DOI: 10.3969/J.ISSN.1000-5137.2022.02.001

基于改进 U-Net 的零件缺陷分割标注

金文倩,朱媛媛*,王笑梅

(上海师范大学信息与机电工程学院,上海201418)

摘 要:提出一种以U-Net为基础,依据零件缺陷的特点对网络进行一系列改进的模型,以提升网络对零件缺陷的分割精度.首先在U-Net结构中的编码阶段,使用改进的残差网络 Res2Net提高该阶段的特征提取能力;然后在网络编码器与解码器的中间部位增加空洞卷积, 在不改变特征图尺寸的情况下增加感受野,降低误检率与漏检率;最后在U-Net的输出阶段与 Mini U-Net进行结合,对原本的输出结果进行二次补丁,提高对微小缺陷的检测精度.实验结 果表明,对MVTec数据集进行分割的F1-Score分数达到87.21%,时间为0.017 s,达到了良好的 检测效果.

关键词:图像分割;缺陷检测;U-Net; Res2Net; 空洞卷积

中图分类号: TP 391 文献标志码: A 文章编号: 1000-5137(2022)02-0129-06

Part defect segmentation and annotation based on improved U-Net

JIN Wenqian, ZHU Yuanyuan^{*}, WANG Xiaomei

(College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 201418, China)

Abstract: A model based on U-Net network and a series of improvements to it according to the characteristics of part defects were proposed to improve the segmentation accuracy of part defects. Firstly, the improved residual network Res2Net was used in the coding stage of U-Net network structure to improve the feature extraction ability during this stage. Secondly, the hole convolution was added in the middle of the network encoder and decoder, and the receptive field was increased without changing the size of the characteristic image, so as to reduce the false detection rate and omission detection rate. Finally, in the output stage of U-Net, Mini U-Net was combined with to patch the original output results, so as to improve the detection accuracy of small defects. The experimental results showed that the F1-Score of MVtec dataset segmentation reached 87.21% and the time was 0.017 s, with which outstanding detection effect could be achieved.

基金项目:上海市自然科学基金(17ZR1419800)

作者简介:金文倩(1995—),女,硕士研究生,主要从事基于深度学习的目标检测方面的研究.E-mail:jwq3310528@ 163.com

*通信作者:朱媛媛(1971一),女,副教授,主要从事新型材料与结构力学特性分析的数学建模、计算方法与工程应用方面的研究.E-mail: zhuyuanyuan@shnu.edu.cn

引用格式:金文倩,朱媛媛,王笑梅.基于改进U-Net的零件缺陷分割标注 [J].上海师范大学学报(自然科学版), 2022,51(2):129-134.

Citation format: JIN W Q, ZHU Y Y, WANG X M. Part defect segmentation and annotation based on improved U-Net [J]. Journal of Shanghai Normal University(Natural Sciences), 2022, 51(2):129-134.

收稿日期: 2022-01-04

Key words: image segmentation; defect detection; U-Net; Res2Net; void convolution

0 引 言

将深度学习和缺陷检测问题相结合是当下十分热门的研究领域.CHEN等^[1]结合 single shot multibox detector(SSD)网络及 you only look once(YOLO)算法构建了一个由粗到细的级联检测网络,对 高铁线路紧固件进行缺陷检测;CHA等^[2]通过 faster region convolutional neural networks(FasterRCNN)检 测混凝土裂缝,中、高两级钢腐蚀,螺栓腐蚀和钢筋分层 5种损伤类型;YU等^[3]提出一种基于YOLOv4的 空心杯电枢表面空洞缺陷检测的方法,解决空心杯电枢表面微小缺陷检测过程中检测精度低、速度慢以 及不能实时检测等问题.上述检测算法对于缺陷仅能进行笼统的方框标注,无法精确到图像的具体像 素.本文作者将图像分割应用到目标缺陷检测领域,将图像中的像素分割出来,对应到不同类别中.首先 在U-Net结构的^[4]编码阶段使用改进的残差网络Res2Net^[5]提高特征提取能力,使用更细粒度进行特征 提取.在网络编码器与解码器的中间部位增加空洞卷积^[6],通过改变扩张率来扩大感受野,在不改变特 征图尺寸的情况下获取不同尺度的特征信息,可以避免下采样所造成的细节信息丢失等问题,降低误检 率与漏检率.最后在U-Net的输出阶段,将之与Mini U-Net^[7]进行结合,对原本的输出结果进行二次补 丁,提高对微小缺陷的检测精度.

1 改进的U-Net模型

1.1 模型结构

对U-Net模型编码器-解码器的基本结构进行改进(图1).



图1 改进的U-Net模型结构

在模型的编码器阶段,将原先的普通卷积结构替换为Res2Net模型^[8],使用更细的粒度获取多尺度特征,扩大图像感受野,增加图像分割的精度.

编码器与解码器之间连接空洞卷积结构,进一步扩大感受野,获得不同尺度的特征图像.如图1所示,先串联空洞率分别为1,2,4的3个空洞卷积,每层的感受野分别为3,7,15^[9].U-Net的编码器部分有4个下采样层,在最后一个特征图上融合了第一个特征图上的特征信息,并以此覆盖第一个特征图上的特征信息.

Mini U-Net 是由 U-Net 的中间部分组成的.最后将 U-Net 的输出部分连接一个 Mini U-Net 对 U-Net 的输出结果作二次补丁,避免零件因存在微小缺陷而导致的模糊和细节丢失等情况.

1.2 Res2Net

第2期

Res2Net结构图如图2所示.



图2 Res2Net结构图

由图2可知,经过1×1的卷积后,特征图被均匀地分成了s个特征图子集,用 X_i 表示,其中 $i \in \{1, 2, \dots, s\}$ (图2中s=4),s即为比例尺寸的控制参数,也就是输入特征通道数.s越大,模型表现多尺度 的能力越强.除了 X_1 之外,每个 X_i 都对应一个 3×3 的卷积,用 K_i (·)来表示,用 y_i 来表示 K_i (·)的输出.特征 子集 X_i 与 K_{i-1} (·)的输出相加,再汇入到 K_i (·)中.模型为了在增加s的同时减小参数,忽略了 X_1 的 3×3 卷 积. y_i 的计算方式如下:

$$y_{i} = \begin{cases} X_{i}, & i = 1 \\ K_{i}(X_{i}), & i = 2 \\ K_{i}(X_{i} + y_{i-1}), & 2 < i \le s \end{cases}$$
(1)

如图2中s=4的情况,

$$y_1 = X_1, \ y_2 = X_2 \times (3 \times 3) = K_2, \ y_3 = [X_3 + X_2 \times (3 \times 3)] \times (3 \times 3) = (K_2 + X_3) \times (3 \times 3) = K_3,$$

$$y_4 = \{X_4 + [X_3 + X_2 \times (3 \times 3)] \times (3 \times 3)\} \times (3 \times 3) = (K_3 + X_4) \times (3 \times 3) = K_4.$$
(2)

由此可以得到不同数量以及不同感受野大小的输出,再经过一个1×1的卷积,将输出进行融合.这种先拆分再融合的策略能够提高特征处理的效率.

1.3 空洞卷积

增加网络特征点的感受野就意味着要增加接触到的图像范围,这样能获得更多语义层次及更好的特征,感受野越大,包含的特征更加趋于全局;反之,包含的特征会趋于局部.由Visual Geometry Group (VGG)网络可知,1个7×7卷积层的正则化等效于3个3×3卷积层的叠加,通过多个小的卷积层的叠加不 仅可以大幅度地减少运算参数,还能具备同样的正则化效果,减少了过拟合的可能性,更好地缓解了U-Net在缩小放大过程所产生的特征图细节丢失和精度下降的问题.

感受野的计算公式如下:

$$R_{i+1} = R_i + (k'-1) \times S_i, \tag{3}$$

2022年

其中,*R*_{*i*+1}表示当前层的感受野;*R*_{*i*}表示上一层的感受野;*k*'表示等效卷积核的大小;*S*_{*i*}表示除了本层外之前所有层的步长的乘积.

等效卷积核 k'的计算方式为:

$$k' = k + (k-1) \times (d-1), \tag{4}$$

其中,k表示卷积核的大小;d表示扩展率,即空洞数.

图3为空洞卷积一维图.由图3可以看出,空洞卷积在增加感受野的同时,不会改变图像输出特征图的尺寸.从图3(b)和图3(c)可知,使用了空洞卷积之后的输出并未发生变化,影响输出特征图尺寸大小的因素只有输入特征图尺寸、卷积核大小、padding和步长,与是否使用空洞卷积以及空洞数的多少无关,所以使用空洞卷积增大感受野不影响输出特征图的尺寸大小.



2 实验与分析

2.1 消融实验

为了验证本研究方法的有效性,设计了消融实验检验Res2Net、空洞卷积和MiniU-Net分别对零件 图像分割标注的影响效果,如表1所示.

模型	准确率/%	召回率/%	F1-Score/%	
U-Net	81.8	82.7	82.2	
Res2Net	84.3	83.8	84.1	
Res2Net+空洞卷积	86.2	87.6	86.9	
Res2Net+空洞卷积+ Mini U-Net	88.5	90.6	89.1	

表1 基于U-Net模型不同分割方法性能比较

由表1可知:原始的U-Net模型在零件数据集上进行分割标注实验的分割效果尚佳,但准确率 较低;将原先的卷积结构改进为Res2Net之后,各项性能都有了一定程度的提升,但是召回率的提升 幅度较小;在此基础上继续增加空洞卷积提高感受野,模型的性能得到进一步的提高;再加入 MiniU-Net进行第二次补丁训练,3个指标都有明显的提升.从原始U-Net模型到最终的模型,每增 加一个改进模块,模型的性能都有不同程度的提升,体现了模型改进的有效性.

2.2 基于 MVTec 螺丝数据集异常区域分割

图 4 为使用改进的 U-Net 模型对 MVTec 螺丝数据集^[10]进行异常区域分割的结果.数据集共有 mainpulated_front 螺丝前端、scratch_head 顶部划痕、scratch_neck 颈部划痕、thread_side 螺纹侧面 4 个部 位的缺陷.图4展示了数据集中的ground_truth标签图,可以看出改进的 U-Net 模型可以将数据集中的异 常区域分割出来,准确率较高,对于 thread_side 螺纹侧面缺陷存在一定的漏检现象.



表2为对螺丝数据集的检测结果.从表2中可以看出,本研究的模型对于各缺陷检测的平均精度均值(mAP)均在80%以上,有良好的分割效果.其中,对于scratch_head顶部划痕的分割效果最好,达到91.19%,对于thread_side螺纹侧面的分割效果最差.

缺陷类别	准确率/%	召回率/%	mAP/%	
manipulated_front	82.71	88.69	87.38	
scratch_head	89.26	90.80	91.19	
scratch_neck	85.93	83.21	85.20	
thread_side	83.88	79.96	81.45	

表2 螺丝数据集不同缺陷检测效果



3 结 论

通过在U-Net的编码器阶段添加改进的ResNet结构,增加检测粒度;在编码器与解码器中间部位连接空洞卷积,在不改变特征图尺寸大小的情况下,增加模型感受野;在U-Net解码器的输出部分连接MiniU-Net,对网络进行二次补丁,提高微小缺陷的检测精度.实验结果表明,改进后的U-Net具有良好的检测效果,检测速度较快,检测精度较高.

参考文献:

- [1] CHEN J W, LIU Z G, WANG H R, et al. Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network [J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(2):257–269.
- [2] CHA Y, CHOI W, SUH G, et al. Autonomous structural visual inspection using region: based deep learning for detecting multiple damage types [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(9):731-747.
- [3] YU C H, HUANG H S, ZENG F, et al. Real-time detection of surface defects of hollow cup armature base on YOLO_v4
 [J]. Modular Machine Tool & Manufacturing Technique, 2021(6):59–62,66.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Strasbourg: Springer, 2015: 234-241.
- [5] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(2):652-662.
- [6] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [C]// International Conference on Learning Representations. ICLR: San Juan, 2016: arXiv:1511.07122.
- [7] CAI Q H, LIU Y H, ZHANG R F. Two-stage retinal vessel segmentation based on improved U-Net [J]. Laser & Optoelectronics Progress, 2021, 58(16):480-490.
- [8] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016:770-778.
- [9] QU Z, CHEN W. Concrete pavement crack detection based on dilated convolution and multi-features fusion [J/OL].
 Computer Science[2021-12-23]. http://kns.cnki.net/kcms/detail/50.1075.TP.20211011.1608.004.html.
- [10] BERGMANN P, BATZNER K, FAUSER M, et al. The MVTec anomaly detection dataset: a comprehensive real-world dataset for unsupervised anomaly detection [J]. International Journal of Computer Vision, 2021, 129(1):1038-1059.

2022年

(责任编辑:包震宇,顾浩然)