



改进YOLOv8的轻量化钢材表面缺陷检测

张冀, 王定邦, 曹锦纲, 杨立然

引用本文:

张冀, 王定邦, 曹锦纲, 等. 改进YOLOv8的轻量化钢材表面缺陷检测[J]. *智能系统学报*, 2026, 21(2): 375-388.

ZHANG Ji, WANG Dingbang, CAO Jingang, et al. Improvement of YOLOv8 for lightweight steel surface defect detection[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(2): 375-388.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202504018>

您可能感兴趣的其他文章

融合视觉显著性再检测的孪生网络无人机目标跟踪算法

Siamese network combined with visual saliency re-detection for UAV object tracking
智能系统学报. 2021, 16(3): 584-594 <https://dx.doi.org/10.11992/tis.202101035>

基于Faster R-CNN的多任务增强裂缝图像检测方法

Multi-task enhanced dam crack image detection based on Faster R-CNN
智能系统学报. 2021, 16(2): 286-293 <https://dx.doi.org/10.11992/tis.201910004>

融合共现推理的Faster R-CNN输电线路金具检测

Integrating co-occurrence reasoning for Faster R-CNN transmission line fitting detection
智能系统学报. 2021, 16(2): 237-246 <https://dx.doi.org/10.11992/tis.202012023>

隔级融合特征金字塔与CornerNet相结合的小目标检测

Small target detection based on a combination of feature pyramid and CornerNet
智能系统学报. 2021, 16(1): 108-116 <https://dx.doi.org/10.11992/tis.202004033>

多视角数据融合的特征平衡YOLOv3行人检测研究

Research on multi-view data fusion and balanced YOLOv3 for pedestrian detection
智能系统学报. 2021, 16(1): 57-65 <https://dx.doi.org/10.11992/tis.202010003>

基于反卷积和特征融合的SSD小目标检测算法

SSD small target detection algorithm based on deconvolution and feature fusion
智能系统学报. 2020, 15(2): 310-316 <https://dx.doi.org/10.11992/tis.201905035>

DOI: 10.11992/tis.202504018

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250924.1242.002>

改进 YOLOv8 的轻量化钢材表面缺陷检测

张冀^{1,2}, 王定邦¹, 曹锦纲^{1,2}, 杨立然^{1,2}

(1. 华北电力大学 计算机系, 河北 保定 071003; 2. 华北电力大学 复杂能源系统智能计算教育部工程研究中心, 河北 保定 071003)

摘要: 针对钢板表面缺陷检测目前存在的严重漏检误检以及边缘设备难部署问题, 提出了一种基于 YOLOv8 的轻量化的钢材表面缺陷检测算法。1) 设计了部分卷积门控线性单元 (partial convolutional gated linear unit, PGBN) 模块来替换 BottleNeck 模块, 只对部分通道进行卷积操作以减少模型的参数; 2) 使用不同膨胀率的卷积改进空间金字塔池化模块, 加强对细粒度特征的提取, 并结合全局到局部空间聚合 (global-to-local spatial aggregation, GLSA) 模块改进 BiFPN (bidirectional feature pyramid network) 结构, 提出了多尺度特征融合网络 GLFPN (global-to-local spatial aggregation bidirectional feature pyramid network), 以保留小目标特征, 提升模型精度; 3) 设计了轻量化的检测头, 使用共享权重的卷积进一步减少模型的参数量和计算量; 4) 用 WIoU (weighted intersection over union) 损失函数来替代原有的损失函数。实验结果表明, 改进模型在 NEU-DET 数据集上 mAP50 达到了 79.6%, 相比 YOLOv8n 提升了 4.2 个百分点, 而模型的参数量和计算量仅有 1.43×10^6 和 4.7×10^9 , 较 YOLOv8n 分别下降了 53.3% 和 41.9%, 在提升准确率的同时更加适合在边缘设备部署。

关键词: 轻量级网络; 缺陷检测; YOLOv8n; 多尺度特征融合; 特征提取; 空间金字塔; 目标检测; WIoU**中图分类号:** TP391.4; TG115 **文献标志码:** A **文章编号:** 1673-4785(2026)02-0375-14

中文引用格式: 张冀, 王定邦, 曹锦纲, 等. 改进 YOLOv8 的轻量化钢材表面缺陷检测 [J]. 智能系统学报, 2026, 21(2): 375-388.

英文引用格式: ZHANG Ji, WANG Dingbang, CAO Jingang, et al. Improvement of YOLOv8 for lightweight steel surface defect detection [J]. CAAI transactions on intelligent systems, 2026, 21(2): 375-388.

Improvement of YOLOv8 for lightweight steel surface defect detection

ZHANG Ji^{1,2}, WANG Dingbang¹, CAO Jingang^{1,2}, YANG Liran^{1,2}

(1. Department of Computer Science, North China Electric Power University, Baoding 071003, China; 2. Engineering Research Center of intelligent Computing for Complex Energy Systems Ministry of Education, North China Electric Power University, Baoding 071003, China)

Abstract: Surface defects of steel plates still face significant issues of missed and false detections, as well as difficulties in deployment on edge devices. To address these problems, a lightweight steel surface defect detection algorithm based on YOLOv8 is proposed. First, a Partial Convolution Gated Linear Unit (PGBN) module is designed to replace the BottleNeck module, performing convolution operations only on part of the channels to reduce the model's parameters. Next, the Spatial Pyramid Pooling module is improved by using convolutions with different dilation rates to enhance the extraction of fine-grained features. Then, GLFPN is proposed by combining the Global-to-Local Spatial Aggregation (GLSA) module with the improved BiFPN structure. Its weighted fusion can more effectively retain small sample features, thereby improving the model's accuracy. After that, a more lightweight detection head is designed to further reduce the model's parameters and computational load. Finally, the WIoU loss function is used to replace the original Clou. Experimental results show that the improved model achieves an mAP50 of 79.6% on the NEU-DET dataset, a 4.2 percent improvement over YOLOv8n, while the model parameters and computational load are only 1.43×10^6 and 4.7×10^9 , respectively, representing a 53.3% and 41.9% reduction compared to YOLOv8n. This makes the model more accurate while being more suitable for deployment on edge devices.

Keywords: lightweight network; defect detection; YOLOv8n; multiscale feature fusion; feature extraction; spatial pyramid; object detection; WIoU

收稿日期: 2025-04-23. 网络出版日期: 2025-09-24.

基金项目: 河北省自然科学基金青年科学基金项目 (A 类) (F2024502002); 中央高校基本科研业务费专项资金面上项目 (2024MS127).

通信作者: 曹锦纲. E-mail: caojg168@126.com.

钢材作为工业生产的重要基础材料, 表面质量直接影响到最终产品的性能和使用寿命。由于钢材产品广泛应用于建筑、汽车和航空航天等领域, 表面缺陷 (如划痕、裂纹和夹杂物等) 检测对

于保证产品质量至关重要。因此, 钢材表面缺陷检测技术的研究与应用具有重要的现实意义。

传统的钢材表面缺陷检测方法主要依赖于人工检测、红外检测法和漏磁检测法^[1]等, 这些方法不仅效率低下, 而且检测精度有限, 还存在主观性强以及对复杂表面模式适应性差等问题, 难以满足现代工业生产对高质量、高效率和高稳定性的要求。因此, 如何在保证高检测精度的同时, 实现模型的轻量化, 仍然是一个亟待解决的问题。深度学习强大的特征提取和模式识别能力, 在视觉检测任务中表现出了卓越的性能。现阶段的目标检测算法主要分为两类。

一种是以 R-CNN(region-based convolutional neural networks)^[2]、Fast-RCNN^[3] 和 Faster-RCNN^[4] 为代表的二阶段目标检测算法, 该类算法先利用候选区域生成提取可能存在目标的区域, 然后在这些候选区域上进行特征提取和分类, 然而, 该类方法计算复杂度较高, 难以满足实时检测的需求, 因此在工业生产线上的应用受到限制。近年来, 为了提升两阶段检测算法的效率, 研究人员提出了多种改进方案。例如, Cascade R-CNN^[5] 通过多级级联网络提高了目标检测的精度, Libra R-CNN^[6] 通过平衡正负样本的损失函数优化目标检测效果, 但这类方法仍然存在计算复杂度高、推理速度慢的问题, 不适用于实时性要求较高的工业场景。陈露远等^[7] 提出一种基于融合 Transformer 和 Faster-RCNN 的多模态特征提取与融合模型更好地进行两种模态特征的提取。Sparse R-CNN^[8] 稀疏化候选区域生成过程, 采用了一种动态实例交互头, 与传统的固定参数头部相比, 它以独特的提案特征为条件, 更加灵活, 并在准确性方面具有显著领先优势。邓慧等^[9] 利用精准的兴趣区域池化和均值偏移聚类算法获得更佳的锚框尺寸, 能很好地满足工业检测的需求。

另一种是以 SSD(single shot multibox detector)^[10]、YOLO(you only look once)^[11-13] 为代表的直接进行位置回归的单阶段检测方法。SSD 通过不同尺度的特征图进行多尺度目标检测, 提高了检测精度, 但对小目标的检测效果有限。YOLO 系列自提出以来, 不断优化网络结构和检测机制, 检测精度和速度均得到了显著提升。YOLOv5 采用 C3(cross stage partial network with 3 convolutions) 结构优化了特征提取能力, YOLOv7 通过 E-ELAN(extended efficient layer aggregation networks) 结构增强了计算效率, 而 YOLOv8 优化动态标签分配策略, 采用无锚框(anchor-free) 在训练上更加

高效, 使得检测性能得到进一步提升。李刚等^[14] 提出了一种轻量级的缺陷检测网络, 主干网络中嵌入了通道注意力层和坐标注意力层, 用于获取丰富的表面缺陷特征信息, 此外, 多尺度特征聚合网络融合了深层语义和浅层语义特征信息, 进一步提升了模型对复杂缺陷的检测能力。GS-YOLO(gather-and-distribute-squeeze-YOLO)^[15] 通过引入动态卷积模块, 显著扩大了网络的感受野, 提升了对多尺度缺陷特征的提取能力, 同时改进损失函数更有效地优化检测框的定位精度, 有效地提升了缺陷检测的准确性和效率。敖思铭等^[16] 通过 K-means 选取更合适的预设锚框, 提高预测框与真实目标的匹配度, 提高模型检测精度的同时加速了模型的收敛。侯玥等^[17] 设计了一种由 Swin Transformer 模块和 PANet(path aggregation network) 模块组成的学习器, 对小样本检测任务有了显著的提高。张上等^[18] 设计了基于 YOLOv8 的轻量化锻件缺陷检测算法, 提出轻量化卷积模块, 以增强特征交互并降低计算量, 利用 LAMP(layer-adaptive magnitude-based pruning) 剪枝策略去除不重要的权重参数, 减少模型体积并提高检测速度。然而, 钢材表面缺陷检测由于缺陷尺寸较小、形状不规则和对比度低等特点, 仍然面临诸多挑战, 且在计算资源受限的工业应用场景中, 轻量化模型更具实际价值。为了解决这些挑战, 本文提出了一种改进轻量化钢材表面缺陷检测方法 PFGD-YOLOv8(PGC2f+FPSC+GLFPN+DSHead), 旨在在降低模型尺寸的同时, 实现高检测精度, 便于实际工业应用。本文主要创新如下:

1) 在主干网络部分, 设计了部分卷积门控线性单元(partial convolutional gated linear unit, PGC2f) 模块, 使用部分卷积捕捉局部特征, 并通过门控机制加权这些特征, 在降低模型参数数量的同时, 提升模型的检测精度。

2) 对空间金字塔池化快速(spatial pyramid pooling fast, SPPF) 模块进行改进提出一种新的特征金字塔共享卷积 FPSC(feature pyramid shared convolution) 模块来增强模型对细小缺陷的特征提取能力, 通过采用不同膨胀率的卷积操作, 在不增加参数数量的情况下, 捕捉更大范围的上下文信息, 有效提取不同尺度的特征。

3) 在颈部网络部分, 用全局到局部空间聚合(global-to-local spatial aggregation, GLSA) 模块取代传统卷积, 并结合 BiFPN(bidirectional feature pyramid network) 设计了一种全局到局部空间聚合双向特征金字塔网络(GLSA-BiFPN, GLFPN), 通过

通道分割, 分别提取全局和局部信息, 再进行融合, 能够更好地保留小目标特征, 提升了模型对复杂缺陷的检测能力。

4) 在头部网络部分, 设计了一个轻量化细节增强共享检测头 (detail-enhanced shared convolution detection head, DSHead), 使用卷积共享权重策略, 并结合归一化策略和细节增强卷积以防止精度下降, 有效地降低了模型的参数量, 最后, 用 WIoU(weighted intersection over union) 损失函数替换原有的损失函数进一步提升模型性能。

1 PFGD-YOLOv8

1.1 YOLOv8

YOLOv8 作为近些年被广泛使用的单阶段目标检测算法, 在检测、分割和分类等任务中都表现出优异的速度和准确率, 其网络结构主要有输入、主干网络、特征融合网络、检测头以及输出端 5 部分构成。主干网络负责从输入图像中提取

特征信息, 通常基于 CSPDarknet(cross stage partial darknet), 通过跨阶段部分连接 (cross stage partial connections, CSP) 来减少计算量并提升特征提取能力。YOLOv8 的特征融合网络通常采用 PANet 模块, PANet 通过自底向上和自顶向下的路径聚合, 将不同层次的特征进行融合, 从而增强模型对不同尺度目标的检测能力。为提升模型的检测能力, YOLOv8 默认使用无锚框检测方式, 相比传统的基于锚框的方法, 减少了超参数的调整, 提高了检测速度和精度。

1.2 PFGD-YOLOv8

钢材表面缺陷检测任务具有其特殊性: 缺陷种类多样, 形态复杂尺度变化大, 且工业场景中计算资源有限, 这对算法的轻量化和鲁棒性提出了更高的要求, 因此提出轻量化缺陷检测模型 PFGD-YOLOv8, 模型结构如图 1 所示, 主要包括主干 (Backbone)、颈部 (Neck) 和头部 (Head) 3 部分。

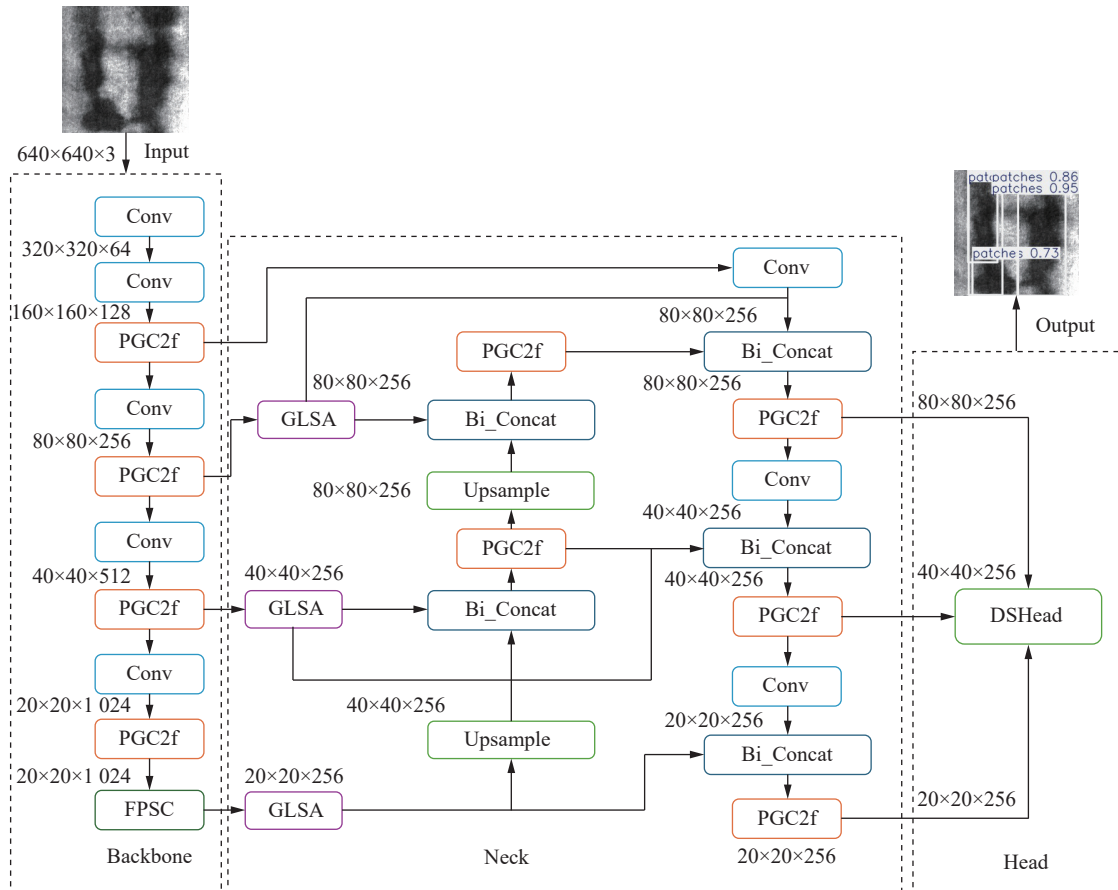


图 1 PFGD-YOLOv8 模型结构

Fig. 1 Structure of PFGD-YOLOv8 model

主干网络包含普通卷积 (Conv)、PGC2f 和 FPSC 模块。Conv 调整通道数与分辨率提取低层边缘、纹理等特征; PGC2f 结合部分卷积, 提高模型对细节的捕捉能力的同时降低模块复杂度;

FPSC 模块通过引入多尺度卷积和残差连接进一步特征选择和融合, 提高模型的检测精度。颈部网络包含 GLSA、PGC2f 和 Bi-Concat 模块, GLSA 模块结合全局与局部注意力的结果, 调整通道数

的同时进一步突出关键区域;抑制背景噪声;Bi_Concat 模块通过加权融合增强细小特征的权重,进一步增强 PFGD-YOLOv8 模型的小目标特征捕捉。在头部使用轻量级细节增强共享卷积检测头 (DSHead)。最后,使用 WIoU 替换原本的 CIoU (complete intersection over union) 作为损失函数。

1.2.1 PGC2f 特征提取模块

在 YOLOv8 特征提取网络中, C2f 模块由 Conv 和多个 Bottleneck 组成,其在特定数据集 (NEU-DET) 特征提取能力效果欠佳,模型训练效果出色,但是测试效果不理想。模块中的标准卷积操作计算复杂度较高,且需要频繁进行内存访问,为了满足模型轻量化的同时提高特征提取能力,受 PConv 启发设计了 PGBN 模块,如图 2 所示。PGBN 的公式为

$$PGBN(x) = x + \text{Drop}(\text{CGLU}(\text{PConv}(x)))$$

式中: x 表示输入特征, $\text{Drop}(\cdot)$ 表示 DropPath 操作, $\text{CGLU}(\cdot)$ 和 $\text{PConv}(\cdot)$ 分别表示经 CGLU(convolutional gated linear units) 模块和 PConv 模块处理。

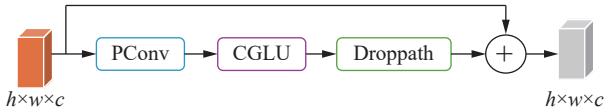


图 2 PGBN 结构

Fig. 2 Structure of PGBN

PGBN 由 3 部分构成,第 1 部分为 PConv^[19] 特征提取模块,只对部分输入通道进行空间特征提取,其余通道数保持不变,当 Pconv 仅使用 1/4 的通道数时,局部卷积的计算量仅为标准卷积的 1/16,能够显著地减少模型的计算复杂度。

PGBN 的第 2 部分为通道混合器 CGLU^[20],通过门控机制动态地控制信息流,使得网络能够自适应地强调重要特征并抑制不重要的特征,增强了局部建模能力和模型的鲁棒性, CGLU 结构如图 3 所示。

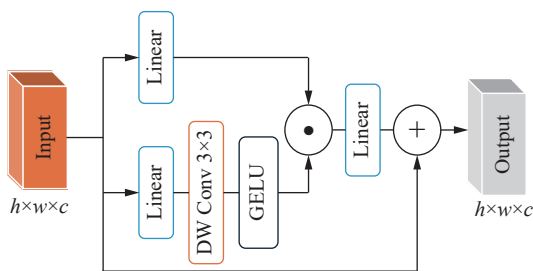


图 3 CGLU 结构

Fig. 3 Structure of CGLU

CGLU 在传统的 GLU(gated linear unit) 门控分支^[21]的激活函数前融入一个简化的 3×3 深度卷积,用来捕捉局部特征并提供位置信息,使其结

构符合门控通道注意的设计理念,每个通道的门控信号能够根据其邻域特征进行调整,增强了模型对局部模式的敏感性,经过激活函数后与其他分支的输出进行逐位相乘,得到最终结果。CGLU 操作表示为

$$\text{CGLU}(x) = x + \text{Lin}(\text{Lin}(x) \odot \sigma(\text{Lin}(\text{DWConv}(x))))$$

式中: $\text{Lin}(\cdot)$ 表示线性变换操作, $\sigma(\cdot)$ 表示 GELU(Gaussian error linear units) 激活函数, $\text{DWConv}(\cdot)$ 表示经过深度可分离卷积 (depthwise convolution, DWConv) 模块处理。

第 3 部分 Droppath 模块在训练过程中随机丢弃整个网络层的输出防止模型过拟合,提高模型的泛化能力。

由 PGBN 模块替换 C2f 中的 Bottleneck 构建成了 PGC2f 模块,如图 4 所示, C 表示拼接操作 Concat。

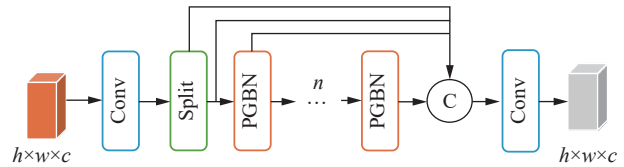


图 4 PGC2f 结构

Fig. 4 Structure of PGC2f

用 PGC2f 替换原本 YOLOv8 的 C2f 模块,通过在前向传播过程中添加 CGLU 模块,结合局部卷积实现更轻的网络架构和更强的特征提取能力。

1.2.2 FPSC

YOLOv8 中的 SPPF 采用多个固定尺度的池化操作 (如 5×5、9×9、13×13) 来提取多尺度特征,虽然可以很好地提取输入特征,但小目标的特征可能在较大的池化窗口下被平均化,导致信息损失,对小目标检测效果较差,影响检测精度。在钢材表面缺陷检测任务中, SPPF 对细粒度缺陷检测 (如划痕、裂纹等) 的目标边界的描述不够精确,影响目标框的回归效果。因此,设计了一种更加细粒度的特征提取模块 FPSC,网络结构如图 5 所示。

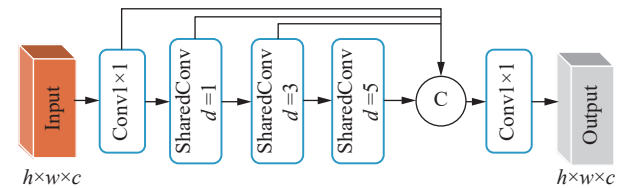


图 5 FPSC 结构

Fig. 5 Structure of FPSC

FPSC 先通过 1×1 的卷积调整通道数,为了防止参数量过大,使用共享卷积层 SharedConv 提取不同尺度特征,其中权重是共享的,不同尺度的

特征提取使用相同的卷积核, 通过调整膨胀因子 d 来改变感受野, 从而捕获更大范围的上下文信息, 最后将所有卷积层的输出在通道维度上进行拼接, 形成一个包含多尺度信息的特征图, 通过共享卷积权重, 该模块在增加感受野的同时保持了计算效率。FPSC 的操作可表示为

$$\begin{aligned} \mathbf{x}_0 &= \text{SharedConv}_{d=1}(\text{Conv}(\mathbf{x})) \\ \mathbf{x}_1 &= \text{SharedConv}_{d=3}(\mathbf{x}_0) \\ \mathbf{x}_2 &= \text{SharedConv}_{d=5}(\mathbf{x}_1) \end{aligned}$$

$$\text{FPSC}(\mathbf{x}) = (\text{Conv}(\text{Concat}(\text{Conv}(\mathbf{x}), \mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2)))$$

式中: \mathbf{x} 为输入特征图; $\text{SharedConv}_{d=i}$, $i=\{1,3,5\}$ 分别表示膨胀因子为 1、3 和 5 的共享卷积; 而 \mathbf{x}_0 、 \mathbf{x}_1 和 \mathbf{x}_2 为对应共享卷积的输出。

1.2.3 GLFPN

在钢板缺陷中, 小目标缺陷较为常见, 可能会导致模型出现误检和漏检现象, 而小目标特征在底层特征图中通常更加明显。YOLOv8 中的特征融合网络通过自上而下的 FPN(feature pyramid network) 块和自下而上的 PANet 块实现了双向特征融合, 增强了底层特征的语义信息, 但其简单的逐元素相加或拼接的特征融合方式可能导致底层细节信息的丢失。

BiFPN^[22] 通过可学习的权重动态调整不同层级特征的贡献, 底层特征在融合过程中不会被简单覆盖, 而是通过权重分配得到增强, 能够更好地保留小目标特征。传统的 FPN 和 PANet 中存

在一些冗余节点(例如只有单一输入边的节点), 这些节点对特征融合的贡献有限, 但增加了计算复杂度。BiFPN 删除了这些冗余节点, 简化了网络结构, 同时提高了计算效率。BiFPN 允许堆叠多个模块, 通过重复使用双向路径和加权融合机制, 逐步优化特征表示, 从而提升模型性能。其特征融合公式为

$$\mathbf{O} = \sum_i \frac{W_i}{\varepsilon + \sum_j W_j} \cdot \mathbf{F}_i$$

式中: \mathbf{O} 是融合后的输出特征图, \mathbf{F}_i 是第 i 个输入特征图, W_i 是对应于第 i 个输入特征图的可学习权重, $\varepsilon=10^{-4}$ 用于避免数值不稳定。以 P4 层为例, 其特征融合表达式为

$$\mathbf{p}_4^{\text{out}} = \text{Conv} \left(\frac{\omega'_1 \mathbf{p}_4^{\text{m}} + \omega'_2 \mathbf{p}_4^{\text{d}} + \omega'_3 \cdot \text{Resize}(\mathbf{p}_3^{\text{out}})}{\omega'_1 + \omega'_2 + \omega'_3 + \varepsilon} \right)$$

式中: $\mathbf{p}_4^{\text{out}}$ 为自下而上第 4 层的输出结果, \mathbf{p}_4^{d} 为中间结果, $\text{Resize}(\cdot)$ 为上采样或下采样操作, $\text{Conv}(\cdot)$ 为普通卷积, $\omega'_1 - \omega'_3$ 为学习到的特征权重。

BiFPN 依赖于多层次的双向特征融合机制, 其中包含多次特征的跨层融合和更新, 导致了较高的计算开销。尤其在高分辨率图像处理时, 计算复杂度会显著增加, 针对该问题本文提出了一种改进的全局到局部空间聚合双向特征金字塔网络 GLFPN, PANet、BiFPN 和 GLFPN 的网络结构如图 6 所示。

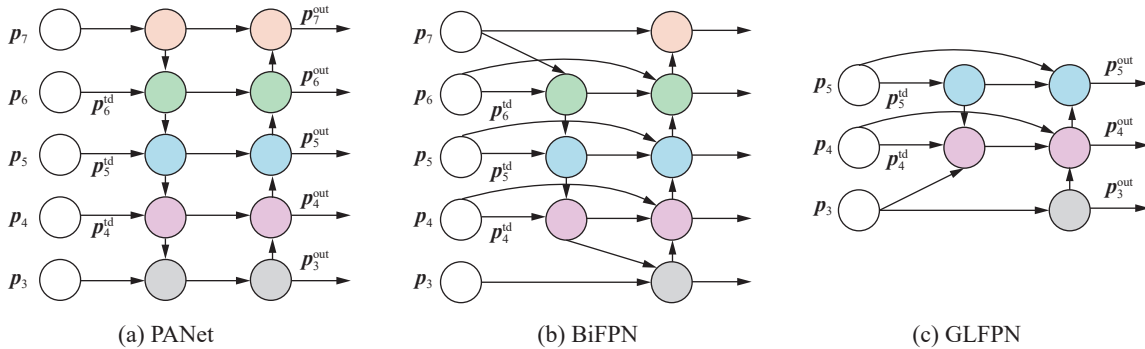


图 6 PANet、BiFPN 和 GLFPN 结构

Fig. 6 Structure of PANet, BiFPN and GLFPN

GLFPN 在保留跨尺度双向特征交互机制的前提下, 对原网络进行层级精简优化, 仅使用 P3、P4、P5 通道进行特征图输出, 降低计算负担的同时避免最高层级(如 P7)带来的噪声干扰。

在颈部部分引入 GLSA(global-to-local spatial aggregation) 模块^[23], GLSA 网络结构如图 7 所示, 输入特征图 F 通过分离通道映射为两组特征图 F^1 和 F^2 , 分别输入到全局空间注意力(global spatial attention, GSA) 模块和局部空间注意力(local

spatial attention, LSA) 模块中。GSA 强调每个像素在空间中的长距离关系, 补充局部空间注意力, 先通过 1×1 卷积层和转置操作, 生成全局空间注意力图, 然后用 Softmax 函数对注意力图进行归一化, 得到的结果与原始输入特征图进行矩阵乘法, 增强全局特征。通过多层感知机(multilayer perceptron, MLP) 第一层将输入映射到扩展比为 2 的高维空间, 经过 ReLU(rectified linear units) 和归一化层处理, 第二层恢复到和输入相同的维

度。GSA 公式为

$$\text{AttG}(F^1) = \text{Softmax}(\text{Transpose}(\text{Conv}(F^1)))$$

$$\text{MLP}(x) = (\text{Conv}(\text{LR}(\text{Conv}(x))))$$

$$\text{GSA}(F^1) = \text{MLP}(\text{AttG}(F^1) \otimes F^1) + F^1$$

式中: AttG(·)表示全局注意力操作, Transpose(·)和

Softmax(·)分别表示转置操作和 Softmax 函数, LR(·)表示经过层归一化和 ReLU 非线性激活函数处理, MLP(·)表示经过 MLP 模块处理, Conv(·)表示 1×1 卷积, ⊗表示矩阵乘法, +表示残差链接操作, GSA(·)表示经过 GSA 模块处理。

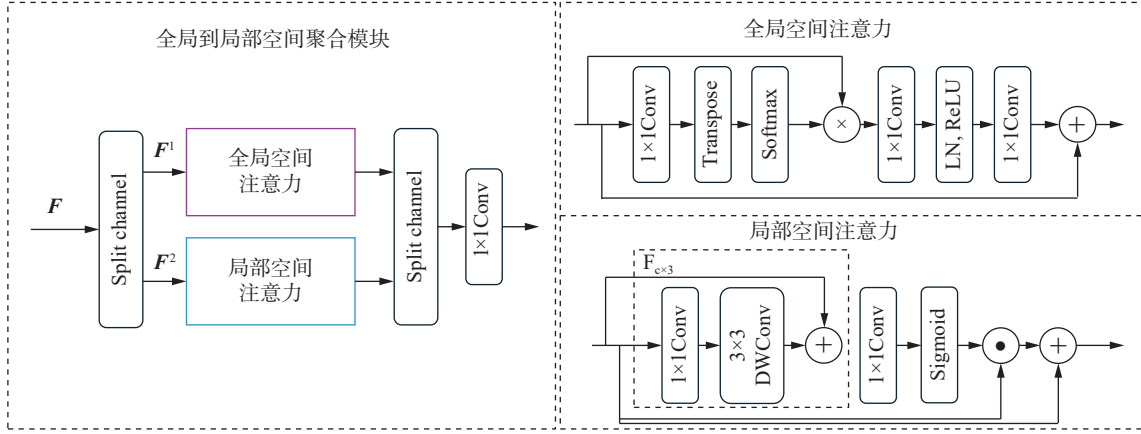


图 7 GLSA 结构

Fig. 7 Structure of GLSA

LSA 通过深度可分离卷积和卷积层提取局部特征, 使用 Sigmoid 函数生成局部空间注意力图, 得到的结果与原始输入特征图先进行逐元素相乘再相加, 公式为

$$F_c(x) = (\text{Conv}(\text{DWConv}(x)) + x)$$

$$\text{AttL}(F^2) = \sigma(\text{Conv}(F_{c \times 3}(F^2))) + F^2$$

$$\text{LSA}(F^2) = \text{AttL}(F^2) \odot F^2 + F^2$$

式中: AttL(·)表示局部注意力操作, σ(·)表示 Sig-

moid 函数, $F_{c \times 3}(\cdot)$ 表示级联的 3 个 F_c 模块, \odot 表示逐元素乘法, LSA(·)表示经过 LSA 模块处理, DWConv(·)表示经过 DWConv 模块处理, 这种结构设计可以用更少的参数高效地聚合局部空间信息。

1.2.4 轻量化解耦头

随着深度卷积神经网络深度的增加, 模型在检测时不可避免地会产生占用大量内存和运行速度过慢的问题, 因此设计了一种新的轻量化的细节增强共享检测头 DSHead, 其结构如图 8 所示。

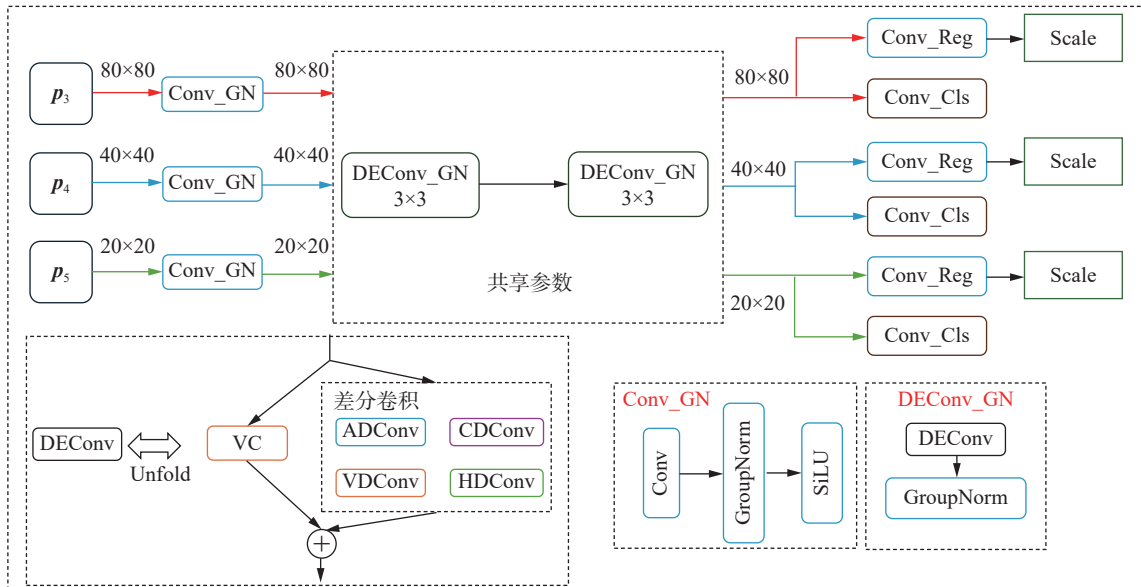


图 8 DSHead 结构

Fig. 8 Structure of DSHead

首先通过基于组归一化的卷积来调整通道数。然后, 使用两组共享参数的基于组归一化的

细节增强卷积 (DEConv_GN) 对跨层空间和语义信息进行聚合。最后, 将提取的信息输入到分类

和回归头中。使用了一个包含可学习缩放因子的缩放层 (Scale) 处理检测头之间目标大小不一致问题并防止小目标特征信息的丢失, 对回归头中的特征进行调整。用 GN(group normalization) 层替换 Conv 中的 BN(batch normalization) 层。BN 层一般用于加速训练并提高模型性能, 然而, BN 层的归一化计算依赖于批次大小, 当批次较小时, 统计均值和方差的不稳定性会导致模型性能下降。GN 层通过将特征通道分组, 并在每个组内独立计算均值和方差来进行归一化操作。这种设计使得 GN 层的计算与批次大小无关, 从而在不同批次大小下都能保持稳定的精度, 尤其在小批次训练中表现优异^[24]。

DEConv 包括原始卷积 (vanilla convolution, VC) 和差分卷积 (difference convolutions)。差分卷积包括 4 种不同类型的卷积操作: 中心差分卷积 (central difference convolution, CDC)、自适应差分卷积 (adaptive difference convolution, ADC)、垂直差分卷积 (vertical difference convolution, VDC) 和水平差分卷积 (horizontal difference convolution, HDC), 能够从不同方向和角度捕获图像的细微特征。这种多维度的特征提取能力帮助模型能够更全面地理解图像内容, 尤其是边缘、纹理等细节信息。此外在 DEConv 上, 当几个具有相同大小的卷积核输入相同的步幅和填充操作时, 对它们的输出求和产生的结果与对相应位置的核求和产生的最终输出相同, 不会增加额外的计算和参数, 使 DEConv 更加适用于轻量级的模型。DEConv 可以利用重参数化技术与普通卷积以相同的计算成本产生 F_{out} , 公式为

$$F_{out} = DEConv(F_{in}) = \sum_{i=1}^5 F_{in} * K_i = F_{in} * \left(\sum_{i=1}^5 K_i \right) = F_{in} * K_{cvt}$$

式中: K_i 表示 VC、CDC、ADC、VDC 和 HDC 的核, $i \in \{1, 2, 3, 4, 5\}$; * 表示卷积操作; K_{cvt} 表示合并并行卷积得到的转换核。

1.2.5 WIou 损失函数

损失函数是影响模型检测性能的重要原因之一, YOLOv8 在边界框预测中采用了 CIoU 损失函数, 以增强边界框的拟合能力和训练效果。CIoU 结合重叠面积、中心点距离和宽高比, 提供更全面的边界框匹配评估, 但在目标重叠面积小或目标边界模糊时, 梯度趋近于零, 导致训练过程中梯度消失, 影响模型对细节的拟合能力。为解决此问题, 本文将 YOLOv8 的损失函数替换为 WIoU。

WIoU 通过非单调聚焦系数和梯度增益分配, 实现了动态聚焦的边界框回归损失, 对离群度大的锚框分配较小的梯度增益来确保锚框质量和减少有害梯度。WIoU 计算公式为

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty)$$

$$L_{WIoU} = \frac{\beta L_{IoU}}{\delta \alpha^{\beta-\delta}} \exp\left(\frac{(x-x_{gt})^2 + (y-y_{gt})^2}{W_g^2 + H_g^2}\right)$$

式中: β 是离群度, 用来评价锚框质量; L_{IoU}^* 表示单调聚焦系数; L_{IoU} 表示动量 m 的滑动平均值; α 和 δ 均为超参数, 随着 IoU 的变化而动态变化; L_{IoU} 表示 IoU 损失函数的值; (x, y) 表示预测框的中心坐标; (x_{gt}, y_{gt}) 表示真实框的中心坐标; W_g 和 H_g 分别表示预测框和真实框最小外接矩形的长和宽。

2 实验结果分析

2.1 实验环境和数据集

本实验使用的操作系统是 Linux, 搭载 PyTorch 深度学习框架, CUDA 版本为 12.6, GPU 型号为 V100-PCIE-32GB, 编译器版本为 3.10, 模型参数设置见表 1。使用东北大学发布的钢材表面缺陷样本数据集 NEU-DET^[25] 和 GC10-DET^[26] 数据集进行实验验证。NEU-DET 包含 1 800 张图片和 6 种常见的缺陷类别, 分别为: 裂纹 (crazing, Cr)、夹杂 (inclusion, In)、斑块 (patches, Pa)、划痕 (scratches, Sc)、氧化轧皮 (rolled in scale, Rs) 和点蚀面 (pitted surface, Ps), 每个类别各 300 张, 分辨率为 200×200。将原始数据集按照 8:1:1 的比例分为训练集 (1 440 张)、验证集 (180 张) 和测试集 (180 张)。GC10-DET 数据集包括 10 种缺陷, 分别为穿孔 (pu)、熔接痕 (wl)、月牙间隙 (cg)、水斑 (ws)、油斑 (os)、丝斑 (ss)、夹杂物 (in)、轧坑 (rp)、折痕 (cr) 和腰部折叠 (wf), 共 2 294 张图片, 按照 8:1:1 将原始数据集分为训练 (1 835 张) 验证 (229 张) 和测试集 (230 张)。

表 1 模型参数设置

Table 1 Model parameter configuration

参数	取值
训练轮次	300
图片尺寸	640×640
批次大小	32
学习率	0.01
优化器	SGD

2.2 评估指标

本实验模型除了要拥有较低的参数量和计算量, 还要兼顾精度要求。因此, 实验结果采用 6 个

评估指标, 准确率 (precision, P)、召回率 (recall, R)、平均精度均值 (mean average precision, mAP)、参数量 (parameters, Pa)、浮点操作次数 (floating-point operations, FLOPs) 和每秒传输帧数 (frames per second, FPS), 计算公式分别为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}$$

$$I_{AP} = \int_0^1 P(R) dR$$

$$m_{AP} = \frac{1}{N} \sum_{i=1}^N I_{AP}^i$$

式中: P 表示准确率, R 表示召回率, I_{AP} 表示平均精度, m_{AP} 表示平均精度均值; N_{TP} 表示预测为正

类且实际为正类的样本数量, N_{FP} 表示预测为正类但实际为负类的样本数量, N_{FN} 表示预测为负类但实际为正类的样本数量; N 表示检测类别的总数; $P(R)$ 是在召回率为 R 时的精度。mAP50 表示 AP 计算时所采用的 IOU(intersection over union) 阈值为 50% 的平均精度, mAP50-95 表示在 IOU 阈值范围在 50%~95% 以 5% 为步长, 所有步长下平均精度的均值。

2.3 对比试验

为验证本文改进模型的有效性, 在相同硬件资源和相同数据集划分的情况下, 与一阶段检测网络 SSD、YOLO 系列模型, 两阶段网络 Faster-RCNN、Cascade-RCNN, 以及近期其他先进算法^[27-29]进行了详细对比, 结果如表 2 所示。

表 2 NEU-DET 数据集模型对比实验
Table 2 Model comparison experiment on NEU-DET dataset

模型	P/%	R/%	mAP50/%	mAP50-95/%	Pa/10 ⁶	FLOPs/10 ⁹	FPS
Fsater-CNN	69.5	76.8	76.9	38.9	56.0	124.2	43
SSD	72.3	70.5	73.2	35.8	26.3	62.7	46
Cascade-RCNN	75.0	72.1	77.4	43.2	75.3	197.5	139
YOLOv8n	73.6	70.2	75.4	42.6	3.0	8.1	147
YOLOv9-t	70.5	71.2	77.6	43.4	2.6	10.7	132
YOLOv10n	70.5	68.2	76.7	41.8	2.2	6.5	140
YOLOv11n	69.1	73.8	70.0	33.7	2.6	6.4	149
YOLOv12n	70.1	70.3	77.5	43.4	2.5	6.3	152
文献[27]	—	—	76.7	—	2.5	6.6	154
文献[28]	68.9	69.8	76.9	—	3.3	7.8	277
文献[29]	—	—	78.1	—	3.4	11.4	173
本文模型	75.6	71.9	79.6	45.8	1.4	4.7	125

注: 加粗表示最优, “—”表示文献中没有给出该数据。

从实验结果可以看出, 改进模型在 mAP50 和 mAP50-95 上表现均为最佳, 分别达到了 79.6% 和 45.8%, 相较于 YOLOv8n 分别提升了 4.2 个百分点和 3.2 个百分点。相比之下, YOLO 系列的其他版本 (YOLOv9t、YOLOv10n、YOLOv11n 和 YOLOv12n) 的 mAP50 依次为 77.6%、76.7%、70.0% 和 77.5%, 虽然有所提高, 但仍不及本文改进模型。经典的 Faster-CNN 和 Cascade-RCNN 模型的 mAP50 分别达到了 76.9% 和 77.4%, 虽然精度略高于 YOLOv8n, 但参数量和计算量较大, 限制了其在实际应用中的效率。这表明, 改进模型对于目标检测的精度提升具有显著的效果。在模型的轻量化方面, 改进模型在参数量和计算量上具有显著优势。改进模型的参数量为 1.4×10^6 , 计算量为 4.7×10^9 , 与 YOLOv8n 相比, 分别降低了 54% 和 42%, 这表明改进模型的轻量化设计有效

减少了计算资源的需求。传统的 Faster-CNN 和 Cascade-RCNN 在参数量和计算量上远高于 YOLO 系列。文献 [27-29] 的检测速度分别达到了 154、277 和 173, 实时性较高, 但是检测精度较低, 参数量和计算量也较高。综上所述, 本文算法在参数量计算量以及检测精度方面均有出色表现, 符合轻量化应用的需求, 使其在实际应用中具有更好的适用性, 特别是对于资源有限的嵌入式和移动设备部署场景。

为了验证提出的改进模型在不同数据集上的泛化能力, 在 GC10-DET 数据集上进行了泛化实验, 实验结果如表 3 所示。改进模型在 mAP50 的指标上, 相比于 YOLOv8n 的 63.7% 提升了 3 百分点, 达到 66.7%。文献 [28] 检测精度较低, 文献 [29] 虽然检测精度较高但是参数量和计算量也较高。对比实验证明了改进模型在 GC10-DET 数据集上

的目标检测精度有明显提高, 验证了改进模型在不同数据集上具有良好泛化能力, 不仅提高了目标检测精度, 而且在轻量化设计方面也表现突出, 具备了更强的实用性和推广性。

2.4 消融实验

2.4.1 各改进模块性能分析

为验证各改进模块的有效性, 以 YOLOv8n 为基准模型, 在 NEU-DET 数据集上进行了 11 组消融实验, 每组消融实验的环境配置和参数设置均保持一致, 结果如表 4 所示。

表 3 在 GC10-DET 上对比实验

Table 3 Experimental comparison on GC10-DET

模型	P/%	R/%	mAP50/%	Pa/ 10^6	FLOPs/ 10^9
YOLOv9	61.1	65.1	65.5	2.6	10.7
YOLOv10	51.7	63.9	60.1	2.2	6.5
YOLOv8	72.5	57.4	63.7	3.0	8.1
文献[28]	70.4	50.9	58.2	3.3	7.8
文献[29]	—	—	66.3	3.1	11.4
本文模型	64.6	64.8	66.7	1.4	4.7

注: 加粗表示最优, “—”表示文献中没有给出该数据。

表 4 模型消融实验

Table 4 Ablation experiment of model

模型	PGC2f	FPSC	GLFPN	DSHead	mAP50/%	mAP50-95/%	Pa/ 10^6	FLOPs/ 10^9	FPS
M1					75.4	42.6	3.0	8.1	147
M2	√				77.5	44.4	2.3	6.2	151
M3		√			76.0	42.8	3.1	8.1	161
M4			√		77.0	43.7	2.1	7.6	133
M5				√	77.2	44.2	2.3	6.5	167
M6	√	√			78.6	45.2	2.3	6.1	164
M7	√	√	√		78.9	45.3	2.3	5.9	157
M8	√			√	76.6	43.5	1.5	4.8	159
M9	√		√		77.2	44.0	1.6	5.9	131
M10			√	√	78.4	45.1	1.7	6.4	158
M11	√	√	√	√	79.6	45.8	1.4	4.7	125

注: 加粗表示最优, “√”表示在基准模型中加入该模块。

M1 为基线模型 YOLOv8n 的实验结果, mAP50 为 75.4%, 参数量和计算量分别为 3.0×10^6 和 8.1×10^9 。M2、M3、M4、M5 分别为在基线模型中加入 PGC2f、FPSC、GLFPN 和 DSHead 检测头, 模型的平均检测精度 mAP50 较基线模型分别提升了 2.1、0.6、1.6 和 1.8 百分点, 结果表明, PGC2f 模块能有效提升目标检测精度; 引入 FPSC 模块后较 M1 提升了 0.6 百分点, 虽然提升幅度较小, 但 FPSC 通过避免池化操作, 成功避免了分辨率损失, 能更好地保留细节信息, 进而提高了模型的准确性; 加入 GLFPN 模块后, mAP50 达到了 77.0%, 较基线模型提升了 1.6 百分点, FPS 降低至 133; GLFPN 的引入进一步提升了检测精度, 证明了其在处理多尺度特征方面的有效性, 但模型推理时间有所下降。M5 中, 加入 DSHead 模块后, 在提升精度的同时, 保持了较低的计算开销和参数量。在 M6 实验中, 结合 PGC2f 和 FPSC 模块, mAP50 提升了 3.2 百分点, 而参数量和计算量分别下降了 23.3% 和 24.6%。

该实验验证了 FPSC 能够在不增加计算复杂

度的情况下有效提升模型准确性, FPSC 通过不进行池化操作, 能够避免分辨率的损失。尤其是在与 PGC2f 结合后, 性能提升尤为明显, 相较于 PGC2f 单独使用时 (M2), 精度提升了 1.1 百分点。M7 中, 加入 PGC2f、FPSC 和 GLFPN 模块, mAP50 提升至 78.9%, 进一步验证了这些模块组合在提升精度上的协同作用, 相比于 M1, 精度提升了 3.5 百分点, 证明了这 3 种模块的有效整合。在 M8 实验中, 加入 PGC2f 和 DSHead 后, mAP50 提升至 76.6%, 较基线模型提升了 1.2 百分点。与此同时, 参数量和计算量分别下降了 50% 和 43.2%, 证明了 DSHead 的引入, 不仅提升了精度, 而且大幅减少了模型的复杂度。M9 的实验结果显示, 加入 PGC2f 和 GLFPN 后, mAP50 提升至 77.2%, 较基线模型提升了 1.8 百分点, 尽管精度提升较为有限, 但参数量和计算量分别下降了 46.6% 和 27.1%, 表明这种组合有效地优化了模型的轻量化程度。在 M10 实验中, 加入 GLFPN 和 DSHead 模块后, mAP50 提升至 78.4%, 较基线模型提升了 3 百分点, 同时, 参数量和计算量分别下

降了 43.3% 和 20.9%，进一步验证了这两种模块在提升模型精度的同时能够有效减小模型的复杂度。M11 的实验结果显示，所有改进模块组合后的模型达到了 79.6% 的 mAP50，相较于基准模型 M1 提升了 4.2 个百分点，mAP50-95 较基准模型提升了 2.8 个百分点，达到 45.8%，与此同时，参数量下降了 54%，计算量下降了 42%。这一结果表明，虽然检测速度有所下降，但仍能满足边缘设备实时性的需求，改进模型在有效进行轻量化的同时，模型的综合性能得到了显著提升，达到了性能和效率的良好平衡。

通过逐一引入不同的改进模块，实验结果表明 PGC2f 和 DSHead 对模型精度的提升具有显著作用，而 FPSC 则通过优化特征提取过程，保持了较低的计算开销。最终，所有模块的组合 (M11) 使得模型的精度和效率均达到了最优，证明了在保证精度的同时，通过合理的轻量化设计可以有效减少计算开销和参数量，从而适应更为复杂和

高效的目标检测任务。

2.4.2 CGLU 性能分析

为了验证 CGLU 注意力机制在 NEU-DET 缺陷检测上的优越性，基于改进的 YOLO 算法，引入了多种注意力机制进行对比实验，采用了以下 4 种注意力机制：1) 基于通道注意力的一维 SE(squeeze-and-excitation)^[30] 模块，2) 结合通道与空间注意力的 CBAM(convolutional block attention module)^[31]，3) 将空间信息嵌入通道的 CA(coordinate attention)^[32] 模块，4) EMA(efficient multi-scale attention)^[33] 模块，以增强模型对多尺度特征的捕捉能力。通过对比实验，系统地评估了这些注意力机制在目标检测任务中的表现。结果如表 5 所示，CGLU 注意力机制的表现最好，mAP50 最高，达到了 77.5%，比基线模型提升了 2.1 百分点，同时，计算量和参数量较基准模型也有一定优化，这是因为通过门控线性动态调整特征响应，能够更好地建模非线性关系。

表 5 注意力消融实验
Table 5 Ablation experiment of attention

模型	P/%	R/%	mAP50/%	mAP50-95/%	Pa/10 ⁶	FLOPs/10 ⁹
YOLOv8n	73.6	70.2	75.4	42.6	3.00	8.1
Pconv+CBAM	71.2	72.8	75.5	42.6	2.12	7.4
Pconv+CA	67.2	72.6	74.8	41.2	2.86	7.1
Pconv+SE	70.6	72.2	75.8	43.1	2.99	7.4
Pconv+EMA	69.3	73.3	75.7	42.8	2.31	6.5
Pconv+CGLU	71.8	74.7	77.5	44.4	2.22	6.2

注：加粗表示最优。

2.4.3 PGC2f 位置对检测性能影响分析

为研究 PGC2f 添加的位置对模型的影响，分别做了 3 组消融实验：+Backbone 表示仅替换主干网络中的 C2f 模块，+Neck 表示仅替换颈部网络中的 C2f 模块，+Both 表示主干和颈部网络中的 C2f 全部替换为 PGC2f，实验结果如表 6 所示。

表 6 PGC2f 消融实验
Table 6 Ablation experiment of PGC2f

模型	mAP50/%	Pa/10 ⁶	FLOPs/10 ⁹
YOLOv8n	75.4	3.0	8.1
+Backbone	77.2	2.6	6.9
+Neck	77.7	2.6	7.3
+Both	77.5	2.3	6.1

注：加粗表示最优。

单独替换主干网络 (+Backbone)，mAP50 提升了 1.8 百分点，参数量和计算量均有所降低，这是因为通过 Pconv 只对局部进行卷积提局部特征，

减少了冗余计算。单独替换颈部网络 (+Neck)，mAP50 提升了 2.3 百分点，参数量和计算量分别下降了 13% 和 10%。同时替换主干和颈部网络 (+Both)，mAP50 提升了 2.1 百分点，参数量和计算量分别下降了 23.3% 和 24.6%，精度相比单独替换颈部网络有所降低，下降了 0.2 百分点，这是因为主干和颈部均使用 Pconv 对部分通道进行卷积，未被处理的通道依赖后续操作间接交互，可能导致关键通道的信息未被充分激活，但模型的复杂度得到了进一步的优化，参数量和计算量对比单独替换颈部网络均有大幅下降，实现更低的资源消耗。尤其是对于边缘设备，PGC2f 的设计非常适合在保证高精度的同时，减少计算量和内存需求。

2.4.4 损失函数对比

为了验证 WIoU 损失函数的有效性，在改进模型的基础上，将 WIoU 损失函数与 CIoU、EIoU、

DIoU、GIoU 和 MPDIoU 损失函数进行了对比实验, 实验结果如表 7 所示。

表 7 损失函数对比实验
Table 7 Comparison experiment of loss functions %

模型	P	R	mAP50
CIoU	74.9	72.8	79.3
EIoU	72.9	71.5	78.0
DIoU	83.7	67.2	78.5
GIoU	70.1	76.2	78.9
MPDIoU	73.7	72.7	78.6
WIoU	75.6	71.9	79.6

注: 加粗表示最优。

实验结果表明, 使用 WIoU 损失函数的模型性能最优。WIoU 通过动态调整高质量和低质量

锚框的权重, 显著提升了模型对高质量锚框的关注度, 同时有效降低了低质量锚框对训练过程的干扰。这种机制不仅加速了模型的收敛速度, 还显著减少了回归定位损失, 从而提高了检测精度。此外, WIoU 能够更好地处理目标重叠、尺度变化以及边界模糊等问题, 具有更强的鲁棒性和泛化能力, WIoU 具有更强的适应性, 能够更精准地优化检测框的质量, 使得模型在各种复杂情况下仍能保持较高的检测性能。

2.5 可视化结果对比

为进一步验证改进模型 PFGD-YOLOv8 的有效性, 在测试集中与 YOLOv8、YOLOv10、YOLOv11 和 YOLOv12 进行了可视化对比, 分析涵盖了 6 个不同的目标类别, 展示了在不同条件下的检测能力, 结果如图 9 所示。

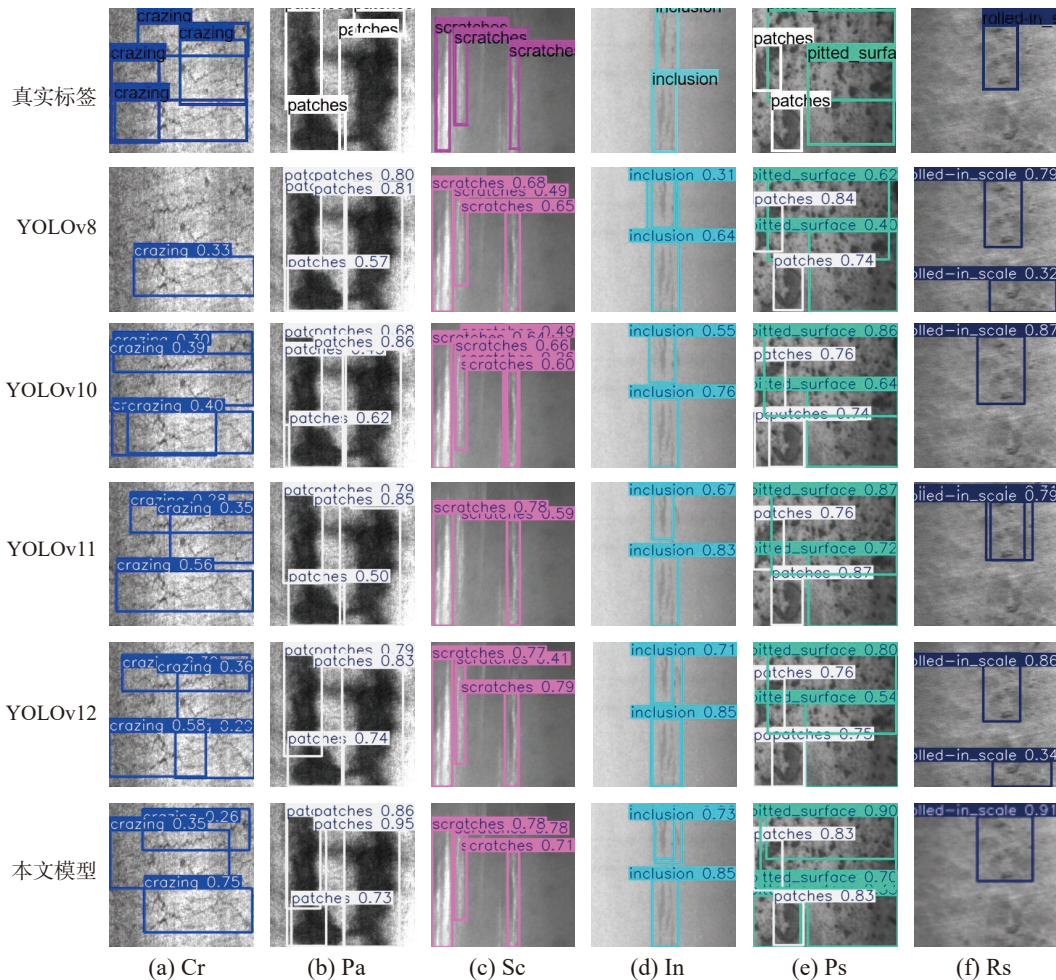


图 9 检测结果对比

Fig. 9 Comparison of test results

在图 9(a) Cr 缺陷的检测中, YOLOv8 存在漏检问题, 检测置信度为 0.33, YOLOv10 和 YOLOv11 成功识别缺陷位置, 但检测精度较低, 而改进模型的置信度提高至 0.75, 成功地识别了缺陷并准

确定位了目标位置, 表明改进模型在细小缺陷的检测上具有更好的敏感性和准确性。在图 9 (b) Pa 和 (c) Sc 的缺陷检测中改进模型对缺陷识别更加准确, YOLOv11 出现了漏检现象, 表明改进模

型具有更强的检测能力。在图 9 (d)In 缺陷的检测中, YOLOv8 的置信度分别为 0.31 和 0.64, YOLOv11 和改进模型的置信度相对较高, 分别为 0.67、0.83 和 0.71、0.85。背景存在一定的纹理干扰, 表明改进模型具有更强的抗干扰能力, 能在复杂背景下更精准地识别缺陷。在图 9 (e) Ps 缺陷的检测中, YOLOv8 的检测框定位存在偏差, 而改进模型的定位更加准确, 检测框的位置准确度有了明显改善, 表明了改进模型在检测框定位方面的优势。此外, YOLOv8 在缺陷与背景相似的情况下存在误检问题。在图 9 (f) Rs 缺陷的检测中, 缺陷特征为表面存在与背景纹理相似的滚压痕迹, 检测结果显示 YOLOv8 在此类目标的检测中出现误

检, YOLOv10 和 YOLOv11 检测精度较低。改进模型通过提高对背景与缺陷的区分能力, 显著减少了误检, 且提高了检测精度。综上所述, 改进后的 PFGD-YOLOv8 模型在检测精度、定位精度和鲁棒性方面均优于 YOLOv8, 特别是在小尺度目标和复杂背景下。通过引入更加高效的特征提取和融合模块, 改进模型能够更好地应对复杂环境中的目标检测任务。为更直观展示改进模型对目标特征的识别效果, 通过对比 YOLOv8、YOLOv10、YOLOv11 和 YOLOv12 与本文方法的热力图分布, 直观反映检测目标特征变化情况, 实验结果如图 10 所示, 其中位置亮度越高, 表明模型更注意该位置。

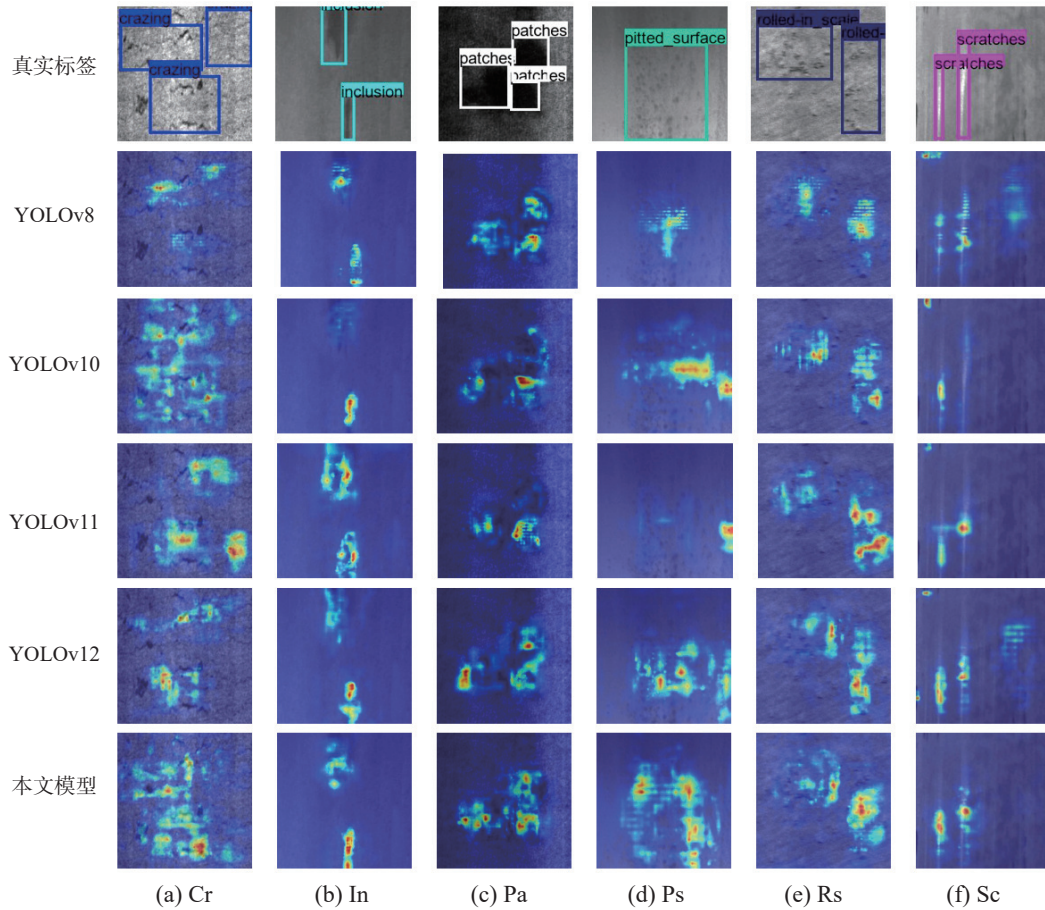


图 10 不同模型热力图对比
Fig. 10 Comparison of heatmaps from different models

通过热力图可视化分析, 本文方法在目标特征完整性、背景噪声抑制和多尺度适应性上优于 YOLOv8、YOLOv10、YOLOv11 与 YOLOv12。改进模型的热力图对 Cr、Pa 和 Ps 缺陷类别给予更显著的关注; 对 In、Rs 和 Sc 缺陷类别聚焦更加精准。YOLOv8 在 Sc 缺陷检测时, 对部分区域存在过度聚焦, 可能引入背景干扰, 说明改进模型在聚焦缺陷特征方面展现出更优的有效性, 能够更

精准、高效地捕捉缺陷特征。

3 结束语

本文针对钢材表面缺陷检测任务中存在检测精度不足、模型计算量大和难以在资源受限的设备上部署等问题, 提出了一种基于改进 YOLOv8 的轻量化检测算法 PFGD-YOLOv8。本文设计的 PGBN 模块通过部分卷积操作显著降低了计算量

和模型参数,同时提高了对局部特征的提取能力。改进的 FPSC 模块通过不同膨胀率的卷积操作扩大了感受野,有效提升了对小目标的检测精度。GLFPN 模块通过全局与局部信息的加权融合,更好地保留了小目标特征。与现有基于 YOLO 系列的检测模型相比, PFGD-YOLOv8 在保持高检测精度的同时,显著降低了模型的计算量和参数量,这主要得益于本文提出的轻量化网络结构和改进的特征提取机制。综上所述, PFGD-YOLOv8 算法在钢材表面缺陷检测领域展示了显著的优势和潜力,为工业自动化检测提供了一种高效、准确的解决方案。未来的研究将在此基础上对模型结构进行优化,以实现更广泛的应用和更优的性能。

参考文献:

- [1] WANG Ling, LIU Xinbo, MA Juntao, et al. Real-time steel surface defect detection with improved multi-scale YOLO-v5[J]. *Processes*, 2023, 11(5): 1357.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580–587.
- [3] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1440–1448.
- [4] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. [S. l.]: IEEE, 2017: 1137–1149.
- [5] CAI Zhaowei, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6154–6162.
- [6] PANG Jiangmiao, CHEN Kai, SHI Jianping, et al. Libra R-CNN: towards balanced learning for object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 821–830.
- [7] 陈远露, 王亮. 基于 Transformer 与 FasterRCNN 的多模态特征提取与融合[J]. *信息技术与信息化*, 2024(5): 111–114.
CHEN Yuanlu, WANG Liang. Multi-modal feature extraction and fusion based on Transformer and Faster-RCNN[J]. *Information technology and informatization*, 2024(5): 111–114.
- [8] SUN Peize, ZHANG Rufeng, JIANG Yi, et al. Sparse R-CNN: end-to-end object detection with learnable proposals[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 14449–14458.
- [9] 邓慧, 曾磊. 基于改进 Faster R-CNN 的热轧带钢表面缺陷检测[J]. *控制工程*, 2024, 31(4): 752–759.
DENG Hui, ZENG Lei. Surface defect detection of hot-rolled strip steel based on improved faster R-CNN[J]. *Control engineering of China*, 2024, 31(4): 752–759.
- [10] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Computer Vision–ECCV 2016. Cham: Springer International Publishing, 2016: 21–37.
- [11] 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.
- [12] WANG Ao, CHEN Hui, LIU Lihao, et al. Yolov10: real-time end-to-end object detection[J]. *Advances in neural information processing systems*, 2024, 37: 107984–108011.
- [13] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024–10–23) [2025–04–23]. <https://arxiv.org/abs/2410.17725>.
- [14] 李刚, 邵瑞, 周鸣乐, 等. 基于注意力的轻量级工业产品缺陷检测网络[J]. *计算机工程*, 2023, 49(11): 275–283.
LI Gang, SHAO Rui, ZHOU Mingle, et al. Lightweight industrial products defect detection network based on attention[J]. *Computer engineering*, 2023, 49(11): 275–283.
- [15] 忻迪晔, 严怀成. 基于 GS-YOLO 模型的带钢表面缺陷检测[J]. *计算机应用*, 2024, 44(S2): 302–308.
XIN Diye, YAN Huaicheng. Surface defect detection of strip steel based on GS-YOLO model[J]. *Journal of computer applications*, 2024, 44(S2): 302–308.
- [16] 敖思铭, 周诗洋, 杨智颖, 等. 基于 KAS-YOLO 的钢板表面缺陷检测[J]. *组合机床与自动化加工技术*, 2024(8): 168–174.
AO Siming, ZHOU Shiyang, YANG Zhiying, et al. Surface defect detection of steel plate based on KAS-YOLO[J]. *Modular machine tool & automatic manufacturing technique*, 2024(8): 168–174.
- [17] 侯玥, 王开宇, 金顺福. 一种基于 YOLOv5 的小样本目标检测模型[J]. *燕山大学学报*, 2023, 47(1): 64–72.
HOU Yue, WANG Kaiyu, JIN Shunfu. A few-shot object detection model based on YOLOv5[J]. *Journal of Yanshan University*, 2023, 47(1): 64–72.
- [18] 张上, 许欢, 张岳. 轻量级锻件表面裂纹检测算法[J]. *电子测量技术*, 2024, 47(11): 123–130.
ZHANG Shang, XU Huan, ZHANG Yue. Lightweight forged part surface crack detection algorithm[J]. *Electron-*

- ic measurement technology, 2024, 47(11): 123–130.
- [19] CHEN Jierun, KAO S H, HE Hao, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 12021–12031.
- [20] SHI Dai. TransNeXt: robust foveal visual perception for vision Transformers[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 17773–17783.
- [21] DAUPHIN Y N, FAN A, AULI M, et al. Language modeling with gated convolutional networks[C]//International Conference on Machine Learning. Sydney: PMLR, 2017: 933–941.
- [22] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10778–10787.
- [23] TANG Feilong, XU Zhongxing, HUANG Qiming, et al. DuAT: dual-aggregation Transformer network for medical image segmentation[C]//Chinese Conference on Pattern Recognition and Computer Vision. Singapore: Springer Nature Singapore, 2023: 343–356.
- [24] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: a simple and strong anchor-free object detector[J]. IEEE transactions on pattern analysis and machine intelligence, 2022, 44(4): 1922–1933.
- [25] YEUNG C C, LAM K M. Efficient fused-attention model for steel surface defect detection[J]. IEEE transactions on instrumentation and measurement, 2022, 71: 2510011.
- [26] WANG Xing, ZHUANG Kaiyu. An improved YOLOX method for surface defect detection of steel strips[C]//2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications. Shenyang: IEEE, 2023: 152–157.
- [27] 梁礼明, 陈康泉, 陈林俊, 等. 改进轻量化的 FCM-YOLOv8n 钢材表面缺陷检测[J]. 光电工程, 2025, 52(2): 117–129.
- LIANG Liming, CHEN Kangquan, CHEN Linjun, et al. Improving the lightweight FCM-YOLOv8n for steel surface defect detection[J]. Opto-electronic engineering, 2025, 52(2): 117–129.
- [28] 李相垚, 侯红玲, 杨澳, 等. 面向钢材表面缺陷检测的 DCS-YOLOv8 算法研究[J/OL]. 机械科学与技术, 2024: 1–10. (2024–10–10). <https://link.cnki.net/doi/10.13433/j.cnki.1003-8728.20240128>.
- LI Xiangyao, HOU Hongling, YANG Ao, et al. Research on DCS-YOLOv8 algorithm for steel surface defect detection[J/OL]. Mechanical science and technology for aerospace engineering, 2024: 1–10. (2024–10–10). <https://link.cnki.net/doi/10.13433/j.cnki.1003-8728.20240128>.
- [29] 赵曙光, 易文, 陆小辰. 基于 YOLOv7-Tiny 的轻量化钢材表面缺陷检测方法[J]. 东华大学学报(自然科学版), 2025, 51(4): 194–202.
- ZHAO Shuguang, YI Wen, LU Xiaochen. Lightweight steel surface defect detection method based on YOLOv7-Tiny[J]. Journal of Donghua University (natural science), 2025, 51(4): 194–202.
- [30] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [31] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 3–19.
- [32] HOU Qibin, ZHOU Daquan, FENG Jiashi. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13708–13717.
- [33] OUYANG Daliang, HE Su, ZHANG Guozhong, et al. Efficient multi-scale attention module with cross-spatial learning[C]//2023 IEEE International Conference on Acoustics, Speech and Signal Processing. Rhodes Island: IEEE, 2023: 1–5.

作者简介:



张冀, 副教授, 博士, 主要研究方向为计算机测控、故障诊断、信息融合、图像处理和深度学习。发表学术论文 20 余篇, 出版规划教材 2 部。E-mail: 72zhangji@163.com。



王定邦, 硕士研究生, 主要研究方向为计算机视觉。E-mail: wangdb951@163.com。



曹锦纲, 讲师, 博士, 主要研究方向为图像处理和模式识别, 发表学术论文 10 余篇。E-mail: caojg168@126.com。