DOI: 10.13382/j.jemi.B2205354

改进 YOLOv5 网络的带钢表面缺陷检测方法*

马燕婷[1] 赵红东[1,2] 阎 超[1] 封慧杰[1] 于快快[3] 刘 赫[4]

(1. 河北工业大学电子信息工程学院 天津 300401;2. 光电信息控制和安全技术重点实验室 天津 300308;
3. 中国电子科技集团公司第五十三研究所 天津 300308;4. 天津金沃能源科技有限公司 天津 300382)

摘 要:带钢表面缺陷检测已成为保证带钢生产质量的重要环节之一。针对当前带钢缺陷检测算法精度有待提高等问题,提出了一种基于 YOLOv5 网络改进的算法模型 MT-YOLOv5。首先在主干网络中引入 Transformer 自注意力机制,使主干网络更聚焦于图像全局特征信息的提取;其次采用 T-BiFPN 网络结构,将 Transformer 层与 BiFPN 网络结构相结合,进一步增强了图像浅层特征信息与深层特征信息的融合;然后引入改进后的轻量化网络 RepVGG 替换主干网络中的部分卷积层,增强主干网络的特征提取能力;最后增加预测层,检测不同尺度的目标。实验结果表明,MT-YOLOv5 算法在 NEU-DET 数据集上的均值平均精度(mAP)达到了 82.4%,较原 YOLOv5s 算法提高了 5.3%,检测速度为 65.4 fps,更好地均衡了检测速度与检测精度。

关键词: 缺陷检测;多尺度特征融合;Transformer;RepVGG

中图分类号: TN98; TP391.4 **文献标识码**: A **国家标准学科分类代码**: 520.2060

Strip steel surface defect detection method by improved YOLOv5 network

Ma Yanting[1] Zhao Hongdong[1,2] Yan Chao[1] Feng Huijie[1] Yu Kuaikuai[3] Liu He[4]

(1. School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China;
2. Key Laboratory of Optoelectronic Information Control and Security Technology, Tianjin 300308, China;
3. The 53rd Research Institute of China Electronics Technology Group Corporation, Tianjin 300308, China;
4. Tianjin Jinwo Energy Technology Company Limited, Tianjin 300382, China)

Abstract: Strip steel surface defect detection has become one of the important links to guarantee the quality of strip steel production. Aiming at the problem of improving the detection accuracy of current strip steel defect detection algorithm, an improved MT-YOLOv5 algorithm based on YOLOv5 is proposed. Firstly, introducing Transformer self-attention mechanism in the backbone network to make the network more focused on the extraction of global image feature information. Secondly, combining the Transformer layer with the BiFPN structure, and the T-BiFPN network is used to further enhance the fusion of image shallow feature information and deep feature information. Then, an improved lightweight network RepVGG is introduced to replace part of the convolutional layers in the backbone network, which can enhance the feature extraction capability of backbone network. Finally, adding a prediction layer to detect objects of different scales. The experimental results show that the value of mean average precision (mAP) of the MT-YOLOv5 algorithm is 82.4% on the NEU-DET dataset, which is 5.3% higher than the original YOLOv5s algorithm, and the detection speed reaches 65.4 fps, which achieves a better balance between detection speed and detection accuracy.

Keywords: defect detection; multi-scale feature fusion; Transformer; RepVGG

收稿日期: 2022-04-06 Received Date: 2022-04-06

*基金项目:天津市科技计划项目(企业科技特派员项目)(21YDTPJC00050)、光电信息控制和安全技术重点实验室基金项目(2021JCJQLB055008)资助

0　引　言

带钢作为重要的工业材料之一,在加工过程中难免会产生裂纹、斑块、划痕等缺陷,这些缺陷不仅会使产品的美观受到影响,还对带钢制品的抗腐蚀性和耐磨性能等性能带来影响,导致企业利润下降。因此,实现高精度的带钢表面缺陷检测算法对于提升带钢产品表面质量具有重要意义[1-2]。

近些年国内外学者在计算机视觉技术上针对瑕疵检测进行了大量研究[3-4],其研究方法主要分为基于机器学习的方法和基于深度学习的方法。基于机器学习的方法主要是基于人工设计提取特征[5]。Celik 等[6]提出一种基于小波变换、灰度共生矩阵(GLCM)与前馈神经网络相结合的简易检测系统,但是该方法不能定位缺陷的具体位置;屈尔庆等[7]针对传统 Gabor 小波滤波器参数多和算法实时性差的问题,利用差分算法,提出了基于复合差分进化的 Gabor 滤波器优化算法,优化后的带钢模型检测速度为 11 fps,仍低于实时检测的要求;刘源等[8]结合双边滤波和 Hilditch 细化算法,提出了一种基于 Sobel 算子的带钢表面缺陷图像边缘检测算法。但是该算法受光环境影响较大,对噪声比较敏感。基于机器学习的检测方法简单且易于实现,但是受人为提取特征影响较大,适应性和鲁棒性较差。

近年来,随着深度学习技术的快速发展,基于深度学习的目标检测算法逐步替代了传统的机器学习方法。基于深度学习的方法主要有一阶段的 YOLO[9]、SSD[10]和二阶段的 Faster-RCNN[11]等算法[12]。He 等[13]基于 Faster-RCNN 提出了一个带钢缺陷检测系统,在 NEU-DET 数据集上缺陷检测任务的 mAP 值为 82.3%,但该方法在检测速度上稍有不足,难以满足工业生产实时检测的需求;李维刚等[14]利用加权 K-means 聚类算法提出了一种改进的 YOLOv3 算法模型,改进后的算法 mAP 达到 80%,检测速度保持在 50 fps,但是该模型参数量较大,不易部署;Cheng 等[15]在 RetinaNet 网络中结合通道注意力机制以减少信息丢失,利用自适应空间特征融合(ASFF)模块对卷积核提取的浅层和深层特征进行有效融合,在 NEU-DET 数据集上 mAP 为 78.25%,但是该模型精度较低。目前应用于带钢表面缺陷检测的算法研究有很多,但是满足实时检测要求的算法,仍然存在精度较低的问题。

本文针对目前带钢表面缺陷检测算法特征提取能力不足,模型检测精度低的问题,提出了基于 YOLOv5 的 Transformer 层多尺度特征融合算法 MT-YOLOv5(multi-scale fusion of Transformer layer YOLOv5),首先将 Transformer 与 YOLOv5 模型相结合;其次引入 T-BiFPN 网络替代特征金字塔网络,增强图像浅层特征信息和深层特征信息融合;接着引入轻量级卷积神经网络 RepVGG,提高网络特征提取能力,并保证速度和精度兼得;最后增加检测层,提升对不同尺度目标的检测性能。本文提出的 MT-YOLOv5 算法在 NEU-DET 数据集上进行测试,实验表明,该方法在带钢表面缺陷检测任务中具有良好的检测性能,可以进一步满足工业部署需求。

1　MT-YOLOv5 模型

图 1 为改进后的 MT-YOLOv5 模型结构图。其中 YOLOv5 是一种基于 YOLOv3、YOLOv4 改进的目标检测算法,是 YOLO 系列的最新版本,且处于不断更新中。YOLOv5 模型结构主要为 5 个版本,分别是 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。它们之间的主要区别是在特定的网络位置上特征提取模块和卷积核的数量不同。5 个版本的模型大小和模型参数的数量依次增加。由于缺陷检测对实时性和轻量化性能要求较高,因此本研究综合考虑了识别模型的准确性、效率和大小,基于 YOLOv5s 模型对带钢表面缺陷检测网络进行了改进设计。改进后的 MT-YOLOv5 模型结构主要由 Backbone、Neck、Head 3 个部分组成。

Backbone,包含 Conv、RepVGG、C3、Transformer 和 SPPF(spatial pyramid pooling-fast)结构,用于提取图像特征。Conv 对图像提取特征,RepVGG 提高网络特征提取能力,之后用 C3 降低计算量和内存,Transformer 聚焦于提取图像全局特征信息,SPPF 网络用于整合特征。

Neck,采用了 T-BiFPN 结构,将经过 Transformer 处理的图像特征信息融入 BiFPN 网络,加强融合了图像深浅层全局和局部的信息。

Head,采用 GIOU Loss 做 Bounding box 的损失函数,用非极大值抑制方式筛选生成目标框。图 1 中 m 为输出特征图的通道数。

1.1　Transformer 模块

Transformer 模型最先应用于自然语言处理(NLP)中,主要由 Encoder 和 Decoder 组成。ViT[16]模型首次将 Transformer encoder 模块应用于计算机视觉中。将 Transformer 与 YOLOv5 相结合,在主干网络中引入 Transformer,用 Transformer encoder 模块替换 YOLOv5 原始模型中的一些 Conv 卷积块和 C3 模块。与 YOLOv5 中的 C3 模块相比,Transformer encoder 模块可以提取图像全局信息和丰富的上下文信息。Transformer encoder 模块由多头注意力机制(multi-head attention)和前馈神经网络(MLP)构成。

具体结构如图 2 所示。首先将图片按给定大小切分为 Patches,结合位置编码得到一维向量,形成 Encoder 模

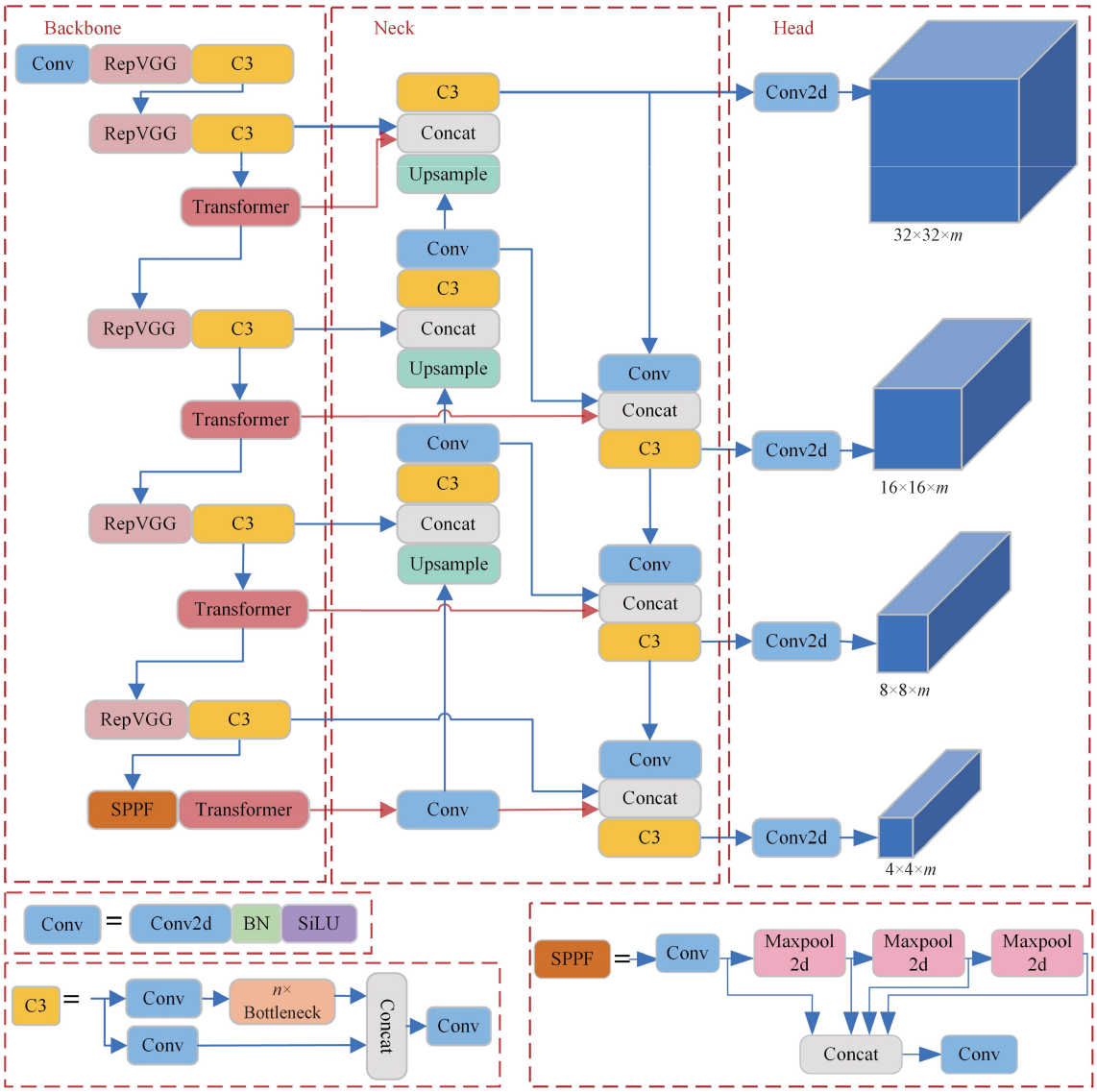


图 1　MT-YOLOv5 模型结构

Fig. 1　MT-YOLOv5 model structure

块的输入。多头注意力机制是 Encoder 模块的重要组成部分,可以平行地计算输入的多组数据。采用自注意力机制,通过点积注意力,计算自注意力特征输出,如式(1)所示。

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{1}$$

式中:$\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}$ 均为输入特征,分别意为查询向量、键向量、值向量,d_k 为输入特征的维度。$\boldsymbol{Q}$ 与 $\boldsymbol{K}$ 的转置点乘后得到向量的相关性矩阵,为了避免激活函数 softmax 造成梯度消失,采用归一化,即除以 $\sqrt{d_k}$。再与 $\boldsymbol{V}$ 矩阵相乘,得到加权后的输出结果。因此多头注意力机制模块不仅关注当前像素点,还可以融合上下文其他像素点的特征。LayerNorm 层和 Dropout 层可以帮助网络更好地收敛,防止网络过拟合。

Transformer encoder 模块不仅增加了提取图像特征信息的能力,还可以利用自注意力机制来挖掘潜在特征信息。在 MT-YOLOv5 模型的主干网络中,多次应用 Transformer encoder 模块,有利于模型对图像深浅特征的全局提取,从而提高检测率。

1.2　RepVGG 模块

2021 年,Ding 等[17]提出了 RepVGG 网络结构,在 VGG 网络中引入 identity 残差分支,通过结构重参数化实现模型训练阶段和推理阶段解耦。本文对 RepVGG 网络进行改进,用激活函数 SiLU 替换原来的激活函数 ReLU,更有利于模型的加深。

RepVGG 网络训练模型和推理模型结构如图 3 所示。其中 RepVGG 网络在训练阶段如图 3(a)所示,采用 3×3 卷积、1×1 卷积分支、Identity 残差结构,形成多分支

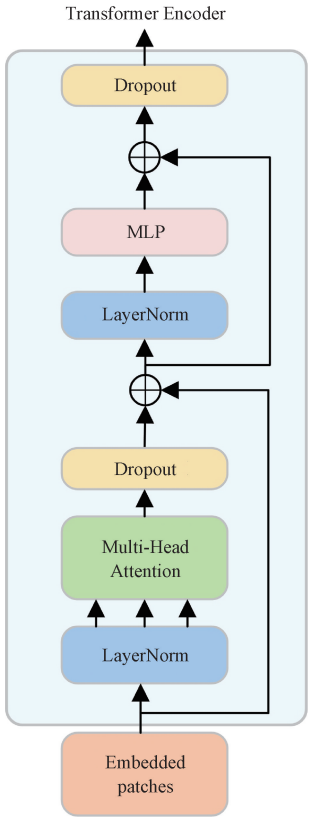


图 2　Transformer encoder 模块结构

Fig. 2　The architecture of transformer encoder

网络结构。引入残差结构后，可以有效地解决梯度消失的问题，使得网络模型以更快的速度收敛。在推理阶段如图 3(b)所示，采用 3×3 卷积和激活函数 SiLU 堆叠而成的类 VGG 结构，该结构比较简单，极大地提高了网络的推理速度。训练模型转化为推理模型采用重参数化的过程，首先将 Conv 层和 BN 层进行融合，将融合后的卷积层都转换成大小为 3×3 的卷积，最后合并所有分支中的 3×3 卷积，即将所有分支的权重 W 和偏置 b 叠加，得到一个新的 3×3 卷积用于模型推理，从而完成推理模型的重参数化过程。

为了满足带钢检测模型精度高、速度快的需求，本文引入了改进后的 RepVGG 模块作为主干网络，提高主干网络的特征提取能力，减少特征信息丢失，在提高网络特征提取能力的同时兼顾速度，有利于模型的部署。

1.3　多尺度特征融合网络

伴随着网络层数的加深，特征信息会造成一定的丢失，因此，多尺度特征融合被广泛应用于目标检测网络中，以提高网络模型的检测性能。目前常用的特征融合网络有 FPN[18]、PANet[19]、BiFPN[20]。FPN 是一种自顶向下的特征金字塔结构，结合深浅层多尺度特征，但受制于单向特征信息传递。PANet 在 FPN 的基础上增加了一

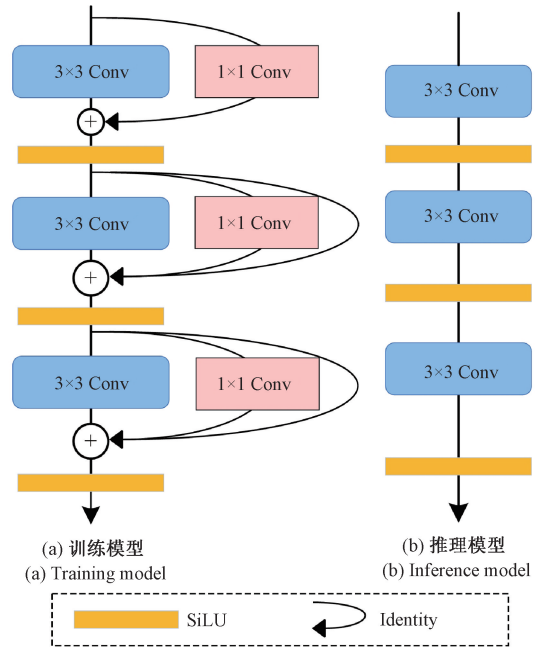


图 3　RepVGG 训练模型和推理模型结构

Fig. 3　RepVGG training model and inference model structure

个自底向上的二次融合路径。BiFPN 是 Google 团队提出的一种加权的双向特征金字塔网络，实现了简单快速的多尺度特征融合。

本文为了减少特征信息的丢失，提高目标的检测精度，提出了 T-BiFPN 网络结构，将 Transformer 和 BiFPN 结构融合，具体结构如图 4 所示，图中 Tr 为 Transformer 模块。首先从主干网络 Transformer 层提取出不同尺度的特征图，接着通过横向连接与下采样层实现第 1 次特征融合，再通过跳跃连接实现与同一尺度的下采样层和上采样层的第 2 次特征融合，最终得到多尺度融合后的特征图。

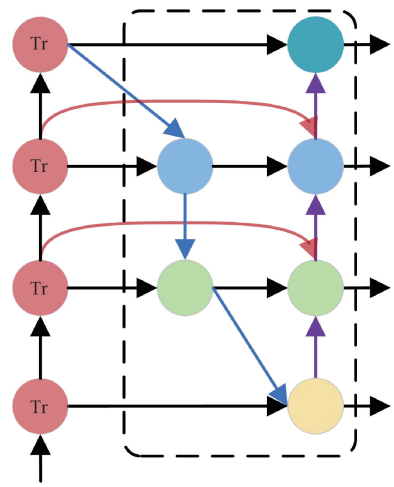


图 4　T-BiFPN 结构图

Fig. 4　The architecture of T-BiFPN

T-BiFPN 网络加强了图像浅层特征信息和深层特征信息的融合,并且每个新加的特征信息层都是经过 Transformer 处理后的,因此更能起到融合深浅层全局和局部信息的核心。总之,T-BiFPN 网络结构增强特征金字塔的表达能力,使网络达到最优的特征融合模式。

本研究将 YOLOv5 中加强特征提取的 PANet 结构改进为 T-BiFPN 结构,将主干网络中 Transformer 处理过的特征图多次输入 BiFPN 结构中,实现双向的多尺度特征融合,提高了模型对整体特征的学习能力,降低了漏检率。

1.4 新增检测层

原始 YOLOv5s 模型有 3 个检测层,当输入图片大小为 256×256 时,检测层特征图的输出尺寸分别为 32×32、16×16、8×8,分别用于小、中、大目标的检测。然而,由于带钢表面瑕疵种类多,缺陷大小不一,为了提高缺陷检测准确率,本文基于 YOLOv5s 网络进行了改进,结合数据集中缺陷目标的大小,增加了 64 倍下采样输出层,对应检测层特征图的输出尺寸为 4×4,满足准确识别各类尺度缺陷目标的需求。最终检测层输出尺度为 32×32、16×16、8×8、4×4。训练时基于训练数据采用自适应锚框计算锚框大小,输出层每层所对应的先验框数量为 3 个,总的先验框数量增加至 12 个。新增检测层对于大目标特征较为敏感,多尺度检测层有利于提高网络的泛化能力,提高模型检测性能,减少漏检、误检的情况。

2 实验结果与分析

2.1 实验数据集

NEU-DET 是由东北大学制作的带钢表面缺陷数据集,包括 6 种类型的常见表面缺陷:轧制氧化皮(rolled-in scale, RS)、斑块(patches, Pa)、裂纹(crazing, Cr)、夹杂(inclusion, In)、表面点蚀(pitted surface, PS)、以及划痕(scratches, Sc)。该数据集共 1 800 张灰度图像,每类缺陷 300 张,分辨率为 200×200,随机抽取训练集、验证集、测试集的比例为 8∶1∶1,即训练集 1 440 张,验证集 180 张,测试集 180 张。考虑到数据集图片的分辨率以及网络模型的特点,将输入图片大小设置为 256×256 的统一尺寸,并采用 Mosaic 数据增强方法。

2.2 实验配置与评测标准

本文实验采用的硬件配置为 Intel Core i7-9700K@3.60 GHz 处理器,显卡为 NVIDIA GeForce RTX 2080Ti 11 GB。软件环境为 CUDA10.2 和 cuDNN8.2.1,操作系统为 Ubuntu18.04。网络模型基于 Pytorch 框架搭建,Python 版本为 3.8,Pytorch 版本为 1.10.1,实验中 batch size 为 64,epoch 设置为 300,动量为 0.937,IoU 阈值设置为 0.5。在训练过程中,初始学习率为 0.01,并采用余弦退火策略降低学习率,衰减系数为 0.000 5。

本文采用 2 个评价指标对模型进行评估,分别是平均精度(AP)、均值平均精度(mAP)。同时需要参考每秒检测帧数(fps)、精确率(precision, P)和召回率(recall, R),表达式分别如式(2)、(3)所示。以 P 为纵轴,R 为横轴构成 P-R 曲线。AP 为 P-R 曲线下的面积,如式(4)所示。P-R 曲线下的面积越大,AP 值越高,说明检测模型越准确。mAP 是所有类别 AP 的平均值,如式(5)所示,n 为类别个数。fps 为每秒检测图片的数量,用来评估模型的检测速度。

$$P = \frac{TP}{TP + FP} \tag{2}$$

$$R = \frac{TP}{TP + FN} \tag{3}$$

式中:TP 为真正例,FP 为假正例,FN 为假反例。

$$AP = \int_0^1 p(R)\,\mathrm{d}R \tag{4}$$

$$mAP = \frac{\sum_{i=0}^{n} AP(i)}{n} \tag{5}$$

2.3 实验结果分析

为了更加直观地展示所提方法在带钢表面瑕疵检测和识别方面的性能,进行了消融实验,实验结果如表 1 所示。

表 1 消融实验结果

Table 1 Results of ablation test

方法	第 1 组	第 2 组	第 3 组	第 4 组	第 5 组
YOLOv5s	√	√	√	√	√
检测层		√	√	√	√
Transformer			√	√	√
T-BiFPN				√	√
RepVGG					√
mAP/%	77.1	79.2	80.2	81.1	82.4
参数量/10^6	7.03	12.33	14.55	15.04	15.37

实验以 YOLOv5s 为基线,表 1 中第 1 组和第 2 组实验对比,检测层由 3 层增加至 4 层,mAP 从 77.1%提高至 79.2%。结果表明增加检测层提高了网络对不同尺度的缺陷目标的检测准确性,降低了漏检率。第 2 组和第 3 组实验对比显示,增加 Transformer 模块后,mAP 从 79.2%提高至 80.2%,提高了 1%。结果表明增加 Transformer 层提高了深度神经网络对图像全局特征信息的提取能力,提高检测精度。第 3 组和第 4 组实验对比显示,采用 T-BiFPN 网络融合特征使 mAP 提高至 81.1%。表明使用 T-BiFPN 网络进行多尺度特征融合可以有效提高网络模型的检测精度。第 4 组和第 5 组结果显示,引用 RepVGG 主干网络使 mAP 提高至 82.4%。表

明引入轻量化网络 RepVGG 作为主干网络提高了模型提取特征的能力。

本文所提出的算法在 NEU-DET 数据集上 mAP 值为 82.4%,检测速度为 65.4 fps。训练过程中的参数量虽然高于原来的 YOLOv5s,但是检测速度仍然满足实时检测需求,且易于部署在移动平台端,可用于钢材表面瑕疵的实时拍摄和识别。总体而言,MT-YOLOv5 对缺陷目标检测具有较高的准确率,并且可以在精度和速度之间取得较好的平衡,具有实际的应用意义。

图 5 为 YOLOv5s 与本文算法在各类缺陷上的检测结果对比。从图 5 中可知,增加检测层、主干网络中引入 Transformer 和 RepVGG 模块、采用多尺度特征融合网络能够有效提高带钢缺陷的检测精度。改进后的算法在多类缺陷检测上均有所提升。如 Cr 类改进后的算法比 YOLOv5s 提高了 13.8%,Ps 类改进后提高了 13%。因此,本文提出的 MT-YOLOv5 缺陷检测模型在精度上取得了优异的效果,符合对缺陷检测精度更高的要求。

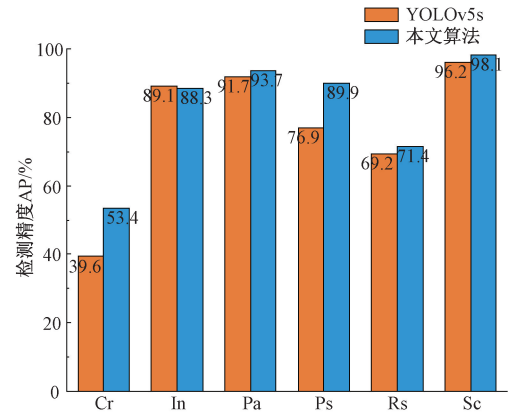
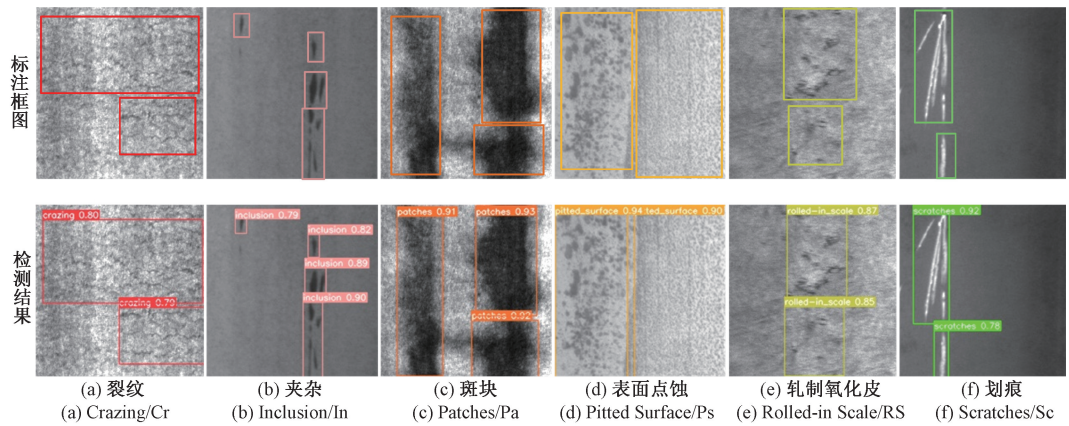


图 5　改进算法检测结果对比

Fig. 5　Comparison of improved algorithms detection results

图 6 为改进后的模型检测结果图,通过与标注框的对比,可以看到 MT-YOLOv5 模型对各类缺陷均有良好的检测效果。



(a) 裂纹 (a) Crazing/Cr　(b) 夹杂 (b) Inclusion/In　(c) 斑块 (c) Patches/Pa　(d) 表面点蚀 (d) Pitted Surface/Ps　(e) 轧制氧化皮 (e) Rolled-in Scale/RS　(f) 划痕 (f) Scratches/Sc

图 6　MT-YOLOv5 模型检测结果

Fig. 6　Detection results of MT-YOLOv5 model

2.4　对比实验

为了证明 MT-YOLOv5 模型在带钢表面缺陷检测方面的优势,将其与几种目前主流的目标检测算法,如 Faster-RCNN、SSD300、YOLOv4、和原始的 YOLOv5 在 NEU-DET 数据集上进行了比较。具体结果如表 2 所示。

表 2　不同算法在 NEU-DET 数据集上检测性能对比

Table 2　Comparison of the different models on NEU-DET

算法	mAP/%	模型大小/M	AP/%					
			Cr	In	Pa	Ps	Rs	Sc
Faster-RCNN	77.2	108	37.8	86.1	92.3	88.5	63.5	95.1
SSD300	72.3	93.1	36.3	81.9	91.3	83.9	62.1	78.2
YOLOv4	69.1	244	35.1	77.6	90.2	78.4	51.2	82.2
YOLOv5s	77.1	14.1	39.6	89.1	91.7	76.9	69.2	96.2
MT-YOLOv5	82.4	29.7	53.4	88.3	93.7	89.9	71.4	98.1

从表 2 可知,本文所提出的改进算法 mAP 值达到了 82.4%,比 YOLOv5s 提高了 5.3%。相比 SSD、YOLO 系列经典算法,Faster-RCNN 在带钢缺陷检测中精度较高,mAP 值达到了 77.2%,这是因为 Faster-RCNN 是二阶段

检测网络,先用特征提取器生成预选框,再进行目标检测,计算量较大,速度较慢。

SSD 在 Ps 类缺陷中表现较好,在 Sc 类缺测中检测性能较差,这是因为 Ps 类缺陷多为大目标,Sc 类缺陷多为小目标,一阶段算法 SSD 在速度上有优势,在小目标检测上略有不足,造成 Sc 类缺陷漏检、误检,检测精度较差。YOLOv4 在 NEU-DET 测试集上 mAP 值为 69.1%,精度较低,且模型权重较大,不易部署。总之,本文提出的方法 mAP 从 76.6%提高到了 82.4%,获得了最佳的 mAP 精度,这是由于改进后的主干网络加强了特征信息提取,多尺度特征融合网络强化了特征融合,增加检测层,降低了漏检率。

综上所述,本文提出的 MT-YOLOv5 缺陷检测算法与经典算法 Faster-RCNN、SSD 相比性能更好,模型权重也较小,在检测精度和速度方面均有较好的表现,更有利于开展工业部署。

3　结　论

为了解决钢铁表面瑕疵检测问题,进一步提高钢铁表面瑕疵检测准确率和识别速度,本文提出了一种端到端的 MT-YOLOv5 模型。MT-YOLOv5 模型采用轻量化的 RepVGG 网络作为主干网络,引入 Transformer 模块加强对全局特征信息的提取,采用 T-BiFPN 网络融合不同尺度的特征信息,增加检测层,降低漏检率。并将提出的模型与经典目标检测算法模型进行对比,实验结果显示,MT-YOLOv5 网络模型在测试集中 mAP 达到了 82.4%,检测速度为 65.4 fps。该模型在检测精度和检测效率上取得了较好的平衡,对钢带表面缺陷识别检测具有重要的现实意义。今后将进一步尝试改进 Anchor-Free 模型,进一步提高模型的准确性。

参考文献

[1] 李维创, 尹柏强. 工业金属板带材表面缺陷自动视觉检测研究进展[J]. 电子测量与仪器学报, 2021, 35(6):1-16.
LI W CH, YIN B Q. Research progress of automated visual surface defect detection for industrial metal planar materials[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(6): 1-16.

[2] 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 168-182.
LAN J H, WANG D, SHEN X P. Research progress on visual image detection based on convolutional neural network[J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 168-182.

[3] 王杰, 陈友兴, 李海洋, 等. 基于激光超声的金属表面裂纹定位研究[J]. 国外电子测量技术, 2019, 38(11): 1-4.
WANG J, CHEN Y X, LI H Y, et al. Research on metal surface crack location based on laser ultrasound[J]. Foreign Electronic Measurement Technology, 2019, 38(11): 1-4.

[4] 周志锋, 万旺根, 王旭智. 基于 YOLO v3 框架改进的目标检测[J]. 电子测量技术, 2020, 43(18): 102-106.
ZHOU ZH F, WAN W G, WANG X ZH. Object detection based on enhanced YOLO v3[J]. Electronic Measurement Technology, 2020, 43(18): 102-106.

[5] ABBAS Q, LI Y. Cricket video events recognition using HOG, LBP and multi-class SVM[J]. Journal of Physics: Conference Series, 2021, 1732(1): 012036.

[6] CELIK H I, DUELGER L C, TOPALBEKIROGLU M. Development of a machine vision system: Real-time fabric defect detection and classification with neural networks[J]. Journal of the Textile Institute Proceedings and Abstracts, 2014, 105(6): 575-585.

[7] 屈尔庆, 崔月姣, 徐森,等. 改进的 Gabor 滤波器带钢表面缺陷显著性检测[J]. 华中科技大学学报(自然科学版), 2017, 45(10): 12-17.
QU ER Q, CUI Y J, XU S, et al. Saliency defect detection in strip steel by improved Gabor filter[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2017, 45(10): 12-17.

[8] 刘源, 夏春蕾. 一种基于 Sobel 算子的带钢表面缺陷图像边缘检测算法[J]. 电子测量技术, 2021, 44(3): 138-143.
LIU Y, XIA CH L. An image edge detection algorithm for strip steel surface defects based on Sobel operator[J]. Electronic Measurement Technology, 2021, 44(3): 138-143.

[9] REDMON J, FARHADI A. YOLOv3: An incremental im-provement[J]. ArXiv Preprint, 2018, arXiv:1804.02767.

[10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. European Conference on Computer Vision, Springer, 2016: 21-37.

[11] 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, 2017, 39(6): 1137-1149.

[12] 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.

[13] 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, 2020, 69(4): 1493-1504.

[14] 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.
LI W G, YE X, ZHAO Y T, et al. Strip steel surface defect detection based on improved YOLOv3 algorithm[J]. Acta Electronica Sinica, 2020, 48(7): 1284-1292.

[15] CHENG X, YU J. RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.

[16] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale [J]. ArXiv Preprint, 2020, arXiv:2010.11929.

[17] DING X, ZHANG X, MA N, et al. RepVGG: Making VGG-style convNets great again[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2021: 13728-13737.

[18] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. Computer Vision and Pattern Recognition, 2017: 936-944.

[19] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[J]. ArXiv Preprint, 2018, arXiv: 1803.01534.

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

作者简介

马燕婷,2019 年于辽宁师范大学获得学士学位,现为河北工业大学硕士研究生,主要研究方向为深度学习、图像处理。
E-mail: 1793829292@qq.com

Ma Yanting received B. Sc. degree from Liaoning Normal University in 2019. Now she is a M. Sc. candidate at Hebei University of Technology. Her main research interests include deep learning and image processing.

赵红东(通信作者),1998 年于北京工业大学获得博士学位,现为河北工业大学教授,主要研究方向为光电信息处理、图像处理、半导体器件。
E-mail: zhaohd@hebut.edu.cn

Zhao Hongdong (Corresponding author) received Ph. D. degree from Beijing University of Technology in 1998. Now he is a professor in Hebei University of Technology. His main research interests include photoelectric information processing, image processing and semiconductor device.