

基于改进 YOLOv8n 的带钢表面缺陷检测方法*

王德伟 刘小芳

(四川轻化工大学计算机科学与工程学院 宜宾 644002)

摘要:针对带钢表面缺陷种类多样、特征不明显,导致漏检和错检等问题,提出一种改进 YOLOv8n 的带钢表面缺陷检测方法。首先,为适应较小尺寸目标,增加 P2 检测层来识别各类缺陷,减少漏检率,以及设计一种高效的 PConv 检测头,维持推理速度;其次,采取将 YOLOv8n 颈部中的 C2f 模块和可变形卷积 DCNv2 融合的方式,增强模型特征提取能力;此外,在骨干网络输出层引入大动态选择性模块 LSKNet,来扩大模型的感受野并提高目标检测的准确性;最后,选择 SIoU 损失函数替换 CIoU 损失函数,增强网络收敛效果,从而提高识别精度。改进后 YOLOv8n 方法在 CSU_STEEL 数据集上测试,实验结果表明,平均精度均值(mAP)mAP@0.5 比原模型提高 8.6%,达到 82.3%,体积只增加 0.5 MB。改进后的方法对带钢表面缺陷有更好检测结果,可为带钢缺陷检测方法的研究提供参考意义。

关键词:缺陷检测;YOLOv8n;LSKNet;损失函数

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.20

Surface defect detection method of strip steel based on improved YOLOv8n

Wang Dewei Liu Xiaofang

(1School of Computer Science and Engineering, Sichuan University of Science and Engineering,
Yibin 644002, China)

Abstract: Aiming at the problems of strip steel surface defects with various types and inconspicuous features, which lead to leakage and misdetection, an improved YOLOv8n strip steel surface defect detection method is proposed. Firstly, to adapt to smaller size targets, a P2 detection layer is added to identify various types of defects and reduce leakage detection, as well as an efficient PConv detection head is designed to maintain the inference speed. Secondly, a fusion of the C2f module in the neck of YOLOv8n and the deformable convolutional DCNv2 is adopted to enhance the feature extraction capability of the model. Furthermore, a large dynamics selective module is introduced in the output layer of the backbone network LSKNet, to expand the sensory field of the model and improve the accuracy of target detection. Finally, the SIoU loss function is chosen to replace the CIoU loss function to enhance the network convergence effect, thus improving the recognition accuracy. The improved YOLOv8n method is tested on the CSU_STEEL dataset, and the experimental results show that the mAP@0.5 is improved by 8.6% the original model to 82.3%, and the volume only increases by 0.5 MB. The improved method has better detection results for strip surface defects, which can provide a reference significance for the research of the defect detection method of strip steel.

Keywords: defect detection; YOLOv8n; LSKNet; loss function

0 引言

带钢作为钢铁产业核心产品之一,是基础设施建设的

原材料。尽管目前带钢产品的尺寸精度基本达到生产需要,但带钢表面质量问题仍需进一步改善。带钢在生产制造过程中,其表面产生辊印、弹性形变、浪形等缺陷^[1],不

收稿日期:2023-12-28

* 基金项目:高层次创新人才培养专项(B12402005)、四川轻化工大学人才引进项目(2021RC16)、教育部高等教育司产学研合作协同育人项目(202101038016)资助

但会影响到钢材表面的外观,而且严重降低其性能,继而影响后续基础设施的安全使用。因此,检测带钢表面缺陷对提升产品生产质量及其重要。

近年以来,以深度学习为基础的目标检测和图像分割逐渐兴起。目标检测算法可以分为两类:一类是基于区域卷积神经网络(region-based convolutional neural network, R-CNN)^[2],另一类是基于单阶段。R-CNN采用两阶段方式进行目标检测。第1步是对图像中的候选框进行提取,这些候选框覆盖了图像中可能存在的所有对象。第2步是对候选框进行划分类别和回归,最后才产生结果。相比之下,单阶段目标检测算法则直接对候选框进行处理,一次推断得到分类和回归结果。因此,这种算法通常比R-CNN更加简单高效。单阶段目标检测算法常用模型包括单次多边框检测(single shot multibox detector, SSD)^[3]和YOLO系列^[4]。两阶段的方法Faster R-CNN^[5]仅在于检测精度相对较高,而单阶段的方法优势在于模型推理速度较快和模型容易部署。

卷积神经网络(CNN)是一种深度学习模型,其卓越的特征提取能力使其在图像处理领域被广泛应用。在钢表面缺陷检测任务中,卷积神经网络可通过学习到的钢表面像素特征,对不同类型的缺陷进行有效检测。Tian等^[6]提出了一种刚表面检测器,该检测器能够定位缺陷中心并回归缺陷属性。为了提高关键点估计的准确性和平衡正负样本,引入一种新的中心度函数权重。在两种带钢缺陷数据上进行实验,解决速度和准确性之间冲突的问题。然而,速度和精度都仍然有待进一步提升。陆尧等^[7]基于Faster R-CNN构建Cascade R-CNN模型,通过使用空洞卷积来扩大特征图的感受野,而不增加参数量。同时优化特征金字塔,提高定位精度,替换损失函数来解决正负样本不平衡。这些改进使得带钢缺陷检测模型的精度得到提升,但是在速度方面并未达到预期。吴健生等^[8]对Faster R-CNN算法采用多尺度检测和融合卷积块注意力模块(CBAM)机制,抑制复杂背景的影响,使网络更专注与缺陷特征的提取。在NEU-DET数据集上测试,虽然平均精度(AP)从0.711提高到0.795,但是检测速度缓慢且模型泛化性低。王素珍等^[9]基于YOLOv5s网络提出一种名为YOLO-DWCSP-CA的算法,专门针对钢铁表面缺陷检测。该算法在骨干网络中使用深度可分离卷积,增强感受野的获取能力。此外,使用极大值抑制处理密集带钢缺陷目标,有效减低漏检率。实验结果表明,该算法的查全率提高6%,但需要进一步改善召回率,从而提高整体检测精度。赵春华等^[10]基于YOLOv7-tiny网络提出一种钢材表面缺陷检测的算法PC-YOLOv7,使用PC-ELAN结构替换主干网络部分的ELAN结构,同时利用双向特征金字塔网络(BIFPN)结构促进语义信息和特征信息融合,并引入SimCS-CA模块加强特征表示能力。改进模型在减少参数量的同时,提高对小目标缺陷的检测结果。Wei

等^[11]将SE(squeeze-and-excitation)注意力机制引入到YOLOv8网络中,对带钢表面缺陷进行检测,较好地解决建模通道间相互依赖的问题。虽然在常见的6种带钢表面缺陷中选择4种进行高精度测试,但忽略了在复杂背景下的另外两种缺陷,这限制了模型的全面性和适应范围。

针对带钢表面缺陷中小目标数量多、区域密集且分布不均、缺陷目标与背景模糊以及尺度差异大的特点,导致缺陷检测准确率低,缺陷目标漏检率偏高,同时还存在模型计算量过大和检测速度慢等挑战。提出了一种基于YOLOv8n的检测方法,做出如下改进。

1) 增加感受野较小的P2检测层,加强特征学习,准确识别密集区域微小缺陷目标。

2) 将DCNV2c与C2f模块融合作为颈部网络下采样阶段,提高网络对带钢缺陷目标关键特征的捕捉能力。

3) 将大型选择性网络(large selective kernel network, LSKNet)模块作用于主干网络的输出层之前,以扩大模型的感受野,进一步强化特征提取能力。

4) 嵌入Siou损失函数,提高模型的损失函数对梯度分配的灵活性,来减小损失自由度,增强网络的鲁棒性。

5) 将PConv与检测头融合,成为一种轻量、高效的检测头,增强检测头的尺度感知,从而更加关注待检测目标,保证模型识别的效率和准确率。

1 YOLOv8 算法

YOLO系列算法是一种单阶段检测算法,目前已优化到YOLOv8版本,其无论是从检测精度还是检测速度方面均有明显优势^[12]。为此,本文基于YOLOv8n模型优化改进,从而提升算法精度。YOLOv8结构主要由输入、Backbone骨干网络、Neck特征融合网络和Head检测模块4个部分组成。YOLOv8网络结构如图1所示。其中输入部分对图像做预处理,调整图像比例,实现Mosaic增强和锚点计算。YOLOv8的主干特征提取模块沿用DarkNet结构,不同之处在于YOLOv8网络使用C2f模块替换C3模块,针对不同尺度模型调整通道数,获得更丰富的梯度流动信息,保留了轻量级特性。通过SPPF模块对输出特征图进行处理,利用不同内核尺寸的池化对特征图进行合并,再将处理结果转移至颈部。YOLOv8的颈部依然采用“双塔结构”。特征金字塔和路径聚合网络相结合,促进了语义和定位特征的转移,进一步巩固了网络特征融合的能力,获取更丰富的特征信息,从而增强其对不同尺度目标的检测性能。检测模块包含3个Detect检测器,采用解耦头的结构进行检测,将回归分支和预测分支进行分离,加速了模型的收敛。

2 改进YOLOv8n算法

针对YOLOv8n模型对带钢缺陷小目标识别率不高,识别准确率低等问题进行模型改进。改进YOLOv8n模型结构如图2所示。改进如下5个方面:1)增加一层P2

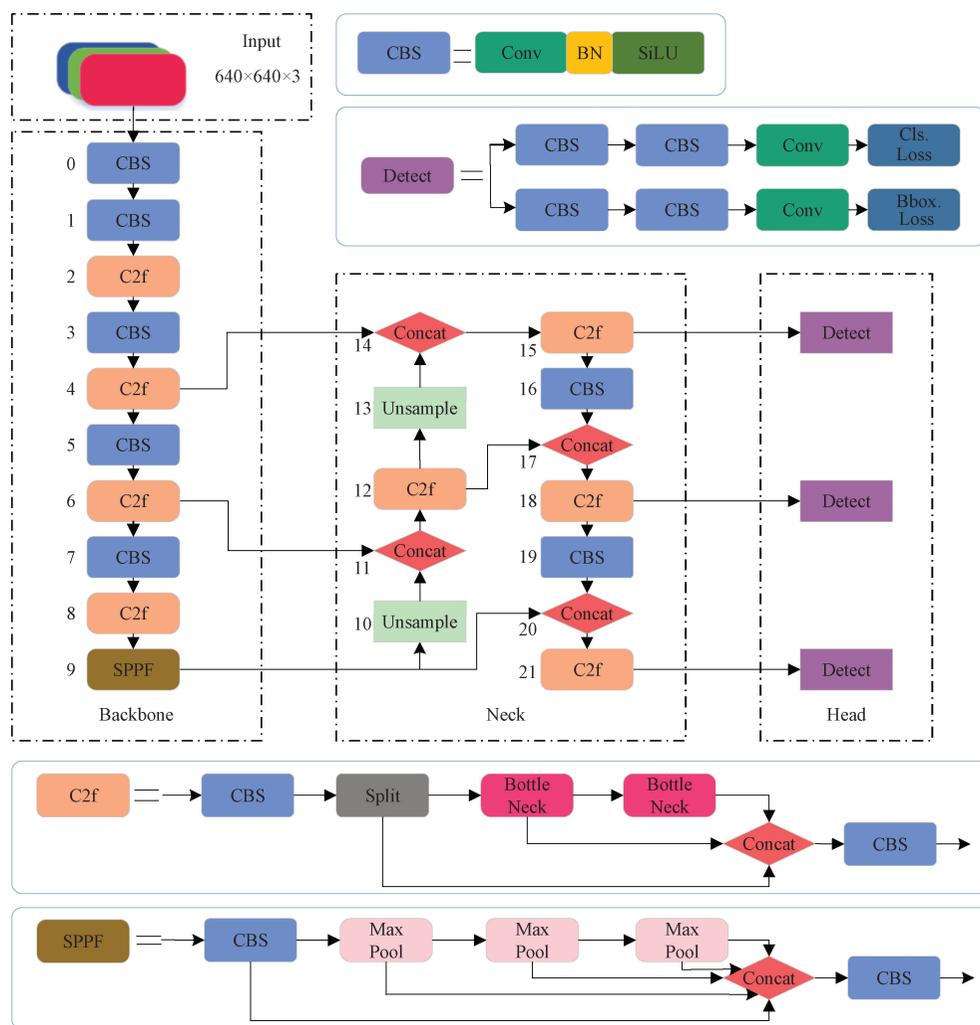


图1 YOLOv8网络结构

Fig. 1 YOLOv8n network structure diagram

检测层,加强特征学习,增强多尺度检测能力,从而降低缺陷小目标的泄漏检测率;2)在Neck端下采样阶段,实现C2f与可变形卷积DCNv2融合(C2f_DCNv2),自适应的调整感知区域,从而捕获更多有用的特征信息;3)替换原始的CIoU损失函数,利用SIoU损失函数使得训练过程更专注,加快模型收敛速度;4)在YOLOv8n的骨干网络输出层中引入大动态选择性模块LSKNet,自适应地学习不同通道的特征,以提高检测算法的准确性;5)将检测头重设计,采用 1×1 的部分卷积PConv,以降低计算参数,保证网络模型的效率。

2.1 小目标检测层

深层特征图的感受野较大,可以提供更多的语义信息,而感受野小的浅层特征图,则注重于图像细节信息^[13]。带钢表面原始图像包含多种小目标,比如辊印、氧化皮、划痕等。面对这些小目标,即使在原始YOLOv8n模型中的 80×80 的浅层特征图也难以完全的覆盖检测,存在严重漏检问题。因此,为提高带钢图像中小目标的检

测精度,在3种尺度检测层中,添加一层感受野更小的检测层P2。主干特征提取网络的P2特征图,被送至颈部的加强特征融合网络中,双塔结构分别再一次上采样与下采样操作。最后,输出生成一组大小为 160×160 的特征图,实现了更深层次的特征传递和特征融合。同时,在检测层中增加一组预测头,处理下采样次数少、分辨率高的浅层次特征图。故对于小目标的检测更加敏感,使模型更敏感地检测小目标,有效解决目标过小引起的漏检和误检问题^[14]。

2.2 可变形卷积DCNv2

传统卷积运算时,卷积核使得像素点位置固定化,并将同一卷积核应用于输入图像的各个位置。然而,在实际处理过程中,不同位置的图像可能具有不同的形变,常规卷积不能很好地你和图像不规则布局和非刚性形变,从而可能找出目标特征提取的精度下降。Dai等^[15]首次在传统卷积中提出可变形卷积(deformable convolutional networks, DCN),DCN通过添加可学习的偏移量的卷积层

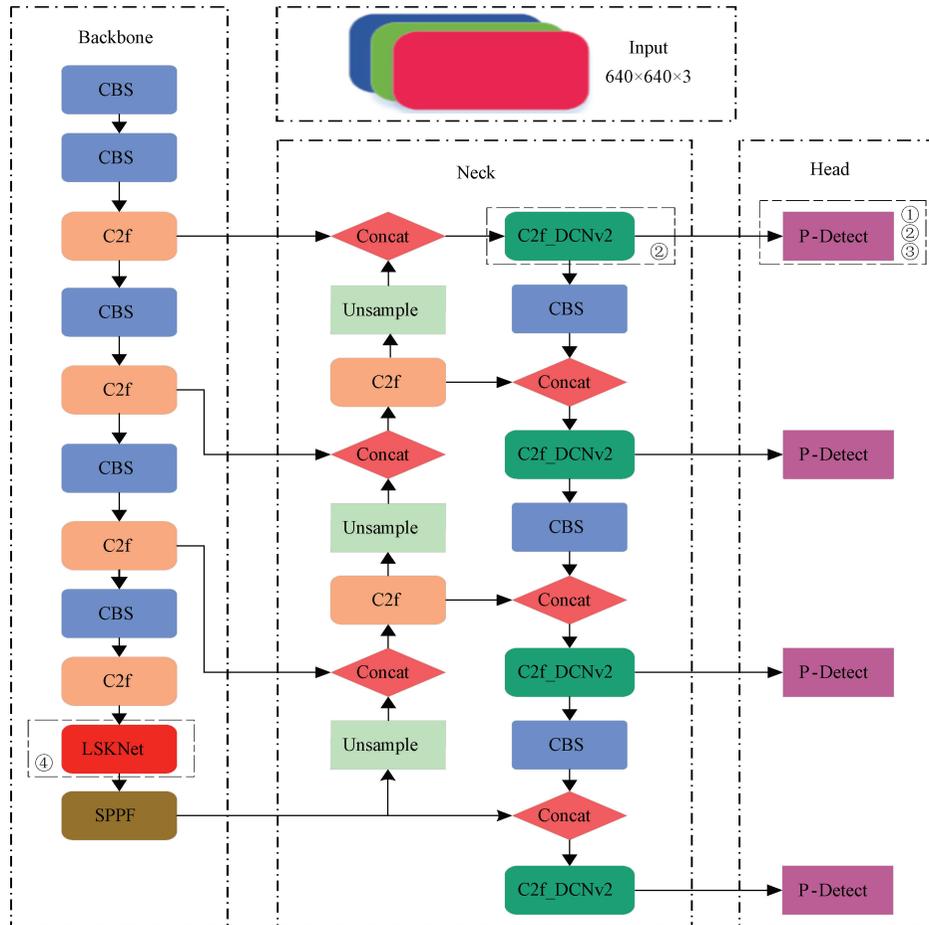


图2 改进 YOLOv8n 模型结构

Fig. 2 Improved YOLOv8n model structure diagram

和全连接层,改进传统的固定卷积核,使模型可以针对不同对象的结构形状变化对卷积核尺寸与形状进行调节。Zhu 等^[16]提出一种可全面集成可变形特性的卷积 DCNv2。DCNv2 相较于 DCN,既能调节感知输入特征的偏移量,又能调整输入特征在不同空间位置范围,使 DCNv2 能够在应对自适应能力受限的情况下,展现出较好的特性。对于定义一个参数为 1 的 3×3 的卷积核有:

$$R = \{x \mid x \in (-1, -1), (-1, 0), \dots, (1, 1)\} \quad (1)$$

式中: R 是网格区域,定义感受野大小和扩张, R 中的元素 x 记录卷积核中每个相对中心点的偏移量位置信息,利用这些偏移量对输入特征图中采样位置进行调整,使其更能适应各种特征分布情况。对于随机位置 p_0 , 输出特征图 $y(p_0)$ 为:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n \quad (2)$$

式中: $w(\cdot)$ 是加权计算函数; $x(\cdot)$ 是采用后的输入特征。当输入特征图 y 经过额外的卷积层时,便会学习每个采样点位置的相关偏移量 Δp_n 和调节参数 Δm_n 。而 p_0 是采样点中心位置, p_n 是采样点位置, Δp_n 是两者的相对位置。

与可变形相关联的还有可变形池化。可变形池化作用是提取输入数据中的关键特征,对于给定 k 个空间特征图 $y(k)$ 为:

$$y(k) = \sum_{j=1}^{n_k} x(p_{kj} + \Delta p_k) \cdot \Delta m_k / n_k \quad (3)$$

其中,每个采样点位置 p_{kj} , 可以表示为第 k 个有效采样中的第 j 个像素点的采样位置,而 n_k 则表示有效采样位置对应的像素点个数^[17]。为了获得更为准确的采样数据,需要利用双线性插值法计算 $x(p_{kj} + \Delta p_k)$, 然后根据偏移量 Δp_k 调整采样点位置,从而得到更准确的像素值。将 C2f 与 DCNv2 模块融合,适应带钢缺陷目标形状变化,使得采样位置更接近缺陷本身的形状和大小,更有利于缺陷特征的提取。可变形卷积实现过程如图 3 所示。通过对传统卷积的特征图进行额外卷积操作,生成可变形卷积的偏置矩阵,实现对卷积核大小和形状的动态调整,从而提升网络的适用性。

2.3 LSKNet 模块

注意力机制具有简单,高效的特点,能够有效强化神经表征信息。LSKNet^[18]是参考选择性核网络(selective

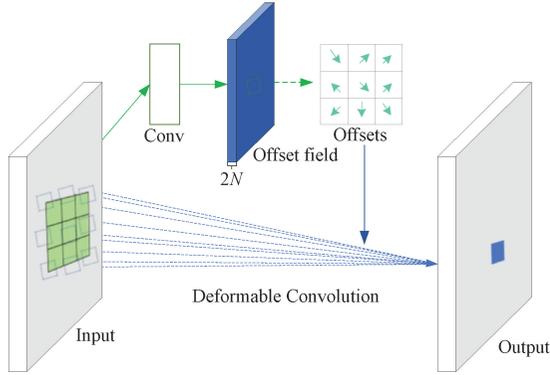


图3 可变形卷积实现过程

Fig. 3 Deformable convolutional realization process

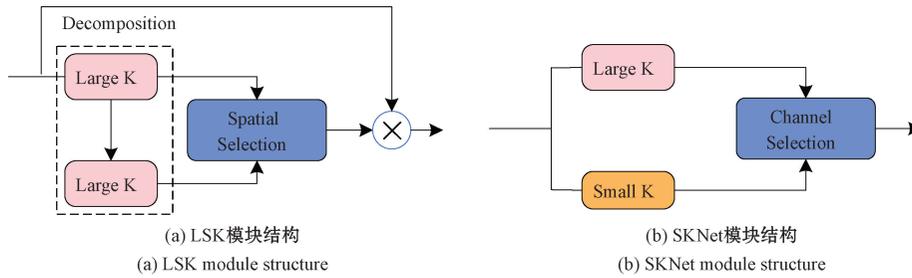


图4 LSKNet 与 SKNet 模块结构比较

Fig. 4 Comparison of LSKNet and SKNet model structure

GELU 激活、LSK 子块和另一个全连接层构成,能够根据 需要动态调整网络的感受域;一个全连接层、一个深度卷 积、一个 GELU 激活和另一个全连接层的序列构成 FFN 子块,用于信道混合和特征细化。

LSKNet 以 LSK 子块为核心组件,由一个大核卷积 序列和一个空间核选择机制组成。更大尺寸的核卷积是 通过明确将其分解为一个带有显著增长核,并引入扩张的 深度卷积序列来构建的。在序列级数中第 i 个深度卷积 核尺寸展开 k , 膨胀速率 d 和感受野 RF 参数定义为:

$$k_{i-1} \leq k_i, d_1 = 1, d_{i-1} < d_i \leq RF_{i-1} \quad (4)$$

$$RF_1 = k_1, RF_i = d_i(k_i - 1) + RF_{i-1} \quad (5)$$

通过增加卷积核的尺寸和膨胀速率,感受野的扩张速 度迅猛。同时,设定了膨胀率的上限,防止膨胀卷积引入 特征图之间的不必要间隙。这不仅使得内核的选择更加 灵活,而且显著减少了模型的复杂度。为了获得输入 X 在不同范围的上下文信息,采用一系列分解深度卷积,这 些卷积具有不同的感受野。

$$U_0 = X, U_{i+1} = F_i^{d_i}(U_i) \quad (6)$$

式中: $F_i^{d_i}(\cdot)$ 是具有卷积核 k_i 和膨胀率 d_i 的深度卷积。

假设有 N 个分解核,每个核通过 1×1 卷积层 $F_i^{1 \times 1}(\cdot)$ 进一步处理,对每个空间特征向量进行通道混合:

$$\tilde{U}_i = F_i^{1 \times 1}(U_i) \quad i = 1, \dots, N \quad (7)$$

鉴于上述背景,LSKNet 引入一种空间选择机制,通

过从各种尺度的大型卷积核中进行特征映射的空间选择, 以加强网络对最相关空间上下文区域的关注,操作步骤如 下。

1) 连接具有不同感受野范围的内核获取的特征进行 融合:

$$\tilde{U} = [\tilde{U}_1; \dots; \tilde{U}_i] \quad (8)$$

2) 利用基于信道的平均池化和最大池化进行有效空 间关系提取:

$$\mathbf{SA}_{avg} = P_{avg}(\tilde{U}), \quad \mathbf{SA}_{max} = P_{max}(\tilde{U}) \quad (9)$$

式中: \mathbf{SA}_{avg} 和 \mathbf{SA}_{max} 是平均和最大池化空间特征描述符, 能够允许不同空间描述符之间的信息交互。

3) 将空间池化特征进行融合,使用一个卷积层 $F^{2 \rightarrow N}(\cdot)$ 将池化特征(具有两个通道)变换成 N 个空间注 意力图 $\hat{\mathbf{S}}\mathbf{A}$:

$$\hat{\mathbf{S}}\mathbf{A} = F^{2 \rightarrow N}([\mathbf{SA}_{avg}; \mathbf{SA}_{max}]) \quad (10)$$

4) 对每个空间注意力图利用 sigmoid 激活函数得到 每一个分解内核独立的空间选取掩码,其表示为:

$$\tilde{\mathbf{S}}\mathbf{A}_i = \sigma(\hat{\mathbf{S}}\mathbf{A}_i) \quad (11)$$

式中: $\sigma(\cdot)$ 为 sigmoid 函数。

5) 将分解后的大核序列的特征在空间选择性上加权 对应,由卷积层 $F(\cdot)$ 融合得到注意力特征 \mathbf{S} 如下:

$$\mathbf{S} = F\left(\sum_{i=1}^N (\tilde{\mathbf{S}}\mathbf{A}_i \cdot \tilde{U}_i)\right) \quad (12)$$

6) LSK 模块最终输出是输入特征 X 与 S 之间的元素积:

$$Y = X \cdot S \quad (13)$$

LSKNet 模块利用大型卷积核多个深度可分离卷积分解的特性,来生成具有宽度感受野的特征,继而减少模型的参数量。此外,LSK 注意力通过考虑输入特征图的局部信息,来动态选择合适的卷积核,以适应各种目标类

型的上下文信息。它还动态调整其感受野,以适应不同检测对象的检测范围。空间选择机制是一种自适应权重分配方法,它从大型卷积核中动态选择最相关的特征图,并在空间上进行组合。LSKNet 模块满足了复杂背景带钢缺陷检测的要求,能够高效关注带钢缺陷相关空间区域,捕获更丰富的信息,利用空间注意提高定位能力,最终提高带钢缺陷检测成功率。

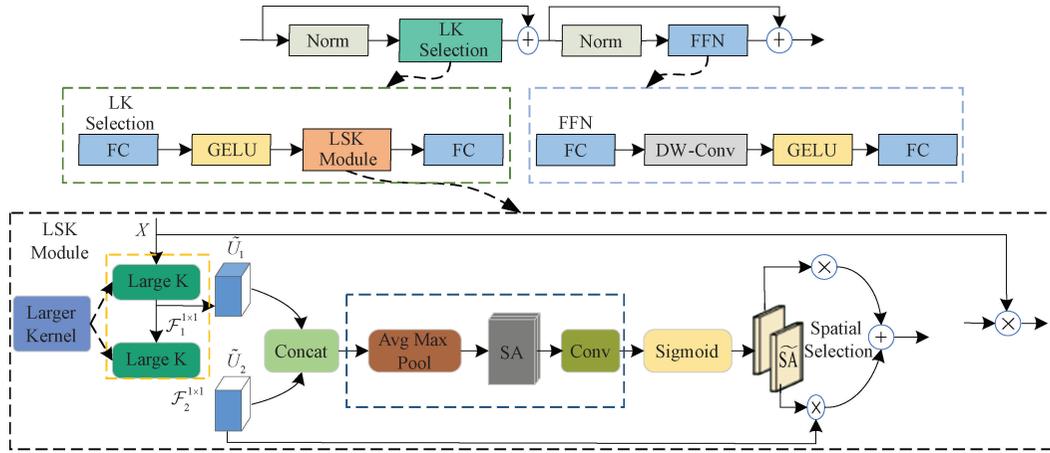


图5 LSKNet 模块结构

Fig. 5 LSKNet module structure diagram

2.4 SIoU 损失函数

预测框与真实框的纵横比,重叠区域以及距离边界以及其他目标框指标是常规目标识别损失函数建立的基础,如 GIoU、DIoU、CIoU^[20]。YOLOv8 网络使用的损失函数为 CIoU。其基于 DIoU 加入检测框的尺度损失和对长宽比的进一步考虑,使目标框返回更平稳且训练过程更专注。但是 CIoU 的缺点也比较明显,一方面难以平衡样本的选取,另一方面 CIoU 将长宽比视为损失函数中的一个惩罚因子,若实框长宽比与预测框长宽比一致,而宽度与高度取值不一致时,惩罚项无法体现两框间的真实差别。CIoU 公式为:

$$\begin{cases} L_{CIoU} = 1 - I_{IoU(A,B)} + \frac{\rho^2(A_{ctr}, B_{ctr})}{c^2} + \alpha v \\ v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \\ \alpha = \frac{v}{(1 - IoU) + v} \end{cases} \quad (14)$$

式中: A 和 B 为预测框与真实框; A_{ctr} 和 B_{ctr} 代表两个框各自中心点; $\rho(\cdot)$ 是两框中心点坐标的欧氏距离;而 c 则是限制它们的最小方框的对角线距离; w, h 是预测框的宽和高; w^{gt}, h^{gt} 则是真实框的宽和高; α 为权衡参数之一; v 用来检查衡量长宽比是否一致。 $v = 0$ 时,长宽比作为惩罚因子不起作用, CIoU 损失函数不能被完全表示^[21]。

Gevorgyan^[22] 提出 SIoU 损失函数,进一步考虑了预

测框和真实框之间的向量角度,可以使预测框快速靠近到最近的轴,随后只需要回归一个 X 或 Y 坐标,这有效地减少了自由度的总数,极大地促进训练过程,进而提高训练的速度和推理的准确性。SIoU 主要考虑了角度、距离和形状三个要素方面的内容^[23]。首先,引入角度变量 Δ ,指的是预测框和真实框的中心连接线在 X 轴方向之间的夹角,目的是为了纠正预测框的偏差角度。预测框沿着 X 轴方向逼近真实框。其次,出于预测框与真实框距离因素的考虑,引入相关变量 Δ 以解决预测框与真实框的位置偏移问题。

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) \quad (15)$$

其中:

$$\begin{cases} \rho_x = \left(\frac{b_{Cx}^{gt} - b_{Cx}}{c_w} \right)^2 \\ \rho_y = \left(\frac{b_{Cy}^{gt} - b_{Cy}}{c_h} \right)^2 \\ \gamma = 1 + 2 \sin^2 \left(\alpha - \frac{\pi}{4} \right) \end{cases} \quad (16)$$

式中: $(b_{Cx}^{gt}, b_{Cy}^{gt})$ 和 (b_{Cx}, b_{Cy}) 为两框中心坐标; c_w 和 c_h 为两框纵横坐标之差; α 为连线水平夹角;当 α 趋近 0 时, Δ 越小; α 趋近 $\pi/4$ 时, Δ 越大。

最后考虑预测框形状并引入形状变量 Ω , 其计算为:

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (17)$$

其中:

$$\begin{cases} \omega_w = \frac{|\tau - \tau^{gt}|}{\max(\tau, \tau^{gt})} \\ \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases} \quad (18)$$

式中： θ 代表对形状优化的参数值，范围取值为 $[2, 6]$ ，最佳取值是 4，避免了因过于关注损失而降低对预测框的移动。 τ, h 和 τ^{gt}, h^{gt} 分别为预测框和真实框的宽和高。

根据以上得出 SIoU 损失函数的表达式：

$$L_{box} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (19)$$

其中：

$$IoU = \frac{|B \cap B^{GT}|}{|B \cup B^{GT}|} \quad (20)$$

式中： B 是预测框面积； B^{GT} 是真实框的面积。

最终完整的 SIoU 损失函数 L 表示为：

$$L = W_{box}L_{box} + W_{cls}L_{cls} \quad (21)$$

式中： W_{box} 和 W_{cls} 分别是边界框和分类概率损失权重； L_{cls} 是 Focal loss。角度、距离和形状组成边界框损失，从而解决了预测框与真实框不匹配问题。SIoU 损失函数有效解决了 YOLOv8n 模型进行带钢表面缺陷预测时的边界框不准确的问题，显著提高模型识别不同尺寸缺陷的精度。

2.5 轻量高效检测头

相较传统卷积，PConv 部分卷积^[24]的优势是无需串联拼接及打乱、冗余计算，不会影响检测的准确性。传统卷积和部分卷积结构如图 6 所示。PConv 利用特征映射的冗余性，系统地对少数输入通道应用常规卷积同时，保持其余通道不变。其计算量 FLOP 为：

$$h \times w \times k^2 \times c_p^2 \quad (22)$$

部分卷积更小的内存访问量为：

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (23)$$

传统卷积的计算量和内存访问量分别为：

$$h \times w \times k^2 \times c^2 \quad (24)$$

$$h \times w \times 2c + k^2 \times c^2 \approx h \times w \times 2c \quad (25)$$

当典型部分比 $r = \frac{c_p}{c} = \frac{1}{4}$ ，PConv 部分卷积与传统

卷积的计算量之比为：

$$\frac{h \times w \times k^2 \times c_p^2}{h \times w \times k^2 \times c^2} = \frac{h \times w \times k^2 \times c_p^2}{h \times w \times k^2 \times (4c_p)^2} = \frac{1}{16} \quad (26)$$

PConv 部分卷积与传统卷积的内存访问量之比为：

$$\frac{h \times w \times 2c_p}{h \times w \times 2c} = \frac{h \times w \times 2c_p}{h \times w \times 2 \times 4c_p} = \frac{1}{4} \quad (27)$$

由式(26)和(27)可知，PConv 的计算量 FLOPs 仅为常规卷积的 1/16，同时内存访问量也仅仅为常规卷积的 1/4。因此，将 PConv 部分卷积融合到 YOLOv8n 的 Head 部分，构成轻量、高效的检测头，对于带钢图像中辊印微小

的缺陷能够高效检测。

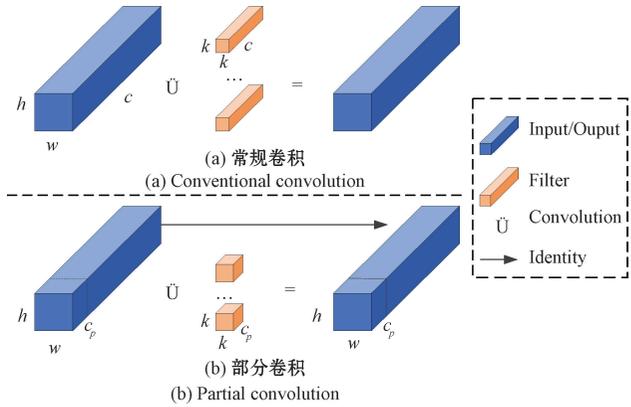


图 6 传统卷积和部分卷积结构

Fig. 6 Convolution and partial convolution structure diagrams

3 实验结果与分析

3.1 数据集

本文研究使用的是中南大学发布的首个热轧带钢表面图像数据集 CSU_STEEL^[25-26]，采集了 968 张工业生产线上热轧带钢表面图像，大小为 4096×1024 pixels。其中包含 6 种表面缺陷类型：弹性形变 (elastic deformation, el)、辊印 (roll marks, rm)、夹杂 (inclusions, in)、划痕 (scratches, sc)、浪形 (wave, wa)、氧化皮 (oxide scale, os) 共 6 类缺陷，接近 6 500 个对象实例。各类曲线示例如图 7 所示。类别缺陷差别较大，存在样本不平衡问题。通过翻转、增加噪声、改变亮度等处理方式，实现数据集扩充，倾向数据集相对更加均衡。将扩充后 2 230 张图像，按照 8 : 1 : 1 比例随机划分，训练集达到 1 784 张，测试集和验证集均为 223 张。各类缺陷扩充前后对比如图 8 所示。

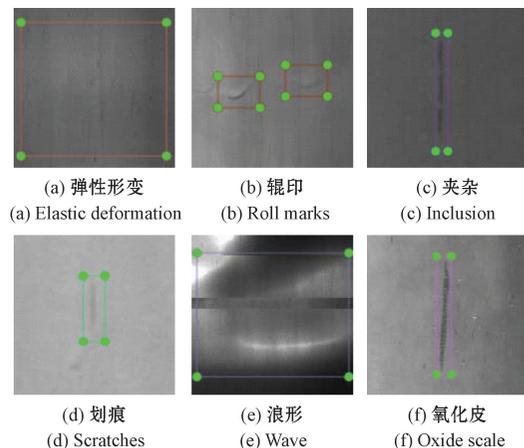


图 7 各类缺陷示例图

Fig. 7 Example diagrams of various types of defects

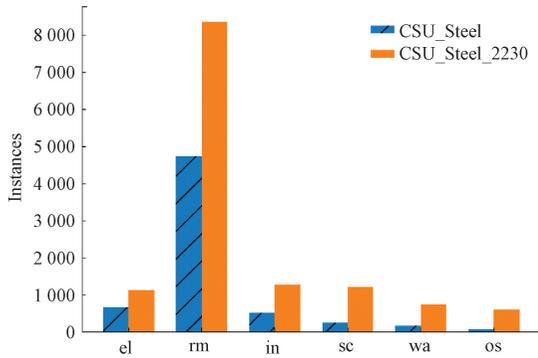


图8 各类缺陷扩充前后对比

Fig. 8 Comparison of various defects before and after expansion

3.2 实验方案

实验使用 Windows 10 操作系统, GPU 为 NVIDIA GeForce RTX3080, 使用 Python3.9 编程语言, 基于 PyTorch 2.1.1 和 CUDA 12.1 深度学习框架。实验参数配置如表 1 所示。实验方案流程如图 9 所示。

表 1 实验参数配置

Table 1 Experimental parameter configurations

参数名	参数值
图片尺寸	640×640
初始学习率	0.001
动量	0.937
权重衰减	0.0005
批处理	16
迭代次数	300

3.3 评价指标

为了缺陷检测模型的效果评估与改进, 选取平均精度均值(mean average precision, mAP)对模型性能进行评价。精准率(precision, P)描述了模型正确预测样本所占所有检测到目标比例, 而召回率(recall, R)描述了模型正确预测的样本占有所有目标的百分比。P 和 R 计算分别为:

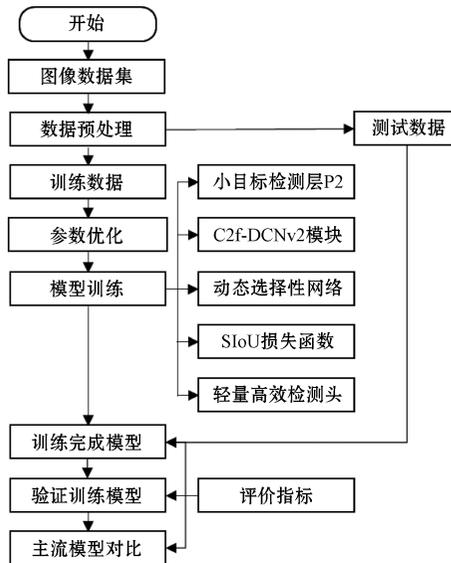


图9 实验方案流程

Fig. 9 Flow chart of the experimental method

$$P = \frac{TP}{TP + FP} \quad (28)$$

$$R = \frac{TP}{TP + FN} \quad (29)$$

式中: TP 表示目标被完全正确检测; FP 表示目标本来是错误样本, 但是被模型检测为正确样本, 即错误检测; FN 表示将目标样本的正类预测为负类, 即漏检。AP 是 P-R 曲线与坐标轴围成的面积, n 为类别数, AP 和 mAP 表示分别为:

$$AP = \int_0^1 P(R) dR \quad (30)$$

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (31)$$

3.4 消融实验

为验证各个模块的性能, 以 YOLOv8n 原始模型为基准, 构建多种模型方法, 使用 CSU_STEEL 数据集设计一系列消融实验进行对比验证。消融实验结果如表 2 所示。

表 2 消融实验

Table 2 Ablation experimental

模型	P2	DCNv2	LSKNet	SIoU	PConv	各类缺陷检测精度/%						mAP@0.5/%	mAP@0.75/%
						el	rm	in	sc	wa	os		
1						85.3	60.4	65.1	56.8	99.1	75.7	73.7	49.7
2	✓					83.6	64.8	72.5	71.6	99.5	88.2	80.0	53.4
3		✓				85.9	61.6	65.8	55.6	98.8	75.6	73.9	48.3
4			✓			87.0	67.7	74.1	63.1	99.5	82.4	79.0	55.7
5	✓	✓				91.7	66.3	70.7	65.4	99.5	88.2	80.3	53.3
6	✓	✓	✓			88.9	64.0	75.0	67.2	99.5	88.3	80.5	53.0
7	✓	✓	✓	✓		88.2	64.9	73.7	75.0	99.5	90.2	81.9	53.8
8	✓	✓	✓	✓	✓	92.2	65.8	74.0	71.9	98.2	91.7	82.3	54.8

选取每组消融实验最优结果进行分析,由表2可知,原始YOLOv8n模型在本文数据集上的mAP@0.5为73.7%,其中对于小目标缺陷检测效果较差;模型2增添小目标检测层P2后,对于小目标缺陷更加敏感,大部分缺陷精度有明显上升趋势;整体精度mAP@0.75提高了6.3%;模型5在此基础上融合可变形卷积DCNv2,可得弹性形变缺陷检测情况得到明显改善,弥补了小目标检测层对多尺度缺陷不敏感的不足。模型6进一步在骨干部分使用LSKNet模块,关注特征细节信息,使得难以区分的夹杂、划痕和氧化皮精度再度提升,整体mAP@0.5精度为80.5%;模型7更换损失函数为SIoU,加快网络收敛,mAP@0.75达到了53.8%;最终,模型8采用轻量高效结构头,尽管大尺度浪形缺陷精度下降,但是其他各类精度平稳增加,特别是辊印缺陷最高为92%,仅次于单独使用LSKNet模型5。所改进的算法在基准YOLOv8n上整体mAP@0.5提升了8.6%,mAP@0.75增加了5.1%。

在消融实验的训练过程中,为了体现模型SIoU损失的优越性。在相同实验条件下,记录了SIoU和CIoU的Loss损失变化曲线。SIoU和CIoU的Loss损失变化曲线如图10所示。

损失值的变化反应模型训练过程的优化结果,损失值越小,代表模型的鲁棒性越好。由图10可见,与CIoU损失函数相比,SIoU损失函数从最高点迅速下降,并且在后续阶段保持相对平稳,维持在较低水平,最终达到了最优

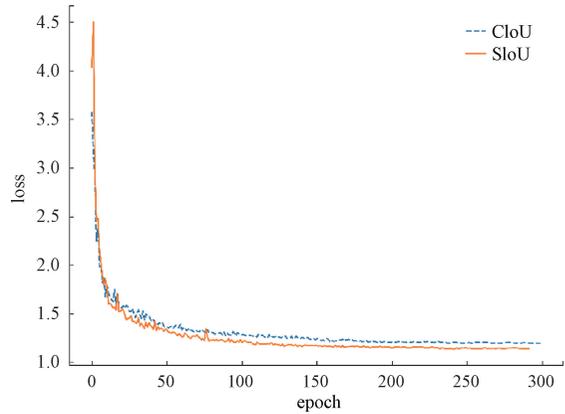


图10 SIoU和CIoU的损失变化曲线
Fig. 10 Loss curves of SIoU and CIoU

值。表明SIoU的收敛性能比相对更好,能够更有效地指导模型训练,并提高模型的性能和准确性。

3.5 对比实验

为验证改进后YOLOv8n模型在带钢表面缺陷检测方面的优越性,与目前表现性能不错的算法进行比较,比如YOLOv5、YOLOv7-tiny、Faster-RCNN和YOLOEPLUS。通过在CSU_STEEL数据集上的训练和测试,选取mAP和推理速度指标为衡量指标。不同算法实验结果如表3所示。

表3 不同算法实验结果

Table 3 Experiments with different algorithms

模型	体积/MB	参数量	计算量/GFLOPs	Backbone	mAP@0.5/%	帧率/fps
Faster-RCNN ^[5]	126	—	—	ResNet50	73.2	10.5
PP-YOLOEPLUS ^[27]	29.2	—	—	CSPRepResNet-S	72.9	17.6
YOLOv5n	3.6	1 772 035	4.2	CSPDarknet53	73.1	103.1
YOLOv7-tiny ^[9]	12.3	6 028 518	13.2	CSPDarknet53	80.2	82.1
YOLOv8n	5.9	3 151 904	8.7	Darknet-53	73.7	125
YOLOv8s	42.8	11 156 544	28.6	Darknet-53	79.0	88.5
本文	6.4	3 657 202	17.3	—	82.3	91.7

改进算法在检测精度、模型体积以及模型检测速度等性能上取得了较大提升,是相对于目前主流检测算法中精度最高,检测速度较快的目标检测模型。尽管速度、计算量和参数量有所下降,可能是加入改进模块后,网络模型复杂度上升,计算量增大,一定程度影响了网络检测速度。但是模型整体表现显著,证明了本文所提出的改进算法的可行性和有效性。

3.6 实验效果

为了更好地体现本文算法,直观感受缺陷的检测结果,将YOLOv5n、YOLOv8n模型和本文改进算法作对比。各类缺陷检测结果对比如图11所示。由图11可知,改后的算法除了在弹性形变缺陷指标略微下降外,其他缺

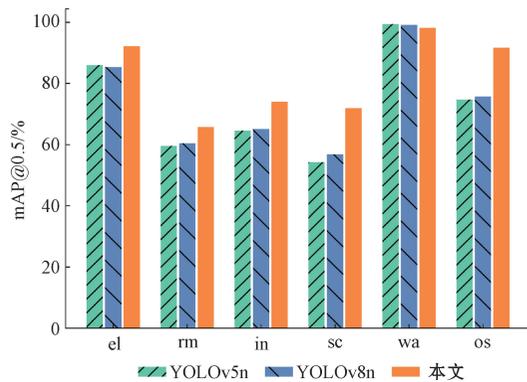


图11 各类缺陷检测结果对比
Fig. 11 Comparison chart of various test results defects

陷指标都得到了明显提升。其原因可能是模型对尺度较大的缺陷识别能力不足。在改进方法中,模型更加关注小尺度目标缺陷,从而导致了在大尺度缺陷的忽视。

为体现模型改进效果,记录图像漏检和精度情况,选取部分带钢图像检测结果和利用 HiResCAM^[28]算法进行特征可视化。部分带钢图像检测效果如图 12 所示。同时为保证其公正性,选取的 4 张图像被裁剪部分均无其他缺陷目标。图 12 中间两层分别是基准 YOLOv8n 和改进算法模型对应的检测效果图,最下层是利用 HiResCAM 可

视化算法得到改进算法模型热力图。由图 12(a)和(b)可知,基准模型均有不同类别缺陷漏检现象,原因可能是缺陷目标太小;由图 12(c)可知,YOLOv8n 基准模型错误检测缺陷,将夹杂错误识别为划痕;尽管,改进的算法(图 12(d))也存在漏检情况,但是属于可接受范围。通过热力图分析,可以观察到改进后算法模型对于特征标准更为准确,并且对于不同尺度缺陷表现出更高的关注度,能够体现缺陷的关键位置。总的来说,本文提出的改进算法在整体检测性能上展现了更优秀的效果。

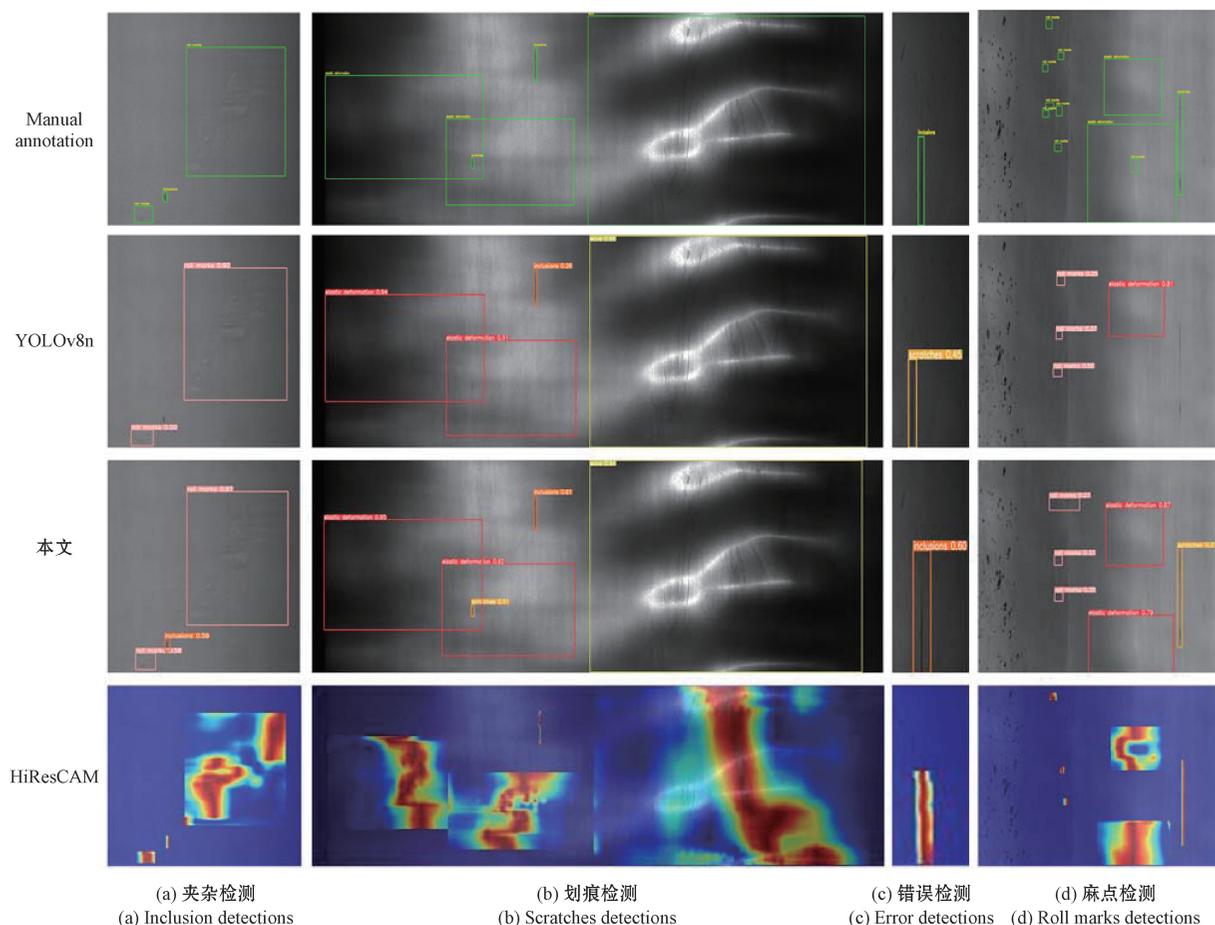


图 12 部分带钢图像检测效果
Fig. 12 Some of the stripes appear in the image

4 结论

本文针对带钢表面缺陷数据密集,缺陷目标较小,缺陷范围分布广,缺陷特征不明确,导致漏检和精度不高的问题,提出一种基于 YOLOv8n 模型改进算法。通过增添检测层、融合具有自适应能力的可变形卷积、引入动态选择核 LSKNet 模块、更换增强收敛速度的 Siou 损失函数以及重设计更轻量的检测头 PConv,实现对带钢缺陷小目标进行精确定位和分类。实验结果表明,改进后算法在密集区域的小目标检测上取得优异的效果,并且在多尺度缺陷检测上的表现也明显改善。即使在缺陷特征不明显情

况下,该算法也能正确检测。此外,在模型参数和计算量仅有限增加的情况下,算法仍实现了各项缺陷精度的提高。接下来研究中还需要采用更轻量级的体系结构,逐步减少模型参数量,提升模型泛化性。进一步提升对密集区域小目标检测能力,避免漏检现象,并且过滤冗余框,加快推理速度,以实现更高效的缺陷检测性能。

参考文献

[1] CAO J, LI Y, SUN H, et al. A survey on deep learning based visual object detection [J]. Image Graph, 2022, 27(6): 1697-1722.

- [2] 李跃, 王子铭, 李鑫林, 等. 带钢表面缺陷检测方法研究进展[J]. 钢铁研究学报, 2023, 35(8): 950-962.
LI Y, WANG Z M, LI X L, et al. Research progress on surface defect detection methods for strip steel[J]. Journal of Iron and Steel Research, 2023, 35(8): 950-962.
- [3] 江苏理工学院. 基于改进的 SSD 算法的冷轧带钢表面缺陷检测方法和系统: CN202211503165.1[P]. 2023-03-28.
Jiangsu University of Technology. Surface defect detection method and system for cold-rolled strip steel based on improved SSD algorithm: CN202211503165. 1 [P]. 2023-03-28.
- [4] 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. IEEE, 2023: 7464-7475.
- [5] 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.
LIU Q, LEI J SH. Steel surface defect detection based on improved deep network [J]. Computer Engineering and Design, 2022, 43(9): 2654-2661.
- [6] TIAN R, JIA M. DCC-CenterNet: A rapid detection method for steel surface defects[J]. Measurement, 2022, 187: 110211.
- [7] 陆尧, 薛林, 王云森, 等. 基于 Cascade R-CNN 的热轧带钢表面缺陷检测[J]. 仪表技术与传感器, 2023(8): 101-106.
LU Y, XUE L, WANG Y S, et al. Surface defect detection of hot-rolled strip steel based on Cascade R-CNN [J]. Instrument Technology and Sensors, 2023(8): 101-106.
- [8] 吴健生, 王健全, 付美霞, 等. 基于改进 Faster R-CNN 算法的带钢缺陷检测[J]. 鞍钢技术, 2022(6): 23-28.
WU J SH, WANG J Q, FU M X, et al. Strip defect detection based on improved Faster R-CNN algorithm[J]. Angang Technology, 2022(6): 23-28.
- [9] 王素珍, 吕基岳, 邓成禹, 等. 基于改进 YOLOv5 的钢铁表面缺陷检测算法[J]. 国外电子测量技术, 2023, 42(10): 43-50.
WANG S ZH, LYU J Y, DENG CH Y, et al. Steel surface defect detection algorithm based on improved YOLOv5 [J]. Foreign Electronic Measurement Technology, 2023, 42(10): 43-50.
- [10] 赵春华, 罗顺, 谭金铃, 等. 基于 PC-YOLOv7 算法钢材表面缺陷检测[J]. 国外电子测量技术, 2023, 42(9): 137-145.
ZHAO CH H, LUO SH, TAN J L, et al. Steel surface defect detection based on PC-YOLOv7 algorithm [J]. Foreign Electronic Measurement Technology, 2023, 42(9): 137-145.
- [11] WEI Z H, ZHANG Y J, WANG X J, et al. A YOLOv8-based approach for steel plate surface defect detection [J]. Metalurgija, 2024, 63(1): 28-30.
- [12] 方涛涛, 王池社, 王洁, 等. 基于 YOLOv8n 的探地雷达图像管线定位方法[J]. 国外电子测量技术, 2023, 42(11): 170-177.
FANG T T, WANG CH SH, WANG J, et al. A pipeline location method for ground penetrating radar images based on YOLOv8n[J]. Foreign Electronic Measurement Technology, 2023, 42(11): 170-177.
- [13] 侯艳丽, 唐博华. 基于改进 YOLOv5 的 PCB 缺陷检测算法[J]. 国外电子测量技术, 2023, 42(11): 24-32.
HOU Y L, TANG B H. PCB defect detection algorithm based on improved YOLOv5 [J]. Foreign Electronic Measurement Technology, 2023, 42(11): 24-32.
- [14] 杨睿宁, 惠飞, 金鑫, 等. 改进 YOLOv5s 的复杂交通场景路侧目标检测算法[J]. 计算机工程与应用, 2023, 59(16): 159-169.
YANG R N, HUI F, JIN X, et al. Improved YOLOv5s algorithm for detecting roadside targets in complex traffic scenes[J]. Computer Engineering and Applications, 2023, 59(16): 159-169.
- [15] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C]. Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2017: 764-773.
- [16] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 9308-9316.
- [17] 丁浩, 王慧琴, 王可. 基于动态形状特征提取及增强的改进 YOLOv3 火焰检测算法[J]. 激光与光电子学进展, 2022, 59(24): 37-45.
DING H, WANG H Q, WANG K. Improved YOLOv3 flame detection algorithm based on dynamic shape feature extraction and enhancement [J]. Progress in Laser and Optoelectronics, 2022, 59(24): 37-45.
- [18] PAN P, GUO W, ZHENG X, et al. Xoo-YOLO: A detection method for wild rice bacterial blight in the field

- from the perspective of unmanned aerial vehicles[J]. *Frontiers in Plant Science*, 2023, 10(14): 1256545.
- [19] 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.
- [20] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的 YOLO 目标检测综述[J]. *电子与信息学报*, 2022, 44(10): 3697-3708.
- SHAO Y H, ZHANG D, CHU H Y, et al. A review of YOLO object detection based on deep learning[J]. *Journal of Electronics and Information Science*, 2022, 44(10): 3697-3708.
- [21] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测[J]. *计算机工程*, 2023, 49(1): 41-48.
- QI L L, GAO J L. Small object detection based on improved YOLOv7 [J]. *Computer Engineering*, 2023, 49(1): 41-48.
- [22] GEVORGYAN Z. SIOU loss: More powerful learning for bounding box regression [J]. *ArXiv preprint arXiv: 2205. 12740*, 2022.
- [23] RAJ G D, PRABADEVI B. Steel strip quality assurance with YOLOV7-CSF: A coordinate attention and SIOU fusion approach [J]. *IEEE Access*, 2023, 11: 129493-129506.
- [24] 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*. IEEE, 2023: 12021-12031.
- [25] 苏姣姣, 罗旗舞. 热轧带钢表面缺陷特征提取方法综述[J]. *冶金自动化*, 2022, 46(5): 1-19.
- SU J J, LUO Q W. Summary of surface defect feature extraction methods for hot rolled strip steel [J]. *Metallurgical Automation*, 2022, 46(5): 1-19.
- [26] LUO Q, JIANG W, SU J, et al. Smoothing complete feature pyramid networks for roll mark detection of steel strips [J]. *Sensors*, 2021, 21(21): 7264.
- [27] XU S, WANG X, LV W, et al. PP-YOLOE: An evolved version of YOLO[J]. *arXiv preprint, arXiv: 2203. 16250*, 2022.
- [28] RACHEL L D, LAWRENCE C. Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks [J]. *arXiv preprint arXiv: 2011. 08891*, 2020.

作者简介

王德伟, 硕士研究生, 主要研究方向为目标检测。

E-mail: 157871813@qq.com

刘小芳(通信作者), 博士, 教授, 主要研究方向为模式识别与智能系统、数据挖掘、大数据分析、人工智能。

E-mail: lxf1969@163.com