

文章编号: 2095-2163(2022)05-0031-06

中图分类号: TP391

文献标志码: A

# 基于铆钉表面缺陷检测的语义分割网络研究

赵娜<sup>1</sup>, 覃志东<sup>1</sup>, 蔡勇<sup>2</sup>, 肖芳雄<sup>3</sup>

(1 东华大学 计算机科学与技术学院, 上海 201620; 2 上海岳展精密科技有限公司, 上海 201614;

3 金陵科技学院 软件工程学院, 南京 211169)

**摘要:** 本文基于 U-net++ 基本网络结构, 结合铆钉表面缺陷检测的特点, 设计了一个语义分割网络模型。首先, 对标注数据集进行模糊标签处理, 增强了网络对缺陷的学习能力, 有效解决了铆钉缺陷边界确定难的问题。然后, 考虑到铆钉语义层次的特点, 对原始网络结构进行剪枝, 减少下采样的次数, 以满足系统对实时性的要求。最后, 设计了复合损失函数, 以促进小样本缺陷的学习速度, 解决了采用单一交叉熵损失函数在缺陷较小的数据集中收敛慢的问题。在对比试验中, 本文提出方法均展现出较好的效果。

**关键词:** U-net++; 语义分割; 缺陷检测

## Research on semantic segmentation network based on surface defect detection for rivets

ZHAO Na<sup>1</sup>, QIN Zhidong<sup>1</sup>, CAI Yong<sup>2</sup>, XIAO Fangxiong<sup>3</sup>

(1 College of Computer Science and Technology, Donghua University, Shanghai 201620, China;

2 Shanghai UAZAN Precision Technology Co., Ltd., Shanghai 201614, China;

3 School of Software Engineering, Jinling Institute of Technology, Nanjing 211169, China)

**[Abstract]** Based on the basic network structure of U-net++ and the characteristics of rivet surface defect detection, this paper designs a semantic segmentation network model. First, fuzzy label processing is performed on the labeled data set, which enhances the network's ability to learn about defects, and effectively solves the problem of difficulty in determining the boundary of rivet defects. Then, considering the characteristics of the rivet semantic level, the original network structure is pruned to reduce the number of downsampling to meet the real-time requirements of the system. Finally, a composite loss function is designed to facilitate the learning speed of few-shot defects, solving the problem of slow convergence in datasets with small defects using a single cross-entropy loss function. In the comparative experiments, the methods proposed in this paper have achieved good results.

**[Key words]** U-net++; semantic segmentation; defect detection

## 0 引言

铆钉作为一种紧固件, 被广泛用于航空、船舶、武器装备等安全关键行业。为杜绝铆钉缺陷所致安全隐患, 铆钉缺陷检测是生产加工过程中必不可少的一环。目前, 大多数缺陷检测方案是基于传统机器视觉技术的。如周永洪等人<sup>[1]</sup>采用阈值法去除噪声对检测的影响, 借助最小二乘法进行边缘拟合, 来提高铆钉缺陷检测的精度。刘春辉等人<sup>[2]</sup>采用了 Blob 算法、金字塔匹配等基于图像的缺陷判别方法来检测铆钉缺陷。但是, 由于受生产环境的干扰, 人工提取到的缺陷特征往往不能完全表达缺陷的特

性。所以, 铆钉表面缺陷具有难以量化的特点, 使得基于规则的传统机器视觉检测方案效能不佳。

而基于深度学习的缺陷检测技术利用卷积神经网络<sup>[3]</sup>针对原始输入图像, 逐层抽取底层特征最终得到抽象的高级语义, 提取到的特征更加合理, 避免了人工提取特征且特征模型后续难以使用的问题。深度学习的发展为产品表面缺陷检测提供了新的思路和替代方案。

考虑到产业界只对铆钉是否存在缺陷、缺陷位置和所占像素进行识别, 而不关注其所属类别, 因此本文将铆钉缺陷检测问题定位为机器视觉中的语义分割问题。FCN 网络<sup>[4-5]</sup>是一种典型的语义分割网

**基金项目:** 国家自然科学基金(6126200); 2021 国家级大学生创新创业项目(112-03-0178010/001)。

**作者简介:** 赵娜(1996-), 女, 硕士研究生, 主要研究方向: 人工智能; 覃志东(1974-), 男, 博士, 副教授, 主要研究方向: 嵌入式系统与人工智能; 蔡勇(1976-), 男, 硕士, 高级工程师, 主要研究方向: 图像处理与机器视觉; 肖芳雄(1972-), 男, 博士, 教授, 主要研究方向: 大数据与人工智能。

**通讯作者:** 覃志东 Email: zdqing@dhu.edu.cn

收稿日期: 2021-12-06

络,是将来自浅层的语义信息与深层的语义信息融合,以产生更为准确的分割结果。

FCN 虽然实现了端到端的卷积网络,但是其卷积过程仍然丢失了很多细节信息,分割精度有待提高。此外,FCN 缺乏空间一致性,分类过程是对独立的像素进行的,忽略了像素之间的关系。因此本文采用了另外一种解码效果更好,更注重保留细节信息的语义分割网络 U-net<sup>[6]</sup>。U-net 是 2015 年被提出用来解决医学中图像分割问题,由于其出色的分割结果,现已广泛应用到不同的语义分割任务中。2018 年,针对 U-net 网络中跳跃连接仅在编码器和解码器的相同比例特征图上进行强制融合以及编解码结构的深度因数据集而异的问题,Zhou 等人<sup>[7]</sup>对 U-net 原始网络进行了改进,提出了 U-net++。U-net++对固定深度的 U-net 进行改进,引入了一个内置的深度可变的 U-net 集合,并重新设计了跳跃连接方式,去除了 U-net 需要相同比例特征图进行融合的限制。本文将 U-net++作为基础网络,结合铆钉缺陷特点进行改进得到新网络 U-net++\*,用于铆钉表面缺陷检测。

## 1 U-net++网络介绍

为了更好地介绍 U-net++网络的结构,首先简单介绍一下原始的 U-net 网络。U-net 是经典的利用全卷积网络进行语义分割的算法,整体网络结构呈现一个“U”型结构<sup>[6]</sup>,左侧部分是网络的编码阶段,右侧是解码部分,如图 1 所示。灰色矩形框代表

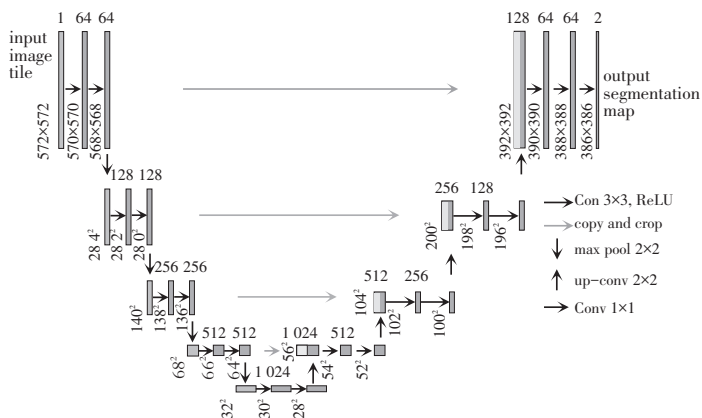


图 1 U-net 网络结构

Fig. 1 Network structure of U-net

的是网络提取的特征图,在编码阶段,每进行一次下采样,特征图尺寸都会减半,通道数加倍;目的是增加图像抗干扰的能力,降低过拟合的风险,减少运算量。在解码阶段是为了将特征图还原至原始图像大小。在进行上采样前,网络先将特征图与编码阶段生成的相同比例的特征图进行融合,用来增加浅层网络提取的细节信息。融合的方式是通过跳跃连接实现的。将编码器阶段的局部的细节特征与解码器阶段全局的抽象特征进行融合,使得网络更加有效地抓取目标的细节特征。

U-net++网络是 U-net 网络的改进版,在整体设计上还是遵循编解码的“U”型结构,如图 2 所示。U-net++与 U-net 原始网络相比,相当于把不同深度的 U-net 进行了拼接,让网络自主学习,极大地丰富了解码过程。为了解决单纯的拼接导致 U-net 网络会存在中间隐藏层的输出没有与最终的输出相连接,造成网络计算的损失无法通过反向传播的方式更新中间隐藏层权重的问题,在保留原始长链接的同时,为中间隐藏层每个卷积之间都增加了短链接。长链接和短链接共同作用于网络,使得解码阶段特征融合得更加灵活,还有助于还原下采样带来的信息损失。除此之外,网络还引入了深监督(Deep Supervision),为每一个隐藏层的输出增加了  $1 \times 1$  的卷积,来监督中间层的输出结果,不仅进一步解决了中间隐藏层无法训练的问题,还加快了网络的收敛速度。

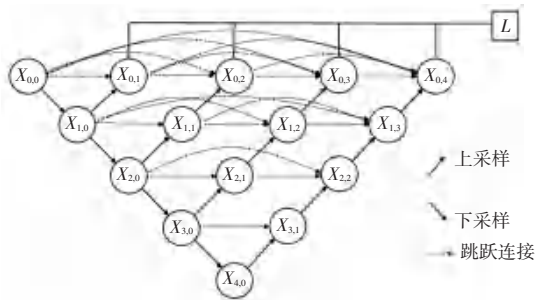


图 2 U-net++网络结构

Fig. 2 U-net++ network structure

## 2 改进的缺陷检测模型 U-net++\*

U-net++原始网络是针对医学图像中的细胞分割,虽然在医学图像上取得了很好的分割效果,但是

由于本文铆钉缺陷与医学细胞特征的差异性,使得原始网络并不能完全适合本文的检测对象。因此本节从铆钉表面缺陷特点出发,分析 U-net++网络结构并对其改进得到用于缺陷检测的模型

U-net++\*。

### 2.1 U-net++网络结构的改进

在深度学习网络中,浅层网络是用来提取一些具体的、局部的特征信息,随着网络层次的加深,所提取到的特征逐渐变得复杂。针对本文铆钉缺陷数据集而言,缺陷特征多为局部特征,语义层次较低,感受野不宜过大。因此铆钉表面缺陷检测网络不需要设计得过深。其次,网络越深,网络参数会呈现直线上升趋势,因此采用合理的网络结构是保证缺陷检测精度和效率的关键。

为了确定更适合铆钉缺陷的网络深度,本文基于U-net++设计了1~4层的网络结构,并在本文数据集上进行了对比试验,实验结果如图3所示,其中圆圈、矩形、三角形、五角星分别表示1~4层的网络。随着网络深度的增加, *mIoU* 和推理时间都有增加的趋势。但是4层的网络结构相比于3层网络结构, *mIoU* 提升0.3个百分点。推理时间需要多消耗19.6个百分点,可见用时间消耗大幅增加换来精度的微弱提升是不值得的,因此本文对原始网络进行剪枝,采用3层下采样的网络结构,如图4所示。

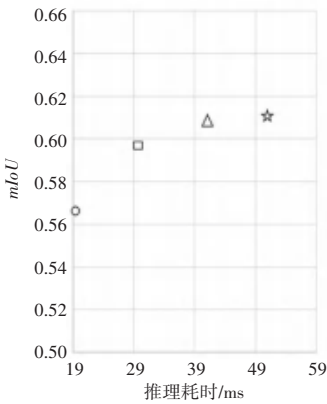


图3 不同深度网络对比图

Fig. 3 Comparison of networks with different depth

网络的整体结构还是遵循典型的全卷积网络结构,编码阶段由3部分组成,每一部分包含2个3\*3的卷积,每个卷积后有一个ReLU激活函数和一个最大池化层。其中,卷积采用same卷积。避免了在解码阶段时为了使不同尺寸的特征图进行融合而采用的裁剪步骤。解码阶段由6个基本单元组成,每一部分包含一个2\*2的反卷积对特征图尺寸进行加倍,2个3\*3的卷积加一个ReLU激活函数,最后通过一个1\*1的卷积映射到二分类进行输出。

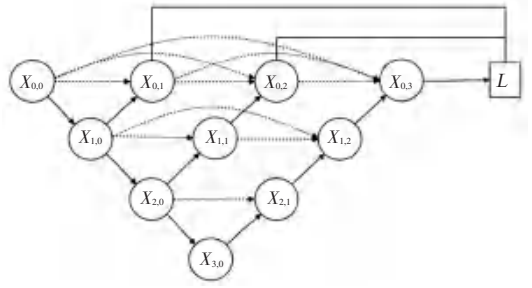


图4 剪枝后的网络结构

Fig. 4 Network structure after pruning

### 2.2 数据增强

工业中缺陷数据集的收集比较困难,而且铆钉的缺陷种类繁多,形状各异,缺陷标注的工作量非常大,标注过程中往往需要根据公司需求反复确认修改标注。因此本文最终共标注了300张缺陷图像。为了在小数据集上训练得到一个较好的模型,避免训练过程的过拟合现象,本文除了使用常规的水平翻转、垂直翻转等数据增强方式外,还针对铆钉缺陷的特点设计了模糊标签的数据增强方式。

工业环境中缺陷产生的随机性导致缺陷并没有一个标准的形状,对于那些位于缺陷中心的像素可以很明确地认为是缺陷像素,但是对于缺陷边界处的像素就难以界定。同时,本文的缺陷检测是像素级别的检测,每个像素的标签值对网络的学习结果影响较大。为此,本文提出了模糊标签的解决方案,即在缺陷边缘区域像素的标签值不仅局限于是缺陷,或不是缺陷这种非0即1的情况,而是通过一个较好的方式根据缺陷边缘像素不确定的程度赋予一个对应的标签权重。本文中标签模糊化是通过一个3\*3的均值滤波来处理的。均值滤波处理思路比较简单,即将这个3\*3大小窗口内的像素根据式(1)来计算平均值:

$$F(x, y) = \frac{1}{N} \sum_{i=1}^N V(x, y) \quad (1)$$

此后,将窗口计算得到的均值设置为锚点上的像素值。均值滤波有个明显的特点,就是通过计算均值的方式会使边缘信息模糊掉,而本文正是基于此而将原本标签边界0到255的跳变通过均值滤波处理成连续的像素值。

### 2.3 损失函数改进

原始网络是针对医学图像细胞的分割,针对细胞之间相互接触问题引入了带权重的交叉熵损失函数,使得越靠近边缘的像素权重越大,从而使网络更

好地学习边缘信息。对于语义分割这种二分类的情况,模型最终只需要输出2种结果,交叉熵损失函数非常适合这种二分类问题。假设每个类别中预测的概率为 $p$ 和 $1-p$ ,那么交叉熵损失函数可以写成式(2)的形式:

$$Loss = \frac{1}{N} \sum_{i=1}^N Loss_i = - \frac{1}{N} \sum_{i=1}^N [y_i \cdot \log p_i + (1 - y_i) \cdot \log(1 - p_i)] \quad (2)$$

其中, $y_i$ 表示样本 $i$ 的标签,前景为1,背景为0; $p_i$ 表示样本 $i$ 预测为正类的概率,在这里是归一化后的概率值。由于本文采用模糊标签增强数据,所以就需要进行归一化处理,本文选择了更适合二分类的 $sigmoid$ 激活函数如式(3)所示:

$$sigmoid(h_i) = \frac{1}{1 + e^{-z_i}} \quad (3)$$

其中, $z_i = \mathbf{w}_i^T \mathbf{X}_i + c_i$ , $\mathbf{w}_i^T$ 是权重矩阵的转置矩阵。

但单一交叉熵损失函数在本文数据集中存在收敛速度较慢的问题。原因在于本文数据集中的缺陷形状大小差别较大,有很多缺陷、如碰伤,所占像素面积很小,但是深度较深仍需要检出,加上交叉熵损失函数的竞争机制,使得损失函数关注于正确标签预测概率的准确性。如果网络把所有的样本都视为无缺陷样本,准确率也可以达到一个很高的结果,因此若只是用交叉熵损失函数,网络训练前期将长时间学习不到缺陷样本。为了提高网络的收敛速度,本文引入了 $IoU$ 损失函数<sup>[8]</sup>。 $IoU$ 即交并比,可以很好反映出预测区域和真实区域的预测效果。 $IoU$ 即交集与并集的比值。 $IoU$ 值越大,说明预测值与真实值越接近,损失函数值应该越小。 $IoU$ 损失函数的定义如式(4)所示:

$$IoU_{loss} = - \ln \frac{Intersection(A, B)}{Union(A, B)} \quad (4)$$

其中, $A$ 表示预测的像素区域; $B$ 表示标注的像素区域。本文采用交叉熵损失函数和 $IoU$ 损失函数相结合的复合损失函数方式,可以有效加快网络的收敛速度。

## 3 实验结果

### 3.1 公开数据集上对比试验

鉴于本文所使用的铆钉表面缺陷数据集是未公开数据集,为了证明本文网络模型改进方案的合理性和有效性,故在公开数据集 Kolektor 对本文网络模型与其它算法进行对比实验。Kolektor 是由

Kolektor Group 提供和标注的电子换向器表面缺陷数据集。图像数据在电子换向器8个不重叠的表面进行采集,共采集了50个有缺陷的电子换向器。在采集过程中保证一个缺陷仅在一张图像中可见。数据集共采集399张图像。其中有缺陷图像52张,无缺陷图像347张,每张图像分辨率为 $500 * 1263$ 。为了适应本文网络训练,图片分辨率被调整为 $480 * 1248$ 。在该公开数据集上评估了语义分割算法 U-net、DeepLabv3+ 以及文献[9]提出的改进算法 SqueezeNet,与本文模型 U-net++ 的数据对比结果如图5所示。

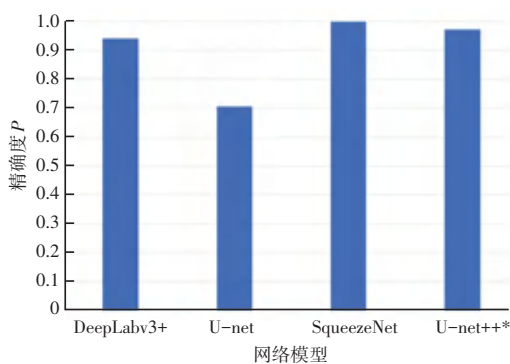


图5 网络模型精确度对比柱状图

Fig. 5 Accuracy comparison histogram of network models

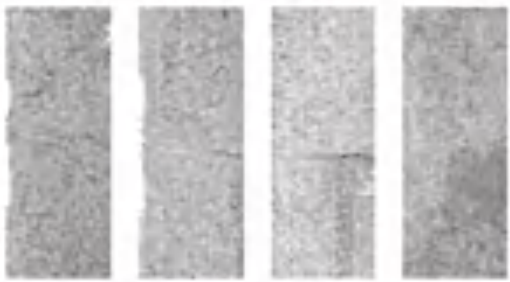
从图5可以看出,文献[9]采用的网络模型 SqueezeNet 精确度最高,但是其网络模型并未对图像进行全分辨率的像素级别的分割,输出图像的分辨率降低了8倍,而本文网络模型输出结果是原比例输出,仅比 SqueezeNet 低了1.8个百分点。此外,本文算法模型在该公开数据集上将缺陷样本识别为无缺陷样本的概率降至0,其它对比算法中仍存在将缺陷样本推理为无缺陷样本的情况。

图6是本文网络模型在 Kolektor 上的预测部分结果展示,图6(a)是人工标注结果,图6(b)是本文模型推理结果。图6(c)是人工标注结果的对应放大图,图6(d)是本文算法推理结果的对应放大图。该结果直观表明了本文网络模型在 Kolektor 数据集上仍有较好的表现。

### 3.2 损失函数的改进实验

为了验证 $IoU$ 损失函数,与交叉熵损失函数构成复合损失共同作用于网络的训练过程,与只使用单一的交叉熵损失函数的网络模型相比,可以加快模型拟合的速度。本文设计了单一交叉熵损失函数和复合损失函数在本文数据集和公开数据集对比试

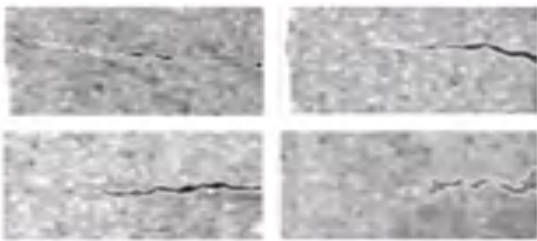
验 Kolektor 上的对比结果。实验数据如图 7 所示。图 7 中,数据集 1 是本文铆钉的表面缺陷数据,数据集 2 是公开数据集 DAGM 2007。横坐标表示模型训练的迭代次数,纵坐标  $mIoU$  用于观测模型训练过程的拟合趋势,其值是训练数据集在当前迭代次数训练得到的模型检测出缺陷面积与标注图像中缺陷面积的比值。



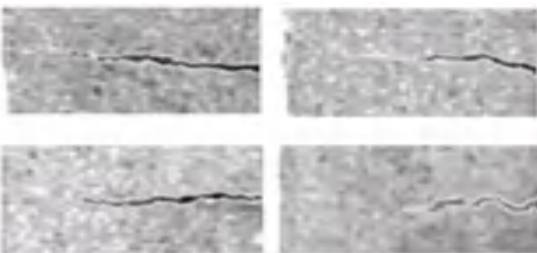
(a) 人工标注结果



(b) 本文模型推理结果



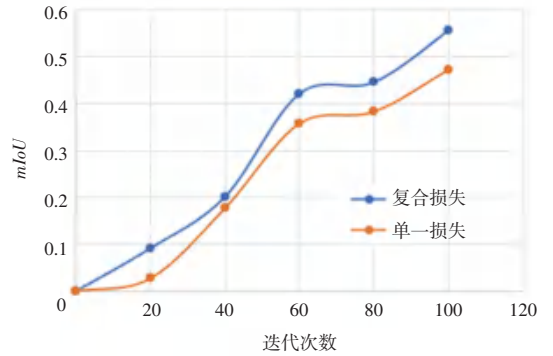
(c) 人工标注结果的对应放大图



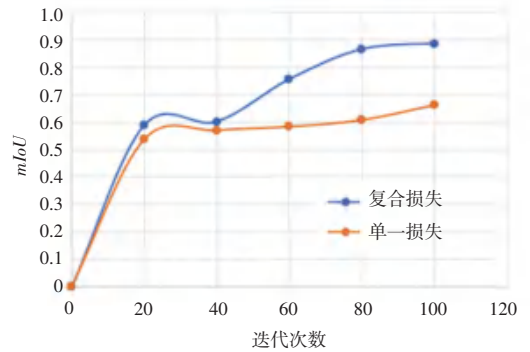
(d) 本文模型推理结果的对应放大图

图 6 U-net++\* 模型在 Kolektor 上的预测结果

Fig. 6 Prediction results of the model U-net++\* on Kolektor



(a) 数据集 1 的实验结果



(b) 数据集 2 的实验结果

图 7 有无  $IoU$  损失函数对比曲线图

Fig. 7 Comparison graph of loss function with and without  $IoU$

从 2 组对比实验不难发现,复合损失函数在前期迭代过程中,都不同程度上加快了网络的收敛速度,分别在数据集 1 上提升了 21.3%、数据集 2 上提升了 24.9%。实验表明,单一的交叉熵损失函数对网络模型的优化方向较为单一,对于背景较多的数据集而言,找到前景所消耗的时间较长。交叉熵损失函数配合  $IoU$  损失函数后,可以促使优化方向发生一定的跳动以加快寻求最好的优化结果的速度。

### 3.3 铆钉缺陷检测结果

图 8 展示的是本文模型 U-net++\* 推理结果,可以看到每一个缺陷都计算了其所占像素。因为本文在缺陷标注过程中,一些细微的缺陷仍然标注在内使得模型将很小的缺陷也被标注出来,这往往会导致最终的精确度会偏低,因为某部分存在一些狭小缺陷但可以被认定为良品的铆钉被判定为了不良品。因此在系统做了根据缺陷所占像素进行筛选的功能操作后,根据用户的需求,可以调高或者放低对缺陷大小的要求,更符合铆钉的生产场景需求。

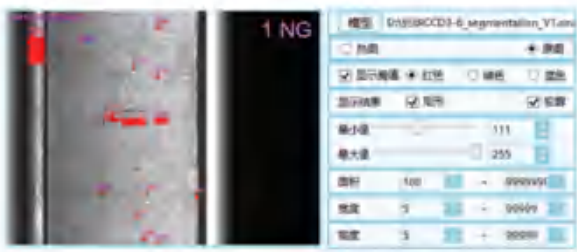


图8 模型推理检测结果

Fig. 8 Model inference detection results

## 4 结束语

目前,缺陷检测已经成为工业生产流水线中的重要一环,随着对产品质量的要求越来越高,基于规则的传统机器视觉检测方案在一些难以量化的缺陷面前表现不佳。本文针对铆钉表面缺陷检测的特点,设计了一种基于U-Net++的语义分割网络模型U-net++\*。首先针对铆钉缺陷语义层次的特点,对U-net++原始网络进行剪枝;此外针对数据集较小、且缺陷边界难以界定的问题,提出了模糊标签的数据增强方式;并针对单一的交叉熵损失函数在网络训练中存在的收敛过慢的问题,本文增加了IoU损失函数与其构成复合损失函数共同作用于网络。本文提出的铆钉表面缺陷检测模型U-net++\*,是在U-Net++的基础上改进优化,满足工业实时性检测需求的网络结构。

## 参考文献

- [1] 周永洪,韩震宇.基于计算机视觉的铆钉缺陷检测[J].科技与创新,2017(13):104-105.
- [2] 刘春辉,徐平.轴承铆钉缺陷检测方法研究[J].杭州电子科技大学学报,2014,34(5):101-105.
- [3] LÉCUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [4] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA:IEEE,2015:3431-3440.
- [5] 欧先锋,晏鹏程,向灿群,等.基于FCN的TFT-LCD表面缺陷快速检测算法[J].成都工业学院学报,2017,20(03):6-10.
- [6] 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. Cham:Springer, 2015:234-241.
- [7] ZHOU Zongwei, SIDDIQUEE M M R, TAJBAKHS N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]// Deep learning in medical image analysis and multimodal learning for clinical decision support. Cham:Springer, 2018:3-11.
- [8] YU Jiahui, JIANG Yuning, WANG Zhangyang, et al. Unitbox: An advanced object detection network[C]// Proceedings of the 24th ACM International Conference on Multimedia. New York, NY, United States:ACM, 2016:516-520.
- [9] 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] 龙真真,张策,刘飞奇,等.一种改进的Chameleon算法[J].计算机工程,2009,35(20):189-191.
- [6] 文俊浩,何光辉,任海军.利用改进变色龙算法进行有障碍物体聚类[J].计算机工程与应用,2005(32):32-33,56.
- [7] 康海燕,李昊.基于Web日志的性格预测与群体画像方法研究[J].郑州大学学报(理学版),2020,52(01):39-46.
- [8] 王鹏宇,王国宇,贾贞,等.一种基于局部特征的层次聚类算法[J].中国海洋大学学报(自然科学版),2019,49(S2):176-184.
- [9] 韩新新.基于K-medoids的Chameleon算法分析[J].现代商贸工业,2019,40(34):195-196.
- [10] 喻新潮,曾圣超,温柳英,等.一种聚类与kNN结合的协同过滤算法[J].小型微型计算机系统,2019,40(04):755-759.
- [11] 徐国天.一种高维向量空间K近邻快速搜索方法[J].小型微型计算机系统,2020,41(11):2274-2281.
- [12] 冷亚军,吴宗育,赵文会,等.基于近邻传播聚类权重的黑启动方案评估方法[J].电力系统自动化,2020,44(13):73-80.
- [13] 白兰,王震,邵元海.回归问题中的k近邻加权方法研究[J].小型微型计算机系统,2016,37(07):1557-1561.
- [14] 唐洪婷,李志宏.基于hMETIS与FP-Growth的协同创新社区领域知识发现方法[J].系统工程理论与实践,2018,38(08):2068-2078.
- [15] 王野.基于Chameleon算法的网络社群用户特征挖掘[D].哈尔滨:哈尔滨工程大学,2018.
- [16] 刘奕志,程汝峰,梁永全.一种基于共享近邻的密度峰值聚类算法[J].计算机科学,2018,45(02):125-129,146.
- [17] 田兴宇,张雄杰,汤彬,等.基于D-D中子源校准中子周围剂量当量仪的最小房间尺寸研究[J].计量学报,2021,42(07):971-976.

(上接第30页)