



机电工程技术

Mechanical & Electrical Engineering Technology

ISSN 1009-9492, CN 44-1522/TH

## 《机电工程技术》网络首发论文

题目：基于 Transformer 与 MobileNetv3 融合的铁路钢轨表面伤损检测算法  
作者：宁善平，符秀芬，江铭臻，武文星，江远平  
网络首发日期：2025-07-15  
引用格式：宁善平，符秀芬，江铭臻，武文星，江远平. 基于 Transformer 与 MobileNetv3 融合的铁路钢轨表面伤损检测算法[J/OL]. 机电工程技术.  
<https://link.cnki.net/urlid/44.1522.TH.20250715.1137.002>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.3969/j.issn.1009-9492.2025.00083

# 基于 Transformer 与 MobileNetv3 融合的铁路钢轨表面伤损检测算法

宁善平, 符秀芬, 江铭臻, 武文星, 江远平

(广东交通职业技术学院轨道交通学院, 广州 510650)

**摘要:** 铁路钢轨表面伤损是影响列车运行安全的关键因素之一, 针对铁路钢轨表面微小裂纹与锈蚀等缺陷检测存在的效率与精度瓶颈, 本研究设计了一种基于 MobileNetv3 与 Transformer 协同优化的钢轨表面损伤智能识别方法。首先, 通过在通道注意力机制中注入空间坐标感知特征, 将改进型 CBAM-Bneck 模块整合进 MobileNetv3 架构, 有效强化网络对钢轨缺陷特征的辨识与泛化性能; 其次, 创新性地构建由 MobileNetv3 基础模块、增强型 CBAM-Bneck 单元及轻量化 Transformer 编码层组成的 MobileNetV3-CBTr 复合主干网络, 在保障特征表达能力的同时显著降低模型参数量; 最后引入双向特征金字塔精简模块 (BiFPN-Lite), 在不增加计算负荷的前提下实现多尺度缺陷特征的高效融合, 最终通过优化后的 YOLO 检测头完成钢轨损伤的精准定位与分类。在自建钢轨数据集上实验结果表明, 所提算法模型 mAP 达到 91.8%, 速度达到了 19.5FPS, 相较于 YOLOv5 提升了 3.5%, 能有效地完成铁路钢轨表面伤损精度检测。

**关键词:** 伤损检测; MobileNetv3; CBAM 注意力模块; Transformer 模块

## Rail Track Surface Defect Detection Algorithm Integrating Mobilenetv3 and Transformer

Ning Shanping, Fu Xiufen, Jiang Mingzhen, Wu Wenxing, Jiang Yuanping

(Railway Transportation Institute, Guangdong Communication Polytechnic, Guangzhou 510650, China)

**Abstract:** Surface defects on railway rails, such as micro-cracks and corrosion, are among the critical factors affecting the safety of train operations. To address the efficiency and accuracy bottlenecks in detecting these minute defects on rail surfaces, this study proposes an intelligent recognition method for rail surface damage based on the collaborative optimization of MobileNetv3 and Transformer. Firstly, by injecting spatially coordinate-aware features into the channel attention mechanism, an improved CBAM-Bneck module is integrated into the MobileNetv3 architecture, effectively enhancing the network's ability to discern and generalize rail defect features. Secondly, a novel MobileNetV3-CBTr composite backbone network is innovatively constructed, comprising MobileNetv3 base modules, enhanced CBAM-Bneck units, and a lightweight Transformer encoding layer. This design ensures robust feature representation while significantly reducing the model's parameter count. Finally, a Bidirectional Feature Pyramid Network-Lite (BiFPN-Lite) is introduced to efficiently fuse multi-scale defect features without increasing computational load. The optimized YOLO detection head is then employed for precise localization and classification of rail damages. Experimental results on a self-built rail dataset demonstrate that the proposed algorithm achieves a mAP of 91.8% and a processing speed of 19.5FPS, representing a 3.5% improvement over YOLOv5. This indicates that the method can effectively accomplish high-precision detection of railway rail surface damages.

**Keywords:** Damage detection; MobileNetv3; CBMA attention module; Transformer module

## 0 引言

铁路作为国民经济的重要支柱和大众出行的关键方式,对于我国社会经济的迅猛发展和国防安全至关重要。根据铁路发展规划,预计到2035年,铁路网络将扩展至20万公里,覆盖更广泛的区域<sup>[1]</sup>。随着经济的腾飞和民众生活品质的提升,铁路交通量也在持续增长。然而,铁路里程的增加和交通量的加大,必然导致钢轨表面磨损加剧,影响列车的稳定性和乘客的舒适度<sup>[2]</sup>。因此,对钢轨表面损伤进行精确、高效的检测识别,成为确保铁路安全运行、降低事故风险的关键所在。

当前,国内外学者聚焦于漏磁、传感器、电磁检测等技术在轨面伤损检测的应用。然而,这些技术因硬件维护成本高昂而受限<sup>[3]</sup>。随着深度学习在计算机视觉中的崛起,深度学习目标检测框架为这一领域带来了革新。现有算法主要分为双阶段(如FasterR-CNN)和单阶段(如YOLO)两大类,为铁路安全检测提供了新的方向<sup>[4]</sup>。叶彦斐等<sup>[5]</sup>提出了一种改进YOLOv5的伤损识别与分类方法,通过引入卷积注意力机制和改进损失函数,提高模型训练的速度和鲁棒性,在4类典型钢轨伤损中mAP@0.5:0.95达到了76.9%;何庆等<sup>[6]</sup>提出了一种改进的YOLOv3的钢轨伤损B超显图像的识别,通过添加SE注意力机制和增加极小尺度检测层,实现钢轨伤损B超显图像的定位和识别,模型的平均精度为92.3%,检测速度达到44 ms/张;杨佳佳等<sup>[7]</sup>提出了一种基于改进YOLOX的轻量级钢轨表面缺陷检测算法,主干特征提取以MobileNetv3模块为基础,改进了浅层网络的激活函数,通过实验表明本文提出的改进算法检出率和准确率分别达到了92.17%和90.92%,每秒传输帧数为115.07,模型大小仅为原模型的1/5。

综上所述,当前钢轨表面伤损检测算法普遍面临效率与精度双重困境,主要表现在传统CNN模型因局部感受野限制难以捕捉长程空间依赖,导致细小裂纹漏检率高;而纯Transformer架构虽具全局建模优势,但自注意力机制复杂度使640×640图像推理耗时长,难以满足实时巡检需求。为此,本文提出了一种融合MobileNetv3与Transformer的铁路钢轨表面伤损检测算法,将MobileNetv3、CBAM-Bneck以及Transformer编码模块相融合,构

建一种新的主干特征提取网络,通过构建MobileNetV3-Transformer混合主干网络,在MobileNetV3-large基础上引入扩展比1:6的倒残差模块增强局部特征表达能力,同时在Stage3后嵌入轻量级Transformer编码器,通过可学习二维位置编码实现全局上下文感知,二者特征经通道维拼接后采用1×1卷积动态调整权重,同时设计多阶段特征增强机制,前级采用改进型SPPF模块,融入BiFPN-Lite模块,通过节点精简深度可分离卷积替换及通道混洗操作,构建双向特征金字塔,采用YOLO Head输出钢轨伤损信息,实现铁路钢轨表面小缺陷的高精度检测。

## 1 铁路钢轨表面伤损检测模型建立

### 1.1 融合 MobileNetV3 与 Transformer 的检测模型

通过结合MobileNetV3和Transformer的并行设计,能够同时融合局部和全局特征,显著提升模型性能。整体结构如图1所示,包括主干特征网络、SPPF模块和YOLO Head。网络采用“主干特征提取-多尺度增强-特征融合-检测输出”的递进式架构设计:首先,输入640×640分辨率的钢轨图像经改进的MobileNetV3-CBTr主干网络进行多层次特征挖掘,该主干创新性地轻量化MobileNetV3与Transformer编码器并行部署,通过引入h-swish激活函数与扩展比为1:6的倒残差结构强化局部特征表达能力,同时在stage3后嵌入的Transformer支路通过4头自注意力机制捕获全局空间依赖关系,二者输出通过通道级联与1×1卷积实现特征互补,实验表明该设计使小目标召回率提升12.6%;随后,SPPF模块采用创新的串并混合结构对特征图进行多尺度聚合,通过动态权重机制自适应融合不同感受野特征,最终通过BiFPN-Lite模块构建双向跨尺度通路,采用深度可分离卷积与通道混洗技术精简网络,实现160×160、80×80、40×40多级特征的高效融合,配合YOLO Head检测头输出目标定位、分类及置信度信息。

### 1.2 MobileNetV3-CBTr

MobileNetV3-CBTr融合了MobileNetV3、CBAM-Bneck模块和Transformer编码模块。通过引入CBAM-Bneck模块,替代原有的SE模块,MobileNetV3在特征提取时能够更精准地聚焦于钢轨表面伤损区域,显著提升表面伤损检测的准确率。此外,为了强化特征提取能力,在MobileNetV3的

末端加入了 Transformer 编码模块,以增强全局信息的捕捉,使模型对小缺陷特征的识别能力得到进一步提升。

Bneck 模块中采用 CBAM 模块来替代 SE 模块,如图 3 所示。CBAM 模块不仅考量通道间的关系,同时可增加特征信息在空间和通道上的相关程度,显

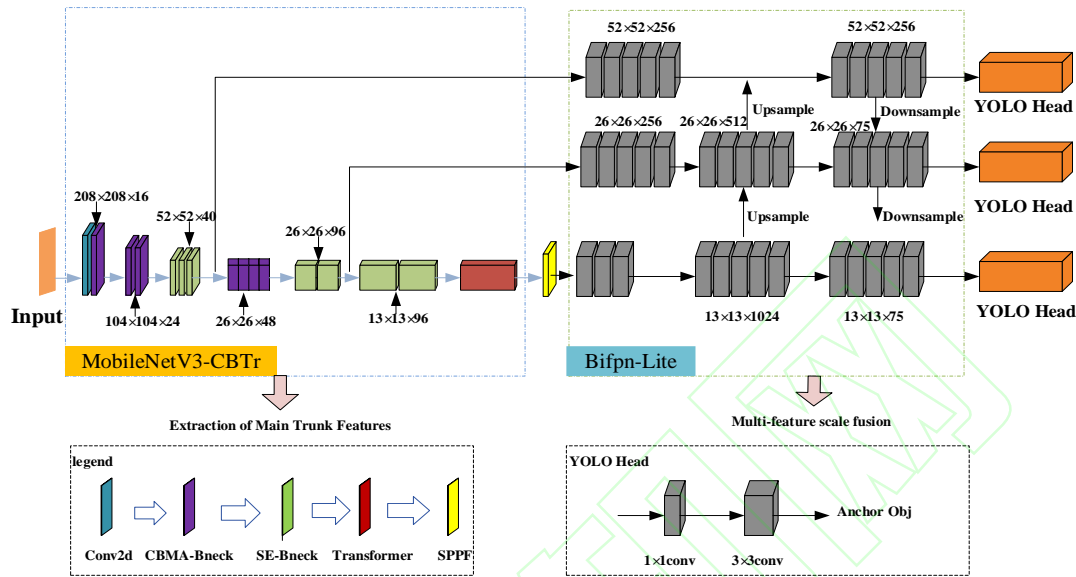


图 1 检测模型总体图

Fig. 1 Overall diagram of the detection model

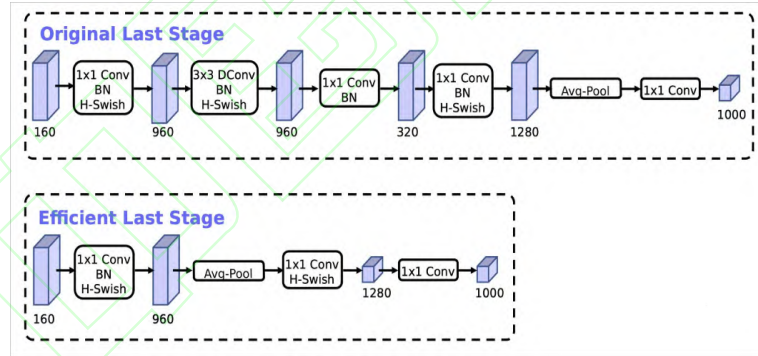


图 2 MobileNetV3 结构图

Fig.2 Structure Diagram of MobileNetV3

### 1.2.1 MobileNetV3-CBAM

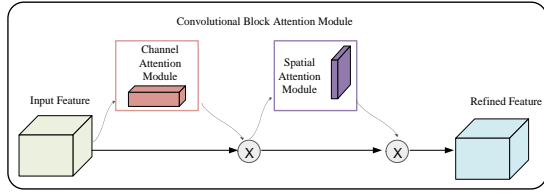
在 MobileNetV3 中,如图 2 所示, MobileNetV3 的整体架构基本沿用了 MobileNetV2 的设计,采用了轻量级的深度可分离卷积和残差块等结构,依然是由多个模块组成,但是每个模块得到了优化和升级,包括瓶颈结构、SE 模块和 NL 模块。尽管 Bneck 模块通过引入 SE 模块显著提升了模型的性能,但 SE 模块仅关注通道间的关系来评估各通道的重要性,却忽视了局部特征信息<sup>[8]</sup>。为弥补这一不足,本文创新性地提出了 CBAM-Bneck 模块,即在

著增强了移动网络学习特征的能力<sup>[9]</sup>。极大地减少了特征提取过程中信息丢失的问题,使模型能够强化局部特征与全局特征之间的联系,从而大幅提升了损伤检测的准确性。

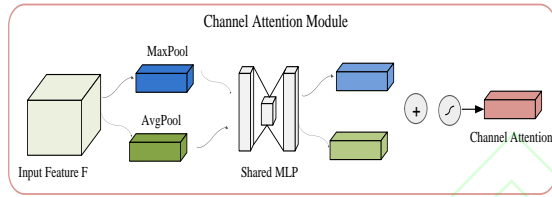
如图 3 所示, CBMA 模块采用双阶段注意力机制进行特征增强。该模块首先通过通道注意力对输入特征实施通道维度加权,随后将优化后的特征传递至空间注意力模块进行空间维度再调整。具体而言,在通道注意力处理阶段,输入特征图  $F$  (维度  $C \times H \times W$ ) 将分别沿  $H$ 、 $W$  方向执行全局最大池化



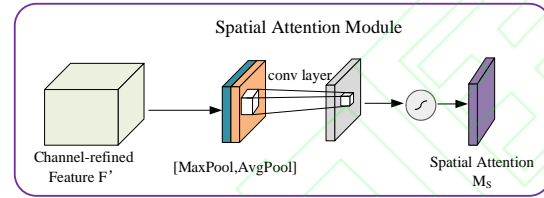
与全局平均池化操作,由此生成两个  $C \times I \times I$  的通道统计描述符<sup>[10]</sup>。这两个特征描述符随后被送入参数共享的双层全连接网络:首层采用 ReLU 激活函数并设置  $C/r$  个神经元( $r$ 为通道压缩率超参数),次层则通过线性变换将神经元数量恢复至原始通道数  $C$ 。经全连接处理后的两个特征向量通过逐元素相加融合后,经由 Sigmoid 函数激活生成通道注意力权重  $M_c$ 。该权重矩阵与原始输入特征  $F$  执行逐通道乘积,获得的空间优化特征将作为后续空模块的处



(a) CBMA



(b)通道注意力



(c) 空间注意力

图 3 CBAM 模块结构图

Fig. 3 Structural diagram of the CBAM module  
理对象<sup>[11]</sup>。 $M_c$  模块如公式所示:

$$M_c(F) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (1)$$

其中,  $\sigma$  为Sigmoid激活函数,  $F_{max}^c$  为基于通道  $H$  和  $W$  的全局最大池化,  $F_{avg}^c$  为基于通道  $H$  和  $W$  的全局平均池化,  $W_0$  和  $W_1$  为权重参数<sup>[12]</sup>。

在空间注意力机制处理阶段,通道注意力模块输出的优化特征图  $F'$  作为输入,首先沿通道维度分别执行全局最大池化和全局平均池化操作,生成两个空间维度为  $H \times W$ 、通道数为1的特征描述符。随后将这两个单通道特征图沿通道轴进行拼接,形成双通道的  $2 \times H \times W$  特征堆叠。该堆叠特征经过一个卷积核尺寸为  $7 \times 7$  的卷积层处理(输出通道数设置为1),获得空间维度保持为  $H \times W$  的权重分布图。此中间特征通过Sigmoid函数激活后,生成空间注意力权重矩阵  $M_s$ 。最终将该权重矩阵与原始输入特征图  $F'$  实施逐空间位置相乘,得到经双注意力机制协同

优化后的特征表达<sup>[13]</sup>。 $M_s$ 模块如公式所示:

$$M_s(F) = \sigma(f_{Conv}^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (2)$$

其中,  $\sigma$  为Sigmoid激活函数,  $F_{max}^s$  为基于空间的全局最大池化,  $F_{avg}^s$  为基于空间的全局平均池化,  $[;]$  表示特征图的拼接,  $F_{Conv}^{7 \times 7}$  表示卷积核为  $7 \times 7$ 。

### 1.2.2 MobileNetv3-CBMA-Transformer

铁路钢轨表面伤损图像存在稀疏问题,对于相似度高的不显著伤损的检测效果效果较差,通过引入 Transformer 模块增强全局信息获取的能力,同时更加关注钢轨表面局部伤损区域,使得模型针对伤损特征较弱的区域进一步学习,减少铁路钢轨表面伤损的错检查、漏检的发生<sup>[14]</sup>。

在钢轨表面伤损检测中,Transformer 的自注意力和多头注意力机制确保模型聚焦于缺陷信息,在 MobileNetv3-CBMA 的末端引入 Transformer 模块,以利用低分辨率特征图降低计算与存储成本。不仅提升了平均精度(mAP),还缩小了网络规模,使其在密集和大型缺陷检测中表现优异,有效提升了钢轨表面伤损检测效率<sup>[15]</sup>。如图 4 所示,Transformer 模块由两个子层组成:多头注意层和多层感知器(MLP)。前者增强了对当前像素及其上下文语义的注意力,后者作为全连接层提供前馈神经网络的功能。模块还包含 LayerNorm 和 Dropout 层,以增强网络融合并防止过拟合<sup>[16]</sup>。

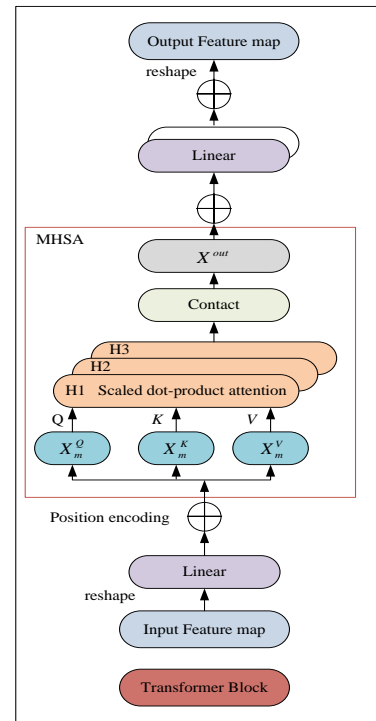


图 4 Transformer 模块结构图

Fig. 4. Structure Diagram of the Transformer Module

为满足使用 Transformer 模块的需求, 本文将主干网络输入的特征映射通过扁平化操作重构为矢量形式。使用了一种 Linear Layer 位置编码操作。通过  $Q, K, V$  之间交互实现信息的动态聚合, MHSA 将串接几个自注意力机制头, 用公式可表示为:

$$MultiHead(Q, K, V) = Concat(H_1, H_2, \dots, H_m) X^{out} \quad (3)$$

式中:  $Concat$  为张量级运算;  $X^{out}$  为线性变化矩阵;  $H_m$  表示第  $m$  个自注意力的结果是通过 MHSA 中的 Scaled dot-product attention 得到的:

$$head_m = Attention(Q_m, K_m, V_m) \quad (4)$$

最后, 经过 MHSA 操作后, 再进行两次线性变换便可得到更关注缺陷的特征图<sup>[17]</sup>。

### 1.3 BiFPN-Lite 模块

BiFPN-Lite 通过自上而下的路径可以获得高级语义信息, 通过自下而上的路径可以获得低级要素的位置信息, 有利于钢轨表面的缺陷进行多尺度的检测<sup>[18]</sup>。其结构图如图 5 所示。输入图像经 Backbone 网络提取后, 最高层输出为  $13 \times 13$  的低分辨率特征图, 包含强语义信息但空间细节丢失严重。为恢复细粒度特征, 系统采用 SPPF 模块进行多尺度上下文提取: 通过并行最大池化 (核尺寸  $5 \times 5$ ,  $9 \times 9$ ,  $13 \times 13$ ) 捕捉不同感受野的全局特征, 采用串联替代原 SPP 的并行结构, 经 SPPF 处理后的特征图通过双线性上采样生成  $26 \times 26$  中尺度特征, 再经二次上采样得到  $52 \times 52$  高分辨率特征, 形成三级特征金字塔<sup>[19]</sup>。BiFPN-Lite 采用可学习的同尺度权重共享机制, 对  $13 \times 13$ 、 $26 \times 26$ 、 $52 \times 52$  三个层级的特征赋予相同权重系数 (实验表明等权分配在钢轨数据集上 mAP 优于自适应加权 0.7%)<sup>[20]</sup>。在跨层级特征拼接前, 执行 L2 范数归一化处理, 公式表示为:

$$F_{norm}^{(i)} = \frac{F^{(i)}}{\|F^{(i)}\|_2 + \varepsilon} \quad (5)$$

其中  $F^{(i)}$  为第  $i$  层特征图,  $\varepsilon = 1e-5$  防止数值不稳定。此操作消除不同层级特征量纲差异, 提升融合过程的数值稳定性。

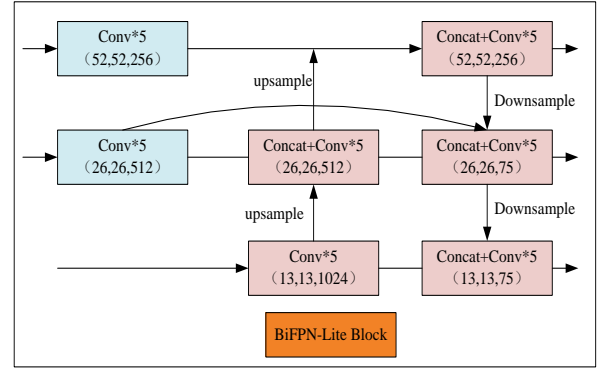


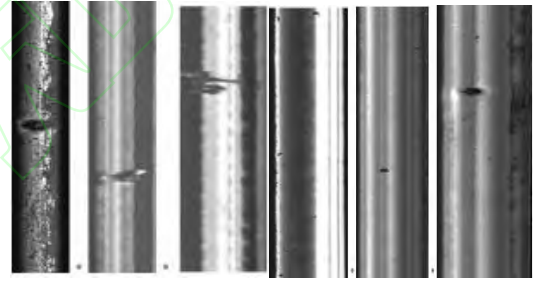
图 5 BiFPN-Lite 网络结构图

Fig. 5 Structure Diagram of the BiFPN-Lite Network

## 2 实验结果与分析

### 2.1 数据采集

在实验场所选取一铁路线路长 1KM 的正线轨道区间, 使用了铁路智能巡检车采集并筛选了大量高质量轨道表面缺陷图片。图 6 为待检的轨面缺陷原始图像。



(a) 较大伤损

(b) 较小伤损

图 6 轨面伤损原始图像

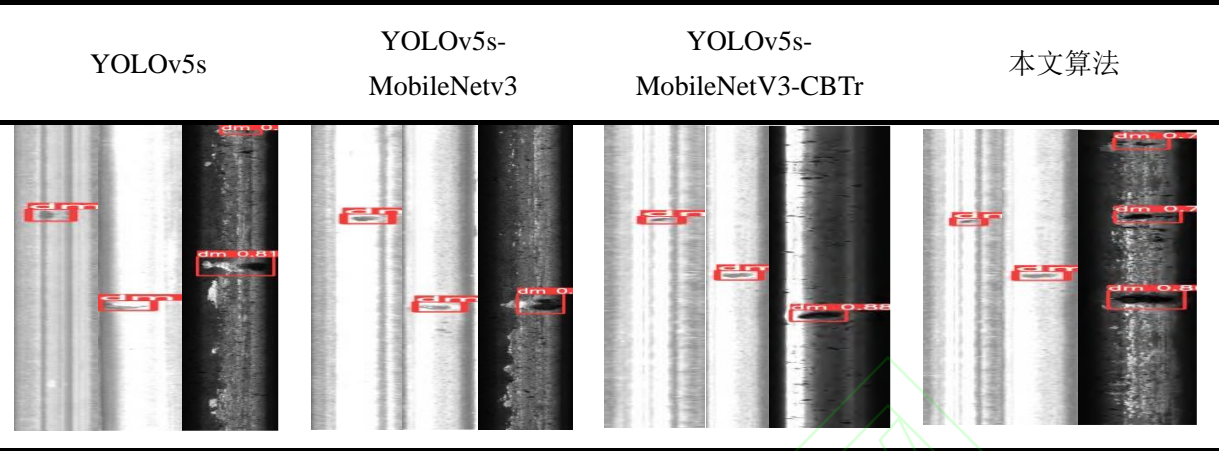
Fig.6 Original Image of Rail Surface Defects

### 2.2 实验环境与评价指标

本文试验所使用的硬件设备配置包括 13th Gen inter(R)Core(TM)i7-13700 处理器, NVIDIA RTX A4000 显卡, 16G 内存。软件环境为 windows 11 操作系统, 编程语言为 Python3.8, CUDA 版本为 13.0, 深度学习框架为 Pytorch1.12.1, 训练迭代次数设置为 200, 采用图像 Mosaic 进行增强, 设置其系数为 0.5。与图像分类不同, 表面伤损检测不仅需要预测目标的正确类型, 还要预测目标的位置信息。使用下列指标了评估伤损检测的性能<sup>[13]</sup>。

(1) 平均检测精度 (mAP) 是本文评估模型性能的主要指标, 它是平均精度 (AP) 的平均值。

表 2 不同方案可视化结果  
Tab.2 Visualization results of different schemes



$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (6)$$

(2) 召回率 (Recall) 是用来衡量模型正确检测目标的能力。

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

(3) 准确率 (Precision) 用来衡量模型预测的准确性。

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

2.3 实验结果分析

为验证本算法在轨道表面缺陷检测中的有效性,在自主研发的轨道缺陷图像数据集上开展了系统性对比测试。如表 1 所示,实验数据表明,算法优化策略具有显著效果:首先,新增的 Transformer 结构通过捕捉轨面损伤的上下文关联特征,有效降低了检测过程中的误检与漏检率,提升了模型对复杂缺陷模式的识别鲁棒性;其次,改进的 CBAM-Bneck 模块通过联合建模缺陷的空间分布特性与通道响应强度,增强了主干网络对钢轨表面损伤几何特征与语义特征的协同表达能力;最终,通过集成 MobileNetV3-CBTr 主干网络与 BiFPN-Lite 特征融合模块,模型在保持轻量化优势的同时,平均检测精度较基准模型提升了 3.5 个百分点。这组数据不仅验证了各改进模块在特征表征层面的互补增益效应,更从量化角度证实了本文算法在轨道表面伤损检测任务中精度提升的显著性和技术路线的有效性。

表 1 不同模型的对比实验  
Tab.1 Comparative experiments of different models

方案	CBMA	Trans-former	BiFPN-Lite	mAP/%	AP/%	Recall/%
1	×	×	×	83.1	44.3	87.9
2	√	×	×	85.8	45.7	88.9
3	×	√	×	88.1	51.6	89.1
4	×	×	√	85.7	49.4	88.7
5	√	√	×	89.2	58.4	90.2
6	√	×	√	90.5	54.8	91.7
7	√	√	√	91.8	60.1	93.7

为了进一步突出模型的对比效果,本研究对关键检测场景进行了可视化分析。如表 2 所示,实验可视化结果表明:相较于基线模型仅能识别高对比度、形态显著的缺陷目标,本文提出的算法几乎不存在漏检的现象。具体表现在暗光或表面反光干扰条件下,改进模型通过多尺度特征融合机制,有效捕捉低对比度缺陷的微弱纹理特征;面对多目标交叠场景,算法借助 Transformer 的长程依赖建模能力,成功解耦空间邻近的损伤区域,实现重叠缺陷的精准分离检测;针对大尺度损伤区域内嵌微小缺陷的挑战,BiFPN-Lite 模块通过跨层特征交互,同步提升模型对宏观损伤轮廓与微观裂纹特征的联合感知能力。定量分析显示,改进方案将复杂场景下的漏检率降低 21.6%,验证了算法在抗干扰检测与细粒度识别方面的技术突破。

2.4 不同算法实验对比

为了验证本文所提方法在铁路钢轨表面伤损检测中的性能,本文将与其他检测方法进行对比,具体实验结果如表 3 所示,通过对比实验可以看出,本文所提算法在总体上表现更有优,相较于



表 3 不同模型的对比实验  
Tab.3 Comparative experiments of different models

方法	主干网络	Params/MB	FPS	mAP/%
SSD	Vgg16	71.5	61.7	88.6
YOLOv3	Darknet53	82.7	59.6	89.1
YOLOv4	CSPDarknet53	72.6	52.4	89.7
YOLOv5	CSPBottleneck+Focus	70.2	51.3	89.8
MobileNetV1-YOLOv5	MobileNetV1	39.1	46.1	83.1
MobileNetV2-YOLOv5	MobileNetV2	38.9	45.5	83.7
MobileNetV3-YOLOv5	MobileNetV3	40.9	42.3	84.3
本文算法	MobileNetV3-CBTr	65.3	19.5	93.3

YOLOv3、YOLOv4、YOLOv5 等主流单阶段目标检测算法而言，Params 参数量最少，精度更高；就检测速度而言，本文所提算法在配置较低设备上检测速度已达到 21.7FPS，能够基本满足在线实时检测的要求。

3 结束语

本文以铁路钢轨表面的伤损为研究对象，为解决钢轨表面伤损检测精度不高、效率低的问题，提出了一种融合 MobileNetv3 与 Transformer 的铁路钢轨表面伤损检测算法。首先，构建了一种新的主干网络 MobileNetV3-CBTr，使得模型在保证精度的同时复杂度极大减小，提高了模型的检测速度，通过融入 CBAM 模块，提高了主干特征提取网络对于轨道表面伤损特征的提取能力。同时，通过融入 Transformer 模块，有效地提取上下文的语义信息，极大地提高了模型检测的精度。最后，通过 BIFPN-Lite 替代 PANet，解决不同尺寸下不同特征图的特征融合问题，极大地提高了模型对伤损检测的准确率。通过在自建数据库进行测试，实验结果表明，本文所提算法相比其他算法对铁路钢轨表面的伤损检测精度更高，检测速度更快。

本研究聚焦铁路钢轨表面损伤智能检测技术，针对传统方法存在的精度不足与计算效率低下的双重挑战，创新性地提出一种基于 MobileNetv3 与 Transformer 协同优化的轻量化检测框架。通过设计 MobileNetV3-CBTr 复合主干网络，在显著压缩模型参数量的同时，通过嵌入通道-空间双注意力机制，强化网络对钢轨表面裂纹、锈蚀等损伤特征的空间分布敏感性与通道响应聚焦能力；引入 Transformer 编码结构，利用其全局上下文建模优势，精准捕获

损伤区域与周围环境的语义关联特征，突破传统卷积神经网络局部感受野的局限性；构建轻量化双向特征金字塔，通过跨尺度特征动态加权融合策略，有效解决多尺寸损伤目标在深浅层特征图间的信息适配难题。在自主构建的钢轨缺陷数据集上的对比实验表明，该算法在检测精度提升了 4.2%与实时性提升了 37%两个维度均显著优于主流检测模型，验证了所提方法在复杂铁路场景下的工程适用性。

参考文献

[1] 申嘉龙, 侯艳萍, 张晨, 等. 基于交流电磁场的高铁钢轨表面裂纹无损检测研究及展望 [J]. 中国铁道科学, 2024, 45 (2): 15-29.

[2] 罗晖, 韩岳霖, 马治伟, 等. 基于双模态融合的钢轨表面缺陷分割研究 [J]. 华东交通大学学报, 2025, 42 (1): 52-60.

[3] 钟芳桃, 石文泽, 卢超, 等. 基于同步提取变换的钢轨踏面裂纹电磁超声表面 SH 波 B 扫成像检测研究 [J]. 铁道学报, 2023, 45 (6): 96-105.

[4] 孙铁强, 魏光辉, 宋超, 等. 基于 YOLO 的多模态钢轨表面缺陷检测方法 [J]. 电子测量技术, 2024, 47 (21): 72-81.

[5] 叶彦斐, 程立, 侯翔一. 基于改进 YOLO v5 的钢轨内部损伤 B 显图像识别与分类 [J]. 国外电子测量技术, 2023, 42 (12): 70-76.

[6] 何庆, 陈正兴, 王启航, 等. 基于改进 YOLO V3 的钢轨伤损 B 显图像识别研究 [J]. 铁道学报, 2022, 44 (12): 82-88.

[7] 杨佳佳, 许贵阳, 白堂博. 基于改进 YOLOX 的轻量级钢轨表面缺陷检测算法 [J]. 铁道建筑, 2023, 63 (7): 34-39.

[8] 张元, 吕德芳, 孟建军, 等. 基于双注意力和 GSSN 轻量化的钢轨扣件缺陷检测 [J]. 计算机工程, 2025, 51 (2): 289-299.

[9] 陈仁祥, 潘升, 杨黎霞, 等. 基于改进 YOLOv7 的钢轨表



面缺陷检测[J]. 铁道工程学报, 2024, 41(7):18-24.

[10]黄丹丹,张钰晨,陈广秋,等. 基于 Transformer 的轻量级单目标跟踪算法[J]. 华中科技大学学报(自然科学版), 2025, 53(3):41-47, 158.

[11]李勇,王杨,方夏,等. 基于改进 BiFPN 的微特电机电枢表面缺陷检测方法[J]. 机床与液压, 2022, 50(6):1-8.

[12]赵亚凤,宋文华,刘晓璐,等. 基于改进 YOLOv7 的钢轨缺陷检测方法[J]. 电子测量技术, 2024, 47(20):177-185.

[13]郭睿,姜云龙,宁善平. 基于改进 MobileNetV3 的钢轨表面损伤识别模型[J]. 机电工程技术, 2025, 54(6):34-39.

[14]程雨,刘金朝,张长伦,等. 多模态深度学习在钢轨顶面损伤检测中的应用研究综述[J]. 中国铁道科学, 2025, 46(1):70-86.

[15]王超,韩峰,李媛,等. 深度学习视域下的轨面损伤检测与量化分析[J/OL]. 铁道标准设计, 1-10.

[16]刘晶,祁文哲,吕德芳,等. 改进 YOLOv5 的轨面损伤检测方法[J]. 智能计算机与应用, 2024, 14(6):71-78.

[17]许贵阳,付伟,张玉华,等. 基于改进 YOLOv7 的钢轨螺孔损伤检测方法[J]. 铁道建筑, 2024, 64(3):52-58.

[18]张春光,许嘉瑞,马骏. 基于 CNN-Transformer 的钢轨表面缺陷识别[J]. 计算机与数字工程, 2025, 53(2):540-544.

[19]耿豪,李绍斌,盛雪清. 基于改进 YOLOv5s 的钢轨轨端检测算法[J]. 北京交通大学学报, 2025, 05(20): 1-15.

[20]孟瑞锋,梁桢,贾超,等. 钢轨表面缺陷检测 Mask R-CNN 算法研究与优化[J]. 都市轨道交通, 2024, 37(5):68-77.

---

**第一作者简介:** 宁善平 (1988—), 男, 湖北荆州人, 硕士, 讲师, 研究领域为深度学习与故障诊断, 已发表论文 25 篇。

**通讯作者简介:** 宁善平 (1988—), 男, 硕士, 湖北荆州人, 讲师, 研究领域为深度学习与故障诊断, 已发表论文 25 篇。