

doi:10.3969/j.issn.1672-4348.2022.04.011

基于改进 UNet 的人造板表面缺陷的图像分割方法

张平均¹, 翁悦¹, 王小红¹, 李稳稳¹, 林艺斌²

(1. 福建工程学院 电子电气与物理学院, 福建 福州 350118;
2. 漳州鑫华成机械制造有限公司, 福建 漳州 363999)

摘要: 为满足人造板表面缺陷图像分割的精度要求, 提出了一种改进的 UNet 语义分割网络模型。在传统的 UNet 网络结构上将编码部分改进为残差网络 ResNet50 并去掉连接层与平均池化层, 网络通过残差块堆叠获取更多特征的底层信息; 同时在跳跃连接中嵌入聚焦注意力机制的模块, 抑制干扰信息, 保留有效位置信息, 聚焦缺陷区域并加强学习。对 4 种 UNet 网络模型的人造板表面缺陷图像分割进行仿真比较, 结果表明, 融合聚焦注意力机制的残差 UNet 网络模型在像素准确率和平均交并比等指标上有较大提升, 分割精度较高。

关键词: 聚焦注意力机制模块; UNet; 残差网络; 人造板表面缺陷; 图像分割

中图分类号: TP391.4

文献标志码: A

文章编号: 1672-4348(2022)04-0373-05

Image segmentation method of surface defects of wood-based panels based on improved UNet

ZHANG Pingjun¹, WENG Yue¹, WANG Xiaohong¹, LI Wenwen¹, LIN Yibin²

(1. School of Electronic, Electrical Engineering and Physics, Fujian University of Technology, Fuzhou 350118, China;
2. Zhangzhou Xinhua Cheng Machinery Manufacturing Co., Limited, Zhangzhou 363999, China)

Abstract: In order to meet the requirements of the precision of image segmentation of surface defects of wood-based panels, an improved UNet semantic segmentation network model was proposed. The coding part of the traditional UNet network was modified into residual network ResNet50, and the connection layer and average pooling layer were removed. The network was stacked with residual blocks to obtain more underlying information of features. At the same time, the module of attention focusing mechanism is embedded in the jump connection to suppress interference information, retain effective location information, focus defect location and enhance learning. The simulation comparison of image segmentation of surface defects of wood-based panels based on four UNet models shows that the residual UNet model integrating the attention focusing mechanism has been greatly improved in pixel accuracy and average intersection ratio, as well as the segmentation accuracy.

Keywords: attention focusing mechanism module; UNet; residual network; surface defects of wood-based panels; image segmentation

人造板在生产过程中由于原料的成分与配比、技术设备工艺等因素, 表面会出现油污、胶斑、裂纹和沙痕等缺陷^[1]。目前人造板表面缺陷检测主要还是依靠人工识别, 容易产生错检和漏检, 亟需开发基于机器视觉的人造板表面缺陷图像检测方法^[2]。

图像分割是缺陷检测过程中一个关键的步骤, 目的是在缺陷分类识别前将缺陷区域从人造板表面图像中分割出来。2018 年, 郭慧^[3]等人提出一种基于灰度共生矩阵分层聚类的缺陷提取算法, 根据缺陷区域与正常区域纹理不同的特点来提取缺陷, 但分割缺陷的过程较长, 需设定待提取

收稿日期: 2022-01-13

基金项目: 福建省科技创新项目(2020C0052)

第一作者简介: 张平均(1969—), 男, 福建福州人, 教授, 博士, 研究方向: 装备智能化。

的缺陷特征。2020 年,郭慧^[4]等人又提出了一种自适应快速阈值图像分割算法,通过改进 Otsu 算法能够根据当前图像给出最佳的类间分割阈值,但对于原料成分复杂、表面粗糙的人造板,缺陷的阈值与背景十分接近,无法达到预期检测效果。近年来,深度学习在图像处理领域表现出色,有不少研究者将图像的语义分割技术应用于材料缺陷检测^[5]。2019 年,He^[6]等人提出了一种混和全卷积神经网络(mix full convolutional net work, Mix-FCN)的方法来检测木材表面的缺陷,准确率达到 91.31%。UNet 是由全卷积神经网络(FCN)改进得到的网络,主要应用于医学图像细胞分割^[7]。2020 年,Rahman^[8]等人提出了一种融合多注意力机制的 UNet 卷积神经网络应用于太阳能电池缺陷检测,提高了检测的准确率。2021 年,谢舰^[9]等人基于 UNet 网络的架构添加辅助损失函数并进行在线数据增强,提升了磁瓦表面缺陷的分割准确率。上述所用 UNet 的特征提取网络都是基于 VGG16 (visual geometry group 16)网络构造的^[10],网络层数较少,上一层提取的缺陷特征不能充分表达达到下一层,并且随着网络加深,梯度不断消失,产生学习效率下降与准确率无法有效提高的问题,同时直连的跳跃连接方法无法突出显示目标区域信息,容易被干扰信息影响,混入噪声。

本研究针对缺陷特征提取与目标区域位置信息学习两个问题,提出一种基于注意力机制的残差 UNet 语义分割网络模型,对人造板的表面图像缺陷提取更深层次特征,增强模型泛化能力,提高缺陷分割的准确率与精度。

1 改进 UNet 的网络设计

1.1 特征提取增强的残差 UNet 网络

UNet 网络是一种从端到端的检测网络系统,主要包含上采样、下采样和跳跃连接。网络由两条对称路径构成,左半部分实现下采样过程,过程中图像不断进行卷积操作和下采样,以提取图像特征,获取特征信息,该过程称为编码过程。本研究选取 ResNet50(residual network 50)网络替换原始的 VGG16 特征提取网络部分,该网络是 2015 年由 He^[11]等人提出的残差网络模型,能有效缓解梯度弥散和网络退化两个问题,增强各层之间的特征融合,更好地保留缺陷区域的特征信息。ResNet50 的残差块包括 3 个卷积层,排序为:

1×1,3×3,1×1,还有线性单元 ReLU 和线性映射 Identity。ResNet50 第二层残差块内部结构如图 1 所示。

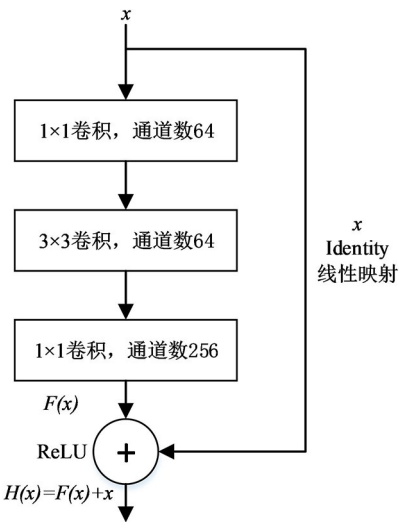


图 1 ResNet50 第二层残差块内部结构
Fig.1 Internal structure of the second layer's residual block of ResNet50

在残差结构中,Identity 可以把先前丢失的信息重新传入网络,此前的权值层信息 x 与此时的输出残差 $F(x)$ 叠加后再通过激活函数得到输出 $H(x)$ 输入到网络下一层,残差的表述如下:

$$F(x) = H(x) - x \tag{1}$$

ResNet50 特征提取部分通过 4 个层数的残差块的叠加构成,每层残差块叠加的数量依次为 3、4、6、3,如表 1 所示。

表 1 ResNet50 特征提取网络结构
Tab.1 ResNet50 feature extraction network structure

卷积层	输出尺寸	层内部结构
第一层	128×128	7×7 卷积,步长 2 3×3 最大池化,步长 2
第二层	64×64	残差块×3
第三层	32×32	残差块×4
第四层	16×16	残差块×6
第五层	8×8	残差块×3

1.2 增强位置信息学习的聚焦注意力机制
UNet 网络结构中的跳跃连接过程是将上采样过程中的特征图与下采样过程中的特征图融

合,融合方式是叠加特征图通道数。但在卷积操作过程中,特征图被不断压缩,缺陷的有效信息逐渐丢失。为了提高准确性,本研究在原有的跳跃连接上嵌入一种聚焦注意力机制模块 (attention focusing mechanism module) ^[12],模块内部如图 2

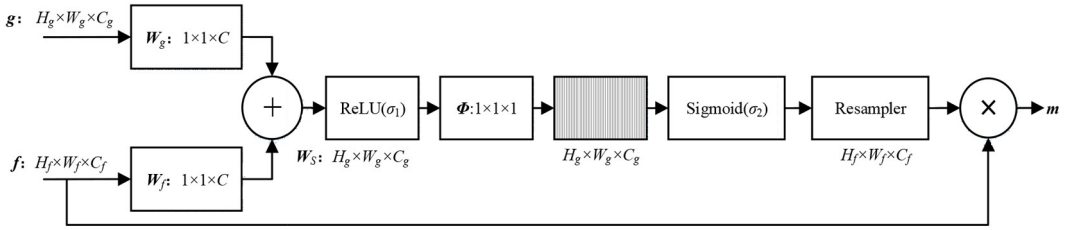


图2 聚焦注意力机制模块内部结构

Fig.2 Internal structure of the attention focusing mechanism module

图2中输入的 $g(H_g \times W_g \times C_g)$ 是下采样层同层特征图矩阵, $f(H_f \times W_f \times C_f)$ 是上采样层上一层的特征图矩阵,其中 H 、 W 分别为输入特征图的高度和宽度, C 为通道数,分别和特征权重矩阵 w_g 和 w_f 进行点乘运算得到 $w_g^T g$ 和 $w_f^T f$,再将得到的特征图相加得到新的特征权重 $w_s(H_g \times W_g \times C_g)$,如式(2)所示:

$$w_s = w_g^T g + w_f^T f \quad (2)$$

加权后的特征经过 ReLU 后再与 $1 \times 1 \times 1$ 卷积进行点乘运算得到注意力中间矩阵 q_{att} ,如式(3)所示:

$$q_{att} = \Phi^T(\sigma_1(w_s)) \quad (3)$$

式中的 σ_1 表示 ReLU 激活函数, Φ 表示 $1 \times 1 \times 1$ 卷积。 q_{att} 经过 Sigmoid 激活函数得到最终基于空间的注意力权重矩阵,如式(4)所示:

$$\alpha = \sigma_2(q_{att}(f; \delta_{att})) \quad (4)$$

式中的 σ_2 表示 Sigmoid 激活函数, δ_{att} 表示特征提取过程中的参数。最后把更新后的注意力权重矩阵和输入的原始特征图 f 点乘得到信息更丰富的输出特征图 m ,如式(5)所示:

$$m = \alpha \cdot f \quad (5)$$

1.3 融合聚焦注意力机制的残差 UNet 网络

本研究通过将左半部分编码过程中的特征提取网络换为提取深层次特征效果更好的 ResNet50 网络,将原始的 ResNet50 网络最后部分的全连接层和平均池化层去掉,保留特征提取结构,并在跳跃连接中嵌入一种聚焦注意力机制模块来抑制干扰信息,聚焦缺陷位置,保留目标区域有效信息,加强编码过程的特征与右半部分解码

所示。该模块能够提取多尺度图像特征,捕捉上下文信息,聚焦目标区域,增强对缺陷区域类别和位置的信息学习。提取的特征图随后通过跳跃连接进行合并,引入针对局部区域的注意力系数,抑制不相关的背景区域信息,突出显示缺陷区域。

过程(上采样)的特征进行对应拼接,扩大特征图尺寸,对目标精确定位,最后进行二分类或多分类的输出。将改进后的网络命名为融合聚焦注意力机制的残差 UNet (attention focusing residual UNet, Att-ResUNet),其网络结构示意图如图3所示。

2 实验分析

2.1 图像数据集

本次实验的人造板表面缺陷数据集由漳州鑫华成机械制造有限公司提供,共 52 张胶斑缺陷图片,为避免训练过程中过拟合现象的出现,利用数据增强对训练集进行扩充,通过对图片旋转角度、镜像翻转、等比例缩放、对比度调节方法,形成 468 张图片,并对所有图片的缺陷位置做好对应标注,训练集和测试集按 4:1 划分。

2.2 评价指标

本次实验选取像素准确率 (pixel accuracy, PA) 和平均交并比 (mean intersection over union, MIoU) 作为实验的评价指标。

像素准确率是一种比较简单的评价指标,它是通过预测类别正确的像素数占总像素数的比例来表示,其表达式为:

$$PA = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

其中, FN 表示被模型错误的分为负例,实际是正例; FP 为被模型错误的分为正例,实际是负例; TP 为被模型正确的分为正例,实际为正例; TN 为被模型正确的分为负例,实际为负例。

平均交并比是对 IoU 指标优化的结果, IoU 表

示一种类别预测结果和真实值的交集与并集的比例,MIoU 是计算每类的 IoU 的均值,其表达式为:

$$MIoU = \frac{TP}{TP + FP + FN}$$

(7)

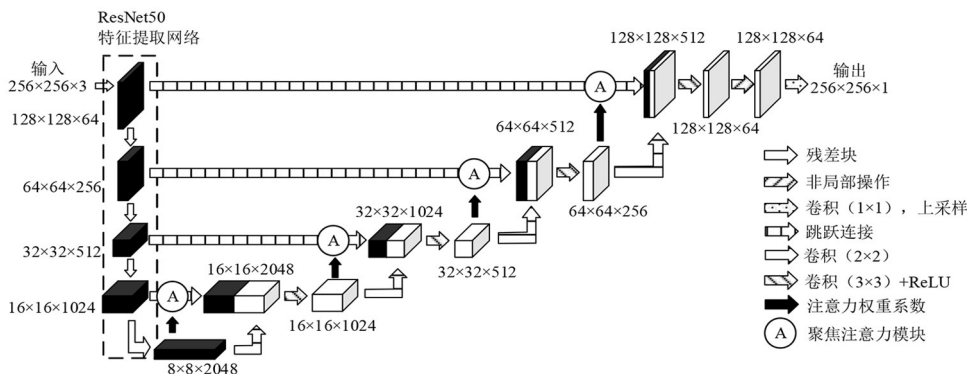


图 3 Att-ResUNet 网络结构

Fig.3 Att-ResUNet network structure

2.3 训练

为确保实验数据的准确性与有效性,本文实验都是在同一台基于 Windows10 操作系统电脑上完成,电脑的硬件环境和软件环境如表 2 所示。

表 2 实验环境

Tab.2 Experimental environment

名称	参数
系统	Windows 10
CPU	Intel(R) Core(TM) i5 2.40GHz
GPU	Nvidia GeForce GTX 1650(4 GB)
软件	PyCharm+Python+Pytorch
RAM	16GB DDR4

训练开始时将输入的图片尺寸统一处理成 256×256,训练过程中使用 Adam 算法进行优化,迭代 300 个训练周期,前 100 个周期学习率调为 0.001,100 个周期之后学习率调整为 0.000 1,动量为 0.9,采样批数据为 2。实验采用的二分类损失函数(binary cross entropy, BCE),公式如下:

BCE =

$$\frac{1}{N} \sum_{n=1}^N [y_n \times \log x_n + (1 - y_n) \times \log(1 - x_n)]$$

(8)

2.4 实验结果分析

为了验证设计模型的效果,本研究选取传统的 UNet、UNet 结合聚焦注意力模块(attention focusing UNet, Att-UNet)和特征提取网络更换为残

差网络的 UNet(residual UNet, ResUNet)3 种网络模型与 Att-ResUNet 模型进行实验对比,研究引入不同模块的作用。

表 3 所示为 4 种模型的实验结果。可见 UNet 分别引入残差网络和聚焦注意力机制模块的 2 种网络在像素准确率和平均交并比两个指标上较传统的 UNet 有小幅提升,而将 2 个模块同时融入到 UNet 中时,像素准确率提升了 4.49%,平均交并比提升了 12.53%,2 项评价指标达到了更优。

表 3 4 种模型实验指标对比

Tab.3 Comparison of experimental indicators of 4 models

模型	PA/%	MIoU/%
UNet	93.83	74.64
Att-UNet	95.36	79.64
ResUNet	95.92	81.14
Att-ResUNet	98.32	86.17

4 种模型对胶斑缺陷的分割效果如图 4 所示,传统 UNet 模型缺陷边界粗糙且模糊,引入残差网络的 ResUNet 模型增强了上下文信息表示,缺陷的边缘轮廓更为清晰,而引入注意力模块的 Att-UNet 模型有效消除了不必要的噪声,去除了非缺陷目标的干扰像素点,同时引入 2 个模块的 Att-ResUNet 模型分割出的区域图像更加连续、平滑。实验结果表明,本研究所提出的 Att-ResUNet 网络显著提高了缺陷分割的精度。

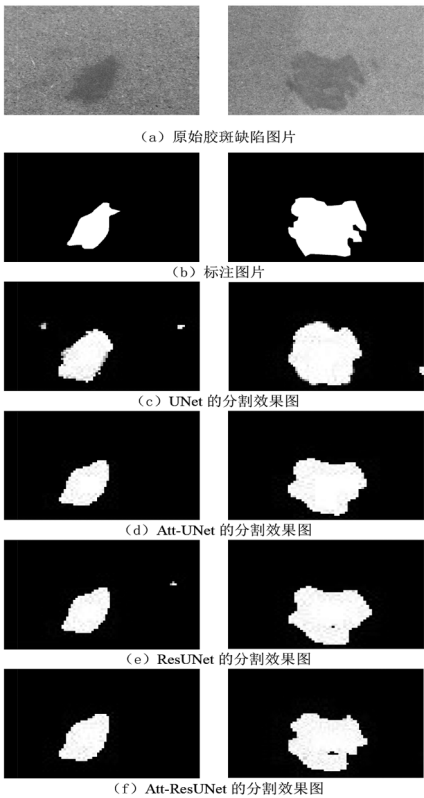


图 4 4 种模型胶斑缺陷分割效果对比

Fig.4 Comparison of the segmentation effect of glue spot defects of 4 models

参考文献:

[1] 朱志, 丁相君. 浅析人造板砂光后的表面缺陷[J]. 农业与技术, 2013, 33(11): 111.

[2] 魏智锋, 肖书浩, 蒋国璋, 等. 基于深度学习的人造板表面缺陷检测研究[J]. 林产工业, 2021, 58(2): 21-26.

[3] 郭慧, 王霄, 刘传泽, 等. 基于灰度共生矩阵和分层聚类的刨花板表面图像缺陷提取方法[J]. 林业科学, 2018, 54(11): 111-120.

[4] 郭慧, 王霄, 刘传泽, 等. 人造板表面缺陷检测图像自适应快速阈值分割算法[J]. 林业科学, 2018, 54(11): 134-142.

[5] 杨传礼, 张修庆. 基于机器视觉和深度学习的材料缺陷检测应用综述[J/OL]. 材料导报, 2022(16): 1-19 [2022-04-15]. <http://kns.cnki.net/kcms/detail/50.1078.TB.20210728.1056.009.html>.

[6] HE T, LIU Y, XU C Y, et al. A fully convolutional neural network for wood defect location and identification[J]. IEEE Access, 2019, 7: 123453-123462.

[7] 宫艳晶, 黄民, 黄小龙. 基于改进 ResNet-UNet 的火焰图像分割方法[J]. 北京信息科技大学学报(自然科学版), 2021, 36(5): 39-44.

[8] RAHMAN M R U, CHEN H Y. Defects inspection in polycrystalline solar cells electroluminescence images using deep learning[J]. IEEE Access, 2020, 8: 40547-40558.

[9] 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.

[10] 徐昭洪, 刘宇, 全吉成, 等. 基于 VGG16 预编码的遥感图像建筑物语义分割[J]. 科学技术与工程, 2019, 19(17): 250-255.

[11] 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, NV, USA: IEEE, 2016: 770-778.

[12] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: learning where to look for the pancreas[EB/OL]. 2018; arXiv: 1804.03999[cs.CV]. <https://arxiv.org/abs/1804.03999>.

(责任编辑: 方素华)

3 结语

针对人造板表面缺陷图像,传统算法对缺陷特征提取能力弱,无法准确学习目标区域位置信息的问题,本研究提出了一种基于改进 UNet 的语义分割网络模型,将原始 UNet 特征提取网络换成 ResNet50 的网络结构并去掉最后的平均池化层和全连接层,增强上下层有效特征的传递,更加有效地提取缺陷的细节特征,同时在原始网络的跳跃连接层中引入一种聚焦注意力机制模块,对人造板表面背景的噪声信息进行抑制、消除干扰,保存有效位置信息。

实验选取原始 UNet、Att-UNet 和 ResUNet 与本研究所提出的网络模型进行对比实验。结果表明,同时引入 2 个模块的 Att-ResUNet 网络在像素准确率和平均交并比两个指标上提升更高,且缺陷图像分割精度更优,验证了改进网络模型有效,达到了人造板表面缺陷图像分割的精度要求。