

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

PIC2f-YOLO: 金属表面缺陷检测轻量化方法

胡依伦, 杨俊, 许聪源, 夏亚金, 邓文斌

引用本文:

胡依伦, 杨俊, 许聪源, 等. PIC2f-YOLO: 金属表面缺陷检测轻量化方法[J]. 光电工程, 2025, 52(1): 240250.

Hu Y L, Yang J, Xu C Y, et al. PIC2f-YOLO: a lightweight method for the detection of metal surface defects[J]. *Opto-Electron Eng*, 2025, 52(1): 240250.

<https://doi.org/10.12086/oe.2025.240250>

收稿日期: 2024-10-23; 修改日期: 2024-12-15; 录用日期: 2024-12-16

相关论文

改进GBS-YOLOv7t的钢材表面缺陷检测

梁礼明, 龙鹏威, 卢宝贺, 李仁杰

光电工程 2024, 51(5): 240044 doi: 10.12086/oe.2024.240044

基于改进YOLOv5s网络的斜拉桥拉索表面缺陷检测

王鹏峰, 李运堂, 黄永勇, 朱文凯, 林婕, 王斌锐

光电工程 2024, 51(5): 240028 doi: 10.12086/oe.2024.240028

局部和全局特征融合的太阳能电池片表面缺陷检测

陶志勇, 何燕, 林森, 易廷军, 张尧晟

光电工程 2024, 51(1): 230292 doi: 10.12086/oe.2024.230292

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



 OE_Journal



Website

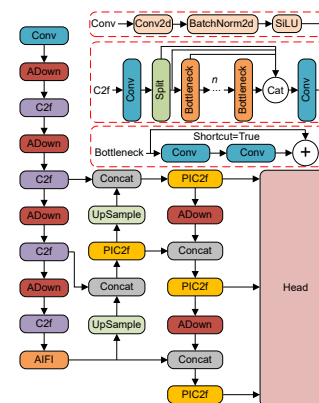
PIC2f-YOLO: 金属表面缺陷检测轻量化方法

胡依伦^{1,2}, 杨俊², 许聪源², 夏亚金³, 邓文斌^{2*}

¹浙江理工大学计算机科学与技术学院, 浙江 杭州 310018;

²嘉兴大学信息科学与工程学院, 浙江 嘉兴 314001;

³海盐中达金属电子材料有限公司, 浙江 嘉兴 314300



摘要: 针对金属表面缺陷检测效率低以及检测算法参数量大、精度低的问题, 本文提出了一种改进 YOLOv8n 的金属表面缺陷检测轻量化方法。首先, 设计局部卷积倒置交叉融合 (partial inverted bottleneck cross stage partial fusion, PIC2f) 模块, 该模块通过构造的局部卷积倒置瓶颈 (partial IRMB bottleneck, PIBN) 模块替换 BottleNeck 模块, 将部分卷积和倒置残差块组合, 从而减少算法的参数量并提升模型的特征提取能力。然后, 采用基于注意力尺寸内特征交互 (attention-based intra-scale feature interaction, AIFI) 模块, 该模块结合位置嵌入和多头注意力机制, 增强了模型对小目标的检测能力。最后, 使用平均池化下采样 (average pooling down sampling, ADown) 模块替换传统卷积作为模型特征缩减模块, 通过池化和卷积操作, 在不降低检测精度情况下, 进一步减少模型的参数量和计算复杂度。实验结果表明, 与 YOLOv8n 算法相比, 在 NEU-DET 钢材缺陷数据集上的 PIC2f-YOLO 方法的 mAP_{50} 增加了 2.7%, 参数量减少了 0.403 M。在铝片表面工业缺陷、PASCAL VOC2012 和带状合金功能材料表面缺陷数据集上的泛化性实验也验证了 PIC2f-YOLO 方法的有效性。

关键词: 表面缺陷检测; 局部卷积; 倒置残差; 注意力机制

中图分类号: TP391.41

文献标志码: A

胡依伦, 杨俊, 许聪源, 等. PIC2f-YOLO: 金属表面缺陷检测轻量化方法 [J]. 光电工程, 2025, **52**(1): 240250

Hu Y L, Yang J, Xu C Y, et al. PIC2f-YOLO: a lightweight method for the detection of metal surface defects[J]. *Opto-Electron Eng.* 2025. **52**(1): 240250

PIC2f-YOLO: a lightweight method for the detection of metal surface defects

Hu Yilun^{1,2}, Yang Jun², Xu Congyuan², Xia Yajin³, Deng Wenbin^{2*}

¹College of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou, Zhejiang 310018, China;

² College of Information Science and Engineering, Jiaxing University, Jiaxing, Zhejiang 314001, China;

³ Haiyan ZhongDA METAL Electronic Material Co., LTD, Jiaxing, Zhejiang 314300, China

Abstract: To address the low efficiency in metal surface defect detection, and the problems related to numerous

收稿日期: 2024-10-23; 修回日期: 2024-12-15; 录用日期: 2024-12-16

基金项目：国家自然科学基金项目 (62302197)；浙江省自然科学基金项目 (LQ23F020006)；嘉兴市科技计划基金项目 (2024AD10045, 2024AY40010)；浙江省教育厅一般科研项目 (Y202455539)

*通信作者: 邓文斌, wbdeng@zixu.edu.cn。

版权所有©2025 中国科学院光电技术研究所

model parameters and low precision, a lightweight detection method based on an improved YOLOv8n was proposed. The partially inverted bottleneck cross-stage partial fusion (PIC2f) module was introduced, replacing the bottleneck module with a partial IRMB bottleneck (PIBN) module. This combination of partial convolution and inverted residual blocks reduced the algorithm's parameters and enhanced the model's feature extraction ability. The attention-based intra-scale feature interaction (AIFI) module was applied, integrating location embedding and multi-head attention to improve the model's small-target detection performance. Lastly, the average pooling down sampling (ADown) module replaced traditional convolution as the feature reduction module, reducing parameters and computational complexity while maintaining detection accuracy. The experimental results show that, compared to YOLOv8n, the PIC2f-YOLO method improves mAP_{50} by 2.7% on the NEU-DET steel defect dataset and reduces parameters by 0.403 M. Generalization experiments on aluminum sheet surface industrial defects, PASCAL VOC2012 and surface defects of strip alloy functional material datasets also confirm the method's effectiveness.

Keywords: surface defect detection; partial convolution; inverted residual; attention mechanism

1 引言

金属材料作为工业领域的常用材料,具有成熟的制造工艺和显著的规模效益,是全球工业化和信息化发展进程中重要材料之一^[1]。金属材料产品表面缺陷检测主要依赖人工检测。随着生产规模扩大和质量要求提高,人工检测存在主观性强、效率低等缺点,如何快速识别金属表面缺陷已成为目前研究的热点。

早期的金属表面缺陷检测技术主要是传统无损检测技术,例如 Pan 等^[2]提出了漏磁检测,通过生成漏磁场检测磁化钢材胚料表面缺陷,但对于较小的缺陷,因漏磁场不足,效果不佳;Xu 等^[3]提出了红外检测,利用红外辐射通过表面温度变化检测钢板缺陷,但仅靠温度分析难以精确分类缺陷区域。为解决模型面对小目标缺陷难以识别的问题,Li 等^[4]提出了一种基于机器学习的钢球表面质量检测方法,利用核极限学习机的高效训练,完成对小钢球表面缺陷的高效检测和精确分类。Lu 等^[5]提出了一种改进的基于机器视觉的钢板表面缺陷检测方法,该方法基于有向无环图的支持向量机 (support vector machines, SVM) 利用无环图对数据进行分层决策,分类钢板表面缺陷,实现高精度的钢板表面缺陷检测。然而,传统的机器学习方法往往无法达到所需的检测精度,不能满足工业生产应用需求。

随着硬件算力的发展,基于深度学习的目标检测应用领域越来越广,例如 Zhang 等^[6]设计的 BS-YOLOv5s 模型应用于进行航空绝缘子图像缺陷检测,Ding 等^[7]提出了 SCD-YOLO 方法实现道路裂缝检测。对于金属表面缺陷检测领域的深度学习方法,主要分为二阶段目标检测和一阶段目标检测^[8]。在二阶段目

标检测领域以 Faster-RCNN^[9] 为代表,Liu 等^[10]提出了一种基于改进 Faster-RCNN 的钢材表面缺陷检测方法,该方法融合残差网络和双路径特征提取网络 (dual path network, DPN) 进行特征提取,具有重复利用已有特征,发掘新的特征以及减少特征冗余的特点,提升模型的表达能力。Shi 等^[11]提出一种改进 Faster-RCNN 的钢板表面缺陷检测方法,通过卷积和注意力模块,抑制复杂背景的特征并提高模型的特征提取能力。Jiang 等^[12]提出了改进 Faster-RCNN 的金属丝网缺陷检测方法,引入 K 均值聚类 (K-means) 算法和遗传算法,使锚框适应金属丝网尺寸,提高模型的缺陷定位能力。在一阶段目标检测领域以 YOLO 为代表, Ma 等^[13]提出了一种改进的钢铁表面缺陷检测算法 MT-YOLOv5,该算法在 YOLOv5 的基础上引入了 Transformer^[14] 模块,以有效提取图像的全局特征信息。Li 等^[15]提出了一种基于 YOLOv5 的改进检测算法用于钢材表面缺陷检测。该算法使用 GhostBottleneck 轻量化神经网络模块,替换原 YOLOv5 网络中的 C3 (cross stage partial bottleneck with 3 layers) 模块和少数卷积结构,减少了模型的参数量和计算量。Xiong 等^[16]提出了基于改进 YOLOX 的钢材表面缺陷检测算法,其引入 Swin Transformer 模块捕获缺陷区域全局上下文信息,采用加权双向特征金字塔网络 (BiFPN) 实现跨尺度特征融合,改进模型的损失函数,具有较高的准确率和实用性。Liang 等^[17]提出了一种改进的钢材表面缺陷算法 GBS-YOLOv7t,该算法设计 GAC-FPN 网络,嵌入双层路由注意力模块且引入 SIOU (SCYLLA intersection over union) 损失函数,优化边界框回归过程在 NEU-DET 数据集上实验效果明显。

积倒置交叉融合 (partial inverted bottleneck cross stage partial fusion, PIC2f) 模块, 如图 3 所示。PIC2f 沿用 C2f 的分支设计, 输入特征经过普通卷积后被拆分为多个分支, 其中部分分支通过 PIBN 模块进行特征提取与增强, 随后通过特征融合 (Cat) 操作聚合多尺度信息, 并最终通过普通卷积生成输出特征。为了避免过多 PIBN 模块堆叠导致的特征冗余和计算开销增加, n 设置为 1。

PIBN 模块分为两个部分, 第一部分为特征提取模块, 通过两个连续的 PConv 对特征图进行复用以加强特征信息的传播。PConv 将输入特征图的部分通道用于特征提取, 其他的通道保持不变, 因此与普通卷积 (Conv) 相比, PConv 在计算量和参数量上存在优势。Conv 和 PConv 结构如图 4 所示, 其中 w 为特征图的宽, h 为特征图的高, c 为特征图的通道数, c_p 为 PConv 的通道数, k 为卷积核大小。其中 c_p 为输入特征图通道数的 25%, 其固定选择索引为 $[0, 1, \dots, c_p-1]$ 的通道, 目的为降低计算复杂度, 保持模型行为的一致性和结果的可复现性。

输入与输出通道数量相同的情况下, Conv 的浮点计算量 F_c 和参数量 P_c 为

$$F_c = h \times w \times k^2 \times c^2, \quad (1)$$

$$P_c = k^2 \times c^2. \quad (2)$$

PConv 的浮点计算量 F_p 和参数量 P_p 为

$$F_p = h \times w \times k^2 \times c_p^2, \quad (3)$$

$$P_p = k^2 \times c_p^2. \quad (4)$$

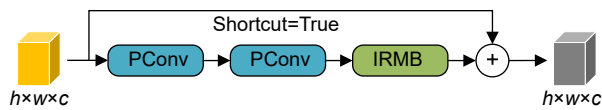


图 2 PIBN 结构
Fig. 2 PIBN structure

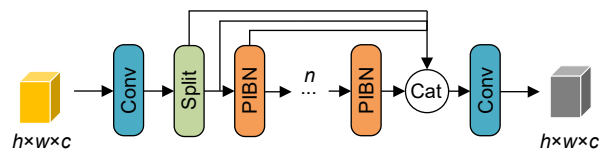


图 3 PIC2f 结构
Fig. 3 PIC2f structure

由于 PConv 利用了部分的通道数, 特征提取能力较弱, 因此在 PConv 后增加点乘卷积 (pointwise Conv, PWConv), 利用所有通道数以更好提取全局信息, 其浮点计算量 F_{p+pw} 和参数量 P_{p+pw} 分别为

$$F_{p+pw} = h \times w \times (k^2 \times c_p^2 + c^2), \quad (5)$$

$$P_{p+pw} = (1 + k^2) \times c_p^2. \quad (6)$$

当 PConv 利用 1/4 数量的通道时, PConv 和 PWConv 组合模块的浮点计算量和参数量分别为 Conv 的 $1/16 + 1/k^2$ 和 $1/16 + 1/(16k^2)$ 。因此, PConv 在计算量和参数量均优于 Conv 操作, 能够显著提升 PIBN 模块的计算效率和资源利用率。

PIBN 模块的第二部分为特征融合模块, 将提取的特征输入倒置残差块 (IRMB), 通过 Transformer 自注意力机制和深度可分离卷积进一步融合特征, 确保信息传递有效和梯度传播顺畅, 结构如图 5 所示。

首先, 对输入特征图进行 1×1 卷积, 生成值 (value, V)、查询 (query, Q) 和键 (key, K)。将 Q 和 K 输入到自注意力模块生成注意力矩阵 (attention matrix, Atten Mat)。该矩阵与 V 进行矩阵乘法, 将注意力分布加权到 V 上, 增强输入图像中相关的区域信息, 形成最终注意力输出。然后, 对加权求和后的特征图进行 3×3 的深度可分离卷积 (DWConv), 进一步提取局部特征。为了确保信息的有效传递和梯度的顺畅传播, 将加权求和后的特征图与深度可分离操作后的特征图通过残差连接相加。最后, 通过 1×1 卷积恢复特征图通道维度, 与输入特征图进行残差连接, 得到输出特征图, 从而增强模型的特征表达能力和整体性能。

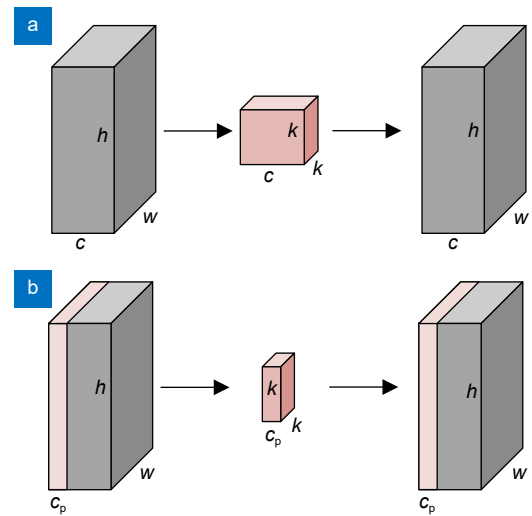


图 4 Conv 和 PConv 结构。(a) Conv; (b) PConv
Fig. 4 Conv and PConv structure. (a) Conv; (b) PConv

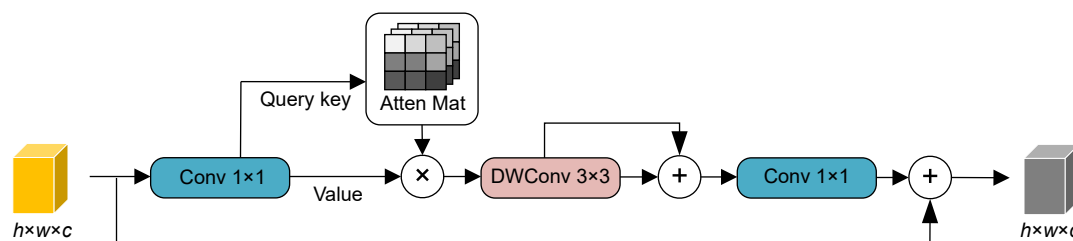


图 5 IRMB 结构

Fig. 5 IRMB structure

2.3 基于注意力的内部尺寸特征交互 AIFI 模块

金属表面缺陷呈现多样化的特征, 包括细长的夹杂物和划痕。在 YOLOv8 网络中, 主干网络末端的 SPPF 模块通过多尺度池化操作进行特征融合, 但由于使用固定大小的池化核, 导致模块泛化性有限且易丢失小目标的特征信息进而难以充分捕捉特定数据集下的小目标缺陷特征。针对这一问题, Zhao 等^[22]提出的基于注意力的尺寸内特征交互 (attention-based intra-scale feature interaction, AIFI) 模块, 通过结合位置嵌入和多头注意力机制, 充分利用空间位置信息和特征间依赖关系的优势, 有效解决了上述问题。AIFI 模块如图 6 所示, 首先对输入特征图进行二维正、余弦位置嵌入, 使模型学习不规则或细小缺陷之间的空间关系及排列方式。嵌入的位置信息 (position embedding) 与输入特征图输入到多头注意力 (multihead attention), 并行计算多个注意力头, 使模型从不同角度、区域捕捉特征间的依赖信息。然后, multihead attention 和输入特征图进行残差连接, 确保

信息的有效传递和梯度的顺畅传播。L2FG 对残差连接后的特征图依次进行归一化 (LayerNorm2d)、全连接 (FC)、激活 (GELU) 及再次 FC, 确保特征分布的一致性, 提高模型的收敛效果和特征表达能力。输入特征图与 L2FG 残差连接后, 通过 LayerNorm 对特征图进行调整, 确保特征分布的均衡性和稳定性, 并得到输出特征图。使得模块在保持计算效率的同时, 提高了模型的精度。

2.4 平均池化下采样 ADown 模块

YOLOv8n 中下采样操作通过步长为 2 的 3×3 卷积对特征图进行变化, 增加了模型的参数量和计算量。Wang 等^[23]提出的平均池化下采样 (average pooling down sampling, ADown) 模块通过池化与卷积操作, 不影响模型的检测性能, 减少模型参数量和计算量, ADown 结构如图 7 所示。

ADown 模块主要由 2D 平均池化层 (AvgPool2d)、2D 最大池化层 (MaxPool2d) 以及卷积层 (Conv) 组成。首先采用 2D 平均池化层计算输入特征图内所有值的

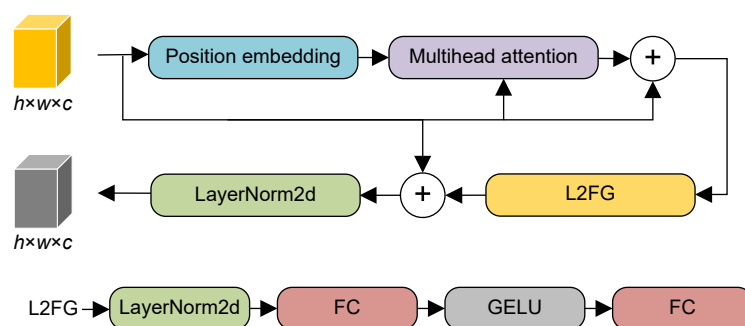


图 6 AIFI 结构

Fig. 6 AIFI structure

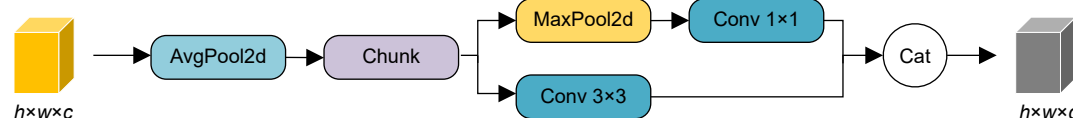


图 7 ADown 结构

Fig. 7 ADown structure

平均值, 以平滑输入特征图, 使模型捕捉图像中的基本结构并去除局部噪声。然后, 在通道数上均匀分割成两组特征图, 分别进行不同下采样。第一组特征图通过 2D 最大池化层, 保留局部最大值特征, 将图像减半, 抑制背景噪声, 再使用 1×1 卷积层融合像素特征。第二组特征图通过步长为 2 的 3×3 卷积层实现下采样。最后将两组特征图进行拼接实现通道数还原。

ADown 模块的参数数量和计算复杂度主要由 3×3 卷积层和 1×1 卷积层决定, 由于在下采样前特征图的通道数经过划分, 其参数数量和计算复杂度得到明显降低。假设特征图和下采样特征图的大小分别为 $h \times w \times c$ 和 $h/2 \times w/2 \times c$, 其中 h 为特征图的高, w 为特征图的宽, c 为特征图的通道数。ADown 模块的参数数量 P_a 、计算复杂度 F_a 、步长为 2 的下采样 3×3 卷积模块的参数数量 P_c 和计算复杂度 F_c 在数学上可分别表示为

$$\begin{cases} P_a = \frac{5}{2}c^2 \\ F_a = \frac{5}{8}c^2 \times h \times w \end{cases}, \quad (7)$$

$$\begin{cases} P_c = 9c^2 \\ F_c = \frac{9}{4}c^2 \times h \times w \end{cases}. \quad (8)$$

由式 (7) 及式 (8) 可知, 传统步长为 2 的下采样卷积的参数数量和计算复杂度是 ADown 模块的 3.6 倍。ADown 模块在参数数量和计算量上明显优于步长为 2 的下采样卷积操作, 使模型保持性能的同时大幅降低参数数量和计算量。

3 实验结果及分析

3.1 数据集

本文使用的 4 种数据集分别为东北大学 NEU-

DET 钢材缺陷数据集^[24]、铝片表面工业缺陷检测数据集^[25]、PASCAL VOC2012 数据集^[26] 和带状合金功能材料表面缺陷数据集^[27], 前三个数据集为公开数据集, 最后一个为自建数据集。对于上述数据集集中的训练集、验证集和测试集, 均采用 8 : 1 : 1 的比例随机划分。

上述四个数据集分为一个通用数据集、一个铝材数据集和两个钢材数据集。其中铝片表面工业缺陷检测数据集总包含 400 张图片, 包含划痕 (scratch)、脏污 (dirty)、褶皱 (fold) 和针孔 (pinhole) 共 4 种缺陷类型, 如图 8 所示。通用数据集 PASCAL VOC2012 总包含 17125 张图片, 共 20 种类型。

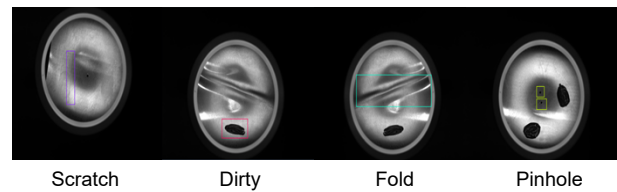


图 8 铝片缺陷
Fig. 8 Aluminum sheet defects

钢材数据集中, NEU-DET 数据集包含 1800 张图片, 包含裂缝 (crazing)、夹杂物 (inclusion)、斑块 (patches)、点蚀表面 (pitted surface)、氧化皮 (rolled in scale) 和划痕 (scratch) 共 8 种缺陷类型。带状合金功能材料表面缺陷数据集获取于企业产线, 使用两台线阵相机实时采集, 共 2942 张缺陷图像, 包含鼓包 (swelling)、凹坑 (dent)、划伤 (scratch)、起皮 (peeling)、缺口 (gap)、穿孔 (perforation)、焊缝 (weld) 和压痕 (snake) 共 8 种缺陷类型, 缺陷图像如图 9 所示, 每张图像的分辨率为 4096 pixel $\times$ 2048 pixel, 共 8 种缺陷类型。由于带状合金功能材料表面缺陷数据集存在样本分配不平衡和数量不足的问题, 因此通过缩放、水平翻转、旋转和裁剪等数据增强方式, 对数据集进

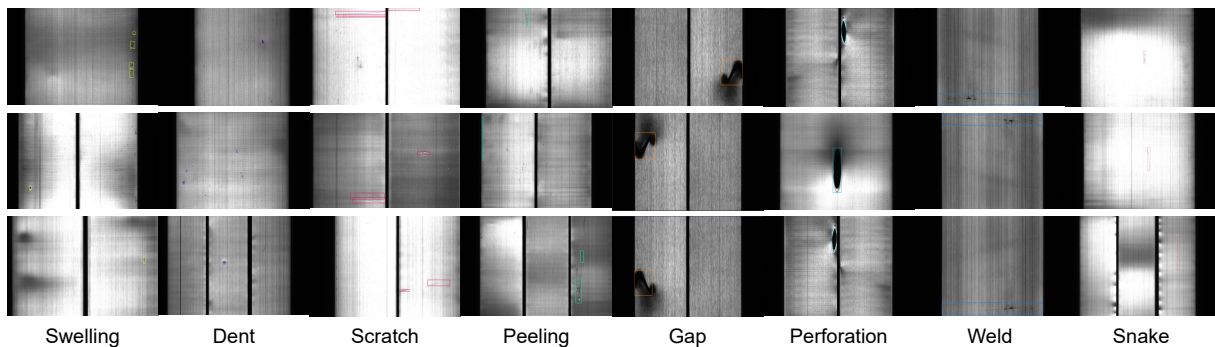


图 9 带状合金功能材料表面缺陷
Fig. 9 Surface defects of strip alloy functional material

行增强。增强后的带状合金功能材料表面缺陷数据集包含 5047 张图像。数据增强前后的缺陷标签分布如表 1 所示。在实际合金功能材料生产中, 起皮和划伤类缺陷更为普遍。因此, 与其他类别相比, 这两类标签数量更多, 有助于提高模型对这类缺陷的检测精度。

3.2 实验环境与参数设置

本文实验使用 Ubuntu 22.04 系统, 采用 PyTorch 2.3.1 框架、Intel Core i9-10900X CPU 及 RTX 3090 GPU (24 GB 显存)。在训练阶段, 采用 AdamW 作为优化器, 初始学习率为 0.001, 动量为 0.9, 衰减系数为 0.0005, 其余超参数均为默认。输入图像尺寸为 640×640、batch size 大小为 32、epochs 为 200。

3.3 评价指标

本文实验采用平均精度均值 (mAP_{50} 及 mAP_{50-95})、参数量 (Param)、计算量 (FLOPs) 和每秒帧数 (FPS) 四个指标对网络性能进行评估。 mAP_{50} 表示在交并比 (IOU) 的阈值为 0.5 时, 所有类别的检测准确率的均

值。 mAP_{50-95} 表示在 IOU 的阈值范围为 0.5 到 0.95 时, 以 0.05 为步长, 所有步长下平均精度 (AP) 的均值, 由精确度 (precision)、召回率 (recall) 和 AP 计算所得。

3.4 数据增强实验

使用 YOLOv8n 和 PIC2f-YOLO 方法在数据增强后的带状合金功能材料表面缺陷数据集上进行实验, 每个缺陷类别的 mAP_{50-95} 如图 10 所示。

其中+表示增强后的带状合金功能材料表面缺陷数据集。YOLOv8n 在数据增强前所有缺陷的 mAP_{50-95} 为 37.8%, 在数据增强后实现了 42.1% 的 mAP_{50-95} , 提高了 4.3%。PIC2f-YOLO 方法在数据增强之前所有缺陷的 mAP_{50-95} 为 36.2%, 在数据增强之后达到了 42.8%, 提高了 6.6%。可以看出, 所提方法在数据增强后性能提升明显, 表明数据增强有效提升模型性能。所提方法在数据集中图像数量较大的情况下, 具有更强的学习能力。

3.5 消融实验

本文 PIC2f-YOLO 方法以 YOLOv8n 为基线模型,

表 1 数据增强前后数据集中各类缺陷标签数

Table 1 Number of defect labels for each class in the dataset before and after data augmentation

Defects	Training set	Validation set	Testing set
Swelling	337/643	38/104	43/73
Dent	493/981	62/120	65/118
Scratch	1752/3006	189/375	228/405
Peeling	1145/1981	141/244	180/237
Gap	417/632	52/71	39/86
Perforation	131/239	11/20	17/27
Weld	285/432	29/51	30/47
Snake	119/254	27/33	19/30

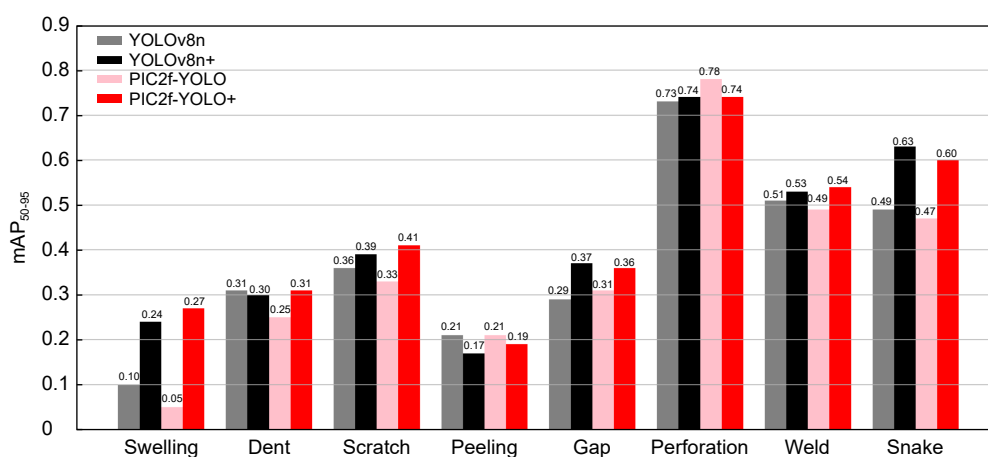


图 10 数据集增强前后缺陷 mAP_{50-95} 值对比

Fig. 10 Comparison of defect mAP_{50-95} values before and after dataset augmentation

围绕三个方面进行改进, 为验证每项改进的有效成分, 本文在 NEU-DET 数据集上进行六组消融实验。其中 M1 表示将颈部网络的 C2f 模块替换为 PIC2f 模块、M2 表示将主干网络的 SPPF 模块替换成 AIFI 模块、M3 表示将主干和颈部网络的下采样模块替换为 ADown 模块。实验结果如表 2 所示。

基线 (Baseline) 为未改进的 YOLOv8n。颈部网络引入 PIC2f 模块后的模型方法与基线相比 mAP_{50} 提高 1.6%, 参数量减少 0.218 M。使用 AIFI 注意力机制后的方法与基线相比 mAP_{50} 提高 1.8%, 计算复杂度保持不变。采用 ADown 模块的方法与基线相比 mAP_{50} 提高 0.9%, 参数量降低 0.415 M, 参数量指标排名第 1, 计算复杂度降低 0.9 G, 计算复杂度指标排名同样第 1。结合 PIC2f 和 AIFI 的方法与基线相比 mAP_{50} 提升 2.3%, mAP_{50} 指标排名第 2。融合 PIC2f、AIFI 和 ADown 的方法与基线相比 mAP_{50} 提升 2.7%, mAP_{50} 指标排名第 1, 参数量减少 0.403 M, 参数量指标排名第 2。

引入 PIC2f 模块后的模型与基线相比, mAP_{50} 提高至 76.9%, FPS 降低至 98 f/s。这表明在颈部网络的多个特征融合模块添加自注意力机制, 能捕获更多的上下文信息, 适用于多尺度缺陷检测任务, 但推理速度受到显著影响。在此基础上进一步引入 AIFI 模块, mAP_{50} 提高至 77.6%, FPS 降至 84 f/s, 说明骨干网络末尾添加多头注意力机制有效增加模型对于小尺寸图像的特征提取能力, 但 AIFI 模块作用于通道数

最多的特征图, 计算成本增加, 导致检测速度下降。进一步加入 ADown 模块后, mAP_{50} 提升至 78.0%, FPS 降低至 82 f/s, 表明下采样模块仅作用于特征图缩放, 对于模型的检测精度和计算速度影响较小。

在实际生产线上, 模型实际检测速度与工业相机采集图像及检测处理速度相关, 工业相机由镜头和图像传感器构成^[28]。其中图像传感器在一定分辨率下的最大帧率决定工业相机的 FPS, 例如 CMOS 传感器 FPS 为 30 f/s^[29]。工业相机的实际检测速度需要 30~60 f/s, 本文设计模块在 NEU-DET 数据集下 FPS 是 82 f/s, 满足生产线上实时检测要求, 基线模型虽然 FPS 达到 163 f/s, 但检测精度较低。因此, 需要在检测精度和处理速度之间达到某种平衡, 通过适度降低检测速度以显著提升检测精度的设计策略在实际应用中是合理的。消融实验结果进一步表明, 本文提出的 PIC2f-YOLO 方法在检测精度、推理速度和模型参数量之间达到了较好的平衡, 具有较高的实际应用价值。

3.6 PIC2f 模块有效性分析

为进一步分析本文 PIC2f-YOLO 方法中 PIC2f 模块的有效性, 本文在 NEU-DET 数据集上对 PIC2f 的 PConv 和 IRMB 进行消融实验, 并对 IRMB 进行深度可分离卷积核大小分析。

PIC2f 实验中基线是不使用 PConv 和 IRMB 的 YOLOv8n 模型, 表 3 中第二行仅使用 PConv 的模型, 与基线模型相比, mAP_{50} 增加 0.3%, 参数量降 0.417

表 2 消融实验结果

Table 2 Ablation results

Methods	$mAP_{50}/\%$	Param/M	FLOPs/G	FPS/(f/s)
Baseline	75.3	3.006	8.1	163
+M1	76.9	2.788	11.5	98
+M2	77.1	3.238	8.1	128
+M3	76.2	2.591	7.2	153
+M1M2	77.6	3.019	11.5	84
+M1M2M3	78.0	2.603	10.6	82

表 3 PIC2f 实验结果

Table 3 PIC2f experiment results

PConv	IRMB	$mAP_{50}/\%$	Param/M
		75.3	3.006
√		75.6	2.589
√	$k=1$	75.8	2.783
√	$k=3$	76.9	2.788
√	$k=5$	75.6	2.797
√	$k=7$	74.9	2.811
√	$k=9$	75.3	2.829

M, 参数量指标排名第 1。其他的实验组引入 IRMB 模块, 用于强化模型的特征融合功能。

为了验证不同大小感受野对模型的特征感知能力的影响, IRMB 模块中的深度可分离卷积核分别设置为 1、3、5、7、9, 从表中可以看出当卷积核设置为 3 时, mAP_{50} 最优, mAP_{50} 指标排名第 1。卷积核设置为 1 的实验组, 其 mAP_{50} 和参数量指标均排名第 2。卷积核设置为 3 与卷积核设置为 1 的实验组相比, 其使用更大的卷积核, 能更好地捕获图像的全局信息。然而卷积核设置为 5 时, 比卷积核设置为 3 的实验组的 mAP_{50} 降低 1.3%, 原因为过大的卷积核导致感受野过大, 丢失特征图局部信息, 进而影响模型性能。

因此选择合适的卷积核大小可以更好地平衡局部与全局信息, 提高模型检测精度。

3.7 可视化对比实验

本文通过比较在不同钢材数据集下的检测效果来展示 YOLOv8n 和本文方法的性能。图 11 表明 YOLOv8n 算法存在裂缝和氧化皮缺陷漏检现象, 难以满足工厂生产高质量钢带的需求。而本文方法能够更准确地检测出有效的缺陷区域, 有效的避免了漏检情况。在裂缝、斑块和划痕缺陷上 PIC2f-YOLO 方法相较于 YOLOv8n 具有较高的缺陷置信度。

图 12 表明 YOLOv8n 算法存在凹坑和划伤缺陷的漏检现象, 而本文算法能够更准确地检测出有效的

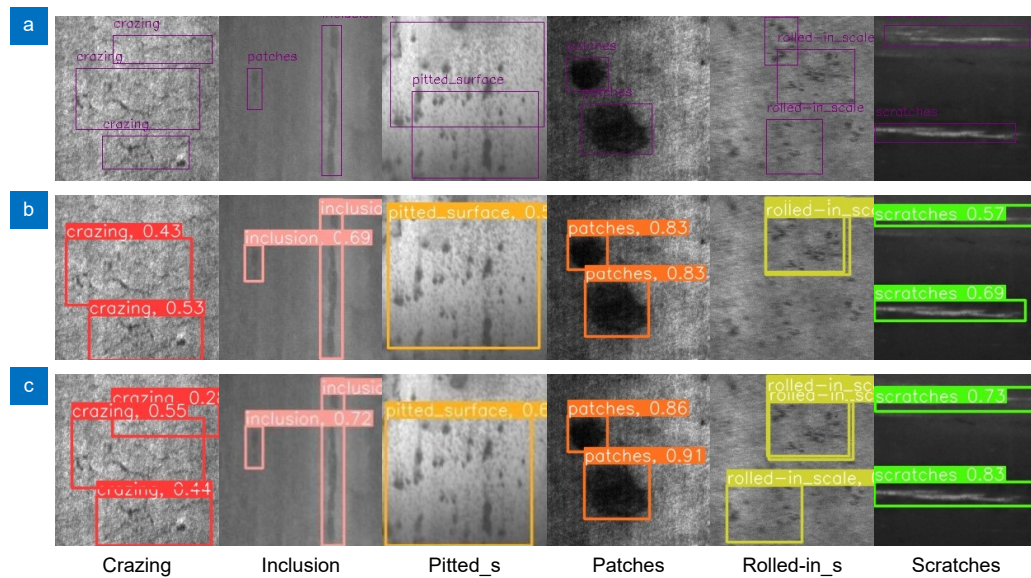


图 11 不同方法在 NEU-DET 数据集下的检测实例。(a) 标注的缺陷图像; (b) YOLOv8n; (c) PIC2f-YOLO

Fig. 11 Detection instance of different methods under the NEU-DET dataset. (a) Annotated defect images; (b) YOLOv8n; (c) PIC2f-YOLO

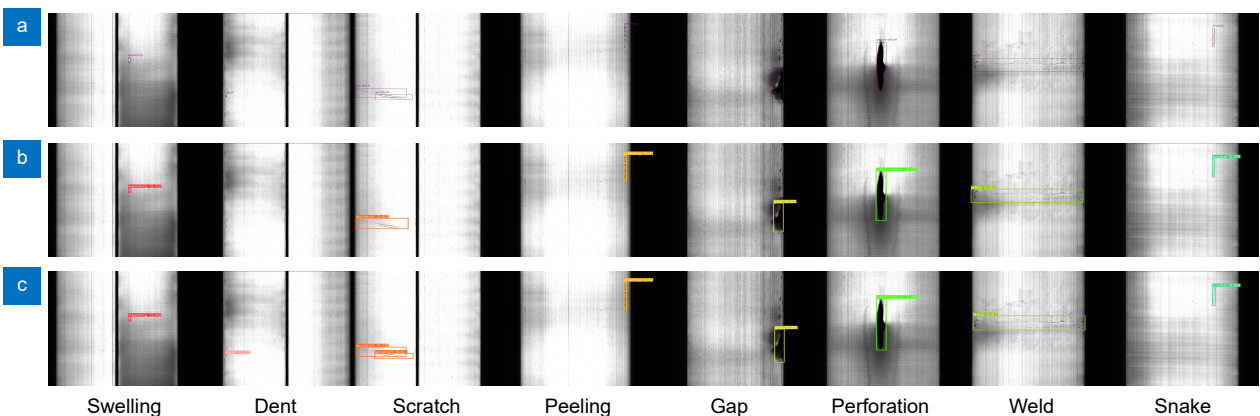


图 12 不同方法在带状合金功能材料表面缺陷数据集下的检测实例。(a) 标注的缺陷图像;

(b) YOLOv8n; (c) PIC2f-YOLO

Fig. 12 Detection instance of different methods under the surface defects of strip alloy functional materials datasets.

(a) Annotated defect images; (b) YOLOv8n; (c) PIC2f-YOLO

缺陷区域, 较好的避免了漏检情况。在鼓包、起皮、缺口、穿孔、焊缝和压痕缺陷上 PIC2f-YOLO 方法相较于 YOLOv8n 具有较高的缺陷置信度。因此本文提出的 PIC2f-YOLO 方法能更好避免漏检和误检, 具备更强的特征提取能力。

3.8 泛化性验证

本文实验采用铝片表面工业缺陷检测数据集^[25]和 PASCAL VOC2012 数据集^[26]对 PIC2f-YOLO 方法进行泛化性验证。将基线模型和本文算法在上述两种数据集上进行实验对比, 其实验结果如表 4 和表 5 所示, 可知 PIC2f-YOLO 方法在铝片表面工业缺陷数据集上的 mAP_{50} 达到 95.0%, 相较于基线模型提升了 1.2%, 参数量减少了 0.403 M。在 PASCAL VOC2012 的数据集上的 mAP_{50} 达到了 59.2%, 相较于基线模型提升了 0.4%, 参数量同样减少了 0.403 M。

实验结果表明, 本文提出的 PIC2f-YOLO 方法在不同类型数据集上均表现出较高的检测性能和较低的模型参数量, 验证了提出的算法具有较好的泛化

能力。

3.9 对比实验

为了进一步验证本文 PIC2f-YOLO 方法的有效性, 在实验设备与实验环境相同的情况下, 在 NEU-DET、VOC2012 和带状合金功能材料表面缺陷数据集上, 将所提方法与其他主流深度学习算法进行实验对比, 实验结果如表 6、表 7 和表 8 所示。

在 NEU-DET、VOC2012 和带状合金功能材料表面缺陷数据集中, Faster R-CNN^[9]为经典的二阶段检测算法, 分别实现 76.7%、56.0% 和 57.2% 的 mAP_{50} , 36.3%、40.7% 和 30.6% 的 mAP_{50-95} , 但其参数量大, 检测速度低。在 YOLO 系列算法中, PIC2f-YOLO 方法与 YOLOv3-tiny^[30]、YOLOv5n、YOLOv6n^[31]、YOLOv7-tiny^[32]、YOLOXs^[33]、YOLOv9-tiny^[23]、YOLOv10n^[34]和 YOLOv8n 相比, 在 NEU-DET 数据集中 mAP_{50} 指标分别高 8.4%、3.9%、8.0%、10.2%、3.4%、2.8%、7.1% 和 2.7%, 在 VOC2012 数据集中 mAP_{50} 指标分别高 6.6%、1.2%、0.3%、1.3%、1.6%、0.1%、0.3%

表 4 铝片表面工业缺陷数据集泛化实验

Table 4 Generalization experiment results on the aluminum sheet surface industrial defect dataset

Methods	$mAP_{50}/\%$	Param/M	FPS/(f/s)
YOLOv8n	93.8	3.006	190
PIC2f-YOLO	95.0	2.603	88

表 5 PASCAL VOC2012 数据集泛化实验

Table 5 Generalization experiment results on the PASCAL VOC2012 dataset

Methods	$mAP_{50}/\%$	Param/M	FPS/(f/s)
YOLOv8n	58.8	3.009	71
PIC2f-YOLO	59.2	2.606	45

表 6 NEU-DET 数据集实验效果对比

Table 6 Comparison experiment results on the NEU-DET dataset

Methods	$mAP_{50}/\%$	$mAP_{50-95}/\%$	Param/M	FLOPs/G	FPS/(f/s)
Faster-RCNN	76.7	36.3	41.37	134.0	32
YOLOv3-tiny	69.6	35.4	12.13	18.9	235
YOLOv5n	74.1	40.6	2.50	7.1	160
YOLOv6n	70.0	36.6	4.23	11.8	180
YOLOv7-tiny	67.8	31.2	6.03	13.2	102
YOLOXs	74.6	39.5	8.94	26.8	97
YOLOv9-tiny	75.2	41.9	2.62	10.7	130
YOLOv10n	70.9	38.9	2.69	6.7	140
YOLOv8n	75.3	40.7	3.01	8.1	163
PIC2f-YOLO	78.0	43.6	2.60	10.6	82

表 7 PASCAL VOC2012 数据集实验效果对比

Table 7 Comparison experiment results on the PASCAL VOC2012 dataset

Methods	mAP ₅₀ /%	mAP ₅₀₋₉₅ /%	Param/M	FLOPs/G	FPS/(f/s)
Faster-RCNN	56.0	40.7	41.43	134.0	13
YOLOv3-tiny	52.6	31.8	12.13	18.9	107
YOLOv5n	58.0	40.0	2.51	7.1	71
YOLOv6n	58.9	43.0	4.24	11.8	72
YOLOv7-tiny	57.9	40.4	6.06	13.2	58
YOLOXs	57.6	39.5	8.96	26.8	52
YOLOv9-tiny	59.1	44.8	2.62	10.7	92
YOLOv10n	58.9	43.1	2.70	6.7	95
YOLOv8n	58.8	40.9	3.01	8.1	71
PIC2f-YOLO	59.2	41.6	2.61	10.6	45

表 8 带状合金功能材料表面缺陷数据集实验效果对比

Table 8 Comparison experiment results on the surface defects of strip alloy functional material dataset

Methods	mAP ₅₀ /%	mAP ₅₀₋₉₅ /%	Param/M	FLOPs/G	FPS/(f/s)
Faster-RCNN	57.2	30.6	41.43	134.0	20
YOLOv3-tiny	63.5	33.7	12.13	18.9	151
YOLOv5n	72.3	38.6	2.50	7.1	108
YOLOv6n	68.9	37.4	4.23	11.8	117
YOLOv7-tiny	69.7	37.4	6.02	13.1	98
YOLOXs	66.7	34.5	8.94	26.8	95
YOLOv9-tiny	71.3	37.9	2.62	10.7	105
YOLOv10n	64.3	36.8	2.69	6.7	110
YOLOv8n	74.1	42.1	3.01	8.1	113
PIC2f-YOLO	75.6	42.8	2.61	10.6	60

和 0.4%，在带状合金功能材料表面缺陷数据集中 mAP₅₀ 指标分别高 12.1%、3.3%、6.7%、5.9%、8.9%、4.3%、11.3% 和 1.5%。对于 mAP₅₀₋₉₅ 指标，PIC2f-YOLO 方法在 NEU-DET 数据集和带状合金功能材料表面缺陷数据集上性能第 1，在 VOC2012 数据集上性能第 4。此外模型的参数量仅次于 YOLOv5n，浮点计算量位居第 4，同时检测速度满足实时需求。此外，在 NEU-DET 数据集中，对于 mAP₅₀ 指标，Faster-RCNN 排名第 2，对于 mAP₅₀₋₉₅ 指标，YOLOv9-tiny 排名第 2。在带状合金功能材料表面缺陷数据集中，对于 mAP₅₀ 和 mAP₅₀₋₉₅ 指标，YOLOv8n 均排名第 2。在 VOC2012 数据集中，对于 mAP₅₀ 指标，YOLOv9-tiny 排名第 2，对于 mAP₅₀₋₉₅ 指标，YOLOv9-tiny 排名第 1，YOLOv10n 排名第 2。此外在上述三种数据集中，对于 Param 指标，YOLOv5n 位列第 1，对于 FLOPs 指标，YOLOv10n 和 YOLOv5n 分别位列第 1 和第 2，对于 FPS 指标，YOLOv3-tiny 和 YOLOv6n

分别位列第 1 和第 2。

4 结 论

针对现有缺陷检测模型检测精度低、容易漏检以及参数大等问题，本文提出的 PIC2f-YOLO 方法能有效解决上述问题。PIC2f 多尺度特征提取融合模块和 AIFI 注意力模块能减少非缺陷区域的环境干扰，提高小目标缺陷的检测精度。同时采用 ADown 下采样模块在保持检测精度的同时减少模型的参数量。实验结果表明，提出的轻量化方法具有较好的性能及泛化性能。

参考文献

[1] Li G Q. Analysis on the development status of bell-less top equipment maintenance for blast furnaces[J]. *Plant Maintenance Eng*, 2019, (6): 35–36.
李国权. 高炉无料钟炉顶设备维护发展现状分析[J]. *设备管理与维修*, 2019, (6): 35–36.
[2] Pan C Y. Research on data compression method of massive

- magnetic flux leakage detection data of surface defects on steel plate based on improved SPIHT algorithm[D]. Shenyang: Northeastern University, 2013.
- 潘晨燕. 基于改进 SPIHT 算法的钢板表面缺陷漏磁检测海量数据压缩方法研究[D]. 沈阳: 东北大学, 2013.
- [3] Xu C H, Chen G M, Xie J. Application of infrared thermography technology in surface defects detection of products in metal[J]. *Manuf Autom Eng*, 2009, **31**(10): 51–54.
徐长航, 陈国明, 谢静. 红外图像处理技术在金属表面缺陷检测中的应用[J]. *制造业自动化*, 2009, **31**(10): 51–54.
 - [4] Li X D, Mao W J, Jiang W. Image recognition for steel ball's surface quality detecting based on kernel extreme learning machine[C]//*Proceedings of the 34th Chinese Control Conference*, 2015: 3727–3731.
<https://doi.org/10.1109/ChiCC.2015.7260217>.
 - [5] Lu J L, Lin M X, Huang Y, et al. A high-accuracy algorithm for surface defect detection of steel based on DAG-SVM[J]. *Sens Transducers*, 2013, **157**(10): 412–418.
 - [6] Zhang Z B, Lv G H, Zhao G X, et al. BS-YOLOv5s: insulator defect detection with attention mechanism and multi-scale fusion[C]//*Proceedings of 2023 IEEE International Conference on Image Processing*, 2023: 2365–2369.
<https://doi.org/10.1109/ICIP49359.2023.10222163>.
 - [7] Ding K Y, Ding Z H, Zhang Z B, et al. SCD-YOLO: a novel object detection method for efficient road crack detection[J]. *Multimedia Syst*, 2024, **30**(6): 351.
 - [8] Zhu K J. Research on object detection based on deep learning[J]. *Mod Inf Technol*, 2024, **8**(13): 76–83.
朱克佳. 基于深度学习的目标检测研究[J]. *现代信息技术*, 2024, **8**(13): 76–83.
 - [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 Trans Pattern Anal Mach Intell*, 2017, **39**(6): 1137–1149.
 - [10] Liu Q, Lei J S. Surface defect detection of steel based on improved deep network[J]. *Comput Eng Design*, 2022, **43**(9): 2654–2661.
刘琪, 雷景生. 基于改进深度网络的钢材表面缺陷检测[J]. *计算机工程与设计*, 2022, **43**(9): 2654–2661.
 - [11] Shi X C, Zhou S K, Tai Y C, et al. An improved Faster R-CNN for steel surface defect detection[C]//*Proceedings of 2022 IEEE 24th International Workshop on Multimedia Signal Processing*, 2022: 1–5.
<https://doi.org/10.1109/MMSP55362.2022.9949350>.
 - [12] Jiang F F, Li N, Qiu C C, et al. Defect detection method of wire mesh based on improved Faster RCNN[J]. *China Sciencepaper*, 2024, **19**(2): 153–159.
姜菲菲, 李宁, 邱翠翠, 等. 基于改进 Faster RCNN 的金属丝网缺陷检测方法[J]. *中国科技论文*, 2024, **19**(2): 153–159.
 - [13] Ma Y T, Zhao H D, Yan C, et al. Strip steel surface defect detection method by improved YOLOv5 network[J]. *J Electron Meas Instrum*, 2022, **36**(8): 150–157.
马燕婷, 赵红东, 阎超, 等. 改进 YOLOv5 网络的带钢表面缺陷检测方法[J]. *电子测量与仪器学报*, 2022, **36**(8): 150–157.
 - [14] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017: 6000–6010.
 - [15] Li X, Wang C, Li B, et al. Steel surface defect detection algorithm based on Improved YOLOv 5[J]. *J Air Force Eng Univ*, 2022, **23**(2): 26–33.
李鑫, 汪诚, 李彬, 等. 改进 YOLOv5 的钢材表面缺陷检测算法[J]. *空军工程大学学报*, 2022, **23**(2): 26–33.
 - [16] Xiong C, Yu A N, Gao X H, et al. Steel surface defect detection algorithm based on improved YOLOX[J]. *Electron Meas Technol*, 2023, **46**(9): 151–157.
熊聪, 于安宁, 高兴华, 等. 基于改进 YOLOX 的钢材表面缺陷检测算法[J]. *电子测量技术*, 2023, **46**(9): 151–157.
 - [17] Liang L M, Long P W, Lu B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. *Opto-Electron Eng*, 2024, **51**(5): 240044.
梁礼明, 龙鹏威, 卢宝贺, 等. 改进 GBS-YOLOv7t 的钢材表面缺陷检测[J]. *光电工程*, 2024, **51**(5): 240044.
 - [18] Yang B C, Li S X, Li Y Z, et al. Improved YOLOv8n algorithm for steel surface defect detection[J/OL]. *Mech Sci Technol Aerosp Eng*, 2024: 1–10. [2024-09-20].
<https://doi.org/10.13433/j.cnki.1003-8728.20240135>.
杨本臣, 李世熙, 李依泽, 等. 改进 YOLOv8n 的钢材表面缺陷检测算法[J/OL]. *机械科学与技术*, 2024: 1–10. [2024-09-20].
<https://doi.org/10.13433/j.cnki.1003-8728.20240135>.
 - [19] Ultralytics. YOLOv8: the latest version of YOLO object detection algorithm[EB/OL]. GitHub, 2023. [2024-08-22].
<https://github.com/ultralytics/ultralytics>.
 - [20] Chen J R, Kao S H, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 12021–12031.
<https://doi.org/10.1109/CVPR52729.2023.01157>.
 - [21] Zhang J N, Li X T, Li J, et al. Rethinking mobile block for efficient attention-based models[C]//*Proceedings of 2023 IEEE/CVF International Conference on Computer Vision*, 2023: 1389–1400. <https://doi.org/10.1109/ICCV51070.2023.00134>.
 - [22] Zhao Y, Lv W Y, Xu S J, et al. DETRs beat YOLOs on real-time object detection[C]//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 16965–16974.
<https://doi.org/10.1109/CVPR52733.2024.01605>.
 - [23] Wang C Y, Yeh I H, Liao H Y M, et al. YOLOv9: learning what you want to learn using programmable gradient information[C]//*Proceedings of the 18th European Conference on Computer Vision*, 2024: 1–21.
https://doi.org/10.1007/978-3-031-72751-1_1.
 - [24] He Y, Song K C, Meng Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. *IEEE Trans Instrum Meas*, 2020, **69**(4): 1493–1504.
 - [25] Lion C. Surface defect detection for aluminum sheets in industrial applications[EB/OL]. Baidu AI Studio, 2022. [2024-09-25]. <https://aistudio.baidu.com/datasetdetail/135640>.
Lion C. 铝片表面工业缺陷目标检测[EB/OL]. Baidu AI Studio, 2022. [2024-09-25].
<https://aistudio.baidu.com/datasetdetail/135640>.
 - [26] Everingham M, Eslami S M A, Van Gool L, et al. The PASCAL visual object classes challenge: a retrospective[J]. *Int J Comput Vision*, 2015, **111**(1): 98–136.
 - [27] Hu Y L. Dataset of surface defects of strip alloy functional materials[EB/OL]. Baidu AI Studio, 2024. [2024-10-22].
<https://aistudio.baidu.com/datasetdetail/299785>.
Hu Y L. 带状合金功能材料表面缺陷数据集[EB/OL]. Baidu AI Studio, 2024. [2024-10-22].
<https://aistudio.baidu.com/datasetdetail/299785>.
 - [28] Yi Z H, Jiang C H, Chen X F, et al. Optimal configuration method for field angle of industrial cameras based on imaging area model[J/OL]. *Control Theory Appl*, 2023: 1–9. [2024-11-29].
<http://kns.cnki.net/kcms/detail/44.1240.TP.20231214.0848.038>.

[html](#).

易遵辉, 蒋朝辉, 陈晓方, 等. 基于成像区域模型的工业相机视场角优化配置方法[J/OL]. 控制理论与应用, 2023: 1–9. [2024-11-29]. <http://kns.cnki.net/kcms/detail/44.1240.TP.20231214.0848.038.html>.

- [29] Shuai J X. The design and implementation of the digital system of a System-on-a-Chip CMOS image sensor[D]. Hunan: Hunan University, 2009.
- 帅金晓. 单芯片 CMOS 图像传感器数字系统的设计与实现[D]. 湖南: 湖南大学, 2009.
- [30] Redmon J, Farhadi A. YOLOv3: an incremental improvement[Z]. arXiv: 1804.02767, 2018. <https://arxiv.org/abs/1804.02767>.
- [31] Li C Y, Li L L, Jiag H L, et al. YOLOv6: a single-stage object

detection framework for industrial applications[Z]. arXiv: 2209.02976, 2022. <https://arxiv.org/abs/2209.02976>.

- [32] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 7464–7475. <https://doi.org/10.1109/CVPR52729.2023.00721>.
- [33] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[Z]. arXiv: 2107.08430, 2021. <https://arxiv.org/abs/2107.08430>.
- [34] Wang A, Chen H, Liu L H, et al. YOLOv10: real-time end-to-end object detection[Z]. arXiv: 2405.14458, 2024. <https://arxiv.org/abs/2405.14458>.

作者简介



胡依伦 (2001-), 男, 硕士研究生, 研究方向为机器视觉、深度学习理论及应用。

E-mail: 2023220603029@mails.zstu.edu.cn



杨俊 (1978-), 男, 工学博士, 嘉兴大学信息科学与工程学院副教授, 研究方向为智能多媒体信息分析与处理、机器学习, 以及计算机视觉。

E-mail: juneryoung@zjxu.edu.cn



许聪源 (1990-), 男, 工学博士, 嘉兴大学信息科学与工程学院讲师, 研究方向为机器视觉、网络空间安全。

E-mail: cyxu@zjxu.edu.cn



夏亚金 (1972-), 男, 本科, 海盐中达金属电子材料有限公司高级工程师, 研究方向为金属材料冶炼、金属材料质量分析与控制, 以及机器学习等。

E-mail: 826018252@qq.com



【通信作者】邓文斌 (1985-), 男, 工学博士, 嘉兴大学信息科学与工程学院讲师, 研究方向为智能装备开发、机器视觉。

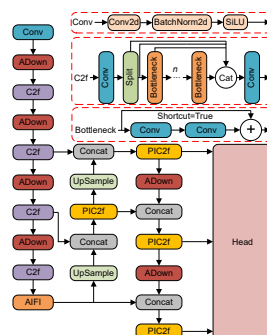
E-mail: wbdeng@zjxu.edu.cn



扫描二维码, 获取PDF全文

PIC2f-YOLO: a lightweight method for the detection of metal surface defects

Hu Yilun^{1,2}, Yang Jun², Xu Congyuan², Xia Yajin³, Deng Wenbin^{2*}



PIC2f-YOLO structure

Overview: Metallic materials, especially steel, are widely used in industry due to their mature manufacturing processes, excellent durability, and significant economic benefits. As a core material in various manufacturing sectors, steel is critical in advancing global industrialization and information technology. In the early stages of industrial metal production, surface defect detection relied primarily on manual visual inspections. However, as production scales expanded and quality demands increased, the limitations of manual inspection methods, such as inefficiency, susceptibility to human error, and high subjectivity, became evident. Thus, quickly and reliably identifying surface defects on metal has become a critical challenge in modern manufacturing processes. Traditional nondestructive testing techniques, such as eddy current testing and infrared inspection, were initially applied to surface defect detection on metal materials. However, these techniques often struggle to detect small defects effectively. Later, machine learning methods were introduced to improve detection accuracy, but limitations in processing power and generalization hindered their performance in large-scale industrial applications. With the rapid development of computing hardware, deep-learning-based methods have become the primary solution for metal surface defect detection. These methods are usually divided into two-stage and one-stage object detection frameworks. Two-stage methods, such as Faster RCNN, are recognized for their high accuracy but have slower processing speeds, limiting their real-time applications. Conversely, one-stage methods, such as YOLO, provide faster detection speeds but often compromise accuracy. To address the low efficiency in metal surface defect detection, as well as problems related to large algorithm parameters and low precision, this paper proposes a lightweight detection method based on the improved YOLOv8n. First, the local convolution inverted cross-stage partial fusion (PIC2f) module was designed. It replaces the Bottleneck module with the constructed local convolution inverted residual bottleneck (PIBN) module, which combines partial convolution and an inverted residual block to reduce algorithm parameters and enhance the model's feature extraction ability. Next, an attention-based intra-scale feature interaction (AIFI) module was adopted. It combines position embedding and multi-head attention to enhance the model's detection capability for small targets. Lastly, average pooling downsampling (ADown) replaced traditional convolution as the feature reduction module, effectively reducing parameters and computational complexity without compromising detection accuracy. Experimental results show that, compared with YOLOv8n, the PIC2f-YOLO method increases mAP₅₀ by 2.7 % on the NEU-DET steel defect dataset and reduces parameters by 0.403 M. Experiments on aluminum sheet surface industrial defects, PASCAL VOC2012, and surface defects of strip alloy functional material datasets also confirm the effectiveness of the PIC2f-YOLO method.

Hu Y L, Yang J, Xu C Y, et al. PIC2f-YOLO: a lightweight method for the detection of metal surface defects[J]. *Opto-Electron Eng*, 2025, 52(1): 240250; DOI: 10.12086/oe.2025.240250

Foundation item: National Natural Science Foundation of China (62302197), the Zhejiang Natural Science Foundation Project (LQ23F020006), the Jiaxing City Science and Technology Project Fund (2024AD10045, 2024AY40010), and the Scientific Research Fund of Zhejiang Provincial Education Department (Y202455539)

¹College of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou, Zhejiang 310018, China; ²College of Information Science and Engineering, Jiaxing University, Jiaxing, Zhejiang 314001, China; ³Haiyan ZhongDA METAL Electronic Material Co., LTD, Jiaxing, Zhejiang 314300, China

* E-mail: wbdeng@zjxu.edu.cn