

DOI: 10.11835/j.issn.2096-6717.2026.019



开放科学(资源服务)标识码 OSID:



融合 BiFPN 与 CBAM 的地铁隧道病害 视觉检测模型

何齐海^{1,2}, 包建新^{2,3a}, 邹坤秘², 杨文钰^{1,3a}, 李梦^{3a,3b}

(1. 重庆大学溧阳智慧城市研究院, 江苏溧阳 213332; 2. 中铁二局集团有限公司, 成都 610036;
3. 重庆大学 a. 土木工程学院; b. 山区土木工程安全与韧性全国重点实验室, 重庆 400045)

摘要:传统人工巡检任务重、强度高,且易受主观因素影响而出现漏检、误检等问题。针对地铁隧道表面病害检测中存在的小目标难识别、背景干扰强、边界框定位不准等问题,提出一种改进的 YOLOv12 模型(YOLOv12-BCM)。引入加权双向特征金字塔网络(BiFPN-Concat)替代传统 Concat 层,增强多尺度特征融合能力;在 Neck 网络中嵌入卷积注意力机制(CBAM),抑制复杂背景干扰;采用基于最小点距离的边界框损失函数(MPDIoU)优化边界框回归,提升定位精度。基于地铁隧道采集装置所采集的隧道表面病害图像验证结果表明,相比原模型,YOLOv12-BCM 在 mAP50 由基础的 0.761 提升到 0.799(+3.8%)、Recall 由 0.749 提升到 0.78(+3.2%)、Precision 由 0.723 提升到 0.740(+1.7%)、F1 分数由 0.736 提升到 0.760(+2.4%),且效果优于经典的 SSD、RT-DETR 目标检测模型,改进模型在小目标病害检测方面表现突出。Grad-CAM++ 可视化分析进一步验证了模型对病害区域的关注准确性。

关键词:地铁隧道;病害检测;计算机视觉;图像采集;目标检测

中图分类号:U457.2 文献标志码:A 文章编号:2096-6717(XXXX)XX-0001-10

Visual inspection model for subway tunnel defects integrating BiFPN and CBAM

HE Qihai^{1,2}, BAO Jianxin^{2,3a}, ZOU Kunmi², YANG Wenyu^{1,3a}, LI Meng^{3a,3b}

(1. Institute for Smart City of Chongqing University in Liyang, Liyang 213332, Jiangsu, P. R. China; 2. China Railway No. 2 Engineering Group Co., Ltd., Chengdu 610036, P. R. China; 3a. School of Civil Engineering;
3b. State Key Laboratory of Safety and Resilience of Civil Engineering in Mountain Area, Chongqing University, Chongqing 400045, P. R. China)

Abstract: Traditional manual inspection tasks are labor-intensive and physically demanding, with potential for missed detections due to subjective factors. This paper addresses challenges in detecting small targets, strong background interference, and inaccurate bounding box localization during surface defect detection in subway

收稿日期:2025-11-21

基金项目:国家自然科学基金(52508368);重庆市博士后研究项目特别资助(2024CQBSHTB1004)

作者简介:何齐海(1984-),男,高级工程师,主要从事地下工程与隧道施工研究,E-mail:409022387@qq.com。

李梦(通信作者),男,博士生,E-mail:202516108382@stu.cqu.edu.cn。

Received: 2025-11-21

Foundation items: National Natural Science Foundation of China (No. 52508368); Special Funding for Postdoctoral Research Projects in Chongqing City (No. 2024CQBSHTB1004)

Author brief: HE Qihai (1984-), senior engineer, main research interest: underground engineering and tunnel construction, E-mail: 409022387@qq.com.

LI Meng (corresponding author), PhD candidate, E-mail: 202516108382@stu.cqu.edu.cn.

tunnels by proposing an improved YOLOv12 model (YOLOv12-BCM). First, a weighted bidirectional feature pyramid network (BiFPN-Concat) replaces the traditional Concat layer to enhance multi-scale feature fusion capabilities. Second, a convolutional attention mechanism (CBAM) is embedded within the Neck network to suppress complex background interference. Finally, a minimum point distance-based bounding box loss function (MPDIoU) optimizes bounding box regression, improving localization accuracy. Validation using tunnel apparent defect images collected by subway tunnel monitoring devices demonstrates that Compared with the original model, YOLOv12-BCM increases mAP50 from 0.761 to 0.799 (+3.8%), Recall from 0.749 to 0.780 (+3.2%), Precision from 0.723 to 0.740 (+1.7%), and F1-score from 0.736 to 0.760 (+2.4%)., respectively, outperforming classic SSD and RT-DETR object detection models. The improved model demonstrates outstanding performance in detecting small-scale defects. Grad-CAM++ visualization analysis further validated the model's accuracy in focusing on defect regions.

Keywords: subway tunnel; defects detection; computer vision; image acquisition; object detection

地铁在中国城市轨道交通体系中占据主导地位^[1]。在地铁长期运营期间,受隧道与车站自身结构材料特性、设备性能劣化、施工质量、隧道结构维护滞后及周边环境影响等多重内外因素的影响,易产生各类结构病害。调研发现,地铁隧道在长期服役阶段易诱发多种典型病害,主要表现为裂缝、渗漏水、剥落等^[2-3]。这些病害若未能及时检测与修复,极易诱发结构失稳甚至突发性灾害事故。因此,构建高效、准确的隧道病害检测方法,是保障地铁运营安全的重要前提。

目前,地铁隧道病害检测主要依赖人工巡检或半自动化方式,人工巡检需在“天窗期”(通常仅3~5 h)内,由技术人员携带照明设备步行检查,依靠经验对病害进行目视判别与记录;半自动化检测则需现场作业人员借助专用测量仪器逐点扫描,并手动存储数据。上述两种方法均易受主观因素影响,存在效率低、强度高及漏检风险大等局限,难以满足大规模线网的快速检测需求^[4]。

基于深度学习的图像识别技术,已在建筑^[5]、桥梁^[6]、隧道^[7]及地质勘探^[8]等场景中有效替代人工巡检作业,且车载式移动检测系统在效率与覆盖范围方面显著优于传统人工巡检^[9]。在地铁隧道表观病害采集装置研发方面,瑞士 AMBERG 公司开发的多功能手推式隧道病害检测车 GPR5000。该设备具有高精度扫描功能,在隧道中以 1.5 km/h 的速度行进可生成 5 mm×5 mm 的检测图像^[10]。Huang 等^[11]设计构建了一套六线阵扫描装置,能够在 5 km/h 速度下以 0.3 mm/pixel 的精度采集隧道表面图像;Li 等^[12]采用线阵相机与光源的采集系统,能在 10 km/h 速度下获得 0.15 mm/pixel 的地铁隧道图像。Menendez 等^[13]提出了一种适用于裂缝检测与评估的机器人框架。然而,现有的手动或电动检测平台主要局限于 0~30 km/h 的中低速检测,难以满

足大规模地铁线网高效检测的实际需求。

在模型算法研究方面,Cord 等^[14]提出一种融合 AdaBoost 算法与纹理描述的道路缺陷自动检测方法,并成功应用于道路图像分类任务。Cha 等^[15]构建基于卷积神经网络(CNN)的裂缝检测模型,为深度学习在隧道病害检测领域的应用提供了思路。Huang 等^[16]设计基于全卷积网络(FCN)的地铁隧道衬砌缺陷识别算法,虽取得一定检测效果,但在检测效率方面存在明显短板,难以适配大规模快速检测需求。Xue 等^[17]提出一种基于区域的全卷积神经网络用于隧道衬砌缺陷识别,该框架由 FCN 网络与 R-FCN 基框架组成,通过 FCN 完成图像初步分类,再利用 R-FCN 基框架实现缺陷定位;然而,受隧道表面图像背景复杂、数据量大等因素制约,此类方法对地铁隧道表面缺陷的检测精度仍有待提升。Xu 等^[18]采用优化后的 Mask R-CNN 模型实现渗漏水与剥落病害的识别。Feng 等^[19]提出改进 YOLO 模型,通过嵌入轻量级注意力机制与自适应特征融合模块,显著增强了对模糊及细小病害区域的检测能力。

现有研究在隧道检测领域多识别单一缺陷,如单一裂缝或渗漏水。在缺陷重叠区域,由于冲突像素的存在,进一步增加了多类型缺陷检测的复杂性与难度。综合来看,现有模型存在三方面关键问题未得到有效解决:一是小目标病害(如细小裂缝、微小渗漏水点)的特征在多尺度传播中易丢失,传统特征融合方式难以实现跨尺度信息的有效聚合;二是隧道背景包含管片拼接缝、螺栓孔等干扰元素,现有模型缺乏对病害区域的针对性聚焦机制,易受背景噪声干扰导致误检;三是隧道病害长宽比差异显著,传统边界框损失函数在处理此类目标时定位精度不足。为解决这些问题,提出一种改进的地铁隧道表观病害检测模型,构建特征增强+干扰抑制

的改进框架,实现各功能模块与检测需求的精准匹配,从而全面提升模型对多类型、小目标病害的检测性能。

1 模型与方法

1.1 现场图像采集

地铁隧道表观病害采集装置用于捕捉高速下地铁盾构隧道表面的特征,由图像采集模块(面阵工业相机、LED光源)、数据收集模块(工业个人电脑、图像采集卡、硬盘驱动器)、同步控制模块(同步器、增量编码器)和巡检小车组成。该平台(图1)搭

载4台1200万像素的面阵工业相机(4112 pixel×3088 pixel),相机焦距为12mm,通过安装在相机支架上的工业相机对隧道衬砌进行连续采集。为了确保图像采集过程中光照强度的稳定性,系统配套LED光源保障光照稳定。通过调整曝光时间及预设参数,实现多相机同一截面同步断面采集。采集图像通过拓展坞传输,存储于工业电脑大容量硬盘上。电脑上安装有专门的相机控制软件,进行数据交换及参数设置。在图像保存过程中,增量编码器安装在检测车的轮毂上随着轮毂一起转动,通过计数器的读数可实现地铁隧道表观病害的无导航卫星系统信号条件下的高精度定位。

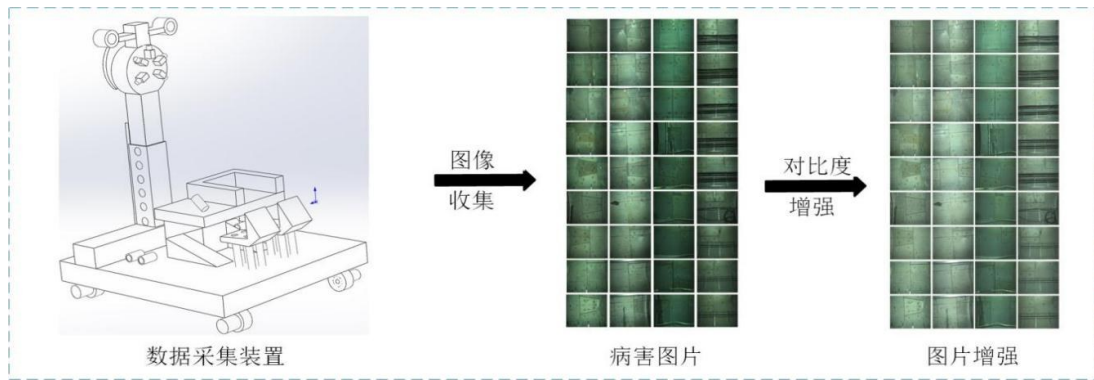


图1 采集装置与采集图像

Fig.1 Acquisition device and image

1.2 模型搭建

采用2025年2月发布的YOLOv12^[20]作为隧道病害检测基础模型,相较于YOLOv8、YOLOv9等经典模型,其以注意力机制为核心的创新架构具备显著优势且高度适配隧道病害检测任务,通过区域注意力模块、残差高效层聚合网络等革新设计,实现检测精度(同规模mAP更优)、实时性(最小推理延迟仅1.64ms/fps)与轻量化的平衡,且兼容性强、小样本适应性强,可有效规避YOLOv8、YOLOv9细小组件捕捉不足、小样本场景过拟合等问题。同时解决隧道较小裂缝漏检问题,低延迟适配监测需求,注意力机制的抗干扰能力可应对隧道光照不均、纹理复杂等环境挑战,小样本适配性还能降低病害标注成本,匹配隧道病害检测目标多样、尺寸差异大、环境复杂的需求,因此确定其为基础模型。针对隧道复杂背景下病害特征易混淆、小目标病害检测效果欠佳的问题。从引入注意力模块、轻量化卷积和改进损失函数三方面对YOLOv12模型进行改进。改进后的模型为YOLOv12-BCM,模型结构见图2。

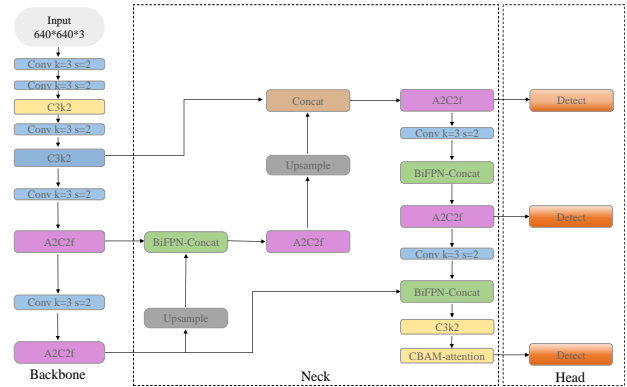


图2 改进模型架构图

Fig. 2 Improved model architecture

单拼接,未考虑特征的重要性差异,导致小目标病害的微弱特征在与大尺度特征融合时被掩盖。将Concat替换为加权双向特征金字塔网络(Bidirectional Feature Pyramid Network)^[21]。其核心创新在于构建自上而下和自下而上的双向特征传播路径与动态权重分配机制,实现小目标特征的强化与跨尺度聚合。

BiFPN的结构如图3所示,主要由输入特征层(P3~P7)、双向传播路径及输出特征层构成。在自上而下传播路径中,高层语义特征(包含目标类别信息)向低层传递,为小目标集中的低层特征图补

- 1)BiFPN-Concat的小目标特征增强机制
传统的Concat层仅对不同尺度特征图进行简

充语义信息;在自下而上的路径中,低层细节特征向高层传播,有效弥补高层特征图的细节丢失。同时,BiFPN为每条特征传播路径分配可学习的自适应权重,对于像素尺寸小的裂缝,BiFPN通过权重调整强化对应的低层细节特征与高层语义特征的融合强度,避免小目标特征在传播过程中被稀释。

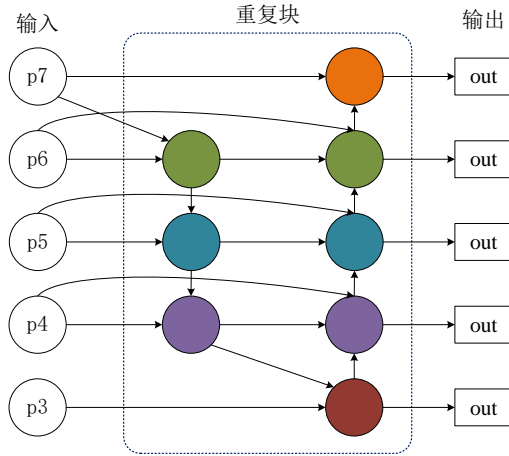


图3 BiFPN结构
Fig. 3 BiFPN structure

2)CBAM的背景干扰抑制原理

卷积注意力模块(Convolutional Block Attention Module)是一种用于前馈卷积神经网络的注意力模块,由通道注意力模块与空间注意力模块两个子模块组成,分别进行通道和空间上的关注^[22]。通过通道特征筛选和空间聚焦的串行协作作用,可动态抑制背景干扰,其具体结构如图4所示。

通道注意力模块的作用是筛选对病害检测具有关键作用的特征通道。隧道图像中,病害区域(如裂缝、渗漏水)与背景干扰(如螺栓孔、拼接缝)的特征在不同通道上的响应存在差异。通道注意力模块通过全局平均池化和全局最大池化获取特征图的通道统计信息,随后利用共享全连接层学习各通道权重,对背景干扰主导的通道赋予低权重,

从而实现通道维度的干扰抑制。

空间注意力模块则专注于定位并强化病害所在的空间区域。在经过通道筛选后,空间注意力模块通过对特征图进行局部平均池化和局部最大池化,得到两个2D特征图,拼接后通过 7×7 卷积运算生成空间注意力权重图,对病害区域赋予高权重,对无病害的背景区赋予低权重,最终实现空间维度的干扰抑制。

$$M_s(F) = \sigma(f^{7 \times 7} [F_{avg}^s, F_{max}^s]) \quad (1)$$

式中: F_{avg}^s 代表全局平均池化特征; F_{max}^s 代表最大池化特征; σ 代表sigmoid函数; $f^{7 \times 7}$ 表示卷积运算,滤波器大小为 7×7 。

CBAM通过通道与空间双重注意力机制,使模型精准聚焦于病害特征区域,有效抑制螺栓孔、管片拼接缝等背景干扰对特征提取的影响,显著提升病害特征表示的鲁棒性和判别能力。

3)MPDIoU对病害的定位优化

边界框回归是目标检测精准定位的关键,传统IoU及其变体(GIoU、DIoU、CIoU等)在处理隧道病害多样长宽比时存在优化问题。针对细长裂缝或不规则剥落区,传统损失函数仅关注边界框重叠区域或中心点距离,难以精准优化四角点,导致定位偏差。提出的Minimum Point Distance based IoU(MPDIoU)损失函数^[23],通过直接优化预测框与真实框的左上和右下关键点距离,兼顾重叠区域、中心点距离及宽高偏差,显著提升不同长宽比病害的定位精度。其核心创新在于将边界框的几何特征与关键点距离相结合,具体计算式为

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

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (3)$$

$$L_{mp} = I - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (4)$$

式中: x_1^{prd} 、 x_1^{gt} 、 x_2^{prd} 、 x_2^{gt} 分别为预测跟真实边界框中心点距

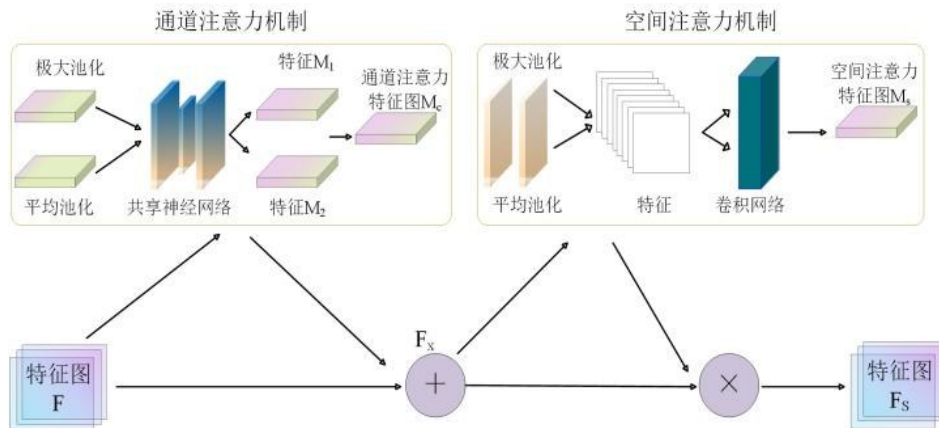


图4 CBAM结构
Fig. 4 CBAM structure

x 轴的距离; y_1^{pd} 、 y_1^{gt} 分别为预测跟真实标注边界框中心点距 y 轴的距离, L_{mp} 表示 MPDIoU 损失, I 是 IoU 表示交并比, 用于衡量两个边界框重叠程度的指标;

MPDIoU 的优势在于: 1) 可通过优化四角点距离贴合细长型病害尺寸; 2) 兼顾重叠区域与关键点距离避免边界框过度偏移; 3) 简化计算过程, 减少模型训练的复杂度, 同时保持高效率 and 准确性。

1.3 数据集

由于面阵工业相机采集的地铁隧道图像尺寸较大 (4 112 pixel \times 3 088 pixel), 计算复杂且识别难度高, 先采用对比度拉伸增强病害信息, 通过滑窗截图将图像分割为 9 份, 每一幅图像尺寸为 1 370 pixel \times 1 029 pixel, 其中小目标病害特指像素面积小于 300 pixel² 的病害区域。利用 Roboflow 专业标注软件对图像中的目标进行标注, 标注后所有图像尺寸被统一设置为 640 pixel \times 640 pixel, 确保输入模型的图像规格一致, 消除尺寸差异对特征提取与模型训练的干扰。这些物体按类别可分为隧道结构件和隧道病害纹理。其中需要检测的隧道病害纹理有裂缝、渗漏水、剥落、剥落修补 4 个类别, 隧道结构件包括管片拼接缝、螺栓孔、管线 3 个类别, 不属于隧道病害, 不会影响隧道的运营使用, 但这类结构件的纹理特征、几何形态与裂缝、微小渗漏水等目标病害存在高度相似性 (如拼接缝线性纹理与细长裂缝近似、螺栓孔深色区域易与微小渗漏水点阴影混淆), 会对病害的检测造成干扰, 因此, 需要标注的共有 7 类 (图 5)。

1.4 评价指标

为全面、客观地评估目标检测模型的性能, 混淆矩阵作为核心分析工具被引入。该矩阵通过量化不同类别的正确和错误预测结果, 可对分类结果进行详细检查。具体而言, 混淆矩阵通过 4 个关键组成部分评估模型性能: T_p 表示真实值和预测值都为病害的样本数; T_n 表示真实值和预测值都为非病害的样本数, 都是预测正确的计数; F_n 表示真实值为病害, 然而预测值却是非病害的样本数; F_p 表示

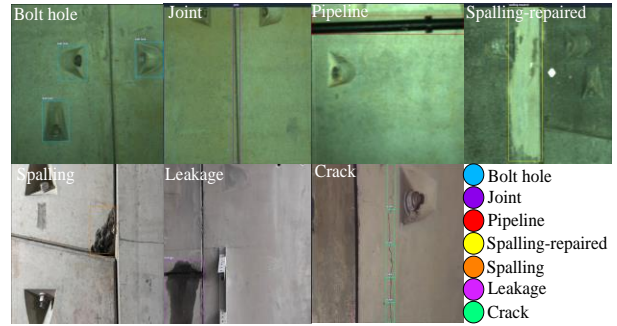


图 5 数据集标注

Fig. 5 Dataset annotations

真实值为非病害, 然而预测值却是病害的样本数, 都是预测错误的计数。在模型精度评价方面, 选取精确率 (Precision, P)、召回率 (Recall, R)、均值平均精度 (mAP, M)、 F_1 分数, 其表达式分别为

$$R = \frac{T_p}{T_p + F_n} \quad (5)$$

$$P = \frac{T_p}{F_p + T_p} \quad (6)$$

$$M = \frac{1}{n} \sum_{i=1}^7 AP \quad (7)$$

$$F1 = \frac{2P \cdot R}{P + R} \quad (8)$$

1.5 超参数选择

构建的目标检测数据集包括 14 029 张地铁隧道图像, 这些图像真实地反映地铁隧道内壁在不同环境条件、光照情况以及各种纹理特征下的实际情况, 具有广泛的代表性。标注了 42 243 个真实标签框。数据集分为训练集包含 11 223 张图像和 34 041 个标签框; 验证集包含 1 403 张图像和 4 009 个标签框; 测试集包含 1 403 张图像和 4 193 个标签框。各数据集中 7 个类别的目标数量分布详见表 1。按照训练集: 验证集: 测试集 = 8: 1: 1 的比例进行划分。需要说明的是, 实际规范维护下的地铁隧道中, 剥落类病害发生率极低, 裂缝类病害虽相对常见, 但其数量仍远少于螺栓孔、管片拼接缝等结构件, 这种类别分布差异是隧道工程的客观实际。所有测试集图像均未参与模型训练过程中的任何参数优化与数据增强操作。

表 1 数据集中 7 类标注目标的数量分布

Table 1 Distribution of the 7 labeled target categories in the dataset

数据集	图片	螺栓孔	裂缝	拼接缝	渗漏水	管线	剥落	剥落修补	总计
训练集	11 223	19 735	1 316	7 146	595	3 913	20	1 316	34 041
验证集	1 403	2 245	319	707	146	433	7	152	4 009
测试集	1 403	2 319	262	798	157	503	8	146	4 193
总计	14 029	24 299	1 897	8 651	898	4 849	35	1 614	42 243

目标检测数据集中的病害图像被统一调整为 640 pixel×640 pixel, 模型的超参数设置为: batchsize=16、epoch=200、learning_rate=0.001、optimizer=SGD、Momentum=0.937。

2 模型改进效果测试

2.1 消融实验

提出的 YOLOv12-BCM 检测模型采用 BiFPN-Concat 替换原始 Concat, 在 Neck 嵌入 CBAM 注意力模块, 并引入 MPDIoU 作为新的损失函数。为量化每项改进的独立贡献及模块间的协同作用, 设计逐步添加改进模块的消融实验, 共开展 7 组实验, 所有实验均在软硬件配置相同的情况下开展, 同时, 为验证结构件标注对病害检测的必要性, 补充设置不含结构件标注(仅 4 类病害标注)的对比试验 YOLOv12-N, 结果表 2 所示。

表 2 不同改进模型检测结果

Table 2 Different results of different improved model

模型	Precision	Recall	mAP50	F1
YOLOv12	0.723	0.749	0.761	0.736
YOLOv12-B	0.771	0.749	0.781	0.760
YOLOv12-BC	0.771	0.722	0.766	0.746
YOLOv12BM	0.784	0.730	0.768	0.756
YOLOv12-C	0.751	0.747	0.783	0.749
YOLOv12-CM	0.757	0.734	0.770	0.745
YOLOv12-M	0.751	0.754	0.781	0.752
YOLOv12-BCM	0.740	0.781	0.799	0.760
YOLOv12-N	0.462	0.566	0.488	0.490

表 3 不同算法的检测结果

Table 3 Detection results of different algorithms

模型	Precision	Recall	AP				mAP50	F1	FPS/(f/s)
			crack	leakage	spalling	Spalling-repaired			
YOLOv12-BCM	0.740	0.781	0.627	0.691	0.841	0.865	0.799	0.760	87
YOLOv9	0.778	0.737	0.453	0.685	0.765	0.850	0.783	0.757	74
YOLOv8	0.759	0.751	0.419	0.741	0.834	0.859	0.791	0.755	87
YOLOv5	0.753	0.754	0.414	0.690	0.839	0.829	0.784	0.753	118
SSD	0.659	0.574	0.399	0.631	0.679	0.691	0.592	0.614	49
Faster Rcn	0.483	0.859	0.294	0.592	0.558	0.687	0.669	0.618	36
RT-DETR	0.561	0.672	0.411	0.639	0.654	0.771	0.613	0.612	72

由表 3 可知, YOLOv12-BCM 的 mAP50 较 YOLOv9 提升 1.6 个百分点, Recall 提升 4.4 个百分点, 尽管其 Precision 低 3.8 个百分点, 但综合 F1 分数持平, 且 FPS 显著高于 YOLOv9, 体现出更好的实时检测能力; 与 RT-DETR 相比, YOLOv12-BCM 在 mAP50、Recall、F1 分数 3 项核心精度指标上均表现出显著优势, 在维持较高检测结果的同时保持更高的推理速度。

实验结果表明, 各改进模块两两组合后的模型, 其核心指标均体现出有效的协同效果。例如, YOLOv12-BM 的 Precision (0.784) 较单一模块 YOLOv12-B (0.771) 和 YOLOv12-M (0.751) 分别提升 1.3 和 3.3 个百分点, 证明 BiFPN 强化的多尺度特征为 MPDIoU 提供了更精准的定位依据。YOLOv12-BCM 的 mAP50 (0.799) 和 Recall (0.781) 为所有模型中的最优值, 相较于最优两两组合模型, YOLOv12-BM 分别提升 3.1 和 5.1 个百分点, 较基础模型提升 3.8 和 3.2 个百分点。其核心优势在于构建了闭环协同机制: BiFPN 先聚合不同尺度病害特征, 解决小目标特征微弱问题; CBAM 对融合后的特征进行通道筛选跟空间聚焦, 剔除背景干扰; MPDIoU 最终针对处理后的病害特征优化边界框, 避免干扰特征导致的定位偏差, 三者形成功能互补, 最大化检测精度与鲁棒性。没有隧道结构件标注的效果明显低于基础模型, 充分证明结构件检测能为模型提供干扰特征区分。

2.2 改进后的模型与其他常用模型对比

为进一步对比改进后模型的性能, 在软硬件配置相同的情况下, 将改进前后模型的检测结果与常用模型 (YOLOv5, YOLOv8, SSD) 及先进模型 (YOLOv9、RT-DETR) 进行对比。表 3 总结了 7 个模型的检测结果, YOLOv12-BCM 的模型检测综合评价指标最高。其中精确率 (Precision) 为 0.740, 召回率 (Recall) 0.781, mAP50 为 0.799, mAP50-95 为 0.578。

图 6 为 YOLOv12-BCM 的真实和预测图像的每个属性的混淆矩阵, 以及它们各自的真阴性 (TN)、假阳性 (FP)、假阴性 (FN) 和真阳性 (TP)。

各项指标全面反映了预测模型的性能: 螺栓孔真阳性率达 0.99, 管线真阳性率 0.94, 表现出极强的检测稳定性; 剥落与剥落修补的真阳性率为 0.82 和 0.85, 显示模型在各类检测任务中具备可靠性。裂缝、渗漏水、拼接缝的检测性能虽达标, 但相较于

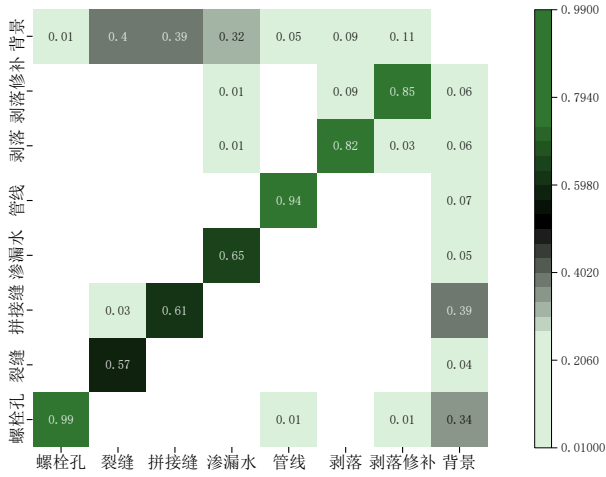


图6 YOLO12-BCM 的混淆矩阵

Fig.6 YOLO12-BCM confusion matrix

其他类别存在一定差异。总体而言, YOLOv12-BCM对真阳性(TP)的准确预测能力,使其可熟练完成各类目标分类,有效提升了其整体预测准确性和可靠性。

图7展示了6组隧道衬砌表面图像的对比实验,用于评估 YOLOv12-BCM在多尺度、多形态缺陷以及复杂背景干扰下的鲁棒性。每组样本左侧

为原始隧道病害图像,右侧依次为 YOLOv12-N、SSD、YOLOv5、YOLOv12、RT-DETR 及 YOLOv12-BCM 的预测结果,各类病害全部用不同颜色框标注,并在框右上角给出置信度。

样本1、2包含规则排列的螺栓孔、纵向拼接缝以及管线。隧道壁周期性纹理极易被误判为缺陷。其中 YOLOv12-N 不涉及此类结构件检测。SSD 因浅层特征图分辨率不足漏检细长拼接缝;其余模型均能有效识别隧道固有纹理。

样本3在螺栓孔中间存在与隧道纹理颜色相近的剥落修补。所有算法均成功检出, YOLOv12-N 检测的置信度分数最低, SSD 跟 RT-DETR 识别效果均不及 YOLO 系列算法。样本4包含两段拼接缝, YOLOv12-N 和 YOLOv5 只识别一个剥落修补漏识别一个, SSD 未检出。而 YOLOv12-BCM 凭借 CBAM 的通道注意力把修补-背景色差通道权重提高,实现了较弱对比度下剥落区与正常隧道背景纹理的有效分割,检测效果良好。

样本5为背景干净的长裂缝。SSD 和 YOLOv12-N 未能识别裂缝, YOLOv12-N 缺少隧道结构件的标注对病害的识别效果较差; YOLOv12

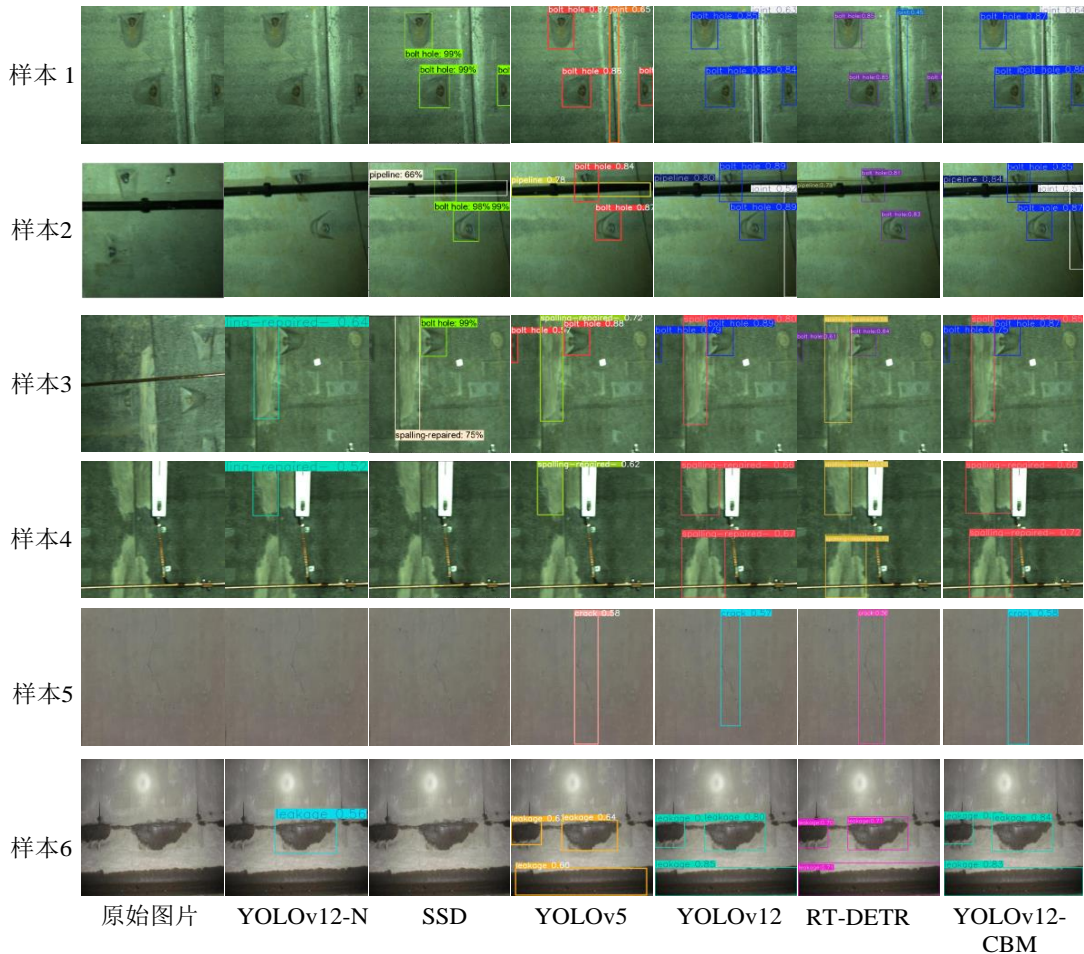


图7 测试集上不同模型的测试结果

Fig.7 Test results of different models on the test set

仅检出部分裂缝。RT-DETR虽能识别整条裂缝但预测分数较低。YOLOv12-BCM可完整检出裂缝且置信度阈值分数最高。得益于MPDIoU的边界框相似性比较指标,其能有效处理预测框与真实框纵横比一致但尺寸不同的情况,在简化计算复杂度的同时,保证了检测效率与准确性。

样本6含多段渗漏水,SSD作为较经典的目标检测模型,对地铁隧道病害的检测效果较差;YOLOv12-N漏检一处。YOLOv12-BCM的BiFPN-Concat通过利用双向路径和加权特征融合策略,有效地整合了不同的特征级别并较好地提取了病害的特征。此外,BiFPN的自适应性,可以根据特征的相关性自动调整权重,显著提高了模型处理不同图像特征的精度。表明提出的加权双向特征金字塔、CBAM注意力与MPDIoU损失的组合,能显著提升隧道表观病害检测的精度与鲁棒性。

2.3 基于Grad-CAM++的训练可视化

为深入探究改进后的YOLOv12-BCM目标检测模型在地铁隧道病害特征识别方面的效果,验证其是否能够准确聚焦病害关键区域,引入梯度加权类激活映射(Gradient-weighted Class Activation Mapping Plus Plus)可视化技术开展分析。

Grad-CAM++作为一种先进的模型可视化技术,核心原理是通过量化特征图中每个位置对最终分类结果的贡献度,生成可直观反映模型关注区域的热图。该方法不仅能够帮助解析模型的决策机制,明确模型在病害检测过程中重点关注的图像区域,还能有效验证模型是否准确识别并定位病害特征,为模型性能的合理性提供直观支撑。

实验选取测试集中渗漏水、剥落修补、裂缝3类典型病害,每类各4张图像,共12张不同场景下的隧道病害图像作为可视化分析对象。采用Grad-CAM++方法对所选图像进行处理,生成特征贡献度热图(如图8所示)。热图中颜色亮度与模型关注程度呈正相关,亮度越高的区域,表明模型对隧道表观病害智能识别任务的特征贡献度越高。

从可视化结果可以清晰看出,热图中的高亮区域清晰与图像中的实际病害区域呈现高度的对应关系,无论是细小裂缝、分散的渗漏水区域还是局部的剥落修补部位,模型都能准确捕捉并重点关注。高亮区域与实际病害区域的高度吻合,充分表明在预测过程中,YOLOv12-BCM模型能有效聚焦病害核心特征,精准定位各类病害的空间位置,同时有效抑制背景纹理等无关信息的干扰,对病害区域的整体关注度和识别准确性较高,进一步验证了模型改进策略的有效性和合理性。

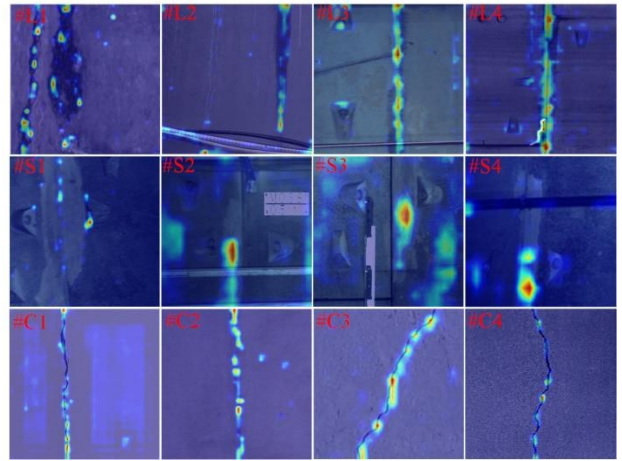


图8 基于Grad-CAM++的病害可视化

Fig.8 Defect visualization of Grad-CAM++

3 结论

针对地铁隧道病害检测中普遍存在的小目标识别困难、背景干扰因素复杂、边界框定位精度不准的核心问题,提出一种融合BiFPN、CBAM与MPDIoU的YOLOv12-BCM改进模型,构建特征增强-干扰抑制-定位优化协同框架,实现模块功能与检测需求的精准适配。具体而言,通过BiFPN-Concat的双向跨尺度特征传播与自适应权重分配机制,强化小目标病害特征表达,解决传统融合中特征丢失问题;借助CBAM的通道筛选与空间聚焦双重机制,可动态抑制螺栓孔、拼接缝等背景干扰,提升病害特征与背景特征判别能力;采用MPDIoU优化边界框四角点距离,适配病害多样长宽比,提升复杂形态病害定位精度。

实验验证表明,与基础YOLOv12模型相比,YOLOv12-BCM模型的mAP50、Recall、Precision和F1分数分别提升3.8%、3.2%、1.7%和2.4%,综合检测性能显著优于SSD、RT-DETR等经典模型,在小目标病害检测和复杂背景鲁棒性方面表现突出。Grad-CAM++可视化进一步证实,模型可精准聚焦病害核心区域,为地铁隧道智能巡检提供可靠技术支撑。

未来研究将围绕模型性能优化、场景适配拓展及应用体系完善展开,结合参数轻量化技术,在保障检测精度的前提下降低模型计算成本;同时深度融合图像语义特征与隧道结构力学参数,构建病害严重程度分级判定模型,实现病害精准检测到风险评估、预警及决策的全过程,进一步提升技术的工程实用性与安全保障效能。

参考文献

[1] 李庆桐,黄宏伟.基于数字图像的盾构隧道衬砌裂缝病

- 害诊断[J]. 岩石力学与工程学报, 2020, 39(8): 1658-1670.
- Li Q T, Huang H W. Diagnosis of structural cracks of shield tunnel lining based on digital images [J]. Chinese Journal of Rock Mechanics and Engineering, 2020, 39(8): 1658-1670. (in Chinese)
- [2] 刘学增, 房茂柳, 吴德兴, 等. 隧道结构病害快速检测技术的现状与发展[J]. 隧道建设(中英文), 2025, 45(4): 657-676.
- Liu X Z, Fang M L, Wu D X, et al. Status and development of rapid detection technology for tunnel structural defects [J]. Tunnel Construction, 2025, 45(4): 657-676. (in Chinese)
- [3] 苏昂, 王士民, 何川, 等. 复合地层盾构隧道管片施工病害特征及成因分析[J]. 岩土工程学报, 2019, 41(4): 683-692.
- Su A, Wang S M, He C, et al. Disease characteristics and causes analysis of segments of shield tunnels in composite stratum during construction [J]. Chinese Journal of Geotechnical Engineering, 2019, 41(4): 683-692. (in Chinese)
- [4] Xu H T, Wang M, Liu C, et al. Automatic detection of tunnel lining crack based on mobile image acquisition system and deep learning ensemble model [J]. Tunnelling and Underground Space Technology, 2024, 154: 106124.
- [5] 杨逸枫, 李康, 刘颖彬, 等. BIM 增强的超大基坑施工场地图像语义分割算法[J/OL]. 土木与环境工程学报(中英文). [2025-11-21]. <https://link.cnki.net/urlid/50.1218.tu.20251120.1530.011>
- YANG Y F, LI K, LIU Y B, et al. BIM-enhanced semantic segmentation algorithm for construction scenes of oversized excavation sites [J/OL]. Journal of Civil and Environmental Engineering. [2025-11-21]. <https://link.cnki.net/urlid/50.1218.tu.20251120.1530.011>. (in Chinese)
- [6] 薛亚东, 高健, 李宜城, 等. 基于深度学习的地铁隧道衬砌病害检测模型优化[J]. 湖南大学学报(自然科学版), 2020, 47(7): 137-146.
- Xue Y D, Gao J, Li Y C, et al. Optimization of shield tunnel lining defect detection model based on deep learning [J]. Journal of Hunan University (Natural Sciences), 2020, 47(7): 137-146. (in Chinese)
- [7] 陈湘生, 曾仕琪, 韩文龙, 等. 机器学习方法在盾构隧道工程中的应用研究现状与展望[J]. 土木与环境工程学报(中英文), 2024, 46(1): 1-13.
- Chen X S, Zeng S Q, Han W L, et al. Review and prospect of machine learning method in shield tunnel construction [J]. Journal of Civil and Environmental Engineering, 2024, 46(1): 1-13. (in Chinese)
- [8] 王婷婷, 王洪涛, 黄志贤, 等. 基于改进 YOLOv7 的露头区岩石裂缝检测识别[J]. 地质科技通报, 2025, 44(1): 1-14.
- Wang T T, Wang H T, Huang Z X, et al. Rock fracture detection and identification in outcrop areas via improved YOLOv7 [J]. Bulletin of Geological Science and Technology, 2025, 44(1): 1-14. (in Chinese)
- [9] Chen P, Huang B, Yuan S S, et al. Automated mobile robot-based detection of in-process tunnel lining deformation during construction [J]. Automation in Construction, 2025, 180: 106560.
- [10] 沈至毅, 谭周. GRP5000 隧道检测车在上海地铁的应用研究[J]. 山西建筑, 2013, 39(27): 158-159.
- Shen Z Y, Tan Z. Application research on GRP5000 tunnel inspection car in Shanghai subway [J]. Shanxi Architecture, 2013, 39(27): 158-159. (in Chinese)
- [11] Huang H W, Sun Y, Xue Y D, et al. Inspection equipment study for subway tunnel defects by grey-scale image processing [J]. Advanced Engineering Informatics, 2017, 32: 188-201.
- [12] Li D W, Xie Q, Gong X X, et al. Automatic defect detection of metro tunnel surfaces using a vision-based inspection system [J]. Advanced Engineering Informatics, 2021, 47: 101206.
- [13] Menendez E, Victores J G, Montero R, et al. Tunnel structural inspection and assessment using an autonomous robotic system [J]. Automation in Construction, 2018, 87: 117-126.
- [14] Cord A, Chambon S. Automatic road defect detection by textural pattern recognition based on AdaBoost [J]. Computer-Aided Civil and Infrastructure Engineering, 2012, 27(4): 244-259.
- [15] Cha Y J, Choi W, Büyükoztürk O. Deep learning-based crack damage detection using convolutional neural networks [J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(5): 361-378.
- [16] Huang H W, Li Q T, Zhang D M. Deep learning based image recognition for crack and leakage defects of metro shield tunnel [J]. Tunnelling and Underground Space Technology, 2018, 77: 166-176.
- [17] Xue Y D, Li Y C. A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(8): 638-654.
- [18] Xu Y Y, Li D W, Xie Q, et al. Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN [J]. Measurement, 2021, 178: 109316.
- [19] Feng Y, Feng S J, Zhang X L, et al. A two-step deep learning-based framework for metro tunnel lining defect recognition [J]. Tunnelling and Underground Space Technology, 2024, 150: 105832.

- [20] TIAN Y J, YE Q X, DOERMANN D. Yolov12: Attention-centric real-time object detectors [J/OL]. arXiv: 2502.12524. [2025-11-21]. <https://doi.org/10.48550/arXiv.2502.12524>.
- [21] Doherty J, Gardiner B, Kerr E, et al. BiFPN-YOLO: One-stage object detection integrating Bi-Directional Feature Pyramid Networks [J]. Pattern Recognition, 2025, 160: 111209.
- [22] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module [C]//Computer Vision - ECCV 2018. Cham: Springer, 2018: 3-19.
- [23] Ma S L, Xu Y. MPDIoU: A loss for efficient and accurate bounding box regression [PP/OL]. V1. arXiv: 2307.07662. [2025-11-21]. <https://doi.org/10.48550/arXiv.2307.07662>.

(编辑 XXX)