

本文引用格式:姚若禹,郑世玲,史怡璇,等. 基于改进 YOLOv8n 的钢材表面缺陷检测算法[J]. 聊城大学学报(自然科学版), 2025, 38(2):177-189.

Citation format: YAO Ruoyu, ZHENG Shiling, SHI Yixuan, et al. Steel surface defect detection algorithm based on improved YOLOv8n[J]. Journal of Liaocheng University (Natural Science Edition), 2025, 38(2):177-189.

文章编号 1672-6634(2025)02-0177-13

DOI 10.19728/j.issn1672-6634.2024070019

基于改进 YOLOv8n 的钢材表面缺陷检测算法

姚若禹, 郑世玲, 史怡璇, 张思启, 张钊飞, 高飞龙, 张霞

(聊城大学 物理科学与信息工程学院, 山东 聊城 252059)

摘要 针对当前钢材缺陷检测中目标特征复杂导致检测精度较低的问题,提出了一种基于改进 YOLOv8n 的钢材缺陷检测算法 GOS-YOLO。首先,使用由轻量级卷积 GSConv 构建的 Slim-neck 范式作为特征融合网络,在减小模型参数量的同时提高精度。其次,将骨干网络的部分 C2f 模块替换为与全维度动态卷积(ODConv)相结合的 C2f_ODConv 模块,以实现模型的多维特征关注进而提高模型的精度。最后,将结合了多分支结构与压缩和激励操作的 SENetV2 注意力机制嵌入颈部网络,增强模型对复杂特征的提取能力。实验结果表明,在 NEU-DET 数据集上,GOS-YOLO 的 R 、 mAP_{50} 、 $mAP_{50\sim95}$ 较 YOLOv8n 分别提高了 3.3%、1.7%、2.3%。在 VOC2007 数据集上 $mAP_{50\sim95}$ 提高了 1%,FLOPs 降低了 16%。

关键词 深度学习;计算机视觉;YOLOv8;缺陷检测

中图分类号 TG142.1;TP391.41;TP183 **文献标志码** A **开放科学(资源服务)标识码(OSID)**



Steel Surface Defect Detection Algorithm Based on Improved YOLOv8n

YAO Ruoyu, ZHENG Shiling, SHI Yixuan, ZHANG Siqi,
ZHANG Xinfei, GAO Feilong, ZHANG Xia

(School of Physics Science and Information Technology, Liaocheng University, Liaocheng 252059, China)

Abstract Aiming at the current problem of low detection accuracy due to the complexity of target features in steel defect detection, a steel defect detection algorithm, GOS-YOLO, based on improved YOLOv8n is proposed. Firstly, a Slim-neck paradigm constructed by lightweight-convolution GSConv is used as the feature fusion network, which improves the accuracy while reducing the number of model parameters. Secondly, some of the C2f modules of the backbone network are replaced with C2f_ODConv modules combined with full-dimensional dynamic convolution (ODConv), to achieve multi-dimensional feature attention of the model and thus improve the accuracy of the model. Finally, the SENetV2 attention mechanism, which combines a multi-branch structure with squeeze and excitation operation, is embedded into the neck network to enhance the model's ability to extract complex features. The experimental results

收稿日期:2024-07-31

基金项目:国家自然科学基金项目(62205136);山东省自然科学基金项目(ZR2022MF284);聊城大学博士科研启动基金项目(318050062);聊城大学横向课题(K23LD94)资助

通信作者:郑世玲,女,汉族,博士,讲师,研究方向:智能信息处理,E-mail:zhengshiling@lcu.edu.cn。

show that on the NEU-DET dataset, the R , mAP50, and mAP50~95 of GOS-YOLO are improved by 3.3%, 1.7%, and 2.3%, respectively, compared with YOLOv8n. On the VOC2007 dataset mAP50~95 improved by 1% and FLOPs decreased by 16%.

Key words deep learning; computer vision; YOLOv8; defect detection

0 引言

现代工业生产中,钢铁作为一种关键材料,其应用范围极为广泛,涵盖了汽车制造、建筑、船舶制造等多个重要领域。钢铁的质量和性能直接关系到最终产品的性能和安全,因此,保证钢铁的高质量是工业生产中的一项重要任务。然而,在钢铁的生产过程中,由于工艺复杂及材料本身的特性,钢铁表面裂纹、夹杂等缺陷的产生往往难以避免。这些缺陷的存在不仅会削弱钢铁的机械性能,如强度和韧性,还可能引发安全隐患,影响产品的可靠性和耐用性^[1]。2023年9月,工业和信息化部编制了《钢铁行业智能制造标准体系建设指南(2023版)》,其中涉及的钢铁表面缺陷检测的检测类标准是钢铁行业智能制造标准建设重点中的基础共性标准之一。因此,研究钢铁表面缺陷智能检测算法具有重要意义。

钢铁表面缺陷智能检测系统通常利用高速工业相机拍摄高清数字图像,通过图像处理、机器学习和深度学习等方法对采集到的钢板图像数据进行处理,完成缺陷检测和缺陷分类工作。随着深度学习的快速发展,基于深度学习的缺陷检测算法与传统的图像处理和机器学习算法相比具有更高的检测精度和更快的检测速度,因此成为当前的研究热点^[2]。基于深度学习的目标检测算法分为两阶段(Two-stage)目标检测算法和单阶段(One-stage)目标检测算法。两阶段检测算法包括 R-CNN(regions with CNN features, R-CNN)^[3]、Fast R-CNN^[4]和 Faster R-CNN^[5]等算法,需要先生成包含缺陷的候选框,再对候选框中的区域进行分类,与单阶段目标检测算法相比,其检测速度较慢。单阶段目标检测算法摒弃了生成候选框的步骤,直接对缺陷的位置和类别等信息进行预测,这种设计使单阶段目标检测算法能够实现快速且高效的目标检测,代表性的算法有 SSD(Single Shot MultiBox Detector)^[6]和 YOLO(You Only Look Once)系列算法^[7-12]等。

YOLO 系列目标检测算法凭借其在缺陷检测上出色的检测速度和准确率,赢得了众多研究者的青睐。Li 等人^[13]提出了一种基于改进 YOLO 的钢材缺陷检测算法,该算法由 27 个卷积层构成,提升了检测精度。Kou 等人^[14]在 YOLOv3 中采用无锚点特征选择机制和密集卷积块使模型实现了更高的精度。针对 YOLOv5,Guo 等人^[15]提出了 MSFT-YOLO 算法,该算法将基于 Transformer 设计的 TRANS 模块融合到 YOLOv5 的骨干网络和检测头中,实现了特征和全局信息的结合并提高了检测精度。Wang 等人^[16]结合多尺度检测块和 SA 注意力机制,实现了 72% 的平均检测精度。Zhao 等人^[17]使用 Res2Net 模块代替 C3 模块,设计双特征金字塔网络(DFPN),最后采用解耦头代替原检测头,使模型实现了更强的性能。曹义亲等人^[18]构造了一种带残差边的 SPP_Res 特征金字塔结构,引入多头注意力机制(C3_MHSA)和多层特征融合机制,增强了模型的速度和特征提取能力。黄友文等人^[19]设计了 C3-ODCBS 模块,引入 WIoUv3-Loss 损失函数和解耦头,使模型在钢铁缺陷上展现出更高的检测精度。张世强等人^[20]引入坐标注意力机制、空洞空间卷积池化金字塔结构、选择性内核注意力机制、SIoU 损失函数,有效提高了模型对钢铁缺陷的检测性能。Qu 等人^[21]提出了一种基于改进 PP-YOLOE 的钢材缺陷检测算法,该算法在骨干网络中引入 CA 注意力机制和特征金字塔结构,实现了 76.4% 的平均检测精度。针对 YOLOv7,高春艳等人^[22]引入 CARAFE 上采样算子、DY-Block、解耦头和 NF-IoU 损失函数,有效降低了模型的漏检率。梁礼明等人^[23]设计了 VoVGA-FPN 网络、引入三重坐标注意力机制和鬼影混洗卷积并添加大目标检测层实现了检测精度和模型体积之间的平衡。针对 YOLOv8,Song 等人^[24]将 C2f_DCNv2 和 Biformer 嵌入骨干网络,再将颈部网络替换为 BiF-PN,最后替换 WIoUv3 损失函数,有效提高了检测精度。邹彦艳等人^[25]在 C2f 模块中引入可变形卷积 DCNv2、在颈部网络嵌入 SE 注意力机制、替换 GIoU 损失函数实现了更高的检测精度和运算效率。马金林等人^[26]通过引入 C2f_DCN 模块和多头注意力机制,拼接 Backbone 和 Head 的特征有效提高了模型在钢铁缺陷上的检测精度。

综上所述, YOLO 系列算法确实能够有效地进行钢铁表面缺陷检测。然而, 当前算法仍存在缺陷目标特征复杂导致检测精度低的问题。为进一步提高钢铁表面缺陷检测的检测精度, 本文提出一种基于改进 YOLOv8n 的钢材缺陷检测算法 GOS-YOLO。该算法引入 Slim-neck 范式作为特征融合网络替换 YOLOv8 的原始颈部网络, 增强模型的特征融合和处理能力, 在保证精度的同时降低模型的复杂程度。在骨干网络引入全维度动态卷积, 提升模型在卷积核数量、空间核尺寸、输入通道数、输出通道数等四个不同维度上对特征的提取能力, 从而增强模型对特征的学习和表达能力, 进而提高识别精度。将结合多分支结构与压缩和激励操作的 SENetV2 注意力机制嵌入颈部网络, 增强网络捕获通道模式和全局知识的能力, 提升模型对特征图的学习能力。最后通过在 NEU-DET 数据集和 VOC2007 数据集上的实验验证了算法的有效性。

1 YOLOv8 算法原理

YOLOv8 算法是由发布 YOLOv5 算法的 Ultralytics 公司于 2023 年发布的^[27]。YOLOv8 在数据预处理方面延续了 YOLOv5 的马赛克增强 (Mosaic)、混合增强 (Mixup) 等多种数据优化策略。在模型架构方面, YOLOv8 在保留 YOLOv5 骨干网络的同时将 C3 模块替换为结合了 ELAN 思想的 C2f 模块。YOLOv8 的颈部网络依然沿用 PAN-FPN^[28, 29] 的思想。FPN 从高分辨率的特征图开始下采样, 同时从低分辨率的特征图开始上采样, 再将它们连接起来形成特征金字塔, 实现浅层和深层特征图的融合, 提高模型对不同尺度目标的检测能力。PAN 从低分辨率的特征图开始上采样, 同时从高分辨率的特征图开始下采样, 再将它们连接起来形成一条路径, 以便更好地保留低分辨率特征图中的细节信息。YOLOv8 结合 PAN 和 FPN 使模型能够更好地适应不同尺度和分辨率的目标, 增强多个尺度上的语义表达和定位能力。头部网络采用解耦头设计, 将回归分支和预测分支分离, 对目标的类别和位置分别进行预测, 并在回归分支使用分布焦点损失 (Distribution Focal Loss, DFL) 策略来解决边界模糊的问题, 进而提高模型的精度。YOLOv8 网络结构如图 1 所示。

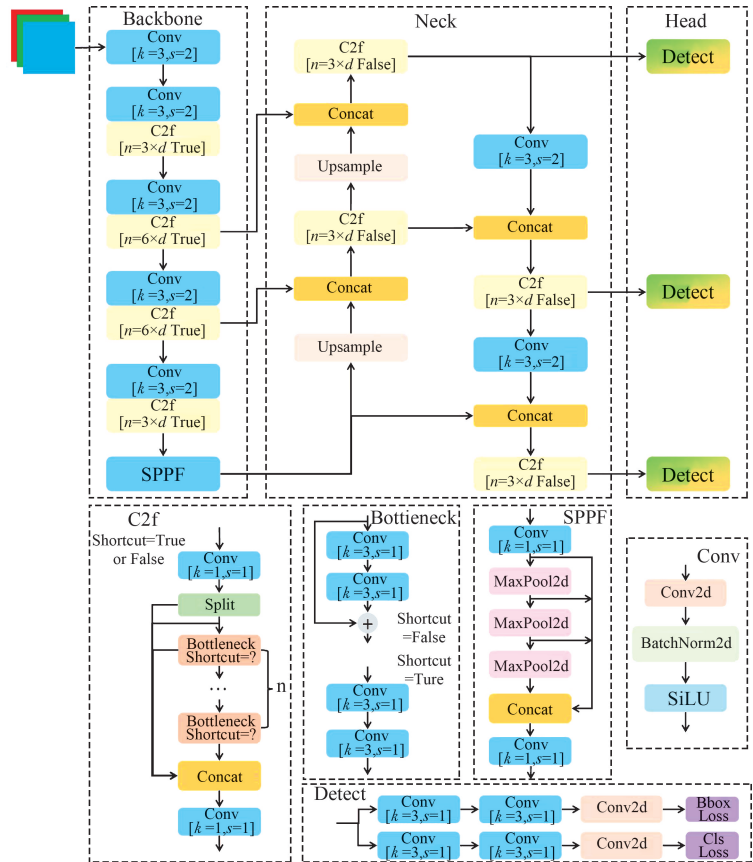


图 1 YOLOv8 网络结构图

Fig. 1 YOLOv8 structure diagram

2 YOLOv8 算法改进

针对当前钢材缺陷检测中目标特征复杂导致检测精度较低的问题, 提出 GOS-YOLO 算法。YOLOv8 原始的颈部网络参数量较高, 计算成本高, 因此我们选择轻量型 Slim-neck 范式^[30]作为特征融合网络, 减少模型参数量的同时增强模型特征融合与表达能力, 提高检测精度。YOLOv8 原始的静态卷积缺乏捕捉多个维度上特征的能力, 因此我们将骨干网络处的 C2f 模块与全维度动态卷积^[31]结合, 增强模型在不同维度上对特征的提取能力。最后引入将压缩和激励操作与多分支结构结合的 SENetV2 注意力机制^[32], 配合 Slim-neck 范式进一步增强模型的特征表达能力。在骨干网络处的 C2f_ODConv 使模型学习到特征图的多维特征后, 再经过 Slim-neck 范式和 SENetV2 注意力机制加强特征的表达, 通过三个模块的配合提升模型的精度。GOS-YOLO 的网络结构如图 2 所示。

2.1 采用 Slim-neck 特征融合网络轻量化改进 Yolov8 的颈部网络

YOLOv8 的颈部网络通过连接骨干网络和头部网络实现特征融合和处理,且颈部网络中的 C2f 模块采用了标准卷积(Standard Convolution, SC),随着网络深度的增加,模型参数量会增加^[30],计算量增大,进而影响 YOLOv8 整体性能。

Li 等人^[30]于 2022 年提出了轻量卷积技术 GSConv 和轻量级结构 VoV-GSCSP,并将 GSConv 和 VoV-GSCSP 联合应用,构成轻量级特征融合网络 Slim-neck 范式。本文将 YOLOv8 的颈部网络替换为 Slim-neck 范式,在保证特征提取能力的同时降低参数量。GSConv 结构如图 3 所示^[30],它由标准卷积(Standard Convolution, SC)、深度可分离卷积(Depth-wise Separable Convolution, DSC)和拼接、混洗操作构成。将 SC 提取的深层语义信息和 DSC 提取的浅层特征信息进行拼接,再对通道信息进行混洗,使 SC 的特征信息渗透到 DSC 的特征信息中,既能够保证模型的精度,同时又能提高速度。

跨级部分网络模块 VoV-GSCSP 是由 GS bottleneck 模块结合一次性聚合思想设计得到的,如图 4 所示^[30]。图 4(a)中的 GS bottleneck 模块由 GSConv 构成,用于提高特征非线性表达和信息复用的能力。VoV-GSCSP 模块的结构如图 4(b)所示,用于不同阶段的特征图之间进行有效的信息融合。

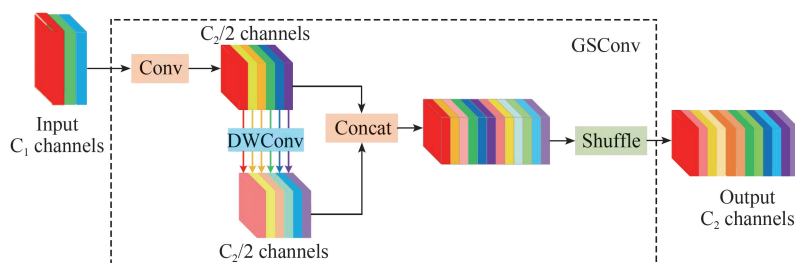


图 3 GSConv 结构图

Fig. 3 GSConv structure diagram

将 GSConv 和 VoV-GSCSP 联合应用,构成 Slim-neck 范式。将 YOLOv8 的特征融合网络替换为轻量级的 Slim-neck 范式,如图 5 所示。不仅能降低模型的参数量和计算量,而且能提高模型的特征融合能力,进而提高检测精度。

2.2 引入全维度动态卷积的 C2f_ODConv

钢材表面缺陷的位置和形状是多样且复杂的。然而,YOLOv8 的 C2f 模块由静态卷积构成,静态卷积核的大小是固定的且感受野相对较小,这限制了它们捕捉局部多样化信息的能力,在提取复杂特征方面存在固有的局限性。因此本文引入全维度动态卷积 ODConv,同时考虑卷积核数量、空间核尺寸、输入通道

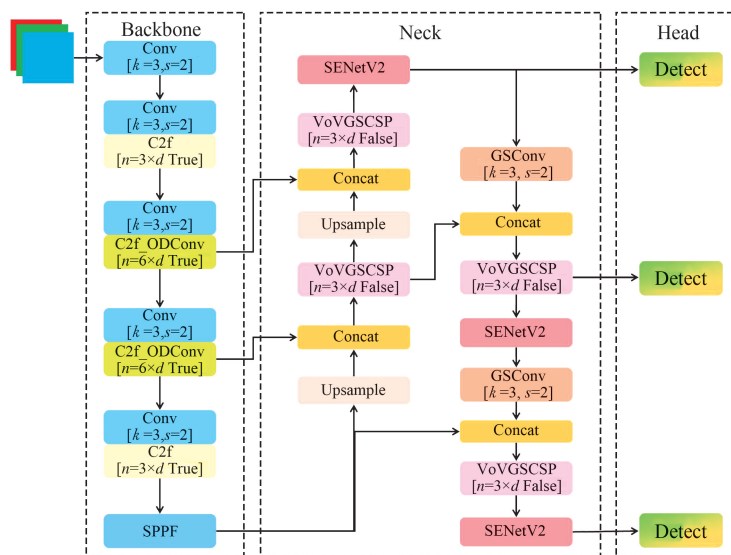


图 2 GOS-YOLO 网络结构图

Fig. 2 GOS-YOLO structure diagram

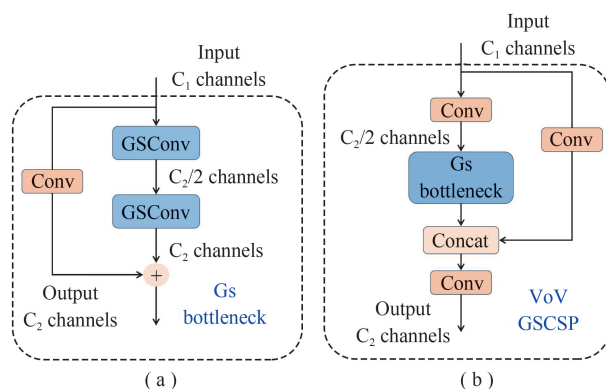


图 4 GS bottleneck 模块与 VoV-GSCSP 模块

(a) GS bottleneck; (b) VoV-GSCSP

Fig. 4 GSConv bottleneck and VoV-GSCSP

(a) GSConv bottleneck; (b) VoV-GSCSP

数和输出通道数四个维度,利用多维注意力机制,以并行的方式沿核空间的四个维度学习四种类型的注意力,并将这些注意力应用到相应的卷积核上,更加精确地定位缺陷的区域复杂特征。其定义为

$$Y = (\alpha_{w1} \odot \alpha_{f1} \odot \alpha_{c1} \odot \alpha_{s1} \odot W_1 + \dots + \alpha_{wn} \odot \alpha_{fn} \odot \alpha_{cn} \odot \alpha_{sn} \odot W_n) * X \quad (1)$$

式中 $Y \in R^{h \times w \times c_{out}}$ 表示输出通道数为 c_{out} , 大小为 $h \times w$ 的特征, $X \in R^{h \times w \times c_{in}}$ 表示输入通道数为 c_{in} , 大小为 $h \times w$ 的特征, W_i 表示第 i 级卷积核, 它由 c_{out} 个 $W_i^m \in R^{k \times k \times c_{in}}$, $m=1, \dots, c_{out}$ 组成。 $\alpha_{wi} \in R$ 表示沿卷积核数量维度的注意力标量, $\alpha_{fi} \in R^{c_{out}}$ 表示沿卷积核输出通道维度的注意力标量, $\alpha_{ci} \in R^{c_{in}}$ 表示沿卷积核输入通道维度的注意力标量, $\alpha_{si} \in R^{k \times k}$ 表示沿卷积核空间核尺寸的注意力标量, $*$ 表示卷积运算, $\odot$ 表示沿核空间不同维度的乘法运算。

全维度动态卷积结构图如图 6 所示^[31]。输入特征 X 经过全局平均池化(GAP)压缩为 c_{in} 个特征向量, 经过全连接层(FC)将特征向量映射到低维空间, 再经过 ReLU 激活函数将特征向量中的负值置为零, 最后将其输出到并行的四个分支, 前三个分支经过全连接层和 Sigmoid 激活函数后分别得到卷积核空间核尺寸、输入通道数、输出通道数上的注意力标量, 最后一个分支经过全连接层和 Softmax 激活函数后得到卷积核数量维度上的注意力标量, 将得到的四个维度上的注意力标量与卷积核相乘累加, 与输入特征 X 卷积得到输出特征 Y 。

图 7 为改进前后的 C2f 模块。本文将 C2f 模块中 Bottleneck 的第二个卷积替换为全维度动态卷积, 从而构成 Bottleneck_ODConv 模块, 用其替换 C2f 模块中原来的 Bottleneck 模块构成 C2f_ODConv 模块, 再用其替换 YOLOv8 中拼接骨干网络与颈部网络的两个 C2f 模块。与之前相比, C2f_ODConv 采用四种动态注意力相互配合, 使模型在面对不同钢材缺陷时能够更加有针对性地对卷积核进行动态调整, 更加有效地学习不同位置的钢材缺陷特征, 捕获钢材图像中的局部特征, 提高检测精度。

2.3 SENetV2 注意力机制

SENet(Squeeze-and-Excitation Network, 压缩和激励网络)^[32] 是一种通道卷积注意力机制模块, Mahendran 于 2023 年结合该模块的压缩激励操作和 ResNeXt 的多分支结构提出了 SENetV2, 进一步提升了特征表达的精细度和全局信息的整合能力。YOLOv8 颈部网络中的 C2f 模块利用两个卷积操作、Split 操作、拼接操作和 Bottleneck 模块来实现特征提取和特征融合, 增强模型的特征表达能力^[24]。为了进一步挖掘特征表达的潜力, 我们利用 SENetV2 将压缩和激励操作与多分支结构结合的特点, 进一步优化特征融合过程。

SENet 模块如图 8 所示^[33], 利用全局平均池化层将特征图 U 在空间维度上由 $H \times W \times C$ 压缩为 $1 \times 1 \times C$ 的通道描述符, 再利用两个全连接层对通道描述符进行非线性变换(激励)得到 $1 \times 1 \times C$ 的通道权重。最后将计算出来的各通道权重值与特征图 U 对应通道相乘得到重新校准的特征图 $\tilde{X}$ 。SENet 模块增加了特征通道维度的敏感度, 强化了重要特征, 抑制了一般特征。

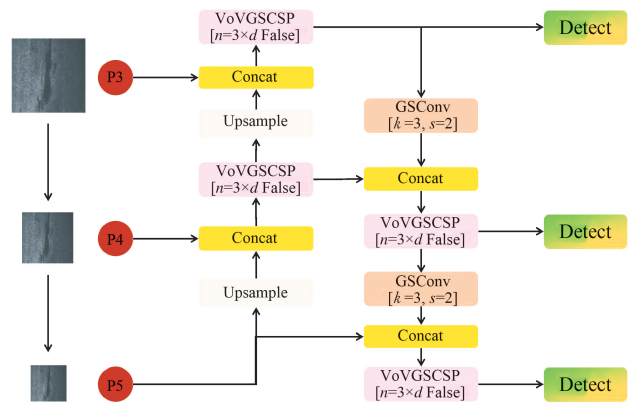


图 5 YOLOv8 中的 Slim-neck 结构

Fig. 5 Slim-neck structure in YOLOv8

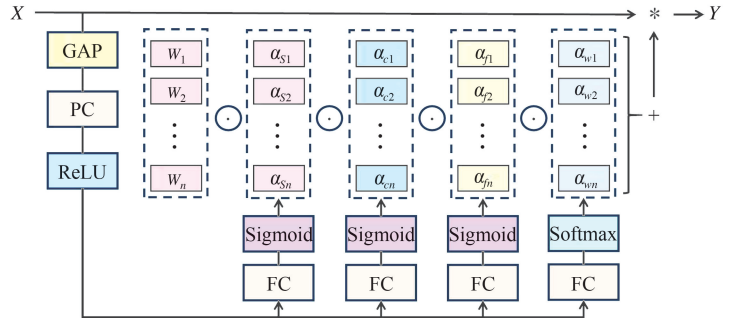


图 6 全维度动态卷积

Fig. 6 Omni-dimensional dynamic convolution

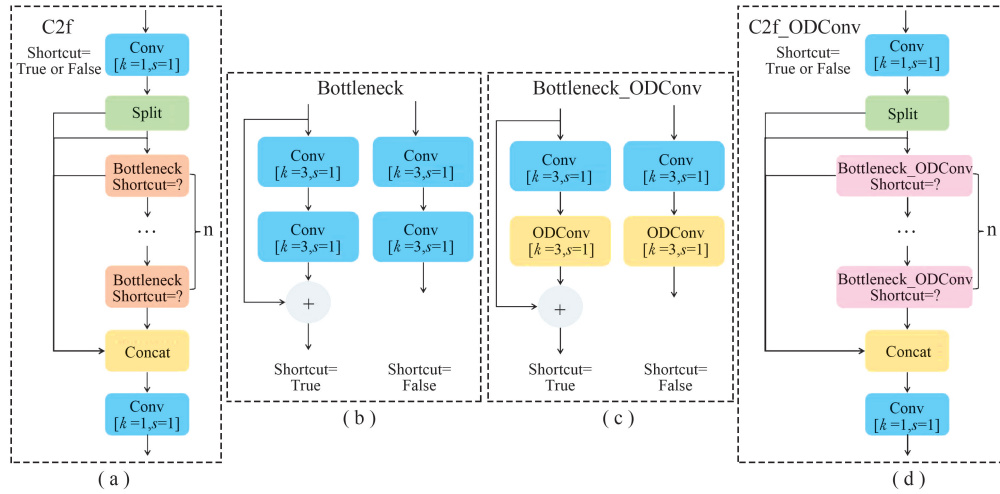


图7 改进前后的C2f模块 (a) C2f; (b) Bottleneck; (c) Bottleneck_ODConv; (d) C2f_ODConv

Fig. 7 C2f module before and after improvement (a) C2f; (b) Bottleneck; (c) Bottleneck_ODConv; (d) C2f_ODConv

SENetV2 结合 SENet 的压缩和激励操作和 ResNeXt 模块的多分支结构,采用多分支全连接层进行压缩和激励操作,最后进行特征缩放,在略微增加模型参数量的同时有效提升模型性能。图 9 为 SENetV2 与 SENet 结构对比图^[32],可以看到两种结构都能选择性地传输关键特征,但 SENetV2 模块与 SENet 相比增加了层与层之间的基数(Cardinality),聚合了多分支的特征信息。因此该模块能够实现丰富且多样化的特征学习,更擅长对复杂特征进行表示^[32]。

本文将 SENetV2 注意力机制嵌入 YOLOv8 的颈部网络(见图 2),能够使网络能够学习到输入数据的不同特征,提升特征表达的精细度,增强模型对复杂钢材缺陷特征的提取能力,进而提高检测精度。

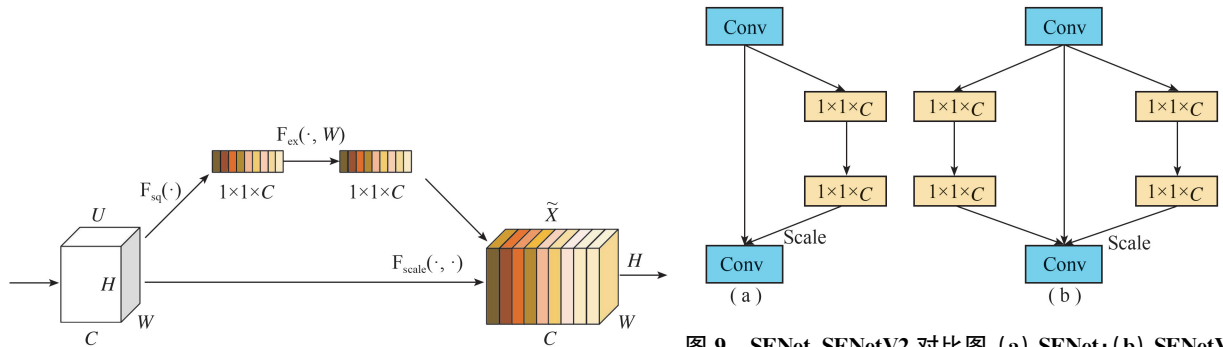


图8 SENet 模块

Fig. 8 SENet module

图9 SENet、SENetV2 对比图 (a) SENet; (b) SENetV2

Fig. 9 SENet、SENetV2 comparison diagram

(a) SENet; (b) SENetV2

3 实验结果与分析

为验证算法的有效性,本节通过实验与其他常用的目标检测算法进行对比。为探究本文算法中各模块对整个算法的贡献,针对提出的三个改进模块进行消融实验。

3.1 数据集

实验使用采用东北大学宋克臣团队制作的钢材缺陷检测数据集 NEU-DET^[34]。该数据集共有 1800 张图片,按照种类的不同划分为开裂(Crazing, Cr)、夹杂物(Inclusion, In)、斑点(Patches, Pa)、点蚀表面(Pitted_Surface, PS)、轧制氧化皮(Rolled-in_Scale, RS)、划痕(Scratches, Sc)等 6 种不同的缺陷,其中每种类型的缺陷图片各有 300 张,图 10 为六种缺陷类别的部分图片。本文按照 7 : 2 : 1 将数据集随机划分为训练集、验证集、测试集。

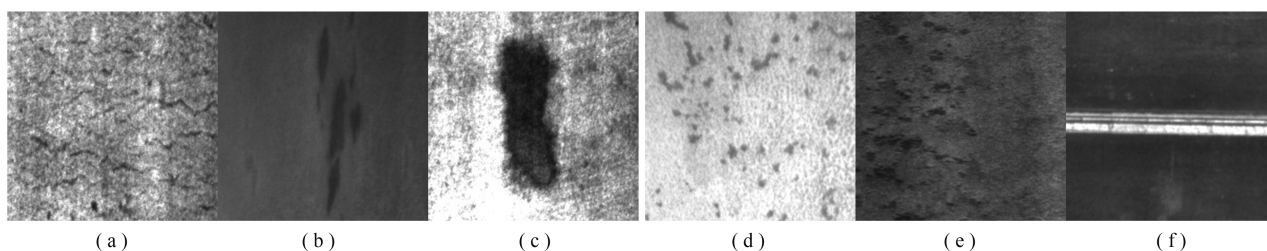


图 10 钢材表面各种缺陷图像 (a) 开裂;(b) 夹杂物;(c) 斑点;(d) 点蚀表面;(e) 轧制氧化皮;(f) 划痕

Fig. 10 Images of defects on the steel surface (a) Cracking; (b) Inclusion; (c) Patches;
(d) Pitted_Surface; (e) Rolled-in_Scale; (f) Scratches

3.2 实验环境设置

本实验软硬件配置见表 1。在模型训练阶段,设置优化器为 SGD,动量因子设置为默认的 0.937,权重衰减因子设置为默认的 0.000 5,输入图像尺寸设置为默认的 640×640 。Batchsize 设置为 8,能够使模型的波动更平稳。为保证模型的收敛,epochs 设置为 300 轮。所有模型的上述参数设置保持相同。实验环境配置如表 1 所示。

3.3 评价指标

本文以准确率(Precision, P)、召回率(Recall, R)、平均精度(Average Precision, AP)、平均精度均值(mean Average Precision, mAP)、参数量(Parameters)和浮点运算次数(FLOPs: Floating point Operations Per second)作为评价模型性能的评价指标。其中 P 、 R 、AP 和 mAP 的计算公式分别为

$$P = \frac{TP}{FP + TP} \times 100\%, \quad (2)$$

$$R = \frac{TP}{FN + TP} \times 100\%, \quad (3)$$

$$AP = \int_0^1 P(R) dR, \quad (4)$$

$$mAP = \frac{\sum_{n=1}^N AP_n}{N} \times 100\%, \quad (5)$$

式中 TP (True Positive)、 FP (False Positive)和 FN (False Negative)分别表示正确分类的正样本数量、误判为正样本的负样本数量和误判为负样本的正样本数量, N 是物体类别总数, n 代表第 n 个类别。因此,准确率 P 是指预测结果中正样本的数量占有所有被预测为正样本的比例。召回率 R 是指预测结果中正样本的数量占有所有正样本的比例,平均精度 AP 是精度和召回(PR)曲线下的面积,平均精度均值 mAP 是对所有 N 个样本类别的 AP 值的平均值,该值越高,代表模型的检测精度越高,因此该指标综合考虑了模型在所有类别中的检测性能。mAP50 表示预测框和真实框交并比阈值为 0.5 的平均精度均值,mAP50~95 表示预测框和真实框交并比阈值范围为 0.5~0.95 的 mAP 值的平均。Parameters 为参数量,指模型在训练过程中所需学习和调整的可变参数的总量,FLOPs 为浮点运算次数。

3.4 消融实验

为了证明各改进模块的有效性,本文以 YOLOv8n 为基础模型进行消融实验,使用 mAP50、参数量、FLOPs 等评价指标进行对比。其中“√”代表模型采用了此改进方式,“—”代表模型没有采用此改进方式。实验结果如表 2 所示。

表 1 实验环境配置

Tab. 1 Experiment-related configurations

Related configurations	Configurations information
Operating system	Windows11
IDE	PyCharm2022. 2. 5
Language	Python3. 11. 5
Deep learning framework	PyTorch2. 1. 0
CPU	Intel i9-13900HX
GPU	NVIDIA GeForce RTX4060
GPU accelerate model	CUDA 12. 1/cuDNNv8. 9. 7

表 2 消融实验

Tab. 2 Ablation Experiments

Algorithm	Slim-neck	C2f_ODConv	SENetV2	SENetV1	mAP50(%)	Parameters/M	FLOPs/G
YOLOv8n	—	—	—	—	77.5	3.01	8.1
YOLOv8n+ Slim-neck	✓	—	—	—	78.5	2.80	7.3
YOLOv8n+ C2f_ODConv	—	✓	—	—	78.3	3.29	7.6
YOLOv8n+ SENetV2	—	—	✓	—	78.1	3.05	8.1
YOLOv8n+ Slim-neck + SENetV2	✓	—	✓	—	79.3	2.84	7.3
YOLOv8n+ SENetV1	—	—	—	✓	77.1	3.02	8.1
GOS-YOLO	✓	✓	✓	—	79.2	3.13	6.8

由于 YOLOv8+Slim-neck 能够将浅层信息与深层信息进行有效融合,提升模型特征表达的能力,所以表 2 中显示其 mAP50 提高了 1%、参数量减少了 0.21 M、FLOPs 降低了 0.8 G。YOLOv8+C2f_ODConv 使模型在空间核尺寸、输入通道数、输出通道数、卷积核数量四个维度上学习特征,从而提高特征提取的能力,表 2 中显示其 mAP50 提高了 0.8%、参数量增加了 0.28 M、FLOPs 降低了 0.5 G。YOLOv8+SENetV2 通过多分支结构与压缩和激励操作的结合增强图像的特征表示,mAP50 提高了 0.8%、参数量仅增加了 0.28 M、FLOPs 降低了 0.5 G。YOLOv8+SENetV2 与 YOLOv8+SENetV1 相比,mAP50 提高了 1%,说明 SENetV2 的多分支结构确实在增强模型特征表达能力方面具有优势,因此提升了检测精度。YOLOv8+ Slim-neck + SENetV2 通过对 Slim-neck 和 SENetV2 的联合应用,使模型对图像特征的表达能力进一步增强,因此 mAP50 提高了 1.8%、参数量减少了 0.17 M、FLOPs 降低了 0.8 G。GOS-YOLO 在引入三种模块后,模型的特征表达和提取能力得到增强,因此 mAP50 提高了 1.7%、参数量仅增加了 0.12 M、FLOPs 降低了 1.3 G,有效实现了检测精度与模型复杂度的平衡。

本文采用热力图可视化作为验证 SENetV2 注意力机制有效性的手段,如图 11 所示,热力图均取自各网络的第 16 层。

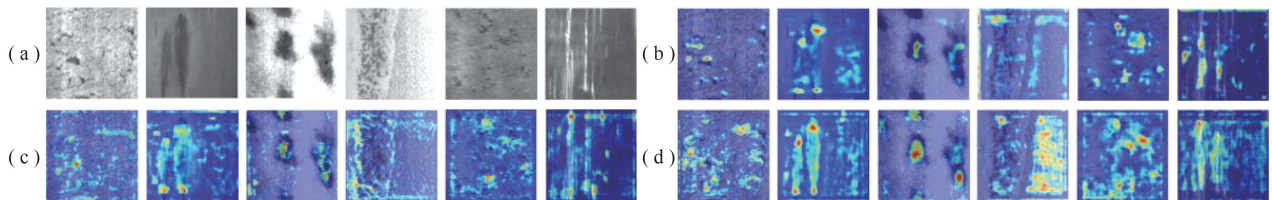


图 11 热力图对比(a) 原始图像;(b) YOLOv8n 的热力图;(c) YOLOv8n 嵌入 SENetV2 的热力图;(d) GOS-YOLO 的热力图

Fig. 11 Heatmap comparison (a) Original image; (b) The heatmap of YOLOv8n; (c) The heatmap of YOLOv8n+SENetV2; (d) The heatmap of GOS-YOLO

图 11(b)和图 11(c)相比,YOLOv8n 在嵌入 SENetV2 后对图像中关键特征的关注程度更高且覆盖范围更大。图 11(d)与图 11(c)相比,GOS-YOLO 不仅增强了对关键特征的关注,而且降低了对非关键特征的关注,从而提高了模型在特征提取方面的准确性。

3.5 对比实验

为了证明 GOS-YOLO 的性能,本文选取了几种流行的目标检测模型进行对比实验。实验结果如表 3 和表 4 所示。

由表 3 可知,与流行的目标检测算法相比,GOS-YOLO 更具优势。与原始算法 YOLOv8n 相比召回率、mAP50、mAP50~95 提升了 3.3%、1.7%、2.3%。与双阶段目标检测算法 Faster-RCNN 相比 mAP50 提高了 4.6%,mAP50~95 提高了 11.1%。与单阶段目标检测算法 SSD、YOLO 系列、RT-DETR-l 相比 mAP50 分别提高了 18.5%、12.1%、2.5%、1.4%、3.3%、11.5%,mAP50~95 分别提高了 22.7%、15%、3%、1.6%、7.2%、7.6%。在模型复杂度方面,GOS-YOLO 的参数量仅略高于 YOLOv5n 和 YOLOv8n,FLOPs

较 YOLOv8n 降低了 1.3 G。由表 4 可知,GOS-YOLO 对开裂(Cr)、夹杂物(In)、点蚀表面(PS)、这几种缺陷的检测效果最佳,且最复杂的开裂(Cr)缺陷的检测精度较 YOLOv8n 提高了 3.2%。由上述数据可以证明 GOS-YOLO 在钢材表面缺陷检测任务中优于其他算法。

表 3 总体算法性能对比实验

Tab. 3 Comparative Experiments of Algorithm Performance

Algorithm	mAP50(%)	mAP50~95(%)	P(%)	R(%)	Parameters/M	FLOPs/G
SSD	60.7	23.8	79.6	27.2	24.30	275.4
Faster-RCNN	74.6	35.4	44.5	85.3	28.30	948.2
YOLOv3-tiny	67.1	31.5	59.8	66.4	12.13	18.9
YOLOv5n	76.7	43.5	71.9	71.5	2.50	7.1
YOLOv6n	77.8	44.9	74.2	71.3	4.23	11.8
YOLOv7-tiny	75.9	39.3	68.3	72.3	6.03	13.2
YOLOv8n	77.5	44.2	75	70.8	3.01	8.1
RT-DETR-l	67.7	38.9	69.4	60.8	32	103.5
GOS-YOLO	79.2	46.5	75	74.1	3.13	6.8

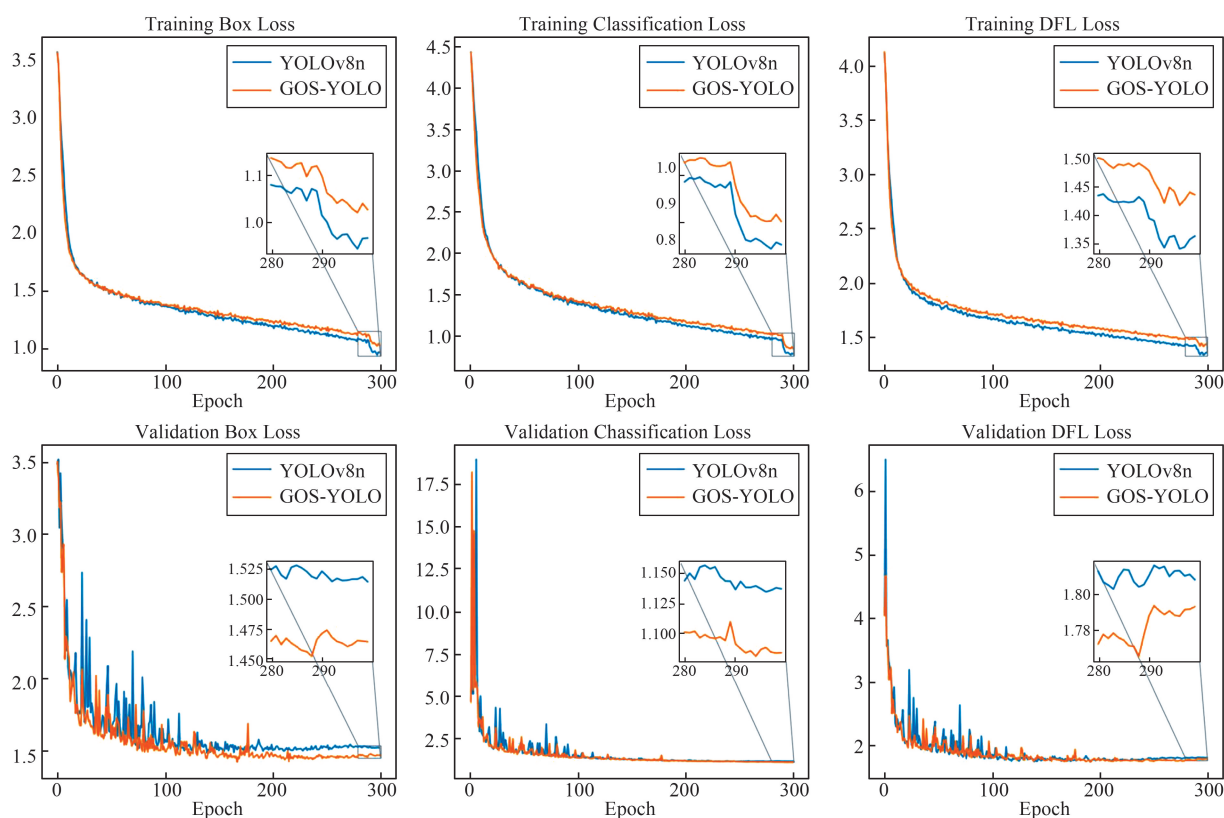


图 12 损失曲线对比

Fig. 12 Loss curve comparison

3.6 可视化分析

为了更加直观地感受 GOS-YOLO 相比于 YOLOv8n 的优势,本文给出两种算法的损失曲线、性能曲线、PR 曲线和检测结果。

图 12 为损失曲线对比。GOS-YOLO 的损失下降速度较 YOLOv8n 更快,证明了其更高效的学习能力。GOS-YOLO 损失曲线的波动较 YOLOv8n 更加平稳,证明模型更加稳定且过拟合的风险更低。

性能指标对比曲线如图 13 所示。

表 4 分类别对比实验
Tab. 4 Category comparison Experiments

Algorithm	AP50(%)					
	Cr	In	Pa	PS	RS	Sc
SSD	27.5	65.5	85.9	61.7	56.1	67.5
Faster-RCNN	46.6	72.1	92.6	76.6	65	94.4
YOLOv3-tiny	39.2	78.3	87.5	61.6	63.2	72.8
YOLOv5n	45.1	81.2	88.9	83.8	71.8	89.3
YOLOv6n	48.3	82	90.1	85.2	66.5	94.8
YOLOv7-tiny	45.7	82.2	90.5	84.5	64.8	87.7
YOLOv8n	48.2	82.4	87.2	85.1	67.9	94.2
RT-DETR-l	23.6	81.4	80.8	74.3	59.2	87.2
GOS-YOLO	51.4	82.9	90.8	87.6	68.8	93.9

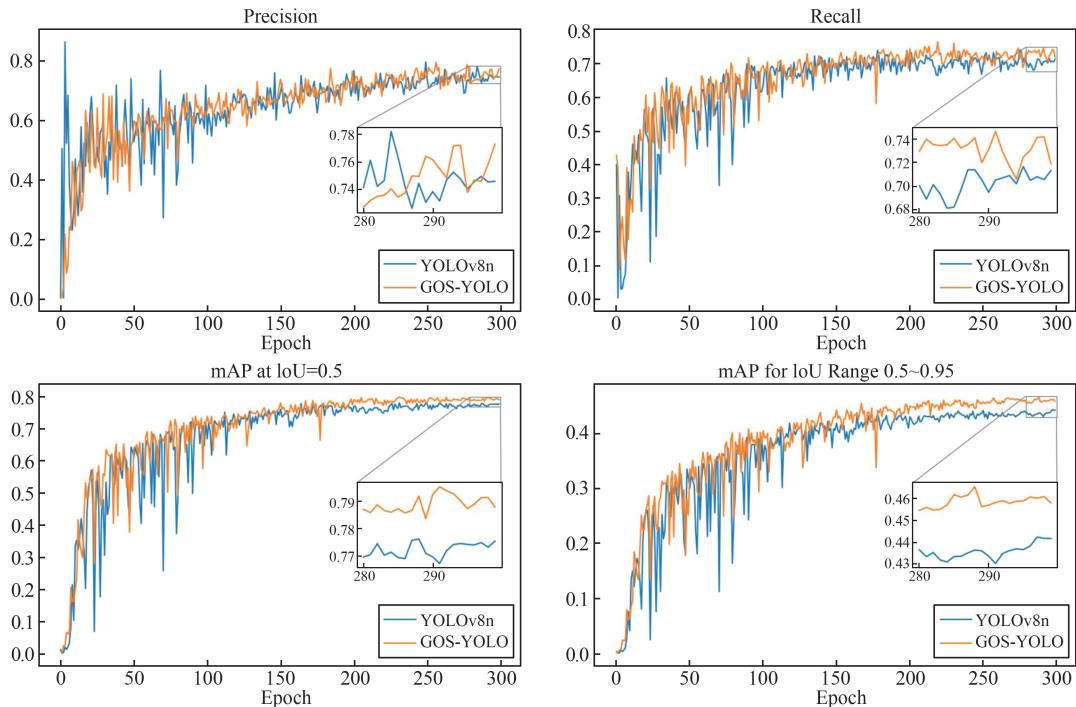


图 13 性能曲线对比
Fig. 13 Performance curve comparison

由图可见,GOS-YOLO 的性能指标上升速度较 YOLOv8n 更快,反映了 GOS-YOLO 更强的学习能力。GOS-YOLO 的性能曲线整体波动较 YOLOv8n 更加平稳,证明该模型稳定性更强。由此可见,GOS-YOLO 无论是在学习能力还是模型稳定性方面均优于 YOLOv8n。

图 14 为 YOLOv8n 与 GOS-YOLO 的 PR 曲线对比。GOS-YOLO 的 PR 曲线比 YOLOv8n 的 PR 曲线更靠近右上角。证明在综合考虑准确率和召回率的情况下,GOS-YOLO 算法的性能要优于 YOLOv8n 算法。

图 15 为检测结果对比。可以看到 YOLOv8n 对钢材表面缺陷检测的置信度相对较低,而较低的置信度容易出现漏检或误检。GOS-YOLO 对缺陷识别的置信度有明显的提升,尤其是在开裂(Cr)和轧制氧化皮(RS)上的提升最为明显。因此 GOS-YOLO 能够更准确地识别和定位缺陷,降低误检漏检出现的概率,提高整体检测效果。

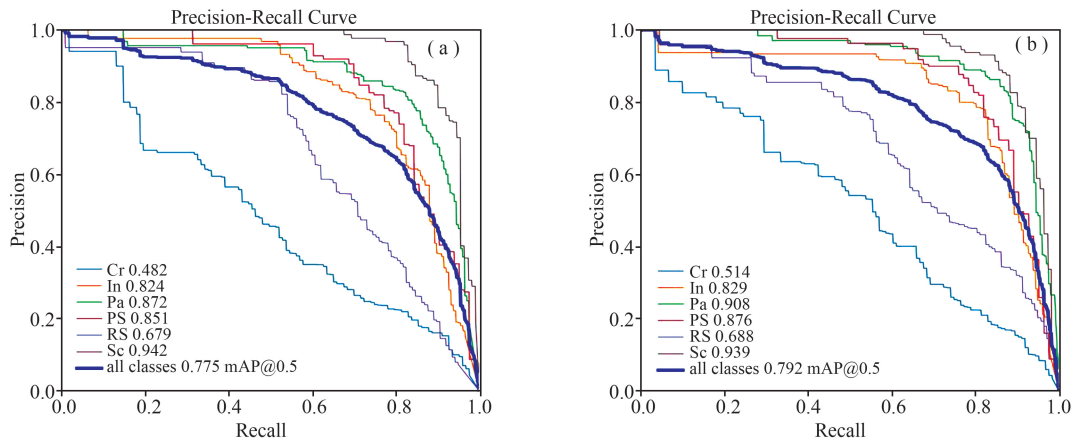


图 14 PR 曲线对比 (a) YOLOv8n 的 PR 曲线; (b) GOS-YOLO 的 PR 曲线

Fig. 14 PR curve comparison (a) PR curve of YOLOv8n; (b) PR curve of GOS-YOLO

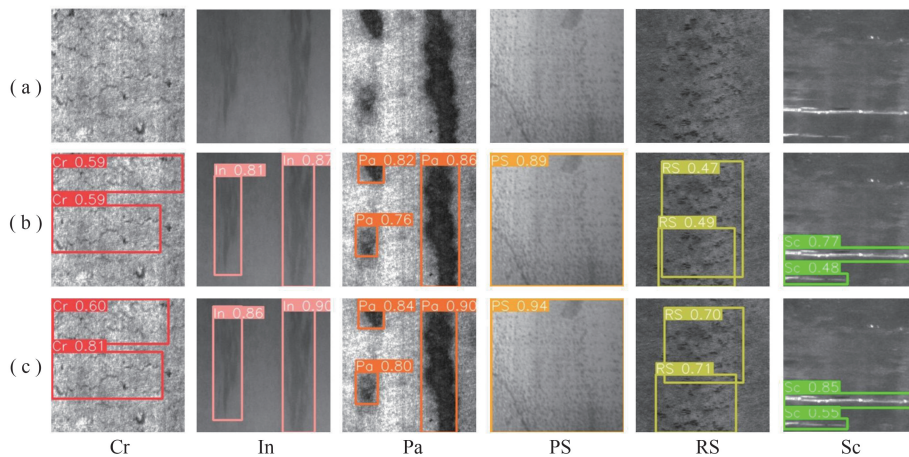


图 15 检测结果对比 (a) 原始图像; (b) YOLOv8n 的检测结果; (c) GOS-YOLO 的检测结果

Fig. 15 Comparison of defect results (a) original image; (b) the defect results of YOLOv8n; (c) the defect results of GOS-YOLO

3.7 泛化实验

为了证明 GOS-YOLO 的泛化能力, 本文使用 PASCAL VOC^[35] 挑战赛的 VOC2007 数据集对 YOLOv8n 和 GOS-YOLO 进行对比实验。参数设置与 NEU-DET 实验保持一致, 实验数据如表 5 所示。从表 5 中的数据可以看出, 在 VOC2007 数据集上, GOS-YOLO 较 YOLOv8n, 虽然召回率降低了 1.2%, 但准确率和 mAP50~95 分别提高了 1.7%、1%。GOS-YOLO 不仅在 NEU-DET 取得性能的提升, 而且在 VOC2007 数据集上也收获效果, 证明 GOS-YOLO 具有一定的泛化能力。

表 5 VOC2007 数据集上的泛化实验

Tab. 5 Generalization Experiments on VOC2007 datasets

Algorithm	mAP50(%)	mAP50~95(%)	P(%)	R(%)	Parameters/M	FLOPs/G
YOLOv8n	67.2	45.6	72.8	60.1	3.01	8.1
GOS-YOLO	67.2	46.6	74.5	58.9	3.13	6.8

4 结束语

本文提出的钢材缺陷检测算法 GOS-YOLO 进一步提高了钢铁表面缺陷检测的检测精度。该算法使用 Slim-neck 范式作为特征融合网络, 使模型在保证精度的同时降低计算复杂度。设计的 C2f_ODConv 模块能够提升模型在四个维度上对特征的关注能力。在颈部网络嵌入 SENetV2 注意力机制, 使模型能够更加高效地学习不同通道之间的关系, 增强网络捕捉通道级模式和全局知识的能力。实验结果表明, 在 NET-

DET 数据集上,GOS-YOLO 的 R 、 mAP_{50} 、 $mAP_{50\sim 95}$ 分别提高了 3.3%、1.7%、2.3%;在 VOC2007 数据集上 P 、 $mAP_{50\sim 95}$ 分别提高了 1.7%、1%。综上所述,GOS-YOLO 在不增加模型复杂度的情况下提高了检测效果且具有良好的稳定性。但是,与其他钢铁缺陷相比,开裂和轧制氧化皮等与背景区分度较低的缺陷的检测精度还相对较低,下一步我们将针对该问题进一步优化算法,提高钢铁缺陷的检测精度和速度。

参 考 文 献

- [1] 吴平川,路同浚,王炎.带钢表面自动检测系统研究现状与展望[J].钢铁,2000,35(6):70-75.
- [2] 伍麟,郝鸿宇,宋友.基于计算机视觉的工业金属表面缺陷检测综述[J/OL].自动化学报,1-24[2024-07-20].<https://doi.org/10.16383/j.aas.c230039>.
- [3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 580-587.
- [4] GIRSHICK R. Fast r-cnn[C]. // Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440-1448.
- [5] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6):1137-1149.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector[C]. // Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I 14. Springer, 2016: 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7263-7271.
- [9] REDMON J, FARHADI A. Yolov3: an incremental improvement[EB/OL]. arXiv preprint arXiv:1804.02767, 2018.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[EB/OL]. arXiv preprint arXiv:2004.10934, 2020.
- [11] LI C Y, LI L L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. arXiv preprint arXiv:2209.02976, 2022.
- [12] 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 the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475.
- [13] LI J Y, SU Z F, GENG J H, et al. Real-time detection of steel strip surface defects based on improved yolo detection network[J]. IFAC-PapersOnLine, 2018, 51(21):76-81.
- [14] KOU X P, LIU S J, CHENG K Q, et al. Development of a YOLO-V3-based model for detecting defects on steel strip surface[J]. Measurement, 2021, 182:109454.
- [15] GUO Z X, WANG C S, YANG G, et al. Msft-yolo: improved yolov5 based on transformer for detecting defects of steel surface[J]. Sensors, 2022, 22(9):3467.
- [16] WANG L, LIU X B, MA J T, et al. Real-time steel surface defect detection with improved multi-scale YOLO-v5[J]. Processes, 2023, 11(5):1357.
- [17] ZHAO C, SHU X, YAN X, et al. RDD-YOLO: a modified YOLO for detection of steel surface defects[J]. Measurement, 2023, 214: 112776.
- [18] 曹义亲,伍铭林,徐露.基于改进 YOLOv5 算法的钢材表面缺陷检测[J].图学学报,2023,44(2):335-345.
- [19] 黄友文,章劲,林志钦. YOLOv5-ODWD 钢材表面缺陷检测算法[J/OL].机械科学与技术,1-11[2024-07-20].<https://doi.org/10.13433/j.cnki.1003-8728.20240063>.
- [20] 张世强,史卫亚,张绍文,等.基于改进 YOLOv5 算法的钢铁表面缺陷检测[J].科学技术与工程,2023,23(35):15148-15157.
- [21] QU Y, WAN B Y, WANG C, et al. Optimization algorithm for steel surface defect detection based on PP-YOLOE[J]. Electronics, 2023, 12(19):4161.
- [22] 高春艳,秦桑,李满宏,等.改进 YOLOv7 算法的钢材表面缺陷检测研究[J].计算机工程与应用,2024,60(7):282-291.
- [23] 梁礼明,龙鹏威,冯耀,等.改进轻量化 VTG-YOLOv7-tiny 的钢材表面缺陷检测[J].光学精密工程,2024,32(8):1227-1240.
- [24] SONG X, CAO S Z, ZHANG J W, et al. Steel surface defect detection algorithm based on YOLOv8[J]. Electronics, 2024, 13(5):988.
- [25] 邹彦艳,曹衍芬,张馨月,等.基于 YOLOv8-DSG 的钢铁表面缺陷检测算法[J/OL].吉林大学学报(信息科学版),1-10[2024-07-20].<https://doi.org/10.19292/j.cnki.jdxxp.20240517.010>.

- [26] 马金林,曹浩杰,马自萍,等.改进 YOLOv8 的带钢缺陷检测[J/OL]. 计算机工程与应用,1-13[2024-07-20]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240710.2154.015.html>.
- [27] JOCHER G. Ultralytics[EB/OL]. [2024-7-25]. <https://github.com/ultralytics/ultralytics>.
- [28] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8759-8768.
- [29] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117-2125.
- [30] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles[EB/OL]. arXiv preprint arXiv:2206.02424, 2022.
- [31] LI C, ZHOU A J, YAO A B. Omni-dimensional dynamic convolution[EB/OL]. arXiv preprint arXiv:2209.07947, 2022.
- [32] NARAYANAN M. SENetV2: aggregated dense layer for channelwise and global representations[EB/OL]. arXiv preprint arXiv:2311.10807, 2023.
- [33] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [34] 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 Transactions on Instrumentation and Measurement, 2019, 69(4):1493-1504.
- [35] EVERINGHAM M, ESLAMI S A, VAN GOOL L, et al. The pascal visual object classes challenge: a retrospective[J]. International Journal of Computer Vision, 2015, 111:98-136.

(责任编辑:刘庆松)

(上接第 176 页)

- [25] LIU Y, LIU Z, YERUDKAR A, et al. Stabilization of probabilistic boolean networks via state-flipped control and reinforcement learning [J]. IEEE Transactions on Automatic Control, 2024,69(3):1858-1865.
- [26] LIU Y, CAO J, SUN L, et al. Sampled-data state feedback stabilization of Boolean control networks[J]. Neural Computation, 2016, 28(4): 778-799.
- [27] LU J, SUN L, LIU Y, et al. Stabilization of Boolean control networks under aperiodic sampled-data control[J]. SIAM Journal on Control and Optimization, 2018, 56(6): 4385-4404.
- [28] YU Y, FENG J, WANG B, et al. Sampled-data controllability and stabilizability of Boolean control networks: nonuniform sampling[J]. Journal of the Franklin Institute, 2018, 355(12): 5324-5335.
- [29] LI R, YANG M, CHU T. State feedback stabilization for Boolean control networks[J]. IEEE Transactions on Automatic Control, 2013, 58(7): 1853-1857.
- [30] VELIZ-CUBA A, STIGLER B. Boolean models can explain bistability in the lac operon[J]. Journal of Computational Biology, 2011, 18(6): 783-794.

(责任编辑:刘庆松)