimes$ 表示矩阵乘法, $S_{softmax}(\cdot)$ 表示激活函数。

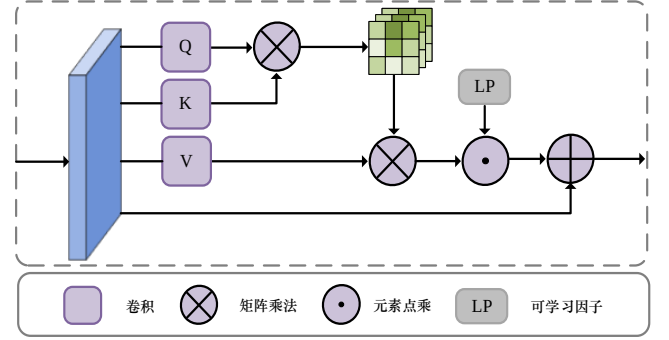


Fig. 3 GCA module structure

图3 GCA模块结构

通过一个可学习的参数(LP)进行调节,参数在训练过程中自动校准,将该全局注意力特征与原始特征相结合,以生成最终输出特征,如式(9)所示。该机制让模型更专注于输入数据的关键信息,从而提升病害识别能力。

$$W = F + (L_N \odot G) \quad (9)$$

其中, W 表示DFE模块的输出, L_N 表示可学习的全局控制参数, $\odot$ 表示元素点乘。

DFE模块第五个分支采用残差网络结构,旨在捕捉并保留小尺寸物体的关键特征,减少因尺度差异导致的信息丢失,确保了检测结果的准确性和可靠性。

1.2 大核选择卷积注意力模块

在目标检测任务中,不同类别的目标往往对上下文信息的需求存在显著差异。为了有效适应这种差异性,本文在YOLOv8模型的颈部(Neck)顶端集成了LSKNe所提出的自适应大核模块(LSK Module)^[15]。这一改进的动机在于,复杂的轨道背景环境,例如密集的道砟纹理、油污残留以及光照不均等因素构成了导致误检(False Positives)和漏检(False Negatives)的另一重要干扰源。为了更好地适应轨道病害检测的复杂背景和微小目标的特点,对LSK模块进行了适配性改进。具体而言,将第二个分支的Depth-wise卷积替换为可变形深度大核卷积。可变形卷积通过动态调整卷积核的形状和范围,能够更好地捕捉病害目标的特征,同时抑制复杂背景干扰。LSK模块的核心创新思想在于赋予模型动态调整感受野尺寸和形状的强大能力。不同于传统卷积固定不变的感受野,LSK模块能够根据输入的具体内容进行自适应学习,使其能够精准地聚焦于与目标高度相关的关键区域,同时有效地抑制无关背景噪声干扰,从而显著提升模型在复杂背景下的鲁棒性和识别精度。LSK模块的结构示意图如图4所示。

该模块以双分支深度卷积构建特征提取网络,第一分支采用标准深度卷积保留空间维度,第二分支使用可变形

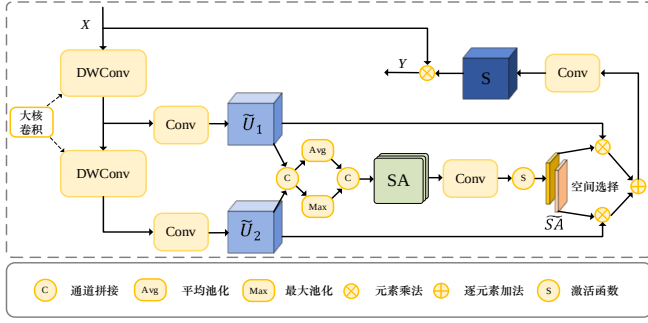


Fig. 4 LSK module structure

图4 LSK模块结构

深度大核卷积。其中,可变形卷积的偏移量通过独立的 1×1 卷积层学习得到。具体而言,对于第二分支,首先通过一个卷积核为 1×1 的卷积层生成偏移量场 $\Delta \in \mathbb{R}^{H \times W \times 2N}$ (N 为卷积核数量,2表示二维空间偏移),该偏移量场与输入特征共同参与可变形大核卷积运算,随后通过 1×1 卷积压缩通道并融合信息,将两分支输出特征拼接,最终通过通道级平均池化与最大池化联合捕获空间上下文关联。相关公式如式(10)一式(13)所示。

$$\tilde{U}_1 = f_{conv}^{1 \times 1}(D_{W1}(X)) \quad (10)$$

$$\tilde{U}_2 = f_{conv}^{1 \times 1}(D_{W2}(D_{W1}(X), \Delta)) \quad (11)$$

$$S_{avg} = P_{avg}(C_{at}(\tilde{U}_1, \tilde{U}_2)) \quad (12)$$

$$S_{max} = P_{max}(C_{at}(\tilde{U}_1, \tilde{U}_2)) \quad (13)$$

其中, D_{W1} 为常规深度卷积, D_{W2} 为带可学习偏移量 Δ 的可变形深度大核卷积, P_{avg} 、 P_{max} 分别表示平均池化和最大池化操作, S_{avg} 、 S_{max} 分别表示平均池化和最大池化得到的空间特征。

将空间池化特征进行拼接,使用卷积层将池化特征转换为多个空间注意力特征图,将Sigmoid激活函数应用于每一个特征图,获得每个解耦的大卷积核所对应的独立空间选择掩膜,如式(14)所示。

$$\tilde{S}_{A_i} = S_{ofimax}(F_{2 \rightarrow N}([S_{avg}, S_{max}])) \quad (14)$$

其中, $F_{2 \rightarrow N}(\cdot)$ 表示将池化特征转多空间注意力图的卷积, $\tilde{S}_{A_i}$ 表示解耦的大卷积核所对应的空间掩膜。

将解耦后的大卷积核序列的特征与对应的空间选择掩膜进行加权处理,并通过卷积层融合以获得注意力特征。最终,LSK模块的输出通过输入特征与注意力特征的逐元素点乘获得,如式(15)和式(16)所示。

$$K = \sum_{i=1}^N \tilde{U}_i \odot \tilde{S}_{A_i} \quad (15)$$

$$Y = K \odot X \quad (16)$$

其中, K 表示将每个空间选择掩膜应用于对应的大卷积核序列并进行特征拼接后的注意力特征, X 表示输入特征, Y 为输入特征经过LSK模块后的输出。

1.3 MPDIoU 损失函数

在轨道病害智能视觉检测系统中,实现高精度的边界框定位是后续进行病害尺寸精确测量与维修位置精准标

定的基石。目前,主流模型YOLOv8默认采用CIoU(Complete Intersection over Union)损失函数,通过融合边界框重叠面积、中心点距离以及宽高比3项几何因素进行联合优化,在一定程度上提升了定位性能^[16]。然而,当面对轨道场景中普遍存在的像素级微小病害目标(如细微裂纹或点状剥落)或具有极端长宽比的条状、线状病害(如纵向长裂缝)时,CIoU的定位能力仍呈现出明显不足。其根本原因在于:一方面,对于微小目标,其预测框与真实框之间微小的绝对中心点偏移距离和宽高差异,在数值计算上导致CIoU损失中相应的惩罚项相对微弱,模型难以感知并修正此类细微偏差;另一方面,CIoU对边界框形状偏差的惩罚主要依赖于长宽比例差异的计算,未能直接建模并强制优化预测框与真实框之间对应4个顶点的精确空间位置关系,这可能导致预测框虽然在整体比例上接近真实目标,但在边界贴合度上,尤其是在边角区域,存在不够紧密的“松散包围”现象。

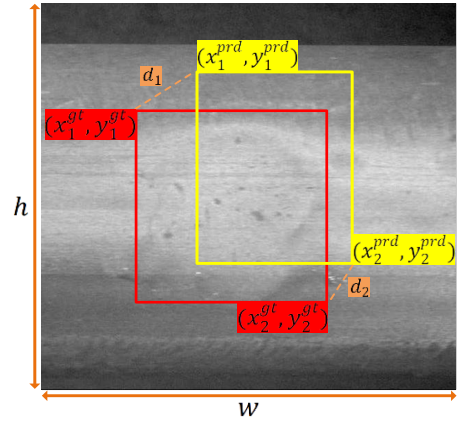


Fig. 5 MPDIoU loss function for border regression

图5 边框回归的MPDIoU损失函数

针对上述由CIoU损失函数对边界框形变敏感性不足以及锚框适应性局限所引发的轨道多态病害(如裂纹、剥落、掉块等)定位精度偏差问题,本文提出采用MPDIoU(Minimum Point Distance based IoU)损失函数作为解决方案。MPDIoU的核心创新在于摒弃了间接的比例约束,转而直接联合优化预测边界框与真实边界框4个对应顶点之间的欧几里得距离总和。这种机制将边界框视为由4个关键顶点定义的空间几何实体,通过最小化所有顶点之间的空间距离,能够更全面、更直接地反映两个边界框在空间覆盖范围(重叠度)和几何形态(形状贴合度)上的综合偏差。这种基于顶点距离最小化的优化策略,有效克服了传统IoU变体(如CIoU)在匹配具有相同宽高比例但绝对尺寸不同的目标时所面临的缺陷。同时,通过显式关联4个顶点的空间位置,显著增强模型对不规则形状和极端长宽比病害(如细长裂纹)的泛化适应能力。值得注意的是,MPDIoU在实现更强几何约束的同时,其计算过程相对简洁高效,未引入复杂的额外项。实际应用验证表明,在轨道多元形态病害检测任务中,集成MPDIoU损失的

YOLOv8 模型显著提升了边界框定位精度,并且该损失函数展现出更优的优化特性,有效促进了模型在训练过程中的收敛速度和稳定性。MPDIoU 损失函数回归框图如图 5 所示,损失函数如式(17)、(18)和(19)所示。

$$(d_1)^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (17)$$

$$(d_2)^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (18)$$

$$M_{PDIoU} = IoU - \frac{(d_1)^2}{h^2 + w^2} - \frac{(d_2)^2}{h^2 + w^2} \quad (19)$$

其中, (x_1^{gt}, y_1^{gt}) 、 (x_2^{gt}, y_2^{gt}) 、 (x_1^{prd}, y_1^{prd}) 、 (x_2^{prd}, y_2^{prd}) 分别表示真实边界框和预测边界框的左上角和右下角点坐标。 d_1 、 d_2 真实与预测边界框的左上、右下角点间距, h 、 w 分别表示检测图片的高度和宽度, IoU 表示真实边界框与预测边界框之间的面积交并比值。

2 实验与结果分析

2.1 数据集与评价指标

本文实验使用了两个自建轨道病害数据集,分别为轨道表面病害数据集和轨道扣件病害数据集。其中,轨道表面病害数据集包含 3 050 张图像,涉及 3 种病害,分别为轨道表面裂纹(crack)、轨道表面凹陷(Squat)和轨道表面剥落(Spalling)。轨道扣件状态数据集包含 2 230 张图像,共覆盖 5 种状态类型,即正常 E 型扣件(fastener)、正常 W 型扣件(fastener-2)、破损 E 型扣件(f-broken)、破损 W 型扣件(f2-broken)和扣件丢失(missing)。两个数据集均按照 8:2 的比例划分为训练集和验证集。

考虑到轨道病害检测中微小目标较多,并且在实际场景中还存在遮挡、姿态变化等情况,训练过程中采用了 YOLOv8 的标准数据增强策略,从而高模型的泛化能力。具体而言,几何变换主要包括随机缩放、随机平移、随机旋转和随机剪切。其中,图像宽高独立缩放的典型比例范围为 0.5~1.5,水平和垂直方向的平移比例通常控制在 $\pm 20\%$ 以内,随机旋转角度范围一般为 $\pm 15^\circ$,水平和垂直剪切因子通常设为 $\pm 15^\circ$,还以 50% 的概率对图像进行随机水平翻转。

模型性能评价方面,本文主要采用平均精度均值(mAP)作为指标,用来衡量模型在轨道表面缺陷检测和轨道扣件病害识别任务中的检测效果。具体而言,采用了 mAP50 和 mAP50:95 两个指标, mAP50 表示在 IoU 阈值为 0.5 条件下得到的 mAP 结果, mAP50:95 则表示在 IoU 阈值从 0.5 到 0.95 范围内计算得到的平均 mAP 值,如式(20)、(21)(22)和(23)所示。对比单一阈值下的评价方式, mAP50:95 能够更全面地反映模型在不同定位精度要求下的检测性能,因而更适合作为目标检测任务中的综合评价指标。

$$P = \frac{TP}{TP + FP} \quad (20)$$

$$R = \frac{TP}{TP + FN} \quad (21)$$

$$A_p = \int_0^1 P(R) dr \quad (22)$$

$$m_{AP} = \frac{\sum_{i=1}^N A P_i}{N} \quad (23)$$

其中, TP 表示模型正确预测为正例的样本数量, FP 表示模型错误地将负例预测为正例的样本数量, FN 表示模型错误地将正例预测为负例的样本数量。 P 表示精确度,为模型预测为正例的样本中实际为正例的比例, R 表示召回率,为所有实际正例中被模型正确预测为正例的比例。 A_p 为 P 、 R 和坐标轴围成的面积,评估每个类别的预测精度, m_{AP} 用于评估所有类别的平均精度。

在考虑模型的轻量化设计时,特别关注两个关键指标:模型的参数量(Parameters)、计算复杂度(GFLOPs)和推理速度(FPS)。参数量的减少可以降低模型的存储需求和内存占用。同时,计算量的优化直接影响到模型推理速度。FPS 作为实际推理速度的衡量标准,直接反映了模型在目标设备上处理输入数据的实时性能。通过优化这些指标,确保模型能够在这些设备上高效运行。

2.2 实验环境与参数配置

实验选用 Ubuntu 22.04.4 LTS 操作系统的计算平台,并配备了 Intel Core i9-14900K 处理器、32 GB 的系统内存,以及 24 GB 显存的 Nvidia GeForce RTX3090 Ti 显卡。在深度学习领域,所用环境为 CUDA 11.3 和 Python 3.8.10,实验采用了 PyTorch 1.11 作为深度学习框架。实验具体参数如表 1 所示。

Table 1 Specific experimental parameters

表 1 实验具体参数

实验参数	参数值
训练轮次	600
优化器	SGD
批量大小	16
图像尺寸	640
学习率	0.01

2.3 对比实验

2.3.1 综合对比实验

为了验证本文方法的有效性和优越性,将其与多种先进的目标检测算法,包括 Convnext、Cst-Yolo、Lite-Yolo、Ca-Auto、Etdnet、Gold-Yolo 以及 YOLOv8 系列等进行了对比实验。实验结果如表 2、表 3 所示。表中加粗为最优结果,下划线为次优结果,“N/A”表示无相应数据。

在轨道表面病害数据集和轨道扣件病害数据集上的实验结果表明,本文模型在关键的 mAP50 和 mAP50:95 指标上均超越了 Convnext、Cst-Yolo、Lite-Yolo、Ca-Auto、Etdnet、Gold-Yolo、Yolov5、Yolov7、Yolov12 以及 Yolov8 系列在内的多种算法,实现了最佳性能。综合考量计算负担、模型大小和推理速度,本文模型展现出卓越的效率平衡。它在保持显著轻量化的同时,其高推理速度亦跻身最快模型

Table 2 Comparative experiment on the track surface disease detection dataset

表 2 轨道表面病害检测数据集对比试验						
模型	发表期刊	mAP50 /%	mAP50:95/%	计算复 杂度/G	参数 量/M	FPS
Convnext ^[17]	CVPR 2022	73.9	44.8	189.10	48.10	12.98
Cst-Yolo ^[18]	ICIP 2024	73.6	43.5	235.40	61.90	8.39
Lite-Yolo ^[17]	TIM 2024	73.0	39.3	9.70	3.76	68.03
Ca-Auto ^[20]	TIM 2023	72.3	34.8	82.79	36.81	16.80
Etdnet ^[21]	TIM 2023	74.0	42.7	23.91	7.01	33.55
Gold-Yolo ^[22]	NeurIPS 2023	75.8	42.7	12.10	5.60	67.03
Yolov8n ^[23]	ADICS 2024	74.9	41.5	8.20	3.04	68.66
Yolov8s ^[23]	ADICS 2024	75.8	45.4	28.80	11.14	20.35
Yolov5s	N/A	74.0	44.6	13.20	7.50	66.57
yolov7-tiny ^[24]	CVPR 2023	76.3	39.7	13.00	6.02	66.66
Yolov12s ^[25]	N/A	76.7	46.1	21.50	9.26	36.19
Ours	N/A	77.2	45.7	11.10	3.82	67.45

Table 3 Comparative experiment on the dataset of track fastener disease detection

表 3 轨道扣件病害检测数据集对比实验			
模型	发表期刊	mAP50/%	mAP50:95/%
Convnext ^[17]	CVPR 2022	96.2	72.8
Cst-Yolo ^[18]	ICIP 2024	93.8	69.4
Lite-Yolo ^[19]	TIM 2024	94.0	70.7
Ca-Auto ^[20]	TIM 2023	94.8	71.4
Etdnet ^[21]	TIM 2023	96.2	72.7
Gold-Yolo ^[22]	NeurIPS 2023	94.6	72.8
Yolov8n ^[23]	ADICS 2024	93.5	73.8
Yolov8s ^[23]	ADICS 2024	95.1	74.2
Yolov5s	N/A	93.9	70.5
Yolov7-tiny ^[24]	CVPR 2023	94.1	65.9
Yolov12s ^[25]	N/A	95.3	74.8
Ours	N/A	96.8	75.6

行列,体现了优异的实时处理潜力,这证明了本文方法是有效且优越的。

2.3.2 不同注意力机制对比实验

为验证本文采用的 LSK 注意力模块的有效性,对其他 4 种不同的注意力机制模块在轨道表面病害数据集上进行了实验,实验结果如表 4 所示。

Table 4 Comparison test of attention modules

表 4 注意力模块对比试验

注意力模块	mAP50/%	mAP50:95/%	计算复杂度/%	参数量/M
CBAM ^[26]	74.2	42.0	8.2	3.01
CA ^[27]	74.7	41.4	8.2	3.01
GAM	73.9	42.4	9.4	3.10
AGCA ^[28]	75.3	41.8	8.2	3.01
LSKblock ^[14]	75.9	43.4	8.2	3.02

从实验结果看,LSK 注意力模块在 mAP50 和 mAP50:95 两个关键指标上的表现都要优于 CBAM、CA、GAM 和 AGCA,分别达到了 75.9% 和 43.4%。也即在这组对比实验中,LSK 的整体效果更好。此外,LSK 在计算复杂度和参数数量上与 CBAM、AGCA 相差不大,基本保持在同一水平,但性能上有一定提升。这说明该模块在改善模型表现的

同时,并没有明显增加额外的计算负担。

2.3.3 不同损失函数对比实验

为了验证 MPDIoU 损失函数在提升模型性能方面的有效性,对模型在轨道表面病害数据集上进行了损失函数对比实验。表 5 展示了不同损失函数在 mAP50 和 mAP50:95 这两个关键指标上的表现。

Table 5 Loss function comparison experiment

表 5 损失函数对比试验

损失函数	mAP50/%	mAP50:95/%
GIOU ^[29]	76.2	44.6
DIOU ^[15]	76.4	43.8
CIOU ^[15]	76.1	45.4
MPDIoU ^[13]	77.2	45.7

实验数据表明,MPDIoU 损失函数在 mAP50 和 mAP50:95 这两个关键性能指标上均优于 GIOU、DIOU 和 CIOU 损失函数。与改进前的 CIOU 损失函数相比,MPDIoU 在 mAP50 指标上提升了 1.1%,在 mAP50:95 指标上提升了 0.3%。这一性能提升证明了 MPDIoU 损失函数在轨道病害检测领域的综合性能优势。

2.3.4 消融实验

本文基于 YOLOv8n 模型,采用轨道表面病害数据集进行消融实验,逐步引入改进模块以验证策略有效性及其对评价指标的影响(见表 6),深入探究性能提升机制并确保实际应用可靠性。

Table 6 Ablation experiment

表 6 消融实验

实验	DFE	LSK	MPDIoU	mAP50/%	mAP50:95/%
1	—	—	—	74.9	41.5
2	√	—	—	75.8	44.7
3	—	√	—	75.5	43.1
4	—	—	√	75.3	42.0
5	√	√	—	76.1	45.4
6	√	√	√	77.2	45.7

本文在 YOLOv8n 模型基础上进行了逐步改进。实验 2 通过引入 DFE 模块中的多分支结构,增强了模型对病害特征的提取能力,使 mAP50 和 mAP50:95 分别提高了 1.9% 和 3.2%。实验 3 仅加入 LSK 模块后,mAP50 和 mAP50:95 分别提升了 0.6% 和 1.6%。实验 4 在不改变其他结构的情况下采用 MPDIoU 损失函数,此两个指标分别提升了 0.4% 和 0.5%。实验 5 和实验 6 进一步在实验 2 的设置上加入了 LSK 注意力模块与 MPDIoU 损失函数,两个评价指标继续提升,其中 mAP50 和 mAP50:95 分别增加了 1.4% 和 1.0%。从整体结果看,上述改进对模型在轨道病害检测任务中的性能提升有效,也说明所采用的组合策略具有合理性。

为进一步探究动态特征增强模块(DFE)中各组件的贡献,对 DFE 模块进行分层拆解,在轨道表面病害数据集上验证其多分支结构的有效性,共设置 4 种方案。①方案 1:基础模型(YOLOv8n);②方案 2:仅保留 DFE 中的分组卷

积分支;③方案3:只保留GCA模块;④方案4:完整DFE模块。实验结果如表7所示。实验中保持其他结构(LSK模块、MPDIoU损失)不变,以隔离DFE组件的影响。

Table 7 DFE module disassembly ablation experiment

表7 DFE模块拆解消融实验

实验	分组卷积	GCA	mAP50/%	mAP50:95/%
1	—	—	74.9	41.5
2	√	—	75.2	42.1
3	—	√	75.4	42.9
4	√	√	75.8	44.7

消融实验结果表明,DFE模块的分组卷积分支和GCA模块均对提升轨道表面病害检测性能具有积极作用。单独使用分组卷积分支或GCA模块(方案2、方案3)相比基础模型(方案1)均能带来性能提升,其中GCA模块的贡献更为显著。更重要的是,完整的DFE模块取得了最佳性能,其提升幅度显著优于两个分支单独作用时的效果之和。这清晰地验证了DFE模块中分组卷积与GCA模块存在显著的协同效应,两者共同作用能更有效地融合多尺度特征并捕获全局上下文信息,从而显著提升模型检测精度和定位。实验充分证明了DFE模块多分支结构设计的有效性及其组件的互补性。

2.3.5 实验结果可视化分析

为验证所提出模块的有效性,本文利用热力图对其激活模式进行了可视化分析。如图6所示(左侧为原始标注图像,中间为YOLOv8n检测热力图,右侧为本文模型检测热力图),引入该模块后生成的热力图在关键语义区域表现出更强烈的聚焦。具体而言,这些区域的激活值分布更集中、边界更分明,而无关背景区域的响应则被有效抑制。这种增强的清晰度与确定性直接反映了模块引导模型更准确地捕捉和学习输入数据中关键信息的能力。

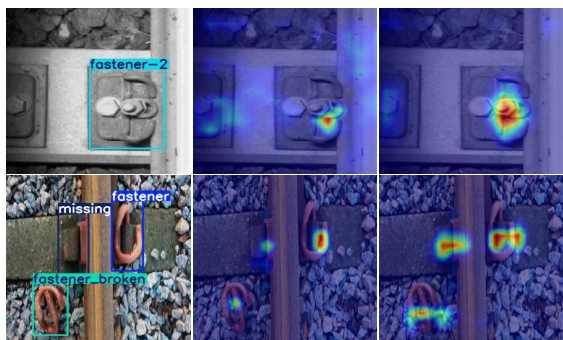


Fig. 6 Heat map of model results

图6 模型结果热力图

为了更直观地展示本文模型相较于YOLOv8n在性能上的提升,在两个关键数据集—轨道扣件病害数据集和轨道表面病害数据集的验证集上进行了深入实验。如图7、图8所示(左侧为原始标注图像,中间为YOLOv8n检测结果,右侧为本文模型检测结果)。

本文模型在轨道扣件病害检测数据集上的置信度有

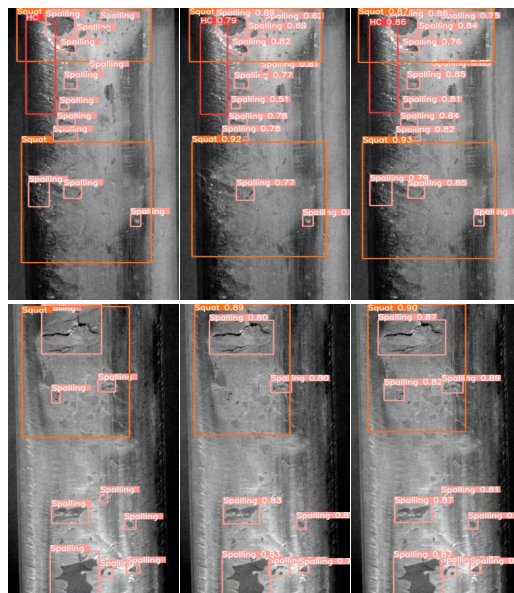


Fig. 7 Visualization experiment of rail surface disease detection dataset

图7 轨道表面病害检测数据集可视化实验

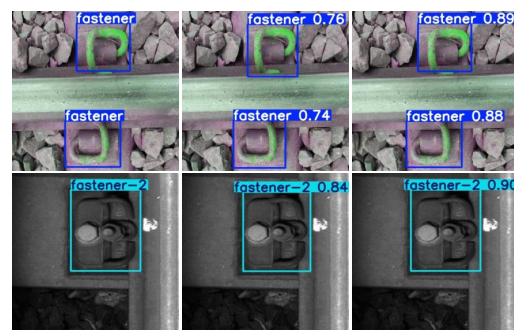


Fig. 8 Visualization experiment of rail fastener disease detection dataset

图8 轨道扣件病害检测数据集可视化实验

显著提升。同时,在轨道表面病害检测数据集上,本文模型不仅在置信度上有所增强,而且在检测精度上也实现了对YOLOv8n的超越。这些成果充分证实了本文模型在定位和识别轨道病害方面的准确性得到了显著提升。

3 结语

基于对轨道病害检测中复杂背景干扰、微小密集目标识别困难和定位精度不足等核心挑战的深入研究,本文系统地提出并验证了多尺度自适应轻量化检测模型。研究通过3方面的创新实现性能突破:首先,提出动态特征增强模块,利用分组卷积实现多尺度特征解耦与融合,结合全局注意力机制强化关键病害细节,显著提升了模型对微小裂纹、剥落等目标的细节捕捉能力;其次,设计大核选择卷积注意力模块,通过可变形深度卷积构建空间选择机制与自适应感受野调节,动态优化卷积核形状与范围,有效抑制道砟纹理、油污等复杂背景干扰,增强病害区域的聚焦能力;最后,引入MPDIoU损失函数替代原回归损失,

通过直接最小化预测框与真实框顶点间距离并强化几何约束,改善了微小目标及不规则病害的边界框定位精度。实验结果表明,改进后的模型在轨道表面病害数据集和轨道扣件病害检测数据集上的性能优于原始模型。

综上,本文研究不仅解决了轨道病害检测中的关键瓶颈问题,为智能化运维提供了高精度、轻量化的实用解决方案,其提出的多尺度特征动态增强、自适应感受野调节及几何约束强化的设计思路,对资源受限环境下的工业视觉检测任务亦具有普适参考价值。需指出,模型在浓雾、强降雪等极端恶劣天气条件下可能存在性能下降,同时在资源受限的边缘设备上实现实时高帧率检测仍需进一步的轻量化优化。未来工作将聚焦于极端天气场景的鲁棒性优化和边缘设备的实时性强化,进一步提升工程适用性。

参考文献:

- [1] REN H Q. Research on the knowledge graph of typical diseases in heavy duty railway track operation and maintenance [J]. Comprehensive Transportation, 2024, 46(7): 95-100, 179.
任海强. 重载铁路轨道运维典型病害知识图谱研究[J]. 综合运输, 2024, 46(7): 95-100, 179.
- [2] HUANG S F, CAO G R, LI B, et al. A review of the application of machine vision in the detection of track fasteners [J]. Rail Transit Materials, 2024, 3(3): 45-53.
黄石甫, 曹广如, 李斌, 等. 机器视觉在轨道扣件检测上的应用综述[J]. 轨道交通材料, 2024, 3(3): 45-53.
- [3] GUO F, QIAN Y, RIZOS D, et al. Automatic rail surface defects inspection based on mask R-CNN [J]. Transportation Research Record, 2021 (11): 036119812110190.
- [4] GUO F, LIU J, QIAN Y, et al. Rail surface defect detection using a transformer-based network [J]. Journal of Industrial Information Integration, 2024, 38: 100584.
- [5] CHANDRAN P, ASBER J, THIERY F, et al. An investigation of railway fastener detection using image processing and augmented deep learning [J]. Sustainability, 2021, 13(21): 12051.
- [6] GIBERT X, PATEL V M, CHELLAPPA R. Deep multitask learning for railway track inspection [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(1): 153-164.
- [7] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information [DB/OL]. <https://arxiv.org/abs/2402.13616>.
- [8] REDMON J, FARHADI A. YOLOv3: an incremental improvement [DB/OL]. <https://arxiv.org/abs/1804.02767>.
- [9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [10] FENG H, JIANG Z, XIE F, et al. Automatic fastener classification and defect detection in vision-based railway inspection systems [J]. IEEE Transactions on Instrumentation and Measurement, 2014, 63(4): 877-888.
- [11] GAO J L, BAI T B, YAO D C, et al. Railway fastener detection based on improved YOLOv4 algorithm [J]. Science, Technology and Engineering, 2022, 22(7): 2872-2877.
高嘉琳, 白堂博, 姚德臣, 等. 基于改进 YOLOv4 算法的铁路扣件检测[J]. 科学技术与工程, 2022, 22(7): 2872-2877.
- [12] FU J, CHEN X, LYU Z. Rail fastener status detection based on MobileNet-YOLOv4 [J]. Electronics, 2022, 11(22): 3677.
- [13] LIU S, HUANG D, WANG Y. Receptive field block net for accurate and fast object detection [C]//Proceedings of the 15th European Conference on Computer Vision (ECCV), 2018: 404-419.
- [14] DING X, GUO Y, DING G, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV) 2019: 1911-1920.
- [15] LI Y, HOU Q, ZHENG Z, et al. Large selective kernel network for remote sensing object detection [C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 16748-16759.
- [16] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 12993-13000.
- [17] LIU Z, MAO H, WU C Y, et al. A ConvNet for the 2020s [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11976-11986.
- [18] KANG M, TING C M, TING F F, et al. CST-YOLO: a novel method for blood cell detection based on improved YOLOv7 and CNN-Swin transformer [C]//Proceedings of the 2024 IEEE International Conference on Image Processing (ICIP), 2024: 3024-3029.
- [19] LI D, LU Y, GAO Q, et al. LiteYOLO-ID: a lightweight object detection network for insulator defect detection [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-12.
- [20] LU H, FANG M, QIU Y, et al. An anchor-free defect detector for complex background based on pixelwise adaptive multiscale feature fusion [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12.
- [21] ZHOU H, YANG R, HU R, et al. ETDNet: efficient transformer-based detection network for surface defect detection [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-14.
- [22] WANG C, HE W, NIE Y, et al. Gold-YOLO: efficient object detector via gather-and-distribute mechanism [J]. Advances in Neural Information Processing Systems, 2023, 36: 51094-51112.
- [23] VARGHESE R M S. YOLOv8: a novel object detection algorithm with enhanced performance and robustness [C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 2024: 1-6.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [25] TIAN Y, YE Q, DOERMANN D. YOLOv12: attention-centric real-time object detectors [DB/OL]. <https://arxiv.org/abs/2502.12524>.
- [26] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the 15th European Conference on Computer Vision (ECCV), 2018: 3-19.
- [27] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [28] XIANG X, WANG Z, ZHANG J, et al. AGCA: an adaptive graph channel attention module for steel surface defect detection [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12.
- [29] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.