doi: 10.3969/j.issn.1003-3106.2020.10.010

引用格式: 段仲静 李少波 胡建军 等.基于 Mask R-CNN 的胶囊缺陷检测方法[J].无线电工程 2020 50(10):857-862.[DUAN Zhongjing LI Shaobo ,HU Jianjun ,et al. Capsule Defect Detection Method Based on Mask R-CNN [J]. Radio Engineering ,2020 ,50 (10):857-862.]

# 基干 Mask R-CNN 的胶囊缺陷检测方法

段仲静<sup>1</sup>,李少波<sup>12\*</sup>,胡建军<sup>2</sup>,杨 静<sup>2</sup>,王 铮<sup>2</sup> (1. 贵州大学 现代制造技术教育部重点实验室,贵州 贵阳 550025; 2. 贵州大学 机械工程学院,贵州 贵阳 550025)

摘 要:针对传统的空心胶囊缺陷检测多用于人工抽样检测,无论是时间成本还是人力成本都较高,并且带有一定的主观性的问题,研究分析了胶囊缺陷的特点,设计并搭建了胶囊制造缺陷在线检测实验平台,设计并开源了药用空心胶囊凹陷缺陷数据集,提出了一种基于 Mask R-CNN 算法的胶囊缺陷检测识别方法,使用 ResNet 作为特征提取网络,ROIAlign对特征图 ROI 提取是完成像素级的对齐,能很好地解决胶囊缺陷的检测难题。经过实验验证,基于 Mask R-CNN 的胶囊缺陷检测识别的正确率较高。

关键词: 深度学习; Mask R-CNN; ResNet; 胶囊; 缺陷检测

中图分类号: TP3914 文献标志码: A 开放科学(资源服务)标识码(OSID):

文章编号: 1003-3106(2020)10-0857-06



## Capsule Defect Detection Method Based on Mask R-CNN

DUAN Zhongjing<sup>1</sup>, LI Shaobo<sup>1,2\*</sup>, HU Jianjun<sup>2</sup>, YANG Jing<sup>2</sup>, WANG Zheng<sup>2</sup>

(1. Key Laboratory of Advanced Manufacturing Technology of Ministry of Education, Guizhou University, Guiyang 550025, China; 2. School of Mechanical Engineering, Guizhou University, Guiyang 550025, China)

Abstract: Traditional hollow capsule defect detection often adopts the manual sampling detection method both the time cost and the labor cost are high and there are certain subjective problems. The characteristics of capsule defects are studied and analyzed. The experimental platform for online detection of capsule manufacturing defects is designed and built. The data set of medicinal hollow capsule depression defect is designed and opened source. A capsule defect detection and recognition method based on Mask R-CNN algorithm is proposed. The method uses ResNet as the feature extraction network. ROIAlign extracts the feature map ROI to complete pixel-level alignment which can solve the problem of capsule defect detection. After the experimental verification, the correct rate of capsule defect detection based on Mask R-CNN is high.

Key words: deep learning; Mask R-CNN; ResNet; capsule; defect detection

### 0 引言

药用胶囊在现实生活中随处可见。药用空心胶囊主要用来包裹固体药物,对药物颗粒起保护作用,并且解决了药剂难以下咽、口感差等问题。胶囊的缺陷主要是双盖、破片、黑斑和凹陷等[1]。

传统对空心胶囊缺陷检测通常采用基于计算机 视觉的模版匹配算法<sup>[2]</sup>或人工抽样检测。在特定 的图像采集装置上采集胶囊图像后,通过对完好胶囊图像的阈值、梯度值的处理 制作成标准模版。当 缺陷胶囊与标准模版不匹配时,标注图像中不匹配位置,即胶囊存在缺陷的位置。文献[3]提出一种

基于极端学习机来提取胶囊的偏心率和圆形度为特征变量的胶囊缺陷检测方法,检出率为90%。文献[4]对胶囊图像边缘进行了轮廓提取与跟踪,实现了胶囊缺陷905%平均识别率。文献[5]分析了透明胶囊的成像特性,结合阈值分割、腐蚀、轮廓提取及区域填充算法,通过最小包围矩形测量胶囊尺寸来判断胶囊尺寸是否有缺陷,检出率为9362%。本研究针对凹陷的胶囊,制作了胶囊数据集,进行缺

收稿日期: 2020-04-22

基金项目: 国家自然科学基金资助项目(51475097;91746116)

Foundation Item: Project Supported by the National Natural Science

Foundation (51475097; 91746116)

陷识别检测,旨在探索验证新算法的实际应用能力。

### 1 缺陷检测算法设计

### 1.1 Mask R-CNN 图像识别算法

针对由 R-CNN 所代表的两阶段方法中的难题,Mask R-CNN 算法<sup>[6]</sup>在 2017 年被何凯明等人提出。文献指出,因为在进行 ROI 池化和下采样的时候 Faster R-CNN<sup>[7-8]</sup>对 feature map 尺寸都采用了量化操作 会对像素级别的计算精度产生较大影响,故 Mask R-CNN 用 ROIAlign 替换 ROIPooling 以至模型不对图片做量化处理。此外 Mask R-CNN 添加了FCN<sup>[9]</sup>层来做语义掩膜预测,该添加层很好地避免了 feature map 与 input image 上的 ROI 不对准的问题,使掩码正确率上升 10%~50%; Mask R-CNN 还采用过 ResNeXt<sup>[10]</sup>作为基础网络,使 Fast R-CNN 在COCO 数据集上的正确率由 19.7%增加到 398%。

Mask R-CNN 对检测精度和实例分割在现阶段都有相当高的水准,即便一些算法模型的性能在不断优化,但基本维持在同一层次。检测速度慢、很难实现实时性工作需求是 Mask R-CNN 的一大短板。此外,实例分割领域的一大难题是对样本的标注成本过高。

### 12 ResNet 网络结构简介

残差神经网络(ResNet) [11]在 2015 年由微软的何凯明等人提出,获得 ILSVRC 第一名。它完成了152 层超深卷积神经网络的训练,取得很好的效果,并且易于和其他的网络结构相集成。ResNet 在 ImageNet 中定位准确率高出第二名 27%,检测准确率高出第二名 16%,分类错误率仅为 357%;在 COCO分割中准确率高出第二名 12%。ResNet 目前在计算机视觉的5 个重要技术邻域都是最好,也是目前最优的 CNN 模型。

ResNet 的基本模块( 残差学习单元) 如图 1 所示。X 为模型输入 F(x)+X 为模型输出 F(x) 是模型中的数据加法或乘法计算。当 H(x)=F(x)+X 为模型的最优拟合结果时 F(x) 最优为 X 和H(x) 的残差 通过残差拟合的途径来优化网络的性能。转换为残差拟合比传统的卷积网络优越 ,这是由于网络在训练过程中保持残差为零,网络性能也不会因增加残差学习单元而受影响。对多个残差学习单元进行叠加就得到了 ResNet 模型 ,理论上不论叠加多少层都不会影响网络性能。

ResNet 模型的一大贡献是实现了超深网络,避免了对神经网络不断深化而导致正确率达到饱和的

瓶颈问题; 此外 能通过直接把输入和输出相连使残差学习成为整体网络的工作以达到简化学习目标与难度的目的; 另外 ,ResNet 也是具有很好迁移性的网络结构,能很方便地与其他网络结构相集成。

·Residual net

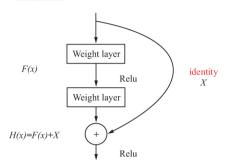


图 1 ResNet 模型的基本模块

Fig.1 Basic modules of ResNet model

#### 13 ROIAlign 层

ROIAlign 是 Mask-RCNN 首次应用的一种区域特征聚合操作,使传统的 ROIPooling 操作中对特征区域 2 次量化引起的区域不匹配问题得到有效处理。实验表明,用 ROIAlign 替换 ROIPooling 可以实现对特征图 ROI 提取,是像素级的对齐,可以提高模型检测正确率。

### 13.1 ROIPooling 的缺陷

在 two-stage 检测模型(如 Fast-RCNN, Faster-RCNN)中,ROIPooling [12]的功能是根据预选框的位置坐标在 feature map 中把相应区域池化为特定大小的 feature map,以便后续的 bbox 回归与分类运算。模型通过回归运算得到浮点数值的候选框位置 特征图经过池化后得到整型尺寸。ROIPooling运算过程会有 2 次量化处理: 一是将候选框的坐标值取整; 二是将取整过的候选框边界区域均分为  $k \times k$  个单元(bin),并对单元边界取整。至此得到的候选框相对最初位置出现了偏差,从而影响分割和检测的正确率。

## 132 ROIAlign 的主要思想和具体方法

针对 ROIPooling 的上述弊端,何凯明等提出了一种改进的 ROIAlign  $^{[13]}$  方法。ROIAlign 的思路很简单: 采用双线性插值方法,得到坐标为浮点数的像素点上的图像值,来替换量化操作,把整个特征汇集过程 转换为不间断的操作。值得关注的是,ROIAlign 不直接填补出候选区域边界上的坐标点,而是有着一套独特的操作流程: 先是遍历全部候选区但对边界坐标值不做取整处理; 接着将候选框边界区域均分为  $k \times k$  个单元(bin) 也不对单元数值做

取整处理;然后采用双线性插值法计算每个单元中固定的 4 个坐标值,最后最大池化操作。实际上ROIAlign的遍历采样点比ROIPooling要少,它却能得到更优的效果,很好地处理了像素不对齐的问题。133 ROIAlign的反向传播

常规的 ROIPooling 反向传播公式为[14]:

$$\frac{\partial L}{\partial x_i} = \sum_r \sum_j \left[ i = i \times (r \, j) \, \right] \frac{\partial L}{\partial y_{rj}} \, , \qquad (1)$$

式中, $x_i$  为池化前特征图上的像素点;  $i \times (r j)$  为点  $y_{ij}$  像素值的来源(  $\max_{pool}$  选出的最大像素值点的位置),第 r 个候选区域的第 j 个点池化后用  $y_{ij}$  表示。由式( 1) 可以看出,在池化过程中只有当某个点的像素值采用了当前点  $x_i$  的像素值( 即  $i = i \times (r j)$  ),才在  $x_i$  处回传梯度。

根据 ROIPooling ,ROIAlign 中对 BP 算法做了优化:  $x_i \times (r,j)$  是 BP 计算得到的采样点浮点型坐标值 在特征图池化前 , $x_i \times (r,j)$  水平和垂直坐标都小于 1 的点都要接受对应点  $y_j$  回传的梯度 ,故 ROIAlign 的法向传播公式:

$$\frac{\partial L}{\partial x_i} = \sum_r \sum_j \left[ d(i \ j \times (r \ j)) \right] < 1 \left[ (1 - \Delta h) (1 - \Delta w) \frac{\partial L}{\partial y_{ij}} \right],$$
(2)

式中  $\mu(\cdot)$  为两点的距离;  $\Delta w$  ,  $\Delta h$  为  $x_i$  与  $x_i$  × (r , j) 水平和垂直坐标的差 .作为双线性插值的系数与原始梯度相乘。

### 2 数据集介绍

#### 2.1 数据采集系统

设计的胶囊制造缺陷在线检测实验平台系统如图 2 所示。组件的胶囊生产表面缺陷在线检测实验平台<sup>[15]</sup> 如图 3 所示。主要包括采集数据的传感器、数据处理器、传送带、光源以及机械支架等构件。采用 32 英寸的工业触摸屏作为输入数据和显示数据的显示屏; 用具备电子滚动快门功能的 MindVision高速工业相机作为图像采集设备 ,因为它能够实时采集快速移动的待检测产品样本。采用树莓派 B3 作为平台的数据处理器; 在箱体内配置可调亮度的带状光源 LED 灯以保证箱体内有合适的光线。另外 ,为了保证获得清晰的图像样本 在相机外部配置了专用环形光源 Microscope LED Ring Light 来对待检测样本补光。平台的传送带由可变速电机驱动 ,为了便于把检测结果可视化 ,在箱体外侧安装了显示器 ,为了减轻数据处理器的运算负荷而配置了戴

尔 GPU1080 显卡工作站。此外 具有无线通信功能的树莓派 B3 数据处理器可以使检测实验平台和工作站之间实现端到端的通信。工作站装载了 SQL Server 2008 R2 数据库 以达到对本地数据实时抓取和对数据自动实时云端存储的目的。

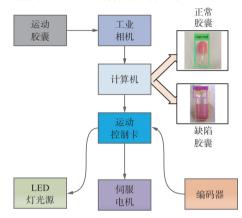


图 2 胶囊制造缺陷检测实验平台系统 g2 Experimental platform system of capsule

Fig2 Experimental platform system of capsule manufacturing defect detection

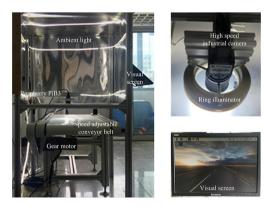


图 3 高速胶囊制造缺陷检测实验平台
Fig.3 High-speed capsule manufacturing defect
detection experiment platform

胶囊通过传送带运送到工业相机镜头视野中,当光纤传感器感受到有样本经过时就会给样本采集部分发送触发脉冲; 样本采集部分按预设指令给照明系统和工业相机发送启动脉冲后实现对样本图像的捕获。专用环形光源 Microscope LED Ring Light起照明作用 照明点亮时间和相机曝光时间之间应该相匹配。在获取样本图像之后 样本采集部分会接收到模拟信号 经模拟数字转换器(Analog to Diglital ,A/D) 转换将图像数字化; 样本采集后数字图像会保存到处理器或计算机的内存里; 最后 处理器将对采集到的样本进行处理、识别、分析 ,以获得检测

2020年 无线电工程 第50卷 第10期 859

### 结果并存储。

### 22 数据集介绍

本实验采集的图像数据大小为512 pixel× 512 pixel , 一共采集了 600 个样本图片。 Mask R-CNN 不能直接输入采集的纯样本图片进行训练 濡 要对样本进行标注,以获得每个样本对应的 label. png 图片和 json 文件。目前主流的制作标签数据集 的工具软件有 labelme、VIA、LabelBox、COCO UI 等[16]。由于labelme简单快速,可以离线运行并且能 直接生产 json 文件 后面对 json 文件的批量转换和 label.png 提取都能用一个很简单的脚本文件实现, 所以采用 labelme 来制作数据集标签。本文主要针 对胶囊凹陷缺陷进行检测识别,因此把数据集分为 3 类 其中 T capsule 代表正常胶囊 F capsule 代表 缺陷胶囊 sunken 代表凹陷胶囊。样本展示如图 4 所示。

按训练集: 测试集按 3:1的比例划分训练集和 测试集,每个数据集都包括 4 个数据文件夹,分别是 放置每个样本 label.png 图片的 cv2\_mask; 放置 json 文件的 ison 文件夹; 放置 ison 文件转换后产生如图 5 所示的 5 个对应文件的文件夹 labelme\_json 和放置原 始数据图片的 pic 文件夹。所用胶囊是白色囊体和 黄色囊盖 其中正常胶囊样本量为 100 黄头缺陷样本 量为 100 白头缺陷样本量为 400 其中有200 个样本 是同一张图像中2粒胶囊 其余的都是每张图像中只 有一粒胶囊。样本图像大小为512 pixel×512 pixel, label.png 的位深为 8 位。





(a) 原始样本

(b) 标签样本

(c) 标签分类可视化

图 4 样本处理展示

Fig. 4 Sample processing display











info.yaml (a) 原始样本

lable.png label\_names.txt label\_viz.png (b) 标签样本 (c) 标签分类可视化

图 5 每个样本对应一个 json 文件夹

Fig5 Each sample corresponds to the file information contained in the json folder

#### 860 2020 Radio Engineering Vol. 50 No. 10

#### 实验与结果分析 3

### 3.1 损失函数与训练方法

本文实验平台为 Windows10 ,实验环境 CPU 为 Intel Core i7-6700 340 GHz×8.

Mask R-CNN 是在 Faster R-CNN 的基础上改进 的 其中一个改进点就是 Mask R-CNN 中添加了 mask 分支 ,所以每个 ROI 的损失函数就是在 Faster R-CNN 损失函数的基础上增加 mask 的损失值[16-17]:

$$L = L_{cls} + L_{box} + L_{mask} , \qquad (3)$$

式中  $L_{ck}$  为分类损失函数;  $L_{box}$  为边界框损失函数;  $L_{mask}$  为掩膜损失函数。训练时,mask 分支对每个 ROI 有  $K \uparrow m \times m$  维度的输出( $K \uparrow m \times m$  维度) 位  $m \times m$  的二值 mask)。针对一个第 K 个类别的 ROI , $L_{mask}$  仅仅取值第 K 个 mask ,这种定义的好处 就是对所有类别都会生成 mask ,而且不会存在类间 竞争。其中一组实验的 loss 值可视化如图 6 所示。

训练数据集一共 450 个样本 ,考虑到样本数量 较小 采取两阶段训练模式 即先训练头部分支 然 后再训练所有层,两阶段的 epochs 都各为 20 \_batch\_ size=1 STEPS\_PER\_EPOCH 设置 150 使在每一个 epoch 都会从训练集中随机抽取 150 个样本来训练。 根据经验值,设置头部分支训练阶段的学习率为 0001 在训练所有层时使学习率衰减 10 倍 ,为 00001 共迭代6000次。

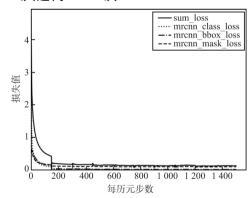


图 6 损失下降曲线 Fig6 Loss reduction curve

### 32 对比试验

通过改变 Backbone 网络和每个样本训练的 ROIs 数量以及单个样本目标数目,一共做了8组对 比实验。分别把经典的 ResNet101 和 ResNet50 作 为获取特征的骨架网络,然后对每个骨架网络做了

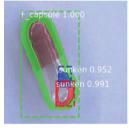
•‡•

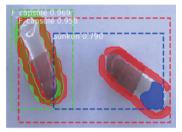
ROIs 分别为 200 和 300 的训练 ,得到每组实验的实例分割 mAP 如表 1 所示。样本测试结果展示如图 7所示。

表 1 目标检测的 mAP 结果对比

Tab.1 Comparison of mAP results of target detection

Method	Backbone	ROIs -	mAP/%	
			A 组	В组
Mask	ResNet50	200	7350	65.67
		300	75.00	68.17
R-CNN	ResNet101	200	9450	8000
		300	9500	8000





(a) 单目标样本测试结果

(b) 多目标样本测试结果

图 7 测试结果展示 Fig.7 Test results display

A 组表示大小为 100 ,且每个样本图片只有一粒胶囊的测试集; B 组表示大小为 150 ,其中有 50 个样本图片有 2 粒胶囊的测试集。从实验结果得出,每个样本仅有一粒胶囊的 mAP 远高于有多粒胶囊的样本 图 7 的结果也显示多目标样本的测试结果不太理想。骨架网络 ResNet101 在 mAP 方面更优于 ResNet50。对于 ROIs 的数量 200 与 300 的差别不大 ,差距在可接受的范围之内 ,但是当 ROIs 为 200 时速度比 300 更快 ,如果太低又会降低准确率 , 所以 ROIs 为 200 是综合最优的。

#### 4 结束语

本文论述了 Mask R-CNN 及其基础网络 ResNet 的结构和优缺点,分析了 ROIPooling 模块的局限性和 ROIAling 替换 ROIPooling 的优越性,并针对胶囊缺陷的检测识别,设计并搭建了胶囊制造缺陷在线检测实验平台。使用了 two stage 类的 Mask R-CNN模型方法对胶囊进行缺陷检测,实验结果表明,本文使用的方法对胶囊缺陷检测有较好的效果。在检测准确率和定位精确度上 two stage 类算法占有优势,而在算法速度上 one stage 相对更有优势<sup>[18]</sup>,对于本实验方法,文献[6]在搭载高性能 GPU 的实验平台

上的运行速度是 5 帧每秒 ,基本可以实现实时性的要求。在今后的研究中 ,对速度提升将展开更多的研究。

#### 参考文献

- [1] ZHU Z T ,HUANG L Q ,YU X Y. Pre-processing Techniques for Online Capsule Inspection Based on Machine Vision [C] // 2011 Fourth International Conference on Intelligent Computation Technology and Automation. Shenzhen: IEEE 2011: 653–656.
- [2] 齐丹阳.基于 SVM 的胶囊缺陷分类识别研究 [D].武汉: 武汉科技大学 2014.
- [3] 赖大虎 黄宴委.基于极端学习机的胶囊缺陷检测[J]. 福州大学学报(自然科学版) 2012 40(4):489-494.
- [4] 吴君才.基于图像处理胶囊缺陷识别与分类研究[D]. 南昌: 南昌航空大学 2013.
- [5] 刘孝星,郑力新,周凯汀, 等.采用图像处理的胶囊尺寸 缺陷检测方法 [J].华侨大学学报(自然科学版), 2017, 38(5):716-720.
- [6] HE K M ,GKIOXARI G ,DOLLÁR P ,et al. Mask R-CNN [J].arXiv preprint arXiv: 1703.06870v3 2018.
- [7] 王婷婷 刘环宇 李君宝 基于机载下视图像的深度学习目标检测系统[J].无线电工程 2019 49(9):759-767.
- [8] 王港 陈金勇 高峰 等.基于深度神经网络的遥感目标 检测及特征提取 [J]. 无线电工程,2018,48(9):
- [9] LONG J ,SHELHAMER E ,DARRELL T. Fully Convolutional Networks for Semantic Segmentation [C] // 2015
  IEEE Conference on Computer Vision and Pattern Recognition. Boston: CVPR 2015: 3431–3440.
- [10] XIE S N ,IRSHICK R ,DLOOAR P ,et al. Aggregated Residual Transformations for Deep Neural Networks [C] // Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE 2016: 5987–5995.
- [11] HE K M ,ZHANG X Y ,REN S Q ,et al. Deep Residual Learning for Image Recognition [C] // Proceedings of International Conference on Computer Vision and Pattern Recognition.Las Vegas: IEEE 2016: 770-778.
- [12] 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 & Machine Intelligence 2017 ,39(6): 1137-1149.
- [13] 李大军 何维龙 郭丙轩 等.基于 Mask-R-CNN 的建筑

2020年 无线电工程 第50卷 第10期 861

物目标检测算法 [J]. 测绘科学,2019,44(10): 172-180.

- [14] 薛冰.基于 Mask R-CNN 的电力设备锈迹检测[J].计算 机系统应用 2019 28(5):248-251.
- [15] 李少波 汪铮 杨静 筹.基于机器视觉的一维和二维条码高速在线识别 [J]. 计算机集成制造系统 ,2020 ,26 (4):910-919.
- [16] 吴金亮 汪港 梁硕 ,等.基于 Mask R-CNN 的舰船目标 检测研究[J].无线电工程 2018 ,48(11):947-952.
- [17] 林相泽 朱赛华 涨俊媛 筹.基于迁移学习和 Mask R-CNN 的稻飞虱图像分类方法 [J]. 农业机械学报, 2019 50(7):201-207.
- [18] 段仲静 李少波 杨静 等.深度学习目标检测方法及其主流框架综述 [J].激光与光电子学进展,2020,57 (12):59-74.

#### 作者简介



段仲静 男,(1992—),就读于贵州大学机械工程专业,硕士研究生。主要研究方向:产品在线无损质量检测。



(\* 通信作者) 李少波 男 (1973—) 博士, 教授。主要研究方向: 智能制造、大数据等。



胡建军 男 (1973一) ,博士 教授。主要研究方向: 大数据、深度学习。



杨 静 男 (1992—),博士研究生。主要研究方向: 机器视觉、智能制造、机器人。



王 铮 男 (1994—),硕士研究生。主要研究方向: 机械产品在线无损质量检测。