

## 基于深度学习的磁瓦表面缺陷分割与识别

谢舰, 姚剑敏, 严群, 林志贤

引用本文:

谢舰, 姚剑敏, 严群, 等. 基于深度学习的磁瓦表面缺陷分割与识别[J]. *液晶与显示*, 2021, 36(5): 713–722.

XIE Jian, YAO Jian-min, YAN Qun, et al. Segmentation and recognition of magnetic tile surface defects based on deep learning[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(5): 713–722.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2020-0232>

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于深度学习的遥感图像地物分割方法

Remote sensing image feature segmentation method based on deep learning

*液晶与显示*. 2021, 36(5): 733–740 <https://doi.org/10.37188/CJLCD.2020-0294>

### 基于深度学习的乳液泵缺陷检测算法

Defect detection algorithm of lotion pump based on deep learning

*液晶与显示*. 2019, 34(1): 81–89 <https://doi.org/10.3788/YJYXS20193401.0081>

### 基于深度学习的木材表面缺陷图像检测

Image detection of wood surface defects based on deep learning

*液晶与显示*. 2019, 34(9): 879–887 <https://doi.org/10.3788/YJYXS20193409.0879>

### 基于U-Net的多图谱标签融合算法

Multi-atlaslabel fusion based on U-Net

*液晶与显示*. 2019, 34(11): 1091–1103 <https://doi.org/10.3788/YJYXS20193411.1091>

### 基于AttentionGAN和形态学重建的TFT阵列缺陷检测

TFT array defect detection based on AttentionGAN and morphological reconstruction

*液晶与显示*. 2020, 35(12): 1270–1277 <https://doi.org/10.37188/YJYXS20203512.1270>

文章编号:1007-2780(2021)05-0713-10

## 基于深度学习的磁瓦表面缺陷分割与识别

谢 舰<sup>1</sup>, 姚剑敏<sup>1,2\*</sup>, 严 群<sup>1,2</sup>, 林志贤<sup>1</sup>

(1. 福州大学 物理与信息工程学院, 福建 福州 350108;

2. 晋江市博感电子科技有限公司, 福建 晋江 362200)

**摘要:**为了满足磁瓦生产工业对表面质量检测的高要求,实现磁瓦缺陷自动分割与识别,本文提出了一种基于卷积神经网络的缺陷分割与分类网络。该网络基于 U-net 架构,通过 U-net 编码部分提取缺陷的深层特征,并使用该深层特征进行缺陷分类,然后通过解码部分输出分割的缺陷区域。为了解决部分缺陷前景面积占比太小,导致网络难以收敛的问题,通过添加差异系数损失以保证网络持续优化。然后在训练阶段添加多层损失和进行在线数据增强进一步提升了分割精度和分类准确率。实验结果表明,添加辅助损失函数和数据增强后,分割网络能够分割出 94.5% 标注的缺陷区域,并且对于缺陷分类的准确率能够达到 98.9%,满足工业生产的高精度要求。该方法能够精准有效地分割和识别磁瓦的表面缺陷,为磁瓦表面质量检测自动化行业提供了一种新的思路。

**关键词:**磁瓦; 缺陷分割; 缺陷分类; U-net; 卷积神经网络

**中图分类号:**TP394.1;TH691.9 **文献标识码:**A **doi:**10.37188/CJLCD.2020-0232

## Segmentation and recognition of magnetic tile surface defects based on deep learning

XIE Jian<sup>1,2</sup>, YAO Jian-min<sup>1,2\*</sup>, YAN Qun<sup>1,2</sup>, LIN Zhi-xian<sup>1</sup>

(1. College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China;

2. Jinjiang RichSense Electronic Technology Co., Ltd., Jinjiang 362200, China)

**Abstract:** In order to meet the high requirements of the magnetic tile production industry for surface quality inspection and realize the automatic segmentation and recognition of magnetic tile defects, a defect segmentation and classification network based on convolutional neural networks is proposed. The network is based on the U-net architecture. The deep features of defects are extracted through the U-net encoding part, and the deep features are used for defect classification, and then the segmented defect areas are output through the decoding part. In order to solve the problem that the proportion of the foreground area of some defects is too small, which makes the network difficult to converge, the

**收稿日期:**2020-09-24; **修订日期:**2020-10-28.

**基金项目:**国家重点研发计划课题(No.2016YFB0401503);广东省科技重大专项(No.2016B090906001);福建省科技重大专项(No.2014HZ0003-1);广东省光信息材料与技术有限公司重点实验室开放基金(No.2017B030301007)

Supported by National Key Research and Development Program of China (No.2016YFB0401503); Science and Technology Major Program of Guangdong Province (No.2016B090906001); Science and Technology Major Program of Fujian Province (No.2014HZ0003-1); Guangdong Provincial Key Laboratory of Optical Information Materials and Technology (No.2017B030301007)

\* 通信联系人, E-mail: yaojm@fzu.edu.cn

continuous optimization of the network is ensured by adding the difference coefficient loss. Then, adding multiple layers of loss and performing online data enhancement in the training phase further improves the segmentation accuracy and classification accuracy. Experimental results show that with the addition of auxiliary loss function and data enhancement, the segmentation network can segment 94.5% of the marked defect areas, and the accuracy of defect classification can reach 98.9%, which meets the high precision requirements of the industry. This method can accurately and effectively segment and identify the surface defects of the magnetic tile, which provides a new idea for the automatic industry of magnetic tile surface quality inspection.

**Key words:** magnetic tile; defect segmentation; defect classification; U-net; convolutional neural network

## 1 引 言

磁瓦是由永磁铁氧体材料制成的瓦形磁体,其在工业上常作为生产永磁电机的核心部分之一。在厂端批量的生产加工磁瓦的过程中,不可避免地会产出表面带有各种缺陷的残次品。这些残次品表面带有不同类型的缺陷,如果用于生产永磁电机,将会影响永磁电机的性能,严重的会引发各种故障。另一方面厂端也需要精准地分辨各种缺陷的边界和大小,来决定这些残次品是否可进行维修或进入下一步工序,因此在磁瓦生产工艺中,表面质量检测是十分重要的一环。传统方法主要依靠人工检查磁瓦的表面,这需要额外的雇佣大量的员工,并且需要专门培养员工来识别各种复杂的表面缺陷。这种方法十分耗时、耗力,不仅效率不高,并且限制了工厂的生产力。因此开发自动缺陷识别系统,将磁瓦表面缺陷识别自动化、智能化,是磁瓦制造企业适应工业化 4.0 发展的重要一步。

早期对磁瓦表面缺陷的自动识别还集中在使用传统图像处理技术上。Li<sup>[1]</sup>提出了一种使用快速离散曲波变换(FDCT)与纹理分析的方法来自动检测磁瓦上裂纹缺陷的机制,可以检测到最小长度为 0.8 mm 的裂纹。杨成立<sup>[2]</sup>等人提出使用小波变换去噪处理图像,通过计算对比阈值提取区域匹配相似度和对比轮廓边缘面积的方式判别磁瓦是否存在缺陷,该方法对磁瓦表面存在的显著缺陷有较好的检测效果。针对经典缺陷检测算法不能很好地提取颜色暗、对比度低的磁瓦图像缺陷问题,李雪琴<sup>[3]</sup>等人提出一种非下采样 Contourlet 域自适应阈值面的磁瓦缺陷自动检测方

法。为了准确定位磁瓦表面缺陷的边界,林丽君<sup>[4]</sup>等人提出一种基于图像加权信息熵与小波模极大值两者结合的缺陷边缘检测算法。以上所描述的方法能够很好地解决部分场景下的磁瓦表面缺陷检测,但面对生产线上背景复杂多变、颜色形状繁多的缺陷,传统图像算法能够起到的作用十分有限。

近几年随着深度学习技术的迅速发展,国内外对基于深度学习的缺陷提取技术进行了不同程度的研究,Wang<sup>[5]</sup>等人尝试对输入图像进行滑窗切片并通过卷积神经网络进行逐位置识别,来实现对布匹缺陷的检测。Mei<sup>[6]</sup>等人通过改进的卷积降噪自编码器网络,使用无缺陷样本进行训练并完成对缺陷的定位。王春哲<sup>[7]</sup>等人引入卷积边缘信息、显著性及位置信息改进了候选区域算法从而有效提升了目标检测算法的召回率。刘畅<sup>[8]</sup>等人通过 U-net<sup>[9]</sup>分割磁瓦表面缺陷,并且将分割的前景裁剪后使用另一个单独的分类网络来进行缺陷分类,实现了缺陷的分割和分类并且都达到了很高的精度。但该方法分割的区域不准时会影响分类准确率,因而限制了准确率的提升。Tabernik<sup>[10]</sup>等人的工作是尝试使用一个网络同时完成分割与分类的任务,仅仅使用少量的训练数据就在 KolektorSDD 数据集上取得了 99.8% 的缺陷识别率。但该方法只能输出原图 1/8 的分割结果,无法分割出原图一致大小的精细掩膜。综上所述,基于深度学习的方法在磁瓦表面缺陷检测识别上已经取得了不错的效果,但在实际生产过程中往往存在着能够获取的总数据少且每种类别分布不均匀的问题,并且想要满足厂端缺陷分割精准、缺陷分类准确率高的要求,还要解决训练

过程中由于缺陷前景面积占比小导致网络难以收敛等技术问题。

结合以上所存在的问题,本文提出了一种基于卷积神经网络的新模型,能够使用少量的训练数据实现对磁瓦表面缺陷的实时分割与分类,很好地契合了工业界对于磁瓦表面质量检测的高要求。

## 2 方 法

图 1 显示了本文设计的网络的主要架构,其中包括了一阶段的分割网络,该分割网络对磁瓦表面的缺陷进行像素级别的定位,通过添加辅助损失函数来提升模型的分割精度。随后将网络提取到的深层特征送入二阶段的分类网络,从而实现了对磁瓦表面缺陷的精准分割和分类。

### 2.1 分割网络

分割网络使用的是经典的 U-net 网络架构。U-net 作为经典的语义分割模型,经常用于医学图像分割、缺陷分割等场景,它能够使用较少的数据获得不错的分割效果。在网络结构上,U-net 采用的是编解码的结构,先对输入进行降采样然后再上采样。

编码结构的存在是使卷积神经网络通过多次

降采样后能够拥有一个足够大的感受野,这个感受野能让网络提取到更加底层的语义信息,更加关注于一个整体目标和其他背景间的区别。解码结构是将网络提取到的抽象特征恢复到原图大小的过程,由于深度神经网络一般会使用池化层或者步长不为 1 的卷积层进行降采样,在这个过程中会逐渐丢失一些细粒度的信息,所以需要使用直连(Skip-connection)结构,这个结构从高层重新补充了语义信息,在底层特征的基础上细化了轮廓和边缘,从而保证了分割的精准度。

分割网络结构如图 1 的红色虚线框部分所示,蓝色的块为每层输出的特征图,黄框内为每层的具体操作。为了加速网络收敛、稳定训练,分割部分使用的卷积除  $1 \times 1$  的以外都通过了批归一化层<sup>[11]</sup>和非线性激活层。将原图经过两个  $3 \times 3$  的卷积进行扩维操作后,一共使用 4 个降采样块进行编码操作,每个降采样块包含一个  $2 \times 2$  的最大池化和两个  $3 \times 3$  的卷积,通过最大池化将输出宽高降为原来的  $1/2$ 。同时为了保证降采样过程中总体信息不损失,每个降采样块都会将图像通道数翻倍,于是在编码部分的底层输出原图宽高  $1/16$  大小且通道数为 512 的特征图。解码部分可以看做编码部分的镜像,不同的是将  $2 \times 2$  的最大池化替换为上采样操作。上采样使用的方法为

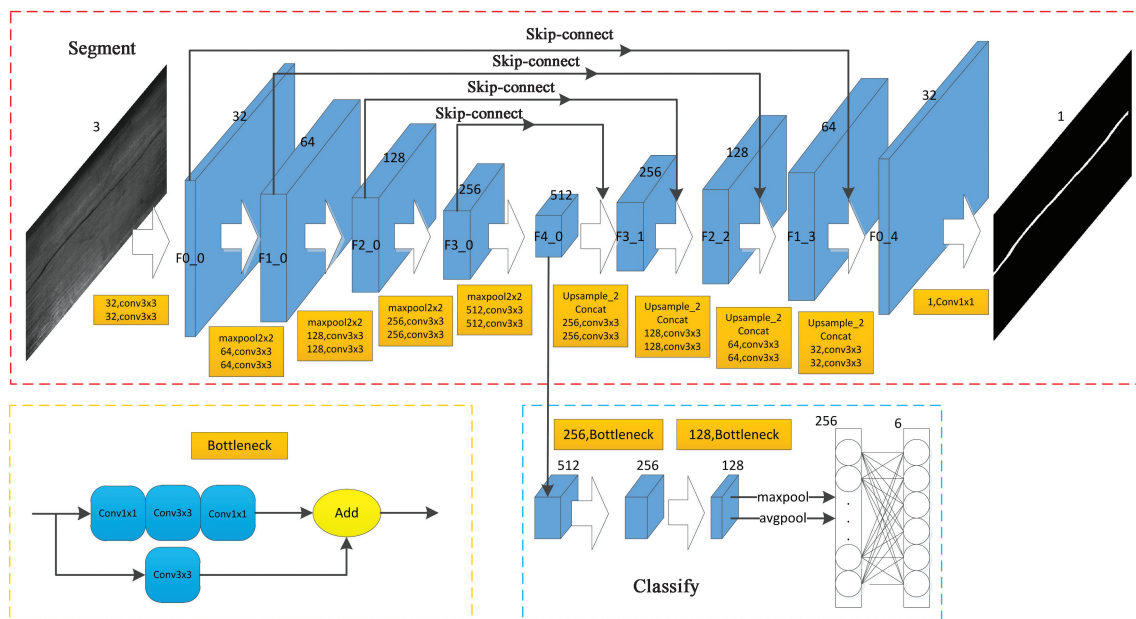


图 1 网络结构示意图

Fig.1 Network structure diagram

双线性上采样,同时将图像编码部分提取到和上采样后一致大小的特征图在通道上进行合并,最后在经过两个  $3 \times 3$  的卷积合并多个通道的信息输出给下一层。最终将编解码得到的特征图通过一个  $1 \times 1$  的卷积进行像素级的分类,输出和原图一致大小的掩膜,于是就完成了磁瓦表面缺陷的分割。

由于在分割上只关注缺陷前景和背景,所以在基础损失函数的设计上本文使用的是二值交叉熵损失。对于每个像素预测为缺陷概率为  $p$ ,背景概率为  $1-p$ ,其公式的定义如下:

$$L_{\text{bce}} = \frac{1}{N} \sum_i - [y_i \ln p_i + (1 - y_i) \ln(1 - p_i)], \quad (1)$$

其中  $y_i$  表示每个像素的标签,缺陷类别为 1,背景类别为 0,  $p_i$  表示像素预测为缺陷类别的概率。

并且在实际磁瓦表面缺陷分割问题上考虑到很多细小缺陷实际占磁瓦表面积比例很小,所以对于损失函数的形式进行修正,提升缺陷正样本所占的损失比重,表达式如下:

$$L_{\text{sce}} = \frac{1}{N} \sum_i L_i + \frac{1}{M} \sum_j L_j, \quad (2)$$

其中  $i$  表示属于标签的所有缺陷类别的像素集合,  $j$  表示属于标签的所有背景类别的像素集合,两者各自求交叉熵的均值后相加。

### 2.1.1 差异系数损失

为了解决部分的磁瓦表面缺陷前景面积占比太小,导致训练阶段梯度太小网络难以收敛,本文添加了差异系数损失<sup>[12]</sup> (Dice loss) 来提升训练速度和分割精度,其定义如下:

$$S = \frac{2 |X \cap Y|}{|X \cup Y|}, \quad (3)$$

从定义上看差异系数是一个相似度度量函数,能够用于计算两个样本间的匹配度。在分割任务上其计算形式是预测的前景区域和标签掩膜前景区域的交并比,所以它是从一个整体角度去考虑和标签的区别。这样即使磁瓦表面缺陷前景面积很小,依然不会影响分割结果和标签计算的差异度,从而维持训练阶段的梯度。并且直接将分割效果评估指标作为损失去监督网络训练,能够忽略背景像素的影响,解决正负样本不均衡的问题,提升网络收敛速度。对应差异系数,其损失函数的形式如下:

$$L_{\text{dice}} = 1 - S = 1 - \frac{2 |X \cap Y|}{|X \cup Y|}, \quad (4)$$

其中,  $X$  是网络预测缺陷区域的像素个数,  $Y$  是标签掩膜的缺陷区域的像素个数。

### 2.1.2 多层损失

为了解决训练过程中深层网络的梯度逐渐消失导致模型无法完全收敛的问题,进一步提升分割精度。本文添加了多层损失<sup>[13]</sup> 维持网络深层的梯度,监督网络的深层继续优化从而降低网络训练难度并提升精度。具体的做法为:对于如图 1 网络编解码部分输出的特征图  $F4\_0$ 、 $F3\_1$ 、 $F2\_2$  和  $F1\_3$ ,在网络训练阶段对每个特征图通过一个  $1 \times 1$  的卷积输出和对应特征图一致宽高的掩膜。于是除网络本身输出外,又得到了原图  $1/2$ ,  $1/4$ ,  $1/8$ ,  $1/16$  大小的多层输出掩膜。并且考虑要得到原图  $1/2$ ,  $1/4$ ,  $1/8$ ,  $1/16$  大小的标签掩膜,就需要对标签的尺寸降低对应的倍率,这会导致一些过小的缺陷标签完全失去边缘信息。于是本文修改方法为对网络多层输出掩膜进行双线性上采样到原图大小,然后计算对应的损失,具体的公式如下:

$$L_{\text{muti}} = L_{\text{sce}}(\bar{Y}_1, Y) + L_{\text{sce}}(\bar{Y}_2, Y) + L_{\text{sce}}(\bar{Y}_3, Y) + L_{\text{sce}}(\bar{Y}_4, Y), \quad (5)$$

其中  $Y$  为标签掩膜,  $\bar{Y}_1$ 、 $\bar{Y}_2$ 、 $\bar{Y}_3$  和  $\bar{Y}_4$  分别为网络解码器部分输出的特征图  $F4\_0$ 、 $F3\_1$ 、 $F2\_2$  和  $F1\_3$ ,通过一个  $1 \times 1$  卷积输出的结果双线性上采样到  $Y$  大小的输出掩膜。它们各自计算公式 (2) 的二值交叉熵后求和,即为最终的多层损失值。

对应添加辅助损失函数后训练的损失下降和缺陷类别的 MIoU 指标上升曲线如图 2 所示,大约训练 75 轮以后基础的交叉熵损失趋近 0,无法再提供梯度信息。此后主要由添加的差异系数损失和多层损失产生梯度,使得网络能够继续优化,使得最终缺陷类别的 MIoU 指标可以提升至接近 70%。

## 2.2 分类网络

分类网络的设计上采用的网络架构类似于 Tabernik<sup>[10]</sup> 等人提出的方法,并且进行了适应性的修改。本文没有设计一个单独的分类网络来重新训练分割出的磁瓦表面缺陷,而是通过在分割网络深层连接卷积层分支来将网络编码阶段提取到的高级语义信息额外用于磁瓦表面缺陷分类。

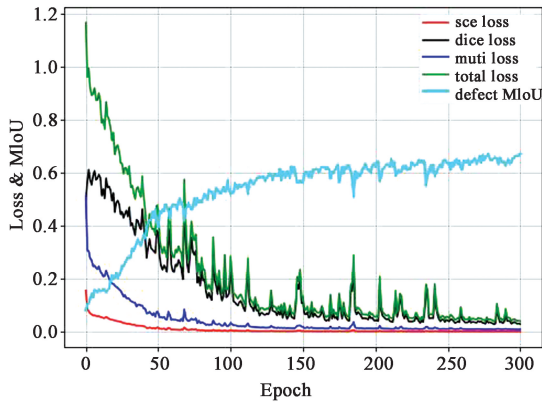


图 2 损失下降与 MIoU 上升曲线

Fig.2 Loss decrease and MIoU rate increase curve

之所以这么做是因为考虑到分割任务本身就拥有比分类任务更强的监督信息,通过分割精确到像素的监督信息网络能够学习到缺陷更细致的轮廓与特征,进一步过滤背景等噪声的影响,从而得到比只用类别标签训练更高的准确率。

参考通常的分类网络<sup>[14-17]</sup>的设计一般会采用 32 倍的降采样倍率,但考虑到部分缺陷很小,在多次的降采样过程中的信息丢失会使得磁瓦表面的小缺陷难以区分,并且由于多层损失的引入使得网络深层对缺陷特征提取能力获得了进一步提升,于是本文在分类网络部分没有再进行降采样。具体的设计是在图 1 的蓝色虚线框内,将一阶段得到的分割网络最终输出的 512 维特征图通过两个残差块分别将最终通道合并为 128 维。此阶段不在进行最大池化操作减少信息丢失,并且使用信息流通更加便捷的残差块<sup>[15]</sup>进行特征的编码。

残差块的结构为在图 1 的黄色虚线框内,将输入分为两条路径后合并。一条为通过  $1 \times 1$  卷积、 $3 \times 3$  卷积和  $1 \times 1$  卷积,其中除最后的  $1 \times 1$  卷积未激活外,其余操作均经过批归一化层和非线性激活层并且通过第一个  $1 \times 1$  卷积合并通道数。另一条路径为直接通过  $3 \times 3$  卷积合并通道数并经过批归一化层,在残差块结尾将两个路径输出的特征图结果按通道求和并通过非线性激活层进行输出。经过两个这样的残差块合并通道特征编码后,最后将输出的 128 维度的特征图通过全局平均池化和全局最大池化,并合并为 256 维的向量通过全连接层进行分类。

分类训练时计算的损失函数为单类别基于 sigmoid 的二值交叉熵而非一般所使用多类别的 softmax 交叉熵,因为考虑到磁瓦表面可能存在多种类别缺陷共存的状态,所以选取各种类别间不相互排斥的单类别交叉熵作为损失函数。网络最终输出的结果为一个 6 维的向量,对向量取 sigmoid 激活后获得每个类别的概率分布,其中概率最大的就是网络预测的类别。损失函数具体的定义如下:

$$L_{cls} = \sum_i - [y_i \ln p_i + (1 - y_i) \ln(1 - p_i)], \quad (6)$$

其中  $y_i$  表示对应类别  $i$  的预测概率,  $p_i$  为类别的标签,样本类别如果是  $i$  则为 1,否则为 0,将样本所有类别的二值交叉熵求和的结果即为分类损失。

### 2.3 数据增强

由于实际场景下能够获取的磁瓦表面图像数据十分有限,在训练集过小时模型的鲁棒性会下降,并且有过拟合的风险。这时通过数据增强能够有效地扩充训练数据,增加数据的多样性,增强模型的泛化能力,提升模型的性能。因此本文在训练的过程中对数据进行在线增强,具体的操作如下:

在训练的过程中对图像和对应标注掩膜进行增强,具体的增强策略包括:(1)有 0.5 的概率随机上下翻转;(2)有 0.5 的概率随机左右翻转;(3)进行  $-5^\circ \sim 5^\circ$  的随机旋转;(4)有 0.5 的概率加高斯白噪声。在训练过程中,每轮对以上 4 种图像增强操作随机选择 3 种,测试集不进行数据增强。得到具体的效果如图 3 所示。

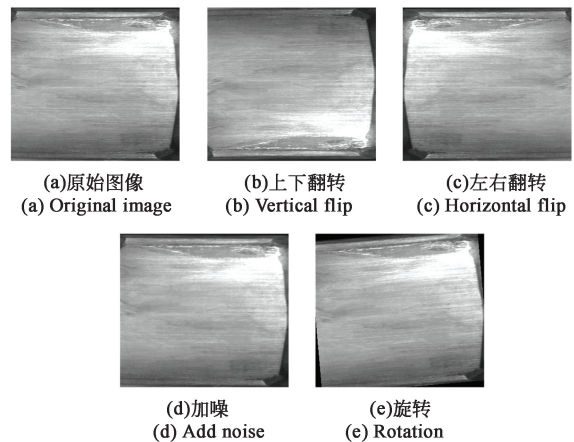


图 3 磁瓦图像增强效果

Fig.3 Magnetic tile image augment effect

### 3 实验和结果分析

实验环境基于开源深度学习框架 pytorch, 编程工具为 vscode。计算机的配置为 Ubuntu18.04 操作环境, 8 G 内存, 显卡为 GTX1060ti, 显存大小为 6 G。模型的训练是分阶段进行的, 将分割网络部分和分类网络部分分开训练。分割网络训练时模型会只输出分割网络的结果, 分类网络的权重不会参与更新。同样分类网络训练时将冻结分割网络部分的权重, 只对分类网络的权重进行更新。在训练阶段不载入任何预训练权重, 而是权重随机正态初始化。考虑到实际场景下的测试样本会远多于训练样本, 为了更好地验证模型的鲁棒性并且降低模型过拟合的风险, 对数据集进行 5 : 5 划分, 训练集包含 670 个样本, 测试集包含 674 个样本, 两者的各类别的数目分布相同。

#### 3.1 数据预处理

本文所使用的数据来自开源数据集 magnetic tile surface defects<sup>[18]</sup>, 其包含 5 种缺陷类别和无缺陷类别, 共 1 344 张。数据集的正负样本比例为大约 2 : 5, 无缺陷图片占多数, 数据分布很符合实际

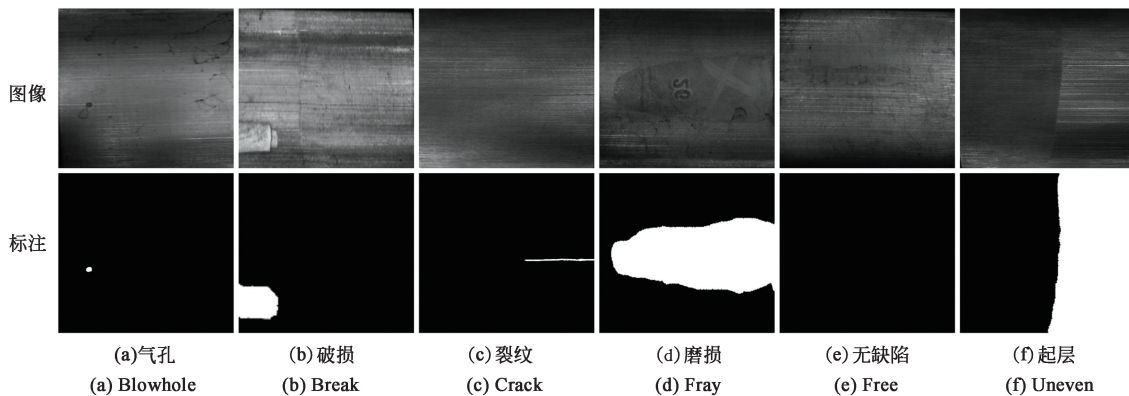


图 5 示例图像

Fig.5 Sample images

#### 3.2 评估指标

对于分割所使用的评估指标, 我们一般会用平均交并比 (Mean Intersection over Union, MIoU) 来衡量某个模型的分割效果。IoU 指的是两个区域的交集面积与并集面积的比例, 在缺陷分割中, IoU 指的是缺陷真实标签和预测值部分的交并的比值, 而 MIOU 指的是该数据集中每个

场景的工厂产出。所有图片的宽高比例均不一, 具体的宽高比例散点图和数目分布如图 4 所示。

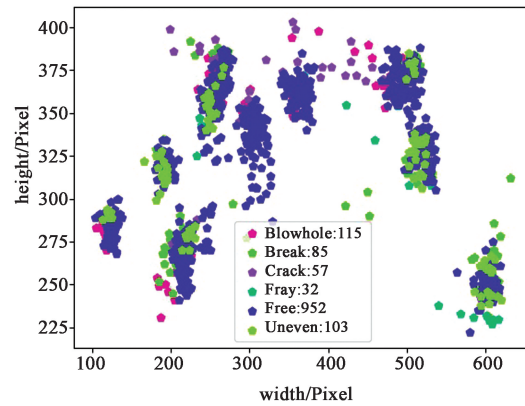


图 4 原始图像宽高分布散点图

Fig.4 Scatter plot of original image width and height distribution

针对原始图像宽高分布和网络的降采样倍率, 本文对所有图片和对应标注掩膜的图片先统一转置使得宽大于高, 然后再双线性插值为宽 352 像素、高 320 像素的图像。对应的预处理结果示例如图 5 所示, 数据集包含的磁瓦表面缺陷有气孔、破损、裂纹、磨损、起层 5 种缺陷和无缺陷类别。

类别样本的交并比的平均值。其定义如下:

$$\text{MIoU} = \frac{|TP|}{|FP| + |TP| + |FN|} \times 100\% \quad (7)$$

其中 TP 表示实际为缺陷区域以及预测正确为缺陷部位的区域, FP 表示实际为背景区域以及预测错误为缺陷部位的区域, FN 表示实际为缺陷区域以及预测错误为背景的区域。

考虑到在缺陷分割任务中, 对于缺陷图片和

无缺陷图片,我们所期望的分割结果是不相同的,对于缺陷图片更希望能够分割出缺陷,我们可以用分割出前景的 MIoU 来衡量分割的效果。但对于无缺陷图,我们更希望模型输出一个正常的结果,也即没有分割出任何前景。又由于实际场景的无缺陷图片是远多于缺陷图片的,如果单纯地放在一起进行统计会对总 MIoU 有很大的影响,从而对模型的指标评估产生偏差。所以本文对无缺陷图片所采取的衡量方法为单独取出评估,并且计算背景部分的 MIoU。

在实际的场景下,一般会更加关注与缺陷分割出的面积占缺陷本身的比例。于是定义缺陷分割准确率 PA(Pixel Accuracy, 像素准确率)为分割出的缺陷部位面积占所有缺陷面积的平均值,计算公式如下:

$$PA = \frac{|TP|}{|TP| + |FN|} \times 100\% \quad (8)$$

对于分类评估的指标除了总分类准确率外,想要衡量对于每个缺陷类别的预测精度,会使用召回率(Recall)和准确率(Precision)来衡量每个类别的预测结果,其各自的定义如下:

$$Precision = \frac{TP}{TP + FP}, \quad (9)$$

$$Recall = \frac{TP}{TP + FN}, \quad (10)$$

其中 TP 表示预测为对应类别实际也是该类别的样本, FP 表示预测为该类别实际并非该类别的样本, FN 表示预测并非该类别实际是该类别的样本。

为了平衡召回率和准确率,本文采用 F1-score 来衡量模型对每个缺陷类别的预测结果。F1-score 同时兼顾了分类模型的精确率和召回率,在形式上表现为每个类别召回率和准确率的调和平均数,其公式定义如下:

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (11)$$

### 3.3 分割效果

分割网络部分训练过程如下:设定初始学习率 0.01,优化器采用使用动量的 SGD,设定权重衰减率为  $5e-5$ 。训练采取动态学习率衰减策略,在训练轮数为 100, 200, 300, 400 时令学习率衰减 1/10,训练 500 轮后网络基本收敛。

为了验证添加辅助损失以及数据增强的效果,本文训练了 4 组模型,并且记录了在测试图片上的各种相关指标,具体结果如表 1 所示。

表 1 添加辅助损失以及数据增强效果对比

Tab.1 Effect comprison of add auxiliary loss and data augment

实验方法	缺陷类别 MIoU%	无缺陷类别 MIoU%	缺陷分割 准确率/%
分割网络	51.4	98.9	66.8
分割网络+差异系数损失	53.8	99.8	67.2
分割网络+差异系数损失+多层损失	60.2	99.9	77.6
分割网络+差异系数损失+多层损失+数据增强	68.4	99.9	94.5

根据消融实验的结果可以看到,原先分割网络对于缺陷分割的准确率较低,通过添加差异系数损失能够使得网络继续优化,从而提升 2.4% 的缺陷类别 MIoU,但是缺陷分割准确率只提升 0.4%,表明该损失函数贡献主要在减少网络对非缺陷部位的误分割。对比添加多层损失的效果,缺陷类别 MIoU 提升 6.4%,缺陷分割准确率提

升 10.4%,表明添加多层损失能够有效优化网络深层,明显提升网络分割效果。最后对比进行数据增强后的效果,缺陷类别 MIoU 提升 8.2%,缺陷分割准确率提升 16.9%,表明通过数据增强能够有效提升模型的鲁棒性,增强模型的泛化效果,解决数据量少、数据分布不均衡的问题,从而大幅提升分割精度。

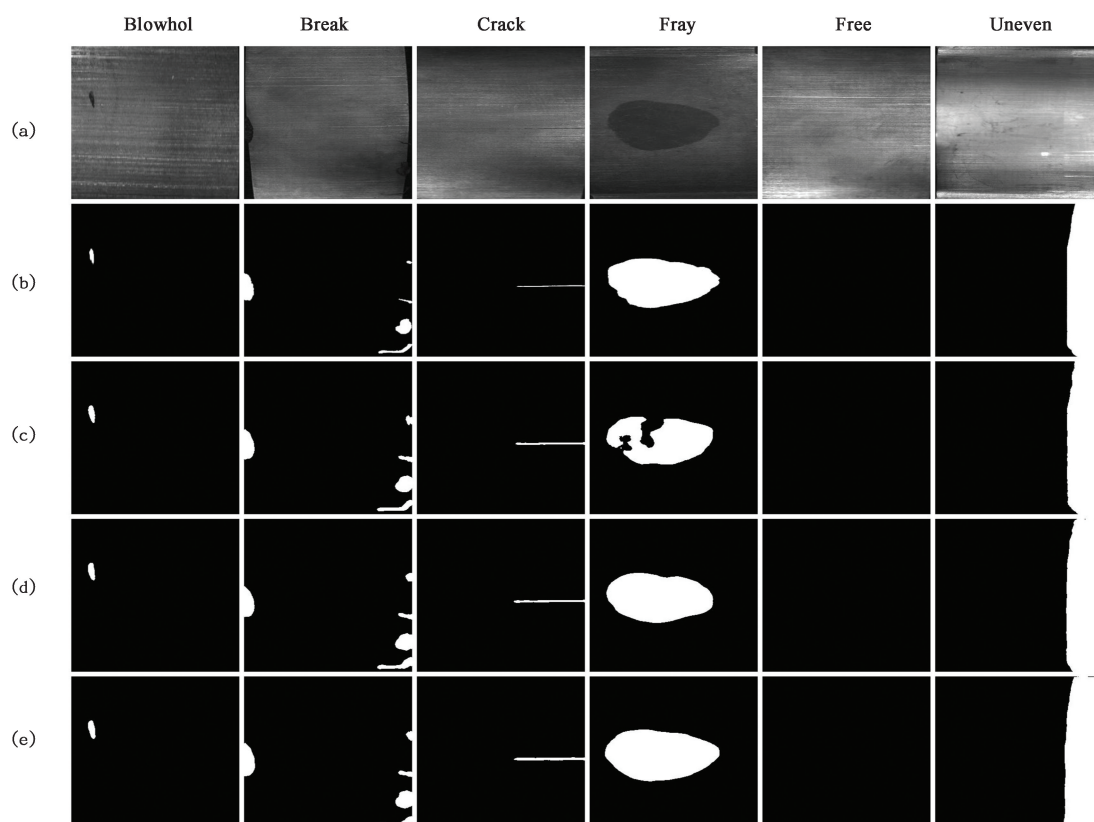


图 6 分割效果对比图。(a)原始图像;(b)标签掩膜;(c)原版分割网络;(d)添加辅助损失;(e)添加辅助损失与数据增强。

Fig.6 Comparison charts of segmentation effect. (a)Original image; (b)Label mask; (c)Original segment network; (d)Add auxiliary loss; (e)Add auxiliary loss and data augment.

分割网络消融实验效果对比如图 6 所示,可以明显看出,通过添加辅助函数后原先分割网络对于缺陷能够有效地分割出整体,通过数据增强训练后,分割的轮廓更加精细。本文训练使用的缺陷图片共 196 张,最后在测试的 196 张缺陷图片上获得了 94.5% 的缺陷分割准确率,表明本文提出的方法能够有效地分割出磁瓦表面缺陷,适应实际场景对于磁瓦表面缺陷分割的高精度要求。

### 3.4 分类效果

分类网络部分训练过程如下:设定初始学习率  $3e-4$ ,优化器采用 Adam,设定权重衰减率为  $5e-5$ 。冻结分割网络部分的权重,训练 50 轮后网络收敛。

为了验证本文通过分割提取特征进行分类方法的有效性,本文设置了对比实验,使用同样的训练集与测试集来训练常用的分类网络。为了验证本文方法相比刘畅等人的方法是否进一步提升了分类准确率,设置了和文献[8]的对比实验。并且

为了横向对比本文所提出的分类网络结构的有效性,与拥有同样类似结构的文献[10]进行对比,最终实验的结果如表 2 所示。

表 2 分类准确率对比

Tab.2 Comparison of classification accuracy

分类网络对比	分类准确率/%
Resnet-50	72.8
X-ception	89.4
文献[8]	93.2
文献[10]	97.1
本文方法	98.9

通过实验 1、2 与 3、4、5 的对比,在同样的图片预处理下没有通过一阶段分割网络直接使用 Resnet-50<sup>[15]</sup>、X-ception<sup>[16]</sup>等分类网络,最终分类准确率只有 70%~89%,表明大量的无缺陷图片对分类网络的精度有很大影响,并且由于大部分

缺陷占前景面积很小,网络很难过滤背景信息的影响,因此很难有效提取到缺陷的特征并进行区分。而本文的方法能够获得 98.9% 的准确率表明通过分割部分的强监督学习,网络能够有效过滤缺陷背景并且学习到缺陷的特征,避免了类别不均衡的影响,获得了很高的准确率。通过对比文献[8]和本文方法的实验结果,可以看到本文方法在分类准确率上提升了 5.4%,表明通过分割网络提取缺陷特征用于分类的方法能够避免因误分割

或分割不全所导致的误分类,从而进一步提升分类的准确率。

本文方法和文献[10]对比的各缺陷类别 F1-score 分布柱状图如图 7 所示。

由图 7 可知,本文方法所有的缺陷类别 F1-score 都能够到达 93% 以上,并且大部分高于文献[10]方法,对比可得应对该场景下本文的分类网络结构设计更加合理,能够获得更高的缺陷分类准确率,满足工业生产的高精度要求。

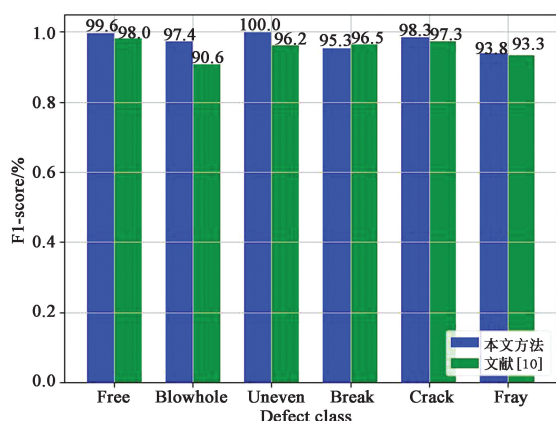


图 7 各缺陷类别 F1-score 对比柱状图

Fig.7 F1-score comparison histogram for each defect category

## 4 结 论

本文针对磁瓦生产工艺对磁瓦表面缺陷分割准确、缺陷分类精度高的要求,提出一种基于深度学习的缺陷分割与分类方法,该方法包括一个缺陷分割网络和分割网络上添加的缺陷分类网络。本文通过添加辅助损失函数有效地提升了分割精度,并使用在线数据增强的方式训练模型。消融实验表明,本文添加辅助损失函数和在线数据增强能够有效提升网络性能,使其分割出 94.5% 标注的缺陷区域。对比实验表明,本文方法相比其他分类方法能够获得 98.9% 的分类准确率,满足工业上磁瓦表面质量检测的高要求。

## 参 考 文 献:

- [1] LI X Q, JIANG H H, YIN G F. Detection of surface crack defects on ferrite magnetic tile [J]. *NDT & E International*, 2014, 62: 6-13.
- [2] 杨成立,殷国富,蒋红海,等.基于小波变换的磁瓦表面缺陷检测方法研究[J].*计算机应用与软件*,2014,31(11): 210-213,274.  
YANG C L, YIN G F, JIANG H H, et al. On magnetic tile surfaces defect detection based on wavelet transform [J]. *Computer Applications and Software*, 2014, 31(11): 210-213, 274. (in Chinese)
- [3] 李雪琴,蒋红海,刘培勇,等.非下采样 Contourlet 域自适应阈值面的磁瓦表面缺陷检测[J].*计算机辅助设计与图形学学报*,2014,26(4):553-558.  
LI X Q, JIANG H H, LIU P Y, et al. Defect detection on magnetic tile surface based on adaptive threshold surfaces in NSCT domain [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2014, 26(4): 553-558. (in Chinese)
- [4] 林丽君,殷鹰,何明格,等.基于小波模极大值的磁瓦裂纹缺陷边缘检测算法[J].*电子科技大学学报*,2015,44(2): 283-288.  
LIN L J, YIN Y, HE M G, et al. Edge detection algorithm of magnetic tile crack based on wavelet modulus maxima [J]. *Journal of University of Electronic Science and Technology of China*, 2015, 44(2): 283-288. (in Chinese)
- [5] WANG T, CHEN Y, QIAO M N, et al. A fast and robust convolutional neural network-based defect detection model in product quality control [J]. *The International Journal of Advanced Manufacturing Technology*, 2018, 94(9/12): 3465-3471.

- [6] MEI S, YANG H, YIN Z P. An unsupervised-learning-based approach for automated defect inspection on textured surfaces [J]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 67(6): 1266-1277.
- [7] 王春哲, 安军社, 姜秀杰, 等. 基于卷积神经网络的候选区域优化算法[J]. 中国光学, 2019, 12(6): 1348-1361.  
WANG C Z, AN J S, JIANG X J, *et al.* Region proposal optimization algorithm based on convolutional neural networks [J]. *Chinese Optics*, 2019, 12(6): 1348-1361. (in Chinese)
- [8] 刘畅, 张剑, 林建平. 基于神经网络的磁瓦表面缺陷检测识别[J]. 表面技术, 2019, 48(8): 330-339.  
LIU C, ZHANG J, LIN J P. Detection and identification of surface defects of magnetic tile based on neural network [J]. *Surface Technology*, 2019, 48(8): 330-339. (in Chinese)
- [9] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [J]. *arXiv*: 1505.04597, 2015.
- [10] 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.
- [11] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [J]. *arXiv*: 1502.03167, 2015.
- [12] MILLETARI F, NAVAB N, AHMADI S A. V-Net: fully convolutional neural networks for volumetric medical image segmentation [C]//2016 *Fourth International Conference on 3D Vision (3DV)*. Stanford, CA, USA: IEEE, 2016: 565-571.
- [13] ZENG G D, YANG X, LI J, *et al.* 3D U-net with multi-level deep supervision: fully automatic segmentation of proximal femur in 3D MR images [C]//8th *International Workshop on Machine Learning in Medical Imaging*. Quebec City, QC, Canada: Springer, 2017: 274-282.
- [14] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv*: 1409.1556, 2014.
- [15] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [J]. *arXiv*: 1512.03385, 2015.
- [16] CHOLLET F. Xception: deep learning with depthwise separable convolutions [J]. *arXiv*: 1610.02357, 2016.
- [17] SZEGEDY C, IOFFE S, VANHOUCHE V, *et al.* Inception-v4, Inception-ResNet and the impact of residual connections on learning [J]. *arXiv*: 1602.07261, 2016.
- [18] HUANG Y B, QIU C Y, GUO Y, *et al.* Surface defect saliency of magnetic tile [C]//2018 *IEEE 14th International Conference on Automation Science and Engineering (CASE)*. Munich, Germany: IEEE, 2018: 612-617.

#### 作者简介:



谢舰(1996—),男,江西赣州人,硕士研究生,2018年于江西理工大学获得学士学位,主要从事深度学习、图像处理方面的研究。E-mail:1653698129@qq.com



姚剑敏(1978—),男,福建莆田人,博士,副研究员,2005年于中国科学院长春光学精密机械与物理研究所获得博士学位,主要从事人工智能、图像处理、信息显示技术等方面的研究。E-mail:yaojm@fzu.edu.cn