

基于改进 YOLOv12 的轻量级焊缝缺陷检测模型研究

戚玉亮^{1,2,3}, 黄柯柯¹, 王建波⁴, 郭嵩⁵

(1. 广州建筑产业研究院有限公司, 广东 广州 510663; 2. 广州建筑股份有限公司, 广东 广州 510288;
3. 广州建筑产业开发有限公司, 广东 广州 223456; 4. 山东大学 齐鲁交通学院, 山东 济南 250061;
5. 山东农业大学 水利土木工程学院, 山东 泰安 271000)

摘要: 焊缝质量与工业装备的结构完整性、运行安全直接相关, 高效、准确地识别焊缝表面缺陷是实现智能制造与质量控制的关键。传统焊缝缺陷检测方法在复杂干扰场景下存在检测精度低、泛化能力不足等问题。基于此, 提出一种基于 YOLOv12 的轻量化改进模型 YOLOv12-DSD, 以实现对接缝表面缺陷的高效精准检测。首先, 在特征提取阶段引入动态卷积机制, 搭建 C3k2_D 模块, 提升模型对复杂纹理特征、多尺度缺陷目标的适配能力; 其次, 在边界框回归阶段, 通过设计 SMPDIOU 损失函数, 提升模型对目标几何结构差异和定位误差的建模能力; 最后, 在检测头引入 Dyhead 模块, 依靠尺度、空间、任务 3 个维度的注意力机制, 加强多类别缺陷的特征表达与判别能力。基于包含 1 414 张焊缝缺陷图像的数据集开展消融实验和对比实验, 结果表明, YOLOv12-DSD 模型参数量为 2.13 M, 精确率为 92.41%、召回率为 74.80%、mAP@0.5 为 87.66%, 整体性能优于现有主流检测算法。该模型在轻量化的基础上提升了检测性能, 拥有较好的工业落地部署价值。

关键词: 焊缝缺陷检测; 动态卷积; 轻量化模型; YOLOv12

DOI: 10.11907/rjtk.251628

中图分类号: TP399

文献标识码: A

文章编号: 1672-7800(XXXX)0XX-0001-09

扫描二维码阅读全文:



Research on a Lightweight Weld Defect Detection Model Based on the Improved YOLOv12

QI Yuliang^{1,2,3}, HUANG Keke¹, WANG Jianbo⁴, GUO Song⁵

(1. Guangzhou Construction Industry Research Institute Co., Ltd., Guangzhou 510663, China; 2. Guangzhou Construction Co., Ltd., Guangzhou 510288, China; 3. Guangzhou Construction Industrial Development Co., Ltd., Guangzhou 223456, China;
4. School of Qilu Transportation, Shandong University, Ji'nan 250061, China; 5. College of Civil and Hydraulic Engineering, Shandong Agricultural University, Tai'an 271000, China)

Abstract: Weld quality is directly related to the structural integrity and operational safety of industrial equipment. Efficient and accurate detection of weld surface defects is essential for intelligent manufacturing and quality control. However, traditional weld defect detection methods often suffer from limited detection accuracy and poor generalization ability in complex interference environments. To address these challenges, this paper proposes a lightweight improved model, termed YOLOv12-DSD, based on YOLOv12 for efficient and accurate weld surface defect detection. First, a dynamic convolution mechanism is introduced into the feature extraction stage, and a C3k2_D module is constructed to enhance the model's adaptability to complex texture features and multi-scale defect targets. Second, an SMPDIOU loss function is designed for the bounding box regression stage to improve the modeling capability of geometric structural differences and localization errors. Finally, a Dy-Head module is incorporated into the detection head, leveraging attention mechanisms across scale, spatial, and task dimensions to strength-

收稿日期: 2025-11-21

基金项目: 国家自然科学基金面上项目(52378332); 广州市建筑集团有限公司科技计划项目([2022]-KJ014, [2024]-KJ061)

作者简介: 戚玉亮(1981-), 男, 博士, 广州建筑产业研究院有限公司正高级工程师, 研究方向为智能建造、装配式建筑; 黄柯柯(1991-), 女, 硕士, 广州建筑产业研究院有限公司工程师, 研究方向为智能建造; 王建波(1994-), 男, 山东大学齐鲁交通学院工程师, 研究方向为机械工程; 郭嵩(2002-), 男, 山东农业大学水利土木工程学院学生, 研究方向为智能建造。本文通讯作者: 王建波。

en feature representation and discrimination for multiple defect categories. Ablation and comparative experiments are conducted on a weld defect dataset containing 1,414 images. The results demonstrate that the proposed YOLOv12-DSD model contains only 2.13 M parameters while achieving a precision of 92.41%, a recall of 74.80%, and an mAP@0.5 of 87.66%, outperforming existing mainstream object detection algorithms in overall performance. The experimental results indicate that the proposed model improves detection performance while maintaining a lightweight architecture, demonstrating strong potential for industrial deployment and practical applications.

Key Words: weld defect detection; dynamic convolution; lightweight model; YOLOv12

0 引言

焊接作为连接金属材料的重要工艺手段,广泛应用于航空航天、能源装备、轨道交通、压力容器及桥梁工程等领域。与此同时,受工艺参数波动、材料组织不均匀及施工操作误差等因素影响,焊缝易出现裂纹、未熔合、气孔、夹渣等缺陷,若不能及时、可靠地检出上述缺陷,将导致结构完整性下降、服役性能退化,严重时可引发重大安全事故。因此,开发高效、可靠的焊缝缺陷检测技术对于保障产品质量与工业装备安全运行具有重要意义。X射线检测因其穿透能力强、成像清晰,已成为焊缝内部缺陷检测的首选方法^[1]。近年来,现代X射线检测逐步向自动化、智能化方向发展,数字成像技术与智能分析方法的融合显著提升了检测效率与结果可靠性^[2]。同时,高帧率成像及同步辐射X射线技术的应用,实现了对焊接过程中内部缺陷演化过程的实时观测,为缺陷形成机理研究提供了重要支撑^[3-4]。

深度学习,尤其是近些年兴起的基于图像的目标检测算法,在焊缝缺陷智能识别领域展现出巨大潜力。与传统图像处理依靠人工设计特征提取器不同,深度神经网络能够从数据自动学习具有判别能力的深层特征表示,大幅提升缺陷识别准确率^[5-6]。针对X射线图像背景繁杂、纹理波动大的问题,深度学习算法也具备优异的鲁棒性^[7-8]。当前主流目标检测算法主要分为两类:一是以FasterRCNN为代表的双阶段算法,通过生成候选框再完成分类,检测精度优异^[9];二是以YOLO系列为代表的单阶段算法,实现图像端到端检测定位,实时性与落地部署优势突出^[10],但该类算法在强噪声、复杂纹理、微小缺陷场景中容易出现漏检、误检。此外,基于Transformer的检测模型逐步发展,全局特征建模性能优于传统算法^[11]。

大量研究表明,深度学习技术与X射线图像分析方法的融合显著推动了焊缝缺陷智能检测发展。现有研究主要围绕数据增强、网络结构优化和检测策略改进等方面展开。针对工业场景中样本获取成本高、缺陷样本分布不平衡等问题,研究人员采用数据增强、迁移学习等方法扩充训练数据并提升模型泛化能力^[12-13]。在网络结构设计方面,引入注意力机制、特征金字塔网络(Feature Pyramid Network, FPN)以及多尺度特征融合策略,能够有效增强模型对复杂背景和不同尺度缺陷目标的特征表征能力^[14-15]。同时,随着边缘计算与智能检测设备的发展,兼顾检测精

度与计算效率的轻量化网络逐渐成为研究热点^[16]。

除网络结构优化外,目标检测性能的提升还依赖于边界框回归机制与特征增强策略的改进。为提高模型定位精度与收敛稳定性,研究者提出了多种改进的边界框回归损失函数,并通过优化目标间的空间关系建模提升检测鲁棒性^[17]。此外,通过融合空间注意机制、动态卷积和高分辨率图像处理策略,模型对微小缺陷和复杂纹理特征的识别能力得到了进一步增强^[18-19]。尽管现有研究在焊缝缺陷检测领域取得了显著进展,但面向实际工业场景的应用仍面临若干挑战。首先,焊缝缺陷类型复杂,不同缺陷在形态特征、尺寸分布及纹理表现方面存在显著差异,对模型特征提取能力提出了较高要求;其次,焊接过程中产生的金属反光、飞溅、噪声干扰及成像质量波动等因素,容易造成缺陷特征弱化甚至淹没,从而影响检测精度;最后,焊缝缺陷具有明显的尺度不平衡特征,其中微小缺陷占比较高,而现有检测模型对小目标特征的表达与定位能力仍存在不足^[20]。

针对上述问题,本文提出一种基于YOLOv12的改进焊缝表面缺陷检测模型YOLOv12-DSD。该模型通过构建动态特征提取模块、设计几何约束增强边界框回归损失函数以及引入多维动态检测头结构,提升复杂工业场景下焊缝缺陷的特征表征能力、多尺度目标检测能力和目标定位精度,进而实现轻量化与高精度兼顾的焊缝缺陷智能检测。

本文主要创新与贡献如下:

(1)构建基于动态卷积的特征提取模块。在主干网络中引入动态卷积机制,设计C3k2_D模块,通过卷积核参数的自适应调整增强模型对复杂纹理特征和多尺度缺陷目标的表征能力,提高不同类型焊缝缺陷的特征提取效果。

(2)提出SMPDIoU边界框回归损失函数。融合最小点距离约束与方向信息约束,增强对目标几何结构差异及定位误差的建模能力,提高边界框回归精度和训练收敛稳定性。

(3)引入基于DyHead的多维动态检测头。利用尺度、空间和任务3个维度的注意力机制建立跨层级特征交互路径,增强多类别缺陷目标的特征表达能力和判别能力,提高复杂场景下的检测性能。

1 YOLOv12基准模型

近年来,YOLO系列目标检测算法凭借端到端检测框架、高检测精度和实时推理能力,在工业视觉检测领域得

到了广泛应用。与 Faster R-CNN 等两阶段目标检测算法相比, YOLO 系列无需候选区域生成过程, 具有更高的检测效率和更低的计算开销, 更适用于对实时性要求较高的工业在线检测场景。作为 YOLO 系列的最新版本, YOLOv12 在检测精度、推理速度和模型轻量化之间取得了较好的平衡, 并通过引入注意力机制与高效特征聚合结构进一步增强了特征建模能力^[12]。因此, 本文选择 YOLOv12 作为基准模型开展焊缝缺陷检测研究。

YOLOv12 的整体结构如图 1 所示, 其主要由主干网络 (Backbone)、特征融合网络 (Neck) 和检测头 (Head) 3 部分组成。其中, 主干网络负责提取输入图像的多层次特征信息; 特征融合网络通过跨尺度特征融合实现语义信息与细节信息的有效交互; 检测头则完成目标分类与边界框回归任务。与传统 YOLO 模型主要依赖卷积堆叠进行特征提取不同, YOLOv12 在主干网络中引入区域注意力机制。该机制通过将特征图划分为若干水平或垂直区域, 在局部区域内建立特征关联关系, 实现感受野的区域化建模。相较于基于滑动窗口或窗口偏移的注意力机制, 该方法仅需简单的特征重组操作即可完成区域划分, 在降低计算复杂度的同时保持对局部上下文信息的建模能力, 从而增强关键区域特征表达能力。

为提高注意力机制下网络训练的稳定性, YOLOv12 进一步引入残差高效层聚合网络。该结构在 ELAN 模块基础上增加跨层残差连接, 通过强化特征复用和梯度传播能力。同时, YOLOv12 采用 Flash Attention 对传统注意力计算过程进行优化, 以降低内存访问开销和计算复杂度, 提高训练与推理效率。此外, 通过适当调整注意力模块深度和前馈网络比例, 使模型能够更好地适应实时目标检测任务。尽管 YOLOv12 在检测精度与推理效率之间取得了良好的平衡, 但面对焊缝缺陷目标尺度差异大、纹理信息复杂和背景噪声干扰严重等问题, 其在细粒度特征提取、边界框回归精度和多尺度特征表达能力方面仍存在进一步优化空间。因此, 本文在 YOLOv12 基础上构建 YOLOv12-DSD 模型, 以提升焊缝缺陷检测性能。

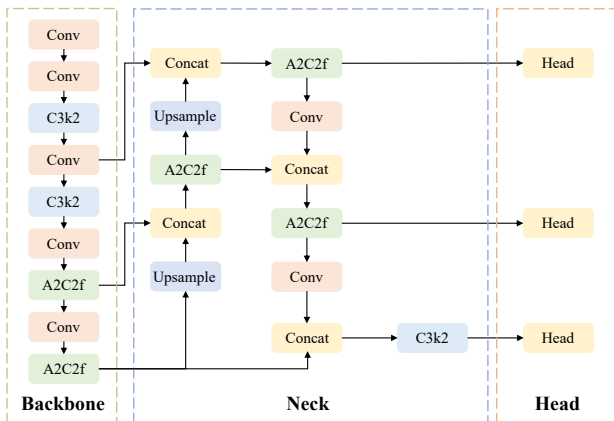


Fig. 1 Architecture of YOLO12

图1 YOLO12网络结构

2 YOLOv12-DSD

本文基于 YOLOv12 模型提出一种改进的焊缝表面缺陷检测方法 YOLOv12-DSD, 其整体结构如图 2 所示。YOLOv12-DSD 旨在解决焊缝缺陷检测中存在的多尺度小目标难以有效捕捉、边界框定位精度受限以及检测头对空间上下文与多任务适应性不足等核心问题。现有方法在轻量化卷积结构下缺乏对复杂特征的自适应能力, IoU 类损失函数对角度和宽高比差异刻画不足, 且传统检测头难以充分整合多维度注意力信息, 从而在小目标、高变异性缺陷场景下表现受限。针对这些局限, YOLOv12-DSD 在特征提取阶段对 C3k2 模块引入动态卷积 (Dynamic Convolution) 机制, 根据输入特征动态生成卷积核权重, 在计算成本近似不变的前提下提升参数空间和特征表达能力, 从而增强对多尺度缺陷的感知; 在边界框回归阶段设计 SMP-DIoU 损失函数, 融合 MPDIoU 与 SIoU 中的角度与形状惩罚项, 全面刻画预测框与真实框的几何差异, 优化回归梯度并提升定位精度; 在检测头阶段采用 Dyhead 模块, 通过逐维引入通道、空间与任务注意力机制, 有效增强模型对复杂缺陷模式的响应能力和多任务适应性。综上, YOLOv12-DSD 在保持实时性和轻量化的同时, 有针对性地解决了焊缝缺陷检测中的核心问题, 有效提升了模型检测性能。

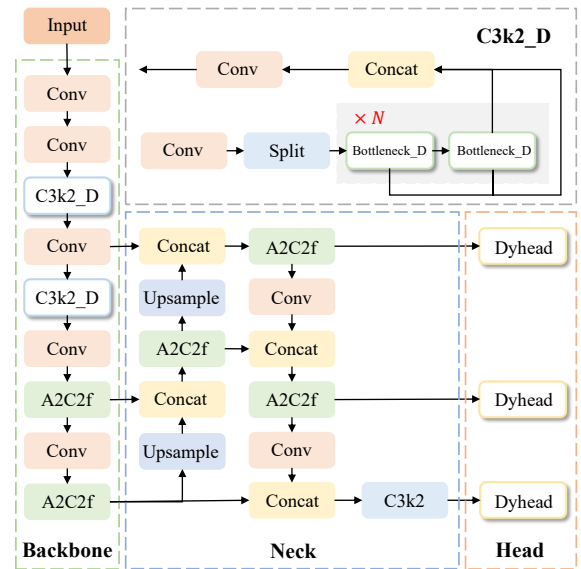


Fig. 2 Architecture of YOLO-DSD

图2 YOLO-DSD网络结构

2.1 C3k2_D 模块

在目标检测领域, 模型结构的轻量化与特征提取能力之间存在显著的权衡关系。为在保证推理效率的同时增强网络表达能力, 本文提出在 YOLOv12 骨干网络的 C3k2 模块中引入动态卷积 (Dynamic Convolution) 机制^[22], 从而在参数可控的前提下提升模型对多尺度与复杂结构的特

征感知能力。改进后的 C3k2_D 模块如图 3 所示。

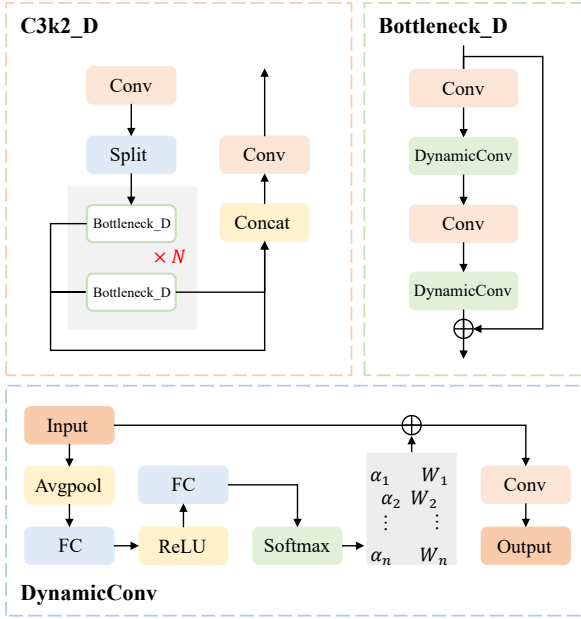


Fig. 3 Architecture of C3k2_D

图3 C3k2_D结构

传统卷积操作依赖静态的卷积核参数,即网络训练完成后,每一层的卷积核在整个推理过程中保持不变。这种静态机制虽然计算效率较高,但缺乏对不同输入图像内容进行自适应调整的能力,难以充分捕捉多样化特征表达。为解决这一限制,DynamicConv通过引入样本依赖的权重生成策略,能够依据当前输入特征动态构建卷积核,从而实现按需调整的特征提取。DynamicConv的核心思想是引入多个“专家卷积核”(Expert Kernels)并利用一个条件生成模块为每个输入样本分配融合系数,以构造最终参与计算的卷积核权重。具体而言,给定输入特征图,首先对其进行全局平均池化以获得通道级别的描述向量,该向量随后被输入到一个由两层全连接层构成的多层感知机(MLP)中,以生成归一化的融合权重。这些权重通过Softmax函数处理后,用于对多个专家卷积核进行线性加权组合,形成适应当前输入的动态卷积核。该机制可形式化为如式(1)所示。

$$W' = \sum_{i=1}^M \alpha_i W_i \quad (1)$$

其中, W' 为动态卷积核权重, M 为动态专家数量, α_i 为第 i 个动态专家的权重系数,且满足 $\sum_{i=1}^M \alpha_i = 1$, W_i 为卷积核权重。

在实际操作中,这一机制引入了3个阶段的计算:①生成融合权重的MLP模块;②对专家核进行加权融合;③基于融合卷积核执行常规卷积操作。尽管新增了权重生成和融合步骤,但由于前两阶段主要涉及低维度的矢量操作,其额外计算开销相较于整体卷积操作可忽略不计。通过将DynamicConv嵌入YOLOv12中C3k2模块的瓶颈结

构,可显著增强模块对特征的表达能力。该模块原为复合卷积结构,集成多个标准卷积与激活单元以建模局部与全局信息,改进后通过动态机制增加了网络对样本内容的敏感度。在不显著增加计算成本的前提下,实现了参数量的线性增长,进而提高了网络的鲁棒性与泛化能力。DynamicConv和标准卷积的参数量和计算量之比可表示为如式(2)、式(3)所示。

$$\begin{aligned} Ratio_{param} &= \frac{M \times C_{out} \times C_{in} \times K \times K + C_{in}^2 + M \times C_{in}}{C_{out} \times C_{in} \times K \times K} \\ &= \frac{C_{in}}{C_{out} \times K \times K} + \frac{M}{C_{out} \times K \times K} + M \end{aligned} \quad (2)$$

$$\begin{aligned} &\approx \frac{1}{K \times K} + M \\ &(M \ll C_{out} \times K \times K, C_{in} \approx C_{out}) \\ Ratio_{FLOs} &= \frac{C_{in}}{W' \times H' \times C_{out} \times K \times K} \\ &+ \frac{M}{W' \times H' \times C_{out} \times K \times K} \\ &+ \frac{M}{W' \times H'} + 1 \\ &\approx 1 \\ &(1 \ll M \ll W' \times H', C_{in} \approx C_{out}) \end{aligned} \quad (3)$$

其中, C_{in} 为输入特征通道数, C_{out} 为输出特征通道数, $K \times K$ 为卷积核的空间大小, $W' \times H'$ 为输出特征图的宽度与高度。以上两式表明,DynamicConv在保持计算量近似不变的同时,可将参数量扩大至原来的 M 倍,为网络提供更丰富的学习能力。这一结构性调整对于目标检测任务中的小目标识别和背景复杂场景具有显著优势。

综上所述,DynamicConv的引入为C3k2模块带来了更强的特征自适应能力,使YOLOv12在保持实时性的同时进一步提升了检测精度。此改进为动态机制在轻量级检测框架中的应用提供了新的思路与可行性验证。

2.2 SMPDIOU 损失函数

损失函数设计对目标检测模型的性能起着关键作用,尤其是在边界框回归过程中。尽管GIoU^[23]、DIoU^[24]、CIoU和EIoU^[25]等方法在优化边界框的重叠程度、中心点距离和尺度差异方面取得了一定成果,但在处理复杂几何关系时仍存在一定的局限性。为进一步提高边界框回归精度,本文在MPDIOU(Minimum Point Distance IoU)基础上,引入SCYLLA-IoU(SIoU)^[26]中关于角度与方向惩罚的思想,构建了一种新的损失函数SMPDIOU,以增强模型对边界框几何结构差异的建模能力。

MPDIOU通过引入最小点对点距离以优化预测框与真实框之间的位置关系,其影响因素如图4所示。

其中, w, h 分别代表边界框的宽和高, d_1 和 d_2 分别代表两框之间的最小点距离。由此可得,MPDIOU的计算公式如式(4)和(5)所示。

$$d_{min}^2 = d_1^2 - d_2^2 \quad (4)$$

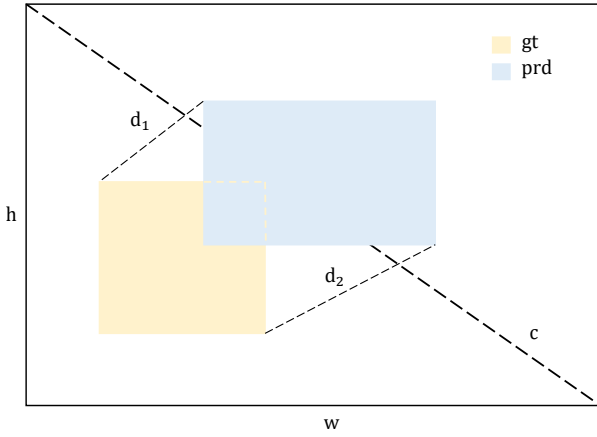


Fig. 4 Factors affecting MPDIoU

图4 MPDIoU 的影响因素

$$\text{MPDIoU} = \text{IoU} - \frac{d_{\min}^2}{c^2} \quad (5)$$

其中, IoU 为标准交并比, $d_{\min}$ 表示预测框与真实框之间的最小点对点距离, c 为包含两框的最小封闭矩形的对角线长度, 用于归一化。

为进一步提升回归过程中对方向性与形状差异的刻画能力, 本文引入了 SIoU 中的两个关键惩罚项: 角度误差项和形状误差项。角度误差项用于度量预测框中心与真实框中心连线方向的偏移程度, 定义如式 (6) 和式 (7) 所示。

$$\Delta_{\theta} = 1 - \cos(\theta) \quad (6)$$

$$\theta = \arctan\left(\frac{y_g - y_p}{x_g - x_p + \epsilon}\right) \quad (7)$$

(x_p, y_p) 与 (x_g, y_g) 分别为预测框与真实框的中心点坐标, ϵ 为防止除零的微小常数, θ 为两中心点连线与水平方向之间的夹角。

形状误差项用于衡量预测框与真实框在宽高比上的差异, 其表达式如式 (8) 所示。

$$\Delta_{\text{shape}} = 1 - e^{-\left(\frac{|w_p - w_g|}{w_g} + \frac{|h_p - h_g|}{h_g}\right)} \quad (8)$$

其中, w_p, h_p 与 w_g, h_g 分别为预测框和真实框的宽度与高度。

最终, 所提出的 SMPDIoU 损失函数可表示为如式 (9) 所示。

$$\mathcal{L}_{\text{SMPDIoU}} = 1 - \text{MPDIoU} + \lambda_{\theta} \cdot \Delta_{\theta} + \lambda_{\text{shape}} \cdot \Delta_{\text{shape}} \quad (9)$$

其中, λ_{θ} 与 λ_{shape} 为调节角度误差与形状误差影响程度的超参数。

为清晰、严谨地说明惩罚项系数的设定依据, 本文对角度误差项权重 λ_{θ} 和形状误差项权重 λ_{shape} 的取值进行了系统性分析。理论分析表明, 角度误差项用于约束预测框与真实框之间的方向偏移, 形状误差项则直接控制边界框的尺度匹配效果, 为保证几何拟合的稳定性, 需合理提升形状项的权重占比。本文用网格搜索法对 λ_{θ} 与 λ_{shape} 作准确调参, 在 $\{0.1, 0.3, 0.5, 0.7\}$ 区间内进行参数组合试验, 并在验证集上对平均精度 (mAP) 和边界框回归误差两者都予以客观评价。实验结果表明, 当 $\lambda_{\theta}=0.3$ 、 $\lambda_{\text{shape}}=0.7$ 时, 模型收敛速度稳定, 检测精度和边界框拟合效果都达到最优。因此, 本文选取该组参数作为模型默认配置, 有效平衡了角度约束与形状约束在整体优化过程中的作用强度。

由于所提损失函数继承了 MPDIoU 对空间位置最小距离的刻画优势, 同时有效改进了边界框方向与几何一致性的建模能力, 因此针对目标偏移大、尺度变化剧烈的检测场景, 本文选取该组参数作为默认配置, 以此平衡角度约束与形状约束在整体优化中的作用强度。

2.3 Dyhead 模块

由于焊缝缺陷检测任务中所处理的图像一般含有尺度不同、种类各异的多种缺陷目标, 故检测模型必然需要有优良的多尺度解析能力, 而缺陷的空间分布特性又对模型性能有直接而重大的影响, 因此模型也必须对空间上下文有充分、可靠的理解。鉴于此, 本文在原始 YOLOv12 架构的基础上, 将检测分支替换为动态感知检测头 (Dyhead)^[27], 由此切实提高模型对细粒度特征的敏感性, 进而改善整体检测性能。具体而言, 所提出的动态检测模块巧妙地融合了尺度、空间、任务 3 类信息感知能力, 以模块化方式对不同维度的注意机制分别建模, 因而特征表达更适应于各种缺陷模式, 判别性也更强。具体而言, 其先对通道维度 (C)、空间维度 (S) 和任务维度 (L) 依次引入专用的注意力单元, 既没有增加过多参数, 又最大限度地提取了关键信息。逐维建模的思路既有利于模型准确地响应复杂缺陷模式, 也使多任务场景下的判别性能得到实质提升。因此, 当面对高变异性的焊缝缺陷时, 所设计的模型能够更稳健、更可靠地完成定位与分类。

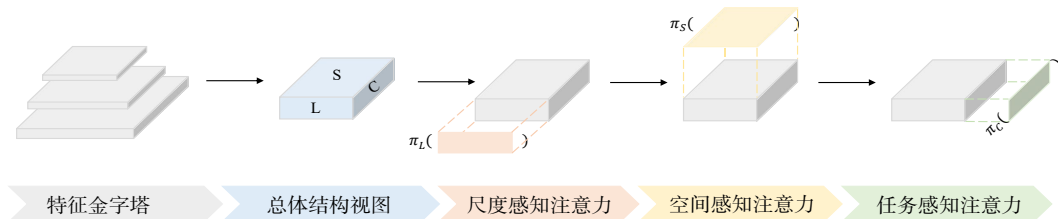


Fig. 5 Architecture of Dyhead

图5 Dyhead 结构示意图

3 实验与结果分析

3.1 数据集

为验证所提出模型在焊缝缺陷检测任务中的有效性,本文构建了焊缝缺陷图像数据集。数据来源包括实验室自主采集图像和公开数据集,共收集原始焊缝图像 986 张。经过数据清洗与质量筛选后,剔除噪声干扰严重、缺陷特征不明显和难以准确标注的样本,最终保留 842 张质量较高的原始图像用于后续实验。为提高模型对复杂场景的适应能力和泛化性能,对筛选后的原始图像进行数据增强处理。具体包括:①随机旋转,旋转角度设置为 $\pm 15^\circ$,用于模拟不同焊接角度下的成像变化;②亮度扰动,亮度调整系数设置为 0.8~1.2,以增强模型对光照变化的适应能力;③噪声增强,通过叠加高斯噪声和椒盐噪声模拟实际工业环境中的成像干扰;④Mosaic 图像拼接,从数据集中随机选取 4 幅图像进行组合,以增强模型对多目标及多尺度目标的学习能力。经数据增强后,共生成 572 张增强图像。为保证实验结果的客观性,增强图像仅用于训练集构建,不参与验证集和测试集划分,从而避免由数据重叠引起的数据泄漏问题。最终数据集包含裂纹、飞溅和无缺陷 4 类样本,各类别原始图像、增强图像及其数量分布如表 1 所示。

Table 1 Category and quantity distribution of weld defect dataset

表 1 焊缝缺陷数据集类别与数量分布 (张)

缺陷类别	原始图像	增强图像	总数
裂纹	210	150	360
气孔	180	120	300
飞溅	220	130	350
无缺陷	232	172	404
总计	842	572	1 414

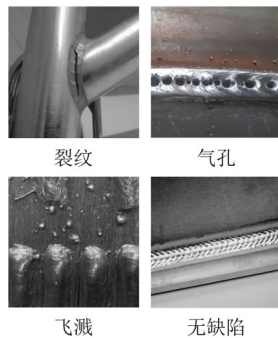


Fig. 6 Real plot of dataset

图 6 数据集真实图

3.2 实验环境

本文是在 Ubuntu 20.04 系统下做的实验,所用 Python 版本为 3.10,深度学习框架为 PyTorch 2.3.0,CUDA 版本为 12.1,配套的硬件为 NVIDIA GeForce RTX 3090 显卡(有 24 GB 显存),硬件条件十分有利于深度学习任务的求解。

模型训练时一共做了 500 个迭代周期,采用了小批量

训练方法,所选优化算法为带动量项的随机梯度下降法(SGD),各参数具体取值为:初始学习率 0.01,动量系数 0.937,权重衰减参数 0.000 5,较好地控制了模型过拟合问题。因此,除特殊说明外,本文后面所列举的诸种实验都严格按此环境及参数设置执行。

3.3 评价指标

为了全面评估检测模型在焊缝缺陷识别任务中的性能,本文采用了基于交并比(IoU)阈值为 0.5 的多个指标进行评价。具体指标包括参数量、精度(Precision)、召回率(Recall)和平均精度均值(mean Average Precision, mAP)。精度用于衡量检测结果的准确性,召回率则反映了模型对真实缺陷的覆盖能力,其计算公式如式(10)和式(11)所示。

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

其中,TP 表示正确识别且预测框与真实框的 IoU 大于 0.5 的样本数,FP 为误判为缺陷的样本数,FN 表示漏检的真实缺陷数量。

平均精度均值(mAP)通过对每一类缺陷的精度—召回率(P-R)曲线下的面积进行积分,并对所有类别的结果取平均值,从而有效地综合评估模型在各类缺陷上的整体检测能力。其计算公式如式(12)所示。

$$mAP = \frac{1}{c} \sum_{i=1}^c \int_0^1 P(R) dR \quad (12)$$

其中,c 表示缺陷类别数量。

3.4 消融实验

由于要考察各模块对模型的实际作用,因此本文设计了 7 组消融实验,并对各组实验数据进行了统计分析。实验结果如表 2 所示。

Table 2 Results of ablation study

表 2 消融实验结果

实验	C3k2_ D	SMPDIoU	Dyhead	Precision	Recall	mAP@ 0.5	Params	FPS
1	√	—	—	88.54	72.96	85.59	2.38	62
2	—	√	—	89.35	72.63	85.68	1.96	59
3	—	—	√	87.91	74.64	86.13	2.42	56
4	√	√	—	88.00	74.83	85.05	3.18	64
5	—	√	√	91.24	73.84	86.85	2.25	66
6	√	—	√	90.34	73.20	86.31	2.36	54
7	√	√	√	92.41	74.80	87.66	2.13	63

由于实验 2 只换了损失函数,但是要获得 SMPDIoU 所需的角点、形状惩罚项,故实验中对回归分支所用的特征通道数做了合理微调,因而参数量从 2.38 M 稍降为 2.36 M。

实验 7 的参数量低于各单模块实验,原因在于 3 个模块集成时对网络作了轻量化整合,共享了部分特征提取层与检测头权重。

本文各实验中所用的参数量都由 PyTorch 自动统计,

故各数值的差别在±0.05M 范围内,属于正常现象。

实验 1~3 依次引入 C3k2_D、SMPDIoU 和 Dyhead 模块,验证了各模块的独立作用。实验结果表明,单独引入 SMPDIoU(实验 2)可将精度提升至 89.35%,mAP@0.5 提升至 85.68%,推理速度为 59 FPS。Dyhead 模块(实验 3)对召回率的提升效果极为显著,召回率可达 74.64%,mAP@0.5 为 86.13%,但推理速度小幅降至 56 FPS,说明该模块在提升检测性能的同时会增加一定的计算负担。实验 4 和实验 6 分别将 C3k2_D 与 SMPDIoU、Dyhead 进行组合,结果表明,两种模块的简单组合虽能优化部分指标,但整体性能仍不及 SMPDIoU 与 Dyhead 的组合方案(实验 5)。实验 5 在参数量仅为 2.25 M 的条件下,取得 91.24% 的精确率、86.85% 的 mAP@0.5,推理速度可达 66 FPS,实现了检测精度与实时性的良好平衡。实验 7 将 3 个模块融合应用,模型的精确率、召回率、mAP@0.5 均达到最优结果,分别为 92.41%、74.80% 和 87.66%,同时参数量仅为 2.13 M,推理速度维持在 63 FPS,有效验证了多模块协同优化的有效性,证明了本文模型的高效性。

3.5 对比实验与可视化结果

为评估 YOLOv12-DSD 模型的性能,将其与当前主流目标检测模型(Faster R-CNN、YOLOv9、YOLOv10、YOLOv11)进行了对比实验,结果如表 3 所示。

Table 3 Comparison results of different models
表 3 不同模型结果比较

模型	Precision	Recall	mAP@0.5	Params	FPS
Faster R-CNN	86.64	73.31	82.92	8.22	35
YOLOv9	88.22	72.00	82.13	3.12	49
YOLOv10	87.78	73.04	81.82	2.70	46
YOLOv11	85.76	70.25	83.42	2.45	53
YOLOv12	88.45	73.38	84.20	2.31	58
PP-YOLO	87.93	72.32	84.52	2.72	55
YOLOv12-DSD	92.41	74.80	87.66	2.13	63

实验结果表明,YOLOv12-DSD 在各项性能指标上均超越了其他对比模型。具体而言,Faster R-CNN 的 Precision、Recall 和 mAP@0.5 分别为 86.64%、73.31% 和 82.92%,但其较大的参数量(8.22 M)和较低的推理速度(35 FPS)使得其计算效率相对较低。相比之下,YOLO 系列模型在保持较低参数量的同时,展现了更优的速度和精度。YOLOv9 的 Precision 为 88.22%,Recall 为 72.00%,mAP@0.5 为 82.13%,参数量为 3.12 M,推理速度为 49 FPS,表现出较好的精度与速度平衡。YOLOv10 进一步优化了网络结构,获得了更高的 Recall(73.04%),但其 mAP@0.5 略有下降(81.82%),参数量降至 2.70 M,推理速度提升至 57 FPS。YOLOv11 则在轻量化设计上有所突破,但其精度有所降低,Precision 为 85.76%,Recall 为 70.25%,推理速度为 53 FPS。

相比之下,YOLOv12 在精度和召回率上取得了显著提升,Precision 达 88.45%,Recall 为 73.38%,mAP@0.5 为 84.20%,并以 58 FPS 的速度展现出优异的实时检测性能。

PP-YOLO 虽在检测精度上(mAP@0.5=84.52%)与 YOLOv12 相近,但其推理速度(55 FPS)略低,导致综合性能相对不足。进一步地,所提出的 YOLOv12-DSD 模型在精度、召回率及综合性能方面均达到更优:Precision 达到 92.41%、Recall 为 74.80%、mAP@0.5 为 87.66%,参数量仅 2.13 M,推理速度达到 63 FPS。

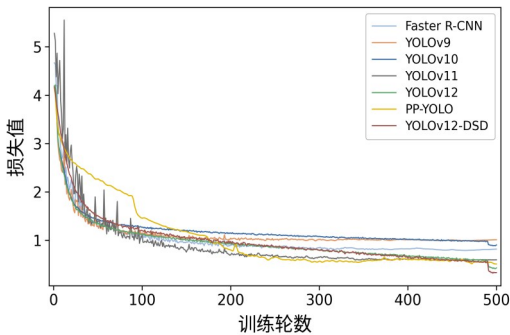


Fig. 7 Loss curve of different models
图 7 不同模型的 Loss 曲线

图 7 展示了 YOLOv12-DSD 与对比模型在训练过程中的 Loss 曲线变化趋势。总体来看,YOLOv12-DSD 在训练初期表现出较快的收敛速度,其总损失值在约 50 个 Epoch 后迅速下降并趋于稳定,表明模型在短时间内已能有效学习关键特征并实现较优的边界框回归与分类判别能力。与之相比,YOLOv12 原始模型在相同训练条件下的 Loss 下降速度略慢,且稳定区间存在一定的震荡,反映出其在特征提取与回归阶段对复杂缺陷的建模能力有限。

从对其他各对比模型(YOLOv9、YOLOv10、Faster R-CNN)的分析中得出,YOLOv12-DSD 在最终 Loss 值上明显优于其他方法,且训练时震荡极小,因此训练稳定性和收敛性都极好。其根本原因是所提出的 SMPDIoU 损失函数对边界框的方向、形状有极强的判别能力,而所引入的 Dyhead 模块又很好地提高了多尺度特征的表示精度,故模型在学习时能够更可靠、更充分地利用差异信息,误差极大减少。

此外,为进一步验证模型在实际焊缝缺陷检测中的效果,本文对 YOLOv12-DSD 在不同类型缺陷样本上的检测结果进行了可视化展示,如图 8、图 9 所示(其中,黄色是标注框,蓝色是预测框)。在图 8 所示的气孔缺陷样本中,YOLOv12-DSD 模型成功识别出多个气孔目标,未出现误检与漏检,验证了 C3k2_D 模块在增强小目标检测能力方面的有效性。而在图 9 的裂纹缺陷中,尽管缺陷具有细长结构与弱对比度等干扰因素,YOLOv12-DSD 模型仍能保持较高的检测准确率,这得益于 Dyhead 模块在任务维度注意机制下对细粒度特征的精确捕捉。

YOLOv12-DSD 在保持低参数量的前提下,显著提升了焊缝缺陷检测任务中的精度、召回率与鲁棒性,展现出优异的检测性能与训练稳定性。通过系统的对比实验、Loss 曲线分析和可视化结果验证,进一步证实了所提模型

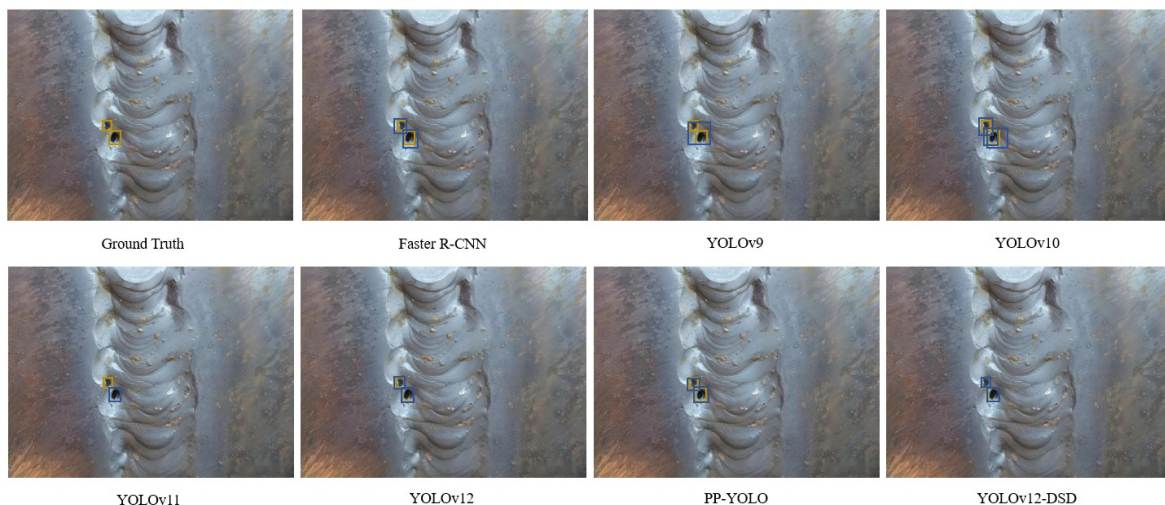


Fig. 8 Visualization results of porosity defect detection

图8 气孔缺陷检测可视化结果

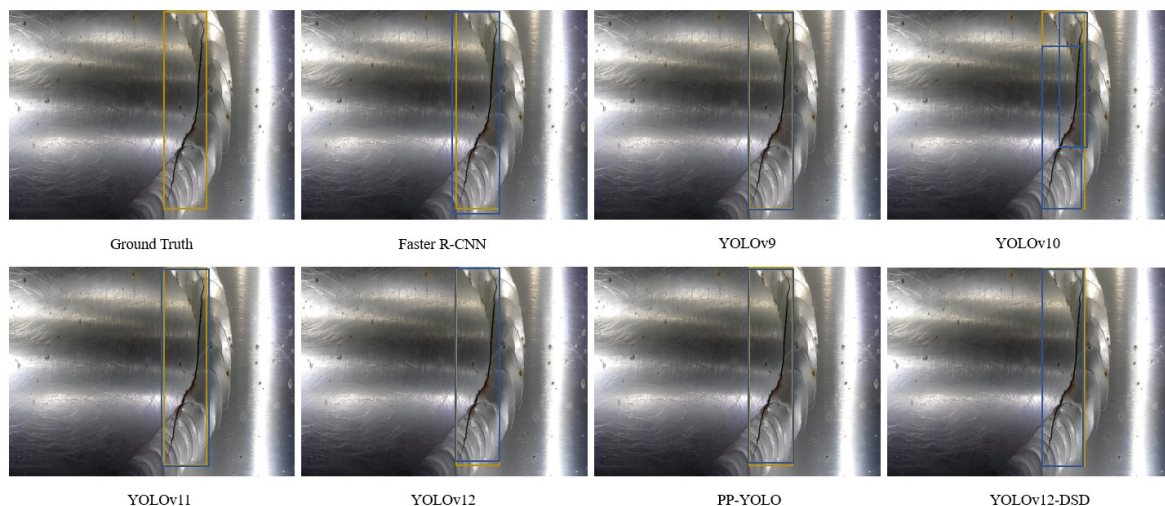


Fig. 9 Visualization results of crack defect detection

图9 裂纹缺陷检测可视化结果

在结构设计、损失函数构建与特征表达能力上的多项改进均具备实际有效性。

4 结语

针对焊缝缺陷形态复杂、尺度差异显著和复杂背景干扰等问题,本文提出了一种基于YOLOv12的改进焊缝表面缺陷检测模型YOLOv12-DSD。该模型通过构建C3k2_D动态特征提取模块、设计SMPDIoU边界框回归损失函数以及引入DyHead动态检测头,分别从特征提取、目标定位和特征表达3个方面对YOLOv12进行了改进,以提升模型对焊缝缺陷目标的检测能力。

基于构建的焊缝缺陷图像数据集开展消融实验和对比实验,对所提出方法的有效性进行了验证。实验结果表明,各改进模块均能够在不同程度上提升模型检测性能。最终YOLOv12-DSD模型参数量为2.13 M,精确率达到92.41%,召回率达到74.80%,mAP@0.5达到87.66%,整体

性能优于对比模型。实验结果表明,所提出方法在保持模型轻量化特性的同时,提高了焊缝缺陷检测准确性和鲁棒性。

研究表明,YOLOv12-DSD能够较好地适应焊缝缺陷检测任务,在检测精度与计算效率之间实现了有效平衡,具备一定的工程应用潜力,可为工业场景下焊缝缺陷智能检测系统开发与部署提供参考。未来,将进一步围绕复杂工况下的模型泛化能力开展研究,重点探索多模态信息融合、小样本缺陷学习和跨场景迁移检测等方法,进一步提升模型在实际工业环境中的适用性和稳定性。

参考文献:

- [1] AGIWAL H, ANSARI MALI, FRANKE D, et al. Material flow visualization during friction stir welding using high-speed X-ray imaging[J]. Manufacturing Letters, 2022, 34: 62-66.
- [2] SCHRICKER K, HELLWIG P, SCHMIDT L, et al. Insight into keyhole and melt pool dynamics in laser welding of zinc-coated steels by means of high-speed synchrotron X-ray imaging[J]. Procedia CIRP, 2024, 124:

- 454-459.
- [3] GEORGE G S, MISHRA P, SINHA P, et al. COVID-19 detection on chest X-ray images using homomorphic transformation and VGG inspired deep convolutional neural network[J]. Biocybernetics and Biomedical Engineering, 2023, 43(1): 1-16.
- [4] UPARKAR O, BHARTI J, PATERIYA R K, et al. Vision transformer outperforms deep convolutional neural network-based model in classifying X-ray images[J]. Procedia Computer Science, 2023, 218: 2338-2349.
- [5] CHEN C Z, LI N, WANG C, et al. Intelligent detection method for bending pipe erosion damage based on multi-scale one-dimensional convolutional neural network[J]. Science Technology and Engineering, 2024, 24(5): 1893-1899.
- 陈传智, 李宁, 王畅, 等. 基于多尺度一维卷积神经网络的弯管冲蚀损伤智能检测方法[J]. 科学技术与工程, 2024, 24(5): 1893-1899.
- [6] WU F, WANG W, LIU W P. X-ray weld defect detection combined with attention mechanism and convolutional neural network[J]. Science Technology and Engineering, 2023, 23(8): 3387-3395.
- 吴昉, 王伟, 刘卫朋. 结合注意力机制和卷积神经网络的X射线焊缝缺陷检测[J]. 科学技术与工程, 2023, 23(8): 3387-3395.
- [7] AJMI C, ZAPATA J, MARTÍNEZ-ALVAREZ J J, et al. Using deep learning for defect classification on a small weld X-ray image dataset[J]. Journal of Nondestructive Evaluation, 2020, 39(3): 68-75.
- [8] MIAO Y X, SUN Z Y, YANG Y, et al. Algorithm of X-ray film digitization and defect detection based on depth learning[J]. Aeronautical Manufacturing Technology, 2023, 66(7): 50-56.
- 缪寅宵, 孙增玉, 杨奕, 等. 基于深度学习的X射线胶片数字化与缺陷检测算法[J]. 航空制造技术, 2023, 66(7): 50-56.
- [9] 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.
- [10] GHEORGHE C, DUGULEANA M, BOBOC R G, et al. Analyzing real-time object detection with YOLO algorithm in automotive applications: a review[J]. Computer Modeling in Engineering & Sciences, 2024, 141(3): 1939-1981.
- [11] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]//Proceedings of the 16th European Conference on Computer Vision-ECCV 2020, 2020: 213-229.
- [12] SONWANE S, CHIDDARWAR S. Enhancing weld defect detection and classification with MDCBNet: a multiscale dense cross block network for improved explainability[J]. NDT & E International, 2024, 142: 103029.
- [13] NING P, JIN J, XU Y P, et al. Enhanced detection of glass insulator defects using improved generative modeling and faster RCNN[J]. Procedia CIRP, 2024, 129: 31-36.
- [14] ZHAO W J, LI D, XU F H. A lightweight weld defect recognition algorithm based on convolutional neural networks[J]. Pattern Analysis and Applications, 2024, 27(3): 94.
- [15] GE Q F, YUAN H, ZHANG Q L, et al. Hyper-Progressive Real-Time Detection Transformer(HPRTDETR) algorithm for defect detection on metal bipolar plates[J]. International Journal of Hydrogen Energy, 2024, 74: 49-55.
- [16] WEI W M, YAN H, ZHENG J C, et al. YOLOv11-based multi-task learning for enhanced bone fracture detection and classification in X-ray images[J]. Journal of Radiation Research and Applied Sciences, 2025, 18(1): 101309.
- [17] WANG R, GAO S Z, LIU W P, et al. A lightweight and efficient defect detection method for X-ray weld images[J]. Transactions of the China Welding Institution, 2024, 45(7): 41-49.
- 王睿, 高少泽, 刘卫朋, 等. 一种轻量级高效X射线焊缝图像缺陷检测方法[J]. 焊接学报, 2024, 45(7): 41-49.
- [18] DAI Z, LIU X J, PAN Q. Weld X-ray image defect recognition algorithm based on CCBFE-RCNN model[J]. Transactions of the China Welding Institution, 2025, 46(1): 24-33.
- 戴铮, 刘晓佳, 潘泉. 基于CCBFE-RCNN模型的焊缝X射线图像缺陷识别算法[J]. 焊接学报, 2025, 46(1): 24-33.
- [19] CUI K B, JIAO J Y. Steel surface defect detection algorithm based on MCB-FAH-YOLOv8[J]. Journal of Graphics, 2024, 45(1): 112-125.
- 崔克彬, 焦静颐. 基于MCB-FAH-YOLOv8的钢材表面缺陷检测算法[J]. 图学学报, 2024, 45(1): 112-125.
- [20] HU L L, DENG C. PCB defect detection algorithm based on YOLO-MCG[J]. Instrumentation Technology and Sensor, 2024(4): 100-106.
- 胡兰兰, 邓超. 基于YOLO-MCG的PCB缺陷检测算法[J]. 仪表技术与传感器, 2024(4): 100-106.
- [21] TIAN Y J, YE Q X, DAVID D. YOLOv12: attentioncentric real-time object detectors[DB/OL]. <https://arxiv.org/abs/2502.12524>.
- [22] HAN K, WANG Y, GUO J, et al. ParameterNet: parameters are all you need for large-scale visual pretraining of mobile networks[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2024: 15751-15761.
- [23] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [24] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 12993-13000.
- [25] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [26] GEVORGYAN Z. Siou loss: more powerful learning for bounding box regression[DB/OL]. <https://arxiv.org/abs/2205.12740>.
- [27] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: unifying object detection heads with attentions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 7373-7382.

(责任编辑:孙娟)