

DOI: 10.19650/j.cnki.cjsi.J2514496

融合动态卷积与可变形注意力的钢材缺陷检测*

赵海丽¹, 狄子隆¹, 景文博², 宋明喆¹

(1. 长春理工大学电子信息工程学院 长春 130022; 2. 长春理工大学光电工程学院 长春 130022)

摘要: 钢材表面缺陷的精准检测是工业质量控制的关键环节。尤其在机械制造、汽车工业、电子设备、航空航天及火炮深管等精密制造领域,其表面质量直接决定终端产品的安全性与可靠性。针对现有钢材表面缺陷检测方法中存在多尺度缺陷检测能力有限,小目标与低对比度缺陷漏检率高,边界框回归精度不佳等问题,提出了一种基于 YOLOv11n 改进的多尺度钢材表面缺陷检测方法。设计多尺度动态卷积,通过并行异构卷积与动态权重融合机制,增强模型对多尺度缺陷的捕捉能力;构建动态残差融合模块,以分组卷积与双重残差结构替换基线 C3K2 模块,在显著降低参数数量的同时提升多尺度特征融合与梯度流通过率,缓解深层网络训练退化问题;改进可变形三元注意力机制,融合可变形卷积与跨维度交互,使注意力感受野随缺陷形态动态调整,精准聚焦微小、低对比度区域,抑制复杂背景干扰;采用 Shape-IoU 损失函数,通过引入形状与尺度因子优化边界框回归精度,解决传统 CIoU 在宽高比一致时惩罚失效的问题。在 NEU-DET 数据集上的实验结果表明:改进模型的 mAP@0.5 达到 81.9%,相比基线 YOLOv11n 提升了 6%,参数数量仅为 2.3 M,计算量降至 5.9 GFLOPs,满足边缘设备部署需求。泛化实验在 GC10-DET 数据集上较基线模型提升 4.1%。可视化分析与泛化性实验进一步验证了其在复杂工业场景下的鲁棒性与实用性。
关键词: 钢材表面缺陷检测;YOLOv11n;多尺度特征提取;可变形三元注意力机制;Shape-IoU
中图分类号: TP391.41 TH165 **文献标识码:** A **国家标准学科分类代码:** 510.40

A steel defect detection with fused dynamic convolution and deformable attention

Zhao Haili¹, Di Zilong¹, Jing Wenbo², Song Mingzhe¹

(1. School of Electronic Information Engineering, Changchun University of Technology, Changchun 130022, China;
2. School of Optoelectronic Engineering, Changchun University of Technology, Changchun 130022, China)

Abstract: Accurate detection of steel surface defects is a critical aspect of industrial quality control. Especially in precision manufacturing fields such as mechanical engineering, automotive industry, electronics, aerospace, and artillery barrel production, surface quality directly determines the safety and reliability of end products. To address the limitations of existing steel surface defect detection methods, including insufficient multi-scale defect detection capability, high missed detection rates for small and low-contrast defects, and suboptimal bounding box regression accuracy, this article proposes an improved multi-scale steel surface defect detection method based on YOLOv11n. A multi-scale dynamic convolution module is designed, which employs parallel heterogeneous convolutions and a dynamic weight fusion mechanism to enhance the model's ability to capture multi-scale defects. A dynamic residual fusion module is formulated, replacing the baseline C3K2 module with grouped convolution and a dual-residual structure. This significantly reduces the parameter count while improving multi-scale feature fusion and gradient flow efficiency, alleviating the degradation issue in deep network training. The deformable triple attention mechanism is enhanced by integrating deformable convolution and cross-dimensional interaction, enabling the attention receptive field to dynamically adjust according to defect morphology, thereby precisely focusing on small, low-contrast regions and suppressing complex background interference. The Shape-IoU loss function is adopted, which incorporates shape and scale factors to optimize bounding box regression accuracy, addressing the failure of the penalty term in traditional CIoU when the aspect ratios are identical. Experimental results on the NEU-DET dataset show that the improved model achieves an mAP@0.5 of 81.9%, representing a 6% improvement over the baseline YOLOv11n. The parameter count is only 2.3 M, and the computational cost is reduced to 5.9 GFLOPs, meeting the requirements for deployment on edge devices. Generalization experiments on the GC10-DET dataset show a

收稿日期:2025-09-13 Received Date: 2025-09-13

* 基金项目:某部靶场测试专项规划项目(2T-2018-06)资助

4.1% improvement over the baseline model. Visualization analysis and generalization experiments further validate its robustness and practicality in complex industrial scenarios.

Keywords: steel surface defect detection; YOLOv11n; multi-scale feature extraction; deformable triple attention mechanism; Shape-IoU

0 引言

钢材作为机械制造、汽车工业、电子设备及航空航天及火炮深管等精密领域的核心材料,其表面质量直接决定终端产品的安全性和可靠性。目前,传统的钢材表面缺陷检测方法,如人工检测、漏磁检测^[1]、涡流检测^[2]等存在成本高,准确性与效率低下问题,无法满足现代工业生产需求。因此,在钢材生产中质量检测时,准确高效的表面缺陷检测方法对保障钢材生产质量至关重要。在智能制造和高质量发展战略的推动下,亟需更先进的检测手段取代人工检测,以提高检测效率和精度^[3-4]。

近年来,随着人工智能的不断发展,深度学习广泛应用到缺陷检测领域。目前,基于深度学习的目标检测算法分为一阶段和二阶段两大类。例如一阶段的单次检测目标算法(you only look once, YOLO)^[5-10]、单发多框检测(single shot multibox detector, SSD)^[11]和二阶段的快速区域卷积神经网络(faster region-based convolutional neural network, Faster R-CNN)^[12]、掩码区域卷积神经网络(mask region-based convolutional neural network, Mask R-CNN)^[13]等算法。由于我国钢铁企业对于表面缺陷的研究起步较晚,现有的技术尚不能完全满足企业实际项目的需要,因此需要研究适用于实际钢材表面缺陷检测的方法。Liu等^[14]提出了一种多尺度检测网络(detection network with multiscale, MSC-DNet)钢材表面缺陷检测算法,构建不同扩张率的膨胀卷积并行架构(parallel architecture of dilated convolution, PADC),增强模型捕捉多尺度缺陷能力以提高缺陷检测精度。张彪等^[15]在YOLOv8n的基础上,引入LiteShiftHead检测头,提升了模型特征提取、目标回归和分类的效率与准确性。凌晓等^[16]基于特征金字塔网络(feature pyramid network, FPN)构建动态稀疏门控特征金字塔网络(dynamic sparse gate feature pyramid network, DSG-FPN),并结合ResNet34结构,通过动态系数门控机制优化特征融合过程,从而提升检测精度与效率。Zhang等^[17]在YOLOv5网络模型基础上引入CSPLayerRes2Attention残差模块,减少分层传输过程中缺陷信息丢失,增强模型对细粒度特征提取能力。Song等^[18]提出了一种RSTD-YOLOv7钢材表面缺陷检测算法,使用Swin Transformer构建Swin Transformer重参数化VGG(Swin Transformer RepVGG, STRVGG)模块合并到颈部网络中,增强模型对特征图深层信息的提取能力。岳剑峰等^[19]采用Focal Loss和加权交叉熵构成的

焦点加权交叉熵损失函数(focal weighted cross entropy loss, FWCE Loss)组合损失函数改进模型,通过调整权重提升模型对不均衡样本的检测精度。Chao等^[20]在YOLOv8的基础上引入信息增强与多尺度(information augmentation and multiscale, IAM)新型骨干网络,减少特征下采样期间的特征信息丢失。徐森等^[21]在YOLOv8模型主干网络中引入空间到深度卷积(space to depth non-strided convolution, SPD-Conv),提高模型的低分辨率图像检测能力。张小龙等^[22]设计高效特征权重聚合模块代替YOLOv11模型颈部特征拼接操作,通过可学习权重实现对不同层级缺陷特征的有效融合。雷富强等^[23]在YOLOv11基础上设计了动态任务对齐检测头,通过分类与定位分支的协同优化策略实现检测性能提升。

上述方法,包括YOLOv8、YOLOv11等最新模型,虽在通用目标检测任务中表现出色,但在应对钢材表面缺陷的特定挑战时仍存在局限:1)对于高密度点状缺陷(麻点、内含物)与稀疏长条划痕等极端多尺度分布的缺陷,其特征提取网络的感受野与特征融合策略仍显不足;2)难以在保持高精度的同时满足严格的实时性要求;3)针对低对比度、微小缺陷的自适应感知能力有限,导致漏检率偏高。因此,亟需一种专为钢材缺陷检测任务设计的、在精度与速度间取得更优平衡的轻量级模型。

针对上述挑战,本文在YOLOv11n基础上进行了一系列面向应用的结构改进。本文的主要贡献可概括为4点,即:

1)提出了一个多尺度动态卷积模块(multi-scale dynamic convolution module, MDC):通过并行异构卷积核与轻量级动态权重生成器,自适应地融合不同尺度的特征,增强模型对形态各异的钢材缺陷的捕捉能力,同时保持了计算的轻量化。

2)构建了一个动态残差融合模块(dynamic residual fusion module, DRF):利用上述MDC模块,设计了分组卷积与双重残差连接的结构,以替换基线模型中的标准瓶颈模块,显著增强了网络的多尺度特征融合与梯度流通能力。

3)改进了可变形三元注意力机制(deformable triplet attention, DTA):将可变形卷积融入三元注意力中,使其在进行跨维度交互时能动态调整感受野,从而更精准地聚焦于微小、低对比度的缺陷区域,有效抑制复杂背景干扰。

4)系统性地整合了上述模块并采用Shape-IoU损失函数:通过将DRF和DTA二者协同集成到YOLOv11n架构中,并结合更关注边界框形状的Shape-IoU损失,构建了一个高效的检测模型。在东北大学表面缺陷数据集(NEU

surface defect database, NEU-DET) 上的实验表明,本方法在仅微增计算成本的条件下,显著提升了检测精度。

1 改进 YOLOv11 模型

本研究以 YOLOv11n 为基线模型进行设计,针对钢材表面缺陷检测中多尺度特征提取不足、小目标漏检率高及低对比度缺陷难识别等问题,旨在进一步提升模型的检测精度与鲁棒性。首先设计一种多尺度动态卷积,

提升模型对不同形态缺陷的捕捉能力,并在其基础上构建动态残差融合模块(DRF),用其替换 YOLOv11 中 C3K2 模块,以提升网络多尺度特征提取能力。其次设计可变形三元注意力机制在嵌入网络 Backbone 末端,在增加极少参数量的情况下,增强网络对钢材表面微小缺陷和低对比度缺陷的感知能力,最后采用 Shape-IoU 替换原网络中完全交并比(complete intersection over union, CIOU),解决当预测框与真实框宽高比相同时惩罚项失效问题。总体结构如图 1 所示。

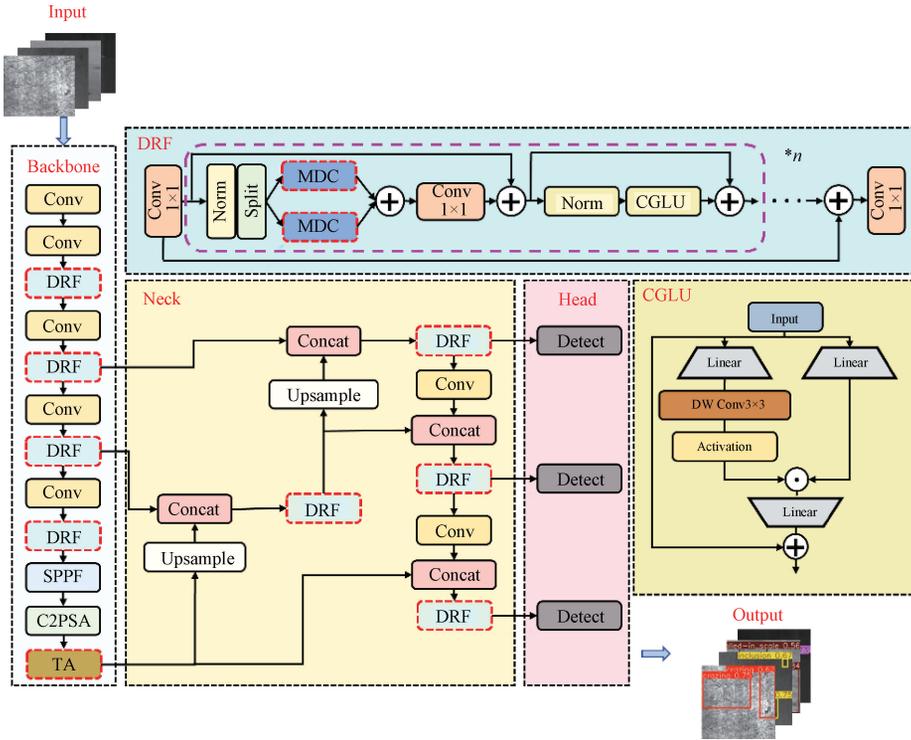


图 1 改进 YOLOv11 网络结构

Fig. 1 Diagram of the improve YOLOv11 network architecture

1.1 多尺度动态卷积模块

钢材表面缺陷(如裂纹、划痕、麻点)具有多尺度分布与形态学异质性的特征,标准卷积核因受限于固定几何结构难以实现对多尺度不同形态缺陷的有效响应。对此,设

计 MDC 多尺度动态卷积模块,通过异构分支与动态权重融合机制,增强模型对多类缺陷的特征捕获能力。多尺度动态卷积模块由方块卷积、水平带状卷积、垂直带状卷积以及轻量级动态权重生成器组成,结构如图 2 所示。

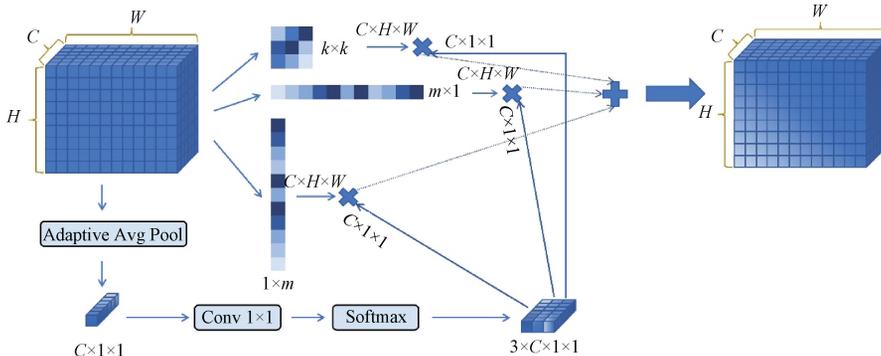


图 2 多尺度动态卷积模块

Fig. 2 Multi-scale dynamic convolution module

假设输入特征图 X 通道数为 C , 空间尺寸为 $H \times W$, 模块通过并行深度卷积分支与轻量级动态权重机制实现输入自适应的特征融合。具体而言, 在多尺度动态卷积模块中, 特征通过 3 条并行深度可分离卷积路径提取异构特征。分别为: 1) 方块卷积, 核尺寸为 $k \times k$, 捕获局部领域特征; 2) 水平带状卷积, 核尺寸为 $1 \times m$, 捕获水平方向特征; 3) 垂直带状卷积, 核尺寸为 $m \times 1$, ($m = 3k + 2$) 捕获垂直方向特征。为进一步提升卷积对不同缺陷的适应性, 模型引入轻量级动态权重生成器。该组件由原输入通过自适应平均池化和 1×1 卷积生成 3 组原始权重, 如式(1)所示。

$$\omega_{\text{raw}} = \mathbf{W}_{\text{conv}} \cdot \text{pool}(x) \quad (1)$$

其中, \mathbf{W}_{conv} 是 1×1 卷积的权重矩阵。将权重矩阵重组并做 Softmax 归一化处理得到各分支对应权重, 如式(2)所示。

$$\omega_i = \frac{\exp(\omega_{\text{reshaped}[i]})}{\sum_{k=0}^2 \exp(\omega_{\text{reshaped}[i]})} \quad (2)$$

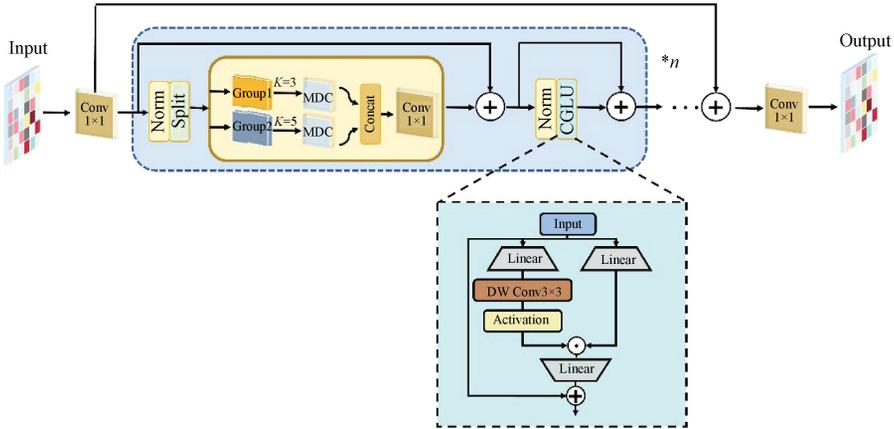


图3 动态残差融合模块

Fig. 3 Dynamic residual fusion module

流程为: 首先输入特征图通过一个 1×1 卷积层进行降维, 随后进入第 1 个残差单元。在该单元中, 特征图经过批归一化处理, 沿通道维度平均分为两组 (Group1 和 Group2)。每组分别使用不同卷积核尺寸 (如 3×3 和 5×5) 的 MDC 模块进行特征提取。当 $kernels = 3$ 时, 模型专注于较细粒度的局部特征, 当 $kernels = 5$ 时, 模型扩大感受野, 以捕捉稍大尺度特征。将提取后的两组特征拼接后通过 1×1 卷积进行融合, 并与原始输入进行残差连接, 形成初步增强特征。将初步增强特征输入第 2 个残差单元, 依次经过批归一化和卷积门控线性单元 (convolutional gated linear unit, CGLU)^[24], 通过动态门控机制进一步优化特征表达。该单元的输出再次与第 1 个残差单元的输出进行残差连接, 实现特征增强。重复上述两个残差单元操作 n 次后, 将最终增强特征与初始输

入进行残差连接, 再通过 1×1 卷积输出最终结果。该模块通过分组卷积降低计算量、异构卷积核增强特征多样性、双重残差结构保障梯度流动, 在提升多尺度特征融合能力的同时, 保持了模型的轻量化特性。

并对加权后结果进行融合得到结果 Y , 如式(3)所示。

$$Y = \sum_{i=0}^2 \omega_i \odot y_i \quad (3)$$

其中, ω 为动态卷积权重, y_i 为三分支卷积结果, \odot 为动态卷积权重与卷积结果逐通道相乘运算。最终通过批归一化和 SiLU 激活输出结果 Z , 如式(4)所示。

$$Z = \text{SiLU}(\text{BN}(\sum_{i=0}^2 \text{softmax}_k(\mathbf{W}_{\text{conv}} \cdot \text{pool}(x)) \odot y_i)) \quad (4)$$

1.2 动态残差融合网络

为进一步增强主干网络对多尺度缺陷的特征的融合与传递, 并保持模型的轻量化, 本文利用 1.1 节提出的 MDC 多尺度动态卷积模块, 构建 DRF 动态卷积残差融合模块, 其结构如图 3 所示。

该模块采用分组卷积与双重残差连接的设计, 既能有效提取多尺度特征, 又能缓解梯度消失问题。具体

入进行残差连接, 再通过 1×1 卷积输出最终结果。该模块通过分组卷积降低计算量、异构卷积核增强特征多样性、双重残差结构保障梯度流动, 在提升多尺度特征融合能力的同时, 保持了模型的轻量化特性。

1.3 可变形三元注意力机制

钢材表面缺陷中裂纹、麻点等目标常呈现尺寸微小、对比度低、形状不规则的特点, 传统注意力机制因卷积核几何结构固定, 难以自适应聚焦缺陷区域, 导致漏检率偏高。针对这一问题, 本文提出可变形三元注意力机制 DTA, 通过融合 Triplet Attention^[25] 的跨维度交互思想与深度可分离卷积 (deformable convolutional networks v2, DCNv2)^[26] 可变形采样能力, 在不增加参数数量的前提下, 显著提升模型对细粒度缺陷的敏感度。DTA 整体采用“三分支并行-跨维度旋转-可变形增强-平均融合”的结

构,依次建模通道-高度、通道-宽度与空间 3 种交互关 系,流程如图 4 所示。

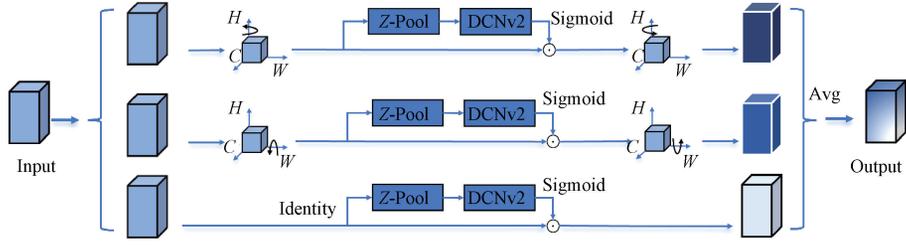


图 4 可变形三元注意力机制

Fig. 4 Deformable triple attention

假设输入特征图的高宽为 H 、 W , ψ 为核为 k 的 DCN2 通道数为 C , 通道-高度分支对输入特征图沿 H 轴逆时针旋转 90° , 得到形状为 $W \times H \times C$ 的特征图 \hat{x}_1 。对 \hat{x}_1 进行 Z-Pool 操作, 将特征图通道维度压缩为 2, 得到形状为 $2 \times H \times C$ 的特征图 \hat{x}_1^* , 再将 \hat{x}_1^* 通过 DCNv2 得到形状为 $1 \times H \times C$ 的中间输出, 最后通过 Sigmoid 激活函数生成注意力权重 ω_1 , 将 ω_1 应用于 \hat{x}_1 , 然后沿 H 轴顺时针旋转 90° , 恢复原始输入特征图形状 $C \times H \times W$ 得到结果 y_1 , 如式(5)所示。

$$y_1 = \hat{x}_1 \sigma(\psi_1(\hat{x}_1^*)) \quad (5)$$

同理, 通道-宽度分支将输入特征图沿 W 轴逆时针旋转得到形状为 $H \times C \times W$ 的特征图 \hat{x}_2 , 随后进行 Z-Pool 操作, DCNv2 和 Sigmoid 激活函数生成注意力权重 ω_2 。将 ω_2 应用于 \hat{x}_2 然后沿 W 轴顺时针旋转 90° 恢复原始输入特征图形状 $C \times H \times W$ 得到结果 y_2 , 如式(6)所示。

$$y_2 = \hat{x}_2 \sigma(\psi_2(\hat{x}_2^*)) \quad (6)$$

空间分支直接在原始特征图 x 进行 Z-Pool 操作, 将通道维数压缩为 2, 得到 $2 \times H \times W$ 的特征图 \hat{x}_3 , 随后通过 DCNv2 和 Sigmoid 激活函数生成注意力权重 ω_3 , 将 ω_3 作用于 \hat{x}_3 得到结果 y_3 , 如式(7)所示。

$$y_3 = x \sigma(\psi_3(x_3^*)) \quad (7)$$

最后, 将 3 个分支生成的注意力权重应用后的特征图进行简单平均, 得到最终输出特征图 y , 如式(8)所示。

$$y = \frac{1}{3}(y_1 + y_2 + y_3) \quad (8)$$

式中: σ 表示 Sigmoid 激活函数; ψ_1 、 ψ_2 和 ψ_3 表示在注意力中 3 个分支中核大小为 k 的 DCNv2。上述过程通过跨维度旋转打破通道与空间维度的隔离, 使模型能够捕获更全面的上下文关系; 同时引入 DCNv2 的可变形采样, 令注意力权重的感受野随缺陷形态动态调整, 从而有效抑制背景噪声、突出微小缺陷。

1.4 边界框损失函数

原始的 YOLOv11 使用完整的交并比损失函数 CIoU

作为边界框损失函数, 其表达式如式(9)所示。

$$L_{\text{CIoU}} = 1 - \text{IOU} + \frac{\rho^2(b, b^{gt})}{c^2} + \beta v \quad (9)$$

式中: β 为权重参数, 用于调整宽高比损失在整个边界框损失中所占的比重, 最大不超过 1, 其表达式如式(10)所示。

$$\beta = \frac{v}{(1 - \text{IOU}) + v} \quad (10)$$

式中: v 为修正因子, 反应预测框和真实框的宽高比一致性, 其表达式如式(11)所示。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (11)$$

可以看出, CIoU 损失考虑了边界框回归重叠面积, 中心点距离和宽高比, 但当预测框与真实框宽高比相同时, 式中 $v=0$, 显然与实际不符。针对此问题, 本文引入 Shape-IoU^[27] 损失函数, 通过引入边界框自身形状和尺度因子来优化损失计算。首先定义权重系数 w 和 h , 它们与真实框的宽度 w^{gt} 、高 h^{gt} 以及尺度因子 $scale$ 相关, 用于体现形状对横纵方向的影响, 其表达式如式(12)和(13)所示。

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (12)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (13)$$

其次, 形状距离 ($distance^{\text{shape}}$) 和形状损失 (Ω^{shape}) 的计算公式如式(14)和(15)所示。

$$distance^{\text{shape}} = hh \times (x_c - x_c^{gt})^2 / c^2 + ww \times (y_c - y_c^{gt})^2 / c^2 \quad (14)$$

$$\Omega^{\text{shape}} = \sum_{t=w,h} (1 - e^{-w_t})^\theta, \quad \theta = 4 \quad (15)$$

最后, Shape-IoU 的总损失函数如式(16)所示。

$$L_{\text{Shape-IoU}} = 1 - \text{IoU} + distance^{\text{shape}} + 0.5 \times \Omega^{\text{shape}} \quad (16)$$

其中, IoU 为交并比, x_c 、 y_c 和 x_c^{gt} 、 y_c^{gt} 分别为预测框和真实框的中心坐标, w 、 h 和 w^{gt} 、 h^{gt} 分别为预测框和真实框的宽高, c 为两框最小外接框的对角线距离, $scale$ 为与

数据集中目标尺度相关的因子。由式可见 Shape-IoU 损失函数由 IoU 损失函数、形状距离损失、形状损失 3 个代价函数组成,并通过引入尺度因子 $scale$ 和权重系数 w, h 更全面地考虑边界框自身形状和尺度对回归结果的影响。预测框与真实框示意图如图 5 所示。

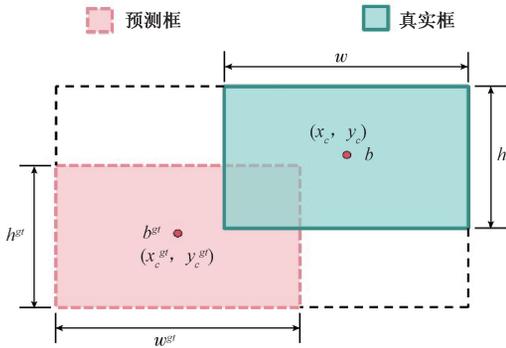


图 5 预测框与真实框

Fig. 5 Diagram of predicted box and real box

2 实验验证

2.1 实验环境

实验环境搭建在 Windows10 系统,采用 Python 编程语言和 Pytorch 框架进行开发。基于 Pytorch 深度学习框架部署模型训练。硬件配置为 24 G 显存的 NVIDIA 3090Ti 显卡。输入图像分辨率为 256 pixels×256 pixels,优化器为 SGD,初始学习率为 0.01,动量参数为 0.937,权重衰减系数为 0.0005,训练轮次为 600,batch size 设置为 128。具体实验环境配置如表 1 所示。

表 1 实验环境配置

Table 1 Experiment setup

软件/硬件	配置环境
GPU	NVIDIA 3090Ti
显存大小	24 G
操作系统	Windows 10
Python	3.8
Pytorch	1.13.0
CUDA	11.6
CuDNN	8.3

2.2 实验数据集

本实验选用东北大学 NEU-DET^[28] 钢材缺陷数据集开展研究。该数据集的标注涵盖裂纹(crazing)、斑块(patches)、内含物(inclusion)、点蚀表面(pitted surface)、氧化皮(rolled in scale)以及划痕(scratches)这 6 类钢材

缺陷。每类缺陷样本均为 300 张,数据集样本总量达 1 800 张。在本文实验中,依据 8:1:1 的比例,将该数据集划分为训练集、验证集与测试集。其中各类缺陷实例如图 6 所示。

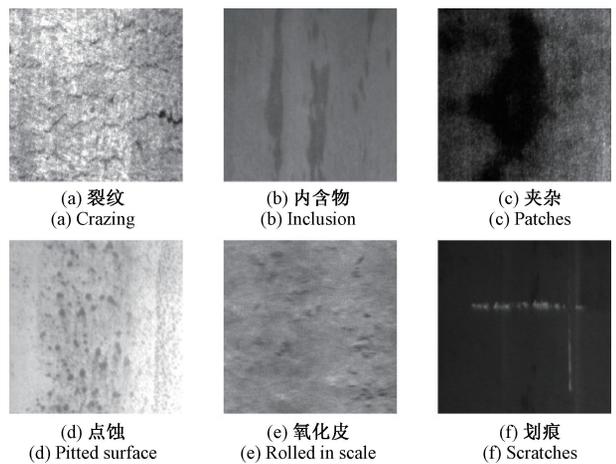


图 6 NEU-DET 中的 6 类缺陷

Fig. 6 Six types of defects in NEU-DET

2.3 评价指标

为全面评估钢材表面缺陷检测模型的性能,本文采用精确率(Precision)、召回率(Recall)和平均精度(mean average precision, mAP)作为评价指标,这些指标能够从不同角度反映模型的检测能力。

精确率衡量的是模型预测为正样本的结果中真正为正样本的比例,体现了模型预测结果的准确性,其计算公式如式(17)所示。

$$Precision = \frac{TP}{TP + FP} \quad (17)$$

其中, TP (true positive) 表示被模型正确预测为正样本的数量; FP (false positive) 表示被模型错误预测为正样本的负样本数量。

召回率衡量的是所有真实的正样本中被模型正确预测出来的比例,反映了模型对正样本的识别能力,其计算公式如式(18)所示。

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

其中, FN (false negative) 表示被模型错误预测为负样本的正样本数量。

平均精度(average precision, AP)是精确率-召回率曲线下的面积,用于综合评估模型在单一类别上的性能。其计算公式如式(19)所示。

$$AP = \int_0^1 P(R) \quad (19)$$

mAP 是一个综合的度量,用于评估跨多个类别的目标检测模型的性能。计算每个类别的 AP,然后求其均值

来衡量模型的性能,其计算公式如式(20)所示。

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (20)$$

其中, C 表示数据集中类别。mAP 值越高,模型性能越好。mAP_{0.5:0.95} 表示在多个 IoU 阈值(从 0.5 ~ 0.95,步长为 0.05)下计算得到的平均 mAP。

表 2 消融实验结果

Table 2 Ablation experiment results

DRF	DTA	Shape-IoU	$R/\%$	$GFLOPs/10^9$	$mAP@0.5/\%$	$Params/M$	$mAP/\%$
			74.3	6.4	75.9±0.3	2.6	-
✓			77.1	5.9	79.8±0.2	2.3	+3.9
	✓		76.4	6.4	78.2±0.4	2.6	+2.3
		✓	75.3	6.4	77.4±0.2	2.6	+1.5
✓	✓		77.3	5.9	80.4±0.3	2.3	+4.5
	✓	✓	76.8	6.4	79.5±0.3	2.6	+4.6
✓		✓	77.4	5.9	80.5±0.2	2.3	+3.6
✓	✓	✓	77.8	5.9	81.9±0.1	2.3	+6.0

本实验以原 YOLOv11n 网络的检测结果作为消融实验基准。表 2 中第 2 行是使用 DRF 模块的结果该模块通过分组 MDC 卷积与双重残差连接结构,在增强多尺度特征融合能力的同时,显著降低了模型复杂度,参数量降至 2.3 M,GFLOPs 降至 5.9。mAP@0.5 提升至 79.8%,召回率提高 2.1%,表明 DRF 在减少参数的同时有效增强了特征表达能力。第 3 行是添加 DTA 注意力机制模块的结果,该模块通过不同维度(通道维度和空间维度)之间的交互来计算注意力权重,同时结合 DCNv2 随缺陷尺度自适应伸缩感受野,提升模型小目标和复杂背景下特征提取能力。在不增加参数量与计算量的前提下,mAP@0.5 提升至 78.2%,召回率提升至 76.4%。第 4 行是使用 Shape-IoU 的检测结果,Shape-IoU 引入边界框自身形状和尺度因子,更关注缺陷本身形状特征,对检测方向敏感,形状多变的缺陷更有优势。添加之后 mAP@0.5 提高至 77.4%,参数量与计算量保持不变,表明其对边界框形状敏感,优化了回归精度。第 5 行展示了使用 DRF 模块和添加 DTA 注意力模块的模型结果,可以发现两个模块的组合强化了 Backbone 层对钢材表面缺陷多尺度特征提取能力,mAP@0.5 进一步提升至 80.4%,参数量与 GFLOPs 仍控制在较低水平。说明 DRF 的多尺度特征提取与 DTA 的注意力聚焦具有互补性,共同提升了模型对复杂缺陷的判别能力。第 6 行是添加 DTA 注意力机制模块和使用 Shape-IoU 模型的检测结果,组合后 mAP@0.5 达 79.5%,参数量与计算量未增加,体现了注意力机制与损失函数在优化检测结果方面

2.4 消融实验

为了验证本文改进模型的优化效果和多尺度动态卷积网络、三重注意力、Shape-IoU 对模型检测精度提升,本文基于 YOLOv11n 网络模型,进行 8 组消融实验,通过选择性增减这 3 种方法来评估他们对模型性能改进的影响。实验结果如表 2 所示。

的协同作用。第 7 行是了使用 DRF 模块和 Shape-IoU 模型的检测结果,该组合在降低模型复杂度的同时,mAP@0.5 提升至 80.5%,进一步验证了结构优化与回归策略结合的有效性。第 8 行为本文模型最终结果,最终模型在参数量为 2.3 M、计算量为 5.9GFLOPs 的轻量级配置下,mAP@0.5 达到 81.9%,较基线提升 6%。召回率亦提升至 77.8%。结果表明,3 个模块在特征提取、注意力机制与边界框回归 3 个层面形成有效协同,共同推动检测性能的全面提升,且在控制计算成本方面表现出色。实验表明,改进后的模型具有更强的多尺度特征提取能力,可以更好地避免漏检误检情况。

将训练后的原始 YOLOv11n 和改进后 YOLOv11n 模型边界框回归损失随训练轮次的变化曲线进行对比,如图 7 所示。

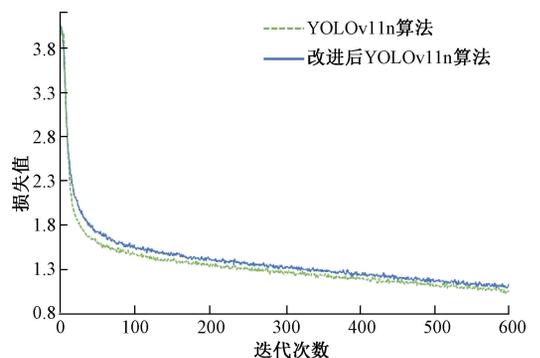


图 7 损失曲线对比

Fig. 7 Comparison of loss curves

在训练初期,Shape-IoU 损失的模型边界框回归损失值下降更为迅速,表明其具有更快的收敛速度。在整个训练周期结束时,Shape-IoU 损失的最终稳定值也低于 CIoU 损失,这一定量结果与表 2 中消融实验显示的 mAP 提升相互印证,证明 Shape-IoU 通过引入形状和尺度因子,能够更有效地指导模型进行边界框回归。

2.5 多种注意力机制对比实验

为了在不增加模型负担的前提下,提高模型对钢材表面缺陷的检测精度,尤其是减少形态差异较大,小目标,低对比度缺陷的漏检率。本文在网络 Backbone 末端融入 DTA 注意力机制。同时,为了说明融入 DTA 注意力机制的有效性,在网络同一位置引入了 4 种主流的注意力机制,分别是压缩-激励注意力机制^[29](squeeze-and-excitation network,SE)、卷积块注意力机制^[30](convolution block attention module,CBAM)、坐标注意力^[31](coordinate attention,CA)以及高效通道注意力^[32](efficient channel attention,ECA)。并在 NEU-DET 数据集上进行对比实验。结果如表 3 所示。

表 3 多种注意力对比实验结果

Table 3 Comparison results of various attention mechanisms

方法	mAP@0.5/%	Params/M	P/%	R/%
原模型	75.9	2.6	77.9	74.3
SE	75.7	2.6	77.6	71.8
CBAM	78.1	2.6	73.8	75.1
CA	75.9	2.6	74.8	74.6
ECA	77.9	2.6	77.5	75.4
DTA	78.2	2.6	76.7	76.4

由表 3 中数据可知,在 5 种注意力机制中 DTA 注意力机制的 mAP@0.5 高达 78.2%,相较于原模型提高了 2.3%,且参数量未出现明显增加;同时,召回率(Recall)同样最高,达到 76.4%,较原模型提升 2.1%。相比之下 CBAM 注意力机制在 mAP@0.5 指标上与三重注意力机制数值接近,但前者召回率仅 75.1%,较三重注意力机制低 1.3%。这表明 CBAM 注意力机制在检测过程中更容易出现漏检现象。

实验结果表明 YOLOv11 中融入 ECA 和 CBAM 注意力机制都可以显著提升模型对钢材表面缺陷的检测能力。具体而言,ECA 作为一种通道注意力,其核心优势在于通过平均池化后的特征施加一维卷积操作,增强跨通道信息交互能力;但由于 ECA 注意力机制仅聚焦于通道维度的特征优化,忽略了空间信息的重要性,一定程度上限制了其性能上限。相比之下,DTA 注意力同时关注通

道与空间双维度的特征信息,强化了模型的全局感知能力,因此在钢材表面缺陷检测任务中比 ECA 注意力机制漏检率更低。CBAM 注意力虽然同样兼顾通道和空间维度信息,但 DTA 注意力采用“旋转+并行”替代 CBAM 的“压缩+串行”,在几乎不增加参数量的前提下,比 CBAM 更好的提升全局化注意力映射。

为更直观地展示不同注意力机制对缺陷特征的表征能力,本文采用 Grad-CAM 方法生成各类注意力机制对应的特征热力图,并进行可视化对比分析,结果如图 8 所示。

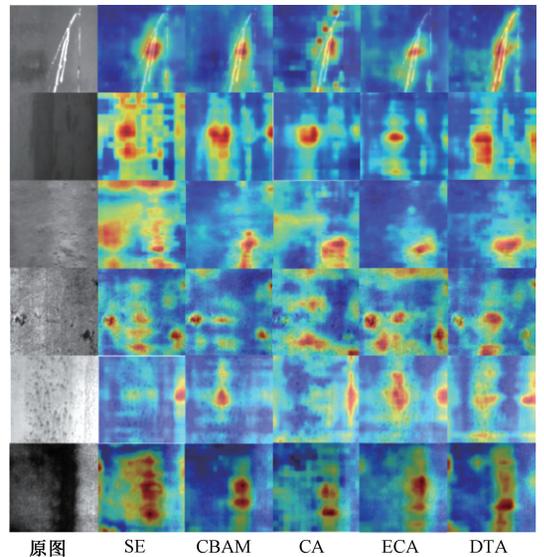


图 8 多种注意力机制可视化热力图

Fig. 8 Visualization heatmaps of various attention mechanisms

从图 8 中可以看出,相较于 SE、CBAM、CA 和 ECA 等主流注意力机制,本文提出的 DTA 机制所生成的热力图与缺陷的真实形状及空间位置高度吻合,尤其是在低对比度、微小或形态不规则的缺陷区域(如细长裂纹、微小麻点等)仍能保持较高的激活响应值。DTA 机制通过跨维度旋转交互与可变形卷积的融合设计,实现了对缺陷区域更精准的注意力聚焦,有效抑制了非缺陷区域的干扰,展现出优异的特征选择与空间感知能力,从而在视觉上验证了其在提升模型对钢材表面缺陷检测精度方面的显著优势。

2.6 模型对比实验

为了进一步验证本研究改进 YOLOv11n 算法的有效性,在相同实验环境及相同数据集划分、输入分辨率、训练轮次及优化器设置下,将本文算法与当前主流的网络进行对比实验。包括 Faster-RCNN、SSD、YOLOv5s、YOLOv7、YOLOv8n、YOLOv10n、YOLOv11n、文献[6]方法和文献[10]方法。实验结果如表 4 所示。

表 4 模型对比实验结果

Table 4 Model comparison results

模型	$mAP@0.5/\%$	Params/M	FPS/s	GFLOPs
Faster-RCNN	71.9	108.0	16.0	350.3
SSD	69.8	92.1	34.0	62.5
YOLOv5s	75.2	7.0	74.0	15.8
YOLOv7	74.3	37.6	104.0	105.7
YOLOv8n	73.8	3.0	107.0	8.2
YOLOv10n	75.4	2.7	113.0	28.9
YOLOv11n	75.9	2.6	118.0	6.4
文献[6]	79.4	104.6	14.1	347.4
文献[10]	79.3	16.5	41.3	110.8
本文	81.9	2.3	109.0	5.9

由表 4 可见,本文改进算法平均精度相较 Faster-RCNN 和 SSD 分别提升了 10% 和 12.1%,参数量 (Params) 和计算量 (GFLOPs) 大大减少。这表明本文模型计算空间复杂度和计算时间复杂度更低,占用设备显存少的同时运算速度更快,在资源受限情况下性能更高。

在 YOLO 系列模型中,本文算法在 NEU-DET 数据集上具有显著优势。虽然 FPS 相对于原 YOLOv11n 网络降低了 8 帧,但平均精度提高了 6%。从参数量上相比也小于 YOLO 系列对比模型。

从钢材表面缺陷具有多尺度分布和形态学异质性的特点,实验结果说明改进后模型在多尺度特征提取多形态特征感知方面得到了极大的改善。精度提高的同时也有着较快的检测速度,极大满足实时性需求。综上所述,本文模型在钢材表面缺陷检测中具有较高的通用性和实用性。

2.7 泛化性能验证实验

为验证本文提出模型在未知数据分布下的泛化能力与鲁棒性,本研究额外选取了公开的金属表面缺陷数据集 GC10-DET 作为外部测试集。GC10-DET 包含 10 大类金属表面缺陷,如划痕、凹坑、焊斑等,其图像背景、缺陷形态及光照条件与 NEU-DET 数据集存在显著差异。本实验将在 NEU-DET 上训练好的模型直接在 GC10-DET 的测试集上进行评估,结果如表 5 所示。

由表 5 可知,在未针对 GC10-DET 数据进行任何再训练的情况下,本文模型的 $mAP@0.5$ 达到了 74.6%,相较于基线 YOLOv11n 提升了 4.1%,同时也优于其他 YOLO 系列轻量级模型。这一结果充分表明,本文所提出的多尺度动态卷积、可变形注意力等模块,有效增强了模型学习通用性特征的能力,使其能够更好地适应不同

来源、不同特性的工业缺陷数据,具有更强的泛化性和实用性。尽管与在 NEU-DET 上测试的绝对精度相比有所下降,但其相对性能优势依然稳定,证明了本文改进方案的有效性。

表 5 模型泛化实验结果

Table 5 Model generalization experiment results

模型	$mAP@0.5/\%$	Params/M	P/%	R/%
YOLOv8n	69.5	3.0	71.9	67.8
YOLOv10n	70.1	2.7	72.5	68.3
YOLOv11n	70.5	2.6	72.0	68.9
本文	74.6	2.3	76.8	72.1

2.8 实验结果可视化分析

为了直观比较基线模型和本文研究算法在钢材表面缺陷检测上的检测性能差异,分别使用 YOLOv11n 模型和本文研究模型在测试集上进行测试,并对检测结果进行可视化分析。如图 9 所示。

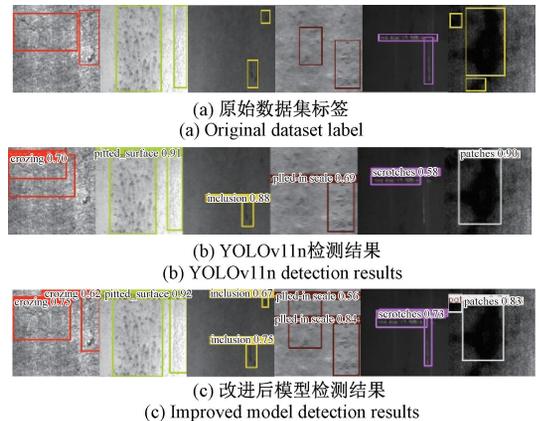
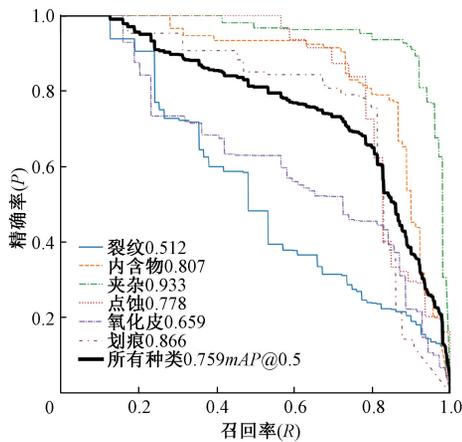


图 9 可视化检测图

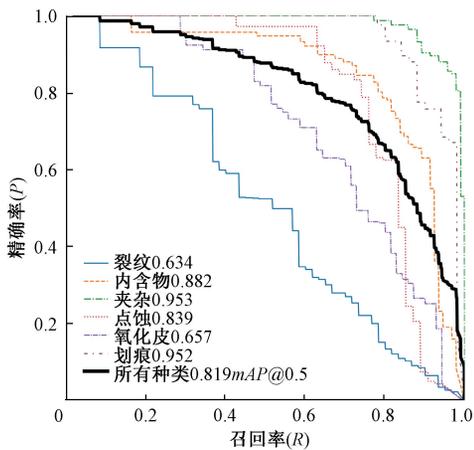
Fig. 9 Visualized detection results

改进后模型能够检测原模型中漏检的缺陷,并对多尺度、对比度低的缺陷检测效果也有很大改善,证明了改善后模型可以较好地避免漏检情况,在检测性能上具备一定优势。

此外,为更全面评估模型在不同置信度阈值下的综合性能,本文绘制了精确率-召回率 (precision-recall, PR) 曲线进行量化分析。如图 10 所示,本文算法的 PR 曲线整体更偏向右上方,所围面积显著大于基线模型,表明其在多个阈值下均能同时保持较高的精确率与召回率,综合检测性能更为优越。该结果与 $mAP@0.5$ 的提升相一致,进一步验证了本文方法在提升模型判别能力方面的有效性。



(a) YOLOv11n 各类别 P - R 曲线
(a) P - R curve of various categories for YOLOv11n



(b) 改进模型各类别 P - R 曲线
(b) Improved model's P - R curves for various categories

图 10 改进前后各类别 P - R 曲线

Fig. 10 P - R curves of various categories before and after improvement

3 结 论

针对钢材表面缺陷中多尺度特征提取不足、小目标漏检率高以及边界框回归精度不佳等问题,本研究基于 YOLOv11n 模型,提出了一种改进的多尺度钢材表面缺陷检测方法。设计 MDC 多尺度动态卷积块和 DRF 动态残差融合模块,有效提升了模型对不同形态缺陷的捕捉能力;改进 DTA 可变形三元注意力机制,增强了模型对微小缺陷和低对比度缺陷的感知能力;使用 Shape-IoU 损失函数优化边界框回归精度。实验表明,改进后的模型在 NEU-DET 数据集上 $mAP@0.5$ 达到了 81.9%,较基线 YOLOv11n 提升了 6%。同时保持了较低的参数和计算量,综合性能对比与复杂度分析表明,该模型在保持高精度的同时,具备了在边缘终端设备上实现实时检测的潜

力。为钢材表面缺陷检测提供了一种高效、可靠的解决方案,具有较高的使用价值。

参考文献

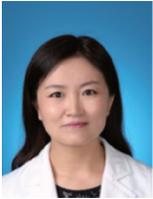
- [1] 马俊杰,张继红,王强,等. PGS-YOLO:一种轻量高效的带钢表面缺陷检测模型[J]. 电子测量与仪器学报, 2025, 39(8): 156-167.
MA J J, ZHANG J H, WANG Q, et al. PGS-YOLO: A lightweight and efficient model for strip steel surface defect detection[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(8): 156-167.
- [2] 许鹏,刘柏霖,陈亚雄. 基于差分柔性涡流探头的铁轨裂纹检测方法[J]. 中国机械工程, 2023, 34(20): 2419-2427.
XU P, LIU B L, CHEN Y X. Rail crack detection method based on differential flexible eddy current probe[J]. China Mechanical Engineering, 2023, 34(20): 2419-2427.
- [3] 赵朗月,吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报, 2022, 43(1): 198-219.
ZHAO L Y, WU Y Q. Research progress of surface defect detection methods based on machine vision[J]. Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.
- [4] 胥姜苗,曹爽,管海燕. 基于改进 YOLOv8 的轻量化钢材表面缺陷检测方法[J/OL]. 电子测量技术, 1-13 [2025-09-02].
XU J M, CAO SH, GUAN H Y. Lightweight steel surface defect detection method based on improved YOLOv8[J/OL]. Electronic Measurement Technology, 1-13[2025-09-02].
- [5] 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, 2016: 779-788.
- [6] JIANG P Y, ERGU D, LIU F Y, et al. A review of YOLO algorithm developments[J]. Procedia Computer Science, 2022, 199: 1066-1073.
- [7] DIWAN T, ANIRUDH G, TEMBHURNE J V. Object detection using YOLO: Challenges, architectural successors, datasets and applications[J]. Multimedia Tools and Applications, 2023, 82(6): 9243-9275.
- [8] SU P, HAN H Z, LIU M, et al. MOD-YOLO: Rethinking the YOLO architecture at the level of feature information and applying it to crack detection[J]. Expert

- Systems with Applications, 2024, 237: 121346.
- [9] CHEN Y M, YUAN X B, WANG J B, et al. YOLO-MS: Rethinking multi-scale representation learning for real-time object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(6): 4240-4252.
- [10] HUSSAIN M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection [J]. Machines, 2023, 11(7): 677.
- [11] ZHANG ZH SH, QIAO S Y, XIE C H, et al. Single-shot object detection with enriched semantics[C]. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018: 5813-5821.
- [12] REN SH 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, 2016, 39(6): 1137-1149.
- [13] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. 2017 IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [14] LIU R Q, HUANG M, GAO ZH M, et al. MSC-DNet: An efficient detector with multi-scale context for defect detection on strip steel surface[J]. Measurement, 2023, 209: 112467.
- [15] 张彪, 荀荣科, 许家忠. 基于深度学习的工业轴承缺陷检测算法研究[J]. 仪器仪表学报, 2025, 46(4): 136-149.
- ZHANG B, XUN R K, XU J ZH. Research on industrial bearing defect detection algorithm based on deep learning[J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 136-149.
- [16] 凌晓, 刘露, 孙宝财, 等. 基于 DSG-ResNet34 的聚乙烯燃气管道电熔焊接缺陷检测 [J]. 仪器仪表学报, 2025, 46(6): 228-240.
- LING X, LIU L, SUN B C, et al. Detection of electrofusion welding defects in polyethylene gas pipelines based on DSG-ResNet34[J]. Chinese Journal of Scientific Instrument, 2025, 46(6): 228-240.
- [17] ZHANG H K, LI S Q, MIAO Q Q, et al. Surface defect detection of hot rolled steel based on multi-scale feature fusion and attention mechanism residual block[J]. Scientific Reports, 2024, 14(1): 7671.
- [18] SONG H Y. RSTD-YOLOv7: A steel surface defect detection based on improved YOLOv7 [J]. Scientific Reports, 2025, 15(1): 19649.
- [19] 岳剑峰, 李伟明, 宁黎华, 等. 基于 YOLO-DEFW 的焊缝缺陷实时检测算法研究 [J]. 中国激光, 2025, 52(8): 64-76.
- YUE J F, LI W M, NING L H, et al. Research on real-time detection algorithm of weld defects based on YOLO-DEFW[J]. Chinese Journal of Lasers, 2025, 52(8): 64-76.
- [20] CHAO CH, MU X Y, GUO ZH H, et al. IAMF-YOLO: Metal surface defect detection based on improved YOLOv8[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74(1): 3548198.
- [21] 徐森, 徐畅, 花小朋, 等. 改进 YOLOv8n 的热轧钢材表面缺陷检测 [J/OL]. 哈尔滨工程大学学报, 1-8 [2025-07-18].
- XU S, XU CH, HUA X P, et al. Improved YOLOv8n for hot-rolled steel surface defect detection [J/OL]. Journal of Harbin Engineering University, 1-8 [2025-07-18].
- [22] 张小龙, 贾渊, 黄杰. 轻量化 YOLO 动态多尺度注意力钢材缺陷检测 [J/OL]. 计算机工程与应用, 1-13 [2025-10-30].
- ZHANG X L, JIA Y, HUANG J. Lightweight YOLO with dynamic multi-scale attention for steel defect detection[J/OL]. Computer Engineering and Applications, 1-13 [2025-10-30].
- [23] 雷富强, 马刘文, 关鹏, 等. 基于 DFD-YOLOv11n 的钢材装备表面缺陷检测算法研究 [J/OL]. 计算机工程与应用, 1-15 [2025-10-30].
- LEI F Q, MA L W, GUAN P, et al. Research on surface defect detection algorithm for steel equipment based on DFD-YOLOv11n[J/OL]. Computer Engineering and Applications, 1-15 [2025-10-30].
- [24] SHI D. Transnext: Robust foveal visual perception for vision transformers[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 17773-17783.
- [25] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: Convolutional triplet attention module [C]. 2021 IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 3138-3147.
- [26] ZHU X ZH, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern

Recognition, 2019: 9300-9308.

- [27] ZHANG H, ZHANG SH J. Shape-IoU: More accurate metric considering bounding box shape and scale [J]. ArXiv preprint arXiv:2312.17663, 2023.
- [28] BAO Y Q, SONG K CH, LIU J, et al. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [29] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [30] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]. Computer Vision-ECCV 2018, 2018: 3-19.
- [31] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13708-13717.
- [32] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11531-11539.

作者简介



赵海丽 (通信作者), 2000 年于长春光学精密机械学院获得学士学位, 2003 年于长春理工大学获得硕士学位, 2009 年于长春理工大学获得博士学位, 现为长春理工大学教授, 主要研究方向为军工光电测试技术及仪器、光电检测与智能信息处理、装备信息化

性能测试。

E-mail: zhljcc@126.com

Zhao Haili (Corresponding author) received her B. Sc. degree from Changchun Institute of Optics and Fine Mechanics in 2000, her M. Sc. and Ph. D. degrees both from Changchun University of Science and Technology in 2003 and 2009, respectively. She is currently a professor at Changchun University of Science and Technology. Her main research interests include military photoelectric testing technology and instrumentation, photoelectric detection and intelligent information processing, and equipment informatization performance testing.



狄子隆, 2022 年于长春理工大学获得学士学位, 现为长春理工大学电子信息工程学院硕士研究生, 主要研究方向为信号与信息处理技术。

E-mail: 15943322398@163.com

Di Zilong received his B. Sc. degree from Changchun University of Science and Technology in 2022. He is currently pursuing his master's degree in the School of Electronic and Information Engineering at Changchun University of Science and Technology. His main research interest includes signal and information processing technology.



景文博, 2003 年于长春光学精密机械学院获得学士学位, 2006 年于长春理工大学获得硕士学位, 2012 年于长春理工大学获得博士学位, 现为长春理工大学教授, 主要研究方向为计算机视觉与图像处理、先进成像与计算光学、机器学习、光电检测、激光性能检测与波前传感技术。

E-mail: wenbojing@cust.edu.cn

Jing Wenbo received his B. Sc. degree from Changchun Institute of Optics and Fine Mechanics in 2003, his M. Sc. and Ph. D. degrees both from Changchun University of Science and Technology in 2006 and 2012, respectively. He is currently a professor at Changchun University of Science and Technology. His main research interests include computer vision and image processing, advanced imaging and computational optics, machine learning, photoelectric detection, laser performance testing, and wavefront sensing technology.



宋明喆, 2022 年于长春理工大学获得学士学位, 现为长春理工大学电子信息工程学院硕士研究生, 主要研究方向为模式识别与人工智能。

E-mail: 18231318834@163.com

Song Mingzhe received his B. Sc. degree from Changchun University of Science and Technology in 2022. He is currently a master's candidate in the School of Electronic and Information Engineering at Changchun University of Science and Technology. His main research interest includes pattern recognition and artificial intelligence.