

文章编号:1671-4229(2024)04-0009-11

基于改进 YOLOv5s 模型的金属表面缺陷检测方法

马 鸽, 邓开宏, 李国章, 李洪伟, 邹 涛

(广州大学 机械与电气工程学院, 广东 广州 510006)

摘要: 精密零件广泛应用于国防、医疗器械、航空航天、电子等产业,是实现“中国制造2025”宏伟蓝图的基础部件之一。然而,在精密零件的加工过程中,零件表面不可避免地会出现裂痕、点蚀表面、划痕等缺陷,导致零件质量和精度下降,进一步地影响设备性能和使用寿命。因此,零件表面缺陷的有效检测至关重要。为解决金属表面小目标缺陷漏检误检、精确度低的问题,文章设计了一种用于金属零件缺陷检测的改进 YOLOv5s 模型。首先,针对零件表面缺陷成像多为条纹状或点状特征、类型相对简单的特点,设计了一种轻量化的 GhostC3 模块,使用更少的参数,通过减少计算量生成特征图;其次,结合点蚀表面、轧制氧化皮等类型缺陷尺寸普遍偏小且分布不均匀的特点,设计一种基于双向特征金字塔网络 BiFPN 的 Concat_BiFPN 模块,充分利用 BiFPN 对不同尺度特征的融合能力,提高小目标检测的精度和稳定性;最后,充分利用 Siou (Spatial Intersection over Union loss) 惩罚函数同时考虑目标形状、空间关系的特点,更好地捕捉不同尺寸的目标在图像中的位置关系,从而提高目标位置的精确性。在 NEU-DET 数据集上的消融实验和对比实验表明,该方法在参数量大幅降低的情况下获得了更高的多类别平均精度。

关键词: 表面缺陷检测; YOLOv5s; 特征融合; Siou

中图分类号: TG178 **文献标志码:** A

Metal surface defect detection method based on an improved YOLOv5s model

MA Ge, DENG Kai-hong, LI Guo-zhang, LI Hong-wei, ZOU Tao

(School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China)

Abstract: Precision components are widely used in various industries, including national defense, medical devices, aerospace, and electronics. They are among the fundamental elements for realizing the grand vision of “Made in China 2025.” However, during the processing of precision components, surface defects such as cracks, pitted-surfaces, and scratches are inevitable, which leads to a decline in the quality and precision of the components. Furthermore, these surface defects will affect the performance and lifespan of the equipment. Therefore, effective detection of surface defects is crucial. To address the issues of missed and false detections of metal surface defects and to improve detection accuracy, this paper proposes an improved YOLOv5s model for surface defect detection in metal components. First, considering that metal surface defects often present as streaks or spots which express as relatively simple types, a lightweight GhostC3 module is designed by generating feature maps using

收稿日期: 2024-03-25; 修回日期: 2024-07-05

基金项目: 广州大学创新训练资助项目(202211078129)

作者简介: 马鸽(1988—),女,副教授,博士. E-mail: m_ge@gzhu.edu.cn

引文格式: 马鸽, 邓开宏, 李国章, 等. 基于改进 YOLOv5s 模型的金属表面缺陷检测方法[J]. 广州大学学报(自然科学版), 2024, 23(4): 9-19.

fewer parameters and resulting in reduced computational complexity. Second, given the characteristics of defects such as pitted-surfaces and rolled-in scale which are generally small and unevenly distributed, a Concat_BiFPN module based on a bidirectional feature pyramid network was designed. This module makes full use of the ability of BiFPN to fuse features at different scales, and improves the accuracy and stability of small target detection. Finally, the SIoU loss function is utilized, which considers the shape and spatial relationships of targets to better capture the positional relationships of different targets among the image, thus resulting in enhancing the precision of target localization. Ablation and comparative experiments on the NEU-DET dataset demonstrate that the proposed method achieves higher multi-class average precision with a significantly reduced number of parameters.

Key words: surface defect detection; YOLOv5s; feature fusion; SIoU

精密零件广泛应用于国防、医疗器械、航空航天、电子、家居用品制造等产业,是实现“中国制造 2025”宏伟蓝图的基础部件之一。然而,在精密零件的加工过程中,零件表面不可避免地会出现缺陷,如裂痕、点蚀表面、划痕等,从而导致质量和精度下降。精密零件的加工质量和精度直接影响设备的质量、性能和使用寿命。因此,零件表面缺陷的有效检测对于提高设备质量、减少次品数量和提高使用寿命至关重要。传统的人工缺陷检测方法效率低下,检测速度和精度高度依赖工人的经验和主观因素,基于机器视觉的缺陷检测方法因其非接触、无损、自动检测等特点被广泛应用于表面缺陷检测领域。

基于视觉的缺陷检测方法主要分为基于传统图像处理的方法和基于深度学习的方法两大类。传统图像处理方法主要通过人工分析表面缺陷图像的纹理、骨架和光谱等结构特征进行检测。文献[1]提出了一种用于分布式缺陷识别的最优多特征集融合的改进随机森林算法:首先,分别利用梯度方向直方图特征集和灰度共生矩阵特征集描述图像的局部和全局纹理特征;其次,针对缺陷图像分布样本少、特征集维数高的特点,引入随机森林算法进行缺陷分类。文献[2]针对板坯纵向裂纹缺陷难检测问题,提出了一种结合 Curvelet 变换和核局部保持投影的特征提取方法。文献[3]针对边界模糊、灰度不均匀、背景纹理复杂等缺陷,提出了一种基于 Active contour 模型的改进分割算法,与 Chan-Vese 模型、局部二值拟合模型和全局符号区域压力模型的结果相比,该方法对模糊缺陷的分割具有更好的准确性和鲁棒性,且计算时

间较短。文献[4]为了获取更好的钢表面缺陷特征,深入研究了分别根据 Haar、Daubechies 2 和 Daubechies 4 特征提取的性能,最终选取 Haar 特征作为特征提取模块,获得了较好的缺陷特征提取效果。

基于深度学习的检测与识别方法通过多层网络结构和大量的数据训练,实现了准确性高、鲁棒性强且高效的金属表面缺陷检测。按照是否需要先生成候选框,基于深度学习的方法可分为二阶段方法和一阶段方法两大类。二阶段方法先生成候选区域再进行特征提取与分类,一阶段方法则在单个步骤内执行特征提取、位置回归和类别分类。

2014 年,文献[5]提出的两阶段检测算法 R-CNN (Region-based Convolutional Neural Network):首先,使用选择性搜索算法生成一系列候选区域;接着,借鉴迁移学习的思想将候选区域输入到预训练的 CNN(Convolutional Neural Network)中提取特征信息;最后,对每个候选区域进行评估以确定其中是否包含感兴趣的目标物体。2015 年,文献[6]提出 SPP (Spatial Pyramid Pooling) 模块可以将不同尺寸的输入向量固定成大小相同的输出,从而提高了网络的推理速度。随后,文献[7]对 SPP 进行改进并且引用到 R-CNN 框架下,提出了 Fast R-CNN,该模型在检测速度方面获得较大提升。然而,R-CNN 和 Fast R-CNN 的预测框都是基于选择性搜索(selective search)产生的,冗余的预选框导致训练需要耗费大量的时间,而 Faster R-CNN 模型^[8]采用 RPN 网络(Region Proposal Networks)自动产生候选区域,有效提升了网络的

训练速度。文献[9]充分考虑特征的小目标特点和复杂性,在 Faster R-CNN 的基础上采用多尺度融合的方法训练网络,同时,将部分传统卷积网络替换为可变形卷积网络,在钢材数据集中获得小目标缺陷检测精度的提升。文献[10]为了解决特征图不匹配和预测边框回归不准确的问题,提出一种基于 Cascade R-CNN^[11]改进的钢轨表面缺陷检测方法,获得分类精度提高的同时,减少了检测时间。尽管二阶段检测算法在分类精度上获得良好的效果,但需要先生成候选区域,然后,再对每个候选区域进行特征提取与分类,该过程大大增加了计算复杂度和时间开销,而一阶段检测算法利用整张图作为网络的输入,直接在输出层回归边界框的位置及其所属的类别,在推理速度和轻量化上比二阶段检测算法更能满足工业检测的快速性要求。

文献[12]提出的单步多框目标检测方法(Single Shot MultiBox Detector, SSD)是典型的一阶段方法,该网络采用金字塔结构对图像进行下采样,融合多尺度信息极大地提升了小目标检测的效果。在此基础上,文献[13]提出一种融合 Channel-Attention 机制的 SSD 模型,该模型应用到铝材表面缺陷检测提升了分类精度。2016年,文献[14]提出了一阶段检测算法 YOLOv1 (You Only Look Once),该方法将图像划分成网格形状,进而为每一个网格赋予目标的预测概率,将寻找目标简化到一次概率计算上,大幅减少了检测时间。随后,文献[15]将 YOLOv1 改为全卷积的形式引入到 NEUDET 数据集检测中,实现了钢表面缺陷实时检测。文献[16]提出 YOLOv2 模型,通过增加 Anchor 机制和添加多尺度融合提高对小目标的检测精度,YOLOv3^[17]则通过引入特征金字塔 (Feature Pyramid Networks, FPN) 的思想捕获更多尺度上下文信息,提高检测精度。在此基础上,文献[18]提出一种改进的 YOLOv3 缺陷检测模型,有效地提高了特征重用、特征传播能力,增强了网络的表征能力。YOLOv4^[19]结合特征融合结构 PAN (Path Aggregation Network),进一步提高了 YOLO 系列方法融合不同尺度大小特征图的能力。为了改善 YOLO 系列模型在钢表面缺陷检测中小目标信息的关注度和精度,文献[20]在 YOLOv5 的基础上,引入注意力机制模块提升了钢表面小

目标缺陷的准确定位与分类精度。文献[21]结合重参数和动态标签分配策略提出 YOLOv7,在不增加推理时间成本的情况下,该算法提高了检测精度,但由于其模型的参数量较大,无法满足工业缺陷检测的快速性需求。2023年1月,Ultralytics 公司发布了 YOLOv8 模型,通过将 YOLOv5 中的 C3 模块替换成 C2f 模块,YOLOv8 在保证模型轻量化的基础上,获得更丰富的梯度流信息,同时,YOLOv8 通过采用无锚框预测目标直接预测对象中心位置,加速推理步骤。该模型具有较好的可扩展性,可方便地应用于目标检测、图像分割、姿态估计等任务。然而,由于 YOLOv8 所采用的无锚框方法对每个位置只预测一个框,从而导致一些重看或遮挡区域无法被检测到,同时该方法精度和稳定性的提高也依赖特殊的损失函数或结构,从而降低召回率或定位精度。YOLOv5s 采用基于锚框的预测方法,该方法使用不同大小和形状的锚框来回归和分类目标,从而产生密集的锚框,使得网络可直接进行目标分类和边界框回归,进而提高目标召回能力,对小目标检测精度的提升有明显的效果。因此,针对金属表面缺陷数据集中小目标缺陷数目繁多,现有一阶段检测算法存在漏检、误检、精确度低等问题,本文基于 YOLOv5s 模型,设计了一种用于小目标金属表面缺陷检测的模型,该模型一方面采用轻量化 GhostC3 模块来压缩模型参数的大小,另一方面通过添加更深层次的特征融合模块 Concat_BiFPN 来增强不同尺度特征的融合能力,进而解决小目标特征难提取的问题。进一步地,充分利用 SiU 惩罚函数的同时,考虑目标形状、空间关系的特点,更好地捕捉不同尺寸的目标在图像中的位置关系,提高目标位置的精确性。

1 YOLOv5 网络模型

根据网络深度和宽度的不同,YOLOv5 分为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 个版本,其中,YOLOv5s 具有最小的网络参数量和最快的计算速度,但受限于较低的平均精度。金属表面缺陷检测算法需要对采集的图像进行实时处理,并返回检测结果以备后续工序进行分拣等操作,因此,对算法的快速性、准确性和模型大小要

求较高,如何有效利用 YOLOv5 快速性的同时,提高检测精度、降低模型大小是本文需要解决的主要问题。

YOLOv5 的框架图如图 1 所示,包括骨干网络 (backbone)、颈部网络 (neck) 和预测输出端 (prediction) 3 部分。其中,骨干网络由卷积层、C3 模块以及 SPPF (Spatial Pyramid Pooling Fusion) 组成。在卷积层中, k^* 表示卷积核大小, s^* 表示卷积步长, p^* 表示边缘填充量, c^* 表示卷积核数量, $*$ 为数字表示相应的卷积核大小、步长、填充量和卷积核数量。C3 模块包含 3 个卷积层和 1 个特征融

合层,通过级联和残差连接来增强特征表示能力。SPPF 模块通过多尺度融合卷积获取更深层次的语义信息,从而提升网络特征提取的准确性。颈部网络将图像特征拼接并传输到预测层,输出端则预测图像特征以生成边界框和预测类别,每一个检测头会输出 $C = (5 + n_{cls}) \times 3$,其中, C 表示输出通道;5 表示每个锚点需要预测的中心点坐标 (x 、 y)、宽度 (width)、高度 (height) 以及置信度 (confidence score) 5 个参数; n_{cls} 表示类别的数量,即目标检测任务中需要分类的不同类别的数量;3 表示每个网格单元预测的 3 个锚点。

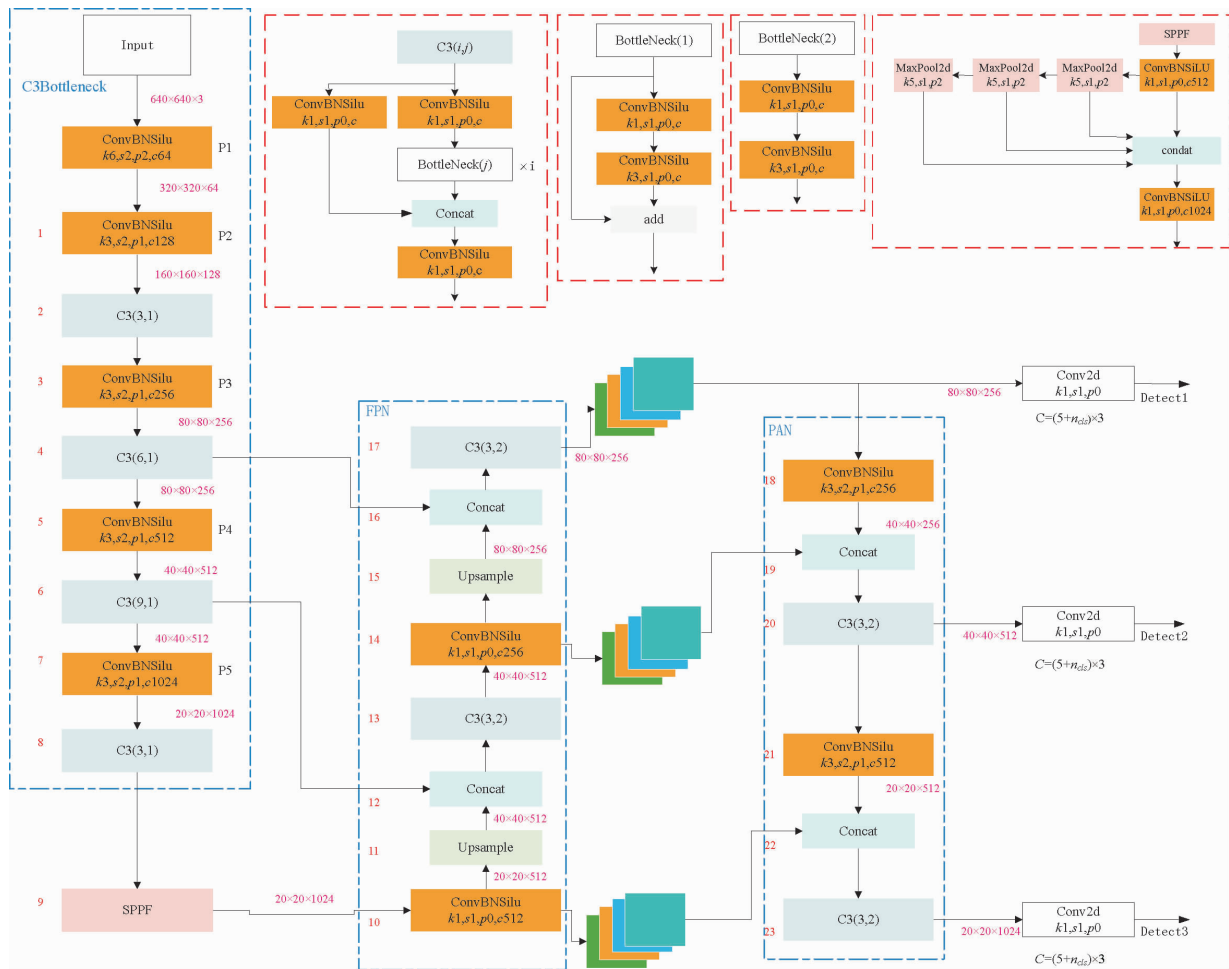


图 1 YOLOv5 结构图

Fig. 1 Structure of YOLOv5

2 基于 YOLOv5s 改进模型设计

2.1 模型结构

为解决金属表面小目标缺陷漏检、误检和精

确度低的问题同时,考虑工业部署的快速性需求,本文设计一种用于金属零件缺陷检测的改进 YOLOv5s 模型。首先,零件表面图像中开裂、刮痕等缺陷表现为条纹特征,点蚀表面、斑块、杂质等缺陷类型多表现为点状特征,特征类型相对简单,

因此,将 FPN 特征融合模块中的 C3 模块替换为轻量化网络 GhostC3,使用更少的参数生成特征图,保持特征提取丰富性的同时,显著减少了计算量;其次,由于点蚀表面、轧制氧化皮等类型缺陷尺寸普遍偏小且分布不均匀,本文在 Concat 特征融合模块中采用双向特征金字塔网络 BiFPN (Bidirectional Feature Pyramid Network),充分利用 BiFPN 对不同尺度特征的融合能力来提高小目标检测的

精度和稳定性;最后,边界框损失函数可以帮助算法有效地训练并获得最佳权重,从而提高定位精度,因此,本文采用 SIoU (Scylla Intersection over Union) 作为惩罚函数,不仅考虑真实框与预测框之间的重叠程度,还引入角度损失、距离损失和形状损失,从而能够更准确地捕捉不同尺寸目标在图像中的位置关系,提高了检测的精确性。改进 YOLOv5s 模型网络整体结构如图 2 所示。

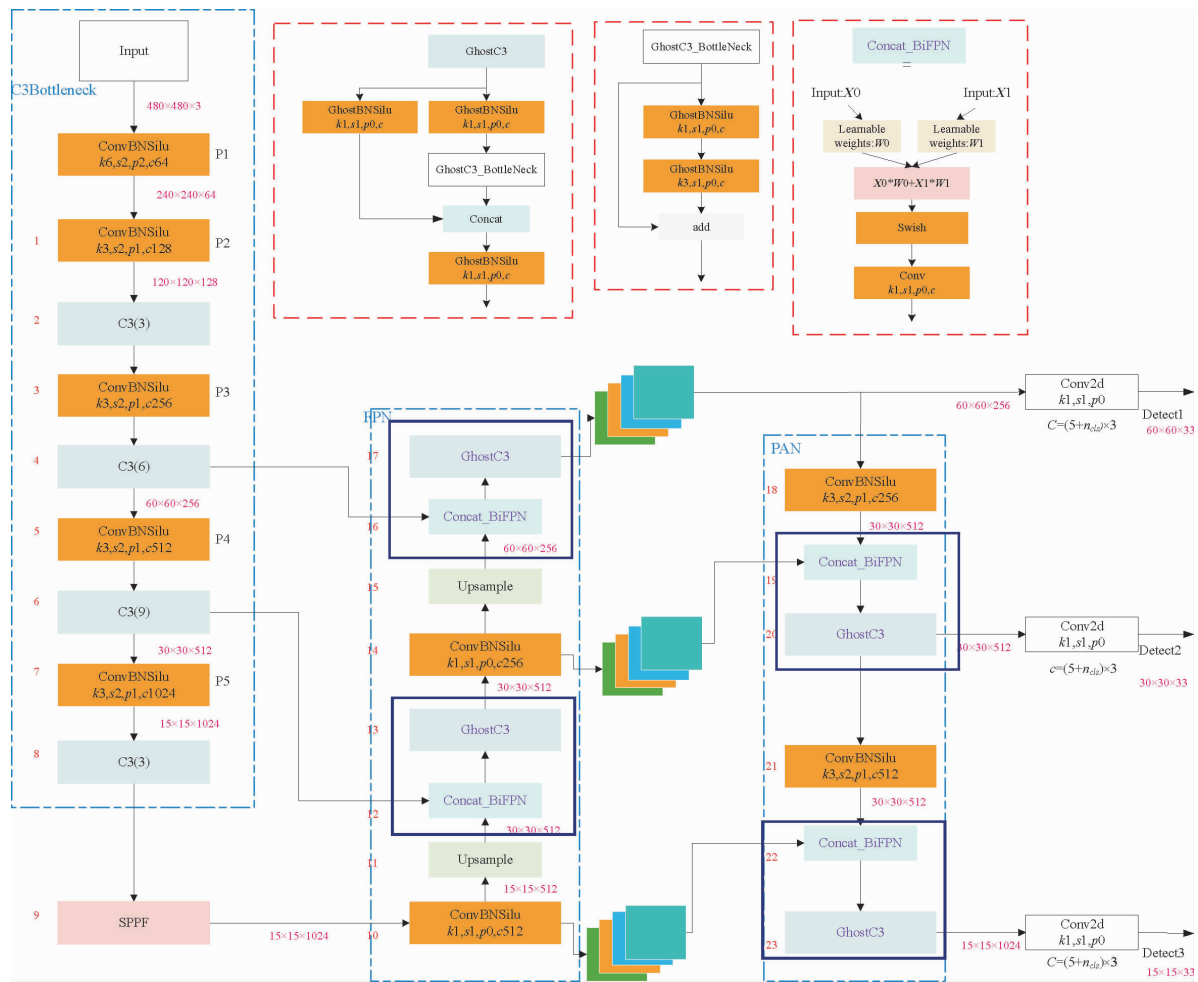


图 2 改进的 YOLOv5s 模型结构

Fig. 2 Structure of the improved YOLOv5s model

2.2 Neck 网络设计

在 Neck 网络中, YOLOv5 系列模型在进行两次上采样后,使用 C3 模块以及传统 FPN 进行特征融合。然而, C3 模块中存在大量卷积计算,其卷积核利用效率低,导致网络计算量大。同时,传统 FPN 在多尺度特征融合时对小目标特征不敏感,导致小目标信息丢失,从而产生漏检和误检问题。针对以上问题,本文在 Neck 网络中将 C3 结构的

Conv 卷积模块替换为 GhostNet^[22],设计一种轻量化的 GhostC3 模块,如图 3(a)和(b)所示。GhostNet 首先,使用较小的卷积核进行基础特征提取,减少输出特征图的通道数,获取一组精简而有效的的基础特征图;其次,通过简单的线性操作,在几乎不增加计算负担的前提下,生成更多特征图,提升网络的特征提取和表示能力;最后,将基础特征图和生成的特征图进行拼接形成最终输出,从而

使网络能够利用更丰富的信息进行学习和预测。GhostC3 模块充分考虑零件开裂、刮痕等缺陷特征相对简单的特点和 GhostNet 有效降低参数数量的优点,在保障特征提取精度的同时,加快了训练速度降低了计算成本。

此外,原始 YOLOv5 的 Neck 网络使用 Concat 操作对特征图进行融合,将多个层按照指定方式端对端地拼接起来,整合上下两层的不同类型和不同级别信息。然而,金属表面点蚀表面、轧制氧化皮等类型缺陷尺寸普遍偏小且分布不均匀,传统的特征融合方法对小目标特征不敏感容易导致误检和漏检。BiFPN 作为一种基于特征金字塔网络(FPN)的对象检测神经网络结构,通过引入双向流动的特征信息,有效地解决了传统 FPN 在提取不同尺度的特征时存在的信息丢失和冗余问题。因此,本文将 BiFPN 融合进 Concat 结构,设计一种 Concat_BiFPN 模块,通过双向路径增强不同尺度特征图之间的信息流动,同时不断迭代加权融合权重优化特征融合过程,从而使得网络更加

重视信息量较大的特征,有效提高特征图质量的同时,捕捉对象的多尺度信息,进而提高目标检测的准确度和速度。Concat_BiFPN 模块对于点蚀表面、轧制氧化皮等尺度变化频繁的场景更加有效。如图 3(c)所示,Concat_BiFPN 模块接收两个输入特征图,分别记为 X_0 和 X_1 。首先,模块初始化两个可学习的权重参数 W_0 和 W_1 用于调整每个输入特征图在融合过程中的权重,迭代过程中,权重参数通过反向传播算法进行优化以获得最优融合结果。在前向传播过程中,输入特征图与对应的权重参数相乘并求和得到加权求和结果;随后,将求和结果归一化处理,通过 Swish 激活函数进行非线性变换;最后,将激活后的特征图进行 1×1 卷积,该卷积层对输入特征图通道数进行调整以整合和压缩特征信息,为后续网络层提供更有效的特征表示。Concat_BiFPN 模块将低级别特征与高级别特征相互传递,形成更深层次的融合关系,从而更好地捕捉小目标缺陷特征^[23],提高小目标缺陷检测的精度。

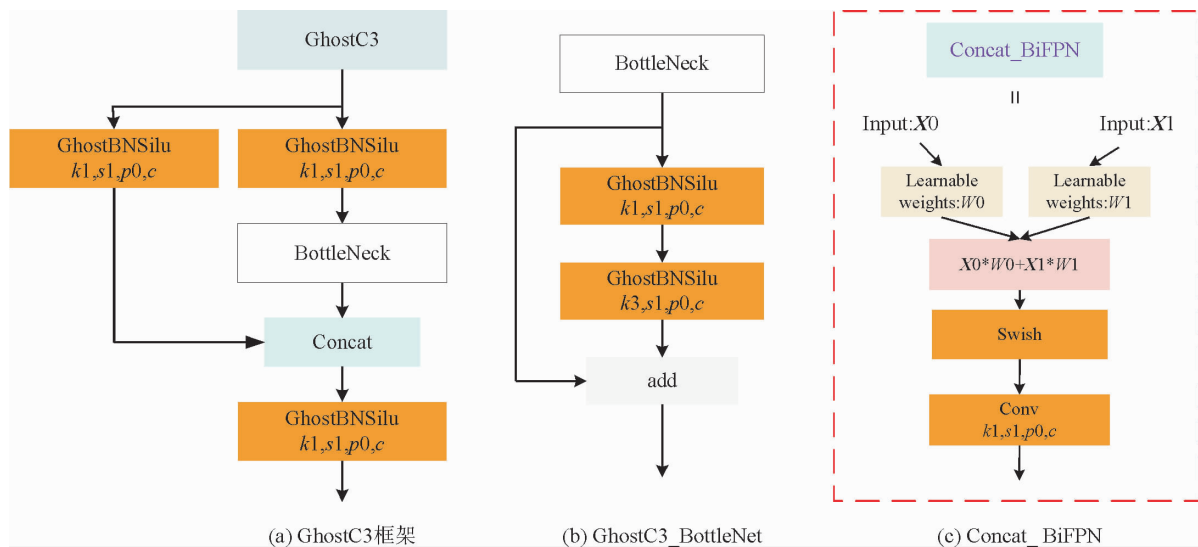


图 3 改进的 Neck 网络结构示意图

Fig. 3 Improved Structure of Neck Network

2.3 SIoU 损失函数

损失函数作为量化模型性能的指标,是训练深度学习目标检测模型的关键组成部分。通过定义损失函数,目标检测模型的训练问题被转化为一个函数最小化问题,通过设计损失函数最小化,优化算法不断学习调整模型参数,使模型的预测结果与实际目标更为接近。因此,选择符合金属

表面缺陷检测任务特征的损失函数尤为重要。

YOLOv5s 网络采用的 CIoU (Complete Intersection over Union) 损失函数主要关注边界框的位置和大小,并没有考虑到标签真实框与网络预测框之间方向不匹配的问题,从而导致对目标方向性信息的不准确建模。而 SIoU 不仅考虑了真实框与预测框之间的重叠程度,还引入了角度损失、距

离损失和形状损失等空间关系的考量,能够更准确地捕捉不同尺寸目标在图像中的位置关系,提高检测的精确性,有助于提高对目标位置的精确性。因此,本文采用 SIoU 作为算法模型预测结果与实际目标接近程度的度量。

SIoU 损失函数由 4 个部分组成的:角度损失 (angle cost)、距离损失 (distance cost)、形状损失 (shape cost) 和 IoU 损失 (Intersection-over-Union cost)。定义如公式(1)所示:

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

其中, Δ 是距离损失, Ω 是形状损失,IoU 是交并比损失。该损失函数通过综合考虑物体的中心位置和预测边界框的中心位置之间的距离以及两个框之间的角度,重新定义距离损失,当两个框的方向不匹配时,即使二者位置非常接近距离损失也会增加,从而提高目标位置定位的准确性。而形状损失则有助于提高模型对于目标的形状变化、扭曲等情况的适应力,有助于解决金属表面图像因图像采集过程中的形变带来的误检问题。另外,SIoU 对目标的尺寸变化不敏感,不会因为目标大小的变化而引入过度的误差,更加适用于各类斑点、杂质、点蚀表面、轧制氧化皮等尺寸偏小且分

布不均匀的情况。尤其地,SIoU 可以直接应用于多类别目标检测问题而无需引入额外的复杂性,从而保障了算法的收敛速度。

3 实验结果及分析

3.1 实验环境配置

本文所使用的数据集是由东北大学 (North-eastern University) 发布的用于表面缺陷检测的公开数据集 NEU-DET (Northeastern University Surface Defects Database) 见图 4,它主要用于金属表面缺陷的检测和分类任务。数据集包含 1 800 幅图像,每种缺陷类别各有 300 幅图像,每幅图像的分辨率为 200×200 像素,其中,1 440 幅用于构建训练集,剩余的 360 幅用于构建测试集,标签文件均为 TXT 格式,训练过程中均采用 Mosaic 数据增强来丰富训练样本增强模型的泛化性能。该数据集含有 6 种钢表面类别缺陷,分别是开裂 (crazing)、点蚀表面 (pitted-surface)、轧制氧化皮 (rolled-in scale)、斑块 (patches)、杂质 (inclusion) 以及刮痕 (scratches)。其中,几种类别的缺陷特征分布不一,同时,缺陷尺度较小,给检测带来了挑战。

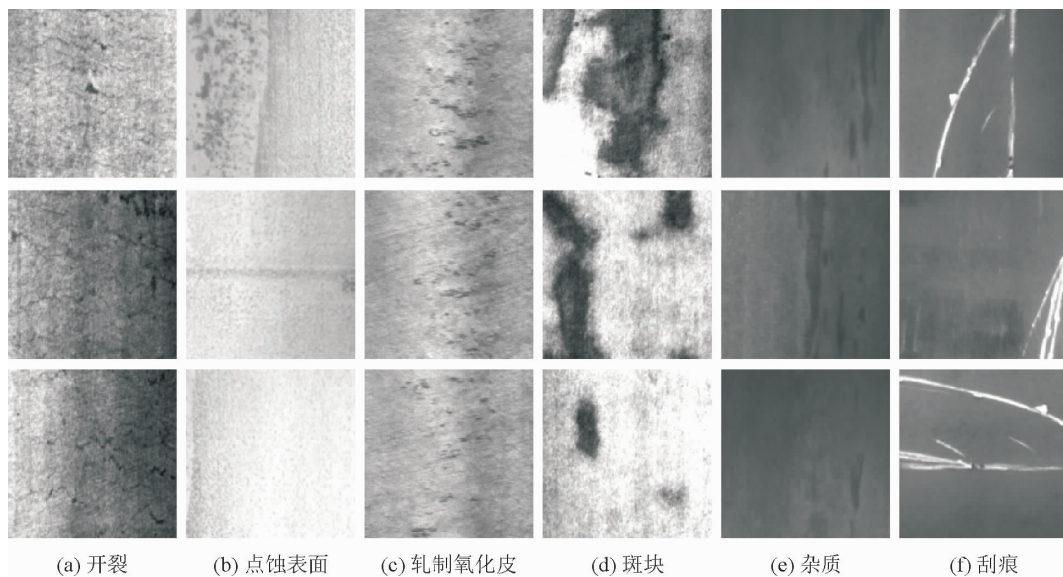


图4 NEU-DET 数据集

Fig.4 NEU-DET dataset

实验评价标准采用:精确率 (precision)、召回率 (recall)、平均精度 (Average-Precision, AP)、多

类别平均精度 (mean Average Precision, mAP) 和每秒帧率 (FPS) 作为客观评价标准。其中,AP 是根

据每个类别的真实标签和预测概率进行计算,依据 Precision 和 Recall 的值作 PR 曲线并计算得到曲线下的面积即为平均精度(AP); mAP 则是将所有类别的 AP 值取平均,其取值范围为 $[0, 1]$,以上两个指标数值越高,则该算法检测性能越好。

Precision 和 Recall 计算公式如式(2)和式(3)所示,实验详细设备环境参数以及实验参数如表 1 所示。

$$Precision = \frac{TP}{TP + FP}, \quad (2)$$

$$Recall = \frac{TP}{TP + FN}, \quad (3)$$

其中,TP:标签值为 True,模型预测为 Positive;

FN:标签值为 False,模型预测为 Negative;

FP:标签值为 False,模型预测为 Positive;

TN:标签值为 True,模型预测为 Negative。

表 2 消融实验对比结果

Table 2 Ablation experimental results of the proposed algorithm

实验	YOLOv5s	Concat_BiFPN	GhostC3	SIoU	mAP/%	参数量	耗时/ms	FPS/(s ⁻¹)
1	√				76.5	7 026 307	8.6	116.2
2	√	√			77.8	6 944 781	8.9	112.6
3	√	√	√		78.1	5 982 887	9.2	108.7
4	√	√	√	√	79.5	5 982 887	9.1	109.9

如表 2 消融实验的数据所示,改进后的模型在多类别平均精度上表现出了明显的提升,模型参数量也出现显著下降,FPS 虽有略微下降,但是相对常规高清视频处理所需的每秒 25~30 帧的标准,本文方法每秒 109.9 帧的 FPS 值仍然满足该应用标准。具体的,实验 2 替换了 Concat_BiFPN 特征融合模块后模型多类别平均精度提高了 1.7%,可见 Concat_BiFPN 对于多类型目标缺陷的特征提取与融合更加有效,但同时因权重学习、非线性变换等增加了部分计算量,导致计算耗时增加。实验 3 将 Neck 网络部分中的 C3 模块替换为轻量化 GhostC3 模块后,相较于消融实验 2,模型精度提升 0.4%、参数量减少 13.9%,可见, GhostC3 模块在保持模型检测精度的同时有效降低了模型参数量,以便于在工业环境中进行部署。实验 4 改用 SIoU 损失函数,模型多类别平均精度相较于实验 3 提高 1.8%,验证了 SIoU 损失函数对于目标检测和定位的有效性。与原始 YOLOv5s

表 1 实验环境详细配置

Table 1 Detailed configuration of the experimental environment

参数	配置说明
操作系统	Ubuntu 20.04
处理器	Intel Xeon Gold 6226R@2.9GHz
GPU	RTX-3090 24G
加速环境	CUDA 11.7、cuDNN 8.5.5
深度学习框架	Pytorch

3.2 实验结果分析

3.2.1 消融实验

为了验证每个更新模块在 NEU-DET 数据集上检测的有效性,本小节对改进算法进行消融实验。通过消融实验分别验证改进后 Concat_BiFPN、GhostC3 以及 SIoU 对模型性能的影响,实验结果如表 2 所示。

相比,本文所提出的模型在多类别平均精度上提高 3.9%,参数量降低 14.8%,表明了本文方法在提高金属表面缺陷识别精度上的有效性。

3.2.2 对比实验

为验证本文算法对金属表面缺陷识别与检测的有效性,选取 SSD、YOLOv3、YOLOv5s、YOLOv8s、YOLOv8n 以及二阶段算法 Faster-RCNN 进行对比实验,实验结果如表 3 所示。两阶段方法 Faster-RCNN 不仅精度低于一阶段检测方法,而且检测速度慢,无法达到实时检测标准。一阶段检测算法 SSD 检测精度和 FPS 相较于 Faster-RCNN 有所提升,但每秒 29.0 帧的处理速度并不适合在工业环境中进行部署。YOLOv3、YOLOv5s、YOLOv8s 和 YOLOv8n 作为 YOLO 系列典型的轻量级检测网络,虽然 FPS 均高于二阶段检测算法以及 SSD 算法,但是在处理金属表面缺陷这一特定应用对象的检测中,其多类别平均精度分别为 72.2%、76.5%、76.3% 和 76.9%,均低于本文算法的 79.5%。

表3 对比实验结果

Table 3 Experimental results of comparable models

Model	mAP/%	FPS/(s ⁻¹)
Faster R-CNN	71.1	23.8
SSD	72.4	29.0
YOLOv3	72.2	64.5
YOLOv5s	76.5	117.6
YOLOv8s	76.3	113.0
YOLOv8n	76.9	119.0
Proposed Method	79.5	109.9

表4列出了本文算法和 YOLOv5s、YOLOv8n、YOLOv8s 的召回率对比结果。相比于 YOLOv5s 模型,本文召回率降低 0.67%,但结合表3数据本文方法的多类别平均精度提升 3.9%,表明改进后的方法可有效识别多种缺陷类型,降低误检。与 YOLOv8s 和 YOLOv8n 相比,本文算法不仅多类别精度分别提升 4.2% 和 3.4%,召回率也获得了不同程度的提升,表明本文方法在处理金属表面缺陷识别任务上的优越性。

为了更加直观地显示本文方法的检测性能,本文进一步对刮痕、轧制氧化皮、斑块等表面类别缺陷的检测结果图像进行分析,实验结果如图5~图7所示。如图5(b)~图(e)所示,针对刮痕缺陷类型 YOLOv5s、YOLOv5n、YOLOv8n 和 YOLOv8s

模型均出现多检或误检的情况,本文方法能够正确检测出杂质缺陷且置信度较高。图6所示为轧制氧化皮表面缺陷检测结果,因轧制氧化皮表面缺陷多为细小点状缺陷,YOLOv5s 和 YOLOv5n 均漏检部分轧制氧化皮表面缺陷,而 YOLOv8n 则将在无缺陷的地方错检为轧制氧化皮表面缺陷。如图6(e)和图6(f)所示,YOLOv8s 和本文算法均能正确检测出钢轧制氧化皮表面缺陷。同样地,由斑块缺陷检测结果如图7所示,可见,YOLOv5n 和 YOLOv8s 未能准确检测出全部斑块缺陷,而 YOLOv5s、YOLOv8n 和本文算法均能较为完整地检测出斑块缺陷,但是本文所提出算法预测框更加贴合真实情况,预测框更加准确,回归效果更好。图5~图7缺陷检测图像进一步说明了本文方法对于小目标缺陷特征提取和预测框推理结果的准确性,表明了本文方法在金属表面缺陷检测任务上的有效性。

表4 各模型召回率对比结果

Table 4 Comparison between recall rates of comparable models

Model	Recall/%
YOLOv5s	74.2
YOLOv8n	67.6
YOLOv8s	72.7
Proposed Method	73.7

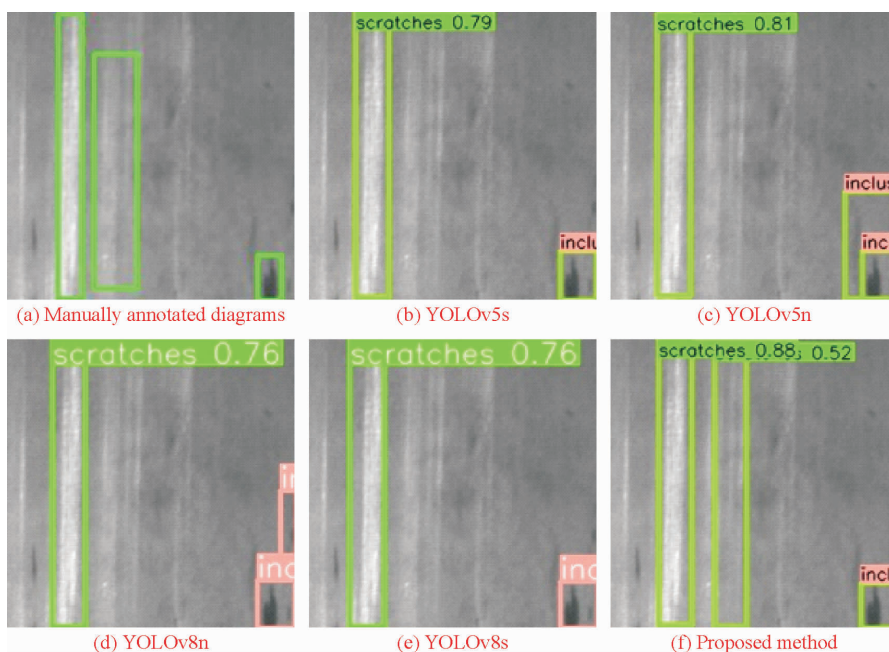


图5 刮痕缺陷检测结果对比图

Fig. 5 Comparison of scratch defect detection results

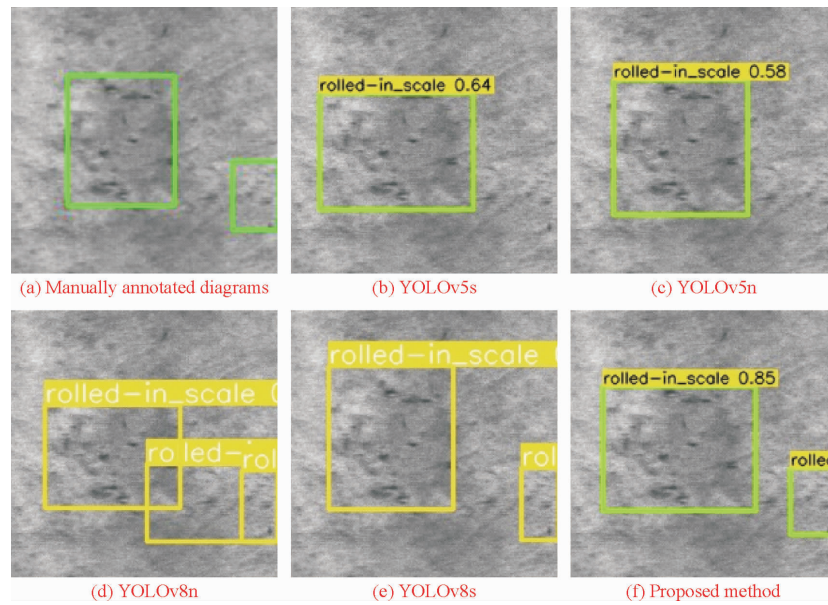


图 6 轧制氧化皮表面缺陷检测结果对比图

Fig. 6 Comparison of rolled-in scale defect detection results

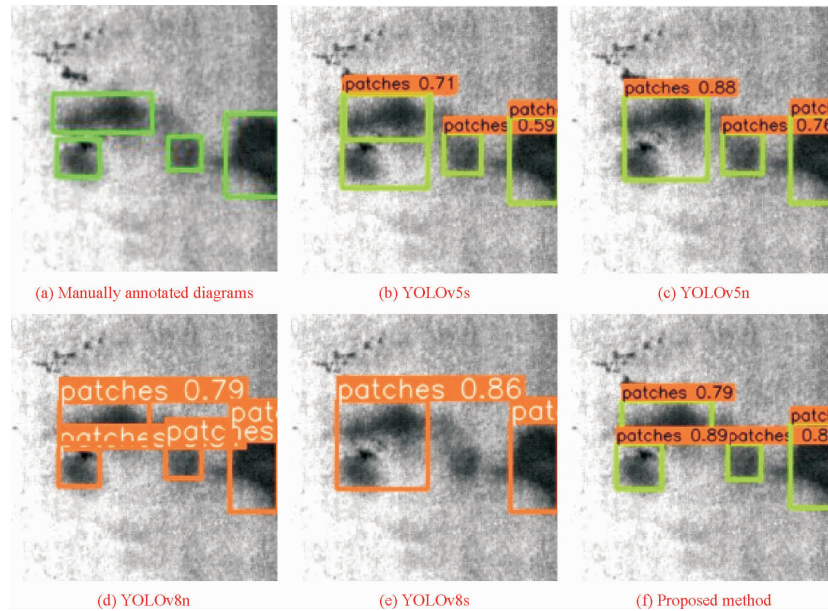


图 7 斑块缺陷检测结果对比图

Fig. 7 Comparison of patch defect detection results

4 总结与展望

针对工业场景下金属工件缺陷检测任务中存在漏检、误检的问题,本文分析零件表面图像中开裂、刮痕、点蚀表面、斑块、杂质、轧制氧化皮等缺陷的成像特点,设计轻量化的 GhostC3 网络以使用更少的参数来生成特征图;针对缺陷尺寸偏小且分布不均匀的情况采用双向特征金字塔网络进行

特征融合,提升小目标检测的精度和稳定性;采用 Siou 惩罚函数更准确地捕捉不同尺寸目标在图像中的位置关系,提升检测的准确性。在 NEU-DET 数据集上进行消融实验的结果表明,本文所提出的模型相比于原始 YOLOv5s 模型在多类别平均精度上提高 3.9%,参数量降低 14.8%。与 FasterRCNN、SSD、YOLOv3、YOLOv5s、YOLOv8s、YOLOv8n 等方法的对比实验表明,本文方法不仅精度高,降低了误检和漏检,同时参数量大幅减少、检测速度

得到提升。未来将针对算法精度和召回率的进一步提升,以及对算法的可移植性进行研究,以期能够满足工业金属缺陷检测准确性和低成本的需求。

参考文献:

- [1] Wang Y, Xia H, Yuan X, et al. Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion[J]. *Multimedia Tools and Applications*, 2018, 77: 16741-16770.
- [2] Ai Y, Ke X. Surface detection of continuous casting slabs based on curvelet transform and kernel locality preserving projections[J]. *Journal of Iron and Steel Research, International*, 2013, 20(5): 80-86.
- [3] Yang J, Li X, Xu J, et al. Development of an optical defect inspection algorithm based on an active contour model for large steel roller surfaces[J]. *Applied Optics*, 2018, 57(10): 2490-2498.
- [4] Ghorai S, Mukherjee A, Gangadaram M, et al. Automatic defect detection on hot-rolled flat steel products[J]. *IEEE Transactions on Instrumentation and Measurement*, 2012, 62(3): 612-621.
- [5] Girshick R, Donahue J, Darrell T, et al. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014[C]. Piscataway: IEEE, 2014.
- [6] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [7] Girshick R. Proceedings of the 2015 IEEE International Conference on Computer Vision and Pattern Recognition, December 07-13, 2015[C]. Piscataway: IEEE, 2015.
- [8] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *Advances in Neural Information Processing Systems*, 2015, 39: 1137-1149.
- [9] Zhao W, Chen F, Huang H, et al. A new steel defect detection algorithm based on deep learning[J]. *Computational Intelligence and Neuroscience*, 2021, 2021(10): 1-13.
- [10] 罗晖, 李健, 贾晨. 基于图像增强与改进 Cascade R-CNN 的钢轨表面缺陷检测[J]. *激光与光电子学进展*, 2021, 58(22): 324-335.
- [11] Cai Z, Vasconcelos N. Cascade R-Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018[C]. Piscataway: IEEE, 2018.
- [12] Liu W, Anguelov D, Erhan D, et al. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, October 11-14, 2016[C]. Berlin: Springer, 2016.
- [13] 符秦沈, 王桂棠, 张巧芬, 等. 融合 Channel-Attention 机制的金属表面缺陷检测算法[J]. *机械设计与制造*, 2023: 1-7.
- [14] Redmon J, Divvala S, Girshick R, et al. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016[C]. Piscataway: IEEE, 2016.
- [15] Li J, Su Z, Geng J, et al. Real-time detection of steel strip surface defects based on improved yolo detection network[J]. *IFAC-PapersOnLine*, 2018, 51(21): 76-81.
- [16] Redmon J, Farhadi A. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017[C]. Piscataway: IEEE, 2017.
- [17] Redmon J, Farhadi A. YOLOv3: An incremental improvement[EB/OL](2018-06-01)[2024-01-21]. <https://arxiv.org/abs/1804.027678>.
- [18] Kou X, Liu S, Cheng K, et al. Development of a YOLOv3-based model for detecting defects on steel strip surface[J]. *Measurement*, 2021, 182: 109454.
- [19] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL](2024-01-21)[2024-04-23]. <https://arxiv.org/abs/2004.10934#>.
- [20] Shi J, Yang J, Zhang Y. Research on steel surface defect detection based on YOLOv5 with attention mechanism[J]. *Electronics*, 2022, 11(22): 3735.
- [21] Wang C Y, Bochkovskiy A, Liao H Y M. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 17-24, 2023[C]. Piscataway: IEEE, 2023.
- [22] Han K, Wang Y, Tian Q, et al. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 13-19, 2020[C]. Piscataway: IEEE, 2020.
- [23] Tan M, Pang R, Le Quoc V. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 13-19, 2020[C]. Piscataway: IEEE, 2020.