



* 文章编号:2096-398X(2023)02-0162-08

基于改进 YOLOv5 的钢材表面缺陷检测

吴 敌¹, 李明辉^{1*}, 马文凯¹, 李睿童¹, 李 艳²

(1. 陕西科技大学 机电工程学院, 陕西 西安 710021; 2. 陕西科技大学 电气与控制工程学院, 陕西 西安 710021)

摘 要:针对当前钢材表面缺陷种类多、形态复杂等原因导致的检测精度低的问题,本文提出了一种基于改进 YOLOv5 目标检测网络的缺陷检测方法. 首先,对于在检测中小目标缺陷易被漏检、错检的问题,增加了小目标检测层;其次,对于缺陷图像表现的背景复杂,且部分缺陷交叉、重叠等问题,引入了 Transformer encoder block 模块和 Convolutional block attention model(CBAM)模块,使网络能更加有效地对抗复杂背景信息,专注于目标对象的检测;最后,使用 NEU-DET 数据集对该改进模型进行了实验. 结果表明,相较于原 YOLOv5 模型,该方法在缺陷检测方面的精度提升了 6.5%;相较于 Faster-RCNN 模型,其精度提高了约 10%. 因此,该方法在钢材表面缺陷检测上,具有较好的检测精度和效率.

关键词:目标检测; 钢材表面缺陷; YOLOv5; 注意力机制

中图分类号:TP391.4

文献标志码:A

DOI:10.19481/j.cnki.issn2096-398x.2023.02.009

Surface defect detection of steel based on improved Yolov5

WU Di¹, LI Ming-hui^{1*}, MA Wen-kai¹, LI Rui-tong¹, LI Yan²

(1. College of Mechanical and Electrical Engineering, Shaanxi University of Science & Technology, Xi'an 710021, China; 2. School of Electrical and Control Engineering, Shaanxi University of Science & Technology, Xi'an 710021, China)

Abstract: In order to solve the problem of low detection accuracy caused by the variety and complexity of the surface defects of steel, we proposed a defect detection method based on improved YOLOV5 object detection network. Firstly, a small target detection layer is added to the problem that small defects are often to be missed and misdetected. Then, the Transformer encoder block module and the Convolutional block attention model(CBAM) are used to deal with the complex background of the defect image, which makes the network more effective against complex background information and more focus on object detection. Finally, the improved model is tested on NEU-DET datasets. The results show that the precision of the proposed method is 6.5% higher than that of the original YOLOv5 model and about 10% higher than that of the Faster-RCNN model. The experimental results show that the method has good accuracy and efficiency in steel surface defect detection.

Key words: object detection; steel surface defects; YOLOv5; attention mechanism

* 收稿日期:2022-06-29

基金项目:陕西省重点研发计划项目(S2023-YF-YBGY-0697); 陕西省咸阳市重点研发计划项目(L2022ZDYFSF047, S2021ZDYF-GY-0244)

作者简介:吴 敌(1997—),男,安徽蚌埠人,在读硕士研究生,研究方向:双目视觉、机械臂控制

通讯作者:李明辉(1972—),男,河南信阳人,教授,博士,研究方向:嵌入式系统与应用,liminghui1972@126.com

0 引言

钢铁行业是国民经济发展的支柱性产业,我国的钢铁年产量逐年上升,目前已位居全球第一,占全球钢铁总产量的半壁江山,2020 年我国粗钢产量达 10.65 亿吨^[1]. 结合 2019 年的数据来看,中国的钢铁行业仍处于上升发展时期,但以生产规模及效益所代表的旧的发展动力已经略显乏力,行业总体盈利水平增长缓慢. 因此,对于钢铁企业而言,需要考虑获取新动力,不断使产业链升级向高端水平迈进. 除了提升装备水平外,更要通过引入新技术使行业从基础制造向智能制造转型,以实现优质、高效、绿色的行业发展目标.

目前,在钢材表面缺陷检测作业方面,常用的检测方法有以下四种:(1)人工检测法;(2)涡流检测法^[2];(3)红外检测法^[3];(4)漏磁检测法^[4]. 其中,后面三种因为技术或成本等原因得不到广泛的应用. 而目前采用最多的还是人工检测方式. 但这种检测方式耗费人力,且效率较低. 近年来,随着机器视觉、深度学习等技术飞速发展,而其相关应用也越来越广. 目前也有了许多有关机器视觉在缺陷检测方面的研究. 基于机器视觉的检测方法不仅成本低,效率高,而且通过逐渐的研究改进能获得接近人工的检测精度.

针对钢板及热轧钢或冷轧带钢的表面缺陷检测问题,目前已有许多使用传统图像检测或深度学习方法进行的研究. 徐科等^[5]通过 Tetrolet 变换将钢板图像分解成子带,并提取子带的矩阵特征,最后通过支持向量机对降维后的特征矢量进行分类. 胡联亭^[6]对于钢板表面缺陷图像首先使用包括滤波算法,直方图均衡化以及 Sobel 边缘检测算子对图像进行预处理,之后使用 Fisher 判别准则提取贡献度较大的特征维度,最后提出 AdaBoost、BK 算法对缺陷进行识别,最高达到了 85.89% 的分类精度. 以上方法虽然在钢材的表面缺陷检测上表现出了一定的效果,但其无论是在图像的特征提取还是在特征识别方面使用的都是传统算法,这导致模型的检测精度较低且提升困难.

近年来,人工神经网络、深度学习等技术的兴起,其在图像处理领域取得了许多惊人的成就. 因此,也有许多的研究学者将深度学习技术与缺陷检测工作相结合,取得了许多较好的结果^[7]. 杨水山^[8]提出基于 CNN 的带钢表面缺陷检测算法,实现了对带钢表面是否具有缺陷的判断,准确率达到了 99.99% 以上. 李鑫灿^[9]在研究中对 U-Net 语义分

割网络进行改进,提出基于分组卷积和特征均匀重组的残差网络结构来实现多尺度的特征提取,通过过渡特征和密集连接改进得到缺陷精细分割网络 Nabl-net,并引入注意力机制使空间能够进行自适应调整以及深度监督机制实现网络剪枝. 师伟婕等^[10]则做了 U-Net 和 Res-UNet 在钢铁图像缺陷检测工作上的对比,结果显示 Res-UNet 在钢铁图像的缺陷位置和缺陷类别的识别方面更准确.

以上方法有的专注于目标缺陷分类,有的专注于图像的分割,却未将二者结合,在进行缺陷分类的同时做到有效定位. 这将极大限制其在工业生产中的应用. 方叶祥等^[11]在金属表面的缺陷研究中使用了改进的 YOLOv3 算法,主要改进包括使用直方图均衡化对图像进行预处理,使用数据增强方法模拟现实工作环境以及优化损失函数来提高模型对表面缺陷分类的准确性. 刘洋^[12]为了提高钢材瑕疵检测的速度,提出了基于 Tiny-YOLOv3 的 R-Tiny-YOLOv3 算法,该算法加入了残差网络以及空间金字塔池化 SPP 模块最后选择 CIOU 作为损失函数,该算法对钢材瑕疵检测精度达到 71.5%,检测速度达到 39.8 帧每秒,能够满足嵌入设备的实时检测需求.

以上研究通过不同的方式分别改进了缺陷检测模型各个方面的性能,但其在针对钢铁表面的独特性缺陷上缺乏针对性的处理,特别是对小目标缺陷及钢材表面复杂纹理背景图像的识别和处理能力不足. 这导致了缺陷检测种类的不全面以及部分缺陷类型的检测精度低,从而导致了模型整体性能的不理想. 本文针对以上问题,首先引入小目标检测层,增强检测网络对小目标缺陷的检测能力. 此外,还将原网络中的部分 CSP 结构替换为 Transformer encoder block,并加入 CBAM (Convolutional block attention module)模块以处理复杂的背景纹理信息.

1 相关工作

1.1 钢材缺陷检测背景

基于机器视觉方法的缺陷检测流程如下,首先通过工业相机获取钢材表面的图像,在对图像进行简单的标号排序等预处理后将处理后的图像输入相应算法模型,由目标检测算法给出图片中缺陷的种类以及位置,再通过之前记录的信息反向推算出相应缺陷在实际钢材中位置. 整个流程可以由一整套自动化程序实现,而本次研究的重点便在于其关

键的目标检测算法。

在钢材生产加工的过程中,由于工艺不足或者操作失误会在钢材表面留下各种缺陷。通常钢材表面的缺陷有以下七种:裂纹缺陷、划伤、折叠缺陷、耳子(突起)缺陷特征、结疤(重皮)缺陷、焊疤缺陷、端部毛刺缺陷。

由以上可见钢材表面的缺陷种类多样。不同种类的缺陷特征差异较大导致不同缺陷在检测时的精度差异较大,部分小目标缺陷检测的精度不理想,同时钢铁表面缺陷图像的背景复杂,纹理混乱,检测目标之间也会出现覆盖交叉的情况,具体情况如图 1 所示。本文针对以上问题,结合钢材检测的实际特点和具体要求,在 YOLOv5 网络模型的基础上进行补充修改和调整,进行实验。并对比其他网络在相同条件下的实验结果。

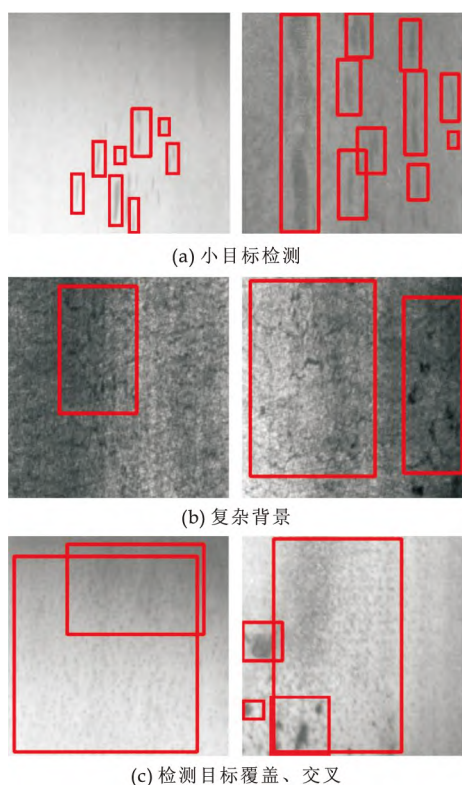


图 1 数据集中的复杂缺陷情况

1.2 目标检测模型

目前,基于深度学习理论的目标检测算法主要可分为两条技术路线:Anchor based(基于候选框)和 Anchor free(不基于候选框)。其中,Anchor based 方法中又可以分为 One-stage(单阶段检测)和 two-stage(二阶段检测)。其中属于单阶段检测的 SSD、YOLO 算法以及二阶段检测方法中的 RCNN、Faster RCNN 等都在表面缺陷检测中有广泛的应用。YOLOv5 网络是在 YOLOv3,

YOLOv4 的基础上进一步优化而来,在保证检测精度的前提下 YOLOv5 迭代出四个不同版本,使 YOLOv5 在使用起来更加灵活方便。王淑青等^[13]的研究表明,YOLOv5 在缺陷检测方面较其他网络如 Faster RCNN、SSD、YOLOv4 有更优异的表现。

以 YOLOv5s 的网络结构为例,如图 2 所示,其中黄色框中的为网络本体,蓝色框中的为对网络中模块的分解说明。其结构可分为 4 个部分:

(1)输入端(input)。YOLOv5 在输入端采用 Mosaic 数据增强,自适应图像缩放来增强网络推理的能力并提高效率,使其对不同的输入具有较强的适应性;

(2)主干网络(backbone)。在主干网络中 YOLOv5 使用 Focus 结构和 CSP 结构来提取输入图像数据中的特征信息,该设计有效降低了计算成本和内存成本并增强了卷积网络的学习能力;

(3)融合网络(Neck)。Neck 部分主要是对之前提取的特征进行整理使其能够被更好的利用从而提取出更有效的特征,此处 YOLOv5 采用 FPN+PAN 的结构,从自顶向下和自底向上两个方向出发,同时增强了语义信息和位置信息的提取,同时还设计的 CSP2 结构,进一步加强了网络的特征融合能力;

(4)检测头(Prediction)。检测头一般都是对损失函数的设计,损失函数又可以分为分类损失函数(Classification Loss)和回归损失函数(Bounding Box Regression Loss),YOLOv5 在此处采用 CI-OU_Loss 作为回归损失函数,最后采用加权 NMS(非极大值抑制)来筛选目标框。

2 基于改进 YOLOv5 的目标检测

2.1 增加小目标检测层

由于钢材表面缺陷种类较多,形态特征较复杂,在检测过程中常有小目标缺陷被漏检或错检。对此,本文在原有的三个检测层的基础上增加小目标检测层,通过对小范围内图像特征的识别,来增加网络对小目标检测的能力。具体方案如图 3 所示。图 3 中的红色锚框为原始锚框,分别为 P3, P4, P5。其中, P3 用于浅层的特征图,由于浅层特征图包含较多的低层级信息,适合用于检测小目标,所以这一特征图所用的锚框尺度较小; P5 应用于深层特征图,其特征图中具有更多高层级的信息,如轮廓、结构等信息,适合用于大目标的检测,所以这一特征图所用的锚框尺度较大;同理,

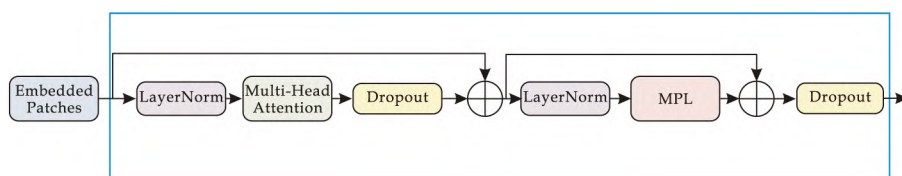


图 4 Transformer encoder block 结构图

CBAM(Convolutional block attention module)是一个简单有效的注意力模块,可以集成到 CNN 体系结构当中,并且可以以端到端的方式训练.对于给定的特征数据,CBAM 沿通道和空间这两个独立维度依次推断注意力图.然后将注意力图和输入的图像特征相乘,进行自适应特征细化.CBAM 模块的结构如图 5 所示,使用两个顺序子模块用于细化通过 CBAM 的特征图,同时还使用了残差网络的结构来增强对特征的提取.该模块主

要应用在 neck 部分,配合 Transformer encoder block 模块和 Conv 模块来加强网络的性能.改进后的网络整体结构如图 6 所示.

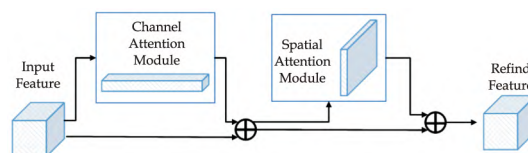


图 5 CBAM 结构图

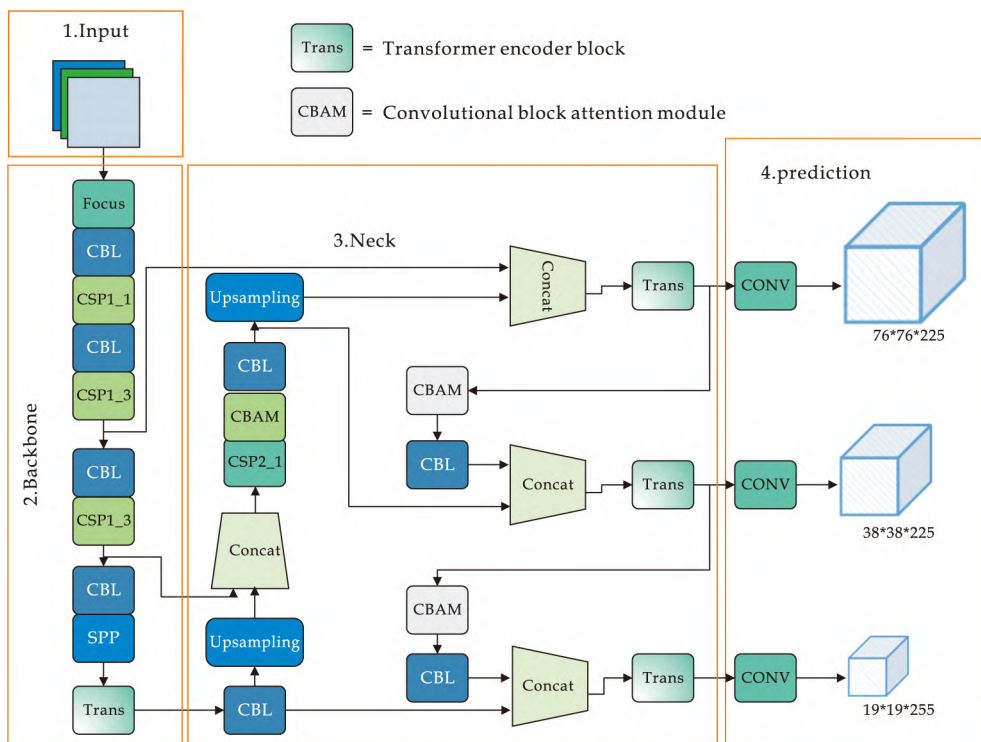
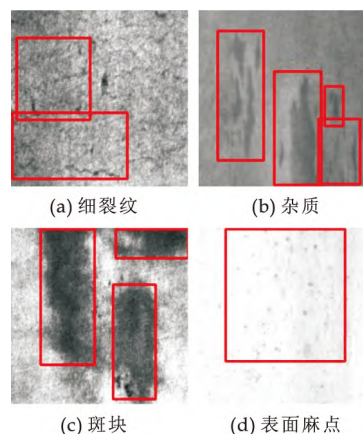


图 6 改进后的 YOLOv5 结构图

3 实验结果与分析

3.1 数据来源

本次钢材表面缺陷检测实验采用东北大学钢材表面缺陷数据集(NEU-DET)^[16],该数据集包含了 6 种常见的钢材表面缺陷:crazing(细裂纹)、inclusion(杂质)、patches(斑块)、pitted surface(表面麻点)、rolled-in scale(滚轧鳞片)、scratches(划痕),具体表现如图 7 所示.数据集中的图像分辨率为 200×200 像素,一共 1 800 张图片,本次实验将其划分为训练集 1 422 张,验证集 378 张.



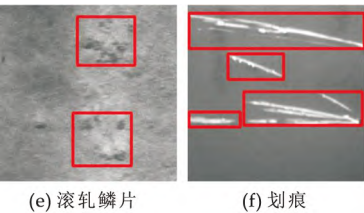


图 7 NEU-DET 数据集缺陷情况

3.2 实验设置

在本文的实验环境,使用 NVIDIA GTX1050Ti GPU 进行训练和测试,CUDA 的版本是 11.3,cuD-NN 的版本是 8.2.1,Pytorch 的版本是 1.10.2,硬件和软件环境的详细信息如表 1 所示.

表 1 实验环境

OS	Win 10 64-bit
CUDA	11.3
CPU	Intel CPU i7-7700
cuDNN	8.2.1
GPU	NVIDIA 1050Ti (4 G)
Python	3.7.0
Pytorch	1.10.2

3.3 实验相关的评价指标

关于目标检测领域的评价指标首先要考虑检测框的准确度,该指标主要由 IoU 指标衡量,即检测框和目标框面积的交并比,如图 8 所示.实际运算中根据 IoU 大小判断预测是否有效,一般当 $IoU \geq 0.5$ 时判定为预测正确.

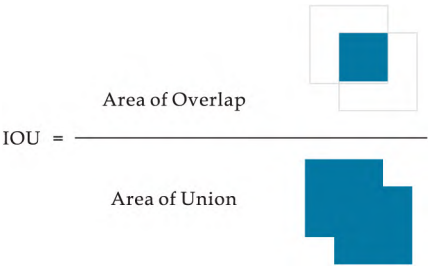


图 8 IoU 计算方法示意图

而对检测结果的划分可以用图 9 中的混淆矩阵表示.根据预测结果和实际情况的组合可分为以下四种情况 TP(实际的和预测的标签都是“真”),FN(实际标签为“真”,预测标签为“假”),FP(实际标签为“假”,预测标签为“真”),TN(实际的和预测的标签都是“假”).根据以上四种情况的统计结果可以得到以下三个指标(公式(1)~(3)).精度(Precision,P)、召回率(Recall,R)和准确率(Accuracy,A).在此基础上的综合评价指标为平均精度(Average Precision,AP)和平均精度均值

(mean AP over classes,mAP).其中 AP 表示同一类别下,所有检测结果的平均精度,而 mAP 则表示所有检测类别平均精度的均值.

		Prediction	
		Positive	Negative
Actual	True	TP	FN
	False	FP	TN

纵列T/F:表示预测的对错
横行P/N:表示预测的结果

图 9 混淆矩阵

$$Percision = \frac{TP}{TP + FP}$$
 (1)

$$Recall = \frac{TP}{TP + FN}$$
 (2)

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$
 (3)

3.4 结果讨论

为证明本文改进后模型的检测效果,以及验证具体改进措施起到的作用,本次实验除了设置原 YOLOv5s 网络组以及 Faster-RCNN 网络组(分别称作 V5-0 和 FF-RCNN)作为对照组外,还设置了三组消融实验组:(1)增加小目标检测层组;(2)增加 Transformer encoder block 模块和 CBAM 结构组;(3)最终优化组.为方便描述,将以上三组消融实验组依次称作 V5-1、V5-2、V5-3.

其中,实验的总体结果如表 2 所示.根据表中的数据显示,优化后的网络(V5-3)相对于原网络(V5-0),mAP 大约有 6%的提升,相对于 Faster-RCNN 网络提升了大约 10%.与消融实验组(V5-1 和 V5-2)的结果进行对比表明,针对于原网络做出的改进措施都起到了良好的效果,相对于原网络大约提升了 5%~6%的精度.

表 2 各网络检测结果

组别	评价指标		
	P	R	mAP@0.5
V5-0	0.577	0.70	0.682
V5-1	0.745	0.726	0.727
V5-2	0.734	0.704	0.732
V5-3	0.754	0.675	0.748
FF-RCNN	0.621	0.723	0.647

具体的消融实验结果见图 10 所示.根据图 10 的检测结果显示,改进方法一通过增加小目标检测

层的方法能够有效提高对于 rolled-in_scale(轧制氧化皮)和 scratches(划痕)这两种缺陷的检测能力;而方法二则通过引入注意力模块的方式有效增强了网络对 crazing(细裂纹),pitted-surface(表面点蚀),轧制氧化皮(rolled-in scale)这些缺陷的

识别,证明了其对复杂背景信息有效的处理能力.最后,最终优化组能够将两种优化方案的优点有效结合,使其在绝大多数的情况下的检测结果都能优于二者,做到了有效的取长补短.

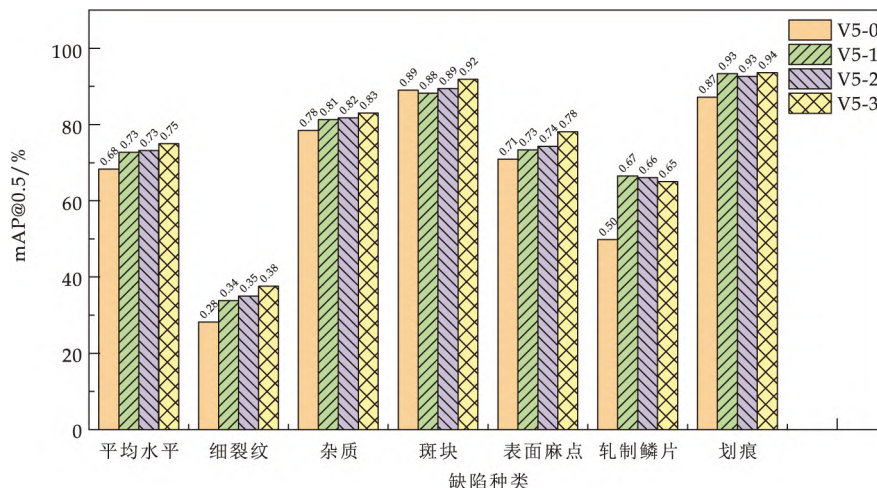


图 10 各种类缺陷检测情况

算法的收敛情况如图 11 所示. 根据图中曲线显示,改进后的网络相对于原网络有明显的改进. 其中最优的权重在 V3 网络第 230 次迭代时产生.

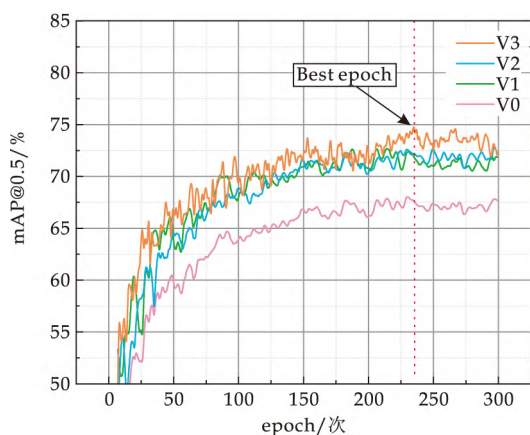
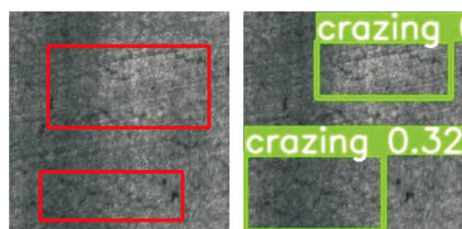
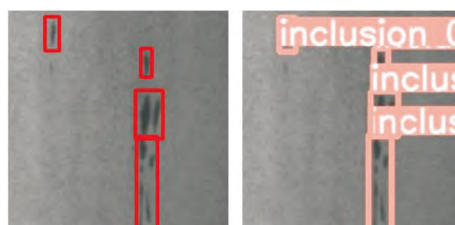


图 11 mAP 随迭代次数变化情况

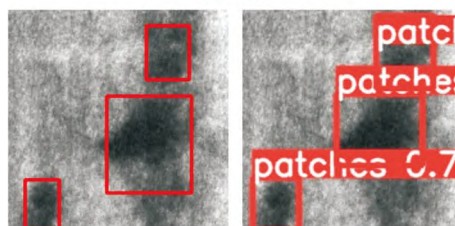
图 12 展示了部分检测结果(左为原标记图片,右为检测结果图片). 图中检测结果显示,优化后的检测网络的检测结果与数据集标注结果高度吻合.



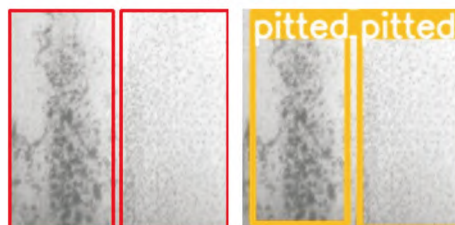
(a) 细裂纹



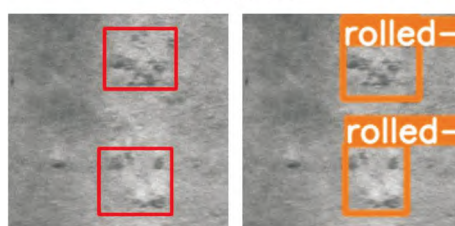
(b) 杂质



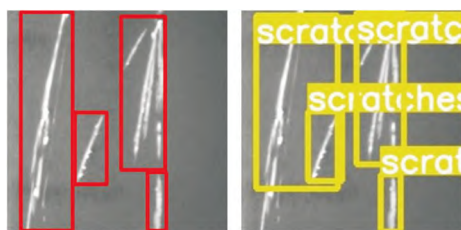
(c) 斑块



(d) 表面麻点



(e) 滚轧鳞片



(f) 划痕

图 12 检测结果对比情况

4 结论

本文使用一种基于 YOLOv5 的改进网络对钢材表面缺陷进行检测,使用 NEU-DET 数据集进行实验验证.该方法通过增加小目标检测层来增强对小目标缺陷的检测效果,同时通过引入 Trans-former encoder block 模块和 CBAM 模块,使网络更有效的对抗复杂背景,更有效聚焦于任务目标的检测.

实验表明,本文中所提的改进模型相较于原模型,在总体平均检测精度 (mAP) 上提升了 6% 左右,相较于以 Resnet50 为骨干网络的 Faster-RCNN 模型,mAP 提升约 10%.同时通过消融实验证明,本文所提的两种改进措施均能对原有网络在钢铁表面缺陷检测方面做出有效提升.并且融合后的网络也能较好的发挥各改进网络的优点,使检测精度有更进一步的提升.然而,该改进模型也存在一些不足.如针对个别特征,优化方案融合效果表现不佳,以及对个别种类缺陷的整体检测精度不足.因此,该优化网络模型还有进一步优化的空间.

依据理论,本文中所提的方法同样可以用于其它工业应用场景中,特别是对于存在小目标以及背景较复杂的目标检测领域.同时,对于其表现的不足,需要进一步的研究改进,从而使模型的总体检测精度能有更进一步的提升,以此来更好的满足具体工业生产的需求.

参考文献

[1] 庞永辉. 基于深度学习的带钢表面缺陷检测与识别方法

研究[D]. 郑州:华北水利水电大学,2021.

- [2] 黄凤英. 钢轨表面裂纹涡流检测定量评估方法[J]. 中国铁道科学,2017,38(2):28-33.
- [3] 徐 欢,殷晨波,李向东,等. 基于有限元的焊缝缺陷红外热像检测研究[J]. 热加工工艺,2019,48(17):122-127,133.
- [4] 田思洋,徐 科,周东东. 钢板表面缺陷图像的跨域翻译[J]. 中国冶金,2019,29(4):54-59.
- [5] 徐 科,王 磊,王璟瑜. 基于 Tetralet 变换的热轧钢板表面缺陷识别方法[J]. 机械工程学报,2016,52(4):13-19.
- [6] 胡联亭. 基于机器视觉的钢板表面缺陷识别研究[D]. 武汉:武汉科技大学,2018.
- [7] 柴 利,任 磊,顾 铿,等. 基于视觉感知的表面缺陷智能检测理论及工业应用[J]. 计算机集成制造系统,2022,28(7):1 996-2 004.
- [8] 杨水山. 冷轧带钢表面缺陷机器视觉自动检测技术研究[D]. 哈尔滨:哈尔滨工业大学,2009.
- [9] 李鑫灿. 基于改进 U-Net 的钢材缺陷检测分割算法研究[D]. 哈尔滨:哈尔滨工业大学,2020.
- [10] 师伟婕,黄静静,王茂发. 基于两种 U 型网络的钢铁图像缺陷检测方法对比[J]. 北京信息科技大学学报(自然科学版),2021,36(1):63-68.
- [11] 方叶祥,甘 平,陈 俐. 金属表面缺陷检测的改进 YOLOv3 算法研究[J]. 机械科学与技术,2020,39(9):1 390-1 394.
- [12] 刘 洋. 基于深度学习的工业钢材瑕疵检测研究[D]. 合肥:安徽建筑大学,2021.
- [13] 王淑青,顿伟超,黄剑锋,等. 基于 YOLOv5 的瓷砖表面缺陷检测[J]. 包装工程,2022,43(9):217-224.
- [14] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, New York; IEEE Transactions on Cybernetics, 2020:12 993-13 000.
- [15] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]// Proceedings of the European conference on computer vision, Munich; Springer International Publishing, 2018:3-19.
- [16] He Y, Song K, Meng Q, 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): 1 493-1 504.

【责任编辑:蒋亚儒】