

基于 AM-YOLOv3 模型的铝型材表面缺陷检测算法

孙连山¹, 魏婧雪^{1*}, 朱登明², 石敏³

¹ 陕西科技大学电子信息与人工智能学院, 陕西 西安 710021;

² 中国科学院计算技术研究所前瞻研究实验室, 北京 100080;

³ 华北电力大学控制与计算机工程学院, 北京 102206

摘要 针对铝型材表面纹理复杂、缺陷尺度差异大导致缺陷检测算法精度不高, 难以满足实际需求的问题, 提出了一种改进的目标检测网络 AM-YOLOv3 (attention-guided multi-scale fusion YOLOv3)。设计了注意力引导模块和 4 个预测尺度, 实现了铝型材表面缺陷的多尺度特征提取。构建了自底向上的特征传输路径, 与原有的特征金字塔网络结合形成双塔结构, 实现了多尺度特征融合。使用 K-medians 算法进行锚框聚类, 更准确地表征了锚框尺寸的分布规律, 提高了网络的收敛速度。在公开的铝型材数据集上对所提算法的性能进行了实验验证。实验结果表明, 所提算法的 mAP (mean average precision) 达到了 99.05%, 比 YOLOv3 模型提高了 6.8%, 帧频达到了 43.94 frame/s。

关键词 机器视觉; 目标检测; 铝型材表面缺陷检测; 注意力引导; 多尺度融合; 双塔结构

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.2415007

Surface Defect Detection Algorithm of Aluminum Profile Based on AM-YOLOv3 Model

Sun Lianshan¹, Wei Jingxue^{1*}, Zhu Dengming², Shi Min³

¹ School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an, Shaanxi 710021, China;

² Foresight Research Laboratory, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China;

³ School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China

Abstract In order to solve the problem that the accuracy of defect detection algorithm is not high and it is difficult to meet the actual requirements due to the complex surface texture of aluminum profile and the large difference in defect size, an improved object detection network AM-YOLOv3 (attention-guided multi-scale fusion YOLOv3) is proposed. The attention guide module and four prediction scales are designed to realize the multi-scale feature extraction of aluminum profile surface defects. A bottom-up feature transmission path is constructed, which is combined with the original feature pyramid network to form a twin-towers structure, and the multi-scale feature fusion is realized. K-medians algorithm is used for anchor box clustering, which more accurately characterizes the distribution law of anchor frame size and improves the convergence speed of the network. The performance of the proposed algorithm is verified by experiments on the public aluminum profile dataset. The experimental results show that the mAP (mean average precision) of the proposed algorithm reaches 99.05%, which is 6.8% higher than that of the YOLOv3 model, and the frame rate reaches 43.94 frame/s.

收稿日期: 2021-03-29; 修回日期: 2021-05-20; 录用日期: 2021-06-02

基金项目: 陕西省科技厅自然科学基金基础研究计划项目(2019JM-354)、陕西省教育厅专项科研计划项目(17JK0087)、中国科学院科研仪器设备研制项目(YJKYYQ20190055)

通信作者: *1779664651@qq.com

Key words machine vision; object detection; surface defect detection of aluminum profiles; attention guidance; multi-scale fusion; twin-towers structure
OCIS codes 150.1135; 100.4996; 100.3008

1 引言

在生产和运输过程中,当铝型材受到环境等多方面因素影响时,其表面容易产生各种各样的缺陷。这些缺陷会严重影响铝型材的质量及后期的装配使用。铝型材表面缺陷检测常常采用的是人工检测的方式,但是人工检测存在准确率低、效率低、易受主观影响的缺点^[1],难以满足大规模的铝型材检测需求。

传统的铝型材表面缺陷检测主要采用的是依赖于手工提取图像特征的机器学习方法,如李超贤^[2]首先利用 Hu 矩提取了铝型材表面的缺陷特征,然后将这些缺陷特征送入 4 个不同的支持向量机(SVM)分类器中进行训练。然而,当铝型材表面缺陷出现尺度变化剧烈、缺陷区域特征和背景纹理特征相似等情况时^[3],传统检测方法无法满足检测需求。传统的机器学习算法需要人工设计以提取特征,而深度学习可以直接从底层数据中自动学习特征,改变了缺陷特征的提取方式,大大提高了缺陷检测性能^[4]。

目前,应用于缺陷检测领域的目标检测算法主要分为两类^[5]。一类是依赖锚框的检测算法,包括以 Faster R-CNN^[6-8]为代表的两阶段网络和以 SSD(single shot multibox detector)^[9]和 YOLO(you only look once)^[10-13]系列为代表的一阶段网络。另一类为不依赖锚框的检测算法,如基于关键点的检测算法 CornerNet^[14]、基于中心的检测算法 CenterNet^[15]和一阶全卷积目标检测(FCOS)算法^[16]等。

在实际应用方面,李东洁等^[17]将 Faster R-CNN 和特征金字塔网络(FPN)^[18]相结合使马克杯缺口和划痕的检测精度提升了 2.485%。Li 等^[19]优化了 SSD 结构,简化了检测模型参数,提出了一种基于轻量级 SSD 的灌装容器表面缺陷检测方法。文生平等^[20]将 YOLOv3 网络与变形卷积技术和密集连接技术相结合,提高了条状缺陷和小目标缺陷的检测精度。同时,改进后的算法对铝型材表面 4 类缺陷样本检测的 mAP(mean average precision, M_{mAP})达到了 88.85%。周天宇等^[21]在 YOLOv3 的基础上改进了多尺度检测方法,并利用 K-means

算法聚类生成了初始化锚框,最终实现了载波芯片(COC)缺陷检测。王璇等^[22]将轻量级 MobileNetV3 作为骨干网络,构建了 AFP-CenterNet,实现了自动纤维铺放铺层表面缺陷的检测。

虽然上述改进网络均能够较好地实现端到端检测,但它们仍难以满足铝型材表面缺陷检测的精度需求。为此,本文建立了注意力引导的多尺度融合网络——AM-YOLOv3,以提高铝型材表面缺陷的检测精度。该网络在 YOLOv3 三个预测尺度的基础上,添加了一个包含更多细节信息的大尺度特征图,解决了三个预测尺度包含小目标信息量不足的问题。通过设计注意力引导模块,将多分支卷积和通道注意力模块(SE block)^[23]结合,在增加少量计算量的情况下提高了网络的表征能力,解决了主干特征提取网络 DarkNet-53 特征提取能力有限的问题。在 FPN 的基础上,构建了自底向上的路径以向高层传递低层的强定位信息,解决了模型对小目标检测效果差的问题。使用 K-medians 聚类算法生成的锚框参数可以准确表征铝型材表面缺陷的分布,解决了 K-means 算法锚框聚类参数精度低引起的网络推理速度减缓问题。实验结果表明,所提算法能够准确识别铝型材表面缺陷,且满足实时性要求。

2 相关理论

2.1 YOLOv3 原理

YOLOv3 的整体网络结构如图 1 所示,其内部基本组件如图 1 中的虚线框所示。其中,DBL 表示基础卷积模块,Conv 表示卷积层,Leaky ReLU 表示激活函数,Res unit 表示残差单元,ResN 表示具有 N 个残差单元的残差模块,Zero Padding 表示填充层,Y1 表示特征图大小为 13×13 的预测分支 1,Y2 表示特征图大小为 26×26 的预测分支 2,Y3 表示特征图大小为 52×52 的预测分支 3, N 表示残差单元个数,concat 表示张量拼接单元,Add 表示张量相加单元。当图像输入 YOLOv3 网络后,首先经过 DarkNet-53 网络的 32 倍、16 倍、8 倍下采样可输出三种尺度的特征图,用于预测不同大小的目标。然后,为了提升多尺度检测性能,利用 FPN 创建自顶向下的路径传递信息,实现了高低层特征的融合。最后,将 FPN 融合后的特征送入预测模块以获取预

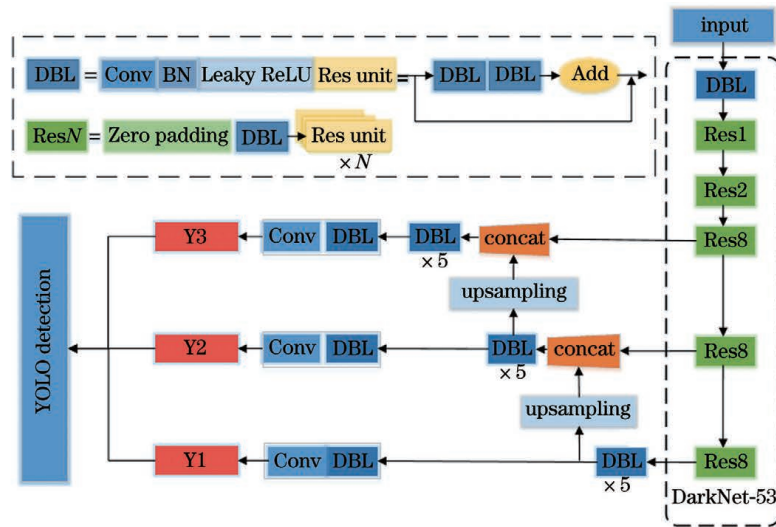


图 1 YOLOv3 的整体网络结构

Fig. 1 Overall network structure of YOLOv3

测结果。

YOLOv3 网络三个预测尺度的输出维度分别为 $N_1, N_1, 3 \times (S + 1 + 4)$ 。其中, N_1 代表特征图划分的 $N_1 \times N_1$ 个网格, 3 代表每个预测尺度设置三种锚框, S 代表待检测目标的类别, 1 代表预测框置信度, 4 代表预测框坐标值。

2.2 通道注意力模块

SE block 通过学习特征图的全局信息自适应地为各通道分配权重。该机制可以使检测网络学习到更多重要的特征信息, 减少无关信息的干扰, 从而合理利用有限资源, 提升网络性能。上述过程可以描述为

$$F_{scale}(v, x) = \sum_{s=1}^C v_s \times x_s, \quad (1)$$

式中: x 为原始特征通道; v 为分配的新权重; C 为

通道数; $F_{scale}(v, x)$ 为利用乘法逐通道加权方式对原始特征通道进行重标定的函数。

3 AM-YOLOv3 模型的设计

针对铝型材表面缺陷检测算法精度不高的问题, 提出了一种注意力引导的多尺度融合网络——AM-YOLOv3, 其整体结构如图 2 所示, 其中 Y_4 表示特征图大小为 104×104 的预测分支 4, AG 表示注意力引导模块。AM-YOLOv3 在 YOLOv3 的 3 个预测尺度的基础上新增 1 个大尺度特征图, 构成了 4 个预测尺度。在将待检测图像输入 AM-YOLOv3 网络中后, 首先通过 DarkNet-53 提取缺陷特征。然后, 将提取后的特征送入四分枝注意力引导模块(FAG)中以进一步提取缺陷特征。接着, 将 FAG 提取到的特征送入双塔结构中, 进而实现各

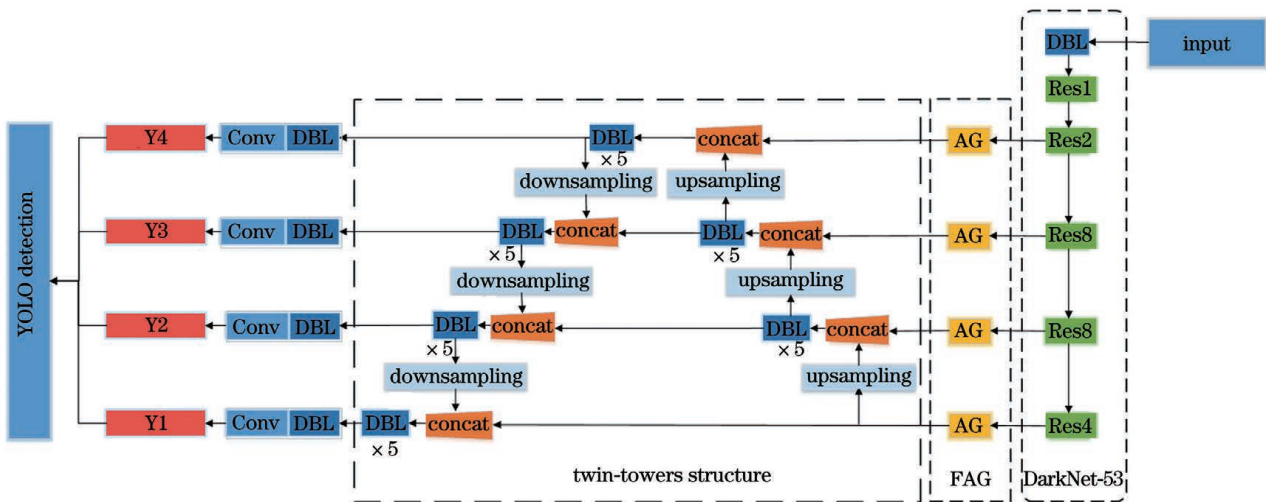


图 2 AM-YOLOv3 的整体网络结构

Fig. 2 Overall network structure of AM-YOLOv3

预测尺度的特征融合。最后,将双塔结构融合后的特征送入预测模块中以获取铝型材表面缺陷的检测结果。

3.1 多尺度预测的改进

铝型材数据集中存在大量的小目标缺陷,而 YOLOv3 模型在检测小目标缺陷时效果较差,这是因为随着主干特征提取网络 DarkNet-53 不断地下采样,有利于检测小目标缺陷的细节信息不断减少。若要提高模型对小目标缺陷的检测精度,就需要大尺度特征图提供更多的细节定位信息。YOLOv3 进行预测的最大尺度特征图为 52×52 ,该特征图所包含的细节信息不足,难以满足小目标缺陷的检测需求,因此需要一个包含更多细节信息的大尺度特征图。

为解决上述问题,AM-YOLOv3 在 YOLOv3 网络的 3 个预测尺度基础上,添加了 1 个尺寸为 104×104 的浅层大尺度特征图,构成了 4 个预测尺度。预测尺度数量的增加扩充了特征信息量,进而提高了铝型材表面小目标缺陷的检测精度。

3.2 注意力引导模块

铝型材表面缺陷尺度变化范围大,而 DarkNet-53 的网络层浅,特征提取能力有限。因此,其他基于 YOLOv3 改进的网络常常通过增加网络深度的方式来提高网络特征提取能力。然而,增加网络深度会造成计算量的增加,进而导致网络效率降低。

针对上述问题,设计了注意力引导模块。该模块不仅可以在引入较少参数的情况下,提高网络特征提取能力,还可以减少无关信息的干扰,提升网络性能。

注意力引导模块由五分支卷积块(FBC)和 SE block 共同构成,其结构如图 3 所示,其中 conv2d 表示二维卷积,FC 表示全连接层,ReLU 和 Sigmoid 表示两类常用的激活函数,Scale 表示将新分配权重与原始特征图进行逐通道乘法加权的单元, $H \times W \times C$ 表示通道数为 C 、大小为 $H \times W$ 的特征图, r 为超参数。将注意力引导模块添加在 AM-YOLOv3 的 4 个预测尺度上,可构成如图 2 所示的 FAG。FBC 是多分支并行结构,除了 shortcut 分支外,其余每个分支由普通卷积层和膨胀卷积层构成。普通卷积层采用不同大小($1 \times 1, 3 \times 3, 5 \times 5, 7 \times 7$)的卷积核来捕捉缺陷的多尺度信息。为了减少计算量,FBC 中使用多个小卷积核替换大卷积核。膨胀卷积层可以在不增加太多计算量的情况下保证分辨率不变并扩大感受野。由于网络对不同尺度缺陷的检测性能与感受野大小呈正相关,因此使用扩张率不同(1,3,5,7)的膨胀卷积^[24]形成不同大小的感受

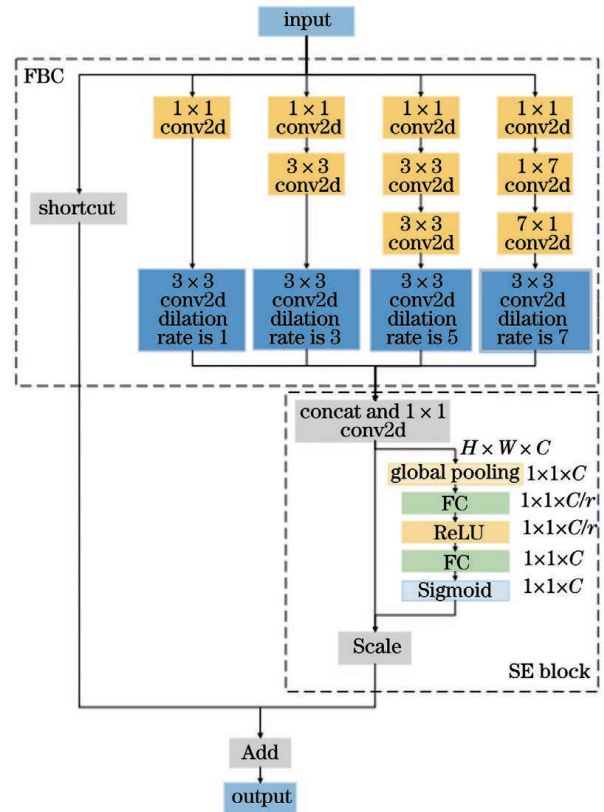


图 3 注意力引导模块的结构示意图

野,以更好地覆盖多尺度缺陷。最后,利用通道拼接方法将多分支融合以构成 FBC。考虑到不同通道特征对网络性能的贡献程度不同,因此在 FBC 后引入 SE block。

由于 FBC 可以在空间维度上增强网络的特征提取能力,SE block 能够在通道维度上提升网络性能,因此两者构成的注意力引导模块可以从空间和通道维度上共同提升网络的表达能力。

3.3 双塔结构

高低层的语义信息和定位细节信息之间存在不均衡情况,即低层包含定位信息多,语义信息少;高层包含语义信息多,定位信息少。针对上述问题,YOLOv3 利用 FPN 融合高低层特征以提高网络性能。FPN 具有横向连接和自顶向下的路径扩展,从而可以让高层的语义信息直接传播至底层。然而,FPN 不具备自底向上的路径扩展,因此低层定位信息不易向高层传递,这增加了网络获取准确定位信息的难度。

为了缩短底层与高层间的信息传递路径,AM-YOLOv3 在 4 个预测尺度的基础上,构建了自底向上的路径。首先,对尺寸为 $104 \times 104, 52 \times 52, 26 \times 26$ 的特征图执行下采样操作。然后,将其与

DarkNet-53 中高一层级的特征图横向连接。自底向上的金字塔结构与 FPN 共同构成了如图 4 所示的双塔结构,弥补了 FPN 对细节信息感知能力差的缺陷,降低了网络获取低层定位信息的难度。双塔结构可以对各预测尺度进行参数聚合,使每个尺度的特征既包含强语义信息又包含强定位信息。

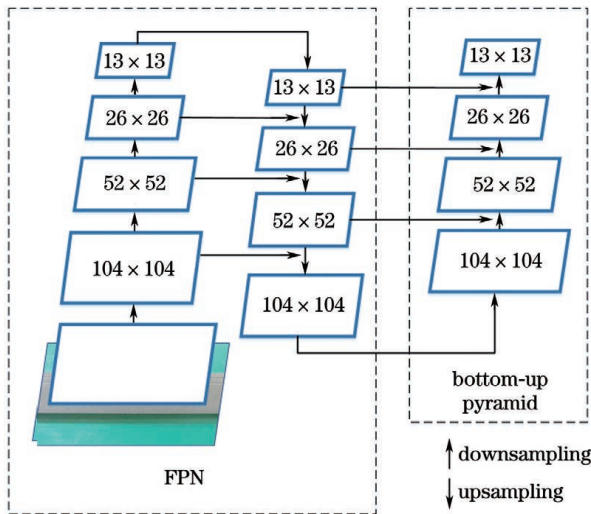


图 4 双塔结构示意图

Fig. 4 Schematic diagram of twin-towers structure

3.4 锚框聚类算法

YOLOv3 需要提前通过聚类算法生成锚框参数,高质量的锚框参数能够准确反应数据分布规律,从而降低网络损失。YOLOv3 的锚框参数是通过公共数据集聚类生成的,但铝型材表面缺陷尺度变化剧烈且存在大量狭长和细小的表面缺陷,故公共数据集聚类得到的锚框参数不适用于铝型材表面缺陷。因此,需要重新聚类生成锚框参数。

原始 YOLOv3 中使用的是 K-means 聚类算法。K-means 算法使用平均值来计算每一簇的聚类中心,但 K-means 算法存在一定的局限性,即平均值容易受到簇中极端值的影响。中位值代表样本数据所占频率的等分线,当数据集中存在几个极端值时,中位值的变化不是很明显甚至可能没有变化,故用中位值来表示铝型材表面缺陷的分布规律更为合适。因此,本文使用 K-medians 聚类算法来生成锚框参数。

AM-YOLOv3 中包含 4 个预测尺度,每个预测尺度需设置 3 个锚框参数,因此聚类后会生成 12 个锚框参数。K-means 和 K-medians 聚类后的锚框参数如表 1、2 所示。其中,锚框尺寸与特征图尺寸大小成反比。若特征图尺寸小,则锚框尺寸大,此时可用于检测大目标。若特征图尺寸大,则锚框尺寸

小,此时可用于检测小目标。从表 1、2 可以发现,K-medians 算法比 K-means 算法的聚类精度提高了 4.59%,说明 K-medians 算法聚类生成的锚框参数能更准确地表征铝型材表面缺陷的分布规律。

表 1 K-means 算法的聚类结果

Table 1 Clustering results of K-means algorithm

| Size | Clustering parameter | | |
|-----------|----------------------|-----------|------------|
| 13×13 | (392, 33) | (408, 91) | (414, 182) |
| 26×26 | (224, 163) | (353, 17) | (365, 57) |
| 52×52 | (87, 90) | (94, 30) | (184, 44) |
| 104×104 | (9, 12) | (21, 25) | (30, 60) |
| Precision | 71.14% | | |

表 2 K-medians 算法的聚类结果

Table 2 Clustering results of K-medians algorithm

| Size | Clustering parameter | | |
|-----------|----------------------|------------|------------|
| 13×13 | (416, 75) | (416, 102) | (416, 190) |
| 26×26 | (179, 44) | (416, 17) | (416, 39) |
| 52×52 | (58, 27) | (94, 30) | (113, 28) |
| 104×104 | (8, 11) | (85, 70) | (25, 46) |
| Precision | 74.41% | | |

4 实验与分析

4.1 实验数据集

铝型材缺陷数据集来自阿里云承办的 2018 年广东工业制造大数据创新大赛——铝型材表面瑕疵识别^[25]。本文主要研究了单一缺陷的铝型材图像,缺陷类别如图 5 所示,包括不导电、划痕、褶皱、喷流、脏点。

由于原始数据分布存在类别间数据不均衡的问题,因此将水平翻转、竖直翻转、明暗变换、旋转缩放等方法随机结合对铝型材图像进行数据增强,数据增强结果如图 6 所示。扩充后的图像共计 10570 张,各类缺陷图像扩充前后的数量如表 3 所示。使

表 3 扩充前后缺陷图像数量的比较

Table 3 Comparison of number of defect images before and after expansion

| Class | Original image | Enhanced image |
|----------------|----------------|----------------|
| Non-conduction | 375 | 2250 |
| Scratch | 104 | 2080 |
| Wrinkle | 151 | 2114 |
| Jet | 66 | 2046 |
| Spot | 104 | 2080 |
| Total | 800 | 10570 |

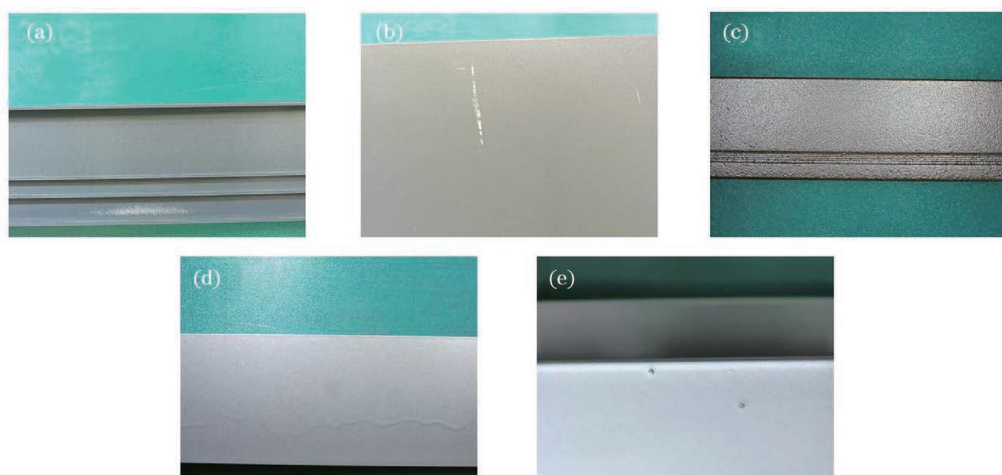


图 5 铝型材数据集中的 5 类缺陷。(a)不导电;(b)划痕;(c)褶皱;(d)喷流;(e)脏点

Fig. 5 Five types of defects in aluminum profile dataset. (a) Non-conduction; (b) scratch; (c) wrinkle; (d) jet; (e) spot

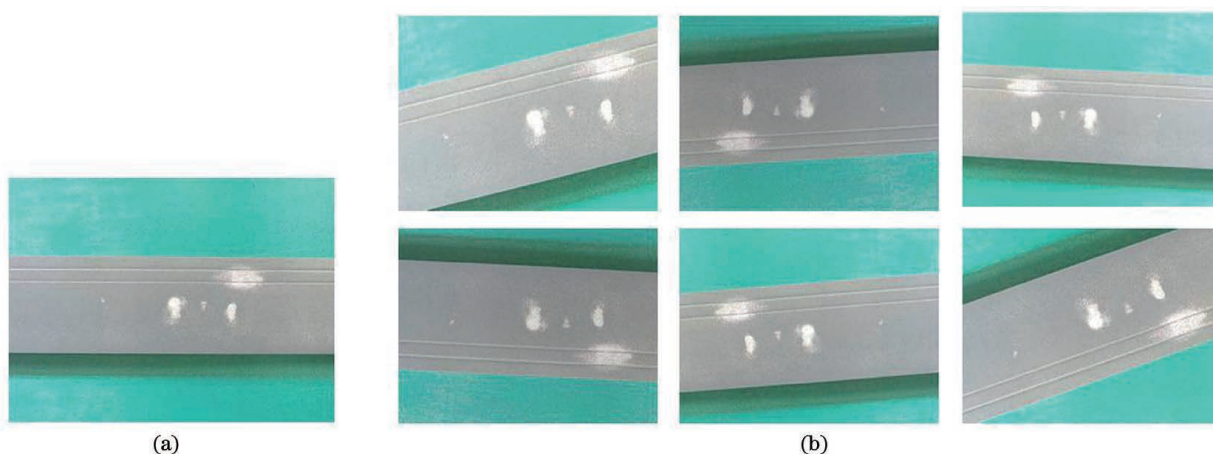


图 6 数据增强结果。(a)原始图像;(b)增强后图像

Fig. 6 Data enhancement results. (a) Original image; (b) enhanced images

用 LabelImg 工具标注铝型材表面缺陷,生成的 XML 文件供目标检测网络使用。

4.2 模型训练与评价指标

本文的深度学习框架为 PyTorch1.3.1,服务器为 NVIDIA GeForce RTX 2080 Ti,选用 10.0 版本的 CUDA,7.6 版本的 cuDNN,0.4.2 版本的 torchvision。

在训练 AM-YOLOv3 网络时,需要调整 AM-YOLOv3 网络的内置参数使其符合铝型材数据集的训练要求。修改后的部分训练参数如表 4 所示。

表 4 修改后的部分训练参数

Table 4 Partial training parameters after modification

| Parameter | Value |
|--------------------------------|--------|
| Epoch | 150 |
| Initial learning rate | 0.0010 |
| Learning rate when epoch is 50 | 0.0001 |
| Momentum | 0.9 |

设置初始学习率为 0.0010,最大迭代次数为 150。采用冻结方式训练,当训练周期(epoch)数达到 50 时,学习率变化为 0.0001,直到 epoch 数达到 150 时结束训练。优化器采用带动量的梯度随机下降(SGD),动量设置为 0.9。

选用 mAP 和帧频(FPS, f_{FPS})来评价模型的性能。mAP 是衡量模型检测精度的指标,它是通过对所有类别的 AP(average precision, M_{AP})求和除以类别数得到的,mAP 越大代表模型检测精度越高。FPS 是衡量模型处理速度的指标,FPS 越大代表模型速度越快。两个指标的表达式为

$$M_{AP} = \int_0^1 p(r) dr, \quad (2)$$

$$M_{mAP} = \frac{1}{n_1} \sum_{i=1}^{n_1} M_{AP, i}, \quad (3)$$

式中: $p(r)$ 表示当分类阈值取不同 r 时计算所得的精确率; n_1 表示待测目标的总类别数; $M_{AP, i}$ 表示

第 i 类的 AP。

图 7 展示了 5 类缺陷的 AP, 此时 AM-YOLOv3 模型的 M_{mAP} 为 99.05%。AM-YOLOv3 模型的 mAP 曲线如图 8 所示, 其中大圆点代表最后一轮的记录点。可以发现, 随着 epoch 的增加, mAP 曲线呈现上升趋势。当 epoch 达到 107 左右时, mAP 开始趋于稳定。

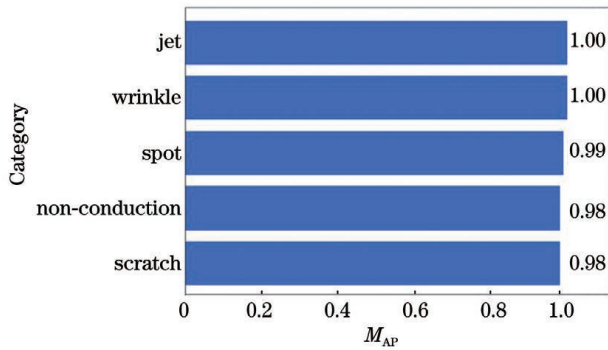


图 7 5 类缺陷的 M_{AP}

Fig. 7 M_{AP} of five types of defects

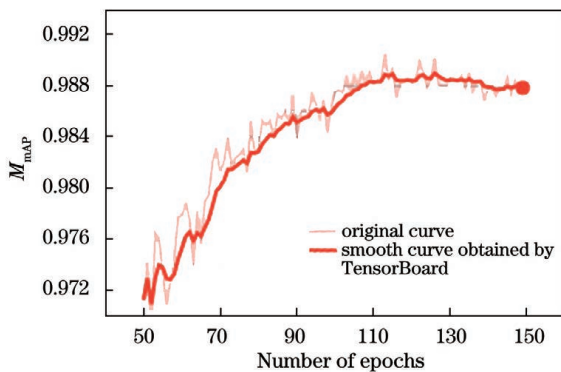


图 8 mAP 曲线

Fig. 8 mAP curve

4.3 实验结果

为了验证所提方法的有效性, 将 AM-YOLOv3 划分为 5 组, 设计了消融实验, 实验结果如表 5 所示, 其中 \times 表示未选用相应算法, \checkmark 表示选用相应算

法。对比 YOLOv3 和 Group A 可知, 采用 K-medians 算法可以使 mAP 提高 0.99%, FPS 提高 15.36%, 说明采用锚框聚类精度更高的 K-medians 算法可以提高检测精度, 同时也可以加快网络的推理速度。对比 Group A 和 Group B 可知, 当预测尺度由 3 个增加到 4 个时, 可以使 mAP 提高 2.85%, 表明增加预测尺度数量能够有效扩充缺陷特征信息量, 增强网络表达能力。通过 Group B 和 Group C 的对比可知, 加入 FAG 可以使 mAP 提高 2.01%, 说明提高网络特征提取能力有利于提升网络精度。最后, 根据 Group C 与 AM-YOLOv3 的对比可知, 构建双塔结构来融合高低层特征, 可以使 mAP 提高 0.95%。

根据表 5 的分析可知, 提出的 4 种改进方法对 AM-YOLOv3 算法都是必要的, 且这 4 种改进方法带来了相当大的收益。4 种改进方法组合后的 AM-YOLOv3 算法可以在增加少量计算量的情况下使原始 YOLOv3 算法的 mAP 提高 6.80%。

为了进一步验证 AM-YOLOv3 算法的有效性, 在相同实验条件下, 将 AM-YOLOv3 算法与多种目标检测算法进行实验对比, 不同算法的性能结果如表 6 所示。

首先, 将 AM-YOLOv3 算法与经典的基于锚框的算法 Faster R-CNN、SSD 和 YOLO 系列进行对比。可以发现 AM-YOLOv3 算法相较于以 VGG16 为主干特征提取网络的 Faster R-CNN, mAP 提高了 37.25%, FPS 提高了 155.32%, 表明 AM-YOLOv3 算法在检测精度和速度方面具有较大优势。出现该现象的原因: 1) VGG16 特征提取能力有限, 难以满足铝型材表面缺陷的检测需求; 2) Faster R-CNN 的锚框是预先设定的, 无法根据缺陷目标大小进行调整, 导致 Faster R-CNN 检测效果差。与 SSD 相比, AM-YOLOv3 明显提高了铝型材表面各类缺陷的 AP, 且使 mAP 提高了 24.39%, 但速度有一定程度的下降。这是因为 AM-YOLOv3 通过双

表 5 消融实验结果对比

Table 5 Comparison of ablation experimental results

| Algorithm | K-medians | Four feature scales | FAG | Twin-towers structure | $M_{mAP} / \%$ | $f_{FPS} / (\text{frame} \cdot \text{s}^{-1})$ |
|-----------|--------------|---------------------|--------------|-----------------------|----------------|--|
| YOLOv3 | \times | \times | \times | \times | 92.74 | 48.83 |
| Group A | \checkmark | \times | \times | \times | 93.66 | 56.33 |
| Group B | \checkmark | \checkmark | \times | \times | 96.31 | 50.28 |
| Group C | \checkmark | \checkmark | \checkmark | \times | 98.17 | 47.85 |
| AM-YOLOv3 | \checkmark | \checkmark | \checkmark | \checkmark | 99.05 | 43.94 |

表 6 不同算法的性能

Table 6 Performance of different algorithms

| Algorithm | Backbone | AP of non-conduction detection / % | AP of scratch detection / % | AP of jet detection / % | AP of wrinkle detection / % | AP of spot detection / % | $M_{mAP} / \%$ | $f_{FPS} / (\text{frame} \cdot \text{s}^{-1})$ |
|-------------------------------|--------------|------------------------------------|-----------------------------|-------------------------|-----------------------------|--------------------------|----------------|--|
| Faster R-CNN ^[8] | VGG16 | 87.41 | 74.56 | 79.97 | 96.65 | 22.25 | 72.17 | 17.21 |
| SSD ^[9] | VGG16 | 89.79 | 79.88 | 88.96 | 98.92 | 40.62 | 79.63 | 77.17 |
| YOLOv3 ^[11] | DarkNet-53 | 93.92 | 95.18 | 98.50 | 99.80 | 76.32 | 92.74 | 48.83 |
| YOLOv3 | ResNet152 | 94.85 | 93.97 | 98.62 | 99.26 | 75.80 | 92.50 | 27.37 |
| YOLOv4 ^[12] | CSPDarknet53 | 95.21 | 95.79 | 98.99 | 99.76 | 89.96 | 95.94 | 45.29 |
| YOLOv5 ^[13] | CSPDarknet53 | 96.83 | 97.34 | 99.95 | 99.98 | 93.58 | 97.53 | 42.37 |
| CenterNet ^[15] | Resnet-101 | 85.87 | 68.33 | 79.89 | 99.15 | 43.93 | 75.43 | 70.61 |
| shuzhilian ai ^[26] | Resnet-101 | 94.48 | 81.80 | 67.73 | 98.30 | 51.12 | 78.69 | 21.43 |
| AM-YOLOv3 | DarkNet-53 | 98.13 | 98.00 | 100.00 | 99.96 | 99.12 | 99.05 | 43.94 |

塔结构融合高低层特征,虽然增加了一定的计算量,但是获取了丰富的特征信息。然而,SSD 高低层特征之间相互独立,获取的特征信息较为单一。相较于 YOLO 系列算法,AM-YOLOv3 算法的 mAP 明显优于 YOLOv3 (DarkNet-53) 网络以及 YOLOv3 (ResNet152) 网络,且检测速度与 YOLOv3 (DarkNet-53) 基本持平。当 YOLOv3 的主干网络使用更深层的 ResNet152 时,mAP 比 YOLOv3 (DarkNet-53) 下降了 0.26%,FPS 降低了 43.95%。然而,根据表 5 中的实验结果可知,加入注意引导模块后的 mAP 比 YOLOv3 (DarkNet-53) 提升了 2.01%,且 FPS 只降低了 4.98%,表明本文提出的注意引导模块能够在增加少量计算量的情况下,提高网络特征提取能力。对比 YOLOv4 可知,AM-YOLOv3 算法使 mAP 提高了 3.24%,脏点缺陷检测的 AP 提高了 10.18%,并且其他尺度的缺陷的 AP 也达到了较高的水平,说明 AM-YOLOv3 算法对尺度具有一定的鲁棒性。将 AM-YOLOv3 算法与 YOLOv5 对比可得,AM-YOLOv3 算法能够在推理速度稍占优势的情况下,使 mAP 提高 1.56%。其中,小目标缺陷检测精度提升最为明显,脏点缺陷检测的 AP 提高了 5.92%。

其次,将 AM-YOLOv3 算法与不依赖锚框的算法 CenterNet 对比。从速度来看,AM-YOLOv3 算法较 CenterNet 算法有所下降。然而,从检测精度来看,CenterNet 算法的 mAP 为 75.43%,与 AM-YOLOv3 相比没有优势。铝型材表面各类缺陷检

测的 AP 均有明显提高,其中,脏点缺陷检测的 AP 提高了约 2.3 倍。

最后,将所提算法与基于 Faster R-CNN 改进的 shuzhilian ai 算法^[26] 进行对比。可以发现,所提算法在检测精度和速度方面都具有较大优势,使 mAP 提高了 25.87%,FPS 提高了 2.1 倍。

由表 6 分析可知,针对铝型材表面的 5 类缺陷,AM-YOLOv3 算法的检测精度对比其他目标检测算法具有很大优势。

为了更加直观地评估算法的优劣,选取算法改进前后的部分检测效果图用于定性分析,如图 9 所示,其中数值表示检测置信度。根据图 9(a)~(e) 的对比可知,改进后的算法提高了多处缺陷的检测置信度。检测置信度的提高说明扩充特征信息量、提高特征提取能力、加强特征融合以及提高锚框聚类精度的改进方法是有效的,能够明显改善算法检测精度。在图 9(e) 所示的脏点缺陷检测中,铝型材表面共有 4 处脏点,但 YOLOv3 算法只识别出了其中的 3 处缺陷,而改进后的算法能够准确地识别出 4 处缺陷。由此可以得出,相比于经典的 YOLOv3 算法,改进后的算法在误检和漏检问题上有明显改善。

4.4 未来改进方向

所提方法主要考虑了检测目标的大小问题,而通过对数据的分析发现,铝型材表面缺陷形态差异较大、分布不规则,即脏点缺陷形态类似小圆点,划痕缺陷大多呈狭长分布,其他缺陷长宽比例分布不

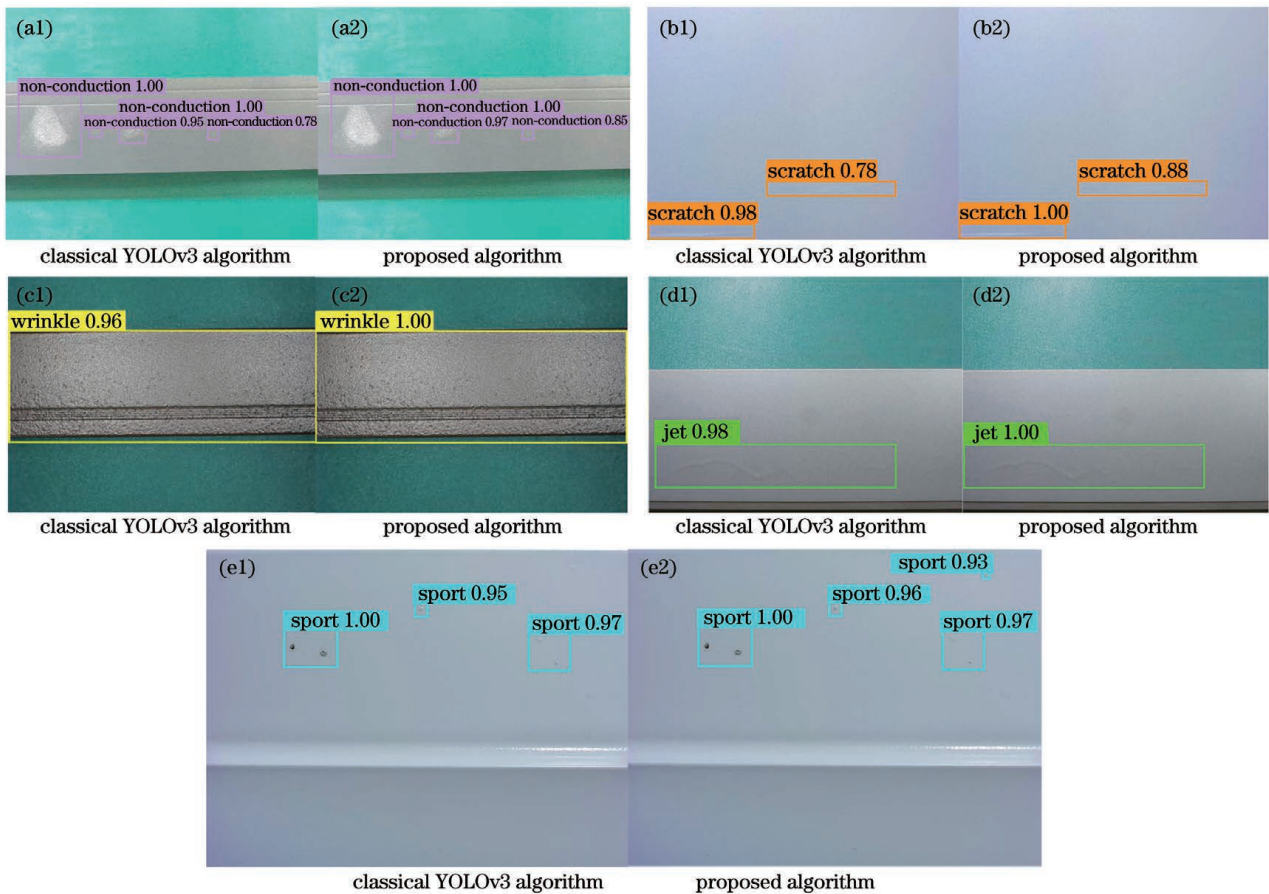


图 9 缺陷检测效果。(a1)(a2)不导电;(b1)(b2)划痕;(c1)(c2)褶皱;(d1)(d2)喷流;(e1)(e2)脏点

Fig. 9 Defects detection effect. (a1)(a2) Non-conduction; (b1)(b2) scratch; (c1)(c2) wrinkle; (d1)(d2) jet; (e1)(e2) spot

均匀。然而,普通的卷积神经网络对不规则的缺陷检测效果不是很好,因此在未来的研究中将尝试通过改进卷积和池化的方式提高算法对不规则缺陷的检测精度。

此外,随着目标检测算法的快速发展,衍生了一系列高性能的检测算法,例如 Double-Head R-CNN^[27]、NAS-FPN^[28]、CentripetalNet^[29]等,这些算法为铝型材表面缺陷检测提供了值得借鉴的新方法与新思路。未来将从实际生产环境中采集数据以丰富铝型材数据集,同时尝试将基于高性能算法的新技术与所提算法相结合,在扩充的铝型材数据集上进行进一步的优化验证。

5 结 论

为了满足铝型材表面缺陷检测的实用性需求,在 YOLOv3 网络的基础上,从扩充特征信息量、提高特征提取能力、加强特征融合以及提高锚框聚类精度方面进行改进,提出了一种注意力引导的多尺度融合网络——AM-YOLOv3。实验结果表明,

AM-YOLOv3 针对铝型材表面缺陷的识别精度可以达到 99.05%。所提模型具有高检测精度优势且检测速度满足实时性需求,为在工业中检测铝型材表面缺陷提供了新思路。

参 考 文 献

- [1] Tang B, Kong J Y, Wu S Q. Review of surface defect detection based on machine vision[J]. Journal of Image and Graphics, 2017, 22(12): 1640-1663. 汤勃, 孔建益, 伍世德. 机器视觉表面缺陷检测综述[J]. 中国图象图形学报, 2017, 22(12): 1640-1663.
- [2] Li C X. Research on surface defect detection algorithm of aluminum profile based on machine learning[D]. Guangzhou: South China University of Technology, 2020: 21-34. 李超贤. 基于机器学习的铝型材表面缺陷检测算法研究[D]. 广州: 华南理工大学, 2020: 21-34.
- [3] Tao X, Hou W, Xu D. A survey of surface defect detection methods based on deep learning[J]. Acta Automatica Sinica, 2021, 47(5): 1017-1034. 陶显, 侯伟, 徐德. 基于深度学习的表面缺陷检测方

- 法综述[J]. 自动化学报, 2021, 47(5): 1017-1034.
- [4] Tabernik D, Šela S, Skvarč J, et al. Segmentation-based deep-learning approach for surface-defect detection[J]. *Journal of Intelligent Manufacturing*, 2020, 31(3): 759-776.
- [5] Yan W X. Research on deep learning and its application on the casting defects automatic detection [D]. Guangzhou: South China University of Technology, 2016: 23-37.
颜伟鑫. 深度学习及其在工件缺陷自动检测中的应用研究[D]. 广州: 华南理工大学, 2016: 23-37.
- [6] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [7] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [8] Ren S Q, He K M, 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.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [10] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788. [LinkOut]
- [11] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2021-03-20]. <https://arxiv.org/abs/1804.02767>.
- [12] Bochkovskiy A, Wang C Y, Liao H YMark. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. (2020-04-23) [2021-03-20]. <https://arxiv.org/abs/2004.10934>.
- [13] Glenn J, Alex S, Jirka B, et al. ultralytics: yolov5 [DB/OL]. (2020-08-13) [2021-03-20]. <http://doi.org/10.5281/zenodo.3983579>.
- [14] Law H, Deng J. CornerNet: detecting objects as paired keypoints [J]. *International Journal of Computer Vision*, 2020, 128(3): 642-656.
- [15] Zhou X Y, Wang D Q, Krähenbühl P. Objects as points [EB/OL]. (2019-04-16) [2021-03-20]. <https://arxiv.org/abs/1904.07850>.
- [16] Tian Z, Shen C H, Chen H, et al. FCOS: fully convolutional one-stage object detection [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea. New York: IEEE Press, 2019: 9626-9635.
- [17] Li D J, Li R H. Mug defect detection method based on improved Faster RCNN[J]. *Laser & Optoelectronics Progress*, 2020, 57(4): 041515.
李东洁, 李若昊. 基于改进 Faster RCNN 的马克杯缺陷检测方法[J]. *激光与光电子学进展*, 2020, 57(4): 041515.
- [18] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [19] Li Y T, Huang H S, Xie Q S, et al. Research on a surface defect detection algorithm based on MobileNet-SSD[J]. *Applied Sciences*, 2018, 8(9): 1678.
- [20] Wen S P, Li C X. Surface defect detection of aluminum profile based on Gaussian-Yolov3 [J]. *Computer Measurement & Control*, 2020, 28(9): 88-93.
文生平, 李超贤. 基于 Gaussian-yolov3 的铝型材表面缺陷检测[J]. *计算机测量与控制*, 2020, 28(9): 88-93.
- [21] Zhou T Y, Zhu Q B, Huang M, et al. Defect detection of chip on carrier based on improved YOLOV3 [J]. *Laser & Optoelectronics Progress*, 2021, 58(12): 1210002.
周天宇, 朱启兵, 黄敏, 等. 基于改进 YOLOV3 的载波芯片缺陷检测[J]. *激光与光电子学进展*, 2021, 58(12): 1210002.
- [22] Wang X, Kang S, Zhu W D, et al. Defect detection of laminated surface in the automated fiber placement process based on improved CenterNet [J]. *Infrared and Laser Engineering*, 2021, 50(10): 20210011.
王璇, 康硕, 朱伟东, 等. 基于改进 CenterNet 的 AFP 铺层表面缺陷检测[J]. *红外与激光工程*, 2021, 50(10): 20210011.
- [23] Hu J, Shen L, Albanie S, et al. Squeeze-and-Excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [24] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [EB/OL]. (2015-11-23) [2021-03-25]. <https://arxiv.org/abs/1511.07122>.

- [25] Shen X H, Li Z H, Li M, et al. Aluminum surface-defect detection based on multi-task deep learning[J]. *Laser & Optoelectronics Progress*, 2020, 57(10): 101501.
沈晓海, 栗泽昊, 李敏, 等. 基于多任务深度学习的铝材表面缺陷检测[J]. *激光与光电子学进展*, 2020, 57(10): 101501.
- [26] Shuzhilian ai. Tianchi aluminum profile surface defect recognition competition [EB/OL]. [2021-03-25]. <https://tianchi.aliyun.com/course/live/1332>.
- [27] Wu Y, Chen Y P, Yuan L, et al. Rethinking classification and localization for object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 10183-10192.
- [28] Ghiasi G, Lin T Y, Le Q V, et al. NAS-FPN: learning scalable feature pyramid architecture for object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020, 7029-7038.
- [29] Dong Z W, Li G X, Liao Y, et al. CentripetalNet: pursuing high-quality keypoint pairs for object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 10516-10525.