

张琴. 基于改进 YOLOv7 的钢材表面缺陷检测[J]. 智能计算机与应用, 2024, 14(10): 182-188. DOI: 10.20169/j.issn.2095-2163.241026

## 基于改进 YOLOv7 的钢材表面缺陷检测

张琴

(福州职业技术学院 信息工程系, 福州 350108)

**摘要:** 针对钢材表面缺陷纹理特征不明显、不同缺陷类间差异不明显和缺陷尺度变化剧烈等问题, 本文设计了一个纹理信息增强模块(Texture Information Enhancement Module, TIEM)来保留主干网络上采样丢失的细节纹理特征信息和加强主干网络对不规则缺陷的空间建模能力; 在颈部网络融入跳跃连接的多尺度自适应卷积模块(Multi-scale Adaptive convolution module with Skip Connections, MASC)来增强网络对不同尺度缺陷目标的感知能力, 进而增强小目标的细粒度特征和大目标的高层语义信息, 增强检测器的全局感知能力。以 YOLOv7 为基线模型, 在公开数据集 NEU-DET 上, 改进后的模型比基线模型  $mAP_{50}$  和  $mAP_{50:95}$  分别提高了 3.0% 和 2.1%, 并优于现阶段其他主流目标检测器。

**关键词:** 目标检测; YOLOv7; 缺陷检测; 纹理信息; 多尺度自适应卷积

中图分类号: TP391.41

文献标志码: A

文章编号: 2095-2163(2024)10-0182-07

### Steel surface defect detection based on improved YOLOv7

ZHANG Qin

(Department of Information Engineering, Fuzhou Polytechnic, Fuzhou 350108, China)

**Abstract:** Aiming at the problems of inconspicuous texture characteristics of steel surface defects, inconspicuous differences between different defect categories, and severe defect scale changes, this paper designs a Texture Information Enhancement Module (TIEM) to preserve the loss of detailed texture feature information in the previous stage of the backbone network and strengthen the spatial modeling ability of the backbone network for irregular defects; a multi-scale adaptive convolution module (multi-scale Adaptive convolution module with Skip Connections, MASC) is integrated into the neck network to enhance the perception ability of the network to targets of different scales, and then enhance the fine-grained features of small targets and the high-level semantic information of large targets, and enhance the global perception ability of the detector. Taking YOLOv7 as the baseline model, on the public dataset NEU-DET, the improved model has improved by 3.0% and 2.1% compared with the baseline model  $mAP_{50}$  and  $mAP_{50:95}$ , respectively, and is better than other mainstream target detectors at this stage.

**Key words:** object detection; YOLOv7; defect detection; texture information; multi-scale adaptive convolution

## 0 引言

钢材表面缺陷检测是保证钢材质量的重要环节。不同产品的钢材表面缺陷种类繁多, 相应的特性因工厂和操作员的不同而存在差异。此外, 由于生产过程的变化多样, 缺陷也有着不同表现。例如, 在热轧带钢的生产过程中, 主要缺陷包括开裂(crazing)、夹杂物(inclusion)、斑块(patch)、点蚀表面(pitted surface)、轧制氧化皮(rolled-in scale)和划痕(scratches), 如图1所示<sup>[1-3]</sup>。面对庞大的数据量, 人工视觉检测的方法不仅耗时而且容易受到

质检员的主观判断影响, 视觉判断容错率低。因此, 开展钢材表面缺陷的自动化检测研究对于钢材的智能化生产意义重大。

近年来, 基于机器视觉的表面缺陷检测技术随着计算机的发展而逐渐成熟, 现已能够在图像中分辨缺陷的具体类型和位置。基于机器视觉的表面缺陷检测方法主要分为传统方法与深度学习方法两种。传统方法依靠有效的特征提取器和分类器来进行缺陷位置的定位和具体类别的区分, 使用经典的特征提取器定向梯度直方图<sup>[1]</sup>、局部二值模式<sup>[2]</sup>或者灰度共生矩阵<sup>[3]</sup>提取不同的特征作为分类器的

基金项目: 2022年度福建省中青年教育科研项目(JAT220652)。

作者简介: 张琴(1986-), 女, 硕士, 讲师, 主要研究方向: 计算机视觉。Email: zqsrhx@163.com

收稿日期: 2024-03-27

输入,后续使用支持向量机<sup>[4]</sup>分类器进行不同类别的区分。Song 等学者<sup>[5]</sup>用完全局部二值模式的方法实现了带钢表面缺陷的分类。Liu 等学者<sup>[6]</sup>使用局部二值模式提取缺陷特征,根据提取的特征类型进行带钢表面缺陷分类。但是传统方法都需要手工制作缺陷特征,面对多场景的应用环境,工作量大、鲁棒性差,不适用于大规模的缺陷检测生产过程。

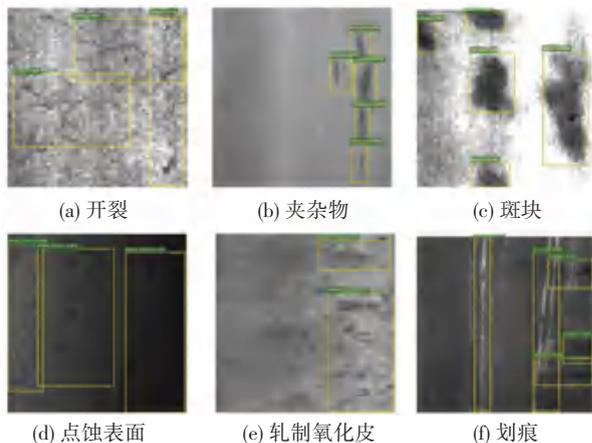


图 1 钢材表面缺陷的不同类别<sup>[1-3]</sup>

Fig. 1 Different categories of steel surface defects<sup>[1-3]</sup>

随着深度学习的不断发展,基于卷积神经网络的目标检测方法不断被应用于表面缺陷检测场景。针对复杂多样的印制电路板检测问题,Ding 等学者<sup>[7]</sup>提出了一种表面缺陷检测网络,该方法的特征金字塔融合了多尺度和金字塔层次结构,缺点是检测速度较低。程婧怡等学者<sup>[8]</sup>提出一种基于改进 YOLOv3<sup>[9]</sup>的缺陷检测算法,融合浅层特征与深层特征,该方法适用于尺寸小和特征模糊的目标。吴桐等学者<sup>[10]</sup>利用 YOLO<sup>[11]</sup>网络实现激光焊缝瑕疵检测,在输入网络之前对数据集加入合适的锚框,结合多尺度特征融合技术提高检测精度,对激光焊缝瑕疵的检测率超过 94%。针对复杂背景下对输电线路销钉的缺陷检测,李瑞生等学者<sup>[12]</sup>提出改进 SSD<sup>[13]</sup>模型,该方法结合经典残差网络和多层级特征融合策略,采用卷积拆分和权值量化减小模型规模,对输电线路销钉缺陷检测的召回率达到 80%以上。吴越等学者<sup>[14]</sup>提出了基于 Faster R-CNN<sup>[15]</sup>的钢板表面缺陷检测方法,采用 ROI Align 作为池化模块,改善对于小目标的检测精度。李庆党等学者<sup>[16]</sup>提出了基于 YOLOv3 的钢板缺陷检测方法,采取了增加尺度的方法来提高缺陷检测的精度。李维刚等学者<sup>[17]</sup>提出了基于 YOLOv3 的缺陷检测算法,通过融合浅层与深层特征,提高网络对缺陷的特征提取能力。然而直接使用通用目标检测器进行钢材

表面缺陷检测受到以下与表面缺陷相关的独特问题的挑战。

(1) 纹理特征信息不明显。由图 1 可知,低对比度的灰度图像不同于 RGB 图像,不同缺陷纹理特征弱化,点蚀表面、轧制氧化皮和开裂缺陷目标与图像背景非常相似,目标检测器很难对相应目标进行精确定位。

(2) 类间差异不明显。由图 1 可知,缺陷类型轧制氧化皮、开裂和点蚀表面纹理特征相似,将不同类别缺陷进行准确分类是挑战性任务。

(3) 目标尺度变化剧烈。图 1 中,不同缺陷类型有不同的长宽比,有些缺陷有着不小的长宽比,例如斑块缺陷,有些检测框高度和宽度有显著差异;不同缺陷类型检测框大小差异明显,例如点蚀表面和夹杂物检测框大小比例差异大,目标尺度的剧烈变化通常会降低目标检测器的性能上限。

为了解决以上问题,本文以 YOLOv7<sup>[18]</sup>为基本框架,在主干网络部分,设计了一个纹理信息增强模块(Texture Information Enhancement Module, TIEM)来保留主干网络上一阶段下采样丢失的细节纹理特征信息,加强主干网络有效地捕捉低层次纹理位置信息,增强目标检测器对不同缺陷类型纹理信息的区分和不同缺陷目标的定位效果;在颈部融合网络,提出一种跳跃连接的多尺度自适应卷积模块(Multi-scale Adaptive convolution module with Skip Connections, MASC)来增强网络对不同尺度目标的感知能力,并通过多层跳跃连接实现多模块特征增强,生成新的增强特征,以加强特征复用性,进一步增强小目标的细粒度特征和大目标的高层语义信息,进而提高目标检测器对多尺度目标的鲁棒性。

## 1 YOLOv7 算法简介

多年来,YOLO 系列的目标检测器一直是工业应用中广受欢迎的目标检测框架,YOLO 系列目标检测器的迭代更新,加上自身出色的运行速度,扩展了其在众多硬件平台(例如云 GPU、智能机器人和手机嵌入式设备等)的使用。YOLO 系列的开创性工作 YOLOv3,首次配设于单级检测器上,并在后来进行了实质性改进。YOLOv4<sup>[19]</sup>将检测框架重组为几个独立的部分(主干、颈部和头部),并设计了适合在单个 GPU 上训练的框架。目前,YOLOv5、YOLOX<sup>[20]</sup>、YOLOv6<sup>[21]</sup>、YOLOv7 都是部署高效检测器的优质选择。

当前的目标检测器根据模型尺寸通过缩放技术

调节网络的宽度因子和深度因子,其中 YOLOv7 指出基于这种复合拼接的模型,必须同时考虑 2 个因子,不能只调整宽度因子或深度因子。如果单独放大网络深度会使得过渡层的输入通道和输出通道的比例发生变化,会导致模型硬件使用量的降低。因此,对于基于复合拼接的模型,必须提出相应的复合模型缩放方法。当缩放一个计算块的深度因子时,还必须考虑该块的网络宽度的变化。最后,在过渡层上进行相同变化量的宽度因子缩放。基于以上分析,YOLOv7 提出了高效的聚合层网络(Efficient Layer Aggregation Networks, ELAN)。ELAN 网络结构如图 2 所示,这种复合缩放方法可以保持模型在初始设计时的属性,并保持最优结构。该方法能有效减少现有实时目标探测器 40% 左右的参数和 50% 的计算量,推理速度更快,检测精度更高。

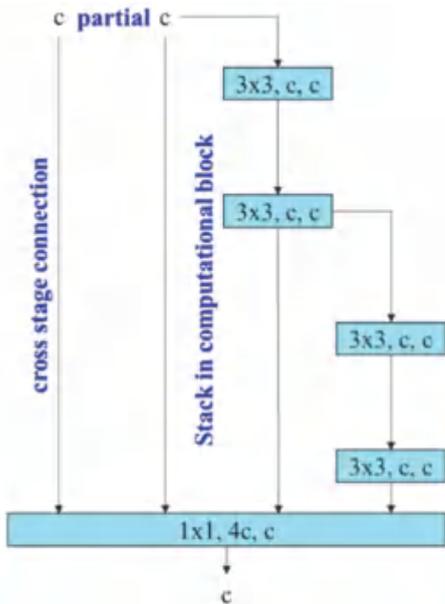


图 2 ELAN 网络结构图

Fig. 2 ELAN network structure diagram

在网络训练的优化策略方面,YOLOv7 引入了结构重参数化和新的动态标签分配方法。对于前者,YOLOv7 在检测头部分引入了结构重参数化卷积,通过这种技术,模型能够学习到更加丰富的预测特征层。对于后者,YOLOv7 在颈部网络的自上而下的信息流向部分增加了额外的辅助检测头计算辅助损失来调整层网络的权重,并且由主要检测头为指导,生成由粗到细的分层标签,分别用于辅助检测头和主要检测头学习。通过让较浅的辅助检测头直接学习主要检测头已经学习的信息,主要检测头将更能专注于学习尚未学习的残留信息。并在训练过程中通过调整辅助检测头和主要检测头粗细标签的

重要程度来避免预测时产生的不良先验影响。

## 2 YOLOv7 算法改进

本文以现阶段流行的目标检测器 YOLOv7 为基线模型,针对钢材图像对比度低、缺陷纹理信息不明显、类间差异小和目标尺度变化剧烈的问题,提出了纹理信息增强模块和跳跃连接的多尺度自适应卷积模块来进一步提高 YOLOv7 在钢材表面缺陷检测的性能。图 3 展示了目标检测器网络结构和检测流程。钢材表面缺陷图像首先经过主干特征提取网络,  $C_i$  表示主干网络提取特征后不同的分层阶段,下角标  $i$  越大特征图越小,共 5 个阶段,之后将  $\{C_3, C_4, C_5\}$  送入颈部融合网络,经多层特征融合后,形成特征金字塔  $\{P_3, P_4, P_5\}$  结构进行多尺度目标的边界框回归和类别分类,最后经过非极大值抑制(Non Maximum Suppression, NMS)检测出最终结果。

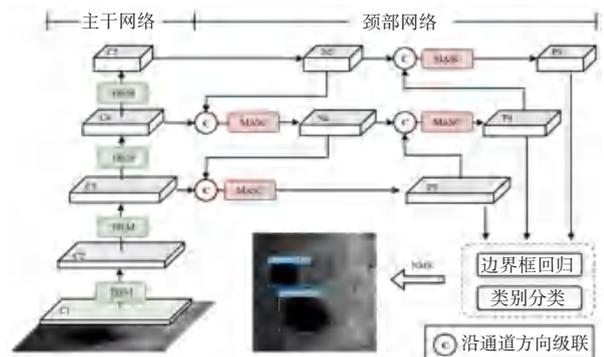


图 3 总体框架图

Fig. 3 Overall framework diagram

本文在 YOLOv7 主干网络部分相邻层阶段的下采样之后插入 TIEM,用来恢复下采样阶段丢失的纹理特征信息和加强对不规则缺陷类型的建模能力,从而提高主干网络对钢材表面缺陷图像纹理信息的特征提取能力;在颈部网络部分,不同层特征沿通道方向级联操作后添加 MASC 来获取多尺度特征表达,MASC 通过动态调整多分支卷积核的分配权重,在单个特征图内获取自适应感受野大小,提高颈部网络对特征图全局特征以及局部特征的感知能力,从而增强尺度变化剧烈目标的检测鲁棒性。

### 2.1 纹理信息增强模块

因为 CNN 中的池化层或下采样操作会丢失图像原本的纹理特征信息,但这些细节纹理信息对于钢材表面的缺陷检测至关重要。本文提出一种纹理信息增强模块,旨在通过恢复主干网络上一阶段的细节纹理特征来补偿因池化或者下采样操作带来的细节丢失问题以及加强对不规则缺陷外观的空间建

模能力。具体来说, TIEM 有 2 个主要分支, 如图 4 所示。图 4 中, 白色背景、黑色边框的矩形框表示了每一个操作后的输出维度,  $C$ 、 $H$ 、 $W$  分别表示通道数、高度和宽度, 假定输入特征图的大小为  $C \times H \times W$ 。不同彩色背景颜色的黑色外边框的矩形框表示不同的卷积操作,  $k$  表示卷积核尺寸,  $s$  表示卷积的步长; 跟随在后面的白色背景外边框的矩形框表示卷积操作后特征的输出维度,  $C$  表示通道数,  $H$  和  $W$  分别表示输出特征的高和宽。

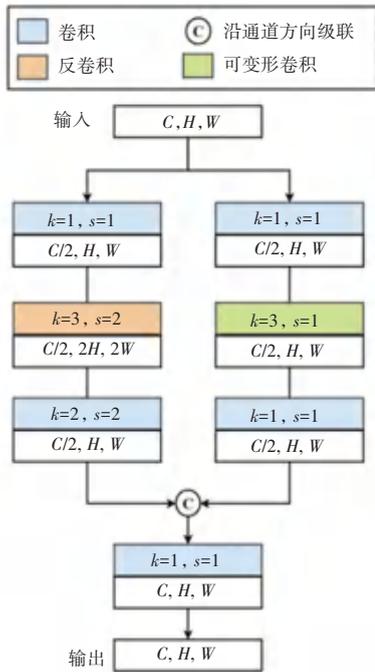


图 4 TIEM 网络结构图

Fig. 4 TIEM network structure diagram

对于左分支, 应用卷积核尺寸为 1、步长为 1 的卷积过渡层, 学习输入特征和输出特征通道之间的线性投射关系, 并将输入特征通道数减半, 这一设计的主要目的是压缩计算空间和减小计算复杂度。之后, 利用卷积核为 3、步长为 2 的反卷积, 将特征图尺寸进行加倍操作处理, 由  $H \times W$  调整为  $2H \times 2W$ , 卷积核大小为 3 的反卷积操作能融合特征像素点周围的局部空间信息, 将局部知识放大并恢复上一阶段原始特征图的细节纹理信息。此后, 应用卷积核为 2、步长为 2 的卷积层, 进一步凝练反卷积操作恢复的纹理语义信息, 并将特征图减半, 从  $2H \times 2W$  调整为  $H \times W$ 。

对于右分支, 同样地和左分支应用相同的卷积过渡层, 学习通道之间的线性关系, 并将输入特征通道数减半。之后, 将特征图送入卷积核为 3、步长为 1 的可变形卷积。相比于普通卷积, 可变形卷积可以加强神经网络的几何变换建模能力, 提高目标检

测器对不规则缺陷目标的提取能力和空间建模能力, 能迅速捕获低对比度下难于定位的缺陷的精确位置和区分不同几何外观类型的缺陷类别。稍后, 送入卷积核为 1、步长为 1 的卷积层, 进一步凝练可变形卷积获取的高级空间语义特征信息。

最后, 将左右分支的输出特征沿通道方向进行级联, 使用卷积核为 1、步长为 1 的卷积层学习 2 个分支的融合特征, 得到 TIEM 最终的输出增强特征, 这不仅强化反卷积和可变形卷积对低级纹理信息的恢复能力和表征能力, 还可以指导不同支路学习更多样化的特征组合。

## 2.2 跳跃连接的多尺度自适应卷积模块

在卷积神经网络中, 上下文信息的使用程度大致取决于感受野的大小。对于缺陷检测, 每种类别的缺陷都具有不同的尺度大小, 有的存在于图像的局部小区域, 有些则贯穿整个图像。目标检测网络需要不同的感受野大小来提高对尺度变化剧烈的缺陷目标的表达能力, 从而提高对图像中不同缺陷的整体感知能力。

提高神经网络感受野大小的方式是多尺度特征表达, 然而现有的大多数方法都是以分层的方式或者并行分支结构来表示多尺度特征, 但是一般的卷积算子结构缺乏特征图高维特征通道之间的内在联系信息, 在单个特征图内部缺乏多样性的表达方式和高级视觉上下文信息<sup>[22]</sup>。本文提出一种在单个特征图内部进行多尺度计算的跳跃连接的多尺度自适应卷积模块, 能够在单个特征图内部具有不同感受野信息流向, 提取输入特征更细粒度的特征, 增强上下文的特征表达能力。

MASC 的网络结构如图 5 所示。图 5 中, AC 表示自适应卷积模块 (Adaptive Convolution module, AC), 输入特征通过一个双分支卷积结构分别获取输出特征  $F_1$  和  $F_2$ , 双分支结构拥有不同的卷积核大小来获取不同的感受野大小, 用来捕获输入特征不同空间域上的上下文信息。之后, 对特征  $F_1$  和  $F_2$  赋值可学习权重  $K_1$  和  $K_2$ , 将  $K_1$  和  $K_2$  进行 *Softmax* 处理并更新, 动态调整双支路不同感受野的加权配比, 从而对输入特征起到动态调节感受野的作用, 加强对输入特征的不同上下文信息的感知能力。而 MASC 是在 AC 的基础上增加了多尺度的跳跃连接功能, 首先输入特征经过卷积核为 1 的卷积层进行通道维度的扩增, 将输入特征拓展到高维空间。紧接着, 对高维空间特征沿通道方向分离, 将特征图均匀地分成 3 个特征图子集, 由  $X_i$  表示, 其中每一个子集的特征图尺寸相同,

对每一子集进行自适应卷积模块处理,对应每一个  $X_i$  有一个输出  $Y_i$ ,对于每一个  $X_{i+1}$  和上一个  $X_i$  的输出  $Y_i$  执行跳跃连接并相加,进而送入自适应卷积模块, MASC 的整个计算过程可描述为:

$$Y_i = \begin{cases} AC(X_i), & i \leq 3 \\ AC(X_i + Y_{i-1}), & 1 < i \leq 3 \end{cases} \quad (1)$$

其中,  $AC$  表示自适应卷积模块;  $X_i$  表示第  $i$  个输入特征图子集;  $Y_i$  表示经过  $AC$  后的第  $i$  个输出特征。

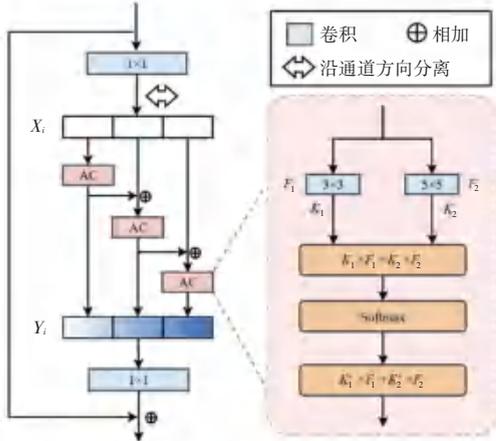


图5 MASC网络结构图

Fig. 5 MASC network structure diagram

最后对  $Y_i$  执行沿通道方向的级联操作,再使用卷积核大小为1的卷积层将级联特征融合多尺度信息,降低到低维空间并恢复到和输入特征相同维度上,引入捷径分支和原始输入特征进行相加以增强训练的稳定性。MASC 通过多尺度的跳跃连接结构和自适应卷积模块丰富了感受野的动态变化,在单个特征图内部能够在更精细的水平上捕获细节和全局特征。

### 3 实验结果与分析

#### 3.1 数据集与实验平台

实验中模型训练和测试所用的钢材表面缺陷检测图像数据集是 NEU-DET<sup>[23]</sup>,由6类缺陷组成,包括开裂、夹杂物、斑块、点蚀表面、轧制氧化皮和划

痕,不同类型缺陷的具体可视化图像在图1中展示。其中,每一个缺陷类型有300张图像,共计1800张图像,以7:3的比例将图像分为训练集和测试集,分别产生1260和540幅图像。

本文模型训练和推理使用的是 Ubuntu 18.04 LTS 系统, Intel (R) Xeon (R) Gold 5218 CPU 处理器, CUDA 11.4.2, cuDNN 8.2.4, 显卡为 GeForce RTX 3090, 24 G 显存大小, 网络开发框架为 Pytorch1.9.1, 集成开发环境是 Pycharm。统一设置训练轮次为300, 批大小为4, 输入图像尺寸为  $640 \times 640$ , 使用 SGD 优化器, 初始学习率设置为0.01, 使用线性衰减学习率调整策略, 最小学习率降低为初始学习率的0.01倍, 权重衰减设置为0.0005。

#### 3.2 评价指标

本文选择精确率、召回率、 $mAP_{50}$ 、 $mAP_{50:95}$ 、浮点操作数 (Floating point OPERations, FLOPs) 和 FPS 作为检测效果的比较标准, 其中 FLOPs 用来衡量模型的计算复杂度, 由此即可确定模型的优点和缺点。以  $IoU$  阈值0.5为标准, 精确率和召回率的计算公式如下:

$$\text{精确率} = TP / (TP + FP) \quad (2)$$

$$\text{召回率} = TP / (TP + FN) \quad (3)$$

其中,  $TP$  表示正样本被正确识别为正样本的个数;  $FP$  表示负样本被错误识别为正样本的个数;  $FN$  表示正样本被错误识别为负样本的个数。  $mAP_{50}$  是由精度和召回率形成的精度-召回 (PR) 曲线下的面积。  $mAP_{50:95}$  是以  $IoU$  阈值从0.50到0.95十等分划分为10个  $IoU$  阈值分别求取精度-召回 (PR) 曲线下的面积加和取平均值。 FPS 表示模型在1s内检测的图片数量。

#### 3.3 消融实验

为了证明 TIEM 和 MASC 在 YOLOv7 目标检测器上的有效性, 本文在 NEU-DET 数据集上进行了消融实验, 表1展示了不同模块的检测结果对比。表1中, G 表示数量级,  $1G = 10^9$ 。

表1 不同模块的检测结果

Table 1 Detection results of different models

TIEM	MASC	精确率/%	召回率/%	$mAP_{50}/\%$	$mAP_{50:95}/\%$	FLOPs/G	FPS
		80.2	78.2	79.3	42.1	104.8	71
✓		81.9	79.5	81.5	43.2	110.3	66
	✓	81.3	79.2	80.8	42.6	109.9	67
✓	✓	82.8	80.2	82.3	44.2	115.4	65

从表1中可以看出, 不论加入 TIEM、还是 MASC 后, 精确率、召回率、 $mAP_{50}$  和  $mAP_{50:95}$  均有提

高, 说明 TIEM 能克服低对比度下纹理特征不明显的问题, MASC 能提高对缺陷尺度变化剧烈的检测

鲁棒性,优化对缺陷的定位准确率。在 TIEM 和 MASC 两个模块的同时作用下,模型性能进一步提升,模型的精确率、召回率、 $mAP_{50}$  和  $mAP_{50:95}$  相较于原模型分别提高了 2.6%、2.0%、3.0% 和 2.1%,充分说明 TIEM 和 MASC 对 YOLOv7 目标检测器的贡献效果,能有效提升对钢材表面缺陷的检测精度。

### 3.4 检测结果可视化

图 6 展示了本文模型改进前后在 NEU-DET 数据集上部分检测结果可视化,不同颜色的矩形框表示不同类的缺陷检测框,共有 6 种不同的缺陷类型。其中,检测框置信度大都保持在 0.8 以上,检测器对每一个检测框均能保持较高的检测置信度。

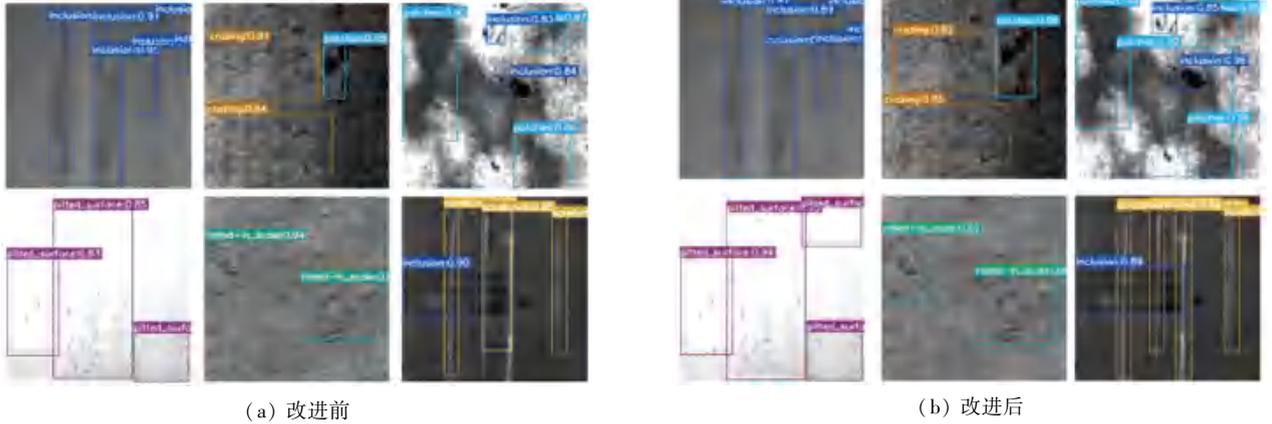


图 6 改进前后检测效果可视化对比

Fig. 6 Visual comparison of detection effects before and after improvement

从图 6 不同类的检测效果来看,面对背景和缺陷相似以及缺陷纹理信息不明显的情况,改进前的原始模型对部分缺陷出现漏检的情况,比如第 1 行第 3 列斑块缺陷和第 2 行第 1 列图像点蚀表面缺陷,而改进后的模型能检测出纹理信息不明显的缺陷类型,例如轧制氧化皮(rolled-in scale)、点蚀表面(pitted surface)和划痕(scratches),这些方面均和图像背景很相似,缺陷分布不规则,具有零散分布特征,并且不同缺陷存在重叠区域。本文模型在 TIEM 的性能提升作用下,能准确检测不同种类的缺陷,实现较好的定位效果。

斑块缺陷和第 2 行第 2 列的轧制氧化皮缺陷也出现定位框的偏移,以及第 2 行第 3 列的划痕缺陷出现了检测框合并的情况,并且未能精确定位,而改进后的模型能检测出不同高宽比的表面缺陷类型,例如夹杂物和开裂类型,均具有很大的高宽比,并且一些小尺寸的夹杂物缺陷和大目标斑块(patches)也能够准确检测出来,说明本文模型能适应目标尺度变化剧烈的不同类缺陷目标检测。

从图 6 的不同类的检测框尺寸来看,面对目标缺陷尺寸变化剧烈的情况,改进前的原始模型对部分缺陷出现定位不准的情况,比如第 1 行第 1 列第 1 个检测框,没有将缺陷全部框出,第 1 行第 2 列的

### 3.5 与现阶段其他模型对比

为突显本文算法在 NEU-DET 数据集上的检测准确性和快速性,表 2 显示了本文算法与其它主流目标检测算法的性能比较。其中, FPS 的计算,包括了数据前处理、模型前向推理和后处理时间,并保持图像长边为 640,短边不失真自适应缩放,并保持短边像素个数为 32 倍数的灰边填充。

表 2 与其他模型对比结果

Table 2 Detection results of different models

模型	精确率/%	召回率/%	$AP_{50}$	$AP_{50:95}$ /%	FLOPs/G	FPS
YOLOv3	74.7	71.2	73.3	38.8	124.8	56
Faster R-CNN	78.2	76.5	76.6	41.2	180.3	32
YOLOv5	80.5	79.2	78.4	42.6	108.4	69
YOLOX	79.9	79.6	79.1	42.2	154.8	60
YOLOv7	80.2	78.2	79.3	42.1	<b>104.8</b>	<b>71</b>
YOLOv6	81.1	80.1	80.9	42.3	150.2	62
本文模型	<b>82.8</b>	<b>80.2</b>	<b>82.3</b>	<b>44.2</b>	115.4	65

从表2中可以看出,相比较两阶段目标检测器 Faster-RCNN,在  $mAP_{50}$  和  $mAP_{50:95}$  指标上分别提高 5.7%和 3.0%,并且 FPS 提高一倍;相较于其他单阶段目标检测器 YOLOv3、YOLOv5、YOLOX 和 YOLOv6,本文模型在精确率、召回率、 $mAP_{50}$  和  $mAP_{50:95}$  指标上均取得了最好的性能,本文模型相对于 YOLOv6 和 YOLOX 模型  $FLOPs$  降低 33.9%以上, FPS 提高了 3 帧以上,相对于 YOLOv5 FPS 低了 4 帧,但是  $mAP_{50}$  和  $mAP_{50:95}$  指标上高出 2.3%和 1.0%,证实了本文模型优于现阶段的其他主流目标检测器,在 NEU-DET 数据集上能表现出最高的检测精度,同时保持较快的检测速度。

#### 4 结束语

本文以 YOLOv7 为基线模型,在 NEU-DET 数据集上开展实验研究,提出纹理信息增强模块和跳跃连接的多尺度自适应卷积模块来解决钢材表面缺陷图像低对比度、缺陷纹理信息不明显以及缺陷尺度变化剧烈等问题,消融实验和检测可视化结果证明了以上方法的可行性,并和现阶段其他主流目标检测器进行对比,本文模型的检测精度优于其他主流目标检测器,并实现了较快的检测速度。

在未来工作中,将探讨研究模型量化、压缩等技术来实现模型容量的进一步减少,减小模型的运行成本和内存消耗,进一步降低硬件成本,为视觉数字化工业缺陷检测提供新思路。

#### 参考文献

[1] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings-2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA:IEEE, 2005: 886-893.

[2] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.

[3] MAKAREMI M, RAZMJOOY N, RAMEZANI M. A new method for detecting texture defects based on modified local binary pattern[J]. Signal, Image and Video Processing, 2018, 12:1395-1401.

[4] CORTES C, VAPNIK V. Support-Vector Networks[J]. Machine Learning, 1995, 20:273-297.

[5] SONG Kechen, YAN Yunhui. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects[J]. Applied Surface Science, 2013, 285(B):858-864.

[6] LIU Yang, XU Ke, WANG Dadong. Online surface defect identification of cold rolled strips based on local binary pattern and extreme learning machine[J]. Metals, 2018, 8(3):197.

[7] DING Runwei, DAI Linhui, LI Guangpeng, et al. TDD-net; a tiny defect detection network for printed circuit boards[J]. CAAI Transactions on Intelligence Technology, 2019, 4:110-116.

[8] 程婧怡,段先华,朱伟. 改进 YOLOv3 的金属表面缺陷检测研究[J]. 计算机工程与应用, 2021, 57(19):252-258.

[9] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[10] 吴桐,杨金成,廖瑞颖,等. 基于线阵图像深度学习的电池组焊缝瑕疵检测[J]. 激光与光电子学进展, 2020, 57(22):307-314.

[11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, USA:IEEE, 2017: 6517-6525.

[12] 李瑞生,张彦龙,翟登辉,等. 基于改进 SSD 的输电线路销钉缺陷检测[J]. 高电压技术, 2021, 47(11):3795-3802.

[13] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. arXiv preprint arXiv:1512.02325, 2015.

[14] 吴越,杨延竹,苏雪龙,等. 基于 Faster R-CNN 的钢板表面缺陷检测方法[J]. 东华大学学报(自然科学版), 2021, 47(3):84-89.

[15] REN Shaoqing, HE Kaiming, 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, 2017, 39(6):1137-1149.

[16] 李庆党,李铁林. 基于改进 YOLOv3 算法的钢板缺陷检测[J]. 电子测量技术, 2021, 44(2):104-108.

[17] 李维刚,叶欣,赵云涛,等. 基于改进 YOLOv3 算法的带钢表面缺陷检测[J]. 电子学报, 2020, 48(7):1284-1292.

[18] WANG C, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE, 2023: 7464-7475.

[19] BOCHKOVSKIY A, WANG C, LIAO H M. YOLOv4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[20] GE Zheng, LIU Songtao, WANG Feng, et al. Yolox: Exceeding YOLO series in 2021[J]. arXiv preprint arXiv:2107.08430, 2021.

[21] LI Chuyi, LI Lulu, JIANG Hongliang, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. arXiv preprint arXiv:2209.02976, 2022.

[22] GAO Shanghua, CHENG Mingming, ZHAO Kai, et al. Res2net: A new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2):652-662.

[23] SONG Kechen, HU Shaopeng, YAN Yunhui. Automatic recognition of surface defects on hot-rolled steel strip using scattering convolution network[J]. Journal of Computational Information Systems, 2014, 10(7):3049-3055.