

DOI: 10.13382/j.jemi.B2407998

基于 ESE-YOLO 的钢带表面缺陷检测研究

沈冰星 黄洪琼

(上海海事大学信息工程学院 上海 200120)

摘要:针对传统钢带表面缺陷检测方法特征提取能力不足、检测精度受限以及计算资源消耗大的问题,提出了一种基于 YOLOv8 的 ESE-YOLO 模型,旨在有效检测钢带表面缺陷。首先,为增强模型对边缘特征的提取能力,引入 EIEStem 高效前端模块,该模块通过 SobelConv 分支提取图像的边缘信息,并结合池化分支捕获重要空间信息,从而提升模型对缺陷区域的感知能力。其次,在主干网络部分,将 shift-wise convolution 与 C2f 模块融合,构建 C2f_SWC 模块,通过位移操作扩展模型的视野,增强其对上下文信息的捕捉能力,进一步提高空间特征的提取精度。此外,为优化特征金字塔网络的结构,采用高效多分支与尺度特征金字塔网络(EMBSFPN)模块,根据不同尺度的特征层自适应选择多尺度卷积核,实现对多尺度感知信息的逐步获取,并通过加权融合不同尺度特征的重要性提升检测精度,同时显著降低模型的参数量和计算成本。实验结果表明,与原始 YOLOv8n 相比,ESE-YOLO 在 NEU-DET 数据集上的平均精度均值(mAP)提高了 4.1%,参数量下降 26.8%,浮点运算量减少 64%;在 GC10-DET 数据集上,mAP 提高了 9.9%。ESE-YOLO 在显著提升检测精度的同时,大幅降低了计算资源需求,更好满足工业场景中资源受限设备的部署需求。

关键词: 缺陷检测; YOLOv8n; EIEStem; shift-wise convolution; EMBSFPN

中图分类号: TP183; TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Research on surface defect detection of steel strip based on ESE-YOLO

Shen Bingxing Huang Hongqiong

(School of Information Engineering, Shanghai Maritime University, Shanghai 200120, China)

Abstract: To address the limitations of traditional steel strip surface defect detection methods, such as insufficient feature extraction capability, restricted detection accuracy, and high computational resource consumption, this study proposes ESE-YOLO, a model based on YOLOv8, designed to effectively detect surface defects on steel strips. Firstly, to enhance the model's ability to extract edge features, an EIEStem efficient front-end module is introduced. This module utilizes a SobelConv branch to extract edge information from images, combined with a pooling branch to capture essential spatial information, thereby improving the model's perception of defect regions. Secondly, within the backbone network, shift-wise convolution is integrated with the C2f module to construct the C2f_SWC module. This integration expands the model's field of view through shift operations, enhancing its ability to capture contextual information and further improving the accuracy of spatial feature extraction. Additionally, to optimize the structure of the feature pyramid network, the EMBSFPN module is employed. This module adaptively selects multi-scale convolutional kernels based on different feature layers, enabling progressive acquisition of multi-scale perceptual information. By weighted fusion of the importance of features across different scales, the detection accuracy is enhanced while significantly reducing the model's parameter count and computational cost. Experimental results indicate that, compared to the original YOLOv8n, ESE-YOLO achieves a 4.1% improvement in mAP on the NEU-DET dataset, with a 26.8% reduction in parameters and a 64% decrease in floating-point operations. On the GC10-DET dataset, ESE-YOLO demonstrates a 9.9% improvement in mAP. Thus, ESE-YOLO significantly enhances detection accuracy while drastically reducing computational resource requirements, better meeting the deployment needs of resource-constrained devices in industrial scenarios.

Keywords: defect detection; YOLOv8n; EIEStem; shift-wise convolution; EMBSFPN

0 引言

钢材是国家建筑中必不可少的重要材料,应用广泛,特别是在基础设施建设中,它发挥着不可替代的作用。目前,大多数钢材是通过对原材料进行压力加工来生产的,以获得具有特定形状和性能的材料。在压力加工过程中,由于成本限制、设备限制、现有技术和表面粗糙度等因素,钢材不可避免地会出现缺陷^[1]。这些缺陷可能包括裂纹、夹杂物、斑块、点蚀和其他缺陷^[2]。一旦钢材中出现这些缺陷,其承受压缩、拉伸、腐蚀等因素的能力必然会受到不同程度的影响。因此,在钢铁材料生产之前,缺陷检测尤为重要。

目前,缺陷检测方法大致可分为传统缺陷检测、基于机器视觉的缺陷检测和基于深度学习的缺陷检测3种类型。传统的缺陷检测方法主要包括人工采样^[3]、红外检测^[4]和漏磁检测^[5]。人工采样方法随机选择样本并通过肉眼进行测试,但存在采样不平衡的问题,导致误差较大。此外,该方法容易受到人为因素及其他问题的影响。红外检测通过监测钢材表面因缺陷引起的温度变化来检测缺陷,但由于钢材的红外吸收能力限制,无法准确分类缺陷。漏磁检测则是通过检测有缺陷的钢表面是否存在磁性来识别缺陷。然而,该技术难以准确识别细小和狭窄的裂纹,限制了其检测能力。

机器视觉检测技术是对传统检测的改进,它减少了员工的工作量,并且不受环境影响。基于机器视觉的缺陷检测方法主要分为基于图像预处理的缺陷检测、分类器、特征提取和图像分割4大类。机器视觉检测方法在识别钢材表面缺陷的任务方面取得了巨大进步,比传统的缺陷检测方法更加高效和实用。但是,它们普遍存在多种缺陷类型分类不佳、需要人工提取特征、实时表现良好但准确率差等问题。

因此,基于深度学习的缺陷检测技术已成为一个重要的研究方向。深度学习通过输入数据并自动提取特征,不仅保留了无需人工操作和避免环境影响的优势,而且省去了手动特征提取的步骤。这种端到端的建模架构简化了工业生产流程。主流的深度学习目标检测算法包括两阶段方法^[6],如区域卷积神经网络(region-based convolutional neural network, R-CNN)^[7]、Fast R-CNN^[8]、Faster R-CNN^[9]和 Mask R-CNN^[10],以及单阶段方法,如单次多框检测器(single shot multiBox detector, SSD)^[11]和 YOLO(you only look once)^[12],以及基于 Transformer 架构的目标检测算法的检测器(detection transformer, DETR)^[13]。两阶段目标检测算法具有较高的检测精度,但由于在检测过程中生成一系列候选框,其检测速度相较于单阶段算法较慢。单阶段目标检测算法则将检测任

务转化为回归问题^[14]。Zhao 等^[15]提出了一种基于可变形卷积和多尺度融合的 Faster R-CNN 算法,尽管其模型在检测钢材表面缺陷方面表现良好,但在检测速度上难以满足工业需求。吴健生等^[16]对 Faster R-CNN 算法采用多尺度检测和融合卷积块注意力模块(convolutional block attention module, 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,虽然平均精度从 0.711 提高到 0.795,但是检测速度缓慢且模型泛化性低。

单阶段目标检测算法虽然精度相对较低,但其检测速度较快,非常适用于钢材表面缺陷的工业检测任务。YOLO 算法由 Redmon 等^[17]于 2016 年提出,是一种将检测问题转化为回归问题的单阶段算法,支持端到端的缺陷检测。经过 YOLO9000^[17]、YOLOv3^[18]和 YOLOv4^[19]等版本的演化。在 YOLOv3 的基础上,YOLOv5 在输入端引入了 Mosaic 数据增强、自适应锚框计算和自适应图片缩放;骨干网络采用 C3 结构与 SPPF 结构;颈部设计为特征金字塔网络(feature pyramid network, FPN)^[20]和路径聚合网络(path aggregation network, PAN)^[21]结构;检测端的损失函数修改为 GIOU^[22]。这些改进提升了检测精度,但在钢材缺陷检测方面仍有进一步优化的空间。

在这方面,Zhao 等^[23]提出的 RDD-YOLO 通过设计双特征金字塔网络(DFPN),增强了特征的融合,提高了特征的丰富性和网络的深度,采用解耦头分离回归和分类任务,提高了检测精度,但在一些模糊和小缺陷的检测上,RDD-YOLO 的性能需要进一步提高。朱成杰等^[24]基于 YOLOv8 进行改进,通过引入 Focal Modulation 和 C2f-MB 结构,增强了模型对多尺度特征的表达能力,虽然通过加入归一化瓦瑟斯坦距离(normalized Wasserstein distance, NWD)损失函数,提升了模型对小目标缺陷的检测能力,但在某些类别的精度上存在偏低的问题,可能需要进一步优化。李思思等^[25]设计的轻量化 FCSP block 增强了主干网络和颈部网络的特征提取和融合能力,提高了模型对缺陷的定位能力,并且通过引入可学习参数促进动态特征融合,增强了小目标浅层特征的表达能力,最后在光照减弱的 NEU-DET 数据集和 GC10-DET 数据集上的实验结果表明,模型具有良好的泛化能力。虽然这些检测方法在检测精度上相比传统的缺陷检测和机器视觉检测有所提升,但其检测方法对特定类型缺陷的检测能力不足以及模型参数和计算量的增加。

综上所述,针对现有算法在工业环境中应用时面临的边缘特征提取能力不足和钢材缺陷检测计算成本高等问题,本文提出了一种基于 YOLOv8n 算法的钢带表面缺陷检测新算法—ESE-YOLO,具有如下优势。

1) 边缘特征提取能力的提升。本文引入了 EIEStem 高效前端模块,通过 SobelConv 分支和池化分支精准提

取图像的边缘信息及关键空间信息,显著增强了模型对边缘特征的提取能力。这一改进对于检测细小、模糊等难以识别的缺陷具有极为重要的意义,而传统算法在边缘特征提取方面往往存在明显的不足。

2) 上下文信息捕捉的高效性。在主干网络架构中,本研究巧妙地将 shift-wise convolution 与 C2f 模块相结合,构建了 C2f_SWC 模块。借助位移操作,模型能够捕捉更广泛的上下文信息,从而显著提升空间信息的捕捉能力和检测精度。这一方法与现有算法在上下文信息捕捉方面的传统手段形成了鲜明对比,展现了其独特的优势。

3) 多尺度特征融合的优化。本文采用了高效多分支与尺度特征金字塔网络 (efficient multi-branch & scalar FPN, EMBSFPN), 根据不同尺度的特征层灵活选择合适的多尺度卷积核,逐步获取多尺度感知信息。同时,通过自适应选择与加权融合不同尺度特征的重要性,不仅提升了检测精度,还有效减少了模型的参数量和计算成本。这一优化成功解决了现有算法在多尺度特征融合时面临的计算成本高和参数量大的问题,为模型的高效运行提供了有力保障。

1 YOLOv8 模型结构

YOLOv8 是基于 Ultralytics 在前几代 YOLO 成功经验

的基础上构建的,融入了多个升级和创新功能,旨在进一步提升性能与灵活性。与前代相比, YOLOv8 引入了全新的主干网络、无锚点检测头部以及新的损失函数等技术创新,从而支持跨平台操作,从 CPU 到 GPU 均可高效运行。根据模型的深度和特征图大小, YOLOv8 提供了多个版本,包括 YOLOv8-s、YOLOv8-n、YOLOv8-m、YOLOv8-l 和 YOLOv8-x,覆盖了 5 种不同的配置。

YOLOv8 的网络结构可以分为 Input、Backbone、Neck 和 Head 4 个主要模块。YOLOv8 网络结构如图 1 所示。Input 模块负责将输入图像按指定比例调整至训练所需的大小,并执行包括缩放和色调整等预处理操作。Backbone 模块则专注于提取图像的核心特征,包含卷积模块 C2f、替代 C3 模块的 CSPLayer_2Cnov 以及 YOLOv5 中使用的空间金字塔池化模块 (SPPF)。Neck 模块通过结合 FPN 和“双流 FPN”路径聚合结构,有效地整合不同尺度的特征,显著提升了效率和速度。Head 模块则与 YOLOv6 和 YOLOX 中的结构相似,但相较于 YOLOv3、YOLOv4 和 YOLOv5 所使用的解耦式检测头, YOLOv8 采用了耦合式检测头。

此外, YOLOv8 继续使用 3 个输出分支,但每个分支进一步细分为两个部分,分别负责边界框的分类和回归。考虑到模型的整体检测能力,实验选用了 YOLOv8-n 作为基线模型,并将在此基础上进行改进。

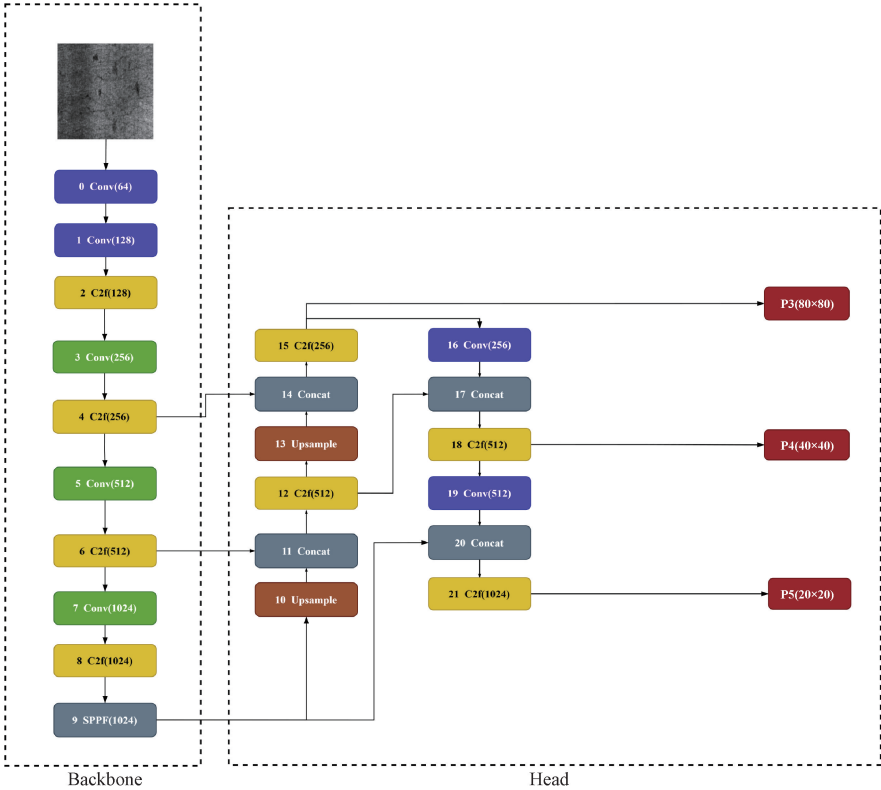


图 1 YOLOv8 网络结构

Fig. 1 Architecture of YOLOv8 network

2 钢带表面缺陷检测模型 ESE-YOLO

2.1 网络结构

钢带表面缺陷检测在实际应用中至关重要,它能够实时、准确地识别和分类钢带表面的各种缺陷,从而提高

产品质量、降低生产成本、确保生产安全,并且通过智能化的检测流程,提升生产效率和自动化水平。针对钢带表面缺陷检测在实际应用中特征提取能力不足,以及限制了其在计算资源受限的设备上部署的问题,本文在 YOLOv8 的基础上,设计了一种有效的钢带表面缺陷检测模型 ESE-YOLO,网络结构如图 2 所示。

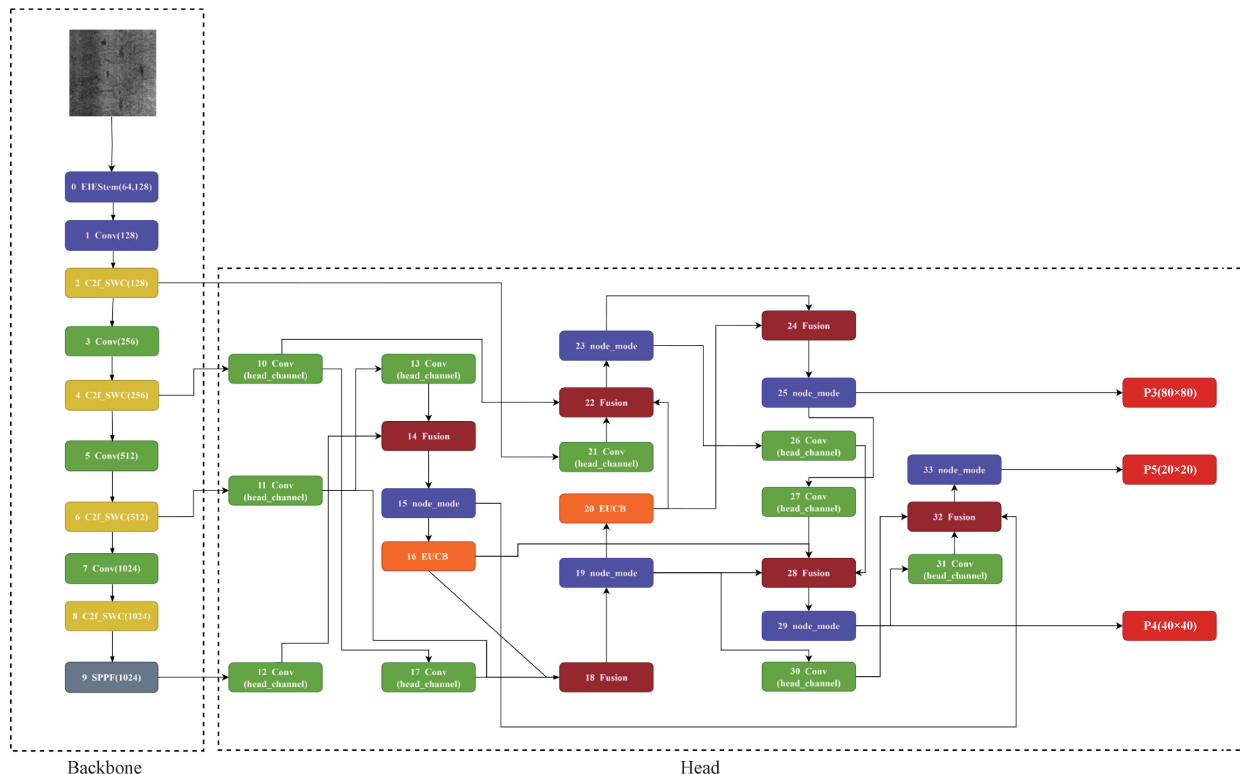


图 2 ESE-YOLO 网络结构

Fig. 2 Architecture of ESE-YOLO network

ESE-YOLO 模型结构通过首先引入了 EIEStem 这一高效的图像处理模块,增强模型对边缘细节的识别能力。其次,在模型的主干网络中,将 shift-wise convolution 与 C2f 模块相结合,创建了 C2f_SWC 模块。这种设计使得模型能够通过位移操作捕捉更丰富的上下文信息,进而提高空间特征的识别精度。最后,通过应用 EMBSFPN,模型能够根据各特征层的不同尺度选择合适的卷积核,逐步整合多尺度的特征信息。此外,模型还能自适应地选择和融合不同尺度的特征,这不仅提高了检测的准确性,还有效降低了模型的参数数量和计算开销。

2.2 EIEStem 模块

YOLOv8 由于卷积操作可能导致信息丢失,特别是在处理小目标时,这种信息丢失会大幅度降低检测能力。小目标在图像中占据的像素较少,如果特征图过小,可能会导致小目标的特征被过度压缩或丢失,从而影响边缘特征的准确提取。针对 YOLOv8 边缘特征提取能力不足

的缺点,本文在特征提取的首层引入了 EIEStem 模块。EIEStem 模块通过结合 Sobel 算子和池化操作,充分利用了边缘信息和局部特征,从而增强了特征提取的多样性和鲁棒性。SobelConv 模块通过 Sobel 算子提取图像的梯度特征,专注于边缘检测,Sobel 算子分为水平和垂直两个方向的梯度计算,其卷积核如下。

水平方向(G_x):

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (1)$$

垂直方向(G_y):

$$K_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (2)$$

假设图像的像素值为 $I(x, y)$, 则梯度计算公式为:

$$G_x = K_x \cdot I \quad (3)$$

$$G_y = K_y \cdot I \quad (4)$$

通过 G_x 和 G_y 计算梯度幅值和方向:

梯度幅值:

$$G = \sqrt{G_x^2 + G_y^2} \quad (5)$$

通常为了简化计算,也可以用绝对值代替:

$$G = |G_x| + |G_y| \quad (6)$$

梯度方向:

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (7)$$

使用 3D 卷积层来实现逐通道的边缘检测,并且不参与训练过程,减少了计算复杂度^[26]。SobelConv 网络结

构如图 3 所示。EIEStem 模块通过将 Sobel 特征和池化特征进行融合,不仅保留了细节信息,还能够通过卷积层进一步提取高层次特征。整体结构能够在多尺度上处理图像的局部特征,对于检测任务具有很好的性能,且减少了计算量。EIEStem 网络结构图如图 4 所示。

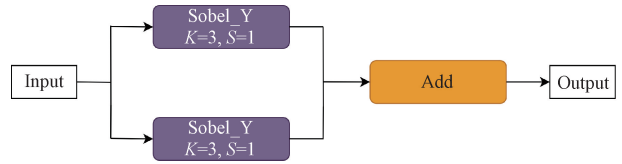


图 3 SobelConv 网络结构

Fig. 3 Architecture of SobelConv network

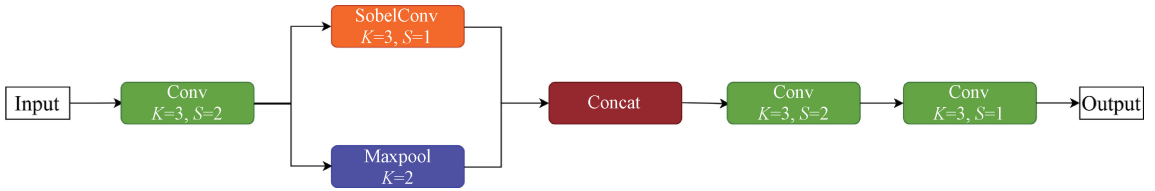


图 4 EIEStem 网络结构

Fig. 4 Architecture of EIEStem network

2.3 C2f_SWC 模块

YOLOv8 因其传统卷积神经网络 (convolutional neural network, CNN) 结构的局限性,尽管 YOLOv8 采用了 C2f 结构来增加网络的感受野并改善梯度流动,但其骨干网络仍然受限于局部感受野的限制,无法从图像的其他位置捕获信息。因此本文将原始 C2f 与 shift-wise convolution 组合,以进一步增强网络提取上下文空间信息的能力。shift-wise convolution 是通过对输入特征图的每个通道进行独立的卷积处理,增强特征的表示能力^[27]。Shift-wise Convolution 的计算公式可以表示为:

$$Y(i, j, c) = \sum_{p=-\frac{KH}{2}}^{\frac{KH}{2}} \sum_{q=-\frac{KW}{2}}^{\frac{KW}{2}} X(i+p, j+q, c) \cdot K(p+\Delta p, q+\Delta q, c) \quad (8)$$

式中: $Y(i, j, c)$ 是输出特征图中的像素值; $\Delta p, \Delta q$ 是由

移位操作产生的偏移量,控制卷积核的相对位置或感受野的移动; p 和 q 分别表示卷积核在水平和垂直方向的索引; $X(i+p, j+q, c)$ 是输入特征图的像素值,表示在坐标 (i, j) 处周围的局部区域; $K(p+\Delta p, q+\Delta q, c)$ 是经过移位的卷积核,表示在应用偏移量后,对应位置的权重。

C2f_SWC 的优点在于通过引入可重参数化的大卷积核和膨胀卷积,有效提升了模型的感受野,并减少了计算量。Bottleneck_SWC 类通过使用 Reparam LargeKernelConv 代替标准卷积层,能够在保持卷积效果的同时,降低计算开销。Bottleneck_SWC 网络结构如图 5 所示。而 C2f_SWC 类通过堆叠多个 Bottleneck_SWC 模块,增加了网络的深度,从而增强了特征提取的能力。C2f_SWC 网络结构如图 6 所示。

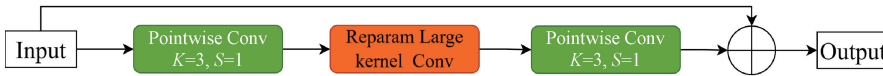


图 5 Bottleneck_SWC 网络结构

Fig. 5 Architecture of Bottleneck_SWC network



图 6 C2f_SWC 网络结构

Fig. 6 Architecture of C2f_SWC network

2.4 EMBSFPN 特征金字塔网络

YOLOv8 在处理小目标时的效果不佳,因为 YOLOv8 将输入图像分成了较小的网格,每个网格只预测一个目标,因此难以处理尺度变化较大的目标。并且 YOLOv8 相对于其他目标检测算法而言,训练速度较慢,由于其骨干网络较深且参数较多,训练过程需要较长的时间。为解决上述问题,本文采用 EMBSFPN 特征金字塔网络,EMBSFPN 网络结构如图 7 所示。EMBSFPN 具有多尺度高效卷积模块和全局异构核选择机制。具有较大感受野的网络更适合检测较大的物体,反之,较小尺度的目标则从较小的感受野中受益,因此在 FPN 阶段,对于不同尺

度的特征层选择不同的多尺度卷积核以适应并逐步获得多尺度感知场信息^[28]。EMBSFPN 采用 BiFPN 中的多尺度特征加权融合,把 Concat 替换成 Add 来减少参数量和计算量的情况下,还能通过不同尺度特征的重要性进行自适应选择加权融合^[29]。EMBSFPN 采用高效上采样模块 置信区间上界(exploration via upper confidence bound, EUCB),EUCB 首先使用上采样操作将特征图尺寸放大,然后通过应用深度卷积和批量归一化来增强这些放大的特征图^[29],因此在保持特征图的空间分辨率和细节信息的同时,显著降低了计算复杂度,提高了运算效率。

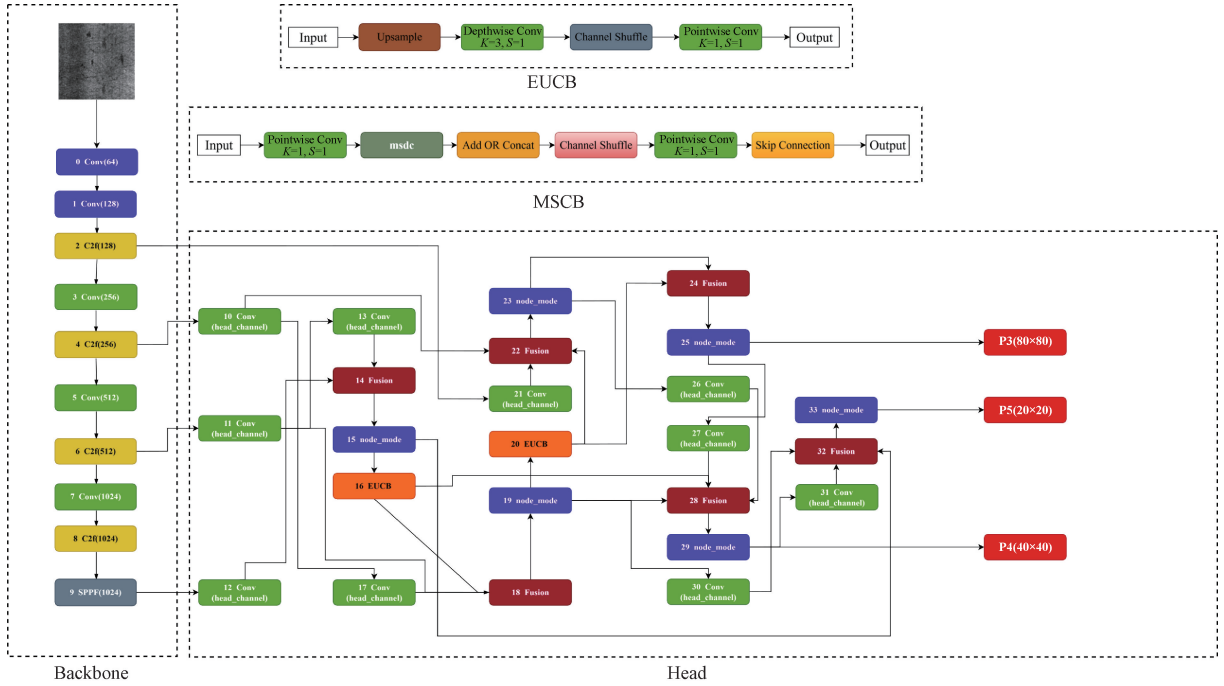


图 7 EMBSFPN 网络结构
Fig. 7 Architecture of EMBSFPN network

3 实验设计与数据分析

3.1 实验环境及参数设置

本实验环境使用 Windows 11 操作系统,16 G 内存,CPU 采用 13thGenIntel(R) Core(TM) i5-13600KF,显卡采用 NVIDIA GeForce RTX4060ti,具有 8 GB 显存,PyTorch2.0.1 作为深度学习框架,Python 版本为 3.8.18,CUDA 版本为 11.7,cuDNN 版本为 11.7。实验参数设置如表 1 所示。

3.2 实验数据集

研究采用由东北大学(NEU)发布的表面缺陷数据集 NEU-DET 和真实工业中收集的表面缺陷数据集 GC10-DET。

表 1 实验参数配置

Table 1 Experimental parameter settings	
实验参数	参数量
Epochs	150
Input image size	640×640
Batch size	8
Initial learning rate	0.01
Optimizer	SGD

NEU-DET 收集了热轧钢带的 6 种典型表面缺陷,即轧制氧化皮(RS)、斑块、开裂、点蚀表面、内含物和划痕。该数据库包括 1 800 个灰度图像,6 种不同类型的典型表面缺陷,每一类缺陷包含 300 个样本,如图 8 所示。数据集图像大小为 200×200,总共 1 800 张图像被随机分成 8:1:1 的比例,以创建 NEU-DET 训练、测试和验证集,即 1 440 个训练样本、180 个测试样本和 180 个验证样本。

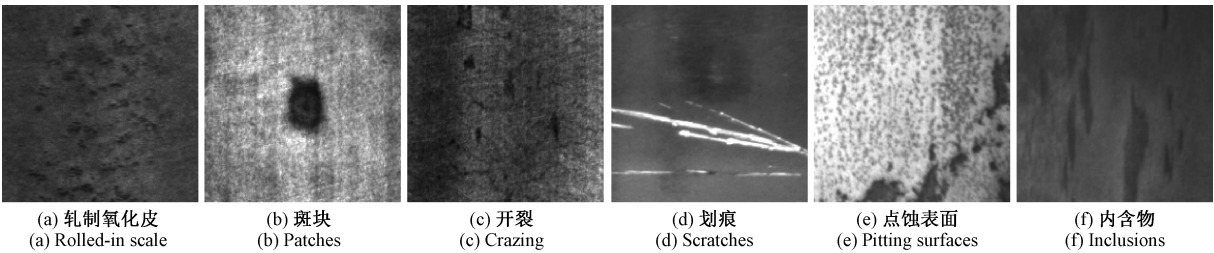


图 8 6 种缺陷样本的示意图

Fig. 8 Schematic diagram of six defect samples

GC10-DET 包含 10 种类型的表面缺陷,即冲孔、熔接线、月牙间隙、水斑、油斑、丝斑、夹杂物、轧坑、折痕、腰部折叠。收集的缺陷位于钢板表面。该数据集包括 3 570 张灰度图像。数据集详细的标注了缺陷的位置信息和类

型信息。实验按照 7 : 2 : 1 对 3 570 张数据图像划分训练集、测试集和验证集。数据集的部分数据图像如图 9 所示。

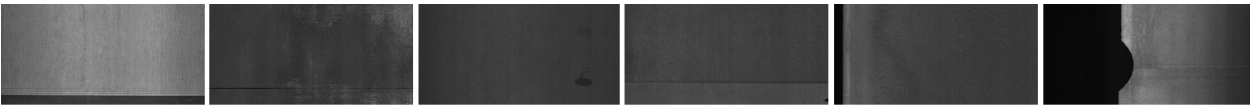


图 9 GC10-DET 部分数据集图片

Fig. 9 GC10-DET selected dataset images

3.3 评价指标

评价钢带表面缺陷检测模型的检测精度的常用评价指标是精度 (precision)、召回率 (recall) 和平均精度均值 (mean average precision, mAP)。具体计算如式(9)~(11)所示。

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \tag{11}$$

式中:TP 代表真实案例;FN 代表假阴性情况;FP 代表假阳性案例;n 代表案例总数。

此外,评价指标还包括模型计算复杂性 (GFLOPs) 和模型参数。模型计算复杂性和模型参数计数分别是评估模型效率和容量的关键指标。

3.4 消融实验

本文提出了一种基于 YOLOv8 的 ESE-YOLO 模型,通过 3 个改进对网络模型进行了优化,为验证 3 个改进对,并进行了消融实验,实验数据如表 2 所示。改进 1 是引入 EIEStem 模式作为高效的特征提取前端,改进 2 则是采用 C2f_SWC 结构替换主干网络中的 C2f 结构,改进 3 采用 EMBSFPN 特征金字塔网络。“√”表示加入,“×”表示不加入。这些改进旨在提升模型的整体性能。

由表 2 可以看出,每次改进加入模型后,mAP@ 0.5 均有所提升,从而改善了模型的检测性能。与原模型相

比,改进后的模型在参数量上下降了 26.8%,计算量下降了 64%,mAP@ 0.5 提高了 4.1%。其中,改进方案 1 通过引入 EIEStem 高效前端模块,有效提升了对目标的检测性能,使 mAP@ 0.5 提高了 3.2%。当改进方案 1 与方案 2 组合时,参数量和计算量略有增加,但帧率提升了 3.4%。虽然改进方案 3 提升的平均精度不高,但参数量下降了 32.6%,并使计算量下降了 20.2%。在 3 个改进方案组合时,mAP@ 0.5 提高了 4.1%,参数量下降 26.8%,浮点运算量减少 64%。改进后的模型通过增强特征提取能力,以及提升模型对目标物体的表达和感知能力,实现了对钢带表面缺陷的更精确检测,整体提升了模型的检测性能。

表 2 消融实验

Table 2 Ablation experiment

改进 1	改进 2	改进 3	mAP@ 0.5	参数量	浮点数/GFLOPs
×	×	×	0.722	3 157 200	8.9
√	×	×	0.754	3 007 634	2.6
×	√	×	0.756	3 183 074	10.9
×	×	√	0.734	2 128 947	7.1
√	√	×	0.762	3 179 042	11
√	×	√	0.756	2 139 011	2.4
×	√	√	0.759	2 300 355	9.9
√	√	√	0.763	2 310 419	3.2

3.5 模型检测效果对比实验

为了更直观地展示本文所提出模型在钢材表面缺陷

检测方面的能力,对改进前后的模型进行了测试,部分测试如图 10 所示。

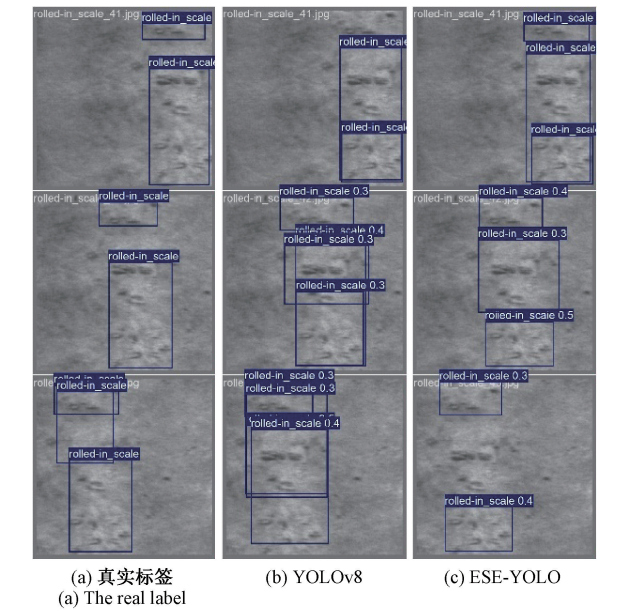


图 10 模型检测效果对比实验

Fig. 10 Comparison experiment of model detection effect

从图 10 可以看出, YOLOv8 模型在检测轧制氧化皮缺陷时,出现了检测框重合重复现象,并且检测框的置信分数普遍低于本文提出的模型,置信分数多为 0.3,而本文模型的置信分数能达到 0.5,明显高于原模型。轧制氧化皮作为钢材缺陷中较难检测的类型,本文模型在这一任务上表现优异。实验表明,改进后的模型有效减少了漏检率并提高了置信度,优于原 YOLOv8 模型的性能。

3.6 泛化实验

为了验证模型的泛化能力,在 GC10-DET 钢材表面缺陷数据集上进行了进一步的实验,实验数据如表 3 所示。在 GC10-DET 数据集的检测结果中, mAP@0.5 值达到 67.4%, 相比于 YOLOv8 分别提高了 9.9%, 参数量和计算量相比于 YOLOv8 分别降低了 23.2% 和 60.5%。综上所述,本文模型不仅在 NEU-DET 数据集上表现出色,在其他数据集上同样具备良好的通用性,证明了改进后模型在工业钢材表面缺陷检测任务中的重要参考价值。

表 3 泛化实验

Table 3 Generalization experiment

模型	mAP@0.5	参数量	浮点数/GFLOPs
YOLOv8	0.575	3 007 598	8.1
ESE-YOLO	0.674	2 311 199	3.2

3.7 对比实验

在对钢材表面缺陷检测的对比实验中, ESE-YOLO 模型在检测钢材表面缺陷方面展现了显著的性能优势。

研究选取了主流目标检测算法作为对照组, 涵盖了 YOLOv3、YOLOv5 的 s、n 和 m 版本、YOLOv6、YOLOv7、YOLOX 以及一些由其他研究者改进的算法。所有实验均在统一的 NEU-DET 数据集和一致的实验条件下进行。对比实验的数据如表 4 所示。

表 4 NEU-DET 数据集上不同算法的性能比较

Table 4 Performance comparison of different algorithms on NEU-DET dataset

算法	mAP@0.5	参数量/($\times 10^6$)	浮点数/GFLOPs
YOLOv3	0.655	63	65.9
YOLOv5n	0.72	1.9	4.5
YOLOv5s	0.707	7.2	16.5
YOLOv5m	0.711	21.2	49
YOLOv6s	0.71	18.5	45.3
YOLOv7	0.709	37.2	104.7
YOLOvXs	0.713	9	26.8
YOLOv8n	0.722	3.2	8.9
YOLOv8s	0.715	11.2	28.6
YOLOv9-c	0.721	2.6	10.7
YOLOv10n	0.724	2.7	8.2
YOLOv11n	0.732	2.6	6.3
文献[30]	0.741	23.9	/
文献[31]	0.733	17.4	/
ESE-YOLO	0.763	2.3	3.2

从表 4 可以看出, 本文提出的改进算法实现了 76.3% 的 mAP 值, 尽管参数量和计算量相比于 YOLOv5n 有所增加, 但精度提高了 4.3%。与检测精度最高的算法^[30]相比, 参数量降低了 90.4%, 准确率提升了 2.2%。在与 YOLOv6 和 YOLOv7 的比较中, 参数量和计算量均降低, 精度分别提升了 5.3% 和 5.4%。与 YOLOX 相比参数量降低 74.5%, 精度提升了 5%, 计算量减少了 88.1%。与 YOLOv8s 相比, 参数量下降 79.5%, 精度提升了 4.8%, 计算量减少了 88.8%。与 YOLOv9-c 和 YOLOv10n 相比精度分别提高了 4.2% 和 3.9%, 参数量和计算量均降低。在与 YOLOv11n 的比较中, 精度提升了 3.1%, 计算量减少了 49.2%。

4 结 论

为解决当前带钢表面缺陷检测算法在实际应用中检测精度低, 计算成本高等问题, 本文在 YOLOv8 的基础上提出了一种有效的钢带表面缺陷检测模型 ESE-YOLO。提出 EIESTem 模块, 通过融合 Sobel 边缘特征与池化空间信息, 增强细节提取能力, 浮点运算量降低 70.8%; 设计 C2f_SWC 模块, 结合位移卷积扩大感受野, 提升上下文信息捕捉精度; 采用 EMBSFPN 特征金字塔, 实现多尺度自适应融合, 参数量减少 32.6%; 在 NEU-DET 与 GC10-DET 数据集上, mAP 分别达 76.3% 与 67.4%, 计算效率

优于主流模型,具备工业部署潜力。后续研究将着重提升钢带表面缺陷的检测速度,有效满足实时检测需求,进一步提升模型检测性能。

参考文献

- [1] DUSPARA M, SAVKOVIĆ B, DUDIC B, et al. Effective detection of the machinability of stainless steel from the aspect of the roughness of the machined surface [J]. *Coatings*, 2023, 13: 447.
- [2] LYU X, DUAN F, JIANG J J, et al. Deep metallic surface defect detection: The new benchmark and detection network [J]. *Sensors*, 2020, 20: 1562.
- [3] AMIN D, AKHTER S. Deep learning-based defect detection system in steel sheet surfaces [C]. *Proceedings of the 2020 IEEE Region 10 Symposium (TENSYP)*, 2020: 444-448.
- [4] DORAFSHAN S, MAGUIRE M, COLLINS W. Infrared thermography for weld inspection: Feasibility and application [J]. *Infrastructures* 2018; 3(4):45.
- [5] WANG G, XIAO Q, GAO Z, et al. Multifrequency AC magnetic flux leakage testing for the detection of surface and backside defects in thick steel plates [J]. *IEEE Magnetics Letters*, 2022, 13: 1-5.
- [6] ZOU Z, CHEN K, SHI Z, et al. Object detection in 20 years: A survey [J]. *Proceedings of the IEEE*, 2023, 111(3): 257-276.
- [7] 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.
- [8] GIRSHICK R. Fast R-CNN [C]. *Proceedings of the IEEE International Conference on Computer Vision*, 2015: 1440-1448.
- [9] 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*, 2015, 39(6): 1137-1149.
- [10] HE K, GKOLARI G, DOLLÁR P, et al. Mask R-CNN [C]. *Proceedings of the IEEE International Conference on Computer Vision*, 2017: 2961-2969.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. *Proceedings of the Computer Vision-ECCV 2016: 14th European Conference*, 2016.
- [12] 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.
- [13] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]. *Proceedings of the European Conference on Computer Vision*, 2020: 213-229.
- [14] ZAIDI S S A, ANSARI M S, ASLAM A, et al. A survey of modern deep learning based object detection models [J]. *Digital Signal Processing*, 2022, 126: 103514.
- [15] ZHAO W, CHEN F, HUANG H, et al. A new steel defect detection algorithm based on deep learning [J]. *Computational Intelligence and Neuroscience*, 2021(10): 1-13.
- [16] 吴健生,王健全,付美霞,等. 基于改进 Faster-RCNN 算法的带钢缺陷检测 [J]. *鞍钢技术*, 2022(6): 23-28,32.
WU J SH, WANG J J, FU M X, et al. Strip defect detection based on improved Faster RCNN algorithm [J]. *Angang Technology*, 2022(6): 23-28,32.
- [17] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition*, 2016: 6517-6525.
- [18] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. *ArXiv preprint arXiv: 1804.02767*, 2018.
- [19] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection [J]. *ArXiv preprint arXiv:2004.10934*, 2020.
- [20] 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*, 2016: 2117-2125.
- [21] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 8759-8768.
- [22] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 658-666.
- [23] ZHAO C, XIN S, XI Y, et al. RDD-YOLO: A modified YOLO for detection of steel surface defects [J]. *Measurement*, 2023, 214(2): 112776.
- [24] 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-NFCM [J].

Foreign Electronic Measurement Technology, 2024, 43(7): 97-104.

[25] 李思思,葛华勇. 改进 YOLOv7 的钢材表面缺陷检测模型[J]. 计算机技术与发展,2024, 34(8): 78-85.

LI S S, GE H Y. Improving the steel surface defect detection model of YOLOv7 [J]. Computer Technology and Development, 2024, 34(8): 78-85.

[26] ZHOU R, LIU D. Quantum image edge extraction based on improved sobel operator [J]. International Journal of Theoretical Physics, 2019, 58(9): 2969-2985.

[27] LI D, LI L, CHEN Z, et al. Shift-ConvNets: Small convolutional kernel with large kernel effects[J]. ArXiv preprint arXiv:2401.12736,2024.

[28] YANG Z, GUAN Q, ZHAO K, et al. Multi-branch auxiliary fusion YOLO with re-parameterization heterogeneous convolutional for accurate object detection [J]. ArXiv preprint arXiv:2401.12736,2024.

[29] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 10778-10787.

[30] 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.

CAO Y Q, WU M L, XU L. Steel surface defect detection based on improved YOLOv5 algorithm [J].

Journal of Graphic Science, 2023,44 (2):335-345.

[31] 王浩然. 基于 YOLOv5 的钢材表面缺陷检测研究[D]. 桂林:广西师范大学,2022.

WANG H R. Research on surface defect detection of steel based on YOLOv5 [D]. Guilin: Guangxi Normal University, 2022.

作者简介



沈冰星, 现为上海海事大学硕士研究生,主要研究方向为深度学习、缺陷检测。
E-mail: 16606262637@163.com

Shen Bingxing is now a M. Sc. candidate at Shanghai Maritime University. His main research interests include deep learning and

defect detection.



黄洪琼(通信作者),2008 年于上海海事大学获得博士学位,现为上海海事大学副教授,主要研究方向为电子信息类、通信与信息系统、信号与信息处理。
E-mail: hqhuang@shmtu.edu.cn

Huang Hongqiong (Corresponding author) received her Ph. D. from Shanghai Maritime University in 2008. She is now an associate professor at Shanghai Maritime University. Her main research interests include electronic information, communication and information systems, as well as signal and information processing.