

# 融合共享参数的 YOLOv10n 钢材表面缺陷检测算法 \*

杨本臣 潘子睿 王春艳 金海波 李世熙

(辽宁工程技术大学软件学院 葫芦岛 125105)

**摘要:**针对钢材表面缺陷检测中的精度低、易受背景干扰的问题,提出一种融合共享参数的 YOLOv10n 目标检测算法。首先,骨干网络引入改进的 FasterNet 轻量网络和通道优先卷积注意力机制,以提升骨干网络对多维信息的表征能力;其次,针对 C2f 模块感受野差的问题,基于部分卷积(PConv)设计了 PConv-C2f 模块;再次,采用小波池化,解决原算法中因上下采样机制引起的图像高频信息混叠和易受背景干扰问题;最后,通过共享参数与动态分布技术融合,提出一种轻量级检测头,以减少模型的计算复杂度并提高边界框预测的准确性。改进算法在 NEU-DET 数据集上的平均精度均值(mAP) mAP@0.5 达到 86.3%,较原算法提升 8.1%,精确率(precision)达到 86.8%,较原算法提高了 18.7%。通过消融、对比实验表明改进算法在钢材和金属材料表面缺陷检测中均具有较好的性能表现,不仅满足了实际应用中对钢材表面缺陷进行高效、准确检测的需求,还显著提升了检测的可靠性和实用性。

**关键词:**钢材缺陷检测;YOLOv10n;轻量级检测头;PConv-C2f;FasterNet;注意力机制

**中图分类号:** TP391.41; TN911.73      **文献标识码:** A      **国家标准学科分类代码:** 520.2060

## Fusion of YOLOv10n steel surface defect detection algorithm with shared parameters

Yang Benchen Pan Zirui Wang Chunyan Jin Haibo Li Shixi

(School of Software, Liaoning Technical University, Huludao 125105, China)

**Abstract:** With an aim to address the issues of low precision and susceptibility to background interference in steel surface defect detection, a YOLOv10n target detection algorithm based on fusion and shared parameters is proposed. Firstly, the backbone network incorporates the enhanced FasterNet lightweight network and the channel-first convolutional attention mechanism to enhance the capacity of the backbone network in representing multidimensional information. Secondly, the PCONV-C2F module is designed based on partial convolution (PConv) to tackle the problem of the disparity in the sensitivity field of the C2f module. Thirdly, wavelet pooling is utilized to address the problem of aliasing and background interference resulting from the up and down sampling mechanism in the original algorithm. Finally, a lightweight detection head is put forward to reduce the computational complexity of the model and enhance the accuracy of bounding box prediction by integrating shared parameters with dynamic distribution techniques. The mean average precision (mAP) mAP@0.5 of the improved algorithm on the NEU-DET dataset attains 86.3%, which is 8.1% higher than that of the original algorithm, and the precision reaches 86.8%, which is 18.7% higher than that of the original algorithm. The ablation and comparison experiments demonstrate that the improved algorithm exhibits excellent performance in the surface defect detection of steel and metal materials, which not only meets the requirement for efficient and accurate detection of steel surface defects in practical applications, but also significantly enhances the reliability and practicability of the detection.

**Keywords:** steel defect detection; YOLOv10n; lightweight detection head; PConv-C2f; FasterNet; attention mechanism

## 0 引言

随着钢铁工业的不断进步,钢铁作为国民经济的重要支柱,对于国家的工业化和现代化进程起到了关键作用。在钢铁材料制造及应用阶段,由于轧制加工与焊接处理等环节的工艺参数调控失当,金属基体在热循环应力、机械摩擦作用及环境氧化效应多重因素影响下<sup>[1]</sup>,易引发表面裂纹、划痕及氧化层形成等质量问题。这些微观缺陷不仅会显著降低钢材的力学性能,也对服役可靠性和安全性构成潜在威胁。现行基于人工目视检验与出钢端传感器监测的技术方案存在明显局限性,且在高温、粉尘等复杂工况条件下监测装置易发生功能性失效<sup>[2]</sup>。

深度学习为目标检测带来了巨大的突破,尤其在复杂背景下识别和检测多种类别目标的能力上表现优异。通用目标检测算法主要分为两类,一类是以 R-CNN<sup>[3]</sup>为基础的两阶段目标检测算法。Ren 等<sup>[4]</sup>在 R-CNN 算法基础上加入特征图共享、端到端训练的理念提出 Fast R-CNN 算法,实现了在卷积特征图上直接分类和回归,显著提高了检测速度。

钢材缺陷检测作为目标检测领域的关键应用之一,当前面临着缺陷小而形态多样、易受背景噪声和光照变化影响等挑战,使得检测任务尤为困难<sup>[5]</sup>。为了克服这些难题,以 YOLO 系列、SSD<sup>[6]</sup>、RetinaNet 为代表的单阶段目标检测算法应需而生。Kou 等<sup>[7]</sup>在 YOLOv3 的基础上引入了无锚点特征选择机制和密集卷积块,使得模型在钢材缺陷检测任务中平均精度均值 (mAP) 达到了 72.2%,显著提高了检测精度。为了提升钢材表面缺陷检测的准确性,Cheng 等<sup>[8]</sup>提出了一种差异通道注意和自适应空间特征融合的 RetinaNet 算法 DEA\_Retinanet,用于钢铁表面缺陷检测,在 NEU-DET 数据集上,该算法的 mAP 达到了 78.25%,比原始的 RetinaNet 提高了 2.92%。Li 等<sup>[9]</sup>将注意力机制模块引入 YOLOv4 的主干网络中,有效提高网络模型的特征提取能力,为日常钢带表面缺陷的检测提供了一种新的方法。在保持检测精度的基础上为减少模型计算量,Cheng 等<sup>[10]</sup>提出了一种适用于嵌入式设备的轻量型 YOLOv5 缺陷检测算法,模型中大量引入深度可分离卷积,最终在降低参数的同时提升了检测精度。为实现对不同缺陷类别的有效识别,Zhang 等<sup>[11]</sup>结合部分卷积、可切换空洞卷积和数据增强技术对 YOLOv7 进行改进,提出了 SSG-YOLOv7 算法。随着双向跨尺度连接和加权特征融合的提出,Wang 等<sup>[12]</sup>通过集成 BiFPN 概念并引入了动态大卷积核注意力机制 LSK-attention 重建了 YOLOv8s 的颈部结构,有效提高了模型对小目标的特征提取能力。

YOLOv10<sup>[13]</sup>在继承 YOLOv8 实时性能的基础上,通

过引入更复杂的网络结构和强化训练策略,显著提升了模型的检测精度和鲁棒性。自 YOLOv10 提出之后,其在缺陷检测领域受到了广泛关注,已开展了大量相关研究。Tian 等<sup>[14]</sup>提出了增强型 YOLOv10 框架的端到端检测模型,这一小目标检测算法专门设计用于快速、精确地检测大片水域中的死亡鱼类,然而,该检测算法并不适用于钢材缺陷检测的场景。YOLOv10n 作为 YOLOv10 的轻量化版本,在进一步减少模型参数量和计算量的同时,仍保持了较高的检测精度,非常适用于工业相机和 5G 模组等实际应用。然而,对于钢材表面缺陷检测需满足至少 64 fps 的速度和 85% 以上的精度标准,而以上算法尚未达到这些要求。

当前,YOLO 系列在目标检测领域的网络架构日趋复杂,但在小目标检测方面,性能提升并不显著。相反,由于网络结构的复杂性增加,模型可能会因过度学习无关信息而损害检测效果。本文针对钢材表面缺陷检测时易受背景影响出现误检、漏检的情况,提出改进 YOLOv10n 的目标检测算法。本文在 Backbone 部分首先引用改进的 FasterNet 轻量网络,其次引用通道优先卷积注意力机制,增强骨干网络在复杂场景中的细节特征识别能力。在 Neck 部分设计了 PConv-C2f 模块,此改进有效扩大了模型的感受野,尤其在检测微小或模糊缺陷时表现更为出色,并采用小波池化解决原算法中因上下采样机制引起的图像高频信息混叠和易受背景干扰问题。在 Head 部分提出一种轻量级检测头 (detect-efficient, DE),共享参数策略不仅减少了检测头的参数量,在降低存储和计算成本,提升推理效率的同时,还提升了模型在多尺度目标检测中的精准度。

## 1 YOLOv10 算法

YOLOv10 算法是一种高效的目标检测模型,旨在各种复杂场景中实现快速而准确的目标检测。YOLOv10 由 Backbone、Neck 和 Head 3 个部分组成。YOLOv10 算法结构如图 1 所示。

YOLOv10n 在前一代的基础上进行了优化,集成了多种先进技术和模块,以适应不断增长的检测需求。该模型主要由 Backbone、Neck 和 Head3 个部分组成,这一结构设计旨在提升目标检测的性能与效率。

在 Backbone 部分引入了 SCDown 和 C2fCIB 模块提出了改进的 CSPNet (cross stage partial network)<sup>[15]</sup> 结构。通过 SCDown 模块将空间维度的下采样和通道维度的下采样解耦,在降低计算成本的同时,最大限度地保留了特征信息。C2fCIB 模块采用紧凑的倒置块结构,利用高效的深度卷积进行空间信息的混合,并通过低成本的点卷积完成通道间的融合操作,优化了特征提取和融合过程,

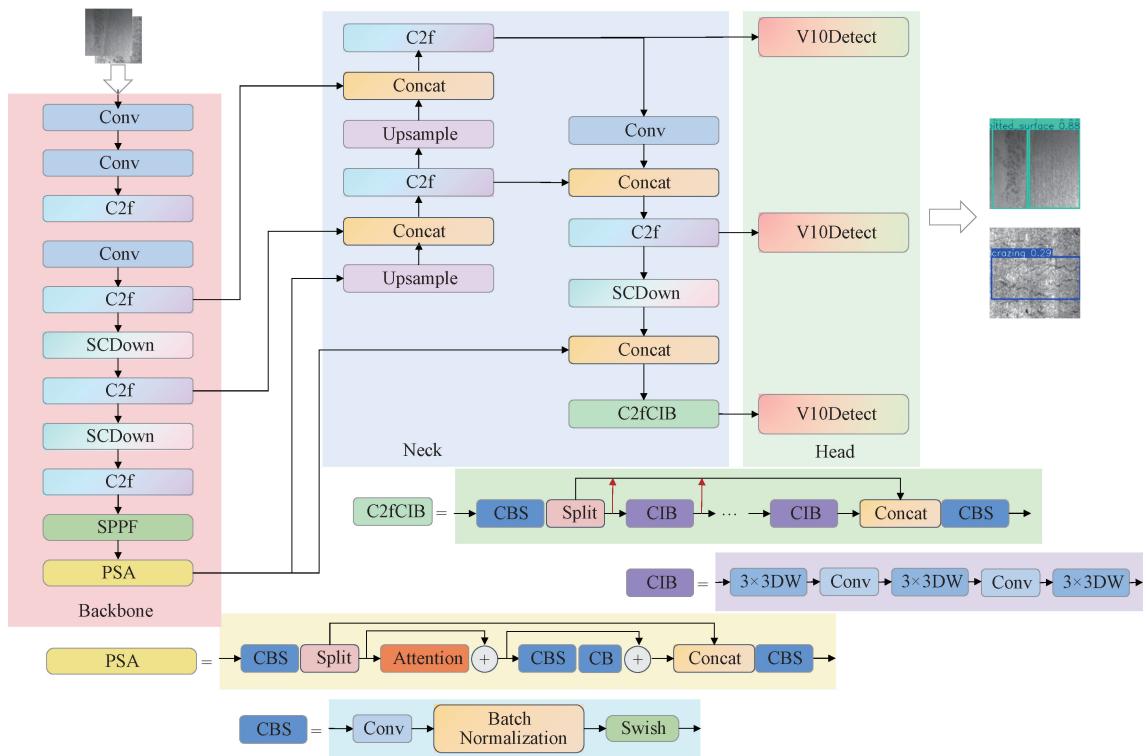


图 1 YOLOv10 算法结构

Fig. 1 YOLOv10 algorithm structure diagram

提升了模型的效率和性能。此外,还额外引入了 PSA (Pyramid Squeeze Attention) 模块,有效捕捉图像的上下文信息,从而提高了在复杂场景下的检测精度,显著增强了特征提取能力;Neck 部分采用了特征金字塔结构,能够有效整合来自不同层的特征图,从而捕捉更丰富的上下文信息;Head 部分使用交并比(IoU)损失函数来优化检测框的精确性,使得模型在处理重叠目标时表现更为优越。

## 2 改进的 YOLOv10n 算法

融合共享参数的 YOLOv10n 目标检测算法主要做出以下改进。首先,在 Backbone 部分引用改进的 FasterNet 轻量网络和通道优先卷积注意力机制;其次,在 Neck 部分设计 PConv-C2f 模并采用小波池化;最后,在 Head 部分提出一种轻量级检测头。改进后的 YOLOv10n 算法结构如图 2 所示,其中红色边框标注的部分为本模型的改进部分。

### 2.1 轻量级检测头

轻量级检测头主要采用了创新的参数共享设计与动态分布 Focal Loss (DFL) 技术,旨在提升模型在处理不平衡数据集时的适应性和性能表现。轻量级检测头的 3 个关键模块,如图 3 所示。

#### 1) 共享参数

回归与分类任务均通过采用轻量级检测头中的 2 个  $3 \times 3$  共享卷积层实现多尺度感知,有效降低了模型的复杂度,减少了计算量和存储需求,同时避免了冗余计算。与 V10Detect 相比,这种设计使模型参数量减少约 50%,降低了训练和推理时的内存占用。共享参数设计如式(1)所示。

$$stem_i = Sequential( Conv(C_i, C_i, 3), Conv(C_i, C_i, 3) ) \quad (1)$$

式中:  $stem$  为共享参数设计模型;  $Sequential$  表示将 2 个具有  $3 \times 3$  卷积核进行线性堆叠;  $C_i$  表示输入和输出的通道数相同。

#### 2) 回归和分类分支设计

该部分继承了 YOLOv10 的 v10detect 检测头的设计精髓,分别配置了回归和分类两个分支,以实现目标边界框和类别预测的独立优化。在回归分支中,模型将回归最大值  $reg\_max$  的 4 倍作为输出通道数,并通过  $1 \times 1$  卷积层进行坐标回归,如式(2)所示。分类分支则采用  $nc$  输出,其卷积操作如式(3)所示。

$$cv2_i = Conv2d(C_i, 4 \times reg\_max, 1) \quad (2)$$

式中:  $Conv2d(C_i, 4 \times reg\_max, 1)$  表示  $1 \times 1$  卷积层,适用于边界框的坐标回归。

$$cv3_i = Conv2d(C_i, nc, 1) \quad (3)$$

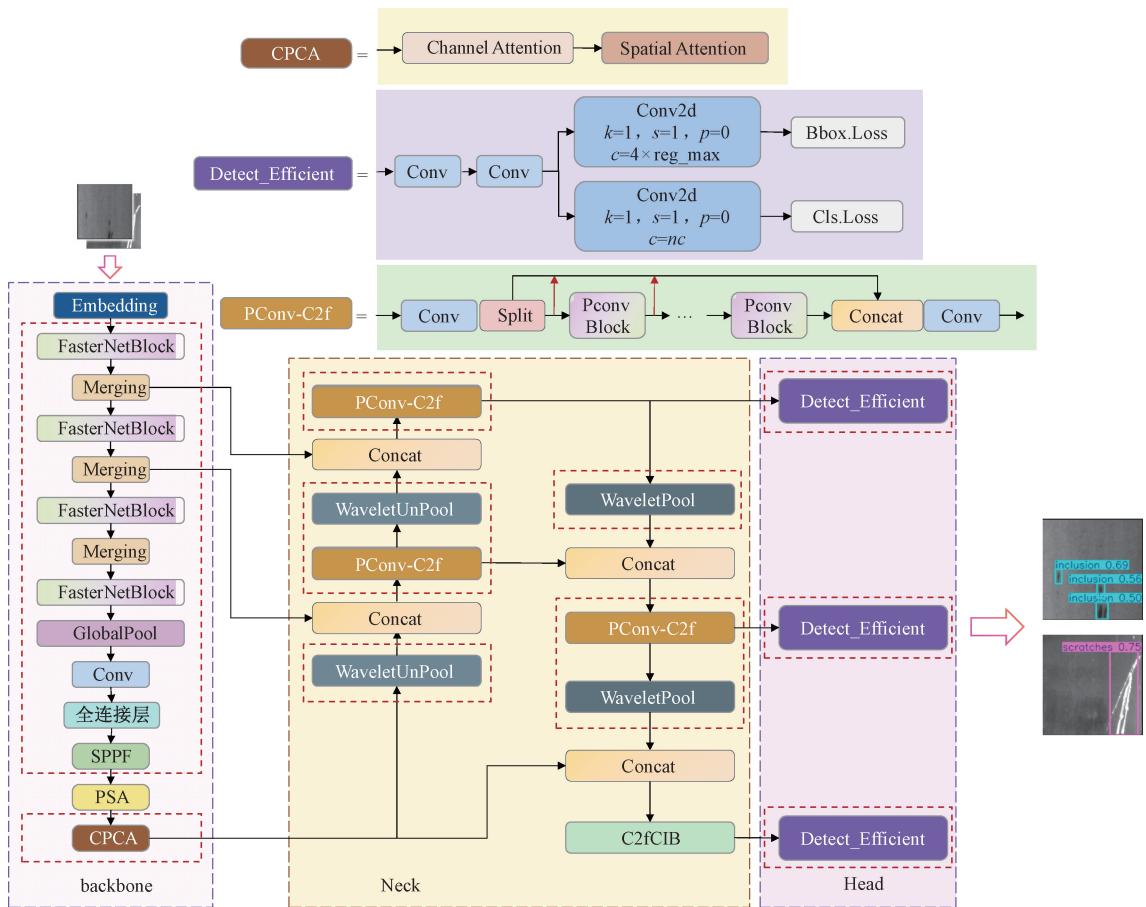


图2 改进后的 YOLOv10n 网络结构

Fig. 2 Improved YOLOv10n network structure diagram

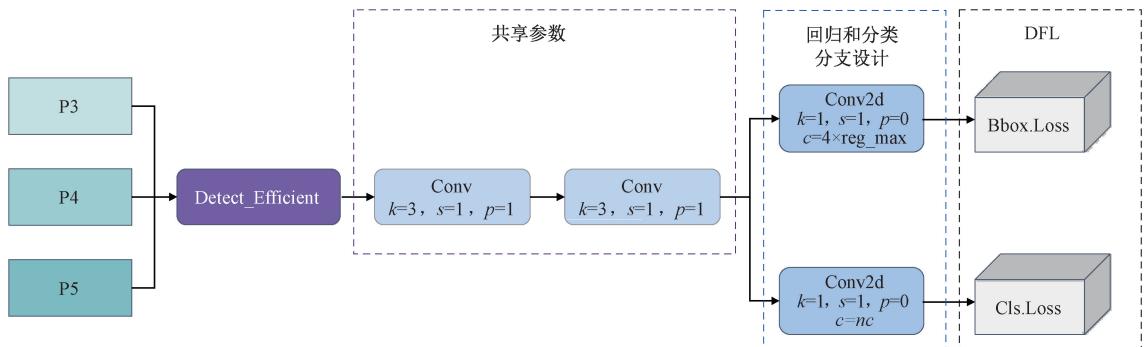


图3 轻量级检测头结构

Fig. 3 Detect\_Efficient structure diagram

式中:  $nc$  为类别数量。此设计使得检测头能够独立处理分类和回归任务,减少特征冲突。

### 3) DFL

由于小目标像素占比低且信息较少,类别分布不均衡,导致模型易出现误检漏检,尤其是在类别分布不均衡的情况下。为此设计 DFL 模块,如式(4)所示。

$$DFL = \begin{cases} DFL(\text{reg\_max}), & \text{reg\_max} > 1 \\ \text{Identity}, & \text{其他} \end{cases} \quad (4)$$

式中:  $\text{reg\_max}$  表示 DFL 的通道数;  $\text{Identity}$  表示恒等函数。当通道数大于 1 时,启用 DFL 层以提升回归效果。

设计的 DFL 模块,在计算损失时根据类别的分布情况自适应调整各类目标的损失分配,对难检测的类别赋予更高的权重,从而使模型对小目标的特征更加敏感,提

高边界框预测的准确性。设计方式即使在复杂背景或目标密集的场景中,也能保持对小目标较高的检测精度。

通过参数共享设计,轻量级检测头能够有效减少模型的参数数量,从而减轻计算负担,提高模型的训练效率。DFL技术使得模型在面对样本不均衡的情况时,能够更加关注那些稀有类别的样本,从而增强了对小类别的检测能力和准确性。

## 2.2 PConv-C2f 模块

为了解决传统神经网络在速度和准确性之间难以平衡的问题,在 Neck 部分设计了 PConv-C2f 模块,结合 PConv 设计的 PConv Block 模块替换了 C2f 模块中的瓶颈残差模块 (Bottleneck)。PConv 仅对输入通道  $X$  中的  $1/4$  部分  $X_1$  进行卷积操作,剩余  $3/4$  的通道  $X_2$  保持不变,如式(5)、(6)所示,这样不仅保留通道信息,还有效减少冗余计算。再将  $Y_1$  和  $X_2$  进行通道维度的拼接,得到拼接后的特征图  $Y_2$ ,如式(7)所示。

$$X = X_1 + X_2 \quad (5)$$

$$Y_1 = Conv3 \times 3(X_1) \in R^{H \times W \times \frac{C_{out}}{4}} \quad (6)$$

$$Y_2 = \text{Concat}(Y_1, X_2) \in \mathcal{R}^{H \times W \times C_{out}} \quad (7)$$

其中,  $X \in \mathbf{R}^{H \times W \times C_{in}}$ ,  $X_1 \in \mathbf{R}^{H \times W \times \frac{C_{in}}{4}}$ ,  $X_2 \in \mathbf{R}^{H \times W \times \frac{3C_{in}}{4}}$ ;  $h$ ,  $w$  为特征图的高和宽;  $k$  为卷积核大小。

此外, PConv 输出后的特征图  $Y_2$  通过  $1 \times 1$  卷积扩展至原通道数的两倍, 并结合 PConv 层未处理的通道, 进一步减少通道信息的损失。最终, 采用了  $1 \times 1$  卷积来替换原算法中的  $3 \times 3$  卷积, 将特征图的通道数恢复至原始规模, 确保了捷径分支与骨干路径的输出维度一致, 并有效避免了信息的丢失。Bottleneck 与 PConv Block 结构对比如图 4 所示。

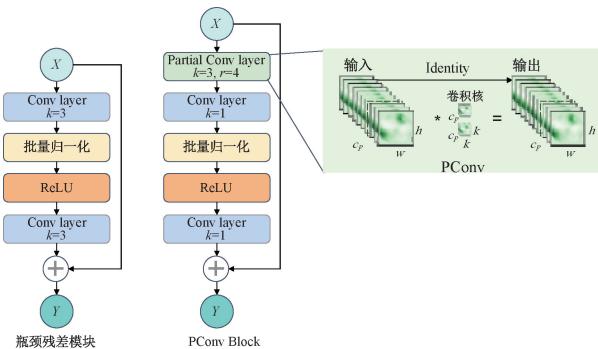


图 4 Bottleneck 与 PConv Block 结构图对比

Fig. 4 Bottleneck and PConv Block structure chart are compared

PConv 仅选择首端或末端连续的  $c_p$  通道进行特征图计算,这种设计不仅减少了计算复杂度,还降低了内存访问量。假设输入特征图的通道数为  $c$ ,而选择的部分通道

数为  $c_p$ ,那么 PConv-C2f 模块的浮点运算次数 (FLOPs) 为  $h \times w \times k^2 \times c_p^2$ 。内存访问需求量为  $h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p$ 。当 PConv 中参与计算的通道数的典型比例为  $r = c_p/c = 1/4$ , PConv 的 FLOPs 仅为常规卷积的  $1/16$ , 同时内存访问量也减少到常规卷积的  $1/4$ 。

PConv-C2f 模块相较于 C2f 模块在提升检测精度的同时,大幅降低了模型的计算开销,并扩大模型的感受野。

### 2.3 小波池化采样

为解决小目标检测中常见的混叠问题以及小目标边缘信息丢失问题。Neck 部分引用小波池化<sup>[16]</sup> (Wavelet Pooling) 替换原有普通上下采样, 小波池化依赖于小波变换定理<sup>[17]</sup>, 并使用离散小波变换 (DWT) 和反离散小波变换 (IDWT) 实现。

Wavelet Pooling 使用基于二维小波变换的小波池化方法, 小波池化是将离散小波变换应用于输入  $X$ , 将反离散小波变换仅应用于低频分量, 以构建下采样输出  $X^*$ 。池化过程如图 5 所示。

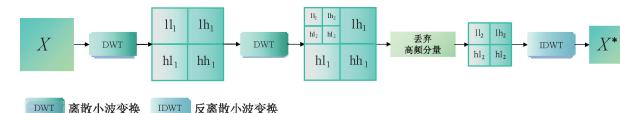


图 5 小波池化结构

Fig. 5 Wavelet pooling structure diagram

改进算法中引入的小波池化技术在下采样过程中，通过小波分解保留低频信息并过滤高频信息，从而增强细节还原和抗噪能力。

## 2.4 Fasternet 骨干网络

钢材表面的微小缺陷在特征提取时容易被忽略, 同时可能引入大量背景噪声或非缺陷特征。此外, 随着神经网络深度的提升, 特征图通道数的增加导致了冗余信息的累积, 这不仅拖慢了模型的检测速度, 还可能降低检测的可靠性。为了解决这一问题, Chen 等<sup>[18]</sup>在提出的 FasterNet 网络中通过引入一种部分卷积, 有效减少冗余信息, 并提高小目标检测中的特征提取效率。FasterNet 网络结构如图 6 所示。

本文模型在 Backbone 部分引入了改进的 FasterNet 网络, 将嵌入维度从 192 降至 64, 路径丢弃率调整为 2%。这一调整显著减少了参数量和计算资源消耗, 尤其在检测钢材表面小缺陷时, 可提高特征提取效率。同时, 适度的路径丢弃率有效降低了过拟合风险, 避免了高丢弃率带来的训练不稳定。

## 2.5 通道优先卷积注意力机制

为优化特征通道间的信息交互,以增强模型的特征识别能力。在 Backbone 中引入通道优先卷积注意力模

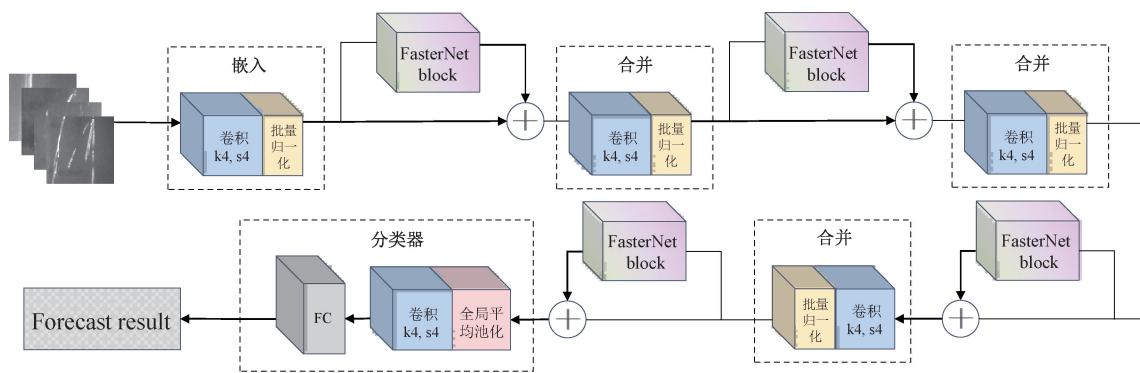


图 6 Fasternet 网络结构

Fig. 6 Fasternet network structure diagram

块(CPCA)<sup>[19]</sup>。CPCA 通过全局聚合和权重学习聚焦关键特征,来提高检测准确性。它通过重标定输入特征优化小目标识别,并结合通道和空间注意力提升模型对显

著特征的捕捉能力。通道优先卷积注意力机制整体结构如图 7 所示。

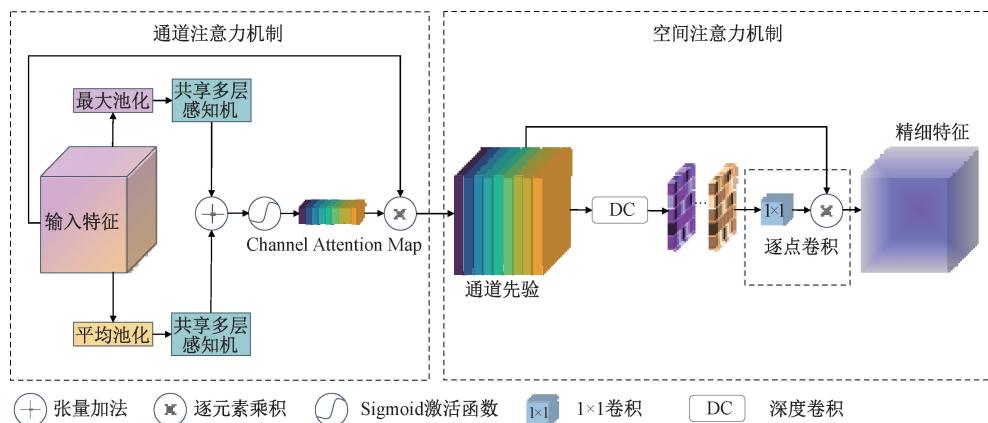


图 7 通道优先卷积注意力机制模块

Fig. 7 Channel priority convolution attention mechanism module

对比标准卷积,标准卷积的计算复杂度是  $D_k^2 \times M \times N \times D_f \times D_f$ ,其中  $D_k$  为卷积核大小,  $M$  为输入通道数,  $N$  为输出通道数,  $D_f$  为特征图的空间维度。而深度可分离卷积的计算复杂度是  $D_k^2 \times M \times D_f \times D_f + M \times N \times D_f \times D_f$ 。通过这种分解,计算量减少了约  $\frac{1}{N} + \frac{1}{D_k^2}$  倍,对于常见的  $3 \times 3$  卷积,计算量减少接近 9 倍,降低了模型计算量。

### 3 实验结果与分析

为验证改进算法的有效性,本文利用东北大学开源的 NEU-DET 钢材表面缺陷数据集进行了训练和测试。

#### 3.1 数据集

该数据集包含冶金工业中 6 类典型表面缺陷,具体

涵盖表面裂纹(crazing, Cr)、非金属夹杂(inclusion, In)、氧化斑块(patches, Pa)、点蚀麻面(pitted\_surface, Ps)、轧制氧化皮压入(rolled-in\_scale, Rs)以及机械划痕(scratches, Sc),每类缺陷包含 300 幅  $200 \times 200$  pixels 分辨率的工业检测图像,样本总量达 1 800 幅。数据集按 8 : 1 : 1 比例划分训练集(1 440 幅)、验证集(180 幅)和测试集(180 幅),确保模型评估的统计学有效性<sup>[20]</sup>。

各类缺陷如图 8 所示,具有显著形貌特征差异。

1) Cr 缺陷通常呈现线性分布,有时会形成 Y 形结构,主要与轧制方向一致,偶尔也会出现在横向或其他方向。

2) In 缺陷通常在钢材表面表现为深度各异的非金属夹杂物,这些夹杂物多以点状、条状或块状的形式存在,颜色通常为暗红、暗黄或灰白<sup>[5]</sup>。

3) Pa 缺陷通常在钢材表面表现为片状或大面积分

布的斑块,某些角度下会显现出辐射状的特征。

4) Ps 缺陷为局部粗糙的凹凸表面,密度不均,严重时有橘皮般的纹理。

5) Rs 缺陷通常呈斑点、鱼鳞状、条状或块状图案,不规则分布于物体表面,深浅不一。

6) Sc 缺陷表面有纵向或横向的划痕,长度、深度和宽度不尽相同,颜色通常为灰白。

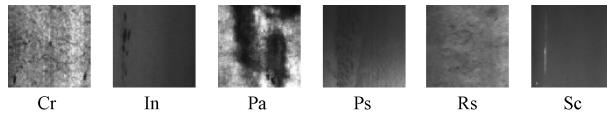


图 8 NEU-DET 数据集缺陷类型

Fig. 8 NEU-DET data set defect type

本文应用该数据集分别进行了消融实验与对比实验,以验证改进模块与改进算法的有效性。

### 3.2 实验环境与参数设置

实验系统搭建采用异构计算架构,硬件平台由以下组件构成:计算单元采用 NVIDIA RTX 2080 Ti GPU(显存容量 11 GB),中央处理器搭载 Intel Xeon Platinum 8255C 12 虚拟核心处理器(基准频率 2.50 GHz),系统内存 30 GB。软件栈构建于 Ubuntu 20.04 LTS 操作系统,编程环境采用 Python 3.8 解释器,深度学习框架选用 PyTorch 1.10.0 版本,并基于 CUDA 11.3 并行计算架构实现 GPU 加速。核心训练参数经网格搜索优化后确定:输入层接收 640×640 分辨率图像,初始学习率配置为  $1 \times 10^{-2}$ ,动量参数设置为 0.937 以平滑梯度更新,L2 正则化系数调整为  $5 \times 10^{-4}$ ,边界框交并比判定阈值保持 0.5,批次规模设为 16,模型经 200 轮迭代完成参数更新。该配置方案通过软硬件协同优化,在提升模型收敛效率与训练稳定性的同时,有效避免了显存溢出等常见计算问题。

### 3.3 实验评测标准

本文采用了 5 种评价指标来对模型进行评估,分别

是 mAP、精确率(precision, P),着重评估误检缺陷对检测系统的影响、召回率(recall, R),重点监测漏检缺陷带来的工业风险、帧率吞吐量(FPS),通过计算单 GPU 环境下处理 1 024×1 024 分辨率图像的推理速度,评估模型在工业在线检测系统中的部署可行性。目标检测中常用 mAP 来评价模型的性能,公式如下:

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (8)$$

$$AP = \int_0^1 p(R) dR \quad (9)$$

式中: $n$  为图像的种类; $i$  为检测次数; $AP$  为单一类别的识别平均准确率;本文设置  $n = 6$ <sup>[5]</sup>。本文使用 mAP@0.5,设 IoU 为 0.5,计算每类图片的  $AP$ ,再对所有种类取平均值<sup>[21]</sup>。

P 表示模型检测的准确程度,在被分类为正例的样本中有多少是真实正例。R 表示正样本中模型检测的程度,在所有实际正例中,正确识别为正例的比例。公式如下:

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

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

式中: $TP$  为正检率; $FP$  为误检率; $FN$  为负样本被检测为正样本的个数<sup>[5]</sup>。

FPS 表示模型每秒检测图片数量,公式如下:

$$FPS = \text{Framenum} / \text{Elapsed Time} \quad (12)$$

式中: $\text{Framenum}$  表示检测图片数量; $\text{Elapsed Time}$  表示模型检测的时间。

### 3.4 消融实验

为了验证所提出方法在钢材表面的检测性能,设计了一系列消融实验,以评估改进点的有效性。实验结果如表 1 所示。

表 1 改进点消融实验

Table 1 Improved point ablation experiments

实验编号	YOLOv10	Fasternet	CPCA	小波池化	PConv-C2f	Detect_Efficient	P/%	R/%	mAP@ 0.5/%
1	√	—	—	—	—	—	68.1	73.5	78.2
2	√	√	—	—	—	—	76	67.1	80.5
3	√	√	√	—	—	—	83.4	72.9	81.9
4	√	√	√	√	—	—	86.8	72.9	82.2
5	√	√	√	√	√	—	80.9	77.9	83.2
6	√	√	√	√	√	√	<b>86.8</b>	<b>84.2</b>	<b>86.3</b>

由表 1 可知,在 NEU-DET 数据集上改进模块在钢材表面缺陷检测中的性能均有显著提升。在骨干网络中引入改进的 FasterNet 网络后, mAP 由 78.2% 提升至

80.5%,P 从 68.1% 提高至 76%,这表明重建后的骨干网络有效增强了缺陷特征的提取能力。在嵌入通道优先卷积注意力机制后,mAP 进一步提升至 81.9%,P 提高至

83.4%，显示出该机制在提升模型对缺陷特征图像全局特征信息的提取能力方面的有效性。此外，采用小波池化替代上下采样后，mAP 由 81.9% 提升至 82.2%， $P$  从 83.4% 提高至 86.8%，表明小波池化更好地保留了高频信息，减少了信息丢失和背景干扰。替换 PConv-C2f 结构后，依托 Pconv 的优势，mAP 提升至 83.2%，显著增强了算法对钢材表面缺陷的检测能力， $R$  也得到明显提升。最后引入以共享参数为基础的轻量级检测头，通过优化特征交互和融合机制，使 mAP 提高 3.1%， $R$  提升 6.3%， $P$  提升 5.9%。这种设计减少了参数量，降低了复杂度，同时增强了分类与定位的协同作用，提升了对不同尺度

目标的检测能力，显著改善了检测性能。综上所述，本文提出的各种改进方法在钢材表面检测方面具有积极效应。

图 9 所示为改进算法对 6 种缺陷检测效果的对比情况。图 9 依次为原始图像、基于原始算法的检测结果以及改进算法的检测结果。通过观察图 9 可以发现，在裂纹、麻点、压入氧化铁皮和划痕的检测中，改进算法能够识别出原 YOLOv10n 未能检测到的缺陷。在杂质、斑块检测中，改进算法不仅检测更加准确，还能同时识别出斑块与杂质缺陷。可见，本研究中所提出的改进算法在针对不同类型的缺陷检测任务中，均呈现出了优异的检测性能。

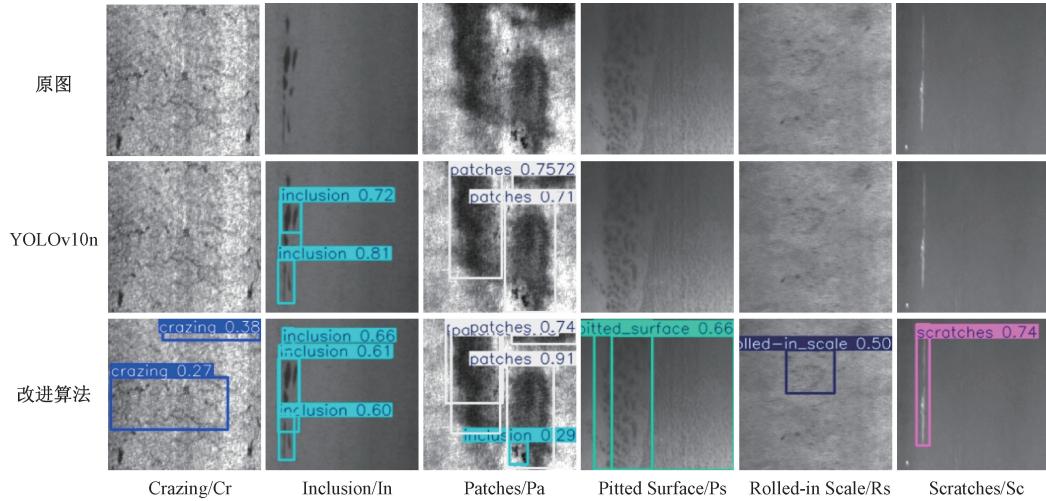


图 9 改进算法效果

Fig. 9 Improved algorithm renderings

### 3.5 不同数据集算法对比

基于多材料缺陷检测的对比研究需求，本实验选取 Aluminum\_ssD 与 GC10-DET 两个工业检测基准库进行跨域验证。其中，Aluminum\_ssD 数据集聚焦汽车轻量化部件制造缺陷，包含 1 400 幅高分辨率工业图像，涵盖 4 大类铝材表面异常：机械擦伤（表面磨痕深度  $>20 \mu\text{m}$ ）、塑性皱曲（冷作硬化区厚度变异系数  $\geq 15\%$ ）、表面污染物沉积（颗粒尺寸  $50\text{~}200 \mu\text{m}$ ）及微孔洞缺陷（孔径分布符合 Weibull 模型）<sup>[22]</sup>。GC10-DET 数据集则针对压力容器用钢表面质量检测，包含 2 292 幅符合 EN 10308 标准的检测图像，其 10 类缺陷包含月牙型冷隔缺陷（铸造工艺参数失当导致的熔融金属融合不良）、水渍氧化斑（相对湿度  $>60\%$  环境下形成的  $\text{FeO(OH)} \cdot \text{nH}_2\text{O}$  沉积）、油性介质残留（动态接触角  $>120^\circ$  的疏水性污染）、轧制腰折缺陷（与轧辊形变曲率呈正相关）等典型工业缺陷形貌。两数据集中的图像采集于光照充足、低干扰的结构化环境，符合工业场景下的检测工况<sup>[22]</sup>。实验结果如表 2、3 所示。

表 2 Aluminum\_ssD 金属缺陷数据集对比实验

Table 2 Comparative experiment of Aluminum\_ssD metal defect data set

算法	mAP@0.5/%	P/%	R/%	模型大小/MB
YOLOv10n	88.3	84.5	80.5	5.8
改进算法	<b>96.2</b>	<b>96.8</b>	<b>92.2</b>	<b>17.5</b>

基于表 2 的实验数据，针对 Aluminum\_ssD 金属缺陷数据集，改进算法的平均精确度提升了 7.9%。同时，Precision 和 Recall 也有显著提升。这反映出了改进算法在检测金属缺陷方面仍然具有良好的性能。

表 3 GC10-DET 钢材表面缺陷数据集

Table 3 GC10-DET steel surface defect data set

算法	mAP@0.5/%	P/%	R/%	模型大小/MB
YOLOv10n	80.2	73.5	71.5	5.8
改进算法	<b>88.4</b>	<b>86.6</b>	<b>80.2</b>	<b>17.5</b>

根据表 3 的实验结果，在 GC10-DET 钢材表面缺陷数据集上，本文改进的算法相较于原始算法尽管模型大

小有所增加,但平均精确度、精确率、召回率具有提高,尤其在加入了以共享参数为基础的轻量级检测头后平均精度得到大幅度提升。

实验结果表明,改进后的算法在多个不同的数据集上展现了卓越的适应性和鲁棒性。

### 3.6 现有算法对比实验

为了验证改进算法在钢材表面检测方面的性能,本文将其与现有目标检测算法在 NEU-DET 数据集上进行了对比分析。算法包括 R-CNN、Faster-RCNN、Retina-Net、YOLOv3、YOLOv4、YOLOv5s、YOLOv7-tiny、YOLOv8n 和 YOLOv10n,对比结果如表 4 所示。结果表明, R-CNN

和 Faster-RCNN 两阶段目标检测算法的 mAP 较低存在大量误检、漏检的情况。Retina-Net 作为单阶段目标检测代表更符合小目标检测的要求但在 Cr 这种易受背景干扰的缺陷时检测能力明显下降。YOLO 系列中 YOLOv8n 在多类别检测上表现稳定,整体 mAP@0.5 达到 81.3% 但仍未达到实际应用中钢材检测的要求。本研究提出的改进算法由于对 YOLOv10n 的 3 个部分均做出了优化,进一步提升了整体 mAP@0.5 到了 86.3%,尤其是 Pa 和 Ps 两类缺陷的检测精度,分别达到 97.6% 和 98.2%。这表明改进算法在检测精度和应用可靠性上展现了明显优势,更适合在工业场景中进行高效目标检测。

表 4 不同算法在 NEU-DET 测试集上的 AP 与 mAP 对比

Table 4 Comparison of AP and mAP of different algorithms on the NEU-DET test set

(%)

算法	mAP@0.5	AP					
		Cr	In	Pa	Ps	Rs	Sc
R-CNN	68.1	12.5	64.1	29.6	74	45	57.1
Faster-RCNN	77.4	48.2	71.7	77.2	86.6	67.4	83.5
Retina-Net	78.25	50.2	68.1	73.8	84.8	71.4	85.3
YOLOv3	75.3	59.8	64	89.7	93.2	60.9	84.3
YOLOv4	67.8	34.7	76.8	87.4	76.5	50.3	81.6
YOLOv5s	73.6	53.5	61.4	87.7	93.1	63.5	83.3
YOLOv7-tiny	66.4	39	86.6	79.2	66.8	55.9	71.3
YOLOv8n	81.3	79.8	74.6	90.3	94.5	63.8	84.7
YOLOv10n	78.2	60	72.7	85.2	96.7	65.4	89.0
改进算法	<b>86.3</b>	<b>76.2</b>	<b>80.5</b>	<b>97.6</b>	<b>98.2</b>	<b>75.2</b>	<b>89.3</b>

## 4 结 论

为解决钢材表面缺陷检测中精度低、易受背景干扰的难题,改进算法在 backbone 部分引入了优化后的 FasterNet 轻量网络和 CPCNA 注意力机制,Neck 部分设计了 PConv-C2f 模块并引入小波池化,Head 部分提出了以共享参数为基础的轻量级检测头。改进后的 YOLOv10n 算法使前景特征点能更全面地捕捉整个物体,而背景特征点集中于富含语义信息的区域,解决了深度增加带来的冗余信息和小目标检测中的混叠问题。改进后特征提取的有效性显著提升,在钢材表面缺陷检测方面,其检测精度和召回率均超越了当前的算法。未来研究将进一步优化模型参数量以提高推理速度,进一步轻量化以减少计算资源消耗、提升小样本数据下的泛化能力和复杂背景中细微缺陷的识别能力。

## 参考文献

- [1] LYU X, DUAN F, JIANG J, et al. Deep metallic surface defect detection: The new benchmark and detection network [J]. Sensors, 2020, 20(6): 1562.
- [2] LI Z, ZHANG J, ZHUANG T, et al. Metal surface defect detection based on MATLAB [C]. 2018 IEEE 3rd Advanced Information Technology, Electronic and

Automation Control Conference (IAEAC). IEEE, 2018: 2365-2371.

- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C]. Advances in Neural Information Processing Systems, NIPS, 2012.
- [4] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [5] 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,1-10 [2024-10-22]. <https://doi.org/10.13433/j.cnki.1003-8728.20240135>.
- YANG B CH, LI SH X, LI Y Z, et al. Improve YOLOv8n steel surface defect detection algorithm [J/OL]. Mechanical science and technology, 1-10 [2024-10-22]. <https://doi.org/10.13433/j.cnki.1003-8728.20240135>.
- [6] 阎馨,杨月川,屠乃威.基于改进 SSD 的钢材表面缺陷检测[J].现代制造工程,2023, 512(5): 112.
- YAN X, YANG Y CH, TU N W. Steel surface defect detection based on improved SSD [J]. Modern Manufacturing Engineering, 2023, 512(5): 112.

- [ 7 ] KOU X, LIU S, CHENG K, et al. Development of a YOLO-V3-based model for detecting defects on steel strip surface [ J ]. Measurement, 2021, 182: 109454.
- [ 8 ] CHENG X, YU J B. RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection [ J ]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [ 9 ] LI M, WANG H, WAN Z. Surface defect detection of steel strips based on improved YOLOv4 [ J ]. Computers and Electrical Engineering, 2022, 102: 108208.
- [ 10 ] CHENG S, YANG H G, XU X Q, et al. Improved lightweight X-ray aluminum alloy weld defects detection algorithm based on YOLOv5 [ J ]. Chinese Journal of Lasers, 2022, 49(21): 2104005.
- [ 11 ] ZHANG J, YANG W, CHEN Y, et al. Fast object detection of anomaly photovoltaic (PV) cells using deep neural networks [ J ]. Applied Energy, 2024, 372: 123759.
- [ 12 ] WANG X, GAO H, JIA Z, et al. BL-YOLOv8: An improved road defect detection model based on YOLOv8 [ J ]. Sensors, 2023, 23(20): 8361.
- [ 13 ] WANG A, CHEN H, LIU L, et al. YOLOv10: Real-time end-to-end object detection [ J ]. Arxiv preprint arXiv:2405.14458, 2024.
- [ 14 ] TIAN Q, HUO Y, YAO M, et al. A method for detecting dead fish on large water surfaces based on improved YOLOv10 [ J ]. ArXiv preprint arXiv: 2409.00388, 2024.
- [ 15 ] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [ 16 ] LI Q F, SHEN L L, GUO S, AND LAI Z H. WaveCNet: Wavelet integrated CNNs to suppress aliasing effect for noise-robust image classification [ C ]. IEEE Transactions on Image Processing, 2020: 7074-7089.
- [ 17 ] 钟明杉,李兆飞,张奕杰,等.基于一维卷积神经网络的天然气管道泄漏检测模型 [ J ].国外电子测量技术,2023,42(5):62-68.
- ZHONG M SH, LI ZH F, ZHANG Y J, et al. Based on one-dimensional convolutional neural network of natural gas pipeline leak detection model [ J ]. Foreign Electronic Measurement Technology, 2023, 42(5): 62-68.
- [ 18 ] CHEN J, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [ 19 ] HUANG H, CHEN Z, ZOU Y, et al. Channel prior convolutional attention for medical image segmentation [ J ]. Computers in Biology and Medicine, 2024, 178: 108784.
- [ 20 ] 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.
- XIONG C, YU AN N, GAO X H, et al. An improved YOLOX-based algorithm for surface defect detection of steel materials [ J ]. Electronic Measurement Technology, 2023, 46(9): 151-157.
- [ 21 ] 冷冰,冷敏,常智敏,等.基于Transformer结构的深度学习模型用于外周血白细胞检测 [ J ].仪器仪表学报,2023,44(5): 113-120.
- LENG B, LENG M, CHANG ZH M et al. Deep learning model based on Transformer structure for peripheral blood leukocyte detection [ J ]. Chinese Journal of Scientific Instrument, 2023, 44 (5): 113-120.
- [ 22 ] 张恒,程成,袁彪,等.基于YOLOv5-EA-FPNs的芯片缺陷检测方法研究 [ J ].电子测量与仪器学报,2023,37(5): 36-45.
- ZHANG H, CHENG CH, YUAN B, et al. Research on chip defect detection method based on YOLOv5-EA-FPNs [ J ]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 36-45.

## 作者简介



杨本臣,分别在1998年、2004年、2013年于辽宁工程技术大学获得学士学位、硕士学位和博士学位,现为辽宁工程技术大学副教授,主要研究方向为大数据、智能数据处理。

E-mail: 297058057@qq.com

**Yang Benchen**, received his B. Sc. degree, M. Sc. degree, and Ph. D. degree from Liaoning Technical University in 1998, 2004, and 2013, respectively. He is now an associate professor at Liaoning Technical University. His main research interests include big data and intelligent data processing.



潘子睿(通信作者),2022年于辽宁工程技术大学获得学士学位,现为辽宁工程技术大学硕士研究生,主要研究方向为图像处理、缺陷检测。

E-mail: 3196300455@qq.com

**Pan Zirui** ( Corresponding author ), received her B. Sc. degree from Liaoning Technical University in 2022. She is now a M. Sc. candidate at Liaoning Technical University. Her main research interests include image processing and defect detection.