

基于改进 Faster R-CNN 的面板缺陷检测算法

陈婉琴¹, 唐清善¹, 黄涛²

(1.长沙理工大学 物理与电子科学学院, 湖南 长沙 410000; 2.中国人民解放军第 3303 工厂, 湖北 武汉 430200)

摘要: 针对面板表面缺陷检测存在精度低、效率低等问题, 提出了一种基于 Faster R-CNN 的优化缺陷检测的算法。该方法通过在特征融合层添加不降维的局部自适应跨通道卷积, 以增加通道交叉的特征映射; 且在骨干特征提取网络后加入 CBAM 注意力网络, 从而捕获特征图的长期特征依赖关系。并分析了缺陷数据集样本宽高比的差异性, 设定锚框生成大小, 结合 DIOU-NMS 建议框筛选机制以提升先验框与目标框的匹配率。实验结果表明, 优化后网络模型的精确率与识别率均得到很大提升。

关键词: 面板缺陷; Faster R-CNN; 目标检测; MobileNetv2

中图分类号: TP301

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.211404

中文引用格式: 陈婉琴, 唐清善, 黄涛. 基于改进 Faster R-CNN 的面板缺陷检测算法[J]. 电子技术应用, 2022, 48(1): 133-137.

英文引用格式: Chen Wanqin, Tang Qingshan, Huang Tao. Panel defect detection algorithm based on improved Faster R-CNN[J]. Application of Electronic Technique, 2022, 48(1): 133-137.

Panel defect detection algorithm based on improved Faster R-CNN

Chen Wanqin¹, Tang Qingshan¹, Huang Tao²

(1.School of Physics and Electronic Science, Changsha University of Science and Technology, Changsha 410000, China;

2.No.3303 Factory of PLA, Wuhan 430200, China)

Abstract: In view of the low precision and low efficiency of panel surface defect detection, this paper proposes an optimized defect detection algorithm based on Faster R-CNN. This method adds local adaptive cross-channel convolution without dimensionality reduction in the feature fusion layer to increase the feature mapping of channel crossing, and adds the CBAM attention network after the backbone feature extraction network to capture the long-term feature dependency of the feature map. It also analyzes the difference in the aspect ratio of the defect data set, sets the generation size of the aiming window, and combines the DIOU-NMS suggestion frame screening mechanism to improve the matching rate of the prior frame and the target frame. Experimental results show that the accuracy and recognition rate of the optimized network model have been greatly improved.

Key words: panel defect; Faster R-CNN; target detection; MobileNetv2

0 引言

在金属板材生产过程中, 由于加工技术、温度控制、杂物掺入^[1]等影响, 金属面板表面会产生各种缺陷, 如划痕、凸粉等。其不仅影响产品的美观和舒适度, 而且会在使用过程中留下潜在隐患。因此, 表面缺陷检测环节对于保障产品的质量非常重要。

人们在缺陷领域的探索分为基于传统目标检测^[2]和基于深度学习^[3]的方法。周神特等人^[4]利用 SIFT 算子提取缺陷特征向量后提出了一种 BP 算法级联 SVM 结合的分类器检测方法检测金属板材表面缺陷, 该方法对缺陷特征要求辨识度高, 对于与背景相似的缺陷无法得到有效处理。李兰等人^[5]提出一种基于空洞卷积融合的 SSD 工件表面缺陷检测方法, 有效地对工件表面的剥落、碎屑、梨沟缺陷进行识别, 但是其缺陷样本是利用电

子显微镜获取, 对缺陷样本的分辨率要求很高。李维刚等人^[6]提出一种基于 K-means 聚类改进的 YOLOv3 算法对带钢表面缺陷进行检测, 较未改进之前在整体上提升检测速度的和缺陷的检测率, 但是其使用的 K-means 对初始设置条件极其敏感, 数据量不够时无法保证结果的有效性。王海云等人^[7]提出了一种利用 FPN 改进 Mask R-CNN 算法来检测工业表面缺陷, 虽然能够准确地定位和识别缺陷, 但是其利用多尺度特征映射融合的方法增加了网络复杂度, 导致其训练速度慢, 识别耗时长。综上所述, SSD 算法^[8]及其变体的一阶检测器^[9]更省时, 在目标检测跟踪方面更具备适用性, 但论检测性能而言, 二阶检测器(如 R-CNN^[10-11])及其变体等在公共的评价基准上取得的检测精度更优。因此, 本文提出了基于改进 Faster RCNN 算法的缺陷检测方法, 提高了检测精度,

并且具有良好的识别速率,适用性更强。

本文实验采用划痕、脏点、凸粉、漆泡四种表面缺陷组成的数据集。从图1中可以看出,划痕的面积占比差异大,分布不规则,深浅不一;而脏点面积占比极小;漆泡面积长宽比极端分布;凸粉缺陷在金属板材表面与板材背景几乎一致。这些缺陷的特征给缺陷的检测带来了很大的困难,因此,选择一种适用于面板缺陷检测的检测方法具有挑战性。

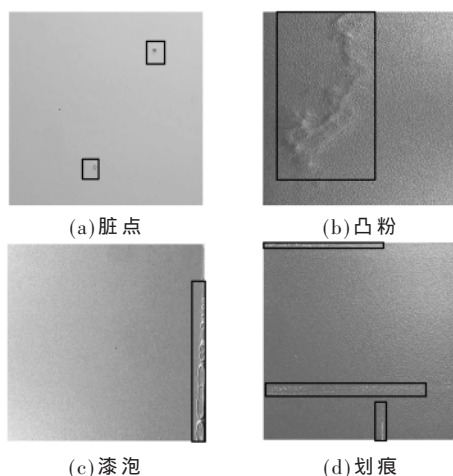


图1 4种缺陷示意图

1 Faster RCNN 算法原理

Faster R-CNN^[12]是基于VGG16骨干特征提取网络组成的二阶目标检测算法,其检测流程如图2所示。

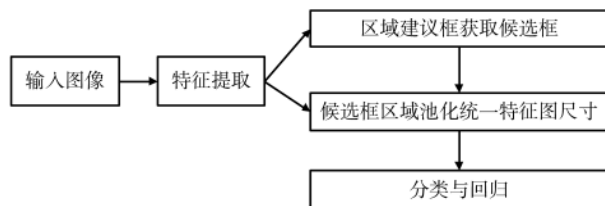


图2 Faster R-CNN 检测流程图

2 算法设计与改进

传统的Faster R-CNN采用多层非线性网络增加网络深度来提取更复杂更具有表达性的特征。尽管其在

COCO数据集^[13]上测试效果极佳,但因为梯度是从后向前传播的,增加网络深度后,模型训练会面临网络退化问题,而且提升精确度是以牺牲时间为代价。由于实验环境限制,本文所采集的缺陷数据图片量较少,同时缺陷目标大小存在巨大差异,为了兼顾缺陷识别准确性和实时性,单纯的学习迁移并不能完全满足缺陷检测的要求。为了解决上述问题,本文在原始的Faster RCNN基础上,主要做了两个方面的改进。首先在Mobilenetv2网络的基础上提出了一种带通道交叉特征映射的残差网络,以此来丰富其特征提取网络对低维特征的敏感性;其次采用新的建议框生成机制和候选框筛选策略提升候选框与目标框的匹配性。改进后的Faster R-CNN检测网络框架如图3所示。

2.1 网络层的改进

本文采用轻量级的Mobilenet网络^[14]来提取缺陷特征。Mobilenet采用可分解的深度可分离卷积,其可以降低模型计算复杂度,并能够极大简化降低模型,满足在移动端部署的应用场景。但其存在检测精确度不高的问题。为了满足在不降低模型计算复杂度的情况下提升整个网络性能的精确度,本文首先基于反残差块结构提出在不减少通道数目的情况下增加通道交叉来平衡通道数目和连通性的新残差网络结构,然后在每个网络块连接之间级联注意力机制,有效集中重点特征信息。优化后的反残差卷积块结构如图4所示。

在模型训练过程中,通过 1×1 的点卷积来扩展维度,进而用 3×3 深度可分离卷积提取特征,最后使用 1×1 点卷积进行数据通道压缩。考虑到使用深度可分离卷积使得网络通道数增加的同时会出现卷积层的连通性降低的问题,优化后的残差模块保留了shortcut结构,以有效避免深度网络中的梯度弥散;卷积核大小为 3×3 的深度可分离卷积分解为两个 3×1 和 1×3 的向量,以减少参数量;其次通过加强每个通道及其 k 个邻域的交叉性,以有效平衡网络通道数与连通性的不足,其中 k 表示本地通道交叉的覆盖范围,卷积核 k 值的确定与每个网络块的输出通道数的大小有关。

为了捕获长期依赖关系,本文在骨干特征提取网络后级联CBAM注意力网络^[15],其原理如式(1)所示:

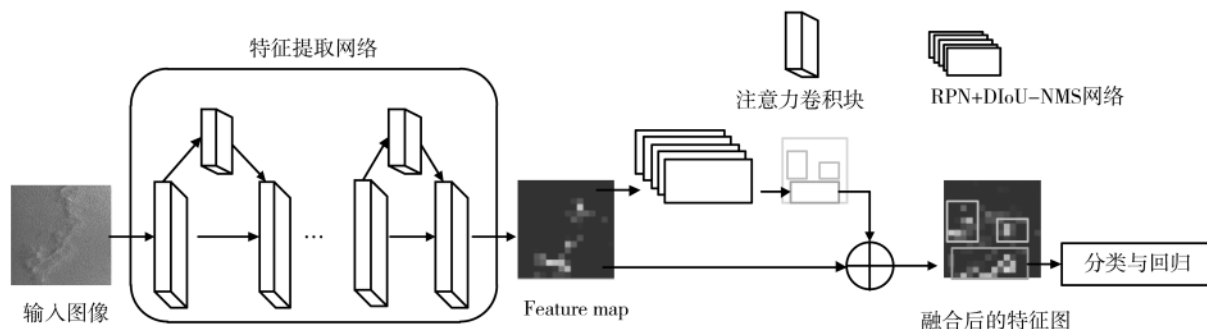


图3 改进后的Faster R-CNN检测网络框

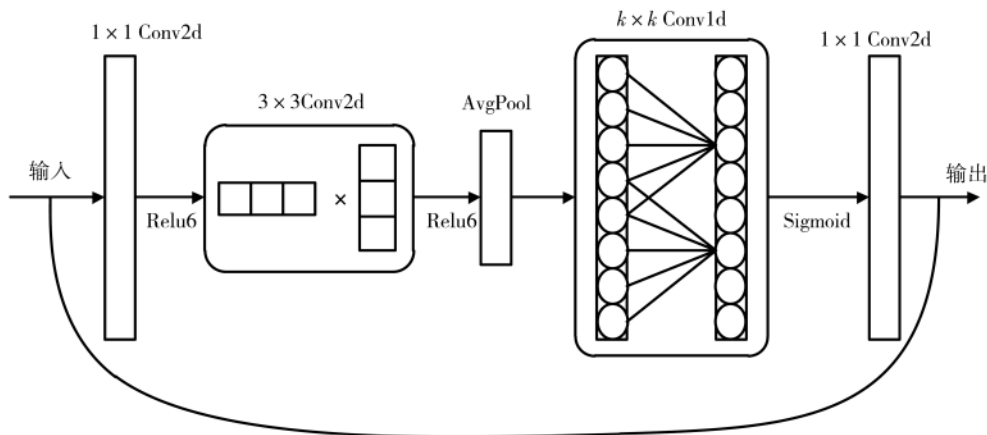


图4 改进型 Mobilenetv2 网络反残差卷积块结构图

$$F'' = M_s(M_c(F) \otimes F) \otimes (M_c(F) \otimes F) \quad (1)$$

其中, $\otimes$ 为逐点乘积, F 为输入特征映射图; M_c 表示在通道维度上做注意力提取的操作, 原理如式(2)所示; M_s 表示在空间维度上做注意力提取操作, 其原理如式(3)所示。

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (2)$$

$$M_s(F) = \sigma(f^{7 \times 7}[\text{AvgPool}(F); \text{MaxPool}(F)]) \quad (3)$$

CBAM 注意力网络由通道注意力和空间注意力两个模块组成, 其中通道注意力模块用于计算特征通道之间的相互依赖性, 学习选择性地增强有意义特征。空间注意力模块用于计算图像上任何两个位置之间的相互作用, 有效地捕获大跨度和大面积缺陷的重要信息^[15]。

2.2 增强型区域推荐生成网络

区域建议框网络中锚框的尺寸设置直接影响候选框的生成, 进而影响后续目标检测预测结果的真实框大小。本文缺陷数据集缺陷尺寸大小差异很大, 若使用原始的锚框尺寸, 可能导致小目标漏检率高, 影响检测效果。

数据集缺陷区域宽高比统计如图5所示。从图5可以看出, 宽高比为1时缺陷数量最大, 综合考虑整体缺陷大小分布数量情况, 选取其宽高比的平均数, 中位数和众数作为锚框尺寸生成比例的考量因素。通过调整锚框的大小及生成比例, 促进候选框更加贴合实际缺陷大小要求, 使得识别更加准确。在统计长宽比时, 剔除极端比例, 设定锚框生成比例为{0.1, 0.2, 0.5, 1.0, 2.0, 5.0, 10.0, 20.0}, 以贴合缺陷大小生成建议框。RPN 网络流程示意图如图6所示。

实际中, 缺陷的产生是随机的, 因此可能同时产生多种缺陷。此时提出的候选框会产生堆叠, 在对候选框进行筛选时会去除得分较低的候选框。如果只考虑重叠区占比的非极大值机制(NMS)方法筛选候选框, 缺陷边缘重合或被其他不规则缺陷包围的小缺陷将被过滤掉, 导致漏检。针对此问题, 本文综合考量重叠率和框选中心矩比例作为筛选条件, 选用 DIoU-NMS 方法去除多余的候选框^[16]。NMS 与 DIoU-NMS 的筛选原理对比如图7所示。

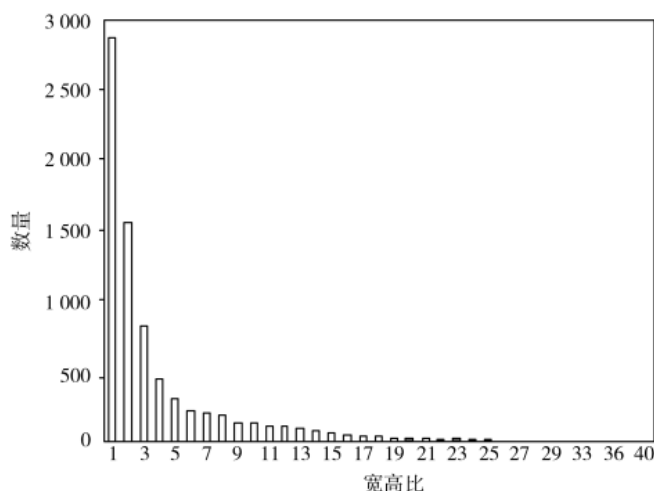


图5 缺陷区域宽高比

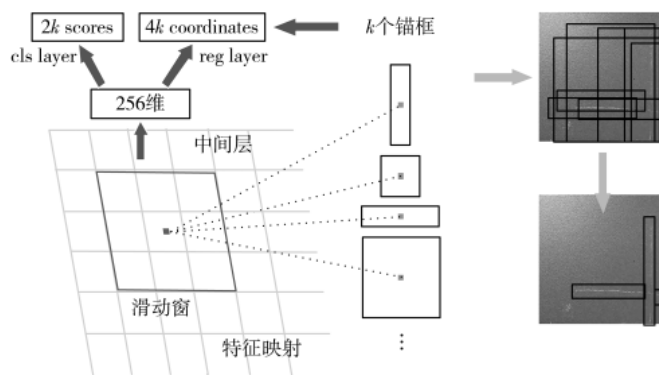


图6 RPN 网络流程示意图

IoU 的计算如下所示:

$$\text{IoU} = \frac{|B \cap B^d|}{|B \cup B^d|} \quad (4)$$

其中, B 表示预测框, B^d 代表真实框。

DIoU-NMS 的计算如式(5)所示:

$$S_i = \begin{cases} S_i, \text{IoU} - R_{\text{DIoU}}(M, B_i) < \varepsilon \\ 0, \text{IoU} - R_{\text{DIoU}}(M, B_i) \geq \varepsilon \end{cases} \quad (5)$$

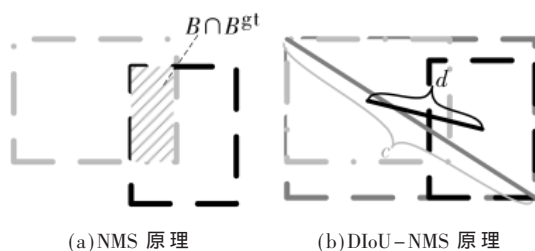


图7 NMS与DIoU-NMS原理对比图

其中, M 指得分最高的预测框, B_i 为考虑重叠区和中心点距离后被删除的候选框, ε 为 DIoU-NMS 中设定的区分正负类的阈值。 S_i 为分类得分, 其惩罚项定义如下:

$$R_{\text{DIoU}} = \frac{\rho^2(b, b^{\text{gt}})}{c^2} \quad (6)$$

其中, (b, b^{gt}) 为预测框和真实框的中心点, ρ^2 指最小化两个边界框中心点之间的标准化距离, c^2 为覆盖两个盒子的最小封闭盒子的对角线标准化长度。

3 实验验证与分析

3.1 金属面板表面缺陷检测数据集

本文使用的面板表面缺陷集采集于机箱钢板表面, 为丰富缺陷种类, 凸粉、漆泡缺陷图片来自于天池比赛提供的铝板缺陷。统一数据集图片尺寸为 512×512 , 所有标签数据均由作者标注完成。本文在实验中使用 COCO 评价准则^[13]衡量整个模型的性能。原始数据集分为脏点、凸粉、划痕、漆泡 4 种共 1 400 张图片, 其各类图片占比约为 3:3:4:4。由于原始数据集的数据量较少, 且不同缺陷的数据数量不均衡, 本文通过数据增强来增加样本的多样性并同时标记框进行相应调整, 数据增强后共得到 5 600 幅图像, 主要增强方式有随机旋转、裁剪、随机颜色和亮度增强。实验过程中对数据集进行随机分割, 训练集和验证集的分割比例为 7:3。

3.2 算法性能分析

本实验所使用开发环境为 Python3.7, GPU 型号为 RTX 2080Ti, 整体利用了开源的 PyTorch 框架进行代码编译, 框架配置为 PyTorch1.6.0 和 CUDA10.2。标注工具采用 LabelImg。

在训练过程中, 利用模型迁移方法固定部分卷积网络的部分参数, 使用随机梯度下降法在每一次迭代过程中进行参数更新。本文算法在训练过程中的损失函数曲线如图 8 所示。从训练的 Train_loss 值可以看出在迭代 6 000 次后, 网络模型损失值趋于稳定。

本文将所提算法与其他算法在相同的数据集上进行了对比实验。在相同的训练 epoch 下, 通过改变骨干特征提取网络及 RPN 网络, 通过迁移学习的方法, 得到不同算法的实例分割评价指标和缺陷平均识别时间, 其对比如表 1 所示, 其 mAP 值对比如图 9 所示。

图 9 表明, 改进后的 Faster R-CNN 网络的 mAP 高于其他网络, 其网络的收敛速度也提升了约 1 个 epoch。

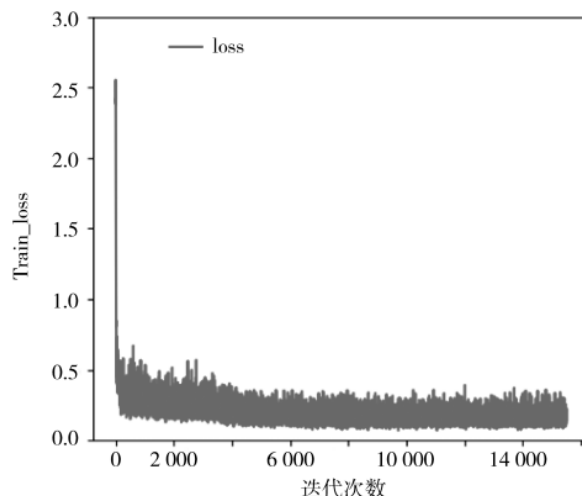


图8 改进后的 Faster R-CNN 算法训练损失下降曲线

表1 改进前后算法性能对比分析

算法	mAP/%		平均识别时间/s
	数据集(原始)	数据集(增强)	
SSD	32.3	56.3	1.9
Faster R-CNN	48.1	76.2	3.5
改进 Faster R-CNN	54.6	78.9	2.5

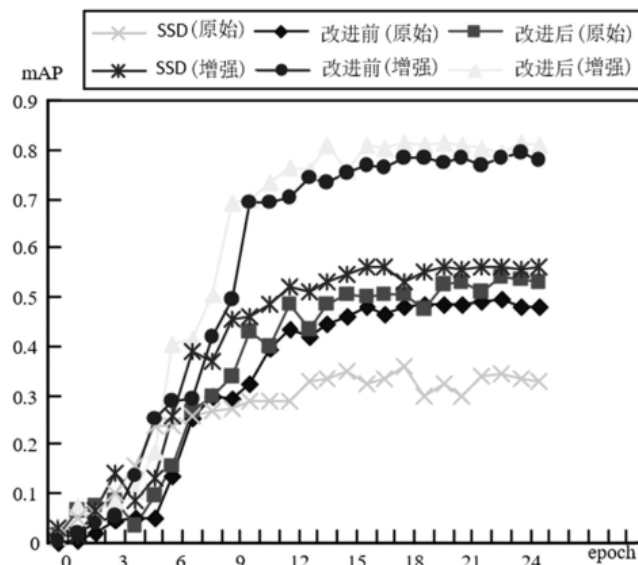


图9 在相同训练 epoch 下 mAP 值对比结果

从表 1 可以看出, 在原始数据集条件下, 本文所优化后网络的 mAP 比未改进前提升了约 6.1%。通过数据增强后, 改进后的网络 mAP 达到 78.9%, 较之前提升了约 2.7%。从检测速度分析, 本文网络模型每张缺陷检测的平均时间是 2.5 s, 低于未改进前的检测时间约 1 s, 其检测速度基本可以与一阶检测算法 SSD 相媲美。缺陷识别结果如图 10 所示, 从图 10 中可以看出, 采用优化后网络训练结果进行缺陷预测时, 4 种缺陷都能够被识别; 而且从划痕的检测结果中可以看出, 对于一些人眼无法

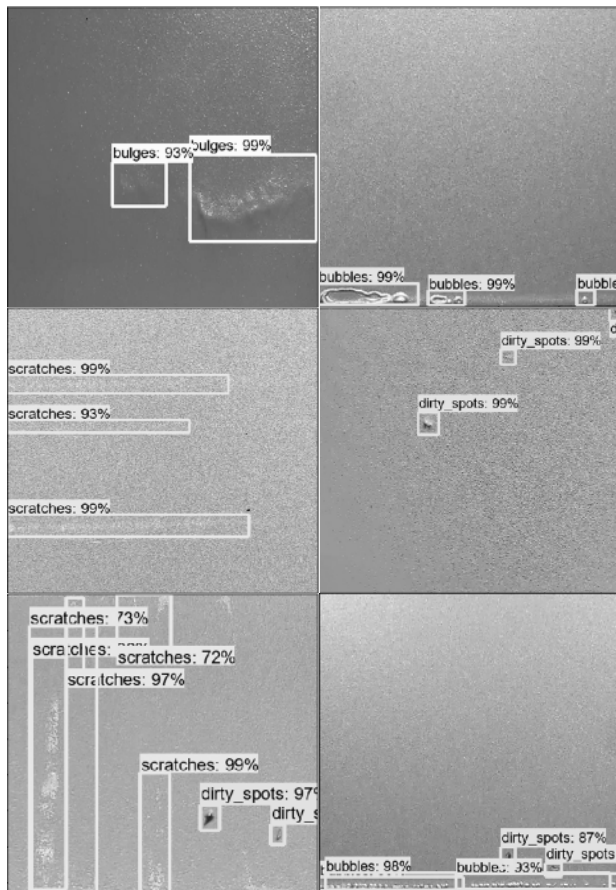


图10 缺陷检测结果图

清晰识别的缺陷,通过该算法也能达到75%以上的识别置信度。因此,本文采取的改进Faster R-CNN网络算法在此数据集上检测效果较好。

4 结论

本文以Faster R-CNN算法为理论基础,结合目标检测领域在工业生产上的应用,针对机箱钢板缺陷,提出了一种基于Faster R-CNN的优化算法。该算法针对于工厂应用环境,采用深度学习算法识别缺陷代替人工检测,在保证识别率高的同时减少成本并提高检测效率。实验结果表明,本文所提算法具有较高的检测精度,同时检测速度得到提升。

由于实验条件有限,数据集样本多样性较少,通过数据增强手段对数据集缺陷样本多样性提升有限,对网络训练阶段缺陷的分割效果还有提升的空间。下一步将采集更多的缺陷样本,探索更加高效提升模型训练速度的方法,进一步增强模型的泛化能力。

参考文献

- [1] 张刘赞.基于机器视觉的手机金属板表面缺陷检测技术研究[D].杭州:浙江大学,2018.
- [2] NG P, HENIKOFF S. SIFT: predicting amino acid changes that affect protein function[J]. Nucleic Acids Research, 2003, 31(13): 3812-3814.

- [3] NOH H, HONG S, HAN B. Learning deconvolution network for semantic segmentation[C]//IEEE International Conference on Computer Vision (ICCV), 2015.
- [4] 周神特, 王宇宇, 张潇, 等. 基于机器视觉的金属板材表面缺陷光学检测技术[J]. 无损检测, 2020, 42(9): 39-44.
- [5] 李兰, 奚舒舒, 张才宝, 等. 基于改进SSD模型的工件表面缺陷识别算法[J]. 计算机工程与科学, 2020, 42(9): 1608-1615.
- [6] 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.
- [7] 王海云, 王剑平, 张果, 等. 改进FPN的Mask R-CNN工业表面缺陷检测[J]. 制造业自动化, 2020, 42(12): 35-40, 97.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//European Conference on Computer Vision. Springer International Publishing, 2016.
- [9] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. Computer Vision and Pattern Recognition, 2018, 36(5): 1658-1672.
- [10] ZHANG N, DONAHUE J, GIRSHICK R, et al. Part-based R-CNNs for fine-grained category detection[C]//Springer International Publishing. Springer International Publishing, 2014.
- [11] KAIMING H, GEORGIA G, PIOTR D, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 26(2): 32-46.
- [12] 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 & Machine Intelligence, 2017, 39(6): 1137-1149.
- [13] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]//European Conference on Computer Vision. Springer International Publishing, 2014.
- [14] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision (ECCV), 2018.
- [16] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//AAAI Conference on Artificial Intelligence, 2020.

(收稿日期: 2021-02-20)

作者简介:

陈婉琴(1997-), 女, 硕士研究生, 主要研究方向: 模式识别、目标检测。

唐清善(1977-), 男, 博士, 讲师, 硕士生导师, 主要研究方向: 模式识别、数字信号处理。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所