

改进基于 YOLOv8n 的轻量化钢材表面缺陷检测算法^{*}

刚 帅 刘培胜 郭希旺

(辽宁石油化工大学人工智能与软件学院 抚顺 113001)

摘要:为了解决钢材表面缺陷检测模型参数量大、计算复杂度高以及对运算平台资源要求高的问题,提出了一种轻量化的改进算法。首先,使用ShuffleNetV2作为改进后的主干层,在降低模型复杂性和计算量上具有显著效果;其次,在SPPF模块后加入足够灵活和轻量的通道注意力机制(CA),同时使用双向特征金字塔网络(BiFPN)改善特征融合,提高了特征信息流动效率;最后,使用轻量级双卷积核(DualConv)替换C2f中的卷积层,通过分组卷积策略实现参数量的减少。实验结果表明,改进后的模型相比于原始的YOLOv8n,在保持检测精度的前提下,实现了轻量化。参数量为原来的56.2%,体积和计算量分别降至3.6 MB和4.8 GFLOPs,相比原模型分别降低了42.86%和41.47%,模型的轻量化降低了部署成本,适合实际部署和应用。

关键词:YOLOv8n;钢材缺陷;轻量化;缺陷检测;注意力机制

中图分类号:TP391.9;TN911 文献标识码:A 国家标准学科分类代码:520.6040

Lightweight improved YOLOv8n model for steel defect detection features

Gang Shuai Liu Peisheng Guo Xi旺

(School of Artificial Intelligence and Software, Liaoning Petrochemical University, Fushun 113001, China)

Abstract: To address the issues of large parameter quantity, high computational complexity, and high resource demands on the computing platform for the steel surface defect detection model, a lightweight improved algorithm has been proposed. Firstly, using ShuffleNetV2 as the improved backbone layer has achieved remarkable results in reducing model complexity and computational load. Secondly, a sufficiently flexible and lightweight channel attention mechanism (CA) was incorporated after the SPPF module, while the bidirectional feature pyramid network (BiFPN) was utilized to enhance feature fusion, thereby improving the efficiency of feature information flow. Finally, the lightweight dual convolution kernel (DualConv) was employed to replace the convolution layer in C2f, and the parameter quantity was reduced through the grouping convolution strategy. Experimental results indicate that, compared with the original YOLOv8n, the improved model achieves lightweighting while maintaining detection accuracy. The parameter quantity is 56.2% of the original, and the volume and computational load have decreased to 3.6 MB and 4.8 GFLOPs, respectively, representing a reduction of 42.86% and 41.47% compared to the original model. The lightweighting of the model reduces the deployment cost and is suitable for practical deployment and application.

Keywords: YOLOv8n; steel defect detection; lightweight; defect detection; attention mechanism

0 引言

在复杂的工业制造过程中,生产流程、设备维护和环境条件等因素可能导致金属表面出现划痕、裂缝和污点等缺陷^[1]。缺陷问题导致产品不同质,进而损害材料本身性能,增加企业开销成本。为解决这类实际问题,钢材缺陷检测显得非常关键,进行缺陷检测能够识别到钢材缺陷,降低缺

陷产品隐患,避免因钢材品质不合格而导致的后续生产问题。降低产品的废弃率,提高企业对生产成本的把控能力,提高产品质量,进行缺陷检测具有现实意义^[2]。

深度学习目标检测算法主要划分为双阶段检测和单阶段检测。双阶段目标检测算法首先对目标进行初步筛选,确定其大致的候选位置,在随后,对这些候选位置进行详细的分类和精确的定位调整。常见的双阶段目标检测算法有

R-CNN^[3]、Faster R-CNN^[4]。典型的单阶段目标检测算法包括 YOLO^[5]。朱成杰等^[6]提出了一种改进 YOLOv8 算法,在原损失 CIOU 的基础上面加入标准化高斯瓦瑟斯坦距离,替换 YOLOv8 模型的空间池化金字塔为聚焦调制,构建新模块 C2f-MB 增强特征表达能力,在主干部分加入卷积块注意力机制。改进后模型在基准模型的基础上提升了 3% 的检测精度。金鑫等^[7]针对煤矿电力设备缺陷检测精度低的问题,提出了一种基于改进 YOLOv5s 的煤矿电力设备缺陷检测的方法,该方法提出了一种多分支的坐标注意力模块来增强模型的获得缺陷信息区域能力,其次,通过修改主干网络和颈部网络之间非相邻的特征信息,进行跨层连接来增强了模型的特征表达及融合能力,最后,提出了一种快速空间金字塔池化平均池化模块,改进后的 YOLOv5s 模型精度提升了 3.1%。曾勇杰等^[8]针对叶片缺陷检测,改进 YOLOv8 算法在风机叶片缺陷检测上的应用,提出了高效多尺度卷积模块替代了传统残差块中的卷积层,通过分组卷积技术提升了检测的精确性,提出了 Inner-Wise-MPDIoU 损失函数替代传统的 CIOU 损失函数,该模型相比原始 YOLOv8n 提升了 2.3%。上述的缺陷检测研究主要聚焦于检测精度的提升,然而追求提升模型精度会加重其运行的复杂程度,不符合在计算资源受限的设备中平稳运行的需求,所以,针对复杂网络进行轻量化设计是非常有必要的。

李相垚等^[9]提出了一种基于 DCS-YOLOv8 的钢材表面检测算法,在输入端提出了 C2f-DCNV2 的模块的改进方法,在网络中引入了一种加权特征 Concat-BiFPN 模块,对特征融合进行了优化,最后,对空间金字塔池化进行了重构。改进后的模型提升精度的同时,计算量降低了 5%。梁礼明等^[10]基于 YOLOv8s 模型,提出一种轻量高效的钢材缺陷检测算法 FMG-YOLOv8s,该方法首先采用轻量级的 FasterNet 网络作为骨干网络,重构特征交互模块 MC2f,最后设计 GS-Detect 模块作为整体模型的检测网络,降低模型复杂度,提升训练和推理速度。冯夫健等^[11]提出了基于自注意特征融合的钢材表面小目标缺陷检测模型 SFNet,模型结合滑动窗口机制和特征融合模块 CSP-FCN 以提升模型特征信息整合能力。吴亚尉等^[12]提出了一种改进 YOLOv5s 算法的轻量化钢材表面缺陷检测模型,通过引入 GhostV2 Bottleneck 轻量化模块作为主干特征提取网络,以减少网络的参数量,将特征融合网络中的普通卷积块替换为深度可分离卷积块进一步降低模型的计算复杂度,加入 RepLK 大卷积核优化大尺度方差的检测效果。

以上研究者提出的改进模型已具备了一定程度的轻量化效果,但改进后的模型轻量化程度仍然不足,导致模型的计算量、参数量仍然比较大,且应关注改进后的模型体积大小并持续优化,以保证模型可以平稳运行的前提下,减少对设备平台的资源开销,降低模型的实际应用成本。

因此,本文提出了一种基于 YOLOv8n 的改进算法,该

算法在确保检测精度不受影响的前提下,实现了显著的轻量化效果并压缩了模型体积。

本文工作如下:1)改进网络的主干网络层为 ShuffleNetV2,显著降低计算量和参数量;2)加入通道注意力机制,注意力机制的引入以增强特征表达能力并提高检测性能;3)改进颈部层的特征融合模块,使用双向特征金字塔网络结构优化特征传递结构;4)加入双重卷积核组成为 C2f-Dual 模块加强特征提取能力并进一步提高模型轻量化。

1 YOLOv8 目标检测算法

YOLOv8^[13]作为一种高效的目标检测算法,其结构设计遵循单阶段架构原则,整体分为三部分:主干网络、颈部网络以及检测头。

主干网络 Backbone 作为特征提取的核心网络结构,为后续检测任务给予坚实基础。具体而言,它融合了 5 个卷积模块、4 个 C2f 模块以及 1 个 SPPF 模块,能够有效提取并整合图像中不同层次的特征信息。

颈部网络 Neck 对特征信息融合和处理。FPN^[14]结合了高层语义信息与低层细节信息,进一步提升模型对于不同尺寸目标的检测能力。PAN^[15]加入自下而上的特征传递路径,丰富低层和高层特征的深度融合,提升了网络在复杂背景下的特征表达能力。

头部网络 Head 是 YOLOv8 网络中最终输出的部分,负责将经过 Neck 层处理后的特征图进行解码,生成目标检测结果。YOLOv8 网络结构图,如图 1 所示。

2 改进 YOLOv8 目标检测算法

为了解决钢材表面缺陷检测模型参数量大、计算复杂度高以及对运算平台资源要求高的问题,本文基于 YOLOv8n 算法做出相应改进,集中在主干网络的特征提取过程和颈部网络的特征融合过程进行改进优化,并加入注意力机制,使改进模型符合精度要求和轻量化效果。

2.1 ShuffleNetV2 网络

ShuffleNetV2^[16]是一种轻量级网络,具备计算量小的特点并兼具优良的提取特征能力。针对 ShuffleNetV1^[17]中关于组卷积和特征信息流动的局限性,从而在移动设备等计算资源受限场景下,提升速度与精度之间的平衡。在 V2 版本中,使用了通道拆分(Channel Split)技术,将输入特征图沿通道维度分开,分别进行后续处理。一部分特征执行简单的恒等映射,而另一部分则通过更复杂的卷积操作来增强表达能力。这样既减少了计算复杂度,又避免了组卷积引入的信息传递瓶颈,并使用了标准的 1×1 卷积,确保所有通道均可参与卷积操作。此外,下采样模块设计为无残差连接,这样设计的好处是在特征空间发生变化的情况下,避免了简单的残差连接导致的特征对齐困难。通道混洗(Channel Shuffle)模块用于重新排列通道顺序,进

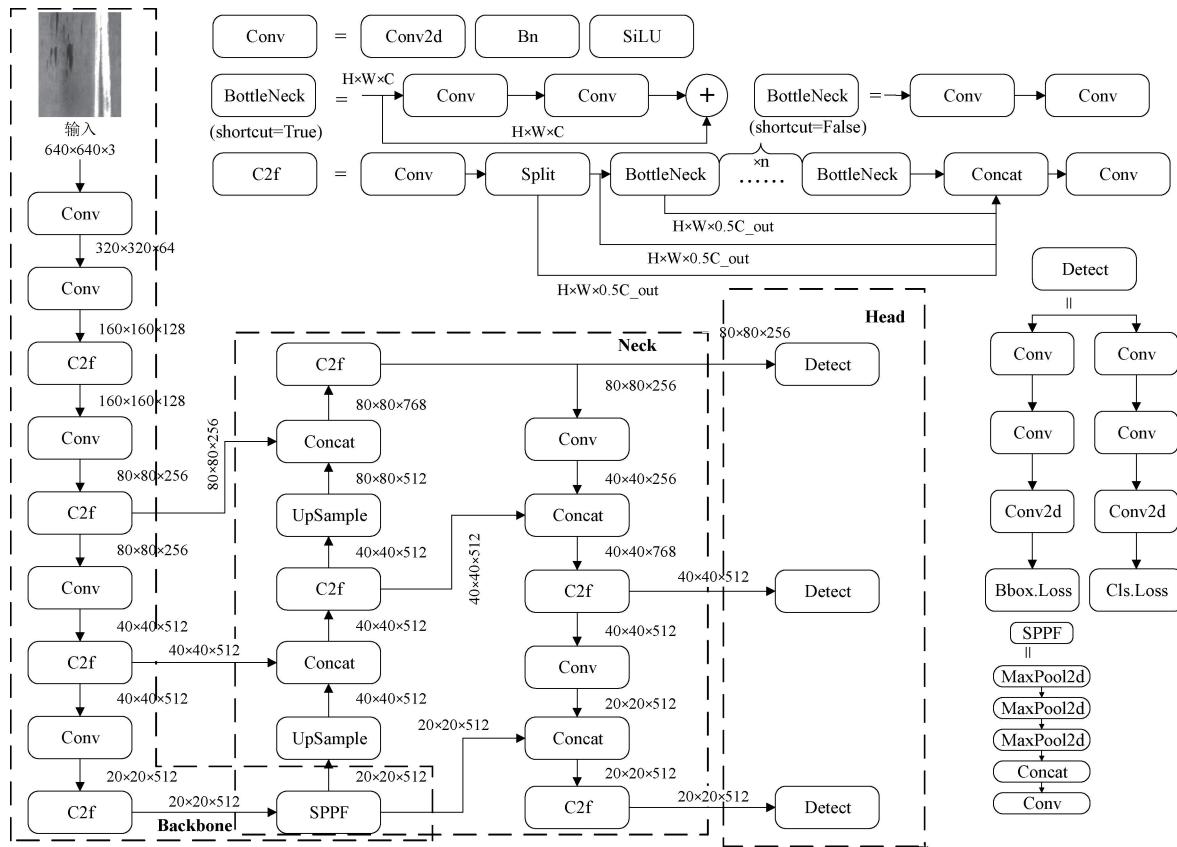


图 1 YOLOv8 结构图

Fig. 1 YOLOv8 structure diagram

而促使分支特征的融合,ShuffleNetV2 网络结构图,如图 2

所示。

准定位和识别目标物体。CA 注意力机制结构,如图 3 所示。

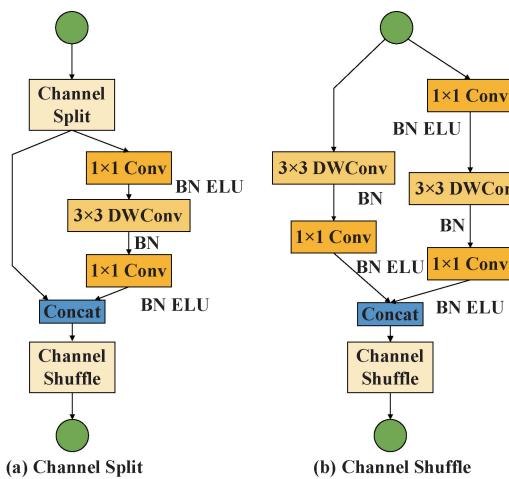


图 2 ShuffleNetV2 网络结构图

Fig. 2 ShuffleNetV2 network architecture diagram

2.2 注意力机制 CA

坐标注意力^[18]机制(coordinate attention, CA)具有优化模型特征表达能力的显著优势。它能够使得模型更精

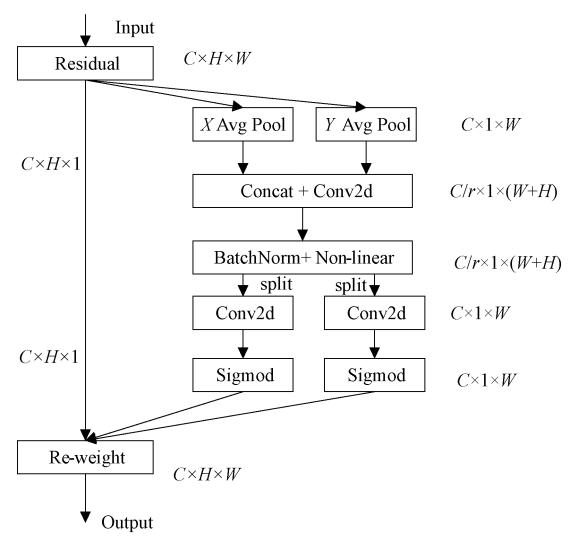


图 3 CA 注意力机制结构图

Fig. 3 The structure diagram of coordinate attention

为明确获取具体位置信息,对全局平均池化进一步分解。具体公式为:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (1)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (2)$$

具有全局感受野的特征图在宽度和高度方向进行拼接操作,并使用共享的 1×1 卷积核将维度压缩至 C/r ,随后对特征图 F_1 进行批量归一化并通过激活函数,得到尺寸为 $1 \times (W+H) \times C/r$ 的特征图 f ,如式(3)所示。

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

对特征图 f 实施 1×1 卷积操作,根据原始的高度和宽度分别恢复通道,生成与原图通道数相匹配的特征图 F_h 和 F_w 。接着,应用 Sigmoid 激活函数对这两个特征图进行非线性变换处理,具体过程如式(4)和(5)。其中,高度方向注意力权重为 g^h ,宽度方向注意力权重为 g^w 。

$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

将初始特征图与通过计算得到的注意力权重融合并进行乘法加权,促使特征图在宽度和高度方向上均融入注意力机制,公式为:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

2.3 双向特征金字塔网络 BiFPN

双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)是一种改进的特征金字塔网络结构,以

更好地捕获多尺度信息和提高模型性能。其中,BiFPN 引入了双向连接,更好地融合低级别和高级别特征;促进了特征的上下文传播,且具有较少的参数和计算复杂度,使其适用于嵌入式设备和实际部署;对于小目标或复杂场景,其性能改进尤为显著,改进前后的特征融合模块结构对比,如图 4 所示。

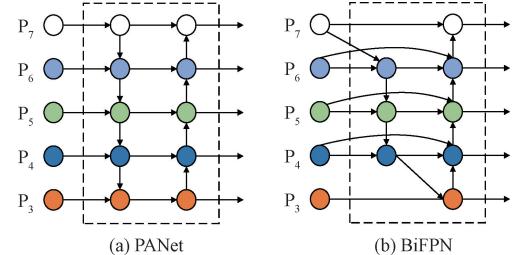


图 4 改进前后的特征融合模块结构对比

Fig. 4 Comparison of feature fusion module structure before and after improvement

2.4 改进 C2f 模块

将低级特征图和高级特征图进行融合时,YOLOv8 中的 C2f 模块因为采用了深层的网络结构,使得模型面临高计算复杂度和庞大的计算量的问题。本文采用双重卷积核 DualConv^[19]重构 C2f 中的卷积来改善计算成本和参数数量。DualConv 结构如图 5 所示。

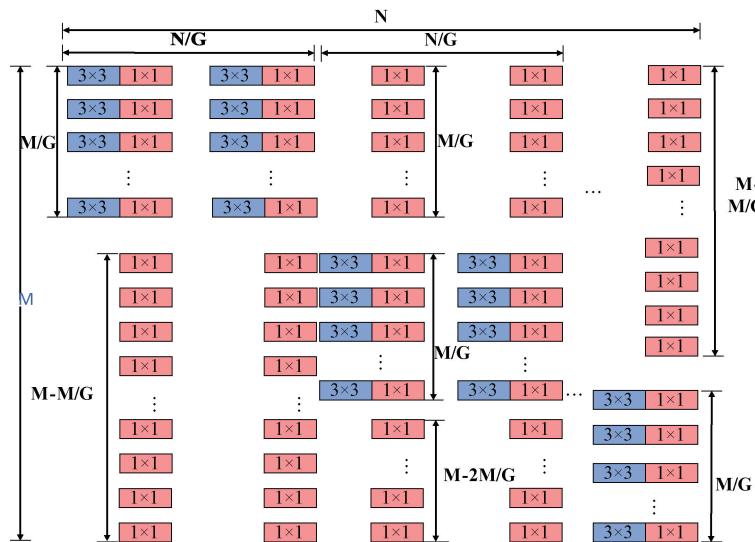


图 5 双重卷积核 DualConv 的结构布局

Fig. 5 Diagram of the dual convolutional kernels structure

对 C2f 模块的改进集中在 BottleNeck 中的卷积层,将原卷积核替换为 DualConv,实现了更高效的卷积计算,除此之外,整体结构依然遵循 CSP 风格设计。BottleNeck 模块结构对比图,如图 6 所示。

2.5 改进 C2f 模块实验分析

为了更好地说明本文改进 C2f 模块的优势,在 YOLOv8n 基础上进行如下的实验分析,如表 1 所示。

根据实验分析结果,本文改进后的 C2f-DualConv 模块在精度、参数量和浮点数计算量方面的实验表现最为均衡,确保模型在没有丢失精度的情况下达到优良的轻量化效果。

2.6 改进后的网络结构

改进后的 YOLOv8 算法在轻量化方面效果显著,在 YOLOv8n 算法的基础上,通过 ShuffleNetV2 特征提取网

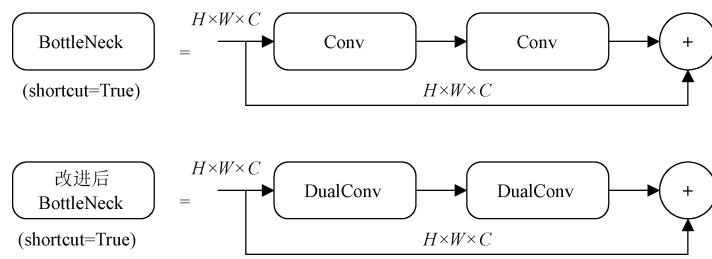


图 6 改进后的 Bottleneck 模块结构对比

Fig. 6 Comparison diagram of the improved Bottleneck module structure

表 1 不同 C2f 模块实验分析

Table 1 Experimental analysis of different C2f modules

方法	Params/ M	mAP50/ %	GFLOPs
C2f-DualConv	2.7	75.7	7.4
C2f-ODConv	4.4	75.7	7.0
C2f-GhostModule- DynamicConv	2.1	74.6	5.8
C2f-Faster-EMA ^[20]	2.9	75.8	7.6

络改进模型主干层;引入 BiFPN 进行特征融合;加入 CA 增强特征信息识别;改进 C2f 模块,将 C2f 的 Bottleneck 模块中的卷积改进为 DualConv,从而在保持准确性的同时减少模型的参数数量和计算成本,其轻量化的特点适合搭建轻量级网络,改进后的 YOLOv8 模型如图 7 所示。

3 实验结果和分析

3.1 实验配置及数据集

实验环境的详细信息,如表 2 所示。

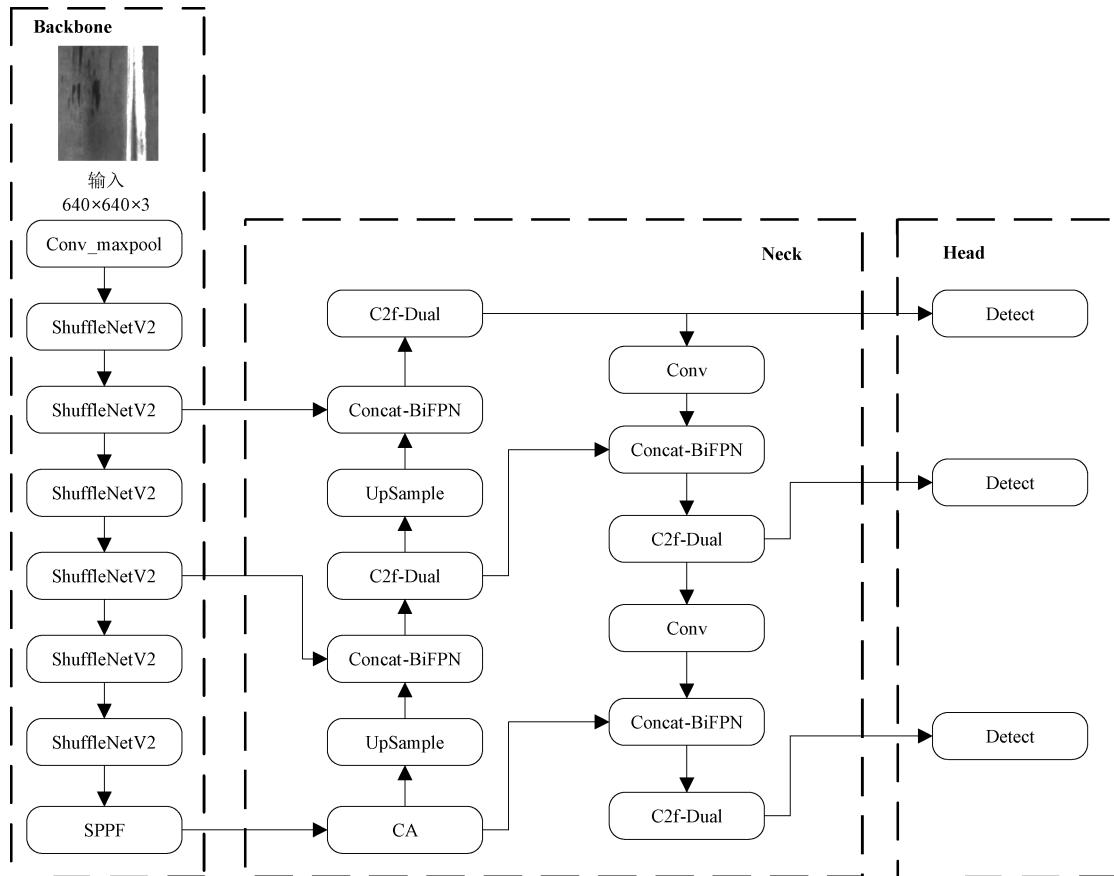


图 7 改进后的 YOLOv8n 模型图

Fig. 7 Improved YOLOv8n model diagram

表 2 实验环境

Table 2 Experimental environment

名称	参数
深度学习框架	Ultralytics YOLOv8.2.77+Python-3.8.10+torch-2.0.0+cuda11.8
操作系统	Ubuntu20.04
GPU	NVIDIA GeForce RTX 3090
CPU	14 vCPU Intel(R) Xeon(R) Platinum 8362 CPU@2.80 GHz

本实验所用数据集为 NEU-DET 钢材表面缺陷数据集,该数据集为公开数据集,包含 1 800 张图像,涵盖 6 种主要缺陷类型:裂纹(crazing, Cr)、夹杂(inclusion, In)、斑块(patches, Pa)、麻点(pitted_surface, Ps)、压入氧化铁皮(rolled-in_scale, Rs)和划痕(scratches, Sc)。部分缺陷类型的示例如图 8 所示。数据集按照训练集:验证集:数据集为 8:1:1 的比例划分。

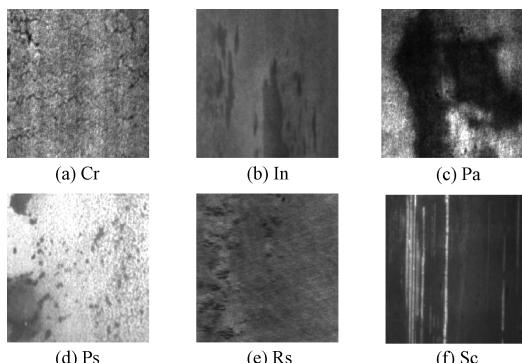


图 8 缺陷种类示例

Fig. 8 Examples of partial defects

实验参数如表 3 所示,包含了模型训练的关键参数设置。

表 3 参数配置

Table 3 Parameter configuration

参数	配置
Epoch	200
Optimizer	SGD
lr	0.01
batch_size	16
close_mosaic	True
momentum	0.937
weight_decay	0.0005

3.2 实验评价指标

为评估改进后模型的性能,实验采取准确率(precision, P)、召回率(recall, R)和平均精度均值(mean average precision, mAP)作为实验评价指标。

$$P = \frac{T_p}{T_p + F_p} \quad (7)$$

$$R = \frac{T_p}{T_p + F_N} \quad (8)$$

其中, T_p 表示真正例,即把正例预测正确; F_p 表示假正例,即假例预测成正例; F_N 表示假反例,即正例预测成反例。

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

$$mAP = \frac{1}{n} \sum_{j=1}^c AP_j \quad (10)$$

其中, AP 表示某个目标类的平均精度, n 表示类别数量。本文采用实验结果的 mAP 评价标准为 mAP50。

除此之外,衡量模型计算效率和模型容量的指标还需考虑到计算复杂度和参数数量。

3.3 消融实验

为了评估各个模块对算法性能的提升作用,本实验采用基础模型 YOLOv8n 作为对比基准。通过逐步添加各个改进模块的方式,分别在 NEU-DET 数据集上进行训练和测试,实验遵循控制变量法,其实验结果如表 4 所示。

表 4 消融实验结果

Table 4 Results of ablation experiment

模型	ShuffleNetV2	BiFPN	CA	C2f-Dual	mAP50/%	参数量/bit	计算量/GFLOPs	模型大小/MB
YOLOv8n					75.2	3 012 018	8.2	6.3
YOLOv8n-S	✓				75	1 830 730	5.1	3.9
YOLOv8n-SB	✓	✓			73.5	1 827 738	5.1	3.9
YOLOv8n-SBC	✓	✓	✓		73.7	1 842 819	5.1	3.9
YOLOv8n-SBD	✓	✓		✓	74.4	1 682 966	4.8	3.6
YOLOv8n-SC	✓		✓		74.7	1 842 807	5.1	3.9
YOLOv8n-SD	✓			✓	74.6	1 681 678	4.8	3.6
YOLOv8n-SCD	✓		✓	✓	74.2	1 695 031	4.8	3.6
YOLOv8n-SCBD	✓	✓	✓	✓	75.5	1 695 043	4.8	3.6

表 4 中的“√”代表 YOLOv8n 网络已采用相应改进方案,与原网络对比,将 ShuffleNetV2 替换为主干网络,使参数量显著减少至 1 830 730,模型大小缩小至 3.9 MB,计算量降为 5.1 GFLOPs,性能表现略有下降,mAP50 为 75%;在改进主干的基础上进一步添加了 BiFPN 结构,使 mAP50 下降至 73.5%,但模型参数量和大小与 YOLOv8n-S 基本保持一致;引入 CA 后,精度提升至 73.7%;用 C2f-Dual 替换原 C2f 结构后, mAP50 提高至 74.4%,参数量和模型大小进一步减少,分别为 1 682 966 和 3.6 MB;最终,改进后的 YOLOv8n-SBCD 模型在性能

上提升至 75.5%,同时显著减少了模型的参数量、计算量和模型体积。

通过消融实验可见, YOLOv8n-SBCD 模型展现了最优的性能与计算效率平衡,其 mAP50 提升 0.3%,确保在保证不丢失精度的前提下,证明了轻量化的有效性。

3.4 实验结果分析

本实验将对未经优化的 YOLOv8n 模型与优化后的 YOLOv8n-SBCD 模型在测试集上的检测能力进行对比评估,在 6 种钢材表面缺陷的检测结果的对比测试如图 9 所示,对比结果表明,改进后的算法具有良好的检测性能。

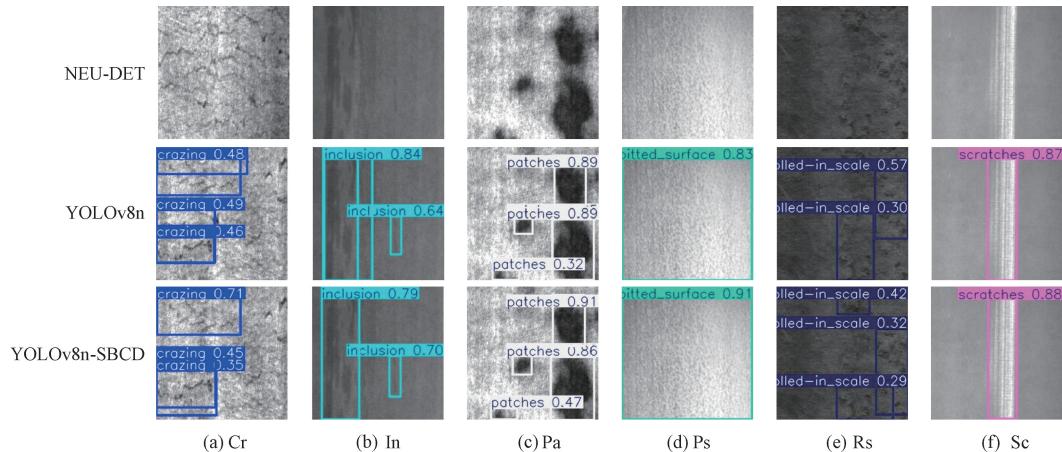


图 9 算法改进前后结果对比

Fig. 9 Comparison of detection results before and after algorithm improvement

为了更准确评估模型检测的表达能力,本文引入了 Grad-CAM 可视化分析技术。通过 Grad-CAM 技术进行更直观的可视化效果展示,它能够通过生成热图来凸显图像中的关键区域,从而直观地揭示模型在做出决策时所关注的目标位置及范围。该方法直观地展示了模型预测的目标区域在图像中的位置,有助于有效评估模型的检测能力。Grad-CAM 可视化结果如图 10 所示。

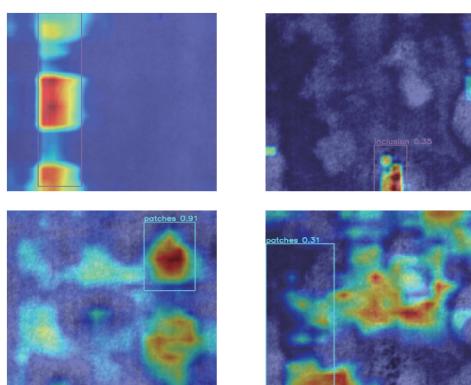


图 10 Grad-CAM 可视化

Fig. 10 Grad-CAM visualization

3.5 对比实验

缺陷检测中的效能,本文在相同的实验框架和统一的数据集划分原则下,进行实验并对比分析。实验选取了当前主流且表现优异的目标检测算法作为参照,具体对比结果如表 5 所示。

根据对比实验结果,经典 YOLO^[20]系列算法中的 YOLOv9^[21]和 YOLO11s 算法在检测精度上略微好于本文方法,但在模型参数、计算量及模型尺寸的评价指标上均要远超本文提出的改进算法,并不符合应用在资源受限的运行环境。YOLOv7^[22]与 YOLO11n 的在精度方面好于本文算法,但二者精度指标的提升均有不同程度的代价,且前者代价过于庞大,相比之下, YOLO11n 是经典 YOLO 系列中综合表现较好的算法,但在轻量化程度上还达不到本文改进算法的效果。

本文还对比了近期研究者提出的改进方法,文献[14]提出的方法在检测精度上占据领先,其 mAP50 达到 77.4%,高于本文 1.9%,但本文的参数量要低于其参数量的 43%,且二者所占模型体积相差 2 倍多,本文在计算量方面的优势更为明显。文献[13]和文献[15]均为轻量化钢材表面缺陷检测的改进模型,其中,文献[13]的改进模型在检测精度方面稍好于本文,文献[15]的精度较好于本文,但二者在轻量化方面的改进,相较于本文仍有不小的差距。与赵佰亭等^[23]提出的改进算法相比,本文改进模型

表 5 对比实验结果

Table 5 Results of comparative experiments configuration

模型	mAP50/%	参数量/10 ⁶	计算量/GFLOPs	模型大小/MB
YOLOv5s	73.9	7.03	16.0	13.8
YOLOv7-tiny	72.3	6.02	13.2	12.3
YOLOv7	78.5	37.22	105.2	74.9
YOLOv8n	75.2	3.01	8.2	6.3
YOLOv8s	74.7	11.13	28.7	22.5
YOLOv9	75.6	60.80	266.2	122.4
YOLOv10s	74.4	8.07	24.8	16.6
YOLOv10n	72.7	2.70	8.4	5.8
YOLO11s	75.8	9.43	21.6	19.2
YOLO11n	77.3	2.59	6.4	5.5
文献[9]	76.9	3.30	7.8	—
文献[10]	75.9	2.90	6.7	—
文献[11]	77.4	3.93	13.2	8.3
文献[12]	77.3	—	14.0	12.0
文献[24]	77.2	33.13	—	—
本文	75.5	1.69	4.8	3.6

精度虽有 1.7% 的不足,但本文在参数量方面具有极为显著的优势,实现了更优的轻量化表现。此外,部分研究者的改进模型中并没有体现出改进后的模型大小,相比之下,本文达到了 3.6 MB 的模型体积,大大减轻了模型对于存储资源的压力。

综合来看,针对资源受限的运行环境,本文所提出的改进算法展现出了出色的性能。与现有主流算法相比,本文算法在资源占用上更为精简,实现轻量化的同时保证检测精度,符合实际应用需求。

4 结 论

针对钢材表面缺陷检测模型参数量大、计算复杂度高,对运算平台资源需求高的问题,提出了一种轻量化的改进算法。采用 ShuffleNetV2 作为主干网络;引入 BiFPN 以实现特征融合;在网络中添加了注意力机制 CA;将 C2f 中 BottleNeck 模块中的卷积替换为 DualConv 并组成 C2f-Dual 模块。

实验结果表明,本文改进后的模型检测精度 mAP50 和原模型相比提升 0.3%,满足在不损失检测精度的同时实现轻量化效果,参数量为原来的 56.2%,体积和计算量分别降至 3.6 MB 和 4.8 GFLOPs,相比原模型分别降低了 42.86% 和 41.47%,改进后,模型对运行平台的资源要求极低,在模型轻量化方面优势明显,大幅减少实际应用成本和运算压力,后续将继续提高模型精度,并尝试部署至硬件设备。

参考文献

[1] 李丹,王慢慢,刘俊德,等.基于轻量级卷积神经网络的

带钢表面缺陷识别[J].仪器仪表学报,2022,43(3):240-248.

LI D, WANG M M, LIU J D, et al. Steel surface defect recognition based on a lightweight convolutional neural network [J]. Chinese Journal of Scientific Instrument, 2022, 43(3):240-248.

[2] 冯丽丹,王闯,祁军,等.基于改进 YOLOv7 的管道缺陷检测技术研究[J].辽宁石油化工大学学报,2024,44(4):82-90.

FENG L D, WANG CH, QI J, et al. Research on pipeline defect detection technology based on improved YOLOv7 [J]. Journal of Liaoning Petrochemical University, 2024, 44(4):82-90.

[3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014:580-587.

[4] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1149.

[5] REDMON J, FARHADI A. YOL09000: Better, faster, stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.

[6] 朱成杰,刘乐乐,朱洪波.基于 YOLOv8-NFMC 的带钢表面缺陷检测算法[J].国外电子测量技术,2024,43(7):97-104.

ZHU CH J, LIU L L, ZHU H B. Surface defect detection algorithm for strip steel based on YOLOv8-NFMC [J]. Foreign Electronic Measurement Technology, 2024, 43(7):97-104.

- [7] 金鑫, 洪彬, 于东升, 等. 基于改进 YOLOv5s 的煤矿电力设备缺陷检测[J]. 电子测量技术, 2023, 46(19): 148-155.
JIN X, HONG B, YU D SH, et al. Defect detection of coal mine power equipment based on improved YOLOv5s[J]. Electronic Measurement Technology, 2023, 46(19): 148-155.
- [8] 曾勇杰, 范必双, 杨涯文, 等. 改进 YOLOv8 算法在风机叶片缺陷检测上的应用[J]. 电子测量与仪器学报, 2024, 38(8): 26-35.
ZHENG Y J, FAN B SH, YANG Y W, et al. YOLOv8 algorithm is improved in the defect detection of wind turbine blades applications [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(8): 26-35.
- [9] 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, 1-10 [2024-11-30]. <https://doi.org/10.13433/j.cnki.1003-8728.20240128>.
LI X Y, HOU H L, YANG AO, et al. Research on DCS-YOLOv8 algorithm for steel surface defect detection[J/OL]. Mechanical Science and Technology for Aerospace Engineering, 1-10 [2024-11-30]. <https://doi.org/10.13433/j.cnki.1003-8728.20240128>.
- [10] 梁礼明, 龙鹏威, 李俞霖. 改进轻量高效 FMG-YOLOv8s 的钢材表面缺陷检测算法[J/OL]. 计算机工程与应用, 1-11 [2024-11-30]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20241030.1315.006.html>.
LIANG L M, LONG P W, LI Y L. Improved lightweight and efficient FMG-YOLOv8s algorithm for steel surface defect detection [J/OL]. Computer Engineering and Applications, 1-11 [2024-11-30]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20241030.1315.006.html>.
- [11] 冯夫健, 罗太维, 谭棉, 等. 基于自注意特征融合的钢材表面小目标缺陷检测[J]. 电子测量技术, 2024, 47(19): 172-180.
FENG F J, LUO T W, TAN M, et al. Defect detection of small targets on steel surface based on self-attention feature fusion [J]. Electronic Measurement Technology, 2024, 47(19): 172-180.
- [12] 吴亚尉, 明帮铭, 何剑锋, 等. 基于 YOLO-GR 算法的轻量化钢材表面缺陷检测[J]. 组合机床与自动化加工技术, 2023(11): 107-111, 115.
WU Y W, MING B M, HE J F, et al. Lightweight steel surface defect detection based on YOLO-GR[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2023(11): 107-111, 115.
- [13] 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 SH, XU H, ZHANG Y. Lightweight forged part surface crack detection algorithm [J]. Electronic Measurement Technology, 2024, 47(11): 123-130.
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2017: 936-944.
- [15] WANG K X, LIEW J H, ZOU Y T, et al. PANet: Few-shot image semantic segmentation with prototype alignment [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 9196-9205.
- [16] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design [C]. European Conference on Computer Vision, 2018.
- [17] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018.
- [18] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
- [19] ZHONG J CH, CHEN J Y, MIAN A. DualConv: Dual convolutional kernels for lightweight deep neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(11): 9528-9535.
- [20] 马磊, 李晔, 王宇翔. YOLOv8-FD: YOLOv8 改进的钢板表面缺陷检测方法[J]. 计算机工程与应用, 2024(11): 1-13.
MA L, LI Y, WANG Y X. YOLOv8-FD: YOLOv8 improved method for detecting surface defects on steel plates[J]. Computer Engineering and Applications, 2024(11): 1-13.
- [21] JOCHER G, STOKEN A, BOROVEC J, et al. Ultralytics/YOLOv5: v3.1-Bug fixes and performance improvements[J]. Zenodo, 2020, DOI: 10.5281/zenodo.4154370.
- [22] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision, 2024: 1-21.
- [23] 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]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [24] 赵佰亭, 张晨, 贾晓芬. ECC-YOLO: 一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2024, 38(4): 108-116.
ZHAO B T, ZHANG CH, JIA X F. ECC-YOLO: An improved method for detecting surface defects in steel [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.

作者简介

刚帅(通信作者), 硕士研究生, 主要研究方向为目标检测。
E-mail: 860988103@qq.com

刘培胜, 硕士, 副教授, 主要研究方向为人工智能、机器学习。
E-mail: 66841954@qq.com

郭希旺, 博士, 副教授, 主要研究方向为人工智能、智能计算。
E-mail: x. w. guo@163. com