

基于机器视觉的轻量化钢材表面缺陷检测模型

刘文钊, 张丹[†]

(南京林业大学 信息科学技术学院, 江苏 南京 210037)

摘 要:针对当前主流缺陷检测模型参数量大、计算复杂度高,难以在计算资源有限的嵌入式设备上部署的问题,提出了一种轻量化钢材表面缺陷检测模型 YOLO-LSNet。首先,为了降低模型的参数量和计算复杂度,提出了一种轻量化卷积模块 MSCConv。其次,提出 M-BiFPN 网络用于深浅层特征信息的融合。最后,用 SIoU 损失函数替换 CIoU 损失函数,加快网络的收敛速度。实验结果表明, YOLO-LSNet 模型在 NEU-DET 数据集上相较于基线网络 YOLOv5, mAP 提升了 1.8%, 模型参数下降了 43.4%, 计算量降低了 36.1%。完成模型轻量化设计的同时,保证了模型的检测精度,具有良好的应用前景。

关键词:缺陷检测;深度可分离卷积;多尺度特征融合;轻量化;注意力机制

中图分类号: TP391.4

文献标识码: A

Lightweight Steel Surface Defect Detection Model Based on Machine Vision

LIU Wenzhao, ZHANG Dan[†]

(College of Information Science and Technology, Nanjing Forestry University, Nanjing, Jiangsu 210037, China)

Abstract: Aiming at the problems of large number of parameters and high computational complexity of current mainstream defect detection models, it is difficult to deploy on embedded devices with limited computing resources. A lightweight steel surface defect detection model YOLO-LSNet is proposed. Firstly, in order to reduce the parameter number and computational complexity of the model, a lightweight convolution module MSCConv is proposed. Secondly, M-BiFPN network is proposed for deep and shallow layer feature information fusion. Finally, CIoU loss function is replaced by SIoU loss function to speed up the convergence of the network. The experimental results show that compared with the baseline network YOLOv5 in NEU-DET data set, the mAP of YOLO-LSNet model increases by 1.8%, model parameters decreases by 43.4%, and computation cost decreases by 36.1%. At the same time, the lightweight design of the model is completed, the detection accuracy of the model is guaranteed, and it has a good application prospect.

Key words: defect detection; depthwise separable convolution; multi-scale feature fusion; lightweight; attention mechanism

钢材是工业生产当中最主要的原材料之一,具有耐腐蚀性、便于加工等优点,在航天航空、机械制造、土木建工等领域皆备受青睐。在实际生产当中,由于生产工艺或生产环境等因素的影响,钢材

表面易出现裂纹、划痕等缺陷,这不仅有损产品外观,甚至严重影响工业产品的安全使用。因此,对钢材表面缺陷进行精准识别,对于后续的生产具有重要意义。

收稿日期: 2023-06-28

基金项目: 江苏省农业科技创新基金资助项目(CX(21)3187)

作者简介: 刘文钊(1998—),男,安徽合肥人,硕士研究生,研究方向:深度学习与图像处理。

[†] 通信联系人, E-mail: zhangdan@njfu.edu.cn

早期的钢材表面缺陷检测方法主要包括人工检测法、红外检测法、激光扫描法等。但上述方法的检测人员容易在持续检测状态下产生疲劳,对检测人员的操作水平存在很大的考验,对生产企业而言成本过高,效率低下,无法满足安全生产的要求^[3]。

随着机器视觉与深度学习的快速发展,基于深度学习的缺陷检测算法逐步取代了传统的缺陷检测方法。当前,基于深度学习的缺陷检测算法主要分为两类。一类是以 R-CNN 为代表的两阶段目标检测算法,包括 Fast R-CNN^[4]、Faster R-CNN^[6]等。另一类是以 YOLO 算法为代表的单阶段目标检测算法等。近年来,众多学者基于深度学习对缺陷检测模型进行了研究。Zhou 等^[10]针对缺陷尺度的多样性,设计了一种残差金字塔模块,丰富了特征图的多尺度信息。Wang 等^[11]提出了一种基于卷积神经网络的带钢表面缺陷检测框架。Shi 等^[12]通过在主干网络中引入注意力机制,有效地提高了对中小目标的关注程度。Wang^[13]等基于 YOLOv7 进行改进,在骨干网络中引入了 ConvNeXt 模块,有效地提高了模型的检测精度。上述模型针对检测过程中遇到问题提出了不同的优化方案,有效地提高了模型的检测精度。但在获取更好特征提取效果以及检测精度的同时,模型参数也开始线性增长,不利于后期轻量化部署的要求。

综上,本文为了在检测性能与计算代价之间取得平衡,提出了一个轻量级钢材表面缺陷检测网络 YOLO-LSNet。首先,为了有效地降低计算成本,

本文提出了一种轻量级卷积模块 MSConv 用以代替标准卷积。其次,提出了改进的多尺度加权特征融合网络 M-BiFPN,提升了网络特征提取能力。最后,本文对边界框损失函数进行优化,用 SIoU^[14] 损失函数替换 CIoU 损失函数,加快了模型的收敛速度,同时提升了网络的检测效果。通过在 NEU-DET 数据集上实验,YOLO-LSNet 表现出良好的检测性能,同时模型参数大幅降低,满足轻量化部署的要求。

1 轻量化钢材表面缺陷检测网络 YOLO-LSNet

本文基于 YOLOv5s 进行了改进,提出了一种轻量型钢材表面缺陷检测模型,命名为 YOLO-LSNet。YOLO-LSNet 网络结构如图 1 所示,由轻量级主干网络、加权特征融合颈部网络和多尺度检测头组成。

在图 1 中,首先用提出的轻量级 MSConv 模块替换原始网络中的标准卷积模块,以减少主干网络的参数量,利用轻量级 MBL 模块和 M-C3 模块进行不同深度特征映射的信息提取,完成轻量级骨干特征提取网络的搭建。其次,在颈部网络中,本文利用加权双向金字塔网络 M-BiFPN 代替特征金字塔网络 PANet,对不同特征进行差异化融合,提高颈部网络的特征融合效率。最后,对边界框损失函数进行优化,用 SIoU 替换 CIoU,解决预测框匹配不稳定的问题,加快网络的收敛速度。

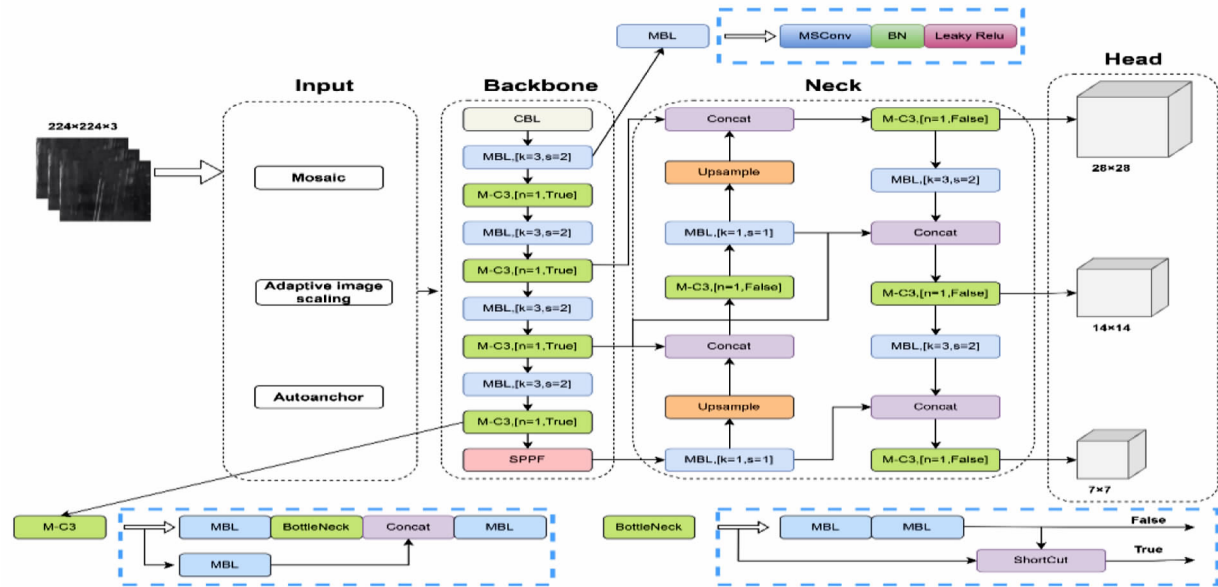


图 1 YOLO-LSNet 网络结构图

1.1 轻量级卷积模块 MSConv

目标检测模型在终端设备中的部署主要需要考虑模型的规模以及计算复杂度。为实现这一目标,越来越多的轻量级网络诞生,如 MobileNet^[15]、ShuffleNet^[16]、GhostNet^[17] 等。为减小网络模型的规模和计算复杂度,本文基于此提出了一种轻量级卷积模块 MSConv,结构如图 2 所示。其核心思想是利用深度可分离卷积来降低模型的规模和计算成本。为了更高效地在每个通道之间传递特征信息,减少特征信息的流失,首先输入一个通道数为 C_1 的特征图,通过 1×1 标准卷积进行降维后,得到通道数为 $1/2 C_2$ 的特征图 F_1 ,再经过 3×3 深度可分离卷积后得到通道数为 $1/2 C_2$ 的特征图 F_2 ,为解决特征通道信息“分离”的问题,在特征图 F_2 之后嵌入通道注意力机制 ECA^[18],接下来将两部分进行拼接,最后利用通道清洗,使特征通道信息进一步充分融合。其数学表达式如式(1)~式(3)所示。

$$F_{1_out} = f_{conv}(X_{in}) \quad (1)$$

$$F_{2_out} = f_{dsc}(f_{conv}(X_{in})) \quad (2)$$

$$X_{out} = f_{shuffle}(ECA(cat(F_{1_out}, F_{2_out}))) \quad (3)$$

$f_{shuffle}$ 表示 shuffle 操作; f_{conv} 由标准卷积、批量归一化、激活函数组成; f_{dsc} 表示深度可分离卷积操作; ECA 表示通道注意力操作。

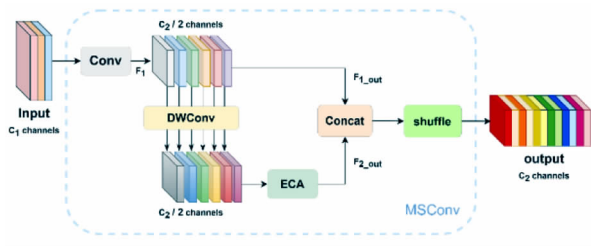


图 2 MSConv 模块结构图

MSConv 模块通过引入深度可分离卷积模块有效地降低了模型参数,同时通过通道注意力模块 ECA 加强了对重要特征通道信息的提取,提高了网络对缺陷信息的获取能力,在轻量化与检测精度之间达到了良好的平衡。

1.1.1 深度可分离卷积

DWConv^[15] 由深度卷积和点卷积两部分组成。深度卷积对特征图的每个通道进行单独的卷积滤波,可以有效地减轻网络的负担。点卷积负责通道方向上的特征映射之间的信息交换,并对通道

进行线性组合以生成新的映射。为验证 DWConv 对降低模型参数的有效性,假设卷积核大小为 $K \times K$, N 表示卷积核的数量, C_i 表示输入通道数量,输出通道数量与卷积核数量一致。图像的宽和高分别设为 W 和 H 。故标准卷积的计算量为:

$$Cost_{conv} = K \times K \times C_i \times N \times W \times H \quad (4)$$

深度可分离卷积由深度卷积与点卷积两部分组成,深度卷积对每个通道进行单独计算,点卷积需要结合加权深度提供的特征,故深度可分离卷积的计算量为:

$$Cost_{DS} = K \times K \times C_i \times W \times H \times C_i \times N \times H \times W \quad (5)$$

深度可分离卷积计算量与标准卷积计算量的比值为:

$$\frac{Cost_{DS}}{Cost_{conv}} = \frac{1}{N} + \frac{1}{K^2} \quad (6)$$

相较于标准卷积,深度可分离卷积的引入大大降低了计算量和模型参数。轻量化的设计为检测模型在移动端以及应用设备上的部署提供了巨大的帮助。

1.1.2 通道注意力模块 ECA-Net

在进行缺陷检测的过程中,除钢材表面的缺陷信息外,还存在大量相似的背景信息。这种无效的特征信息极大地影响了检测精度。针对上述问题,本文引入了通道注意力模块 ECA,结构如图 3 所示。ECA 注意力模型首先对原始输入特征图像进行全局平均池化,在此基础上,通过大小为 k 的一维卷积得到局部跨通道交互,最后通过 sigmoid 函数生成通道权值,将原始输入特征与通道权值结合,有效地帮助检测网络对缺陷特征信息的提取。

大小为 k 的自适应选择函数如等式(7)所示:

$$k = \Psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{odd} \quad (7)$$

其中, C 为输入的通道数; γ 等于 2; b 等于 1; $\lfloor t \rfloor_{odd}$ 表示最接近 t 的奇数。

相较于目前主流的注意力机制, ECA-Net 改变了以模型复杂度换检测性能提升的现状,只需要添加少量的参数就可以取得较大的性能提升。非常符合轻量化网络的设计初衷,有利于嵌入到移动工程当中。

1.2 改进的多尺度加权特征融合网络 M-BiFPN

为了区分背景信息与缺陷信息,提取更多有效的特征信息,本文提出将 M-C3 模块与 BiFPN^[19]

网络组合,构成 M-BiFPN 多尺度特征融合网络,结构如图 4 所示。首先,从骨干网络中 M-C3 模块提取不同尺寸的特征信息,通过水平连接与下采样层完成第一次特征融合,接下来通过跳跃连接与相同尺寸上采样层和下采样完成第二次特征融合,最后得到多尺度特征融合的特征图。

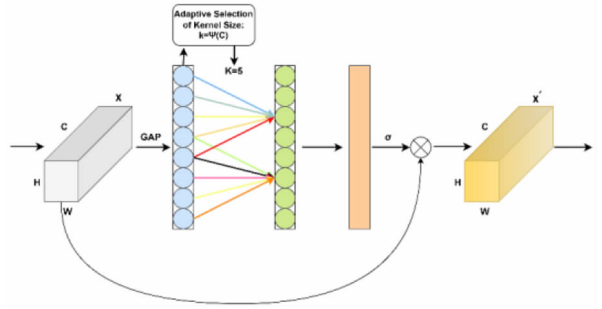


图 3 ECA-Net 模块结构图

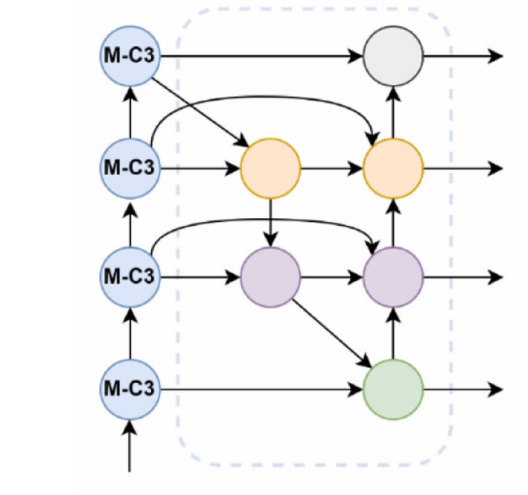


图 4 M-BiFPN 网络结构图

M-C3 模块将 C3 模块中的标准卷积模块全部替换完成轻量化重构,受益于通道注意力模块,局部特征提取网络更加聚焦于表面缺陷特征信息的提取。同时依赖于 BiFPN 网络的差异化融合,对每个输入特征添加额外的权重,区分了不同特征的重要程度。有效地增强了特征金字塔网络的信息融合与表达能力,提升了模型的检测精度。M-BiFPN 采用 Fast normalized fusion 进行加权特征融合,如等式(8)所示。

$$O=\sum_i \frac{\omega_i}{\epsilon+\sum_j \omega_j} \cdot I_i \tag{8}$$

其中, O 表示输出特征; ω_i 是一个可以学习的标量, $\omega_j \geq 0$; $\epsilon=0.0001$, I_i 表示输入特征。

1.3 损失函数优化

在钢材表面缺陷检测过程中存在缺陷密集、目标小等特点。在密集分布的区域,由于 CIoU 没有考虑到预测框与标记框之间的取向差异,会出现预测框自由度大、预测框与标记框收敛速度差的情况,从而导致错检、漏检。因此本文提出使用 SIoU 损失函数代替 CIoU 损失函数完成预测框匹配。

SIoU 损失函数由四个部分组成:角度成本、距离成本、形状成本和 IoU 成本。SIoU 的定义如式(9)所示。

$$L_{box} = 1 - IoU + \frac{\Delta + \Omega}{2} \tag{9}$$

其中, IoU 、 Δ 、 Ω 分别为交并损失、距离损失和形状损失。距离损失同时考虑到了两个框之间的角度损失。

2 实 验

为了验证 YOLO-LSNet 模型的有效性,本次实验选择采用东北大学带钢表面缺陷数据集 (NEU-DET) 进行测试。实验环境配置信息如表 1 所示。实验中,初始模型训练选用 SGD 优化器, batchsize 设为 64, epoch 设为 300,其余基本保持 YOLOv5 的默认参数。

表 1 实验环境配置信息

配置名称	版本信息
处理器	Inter Core i9-12900KF
显卡	NVIDIA GeForce RTX 3090
操作系统	Windows 10
编译器	Pycharm 2022
框架	Pytorch 1.12.1

2.1 数据集准备

根据国际统一规定,钢材表面缺陷共分为 33 种类型。本文根据东北大学钢材表面缺陷数据集 (NEU-DET) 选择了 6 种常见的钢材表面缺陷:裂纹 (Cracking, Cr)、夹杂 (Inclusion, In)、斑块 (Patches, Pa)、表面点蚀 (Pitted_surface, Ps)、轧制氧化皮 (Rolled-in_scale, Rs) 和划痕 (Scratches, Sc)。数据集总共有 1800 张灰度图片,每种缺陷含 300 张图片。样图如图 5 所示。首先将数据集按 8 : 1 : 1 的比例划分为训练集、验证集、测试集,训练集为 1440 张,验证集为 180 张,测试集为 180 张。其

次,由于图片本身分辨率等原因,将输入图片的尺寸设为 224×224 进行训练。

2.2 评价指标

为了多方面评估模型的性能,本次实验引入均值平均精度(mAP)、平均精度(AP)、召回率(Recall)、参数量(Parameters)、计算量(FLOPS)、检测速度(FPS)等参数。其中 Parameters 和 FLOPS 分别用于评价模型的体积大小和计算复杂度。均值平均精度(mAP)用于对模型检测精度进行评价,计算公式如下:

$$P = \frac{TP}{TP + FP}$$
(10)

$$R = \frac{TP}{TP + FN}$$
(11)

其中:TP 表示正样本被识别正确的数量;FP 表示预测错误的负样本数量;FN 表示负样本被预

测错误的数量。

$$AP = \int_0^1 p(R) dR$$
(12)

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n}$$
(13)

其中:n 表示训练样本的类别数,i 表示当前类别的编号。

2.3 实验结果分析

2.3.1 消融实验

为验证文中所提改进方法的有效性,本文通过将不同方法进行叠加,来验证上述方法对模型性能的提升。消融实验结果如表 2 所示,其中“√”表示使用此方法进行改进。

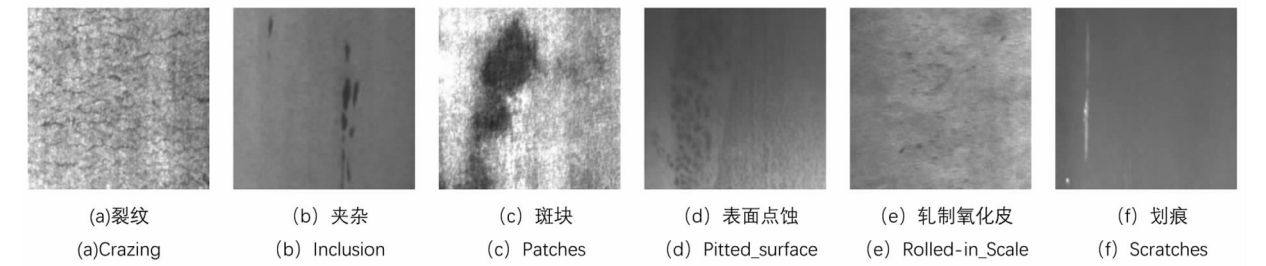


图 5 钢材表面缺陷样本图

表 2 消融实验结果				
Method	Exp1	Exp2	Exp3	Exp4
Yolov5s	√	√	√	√
MSConv		√	√	√
M-BiFPN			√	√
SIoU				√
mAP/%	76.3	75.6	77.4	78.1
Parameters/M	7.036	3.725	3.980	3.981
GFLOPs	15.8	9.1	10.2	10.1

如表 2 所示,以 YOLOv5s 为研究基线,对改进后模型的性能进行考量。第一组实验用基线网络 YOLOv5 进行检测,其 mAP 值为 76.3%、参数量为 7.036M。第二组实验将提出的轻量化卷积模块 MSConv 替换标准卷积,相较于基线网络,模型参数下降了 47.1%,计算量降低了 42.4%,mAP 仅降低了 0.8%,以极小的代价降低了计算成本和空间成本。第三组实验中用改进的 M-BiF-

PN 网络替换 PANet 网络,对比第二组实验,mAP 提升了 1.8%,同时检测精度完成了对基线网络的反超。第四组实验中,用 SIoU 损失函数替换 CIoU 损失函数,相较于实验三,模型参数量与计算量几乎保持不变,mAP 提升了 0.7%。改进后的模型在检测精度、模型参数量、计算复杂度上均优于基线网络,完成模型轻量化的同时,保证了模型的检测精度,符合模型轻量化设计的要求。

图 6 显示了 YOLOv5 与 YOLO-LSNet 模型对各类缺陷的检测情况,由图可知,改进后的模型对同一缺陷置信度更高。同时,原始 YOLOv5 网络检测过程中出现的检测框冗余也有所改进。

2.3.2 对比实验

为验证本文所提的 YOLO-LSNet 钢材表面缺陷检测模型与当前主流模型的性能差异。将该模型与 YOLOv5、YOLOv7^[20]、SSD、Faster R-CNN、YOLO-Steel 在带钢表面缺陷数据集上进行对比。实验结果如表 3 所示。

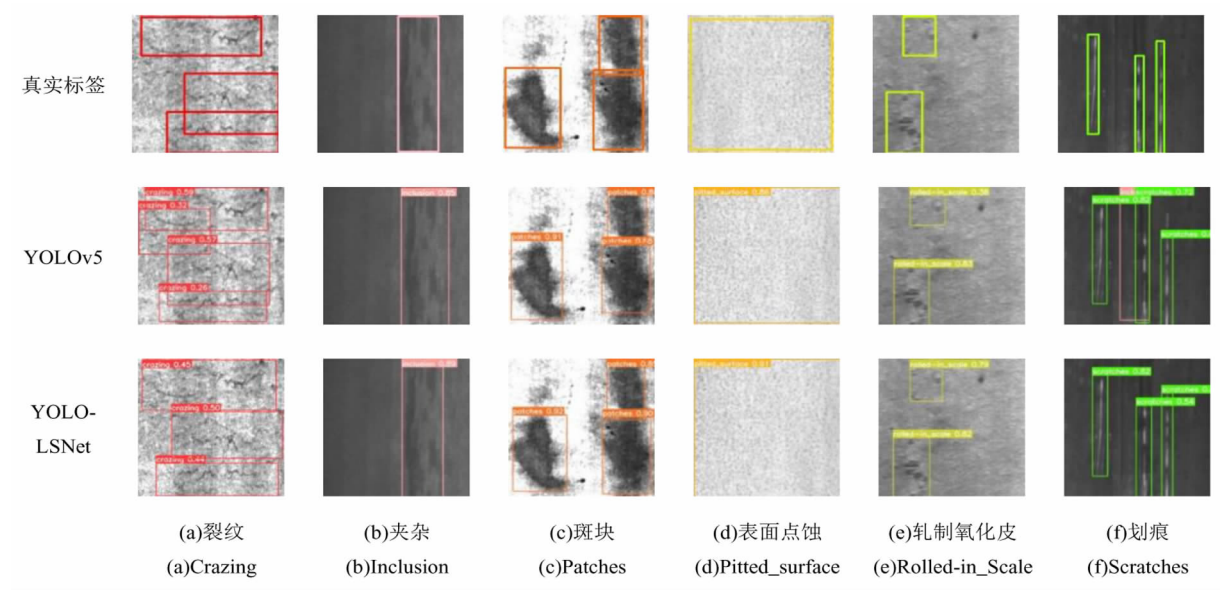


图 6 YOLO-LSNet 模型检测图

表 3 不同算法模型检测性能对比

Module	AP/%						mAP /%	Parameters /M	GFLOPs	FPS
	Cr	In	Pa	Ps	Rs	Sc				
YOLOv5	41.9	86.0	96.0	81.9	56.2	96.6	76.3	7.036	15.8	122.8
YOLOv7	37.9	81.3	96.6	89.9	61.7	95.1	77.1	36.503	103.2	116.9
SSD	37.6	73.5	85.3	63.7	63.2	67.1	65.07	24.146	75.4	60.02
Faster R-CNN	38.2	79.1	96.3	80.3	62.2	91.1	74.5	31.570	67.2	28.51
YOLO-Steel	46.2	84.0	91.8	83.8	64.8	92.8	77.2	7.038	15.9	115.0
YOLO-LSNet	38.9	81.9	97.2	88.6	71.4	90.7	78.1	4.013	10.1	117.3

由表 3 可知,本文所提的轻量化钢材表面缺陷检测模型 YOLO-LSNet,mAP 为 78.1%,参数数量为 4.013M,计算量 GFLOPs 为 10.1,检测速度 FPS 可以达到 117.3。相较于当前 YOLO 系列最新的 YOLOv7 模型,mAP 提升了 1%,同时模型参数降低了 89%,计算量降低了 90.2%。相较于经典的二阶段目标检测模型 Faster R-CNN 和单阶段目标检测算法 SSD,mAP 分别提升了 3.6%、13.03%,同时在计算复杂度和模型参数方面都展现出较大的优势。针对轻量化钢材表面缺陷检测,文献[21]提出了一种 YOLO-Steel 模型,相较于 YOLO-Steel,本文所提出的轻量化检测模型在提升检测精度的同时,进一步降低了模型以及计算量。

实验结果表明,YOLO-LSNet 在检测精度、模型参数大小、计算量这三项指标上均优于当前主流的缺陷检测模型,在检测精度、模型规模、计算复杂

度之间实现了良好的均衡,符合轻量化设计的要求。

3 结 论

针对当前钢材表面缺陷检测过程中存在模型参数量大、计算复杂度高,难以在终端设备部署的问题,提出了一种轻量化钢材表面缺陷模型 YOLO-LSNet。实验结果表明,改进后的模型在保证检测性能的同时,降低了对终端硬件的要求,对实际生产中的部署具有重要意义。

参考文献

[1] 米春风,卢琨,汪文艳,等. 基于机器视觉的带钢表面缺陷检测研究进展[J]. 安徽工业大学学报(自然科学版),2022,39(2):180—188.

[2] 李跃,王子铭,李鑫林,等. 带钢表面缺陷检测方法研究进展

- [J/OL]. 钢铁研究学报:1—17[2023—07—28]. DOI:10.13228/j.boyuan.issn1001—0963.20220363.
- [3] 程锦锋,方贵盛,高惠芳.表面缺陷检测的机器视觉技术研究进展[J].计算机应用研究,2023,40(4):967—977
- [4] 马燕婷,赵红东,阎超,等.改进 YOLOv5 网络的带钢表面缺陷检测方法[J].电子测量与仪器学报,2022,36(8):150—157.
- [5] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440—1448.
- [6] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster r-cnn: towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28.
- [7] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[J]. Springer, Cham, 2016. DOI:10.1007/978—3—319—46448—0_2.
- [8] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [10] ZHOU Xinwen, WEI Mengen, LI Qianglong, et al. Surface defect detection of steel strip with double pyramid network[J]. Applied Sciences, 2023, 13(2): 1054.
- [11] WANG Han, YANG Xiuding, ZHOU Bei, et al. Strip surface defect detection algorithm based on YOLOv5[J]. Materials, 2023, 16(7): 2811.
- [12] SHI Jianting, YANG Jian, ZHANG Yingtao. Research on steel surface defect detection based on YOLOv5 with attention mechanism[J]. Electronics, 2022, 11(22): 3735.
- [13] WANG Rijun, LIANG Fulong, MOU Xiangwei, et al. Development of an improved YOLOv7-based model for detecting defects on strip steel surfaces[J]. Coatings, 2023, 13(3): 536.
- [14] GEVORGYAN Z. SIOU loss: more powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.
- [15] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [16] ZHANG Xiangyu, ZHOU Xinyu, LIN Meng-xiao, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6848—6856.
- [17] HAN Kai, WANG Yunhe, TIAN Qi, et al. Ghostnet: more features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1580—1589.
- [18] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 11534—11542.
- [19] TAN Mingxing, PANG Ruoming, LE Q V. Efficientdet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 10781—10790.
- [20] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464—7475.
- [21] 周彦,孟江南,王冬丽等.基于多尺度轻量化注意力的钢材缺陷检测[J/OL].控制与决策:1—9[2023—07—28]. DOI:10.13195/j.kzyjc.2022.0997.