

基于改进 YOLOv10n 的钢轨表面缺陷检测*

刘 珮 刘伯鸿

(兰州交通大学自动化与电气工程学院 兰州 730070)

摘要: 为提升钢轨表面缺陷检测的准确性与小目标识别能力,设计了一种基于 YOLOv10n 的轻量化目标检测模型。该模型在浅层骨干网络引入 C2f_CGBlock,以增强局部上下文感知与特征表达能力;特征融合部分采用 RepGFPN,并在部分回传路径中引入 SimAM 以突出关键信息。模型训练中使用 Inner-SIoU 损失函数以优化定位精度。实验在钢轨表面缺陷数据集上进行验证,结果显示,改进模型在 Precision、Recall、F1 和 mAP@0.5 等指标上分别提升了 3.38%、3.72%、3.55% 和 4.01%,相比基准模型在小尺寸缺陷和复杂背景下的检测效果有明显提升。该模型在兼顾检测精度与实时性的同时,显著提升了钢轨缺陷检测性能,具备良好的工程应用前景。

关键词: 钢轨表面缺陷;目标检测;YOLOv10n;CGBlock;Inner-SIoU

中图分类号: TP391.7;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Rail surface defect detection based on improved YOLOv10n

Liu Pei Liu Bohong

(School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: This paper proposes a lightweight object detection model based on YOLOv10n to improve the accuracy of rail surface defect detection and enhance the recognition of small targets. The model incorporates C2f_CGBlock into the P3 and P4 layers of the backbone network to strengthen local context perception and feature representation. The feature fusion part uses RepGFPN and integrates SimAM into some feedback paths to emphasize critical features. The training process adopts Inner-SIoU loss function to optimize localization accuracy. Experimental results on a rail surface defect dataset showed that the improved model outperformed the original one, with improvements of 3.38%, 3.72%, 3.55% and 4.01% in Precision, Recall, F1 and mAP@0.5. The model demonstrates clear advantages over the baseline in detecting small-size defects and challenging backgrounds. It effectively enhances the performance of rail defect detection while maintaining a balance between accuracy and real-time efficiency, and has good potential for engineering applications.

Keywords: rail surface defects; object detection; YOLOv10n; CGBlock; Inner-SIoU

0 引言

钢轨作为承担车辆车轮巨大压力、保障运行安全性和有效性的关键部件,容易因环境因素和外力作用发生表面缺陷。轻微缺陷影响旅行舒适度,严重缺陷则会造成经济损失和人员伤亡,因此,钢轨表面缺陷的研究具有重要意义。目前常用的钢轨表面缺陷检测方法包括超声波检测、漏磁检测、涡流检测等。超声波检测基于脉冲反射法,主要探测钢轨内部缺陷,但对钢轨表面的微小裂纹检测效果不佳^[1];漏磁检测易受趋肤效应影响,检测精度下降^[2];涡流检测利用电磁感应原理,根据磁场变化信息进行判断,受提离效应影响存在漏检现象^[3]。上述提到的钢轨表面缺陷检

测方法在一定条件下会影响缺陷检测的准确性和效率。

随着计算机技术的发展,基于深度学习和视觉检测的检测手段应运而生并得到大力发展,并已在钢材、带钢等表面缺陷检测中取得一定成果。张航等^[4]引入注意力机制、改进骨干网络和优化特征融合策略,提高了微小裂纹和凹陷的检测精度和速度。针对钢材表面缺陷检测效率低和精度不足的问题,赵佰亭等^[5]引入 ConvNext、C2fFB 及 MPCE 模块,在保持轻量化的同时显著提升了微小裂缝检测精度。然而由于钢轨服役环境复杂且纹理背景多样,其缺陷检测对小目标识别能力提出了更高要求。

在钢轨表面缺陷检测领域中,刘俊博等^[6]针对少样本数

据集设计样本随机组合策略,并改进损失函数以提升少样本情况下钢轨表面缺陷检测性能。吴永军等^[7]结合傅里叶变换进行语义缺陷增广,针对小目标缺陷增加了低层特征融合与额外检测头,显著增强了网络对小缺陷的定位能力。闵永智等^[8]提出 Attention-DCGAN 算法扩充钢轨表面图像,提升图像多样性。陈仁祥等^[9]针对图像降噪,提出了基于注意力引导的多尺度降噪卷积神经网络,采用多尺度卷积、跳跃连接、注意力机制等方式构建去噪模块。罗晖等^[10]使用改进后的 Retinex 算法处理钢轨表面缺陷样本,通过图像处理技术来提升缺陷与背景的对比度,进而实现图像增强。杜少聪等^[11]在 YOLOv5 的 Backbone 末端嵌入多头自注意力层,构建跨层加权级联结构,实现对裂纹、剥落、磨损等缺陷的精准检测。

尽管上述算法在图像增强、缺陷分类与识别方面有所改善,但钢轨表面小目标缺陷仍存在检测准确率低、检测器难以适应不同缺陷类型以及模型的泛化性能与鲁棒性有待提升等问题。因此在复杂背景下如何提升钢轨小目标缺陷的检测性能,仍是当前亟需解决的关键问题,也是本文研究

工作的主要出发点。

1 YOLOv10 概述

YOLOv10^[12]于 2024 年发布,具有端到端检测和实时高效的优点,广泛应用于目标检测领域。该算法继承 YOLOv8 的基本思路,网络结构包括输入(Input)、骨干网络(Backbone)、颈部网络(Neck)和检测头(Head)。相比 YOLOv8, YOLOv10 在骨干网络中引入空间-通道解耦下采样(spatial-channel decoupled downsampling, SCDown)和部分自注意力机制模块(partial self-attention, PSA),以增强下采样与特征提取能力,同时保持模型轻量化。颈部网络中的 C2f 结构引入紧凑倒置块(compact inverted block, CIB),优化梯度流传递而不增加计算负担。检测头部分取消了非极大值抑制(non-maximum suppression, NMS),采用一对一(one-to-one)和一对多(one-to-many)的检测策略,提高检测效率和精度。本研究选用 YOLOv10n 作为基准算法,其主要模块结构如图 1 所示。

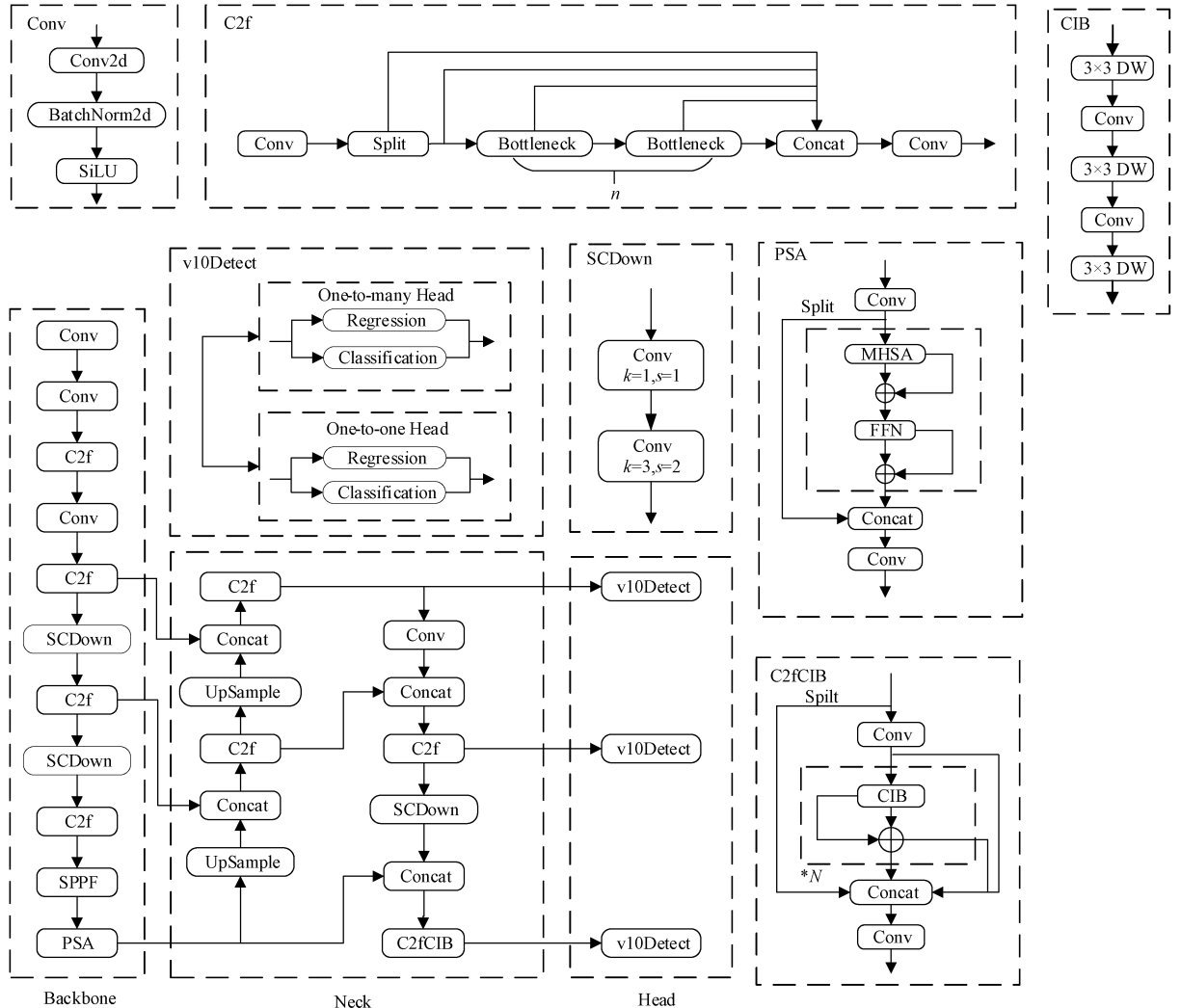


图 1 YOLOv10n 及各部分结构
Fig. 1 YOLOv10n and its structures

2 YOLOv10n 改进

YOLOv10n 在保持轻量化的同时,具备高效捕捉小目标特征的能力,适用于钢轨表面小尺寸缺陷的检测任务。为解决现有检测方法在准确率和模型适应性方面的不足,本研究对 YOLOv10n 进行结构改进,具体为:

1)对 Backbone 中 P3、P4 的 C2f 模块进行结构轻量化优化,引入上下文引导模块(context guided block, CGBlock),构建 C2f_CGBlock 模块,从而增强模型对局部细节与全局语义的建模能力,提高模型对中小尺寸缺陷的感知能力。

2)颈部网络借鉴 DAMO-YOLO 的重参数化泛化特征金字塔网络(re-parameterized generalized FPN, RepGFPN)进行重构,在 P4→P3 的融合路径中对相应的 CSPStage 进行修改,在输出端引入简单无参数注意力模块(simple, parameter-free attention module, SimAM),通过残差加权方式强化小目标保持,从而提升小目标检测表现。

3)将完全交并比(complete intersection over union, CIoU)替换为 Inner-SIoU(Inner Scylla-IoU),以提升目标框回归的精度与收敛速度,使检测器在不同类型缺陷目标下具备更强的泛化能力和自适应性。

改进后的模型结构如图 2 所示。

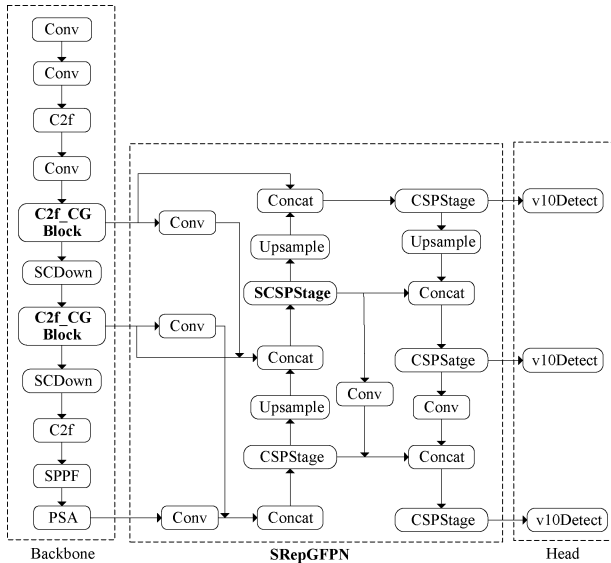


图 2 YOLOv10n 改进结构

Fig. 2 Improved structure of YOLOv10n

2.1 C2f_CGBlock

钢轨表面缺陷通常尺寸较小,纹理细微且分布不均匀,容易出现漏检或误检。CGBlock^[13]是上下文引导网络(context guided network,CGNet)中的核心模块,能够对局部特征、周围上下文信息及全局语义信息进行统一建模,从而实现更具判别性的特征表达。

输入特征经过 1×1 卷积进行通道压缩,一路通过标准 3×3 卷积与通道卷积组成的局部特征提取模块 $f_{loc}(*)$,精准捕获钢轨表面细小裂纹、凹痕等局部缺陷细节;另一路通过 3×3 空洞卷积与通道卷积构成的周围上下文提取模块 $f_{sur}(*)$,在不损失分辨率的前提下扩大感受野,以建立缺陷区域与周围纹理的空间联系,从而降低误检率。两路特征在通道维度上进行拼接(concatenation, Concat),通过批归一化(batch normalization, BN)与激活函数(parametric rectified linear unit, PReLU)构成的联合特征提取模块 $f_{join}(*)$,融合处理后获得兼具局部细节和上下文语义的复合特征。为进一步提升全局信息的调控能力,CGBlock 引入全局上下文提取模块 $f_{glo}(*)$,通过全局平均池化(global average pooling, GAP)和全连接层(fully connected layer, FC),对融合特征进行加权重标定,自适应学习并强化与各类缺陷最相关的特征通道,从而强化关键区域响应、抑制冗余信息。这种由局部至全局的多尺度特征融合机制,从理论上契合了钢轨缺陷尺度多样、纹理复杂、上下文依赖强的特性,能够有效提升模型对多类缺陷的检测性与鲁棒性。CGBlock 结构如图 3 所示。

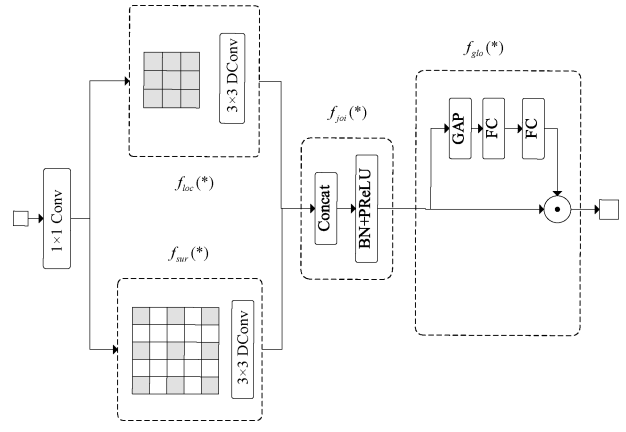


图 3 CGBlock 结构

Fig. 3 CGBlock structure

相较于原始跨阶段特征模块(cross stage partial bottleneck with 2 convolutions, C2f)中采用 Bottleneck 所带来的模型复杂度问题,将其替换为 CGBlock 不仅有效降低参数量和计算量,同时增强了对钢轨中小目标的表征能力与上下文适应性。C2f_CGBlock 结构如图 4 所示。

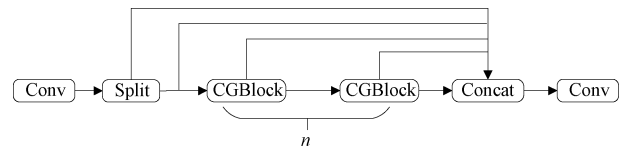


图 4 C2f_CGBlock 结构

Fig. 4 C2f_CGBlock structure

2.2 SRepGFPN

钢轨表面小目标缺陷容易在特征下采样和多层融合

过程中被背景信息掩盖,导致漏检或定位不准。为解决在复杂背景下小目标检测信息在回传过程中衰减的问题,本文以 DAMO-YOLO 的 RepGFPN^[14] 为基础,对内部特定的 CSPStage 结构进行优化。

针对钢轨缺陷在深层特征中表现为低响应、低对比的特点,传统卷积特征在上采样回传阶段容易被背景强响应淹没,难以保持小目标信息的显著性。为此,在 P4→P3 的回传融合路径中,于 CSPStage 输出端引入 SimAM,构建 SCSPStage 模块,进而形成优化结构 SRepGFPN。

SimAM^[15] 基于能量函数建模神经元响应强度,无需额外参数即可在空间维度显著增强特征响应,其核心思想是通过能量最小化机制自适应突出关键区域响应、抑制背景干扰,对每个通道内二维特征图独立建模,从而有效提升小目标特征的保持与区分能力。能量函数定义如式(1)所示。

$$e_i(\omega_i, b_i, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} [-1 - (\omega_i x_i + b_i)]^2 + [1 - (\omega_i t + b_i)]^2 + \lambda \omega_i^2 \quad (1)$$

式中: ω_i, b_i 为某个神经元进行线性变换时的权重与偏差, y 为标签,代表是否为重要神经元, x_i 为其他神经元, i 为空间维度上的索引, M 为某个通道上所有神经元的个数, λ 为正则化系数。此公式具有如下解析解,如式(2)、(3)所示。

$$\omega_i = -\frac{2(t - \mu_i)}{(t - \mu_i)^2 + 2\sigma_i^2 + 2\lambda} \quad (2)$$

$$b_i = -\frac{1}{2}(t + \mu_i)\omega_i \quad (3)$$

其中, μ_i 为均值, σ_i^2 为方差。

将式(2)、(3)代入式(1),可得最小能量如式(4)所示。

$$e_i^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (4)$$

$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i \quad (5)$$

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2 \quad (6)$$

其中, e_i^* 为最小能量。从统计学的角度来看, μ_i 可由式(5)代替, σ_i^2 可由式(6)代替,以减少计算量。上述公式证明:能量越低,神经元 t 与周围神经元区别越大,重要性越高。

SimAM 结构如图 5 所示。

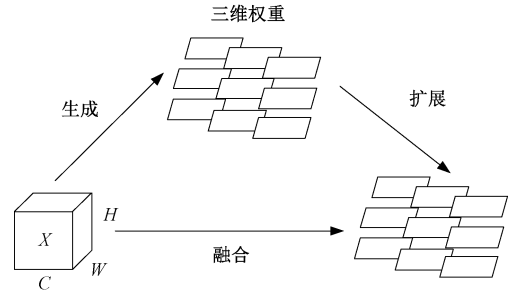


图 5 SimAM 结构

Fig. 5 SimAM structure

按照注意力机制的定义,对相应特征进行增强处理,如式(7)所示。

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (7)$$

式中: E 为特征图 X 上的最小能量。

SimAM 的引入提升了 CSPStage 对小目标特征的保持与增强能力,同时几乎不增加计算开销,保持了结构的高效性与轻量化。SCSPStage 的结构如图 6 所示。

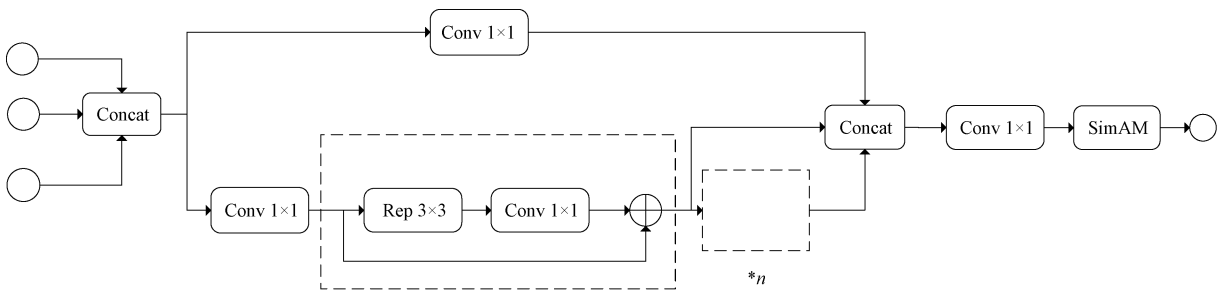


图 6 SCSPStage 结构

Fig. 6 SCSPStage structure

同时在 CSPStage 中引入残差加权结构,对 SimAM 进行自适应调控,避免注意力机制在训练前期存在过强扰动导致训练不稳定的现象,从而进一步提升注意力在不同检测环境下的适应性。模型可通过可学习标量参数 α ,以残差形式融合注意力增强特征与主干输出,如式(8)所示。

$$y = x + \alpha \times (\text{attention} - x) = (1 - \alpha)x + \text{attention} \times \alpha \quad (8)$$

式中: α 为标量参数,attention 为注意力机制算子, x 为输入特征, y 为输出特征。

2.3 损失函数

钢轨表面缺陷的自身特性使其检测易被背景信息干扰,导致预测框偏离或收敛速度慢。传统 CIoU 在这种情况下难以充分聚焦小目标特征,影响检测精度。因此本文采用 Inner-SIoU 损失函数,以增强模型对小目标缺陷的回

归能力。

从理论上分析,钢轨小目标缺陷的检测框面临两个核心挑战:一是“中心点敏感问题”,即预测目标点的微小偏移会显著降低交并比;二是缺陷形状多样,需要损失函数能更精细地引导预测框的尺寸和形状收敛。

在 Inner-SIoU 中的 SiOU^[16],引入预测框和真实框之间的向量夹角并考虑二者方向,通过角度损失(angle cost)、距离损失(distance cost)、形状损失(shape cost)和 IoU 损失(IoU cost)来进行损失函数的计算,能够针对小目标缺陷的中心敏感性和形状多样性问题提供有效回归指导。

针对中心点敏感的问题,角度损失计算真实框与预测框中心点的角度差异,能够有效约束预测框方向,指导模型更精准地回归小目标的位置。角度损失如图 7 所示。

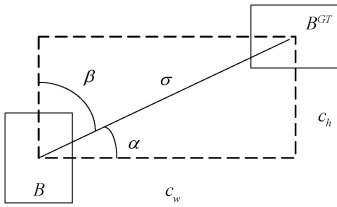


图 7 角度损失
Fig. 7 Angle cost

计算如式(9)~(12)所示。

$$\Lambda = 1 - 2 \times \sin^2 \left(\arcsin \left(\frac{c_h}{\sigma} \right) - \frac{\pi}{4} \right) \quad (9)$$

$$\frac{c_h}{\sigma} = \sin(\alpha) \quad (10)$$

$$\sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2} \quad (11)$$

$$c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y}) \quad (12)$$

式中: b 为真实框, b^{gt} 为预测框, σ 为真实框和预测框中心点的距离, $(b_{c_x}^{gt}, b_{c_y}^{gt})$ 为真实框的中心坐标, (b_{c_x}, b_{c_y}) 为预测框的中心坐标, c_h, c_w 为真实框与预测框中心点的高度差与宽度差。

距离损失指预测框与真实框中心坐标的欧式距离,可减少小目标中心点偏移造成的误差,如图 8 所示。

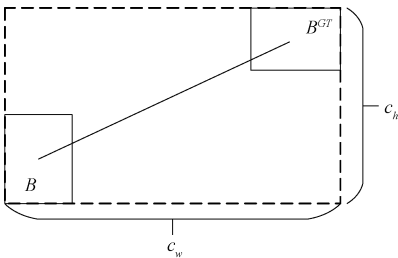


图 8 距离损失
Fig. 8 Distance cost

计算如式(13)、(14)所示。

$$\Delta = \sum_{t=x,y} (1 - e^{-(2-\Lambda) \times \rho_t}) \quad (13)$$

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2 \quad (14)$$

式中: c_w, c_h 是真实框与预测框所形成的最小外接矩形的宽和高。

形状损失则关注宽高比不匹配的问题,如式(15)、(16)所示。

$$\Omega = \sum_{t=w,h} (1 - e^{-W_t})^\theta \quad (15)$$

$$W_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})}, W_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (16)$$

式中: ω, h 分别为预测框的宽和高, ω^{gt}, h^{gt} 分别为真实框的宽和高, θ 代表对形状损失的关注程度。

SiOU 将距离损失与形状损失同时计算,有助于模型独立优化预测框的位置和形状,使其更灵活地适应钢轨表面不同缺陷的几何特性,实现更精准的定位。

交并比(intersection over union, IoU)定义为真实框、预测框的交集部分与并集部分的比值,如图 9 所示。

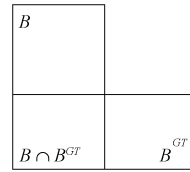


图 9 IoU 损失
Fig. 9 IoU cost

IoU 计算如式(17)所示。

$$IoU = \frac{|B \cap B^{GT}|}{|B \cup B^{GT}|} \quad (17)$$

式中: $B \cap B^{GT}$ 为真实框与预测框的交集面积表示, $B \cup B^{GT}$ 为真实框与预测框的并集面积表示,可得损失函数 SiOU 计算公式,如式(18)所示。

$$L_{SiOU} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (18)$$

尽管 SiOU 能综合考虑角度、距离、形状和 IoU 损失,对预测框进行精细回归,但在实际钢轨缺陷检测中,它无法根据不同目标自动调整对小目标或重叠目标的敏感性,容易导致收敛速度不够快或小目标回归不够准确。针对这一问题,本研究在 SiOU 的基础上引入 Inner 思想进行改进,形成新损失函数 Inner-SIoU。

Inner-IoU^[17] 的核心思想是借助比例因子 $ratio$ 来控制辅助边界框的尺寸,从而针对不同 IoU 样本采用尺度不同的辅助边界框进行损失计算。该方法不仅提高了对重叠部分的评估,而且加快了模型收敛速度并提升性能。Inner-IoU 如图 10 所示,其中图 10(a)和(b)分别表示 $ratio \in [0.5, 1)$ 与 $ratio \in (1, 1.5]$ 时,真实框、辅助边界框、预测框与辅助预测框的相对大小关系。 (x_c^{gt}, y_c^{gt}) 为真实框的中心坐标, ω, h 分别代表真实框的宽、高; $\omega_{inner}^{gt}, h_{inner}^{gt}$ 为辅

助边界框的宽、高; (x_c, y_c) 为预测框的中心坐标, w, h 为预测框的宽、高; w_{inner}, h_{inner} 为辅助预测框的宽、高。

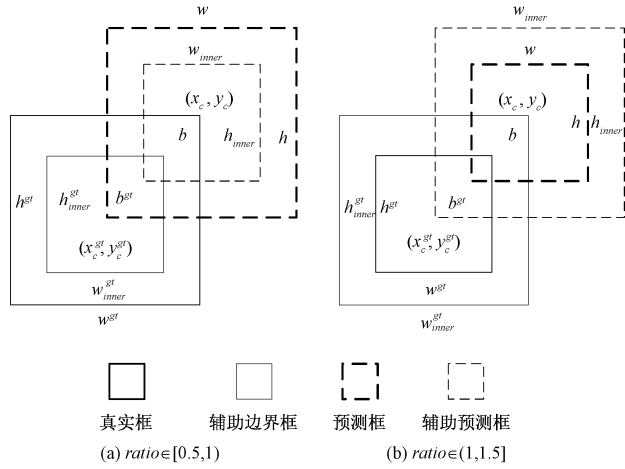


图 10 损失函数 Inner-IoU

Fig. 10 Loss function Inner-IoU

当 $ratio \in [0.5, 1.5]$ 时,根据式(19)、(20)可得辅助边界框的 4 个顶点坐标即 (b_l^{gt}, b_r^{gt}) 、 (b_t^{gt}, b_b^{gt}) 、 (b_r^{gt}, b_t^{gt}) 、 (b_r^{gt}, b_b^{gt}) :

$$b_l^{gt} = x_c - \frac{w^{gt} \times ratio}{2}, b_r^{gt} = x_c + \frac{w^{gt} \times ratio}{2} \quad (19)$$

$$b_t^{gt} = y_c - \frac{h^{gt} \times ratio}{2}, b_b^{gt} = y_c + \frac{h^{gt} \times ratio}{2} \quad (20)$$

根据式(21)、(22)可得辅助预测框的 4 个顶点坐标即 (b_l, b_t) 、 (b_l, b_b) 、 (b_r, b_t) 、 (b_r, b_b) :

$$b_l = x_c - \frac{w^{gt} \times ratio}{2}, b_r = x_c + \frac{w \times ratio}{2} \quad (21)$$

$$b_t = y_c - \frac{h^{gt} \times ratio}{2}, b_b = y_c + \frac{h \times ratio}{2} \quad (22)$$

根据所计算的辅助边界框、辅助预测框顶点坐标,可得二者重叠区域,如式(23)所示。

$$Inter = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) \times (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)) \quad (23)$$

亦可计算辅助边界框与辅助预测框的合并区域,如式(24)所示。

$$Union = (w^{gt} \times h^{gt}) \times (ratio)^2 + (w \times h) \times (ratio)^2 - Inter \quad (24)$$

根据式(23)、(24),结合 IoU 定义,可得 Inner-IoU,如式(25)所示。

$$IoU^{inner} = \frac{Inter}{Union} \quad (25)$$

根据式(18)、(25)可计算 Inner-SIoU,如式(26)所示。

$$Loss_{inner-SIoU} = L_{SIoU} + IoU - IoU^{inner} \quad (26)$$

3 实验环境与超参数设置

3.1 实验数据集

目前暂无权威的钢轨表面缺陷数据集,因此本实验自

建缺陷数据集,样本来源于北京交通大学公布的 RSDDs 数据集和 Kaggle 数据集。

RSDDs 数据集^[18]的拍摄以工业 CCD 相机为主,图像分为两类:Type-I RSDDs 数据集,采集自 express rails(高速铁路),包括 67 张图像,均为 $160 \text{ pixel} \times 1\,000 \text{ pixel}$; Type-II RSDDs 数据集,采集自 common/heavy haul rails(常规铁路/重载铁路),包括 128 张图像,均为 $55 \text{ pixel} \times 1\,250 \text{ pixel}$ 。Kaggle 数据集对荷兰沿线铁路的状态进行拍摄。由于 RSDDs 缺乏 crack 类样本,为实现 3 类缺陷样本均衡,本研究从 Kaggle 数据集中补充 crack 类样本,筛选后合并共得到 393 张原始样本图像。

鉴于缺陷图像在尺寸上的显著差异,本文采用随机裁剪后统一缩放 $100 \text{ pixel} \times 500 \text{ pixel}$ 的方式进行尺寸归一化处理。一方面,随机裁剪可增加样本的多样性,使模型学习不同局部区域的缺陷多尺度特征;另一方面, 100×500 的分辨率兼顾了钢轨表面的纵向长条结构与缺陷细节保真度,且尺寸适中,同时与主流检测模型(如 YOLO 系列)的输入分辨率匹配,从而在保证特征完整的前提下提升模型训练效率与检测稳定性。随后对图像进行灰度统一、图像尺寸统一、图像镜像、图像滤波、图像加噪等增强操作,共获得 1 530 张缺陷图像,缺陷划分为 indentation(压痕)、crack(裂纹)、scratch(划痕)这 3 类,使用 Labelimg 进行标注,保存为 YOLO 风格的 txt 文件,并按照训练集:验证集:测试集 = 8:1:1 的比例划分样本。RSDDs 与 Kaggle 数据集均为典型的缺陷导向型数据集,其原始图像中未包含专门采集的无缺陷样本,但其图像中广泛分布着大量完好的钢轨背景区域,呈现出复杂的纹理结构和显著的光照变化,为模型训练提供了丰富且真实多样的背景上下文。鉴于这一特性,本研究在训练阶段未额外引入独立的正常样本。这一“以缺陷为中心”的数据构建策略具有多重优势:它迫使模型聚焦于从复杂背景中挖掘缺陷的细微判别特征,从而避免对背景模式的过拟合;缺陷图像内部天然存在的无缺陷区域可作为隐式负样本,在空间邻近性和视觉连续性上未特征学习提供强对比信号,有助于提升模型在局部纹理扰动下的鲁棒性。数据集最终构成如表 1 所示。

表 1 数据集最终构成

Table 1 The final composition of the dataset 个				
类别	训练集	验证集	测试集	小计
缺陷图像	1 224	153	153	1 530

由于多数图像包含多类缺陷目标,数据集分布统计以标签框数量为单位,3 类缺陷的具体数量如表 2 所示。在测试集中 3 类缺陷目标均有分布且每类缺陷不少于 65 个,以保证评估结果的代表性与稳定性。

进一步统计发现,按目标面积占整幅图像比例不超过

表 2 每类缺陷对应的标签框数量

Table 2 Number of bounding boxes per defect type 个

种类	训练集	验证集	测试集	总个数
crack	659	78	85	822
indentation	640	65	87	792
scratch	511	79	65	658

2%的标准划分,小目标在数据集中占据显著比例。其中,indentation(压痕)目标几乎全部为小目标,约96%,scratch类(划痕)约占30%,而crack(裂纹)约占5%。因此,本研究的数据集具有典型的小目标特征,对检测模型的多尺度特征提取与精细回归能力提出了较高要求。

从视觉特征与检测难点上分析:indentation(压痕)通常尺寸较小、局部凹陷明显,其检测难点在于细小的目标尺寸易在深层特征图中丢失,且低对比度的凹陷特征易与背景杂波混淆;crack(裂纹)呈线状或曲折状、长度不一,其检测难点在于像素占比低且断续的现状结构难以被有效提取,对模型的细粒度感知能力要求高;scratch(划痕)纹理细长、方向分布多样,检测难点在于微弱、浅淡的纹理特征一被忽略,且其多样的方向性要求模型具备良好的旋转不变性。

综上所述,本数据集对检测模型的多尺度特征提取、细微特征捕捉与复杂背景下的精细定位能力均提出严峻挑战。针对这些关键特性,本文提出的3项改进协同作用,为模型在复杂背景下的小目标钢轨缺陷检测提供了理论依据与实践支持。

3.2 实验环境

实验环境为Linux操作系统,使用Python编程,基于Visual Studio Code平台,Pytorch 2.2.2为深度学习框架,CPU型号为AMD EPYC 7502,GPU型号为NVIDIA GeForce RTX 3090 24 G,CUDA版本为12.1。具体的实验环境配置如表3所示。

表 3 实验环境

Table 3 Experiment environments

环境	规格
操作系统	Linux
编程语言	Python 3.10.14
编程平台	Visual Studio Code
深度学习框架	Pytorch 2.2.2
CUDA	12.1
CPU	AMD EPYC 7502
GPU	NVIDIA GeForce RTX 3090 24 G
cuDNN	cuDNN 8.9.2

3.3 超参数设置

原始图像尺寸 100×500 ,但由于该尺寸不符合YOLO

系列网络的结构设计要求,若直接以该尺寸输入可能导致特征图尺寸异常,影响模型稳定性与推理效率。为保证网络兼容性与训练稳定性,本研究采用YOLO默认的 640×640 输入尺寸,由模型自动完成letterbox填充,保持图像比例的同时适配网络结构。同时为了保证实验的可复现性,选择随机种子数为42,早停机制patience为50,初始SimAM学习参数为0.3,在Inner-SIoU中采用比例因子0.75。超参数设置如表4所示。

表 4 超参数设置

Table 4 Hyperparameter settings

参数	数值
训练批次	200
批量大小	16
工作线程	4
优化器	SGD
优化器动量	0.937
初始学习率	0.01
权重衰减	0.0005
图片尺寸	640
随机种子	42
早停机制	50
学习参数	0.3
比例因子	0.75

3.4 评价指标

评价指标包括准确率(precision, P)、召回率(recall, R)、F1指数(F1-score, F1)、平均精确度(mean average precision, mAP)、每秒帧率(frames per second, FPS)、浮点运算量(giga floating point operations per second, GFLOPs)、参数量(parameters, Params)、模型大小(Model Size)等。其中,P、R、F1和mAP等指标均以百分数形式表示;FPS、GFLOPs、参数量和模型大小则采用对应的单位进行描述。

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (27)$$

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (28)$$

$$F1 = \frac{2 \times P \times R}{P + R} \times 100\% \quad (29)$$

$$mAP = \frac{\sum_{i=1}^N \int_0^1 P(R) dR}{N} \times 100\% \quad (30)$$

式中: T_p 为被正确识别的样本数目, F_p 为被误检测的样本数目, F_n 为未被正确检测的样本数目。 N 为样本类别数目,本研究对钢轨表面缺陷划分为indentation、crack、scratch, $N=3$ 。

4 实验结果分析

4.1 消融实验

为验证各结构改进模块的有效性及其对整体性能的贡献,本文设计了多组消融实验,主要包括以下 3 类内容:

1) 模块细粒度消融:针对 C2f_CGBlock、SRepGFPN 和 Inner-SIoU 等核心模块,进一步探究其不同配置对检测性能的影响;

2) 逐步组合实验:基于 YOLOv10n 原始结构,逐步叠加各改进模块,评估整体结构在引入各模块后的性能提升过程,结合 Shapley 值分析,定量衡量各模块对模型性能的相对贡献;

3) 类别敏感性分析:以 indentation 等难检测类别为代表,考察不同改进对特定类别检测能力的提升效果。

(1) C2f_CGBlock

为验证 C2f_CGBlock 模块对模型性能的影响,本文基于 YOLOv10n 设计两种替换策略进行对比实验。2C2f_CGBlock:仅将 Backbone 中 P3、P4 的 C2f 模块替换为 C2f_CGBlock;4C2f_CGBlock:将 Backbone 中全部 C2f 模块替换为 C2f_CGBlock。实验选取 P、R、F1、mAP@0.5、mAP@0.5:0.95 等指标进行评估,如表 5 所示。

表 5 C2f_CGBlock 对比

Table 5 Comparisons of C2f_CGBlock

模型	P	R	F1	mAP@ 0.5	mAP@ 0.5:0.95
YOLOv10n	84.80	84.25	84.51	87.26	47.22
2C2f_CGBlock	86.67	80.30	83.32	88.61	46.43
4C2f_CGBlock	86.17	77.78	81.58	85.59	46.63

适量引入 C2f_CGBlock 能显著提升模型精度,P 提升 1.87%、mAP@0.5 提升 1.35%,表明该模块有助于增强浅层、中层特征的表达能力。将 C2f 全替换为 C2f_CGBlock 后,P 较高,R 由 84.25% 降至 77.78%,F1 和 mAP@0.5 有所下降,说明过度堆叠可能引入冗余语义信息,影响模型的召回能力和整体泛化性能。

局部替换策略在精度与召回之间实现了更优平衡,为后续结构优化提供了合理的模块插入方式。

(2) SRepGFPN

针对多尺度特征融合能力,本研究将 YOLOv10n 的 Neck 替换为 SRepGFPN,并引入学习参数以控制信息的融合比例。实验在不同参数设置下进行对比,评估 P、R、F1、mAP@0.5 与 mAP@0.5:0.95 等指标,其中 SRepGFPN 中参数 0.0 表示未启用 SimAM,参数 0.3 表示以 30% 的权重融合 SimAM 与原始特征,参数 1.0 表示完全采用 SimAM,记为 0.0、0.3、1.0。结果如表 6 所示。实验结果表明,SRepGFPN 能有效提升模型检测性能。在

不同参数设置下,P、R、mAP 等核心指标上均优于原始 RepGFPN。单模块参数调优实验显示,参数为 0.0 时模型在部分指标上取得最佳表现,而参数为 1.0 时性能有所下降,表明空间引导机制强度对效果存在一定影响。

表 6 SRepGFPN 对比

Table 6 Comparisons of SRepGFPN

模型	P	R	F1	mAP@ 0.5	mAP@ 0.5:0.95
YOLOv10n	84.80	84.25	84.51	87.26	47.22
RepGFPN	82.77	88.53	85.41	89.75	47.34
0.0	88.66	84.45	86.43	91.56	48.49
0.3	84.12	88.60	85.15	90.57	48.15
1.0	83.64	86.64	84.97	88.05	46.82

为进一步验证参数对整体模型的影响,本文进行了多模块组合实验。组合结果如表 7 所示。当参数为 0.3 时,模型在 F1、mAP@0.5 等指标上取得最佳整体性能,优于单模块最优参数配置。该现象表明,适度的空间引导强度更有利于多模块协同优化,提升模型的泛化能力和稳定性。

表 7 SRepGFPN 与其他模块的协同对比

Table 7 Comparisons of SRepGFPN collaborations with other modules

模型	P	R	F1	mAP@ 0.5	mAP@ 0.5:0.95
0.0	82.08	83.43	82.65	86.18	44.84
0.3	88.18	87.97	88.06	91.27	47.70
1.0	85.20	89.45	87.11	90.51	49.06

因此,综合 P、R 及与其他模块的兼容性,本文最终选择 SRepGFPN($\alpha = 0.3$)作为 Neck,为后续模型改进奠定基础。

(3) Inner-SIoU 分析

为探索 Inner-SIoU 中比例参数 *ratio* 对模型性能的影响,根据数据集缺陷目标的尺度分布特征即小目标占比高且部分样本形态较为狭长,设置 *ratio* 在 $[0.7, 0.8]$ 进行对比试验。本研究分别测试了 *ratio* = 0.7、*ratio* = 0.75 和 *ratio* = 0.8 这 3 组配置,在钢轨缺陷检测任务上进行实验。结果如表 8 所示,可用于分析不同 *ratio* 下 P、R、F1 和 mAP 指标平衡关系。

表 8 Inner-SIoU 的 ratio 对比

Table 8 Comparisons of Inner-SIoU ratio

ratio	P	R	F1	mAP@ 0.5	mAP@ 0.5:0.95
0.70	85.34	81.15	82.99	88.65	46.23
0.75	85.17	87.01	86.06	90.16	48.10
0.80	89.51	82.42	85.81	89.55	47.06

由实验数据可知, $ratio = 0.75$ 配置在 R、F1、mAP@0.5、mAP@0.5:0.95 指标上表现最优, 且 R 明显优于其他配置, 体现了更好的缺陷召回能力。虽然 $ratio = 0.8$ 在 P 上表现最高, 但 R 下降导致整体 F1 分数略低于 0.75。 $ratio = 0.7$ 未能超越 0.75 的综合性能。

综上, $ratio = 0.75$ 能较好地平衡检测的准确率与召回率, 提升模型整体表现, 因此被选为 Inner-SIoU 的最佳超参数配置。

(4) 组合性能分析

为验证各模块对模型性能的贡献, 本研究记改进 A (C2f_CGBlock)、改进 B (SRepGFPN, $\alpha = 0.3$) 及改进 C (Inner-SIoU, $ratio = 0.75$)。实验结果如表 9 所示。消融实验显示, 相较原模型, 改进后的模型在 P、R、mAP@0.5、mAP@0.5:0.95 等指标分别提升了 3.38%、3.72%、3.55%、4.01%, 且 FPS 满足实时检测需求。参数量与模型大小分别增加约 25.6%、21.8%, GFLOPs 仅增大 0.1 G,

表 9 消融实验

Table 9 Ablation experiments

模型	P/%	R/%	F1/%	mAP@0.5/%	mAP@0.5:0.95/%	GFLOPs/G	FPS	Params/MB	模型大小/M
YOLOv10n(Baseline)	84.80	84.25	84.51	87.26	47.22	6.5	190.09	8.64	5.5
YOLOv10n+A	86.87	80.30	83.32	88.61	46.43	6.1	251.45	8.11	5.2
YOLOv10n+B	84.12	88.60	86.15	90.57	48.15	7.5	249.90	11.53	7.0
YOLOv10n+C	85.17	87.01	86.06	90.16	48.10	6.5	237.65	8.64	5.5
YOLOv10n+A+B	85.51	84.66	85.08	88.58	46.82	6.5	226.65	10.36	6.4
YOLOv10n+A+C	80.48	86.38	82.65	88.34	45.14	5.6	315.81	7.97	5.1
YOLOv10n+B+C	86.82	78.90	82.65	89.67	46.20	7.5	205.93	11.53	7.0
YOLOv10n+A+B+C	88.18	87.97	88.06	91.27	47.70	6.6	212.02	10.85	6.7

表明模型仍保持较轻量特性。

钢轨表面缺陷检测对精度、计算资源及部署效率有较高要求, 单一性能指标难以全面反映改进价值。本文引入基于 Shapley 值的贡献度分析^[19], 综合考虑 3 类指标 F1、mAP@0.5 与 GFLOPs。前两项指标越大越优, GFLOPs 越小越优。F1 综合准确率与召回率, 全面反映检测性能; mAP@0.5 为目标检测主流指标, 衡量识别能力; GFLOPs 则代表模型计算开销, 尤其关乎工业边缘设备的实时性。Shapley 计算如式(31)所示。

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (n - |S| - 1)!}{n!} \times [v(S \cup \{i\}) - v(S)] \quad (31)$$

式中: i 为模型所添加的模块, 在本研究中依次添加了 C2f_CGBlock、SRepGFPN、Inner-SIoU 这 3 个模块。N 为所有模块的集合, S 为不包含模块 i 的子集, n 为模块数量, 在本研究中可知 $n = 3$ 。|S| 为在子集 S 中的模块数量, $v(S \cup \{i\})$ 为子集 S 与模块共同作用所展现出的性能指标数值, $v(S)$ 为仅有子集 S 展现出的性能指标数值。为统一数值尺度并遵循“数值越大性能越优”原则, 对 GFLOPs 采用归一化至 [0, 1], 并取相反数转化为正向贡献度。归一化公式如式(32)所示。

$$GFLOPs_{norm} = 1 - \frac{GFLOPs - \min(GFLOPs)}{\max(GFLOPs) - \min(GFLOPs)} \quad (32)$$

式中: $\max(GFLOPs)$ 、 $\min(GFLOPs)$ 分别代表在消融实验中 GFLOPs 的最小值与最大值。

在 Shapley 分析中, 性能指标的正值表示模块对性能提升有贡献, 负值表示下降; 复杂度指标的负值表示降低计算成本, 为正向贡献, 正值则反之。求取综合 Shapley 值则对指标进行加权处理, 对应的加权系数为 0.34、0.33、0.33, 可同时兼顾高精度和低计算开销的需求, 如式(33)所示。

$$\phi_i^{综合} = 0.34 \times \frac{\phi_i^{F1}}{\max(F1)} + 0.33 \times \frac{\phi_i^{mAP@0.5}}{\max(mAP@0.5)} + 0.33 \times \frac{\phi_i^{GFLOPs}}{\max(GFLOPs)} \quad (33)$$

式中: $\max(F1)$ 、 $\max(mAP@0.5)$ 分别为在消融实验中 F1 和 mAP@0.5 的最大值。

根据式(31)、(33)进行计算, Shapley 分析如表 10 所示。

根据上述信息可知, 模块 B 在 3 类模块中具有最显著的综合贡献为 0.670 0, 在 F1、mAP@0.5 上均为正向贡献, 为三者最高, 意味着显著提高了检测的准确性, 虽然表示此模块引入了一定的计算复杂度, 但其性能提升足以抵消复杂度增加带来的负面影响。模块 C 在 GFLOPs 几乎不变的前提下, 依然带来了较明显的性能增益, 在 F1、mAP@0.5 的 Shapley 值为 0.815 0、1.668 4, 综合值为 0.492 1, 与 SRepGFPN 相比效果次之。模块 A 在 GFLOPs 贡献最为显著, 但在 F1 和 mAP@0.5 的 Shapley 值为 0.660 0、0.348 3, 证明 A 在此两类指标上贡献有限, 导致其综合评分相对较低为 0.330 0。

表 10 Shapley 分析
Table 10 Shapley analyze

模块	子集 S	权重因子 $\frac{ S !(n- S -1)!}{n!}$	指标增益 $[v(S \cup \{i\}) - v(S)]$			ϕ_i^{F1}	$\phi_i^{mAP@0.5}$	ϕ_i^{GFLOPs}	$\phi_i^{综合}$
			$\Delta F1$	$\Delta mAP@0.5$	$\Delta GFLOPs$				
A 模块 (C2f_CGBlock)	\emptyset	0.333	-1.19	+1.35	-0.4	0.660 0	0.348 3	-0.750 0	0.330 0
	{B}	0.167	-1.07	-1.99	-1.0				
	{C}	0.167	-3.41	-1.82	-0.9				
	{B,C}	0.333	+5.41	+1.60	-0.9				
B 模块 (SRepGFPN)	\emptyset	0.333	+1.64	+3.31	+1.0	2.075 0	1.993 3	0.900 0	0.670 0
	{A}	0.167	+1.76	-0.03	+0.4				
	{C}	0.167	-3.41	-0.49	+1.0				
	{A,C}	0.333	+5.41	+2.93	+1.0				
C 模块 (Inner-SIoU)	\emptyset	0.333	+1.55	+2.90	+0.0	0.815 0	1.668 4	-0.050 0	0.492 1
	{A}	0.167	-0.67	-0.27	-0.5				
	{B}	0.167	-3.50	-0.90	+0.0				
	{A,B}	0.333	+2.98	+2.69	+0.1				

整体来看,模块 B 在提升检测性能方面的边际效益最高,而模块 A 在模型轻量化方面具有突出优势。模块 C 则在两者之间取得了一定平衡。该方法有效解决了多指标在量纲和优化方向上的差异,实现了多角度、统一标准的模块贡献评估,提升了评估结果的科学性与解释力。

(5)对较难检测类别的性能分析

indentation 类缺陷检测难度明显高于其他缺陷类型。

在前述数据集分析中,该类目标几乎全部为小目标,且在复杂背景下难以精准定位,因此本文将其实划分为“难检测类”单独分析。表 11 列出了该类别的指标数据,便于评估各改进模块的效果。单独引入改进 A、B、C 时,indentation 类 F1 均有所提升;YOLOv10n+A+B+C 在 F1 和 mAP@0.5 上表现最佳,显示出融合方案对难检测目标识别能力的增强。参数组合 YOLOv10n+B 在 R 上达到最高,虽 P 较低,但体现了该配置在缺陷捕获上的潜力。

表 11 难检测类别 indentation 的性能比较

Table 11 Performance comparisons for the hard-to-detect category: indentation

模型	P	R	F1	mAP@0.5	mAP@0.5:0.95	%
YOLOv10n(baseline)	68.28	70.11	69.18	73.18	31.98	
YOLOv10n+A	77.46	67.16	71.94	76.30	30.37	
YOLOv10n+B	66.44	79.63	72.44	79.73	29.85	
YOLOv10n+C	69.52	72.41	70.94	77.23	29.92	
YOLOv10n+A+B	72.35	70.11	71.21	73.85	29.25	
YOLOv10n+A+C	59.00	84.37	69.44	75.86	27.90	
YOLOv10n+B+C	77.83	67.82	72.48	79.35	30.72	
YOLOv10n+A+B+C	76.73	79.31	78.00	80.07	29.12	

综上所述,各个改进模块不仅能提升模型整体检测性能,在较难检测目标 indentation 类也具有显著的补强作用,尤其在 R、F1 上的提升尤为明显,验证了改进结构对钢轨微小表面缺陷的敏感性增强。

4.2 对比实验

本研究将所提出的改进模型与当前主流的 YOLO 系列不同版本及 RTDETR-Resnet50 进行系统比较,结果如表 12 所示。从表中可知,本模型在关键指标上表现突出:

P 为 88.18%,R 为 87.97%,F1 为 88.06%,均优于绝大多数对比模型,尤其在检测准确性和召回能力之间取得了良好平衡。同时,mAP@0.5 为 91.27%,表现出较强的整体检测性能。在计算复杂度方面,本模型 GFLOPs 为 6.6 G,显著低于较大规模模型如 YOLOv3 的 282.2 G、RTDETR-ResNet50 的 125.6 G,且接近轻量级模型 YOLOv10n 的 6.5 G。此外,FPS 指标适中,能够满足实时检测需求。参数量及模型大小方面,本模型参数量 10.85 MB,模型大小

表 12 对比实验

Table 12 Comparative experiments

模型	P/%	R/%	F1/%	mAP@ 0.5/%	mAP@ 0.5:0.95/%	GFLOPs/ G	FPS	参数量/ M	模型 大小
YOLOv3 ^[20]	82.42	85.63	83.78	89.05	46.31	282.2	173.48	395.46	198.1
YOLOv5n	88.41	84.80	86.43	90.47	48.51	7.1	274.48	9.55	5.0
YOLOv5s	84.42	88.83	86.52	90.07	46.48	23.8	253.05	34.76	17.6
YOLOv6n ^[21]	80.31	84.72	82.18	86.20	46.25	11.8	228.04	16.15	8.3
YOLOv7-tiny ^[22]	84.00	74.50	79.00	82.30	38.50	13.0	227.78	22.94	12.2
YOLOv8n	87.72	86.21	86.69	89.24	46.29	8.1	276.38	11.47	5.9
YOLOv8s	87.34	86.20	86.74	89.56	46.57	28.4	254.07	42.45	21.5
YOLOv9-tiny ^[23]	82.90	82.16	82.44	87.54	46.47	7.6	203.57	7.31	4.4
YOLO10n	84.80	84.25	84.51	87.26	47.22	6.5	190.49	8.64	5.5
YOLOv10s	85.69	85.00	85.34	89.44	46.65	24.5	297.11	30.66	15.7
YOLOv11n ^[24]	84.46	84.93	84.63	87.62	46.06	6.3	241.86	9.85	5.2
YOLOv12n ^[25]	81.41	87.90	84.34	87.34	45.87	5.8	203.81	9.57	5.2
RTDETR-Resnet50 ^[26]	79.48	79.49	79.08	83.53	39.70	125.6	92.28	159.99	82.0
本模型	88.18	87.97	88.06	91.27	47.70	6.6	212.02	10.85	6.7

约 6.7 M,属于轻量级模型范畴,有利于部署在资源受限的工业设备上。

综合来看,本模型在保持较低计算资源消耗和较快推理速度的同时,实现了优异的检测性能,体现了设计在精度与效率上的良好平衡,适合应用于钢轨表面缺陷检测的工业场景。

4.3 可视化分析

YOLOv10n与改进后模型的检测结果对比如图 11 所示,其中上排为 YOLOv10n 结果,下排为改进模型结果,从左到右依次为 crack、indentation、scratch 类。由于钢轨表面存在背景复杂等因素,原模型在置信度方面整体表现一般,3类缺陷的平均置信度分别为 0.74、0.75、0.72。经过结构改进后,模型在 3 类缺陷上的置信度分别提升至 0.81、0.80、0.77,说明改进模型在复杂背景下对目标的判别更加稳定。同时,这一置信度提升趋势也与各类缺陷在 P、R、F1 等指标上的改善情况一致。与此相对应,scratch

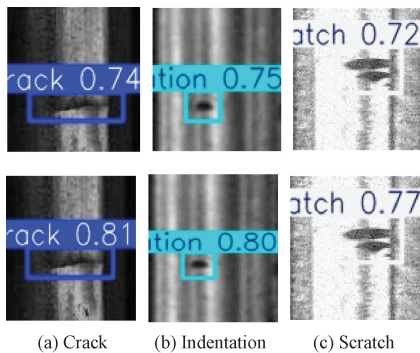


图 11 检测效果对比

Fig. 11 Comparisons of detection effects

和 indentation 两类缺陷的 F1、mAP@0.5 均有所提升。scratch 类 F1 从 85.58% 提升至 88.17%,mAP@0.5 从 89.58% 提升至 94.38%;indentation 类 F1 从 69.18% 提升至 78.00%,mAP@0.5 从 73.18% 提升至 80.07%。尽管 indentation 类在 mAP@0.5:0.95 上略有下降,但整体趋势仍显示出改进模型在复杂背景下更强的识别能力。

5 结 论

本研究针对工业场景下的钢轨表面缺陷检测任务,提出了一种基于 YOLOv10n 的轻量化检测框架,通过引入 C2f_CGBlock、SRepGFPN 以及 Inner-SIoU 等模块,显著提升了模型在低复杂度前提下的检测精度。多组消融实验表明,适度引入 CGBlock 可增强浅层特征类别区分能力,SRepGFPN 中学习参数的合理设计也有助于抑制误检。所提模型在 P、R、F1 以及 mAP 等关键指标上均优于现有主流方法,在小目标类别上表现出更高的识别能力,验证了其在复杂背景下的实用性与鲁棒性。但当前方法仍存在一定局限,后续研究将围绕以下方向展开:

受限于缺陷样本数量偏少,后续计划结合图像生成方法(如生成对抗网络 GAN)扩充小样本类别,以缓解类别不平衡问题;

引入更强的注意力机制与多尺度特征增强模块,提升模型对复杂背景中微小缺陷的感知能力;

探索模型在不同工业场景中的泛化性能及部署优化策略,进一步提升其实用性与稳定性。

参考文献

- [1] 魏新园,周京欢,张楠,等. 基于深度迁移学习的超声缺陷识别方法研究[J]. 仪器仪表学报,2024,45(12):

- 256-263.
- WEI X Y, ZHOU J H, ZHANG N, et al. Research on ultrasonic defect recognition method based on deep transfer learning [J]. Chinese Journal of Scientific Instrument, 2024, 45(12): 256-263.
- [2] 朱玥,王平,张兆珩,等. 基于最小熵解卷积的轨面缺陷漏磁信号处理[J]. 电子测量技术, 2022, 45(17): 167-170.
- ZHU Y, WANG P, ZHANG ZH H, et al. Magnetic flux leakage signal processing of rail surface defects based on minimum entropy deconvolution [J]. Electronic Measurement Technology, 2022, 45(17): 167-170.
- [3] 陈涛,尹永奇,吕程,等. 脉冲涡流缺陷检测理论与技术研究综述[J]. 电子测量与仪器学报, 2024, 38(9): 1-10.
- CHEN T, YIN Y Q, LYU CH, et al. Review of research on theory and technology of pulsed eddy current testing for detection of defects[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(9): 1-10.
- [4] 张航,周毅,邱宇峰. 融合 Hgnetv2 和注意力机制的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2025, 39(1): 36-49.
- ZHANG H, ZHOU Y, QIU Y F. Detection method of steel surface defects with fusion of Hgnetv2 and attention mechanism [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(1): 36-49.
- [5] 赵佰亭,张晨,贾晓芬,等. ECC-YOLO: 一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2024, 38(4): 108-116.
- ZHAO B T, ZHANG CH, JIA X F, et al. ECC-YOLO: An improved method for detecting surface defects in steel[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [6] 刘俊博,杜馨瑜,王胜春,等. 基于少样本学习的钢轨表面缺陷检测方法[J]. 铁道学报, 2022, 44(7): 72-79.
- LIU J B, DU X Y, WANG SH CH, et al. Rail surface defect detection method based on few-shot learning[J]. Journal of the China Railway Society, 2022, 44(7): 72-79.
- [7] 吴永军,崔灿,何永福. 基于语义增广与 YOLOv8 的钢轨表面缺陷检测方法[J]. 铁道科学与工程学报, 2024, 21(9): 3864-3875.
- WU Y J, CUI C, HE Y F. Rail surface defect detection based on semantic augmentation and YOLOv8 [J]. Journal of Railway Science and Engineering, 2024, 21(9): 3864-3875.
- [8] 闵永智,李嘉峰,王果. 基于改进 DCGAN 的钢轨表面缺陷图像扩充方法[J]. 铁道学报, 2023, 45(12): 123-130.
- MIN Y ZH, LI J F, WANG G. Image expansion method for rail surface defects based on improved DCGAN[J]. Journal of the China Railway Society, 2023, 45(12): 123-130.
- [9] 陈仁祥,潘升,杨黎霞,等. 基于注意力引导多尺度降噪卷积神经网络的钢轨表面缺陷图像降噪[J]. 铁道学报, 2024, 46(5): 123-131.
- CHEN R X, PAN SH, YANG L X, et al. Noise reduction of rail surface defect images based on attention-guided poly-scale denoising convolutional neural networks [J]. Journal of the China Railway Society, 2024, 46(5): 123-131.
- [10] 罗晖,李健,贾晨. 基于图像增强与改进 Cascade R-CNN 的钢轨表面缺陷检测[J]. 激光与光电子学进展, 2021, 58(22): 324-335.
- LUO H, LI J, JIA CH. Rail surface defect detection based on image enhancement and improved Cascade R-CNN[J]. Laser & Optoelectronics Progress, 2021, 58(22): 324-335.
- [11] 杜少聪,张红钢,王小敏. 基于改进 YOLOv5 的钢轨表面缺陷检测[J]. 北京交通大学学报, 2023, 47(2): 129-136.
- DU SH C, ZHANG H G, WANG X M. Rail surface defect detection based on improved YOLOv5 [J]. Journal of Beijing Jiaotong University, 2023, 47(2): 129-136.
- [12] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[C]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [13] WU T Y, TANG SH, ZHANG R, et al. CGNet: A light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2020, 30(1): 1169-1179.
- [14] XU X ZH, JIANG Y Q, CHEN W H, et al. DAMO-YOLO: A report on real-time object detection design[J]. ArXiv preprint arXiv:2211.15444, 2023.
- [15] YANG L, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C]. International Conference on Machine Learning, PMLR, 2021: 11863-11874.
- [16] GEVORGYAN Z. SiU Loss: More powerful

- learning for bounding box regression [J]. ArXiv preprint arXiv:2205.12740, 2022.
- [17] ZHANG H, XU C, ZHANG SH J. Inner-IoU: More effective intersection over union loss with auxiliary bounding box [J]. ArXiv preprint arXiv: 2311.02877, 2023.
- [18] GAN J R, LI Q Y, WANG J ZH, et al. A hierarchical extractor-based visual rail surface inspection system[J]. IEEE Sensors Journal, 2017, 17(23): 7935-7944.
- [19] ROZEMBERCZKI B, WATSON L, BAYER P, et al. The shapley value in machine learning[C]. Thirty-First International Joint Conference on Artificial Intelligence. Vienna, Austria, 2022: 5572-5579.
- [20] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv: 1804.02767, 2018.
- [21] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial application [J]. ArXiv preprint arXiv: 2209.02976, 2022.
- [22] WANG C Y, BOCHKOVSKIY A, LIAO H M, et al. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [23] 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.
- [24] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.
- [25] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors[J]. ArXiv preprint arXiv:2502.12524, 2025.
- [26] ZHAO Y AN, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

作者简介

刘伯鸿(通信作者), 硕士, 副教授, 主要研究方向为故障诊断等。

E-mail:2497741476@qq.com

刘珮, 硕士, 主要研究方向为故障检测等。

E-mail:1170556037@qq.com