

基于改进 YOLOv13 的铁轨紧固件缺陷检测算法

周会祥, 曹菀菁, 李柯霖, 汪舰

(华东交通大学信息与软件工程学院, 江西 南昌 330013)

摘要: 为解决铁轨紧固件缺陷检测中特征提取不充分及多目标检测精度不足的问题。在 YOLOv13 算法基础上, 引入上下文联合特征提取模块增强多尺度特征捕获能力, 采用多样化分支块提升卷积表征多样性, 并嵌入倒置残差注意力机制以优化特征融合与推理效率。实验表明, 改进后的 YOLO-CTGI 模型在 fastener-defect-detection 数据集上 mAP₅₀ 达到 95.9%, 较原 YOLOv13 提升 1.4%, 参数量为 5.7M, 推理速度基本保持不变的情况下, 缺失与异物类缺陷上检测效果提升。本文方法在保持较高检测速度的同时有效提升了复杂场景下紧固件缺陷的识别精度, 为轨道运维智能化提供了可靠技术方案。

关键词: YOLOv13; 铁轨紧固件; 缺陷检测; 注意力机制

中图分类号: U283

文献标志码: A

Defect Detection Algorithm for Railway Fasteners Based on Improved YOLOv13

Zhou Huixiang, Cao Wanjing, Li Kelin, Wang Jian

(School of Information and Software Engineering, East China Jiaotong University, Nanchang, Jiangxi 330013, China)

Abstract: This study aims to address the issues of insufficient feature extraction and low multi-target detection accuracy in rail fastener defect detection. Based on the YOLOv13 algorithm, ContextGuidedBlock_Down (CGBlock_Down) was introduced to enhance multi-scale feature capture capability, a Diverse Branch Block (DBB) was adopted to improve convolutional representation diversity, and an iEMA was embedded to optimize feature fusion and inference efficiency. Experimental results indicate that the improved YOLO-CTGI model achieves an mAP₅₀ of 95.9% on the fastener-defect-detection dataset, representing a 1.4% improvement compared with the original YOLOv13, and the parameter count is 5.7M. while the inference speed remains basically unchanged, the detection effect on missing and foreign object defects has been improved. The proposed method effectively improves the recognition accuracy of fastener defects in complex scenarios while maintaining high detection speed, providing a reliable technical solution for intelligent railway maintenance.

Key words: YOLOv13; Railway Fasteners; Defect Detection; Attention Mechanism

轨道交通作为我国交通运输体系的核心组成部分, 承担着大量客货运输任务, 对国家经济发展与民生保障具有不可替代的战略意义。近年来, 随着高速铁路技术的飞速发展和路网规模的持续扩张, 铁路运营安全已成为关注焦点。轨道结构作为铁路系统的关键基础设施, 其状态完好性直接决定行车安全; 其中, 铁轨紧固件(包含弹条、螺栓、扣板等)作为连接钢轨与轨枕的核心部件, 对维持轨距、传递载荷、缓冲振动至关重要。然而, 紧固件在长期服役中受循环载荷、环境腐蚀与机械磨损影响, 易出现松动、断裂、丢失等缺陷。此类缺陷若未及时发现处理, 轻则造成轨道几何尺寸超限, 重则引发列车脱轨等重大事故, 对铁路运营安全构成严重威胁。因此, 对紧固件状态进行定期精准检测, 是保障铁路运营安全的关键环节。

传统轨道检测主要依赖人工巡检[1], 存在效率低、劳动强度大、漏检率高及主观性强等问题, 难以

满足大规模精细化检测需求。随着计算机视觉与深度学习技术的发展，基于图像分析的智能检测方法逐渐成为轨道运维的研究热点。此类方法通过自动采集图像并利用深度学习算法实现缺陷的自动识别与定位，显著提升了检测效率与准确性。

在扣件缺陷检测中，主流深度学习方法主要基于目标检测模型，以 Faster R-CNN^[2]和 YOLO^[3](You Only Look Once)系列为代表。Faster R-CNN 作为典型的两阶段方法，在扣件缺陷识别中取得了一系列进展。例如，裴莹玲等^[4] (2022) 引入可变形卷积改进 Faster R-CNN，对扣件偏移和丢失状态的检测精度分别达到 99.34%和 76.80%，mAP (平均精度均值) 为 88.07%；梁楠等^[5] (2024) 通过多层卷积结构增强对缺陷特征的提取，在复杂光照条件下提升了检测鲁棒性；Zhao Q 等^[6] (2025) 在轨道缺陷检测上实现了 97.3%的检测准确率。此类方法虽精度较高，mAP 超过 85%，但其检测速度较慢，在实时应用中存在局限。现有 YOLO 轻量化模型中，鄢化彪等人^[7]基于 YOLOV8(n)模型提出的剪枝的轻量级缺陷检测方法，通过 DepGraph 剪枝对改进后的模型进行优化，该模型侧重通用场景特征提取效率，但是在严重遮挡下的极端环境中，检测精度稍显不足。Shuangbao Ma 等人^[8]提出了一种基于 YOLOv8 改进的轻量级目标检测算法。该算法通过引入轻量级骨干网络 GhostNet^[9]，利用其少量核心卷积 +廉价线性变换生成冗余特征的核心设计，有效降低了模型的参数量与计算复杂度；同时在主干网络中融入 MPCa (多路径坐标注意力 Multi-Path Coordinate Attention) 机制，强化了模型对目标关键特征的针对性提取能力。该方法为计算资源受限的边缘计算场景提供了切实可行的部署方案，其在钢材表面缺陷检测标准数据集 NEU-DET 上的实验结果表明，检测精度 (mAP) 达到 78.6%，实时推理速度为 171.5 FPS。然而，该算法在检测精度与实时推理速度的协同优化方面仍有进一步提升的空间，尚未完全满足复杂工业场景下的高精度、高实时性检测需求。

YOLO 系列作为单阶段检测算法的代表，在速度与精度之间实现了良好平衡，已广泛应用于交通领域多种视觉任务。例如：朱泓宇等^[10] 在改进 YOLOv8 网络在绝缘子缺陷检测中的应用的场景中，在复杂背景下对低像素占比缺陷的检测 mAP 仍可达 93%以上，处理单张图像仅需数毫秒，满足高实时性要求。此外，生成式数据增强策略的引入有效缓解了样本不平衡问题，提升了模型泛化能力。因此，YOLO 系列算法已成为铁轨扣件智能检测中的重要技术路径。

针对 YOLOv13^[11]算法在铁轨紧固件检测中存在的特征提取不充分及多目标缺陷检测能力较弱的问题，本文以 YOLOv13 为基准算法，提出一种改进模型 YOLO-CTGI。主要改进包括：

- 1) 设计上下文联合特征提取模块 (ContextGuidedBlock_Down)^[12]，通过引入空洞卷积扩大感受野，以获取更广泛的上下文信息，使模型能够捕捉更大范围的特征而非局限于局部细节；
- 2) 采用多样化分支块 (Diverse Branch Block)^[13]结构，通过多尺度与多结构的卷积分支增强卷积层的表征多样性，提升特征表达能力；
- 3) 在主干网络中嵌入倒置残差注意力机制 (iEMA)^[14]，进一步强化特征提取性能；

1 YOLOv13 算法简介

YOLOv13 作为 YOLO 系列于 2025 年发布的最新目标检测算法，在延续经典架构的基础上实现了多项重要革新。其网络结构包含三个主要部分：轻量化主干网络 (Backbone)、增强型颈部网络 (Neck)

以及检测头 (Head)。其中, 主干网络采用深度可分离卷积进行优化, 以 DS-C3k2 模块与 DSConv^[15] 模块替代传统大核卷积, 显著提升了计算效率; 在 YOLOv13 架构中, B3、B4、B5 是骨干网络最后三个特征提取阶段输出的核心多尺度特征图, 其中 “B” 对应 Backbone, 下标 “3、4、5” 代表特征提取阶段序号, 三者分别对应 1/8、1/16、1/32 下采样比例, 是连接特征提取、关联增强 (HyperACE) 与特征融合的关键载体, 其决定模型对多尺度目标的检测能力与复杂场景适应性。颈部网络引入了超图关联增强机制与全流程聚合分发范式 (FullPAD) 两大核心模块 —— H3、H4、H5 是经过 全流程聚合分发范式处理后生成的关联增强特征图。

2 YOLO-CTGI 目标检测算法

本文在 YOLOv13 算法的基础上提出改进算法 YOLO-CTGI, 重点优化铁轨紧固件缺陷检测性能, 改进后网络结构如图 1 所示:

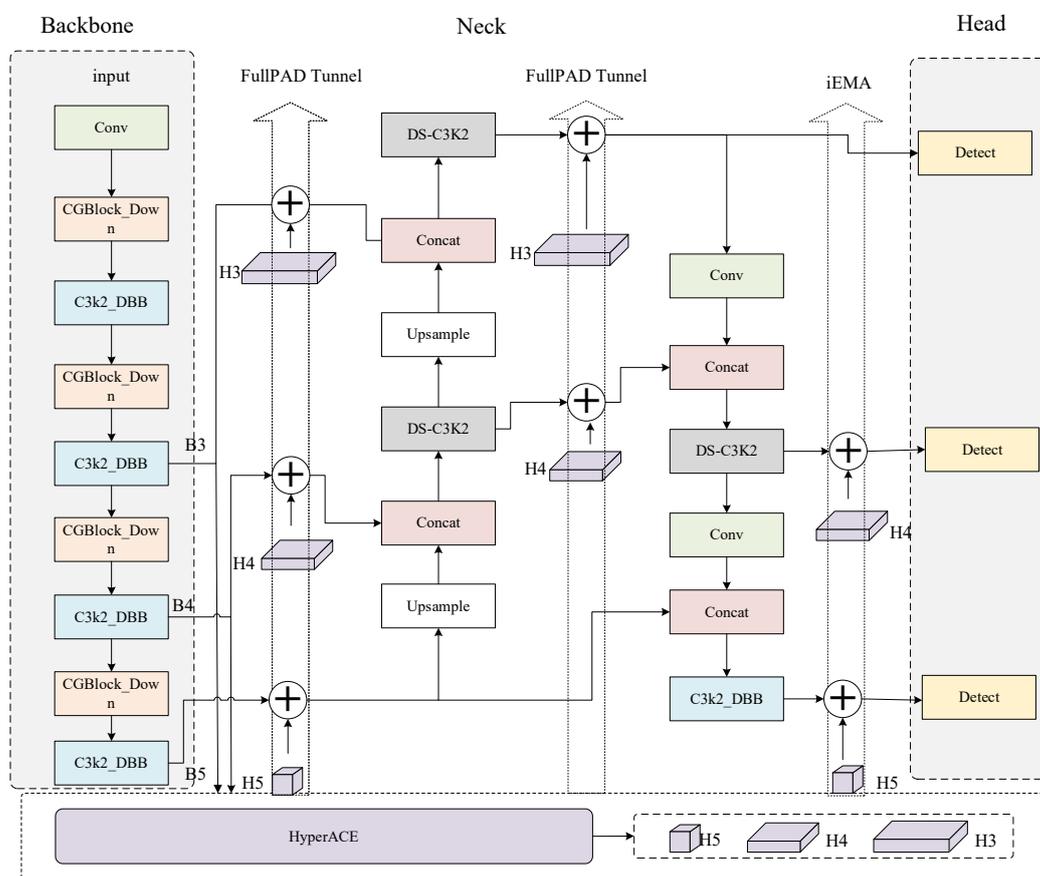


图 1 改进 YOLOv13 的网络结构图

Fig. 1 Improved network model structure of YOLOv13

1) 特征提取网络 (Backbone): 输入图像首先经由 CGBlock_Down 模块进行初步下采样, 随后通过改进的 C3k2_DBB 卷积结构实现多尺度特征提取与融合, 生成大、中、小三种不同尺度的细粒度特征图, 以有效捕捉不同尺寸的缺陷区域。

2) 特征融合网络 (Neck): 将来自特征提取网络的多尺度特征分别输入 FullPAD Tunnel(全流程聚合与分配隧道), 嵌入与缺陷相关的通道、空间及位置信息。进一步借助 iEMA 注意力机制实现特征

拼接与卷积融合，增强算法对缺陷判别性特征的提取能力，促进浅层细节与深层语义信息的高效交互，最终输出三个语义更加丰富的特征层。

2.1 改进 Backbone 模块

为增强模型对铁路扣件病害区域的感知能力与关键特征的捕获效率，本研究对 YOLOv13 算法的原始主干网络进行了针对性结构改进：将 YOLOv13 的初始卷积层与 DS-C3k2 模块依次替换为本文提出的上下文联合特征提取模块 (CGBlock_Down)，并将原网络中的 DSConv 与 A2C2f 模块替换为 C3k2_DBB 结构。C3k2_DBB 结构的核心设计逻辑是将多样化分支块 DBB 的多尺度表征能力与 YOLOv13 原生 C3k2 模块的残差特征提取能力深度融合。可增强模型在复杂环境下的抗干扰能力与多目标判别能力，使其更适用于实际铁路场景中多干扰条件下的扣件缺陷检测任务。

2.1.1 上下文联合特征提取模块

为提升模型在下采样过程中对多尺度上下文信息的捕获能力，本文提出一种新型下采样模块—CGBlock_Down (Context-Guided Block Down)。该模块集成了上下文引导机制 (Context Guided Mechanism)，用以替代主干网络中步长为 2 的标准卷积下采样层。其核心优势是：在降低特征图分辨率的同时，能自适应地聚合并优化局部细节与全局上下文信息，从而在减少信息损失的前提下增强语义表达能力。

本模块的设计灵感来源于上下文引导网络 (CGNet^[16])，该网络在语义分割任务中强调了上下文信息对密集预测的重要性。尽管原始的 CGBlock^[17]面向语义分割设计，本文依据 YOLO 框架在高效率目标检测中的需求，对 CGBlock 结构进行了重构与优化，使其在保持较高推理速度的前提下，提升对复杂背景下缺陷目标的检测准确率。

如图 2 所示，CGBlock_Down 模块包含三个核心组件：局部特征提取器、上下文特征提取器，以及特征融合与下采样单元。该结构通过并行与串行组合方式，实现对不同感受野范围内特征的有效提取与融合，从而增强模型对铁路扣件缺陷的敏感性与判别力。

局部特征提取器 $f_{loc}(\ast)$ 采用卷积操作捕捉目标像素的局部空间特征，提取以目标像素为中心的 3×3 邻域内的空间结构信息，包括边缘、纹理与颜色等低级视觉特征。该模块使用标准卷积（而非步长卷积）执行下采样，能够在计算过程中更好地保持特征图的空间分辨率。

周围上下文提取器 $f_{sur}(\ast)$ 采用膨胀率为 r 的 3×3 膨胀卷积构建，通过调整膨胀率可指数级扩大感受野，从而有效捕获目标与其周围环境之间的中级上下文语义。该设计能够在避免下采样造成信息损失、且不增加参数量的前提下，使特征点覆盖更大范围的输入区域。

特征融合与全局上下文 $f_{join}(\ast)$ 整合模块将局部特征与周围上下文特征沿通道维度进行拼接，生成联合特征表示，随后通过批归一化^[18] (Batch Normalization, BN) 与参数化修正线性单元^[19] (Parametric Rectified Linear Unit, PReLU) 进行归一化与非线性激活，以增强特征的判别能力。

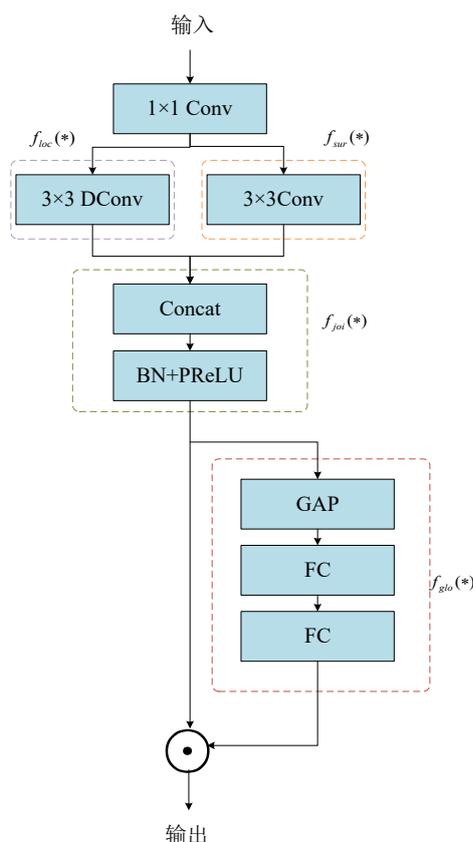


图 2 CGBlock_Down 模块结构

Fig. 2 Structure of the CGBlock_Down module

全局上下文引导 $f_{glo}(*)$ 基于输入图像的全局语义信息，自适应地优化联合特征，突出与场景最相关的特征通道。具体而言，首先对联合特征进行全局平均池化，生成通道统计描述符；随后通过一个轻量级多层感知机（含降维与升维层）计算通道权重向量；最终通过通道重加权机制增强关键通道的响应，抑制冗余或噪声特征。

2.1.2 多样化分支块

YOLOv13 的 Backbone 主要负责从输入图像中提取多层次、多尺度的特征，其基础组件为标准卷积层，但标准卷积的感受野固定、计算模式单一，难以捕捉多样化的局部特征。尽管 YOLOv13 的 HyperACE 机制在宏观上实现了跨尺度的高阶关联建模，但在微观层面，缺乏多路径特征交互、非线性能力有限，每个卷积层的特征提取能力仍有提升空间。为此，本文引入多样化分支块 (Diverse Branch Block, DBB)，其核心设计理念源于结构重参数化思想，通过解耦训练与推理阶段的网络结构：训练时构建多分支拓扑以增强特征多样性，推理时依托卷积的线性特性将多分支等价融合为单一 $K \times K$ 卷积核，最终在不增加推理开销时间的前提下，提升 Backbone 的微观特征表征能力。具体分支设计如下：

主分支：一个标准的 $K \times K$ 卷积层 + BN (batch norm) 层，保留基础的局部空间特征提取能力，其作为 DBB 特征提取的核心骨架。

1×1 卷积分支：一个 1×1 卷积层 + BN 层，通过通道维度的线性变换实现跨通道信息交互，并灵活压缩或扩张通道数以降低后续计算量，同时补充细粒度的通道特征。

1×1 到 K×K 序列分支：一个由 1×1 卷积 + K×K 卷积层（每层后接 BN 层）构成微型两级结构，先通过 1×1 卷积调整通道维度、筛选关键通道特征，再通过 K×K 卷积捕捉局部空间关联，两级变换协同增强特征的非线性表达与抽象能力。

平均池化分支：核大小为 K 的平均池化层（当步长=1 时为平滑操作以抑制噪音，步长>1 时为下采样）+ BN 层，通过聚合局部区域的全局统计信息，与卷积的局部特征提取形成互补。

具体流程如图 3 所示：

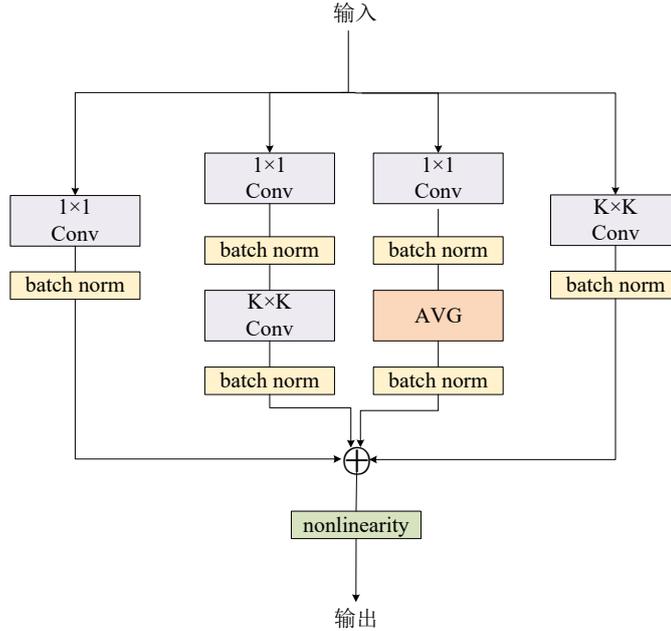


图 3 DBB 模块结构

Fig.3 Structure of the DBB module

多样化分支块 DBB 的分支融合依赖卷积的两大核心线性特性（齐次性与可加性），通过这两个特性可将训练阶段的多分支结构严格等价融合为单一 K×K 卷积核，最终实现训练时复杂化以提升表征能力，推理时轻量化以保持效率的目标，其数学推导如下：

齐次性：对卷积核 F 的参数乘以一个常数 p 后与输入 I 卷积，结果等同于于 I 与 F 卷积后乘以 p ，即：

$$I \circledast (pF) = p(I \circledast F), \forall p \in \mathbb{R} \quad (1)$$

式 (1) 中： I 表示输入特征图 ($C \times H \times W$ 的张量)， F 代表卷积核 ($D \times C \times K \times K$ 的 4 阶张量)， \circledast 代表卷积操作符号， p 表示实数标量。

可加性：两个卷积核 $F^{(1)}$ 、 $F^{(2)}$ 分别对与 I 卷积后进行求和，结果等同于 I 与 $(F^{(1)} + F^{(2)})$ 的卷积，即：

$$I \circledast F^{(1)} + I \circledast F^{(2)} = I \circledast (F^{(1)} + F^{(2)}) \quad (2)$$

基于上述操作，训练结束后可将 DBB 的多分支参数融合为单个 K×K 卷积核，实现尺度卷积、序列卷积、平均池化，捕捉丰富的特征模式，如紧固件缺陷的边缘纹理、局部破损与全局缺失特征。

2.2 改进 Neck 模块

2.2.1 融合 iEMA 注意力机制

iEMA 模块的设计与工作机制，旨在解决铁轨紧固件缺陷检测中“多尺度缺陷特征精准捕捉”与“实

时推理效率”难以兼顾的核心问题。本文设计的倒置残差注意力机制（iEMA）模块，是融合了多尺度注意力模块（EMA）^[20]与倒残差移动块（iRMB）^[21]的高效特征增强模块。

EMA 注意力机制可捕获特征的跨通道关联信息如紧固件缺陷区域与轨枕、铁轨背景的通道响应差异与像素级空间依赖关系如弹条断裂边缘与相邻螺栓的空间关联性，而 iRMB 依托轻量化架构设计，能在低计算开销下实现特征空间变换，二者深度融合使模块兼顾“高特征判别性”与“低推理延迟”，达成性能与效率的协同优化。

iEMA 模块通过双分支协同架构实现 EMA 与 iRMB 的优势互补，具体流程如下：

第一分支为注意力引导分支，接收来自 Neck 层的多尺度特征（涵盖高分辨率浅层细节特征与低分辨率深层语义特征），这些特征对应小尺寸异物、大区域破损等不同缺陷，送入 EMA 注意力机制模块后，经动态计算生成通道权重矩阵——通过强化缺陷相关特征通道的响应强度、抑制冗余背景信息的干扰，完成特征的初步筛选。

第二分支为空间变换分支，借助 iRMB 对输入特征执行轻量化空间变换，采用深度可分离卷积等设计控制计算开销，同时精准保留缺陷边缘、纹理等关键局部视觉信息（如螺栓螺纹结构、弹条细微裂痕），避免特征变换过程中细节丢失。最终，模块通过通道权重矩阵与空间变换后特征的逐元素相乘运算完成特征校准，使缺陷特征得到进一步强化。该设计一方面解决了传统注意力机制计算复杂度高、参数量大的痛点，另一方面弥补了单纯轻量化模块对缺陷特征判别能力不足的短板，精准适配铁轨紧固件缺陷检测场景下“高精度与高实时性兼具”的核心需求。其网络结构如图 4 所示：

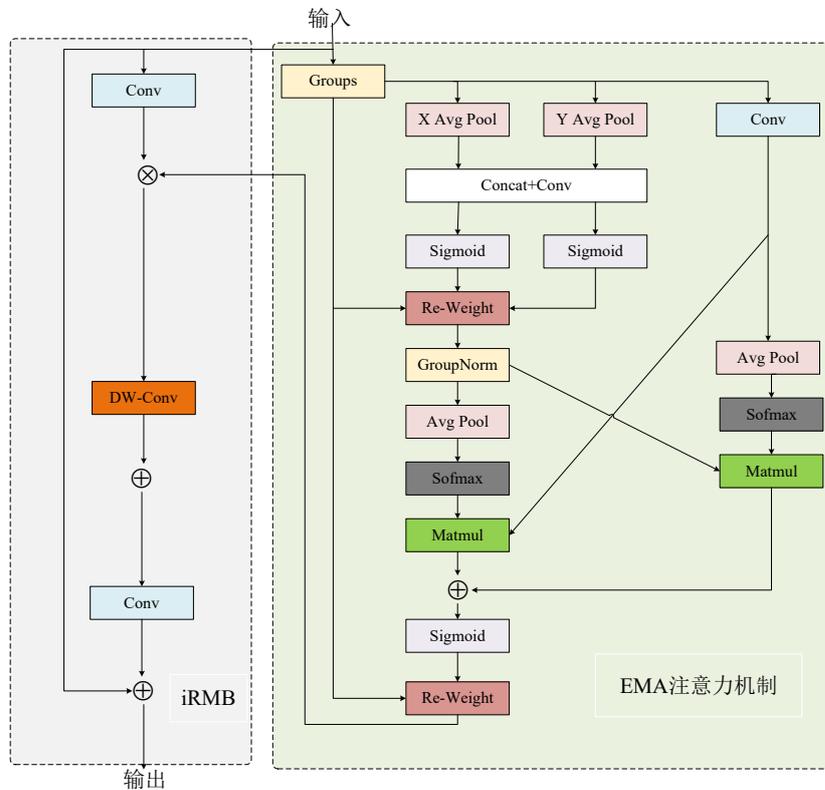


图 4 iEMA 模块结构

Fig.4 Structure of the iEMA module

在 iEMA 模块中, 首先将原始输入中的 3×3 卷积特征一并送入 EMA 注意力机制, 以重新校准通道权重。该机制通过全局池化分别捕捉水平方向编码定义公式所示与垂直方向编码的像素级成对关系。定义如下:

$$z_c^W(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \quad (3)$$

$$z_c^H(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (4)$$

式 (3) (4) 中: c 代表通道数, H 代表特征图的高度, W 是特征图的宽度, $z_c^W(w)$ 表示第 c 通道、宽度 w 上所有高度方向特征的平均值, $x_c(j, w)$ 表示第 c 通道、高度 j 、宽度 w 处的特征值。

$z_c^H(h)$ 表示第 c 通道、高度 h 上所有宽度方向特征的平均值, $x_c(h, i)$ 表示第 c 通道、高度 h 、宽度 i 处的特征值。从而建立跨位置依赖, 增强语义融合能力。具体而言, 特征在不同分支分别处理: 一支通过 1×1 卷积与一维全局池化提取全局上下文, 另一支利用 3×3 卷积提取局部特征经过一个 1×1 卷积层定义嵌入到倒置残差结构 (iRMB) 中进行初步特征变换。随后通过深度可分离卷积 (3×3 卷积) 进一步提取空间特征, 并再次使用 1×1 卷积进行降维。该结构有效压缩了全局通道维度中的冗余信息, 提高了特征表示的紧凑性。

两组特征经过 Sigmoid 激活函数^[22]与归一化操作进行调制, 最终通过乘法实现通道注意力加权, 促进跨通道信息交互。该双分支架构不仅减少了通道冗余和参数量, 还强化了缺陷目标的判别性特征表达, 从而有效提升了模型在缺陷检测任务中的表征能力和检测精度。

3 实验结果与分析

3.1 实验环境与参数设定

表 1 实验环境

Tab.1 Experimental environment

名称	环境配置
GPU	RTX 3090(24GB)
CPU	Xeon(R) Gold 6330
System	Ubuntu
DL framework	PyTorch1.8.1+CUDA11.1
Language	Python 3.8

训练参数设置如下: 采用随机梯度下降 (SGD) 优化算法更新参数, 初始学习率 0.001, 权重衰减系数设置为 0.0005, 以抑制过拟合; 输入图像统一标准化为 640×640 , 批处理大小 (Batch size) 为 64, 训练轮次 (Epochs) 为 300 轮; 配置 10 个工作线程 (Workers) 加速数据读写与预处理, 保障训练高效稳定。

3.2 数据集介绍

本文数据集源自 Roboflow 平台开源的 fastener-defect-detection 数据集, 含 6 类目标: 扣件 1 正常 (fastener)、扣件 2 正常 (fastener2)、扣件 2 破损 (fastener2_broken)、扣件 1 破损 (fastener_broken)、

扣件缺失 (missing)、铁轨异物 (trackbed_stuff)，数据集类别具体分布如下：

表 2 铁轨紧固件缺陷数据集类别数量分布

类别	总数	训练集	验证集	测试集
fastener	1746	1387	181	178
fastener2	462	371	45	46
fastener2_broken	334	270	33	31
fastener_broken	290	231	31	28
missing	921	739	90	92
trackbed_stuff	175	138	18	19

预处理阶段采用 Python 对 20% 的随机样本施加旋转、翻转、对比度随机调整及添加黑色遮挡块等数据增强操作，以提升模型泛化能力；最终生成 2243 张图像，按 8:1:1 随机划分为训练集 1795 张、验证集 224 张与测试集 224 张。部分数据集图片展示如下：

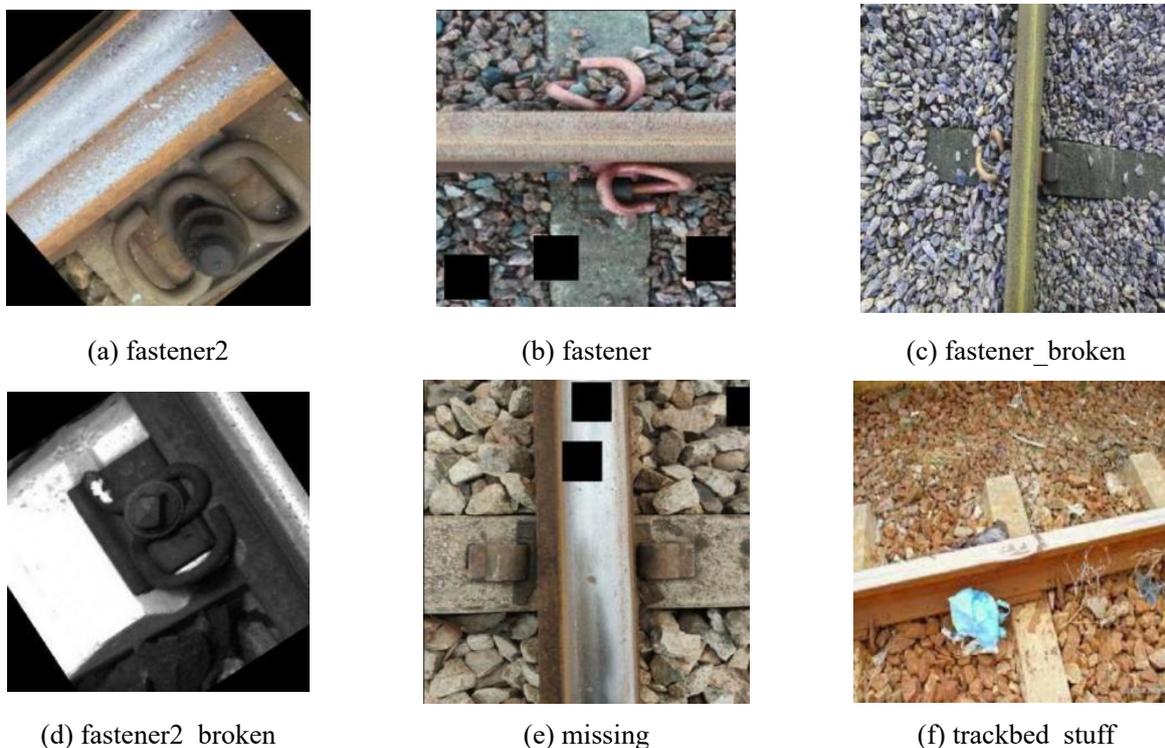


图 5 数据集部分样本

Fig.5 Partial samples of the dataset

3.3 评价指标

本研究采用精确率 (Precision, P)、召回率 (Recall, R)、平均精度 (AP)、均值平均精度 (mAP₅₀)、每秒帧数 (FPS) 作为评价指标，具体定义如下：

$$P = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

其中，TP 为真正例（正确检测的紧固件缺陷数量），FP 为假正例（错误标记为紧固件缺陷数量）。

召回率 R 指被正确预测的正类样本占有所有真实正类样本的比例，具体定义如下：

$$R = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

其中 FN 为假负例（实际缺陷却未被检测到的数量）。

平均精度（Average Precision, AP）是目标检测领域中广泛使用的性能评价指标，其计算基于精确率（Precision, P）与召回率（Recall, R）之间的关系。AP 即通过对 Precision-Recall 曲线进行积分或插值计算得到，具体如下所示：

$$AP = \int_0^1 P(R) dR \quad (7)$$

本文采用全类别平均精度均值（mean Average Precision, mAP）综合评估模型整体表现，其中 mAP_{50} 对应交并比（Intersection over Union, IoU）阈值为 0.5 的指标（即 $mAP@0.5$ ）：

$$mAP_{50} = \frac{1}{C} \sum_{i=1}^C AP(i) \quad (8)$$

其中， C 为类别总数， $AP(i)$ 为第 i 个类别的 AP 值。

同时，采用每秒帧数（Frames per Second, FPS）衡量模型的检测速度，如下所示：

$$FPS = \frac{FrameNum}{ElapsedTime} \quad (9)$$

其中，FrameNum 为检测图片数量，ElapsedTime 为算法检测运行的总时间。

3.4 对比实验及结果分析

网络模型的性能可通过精确率-召回率（Precision-Recall, P-R）曲线进行准确而直观的评估。如图 6 所示，该图展示了改进后 YOLOv13 算法的 P-R 曲线。

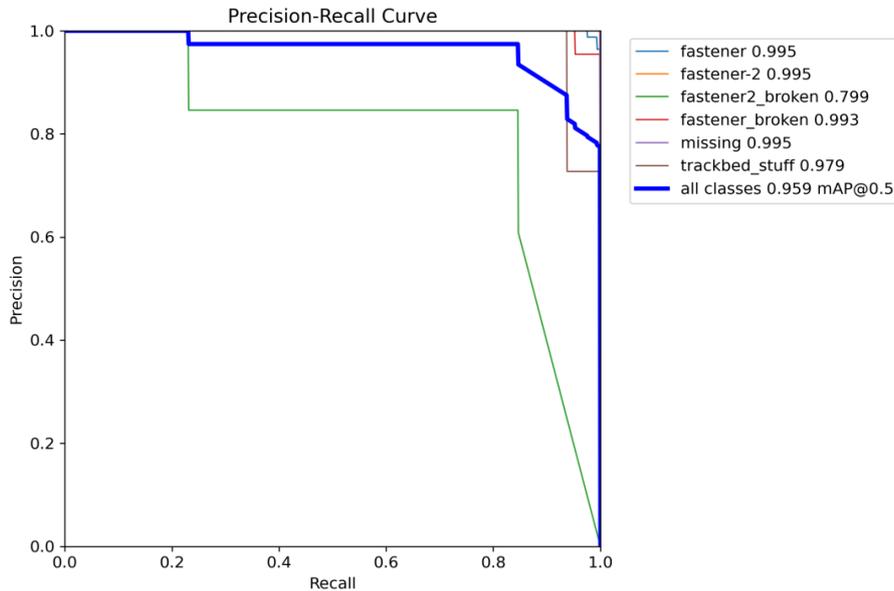


图 6 改进后的 YOLOv13 算法检测 P-R 曲线

Fig.6 The improved YOLOv13 algorithm detects the P-R curve

为验证本文算法在 fastener-defect-detection 数据集上的检测效果，在相同的实验环境和数据集下，选取了不同的目标领域检测算法进行对比实验，具体实验结果如表 3 所示。

表 3 不同模型的实验对比

Tab.3 Experimental comparison of different models

Model	AP/%						mAP_{50}	FPS	
	fastener	fastener2	fastener_broken	fastener2_broken	missing	trackbed_stuff	Params /%	/(f/s) /M	
Faster R-CNN	92.5	94.2	95.4	80.4	97.8	90.2	93.1	41.2	25.5
Co-Deformable-DETR-R50	98.1	92.7	97.2	75.3	96.9	89.8	93.2	21.6	64.2
SSD	91.9	91.7	90.3	74.7	90.4	85.5	88.2	24.4	80.3
YOLOv3	94.9	94.5	95.1	76.3	91.0	85.4	92.4	51.6	77.6
YOLOv5	95.2	95.5	95.6	81.4	93.6	87.0	94.6	21.2	102.8
YOLOX	97.8	97.5	98.0	90.8	97.6	90.3	95.1	9.49	99.8
YOLOv7	96.1	96.3	96.1	83.4	94.7	90.0	96.0	17.9	150.9
YOLOv8 (n)	99.5	99.3	99.5	97.9	99.5	93.9	98.3	6.0	330.4
YOLOv8-CTGI(Our)	99.5	99.5	99.6	98.0	99.5	94.1	98.4	5.9	336.8
YOLOv9 (n)	99.5	99.5	99.5	85.3	99.5	90.9	96.3	5.8	360.4
YOLOv9-CTGI(Our)	99.5	99.5	99.5	85.8	99.5	91.2	96.5	5.8	369.8
YOLOv11	99.2	99.5	99.4	82.3	99.4	92.4	95.7	5.5	390.8
YOLOv13(n)	99.4	99.5	99.2	80.0	99.3	89.3	94.5	5.2	412.4
YOLO-CTGI	99.5	99.5	99.3	79.9	99.5	95.9	95.9	5.7	405.8

根据表 3 结果分析, YOLO-CTGI 模型整体性能较基线 YOLOv13 模型有提升: mAP_{50} 达 95.9%, 较 YOLOv13 的 94.5% 实现 1.4% 绝对提升、约 1.5% 相对提升。该模型在 fastener2_broken (扣件 2 破损)、missing (缺失件)、trackbed_stuff (道床异物) 等关键类别的 AP 值上优于原 YOLOv13 基线模型, 其中 missing 类 AP 提升 0.2%、trackbed_stuff 类 AP 提升 6.6%, 证明其对复杂背景干扰下缺陷目标的识别能力更强。YOLO-CTGI 以较小代价解决了基线在 missing、trackbed_stuff 等缺陷上的检测精度痛点, 且保持高效推理速度, 验证了改进的有效性与实用性。

图 7 为 YOLOv13 与 YOLO-CTGI 的检测结果对比。图中, 蓝, 紫, 粉, 白色边界框表示模型预测结果, 框上的数字为检测置信度。可视化对比显示, YOLO-CTGI 可检测出 YOLOv13 漏检的道床异物小目标, 检测框与 fastener2_broken (扣件 2 破损) 真实缺陷区域置信度更优; 进一步证实该改进有效提升检测性能, 更适用于铁轨紧固件缺陷检测任务, 具备较强实用价值。



图 7 缺陷检测对比

Fig. 7 Comparison of defect detection results

3.5 实验及分析

为验证本文算法的泛化能力，在同类的铁轨紧固件缺陷 Inspection_Railway 数据集与公共数据集 PASCAL VOC2007 上，对 YOLOv13 算法与 YOLO-CTGI 算法进行了对比实验。Inspection_Railway 数据集包含弹条、锚固螺栓、弹条缺失、锚固螺栓缺失四类目标；PASCAL VOC2007 涵盖多种常见动物类别。实验结果如表 3 所示：

表 4 在两个数据集上的实验结果对比

Tab.4 Comparative of experimental results on two datasets

Datasets	Model	$mAP_{50}/\%$
Inspection_Railway	YOLOv13	88.6
	YOLO-CTGI	90.3
VOC2007	YOLOv13	65.3
	YOLO-CTGI	65.5

本文提出的 YOLO-CTGI 在 Inspection_Railway 上 mAP_{50} 达到 90.3%，较 YOLOv13 提升 1.7%；在 PASCAL VOC2007 上 mAP_{50} 为 65.5%，提升 0.2%。实验结果表明，YOLO-CTGI 在同类型数据集与公共数据集上均具有良好的泛化能力与检测性能。

为验证本文算法中各改进点对提升缺陷检测精度的有效性，在 *Inspection_Railway* 数据集上进行了消融实验。消融实验均在统一实验环境下进行，所有改进方案的实验条件保持一致，实验结果如表 4 所示：

表 5 消融实验

Tab.5 Ablation experiment

group	CGBlock_Down	C3k2_DBB	iEMA	mAP ₅₀ /%	Params/M	FPS/(f/s)
1	×	×	×	88.6	3.02	300.4
2	✓	×	×	89.3	2.84	281.6
3	×	✓	×	89.5	3.11	267.9
4	×	×	✓	88.8	3.02	313.4
5	✓	✓	×	90.1	3.14	274.5
6	✓	✓	✓	90.3	3.14	286.1

实验 1 以 YOLOv13 为基线模型，通过系统消融实验验证了各改进模块的有效性及其协同机制：

单模块消融结果显示：CGBlock_Down 模块通过增强多尺度特征提取能力，使 mAP₅₀ 提升 0.7%，参数量优化至 2.84M，但引入轻微计算冗余导致帧率降至 281.6 FPS；C3k2_DBB 模块进一步强化高阶特征表征，将 mAP₅₀ 提升至 89.5%(+0.9%)，参数量微增至 3.11M，推理延迟同步增加(267.9 FPS)；iEMA 模块在保持参数量不变(3.02M)的前提下，通过动态特征更新机制优化推理效率，使帧率提升至 313.4 FPS，同时维持 88.8% 的 mAP₅₀，验证了其高效的性能优化特性。

模块协同验证表明：CGBlock_Down 与 C3k2_DBB 构建的多尺度特征提取与高阶特征表征功能通路，实现了检测精度的协同提升，但伴随计算复杂度累积；引入 iEMA 模块后形成闭环优化机制，其通过提升特征传播效率抵消双模块带来的计算开销，在不增加参数量的前提下实现精度与速度动态均衡。

最终集成模型达成性能突破，在保持高推理效率的同时实现 mAP₅₀ 提升，验证了所提模块组合的合理性与工程实用价值，为兼顾检测精度与部署效率的目标检测模型设计提供了有效范式。

4 结论

本文通过引入 CGBlock_Down、C3k2_DBB 与 iEMA 模块对 YOLOv13 基线检测算法进行优化，构建出适用于铁轨紧固件缺陷检测的改进算法 YOLO-CTGI，在铁轨紧固件缺陷数据集上开展了系统的实验评估，得出如下结论：

1) 所提改进模块可有效增强模型的多尺度特征提取与表征能力。在 *fastener-defect-detection* 数据集上，YOLO-CTGI 的 mAP₅₀ 较 YOLOv13 提升 1.4%，且通过结构优化与重参数化技术，使参数量仅增加 0.5M，维持近 400 f/s 的高推理速度，实现检测精度与效率的协同优化。

2) 消融实验表明，CGBlock_Down 与 C3k2_DBB 可分别提升 mAP₅₀ 0.7% 与 0.9%，但会引入少量计算开销；iEMA 模块可在不增加参数量的情况下，同时进一步提升 0.2% 的 mAP₅₀，有效平衡精度与速度。

3) 改进算法在极端场景（如微小异物、严重遮挡场景）下的缺陷检测中仍存在漏检与定位偏差。后续可研究基于注意力机制的多尺度特征增强策略，并进一步探索模型的轻量化设计方案，进一步提升实际工程适用性与鲁棒性。

参考文献:

- [1] 陈东生, 田新宇. 中国高速铁路轨道检测技术发展[J]. 铁道建筑, 2008, 48(12): 82-86.
CHEN D S, TIAN X Y. Development of check and test technologies for track structure of domestic high-speed railway[J]. Railway Engineering, 2008, 48(12): 82-86.
- [2] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [4] 裴莹玲. 基于改进 Faster R-CNN 的高铁扣件检测算法研究[D]. 南昌: 华东交通大学, 2022.
PEI Y L. A railway fastener detection method based on modified faster R-CNN[D]. Nanchang: East China Jiaotong University, 2022.
- [5] 梁楠, 张伟, 刘洋龙, 等. 基于改进 Faster R-CNN 的高铁扣件弹条缺陷检测[J]. 太赫兹科学与电子信息学报, 2024, 22(11): 1221-1227, 1269.
LIANG N, ZHANG W, LIU Y L, et al. Fastener clips defect detection based on improved Faster R-CNN in high-speed railway[J]. Journal of Terahertz Science and Electronic Information Technology, 2024, 22(11): 1221-1227, 1269.
- [6] ZHAO Q J, FANG S C, LI Y Y, et al. Track defect detection based on improved YOLOv5s[J]. IEEE Transactions on Industrial Informatics, 2025, 21(4): 3346-3355.
- [7] 鄢化彪, 刘秋平, 黄绿娥, 等. 基于模型剪枝的轻量级缺陷检测方法研究[J]. 激光与光电子学进展, 2025, 62(20): 2015002.
YAN H B, LIU Q P, HUANG (L /Lü)E, et al. Lightweight defect detection method based on model pruning[J]. Laser & Optoelectronics Progress, 2025, 62(20): 2015002.
- [8] MA S B, ZHAO X, WAN L, et al. A lightweight algorithm for steel surface defect detection using improved YOLOv8[J]. Scientific Reports, 2025, 15: 8966.
- [9] WANG Z, LI T. A lightweight CNN model based on GhostNet[J]. Computational Intelligence and Neuroscience, 2022, 2022(1): 8396550.
- [10] 朱泓宇, 程换新, 骆晓玲. 改进 YOLOv8 网络在绝缘子缺陷检测中的应用[J]. 水电能源科学, 2024, 42(5): 183-187.
ZHU H Y, CHENG H X, LUO X L. Application of improved YOLOv8 network in insulator defect detection[J]. Water Resources and Power, 2024, 42(5): 183-187.
- [11] LEI M Q, LI S Q, WU Y H, et al. YOLOv13: real-time object detection with hypergraph-enhanced adaptive visual perception[EB/OL]. 2025: arXiv: 2506.17733. <https://arxiv.org/abs/2506.17733>
- [12] WU T Y, TANG S, ZHANG R, et al. CGNet: a light-weight context guided network for semantic segmentation[J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [13] DING X H, ZHANG X Y, HAN J G, et al. Diverse branch block: building a convolution as an inception-like unit[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE, 2021: 10881-10890.
- [14] 朱强军, 胡斌. 基于 IEMA 和混合小目标的 YOLOv8s 交通标志检测方法[J]. 湖北民族大学学报(自然科学版), 2025, 43(2): 202-209.
ZHU Q J, HU B. The YOLOv8s traffic sign detection method based on IEMA and mixed small targets[J]. Journal of Hubei Minzu University (Natural Science Edition), 2025, 43(2): 202-209.
- [15] ZHANG Z Y, LIU Y P, LIU T C, et al. DAGN: a real-time UAV remote sensing image vehicle detection

- framework[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 17(11): 1884-1888.
- [16] YAO Y, ZHANG Z X, NI X, et al. CGNet: detecting computer-generated images based on transfer learning with attention module[J]. Signal Processing: Image Communication, 2022, 105: 116692.
- [17] 魏梓栩. 变电设备带电状态可视化检测关键技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2025.
WEI Z X. Research on key technologies for visualized detection of energized state in substation equipment[D]. Harbin: Harbin Institute of Technology, 2025.
- [18] 张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报, 2019, 42(3): 453-482.
ZHANG S, GONG Y H, WANG J J. The development of deep convolution neural network and its applications on computer vision[J]. Chinese Journal of Computers, 2019, 42(3): 453-482.
- [19] 孙俊, 谭文军, 毛罕平, 等. 基于改进卷积神经网络的多种植物叶片病害识别[J]. 农业工程学报, 2017, 33(19): 209-215.
SUN J, TAN W J, MAO H P, et al. Recognition of multiple plant leaf diseases based on improved convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33(19): 209-215.
- [20] 陈思雨, 付章杰. 融合高效注意力的多尺度输电线路部件检测[J]. 计算机工程与应用, 2024, 60(1): 327-336.
CHEN S Y, FU Z J. Multi-scale transmission line component detection incorporating efficient attention[J]. Computer Engineering and Applications, 2024, 60(1): 327-336.
- [21] 余卓凡. 面向小目标检测的 RT-DETR 轻量化与检测精度优化研究[D]. 武汉: 湖北工业大学, 2024.
YU Z F. Optimization of RT-DETR Lightweighting and Detection Accuracy for Small Target Detection. Wuhan: Hubei University of Technology, 2024.
- [22] 张舞杰, 李迪, 叶峰. 基于 Sigmoid 函数拟合的亚像素边缘检测方法[J]. 华南理工大学学报(自然科学版), 2009, 37(10): 39-43.
ZHANG W J, LI D, YE F. Sub-pixel edge detection method based on sigmoid function fitting[J]. Journal of South China University of Technology (Natural Science Edition), 2009, 37(10): 39-43.



第一作者: 周会祥 (1983—), 男, 副教授, 博士, 硕士生导师, 研究方向为深度学习、机器视觉、目标检测。E-mail: zhx@ecjtu.edu.cn。



通信作者: 曹莞菁 (1999—), 女, 硕士研究生, 研究方向为目标检测。E-mail: caowanjing0508@163.com。