

基于卷积模块注意力机制深度学习模型的 激光熔覆裂纹识别

崔陆军, 李海洋, 郭士锐*, 李晓磊, 崔英浩, 郑博, 孙满盈

中原工学院机电学院, 河南 郑州 450007

摘要 为改进传统人工方法对熔覆区域裂纹检测耗时、准确率低的现状,提出了一种融合注意力模型的熔覆区裂纹自动识别方法,以便对裂纹进行标识和检测。基于 U-net 网络构造的熔覆裂纹语义分割网络存在对局部小特征提取能力不足的问题,而通过增加注意力模型(CBAM)层,提取特征空间和特征通道的权重信息,就可以对激光熔覆区微观裂纹进行实时的像素级标注和检测。实验结果表明:引入注意力模型的深度学习模型可使熔覆裂纹的识别和检测准确率提升 2.7 个百分点;融合注意力模型的网络在熔覆区域裂纹测试集上的准确率为 79.8%。深度学习模型标注的准确度和速度均已超过人工标注,为激光熔覆裂纹的识别提供了有效方法。

关键词 图像处理; 熔覆区裂纹; 卷积模块注意力机制; 语义分割; U-net 网络

中图分类号 TP302.1

文献标志码 A

doi: 10.3788/LOP202158.2014001

Laser Cladding Cracks Recognition Based on Deep Learning Combined Convolutional Block Attention Module

Cui Lujun, Li Haiyang, Guo Shirui*, Li Xiaolei, Cui Yinghao, Zheng Bo, Sun Manying
School of Mechanical and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou, Henan 450007, China

Abstract To improve the time consumption and low accuracy of traditional manual methods of laser cladding crack detection, an automatic identification method combined with the attention model is proposed to identify and detect laser cladding cracks. The semantic segmentation network of laser cladding cracks based on the U-net network cannot sufficiently extract small local features. By adding the convolutional block attention model (CBAM) layer to extract the feature space and feature channel weight information, we can label the microscopic cracks of the laser cladding zone without any time difference in the pixel level. Experimental results show that the deep learning model combined with the CBAM can improve the accuracy of cladding crack identification and detection by 2.7 percentages. The network fused with the CBAM achieves an accuracy of 79.8% on the cladding area crack test set. Both the labeling accuracy and speed of the deep learning model exceed those of manual labeling, providing an effective method for identifying laser cladding cracks.

Key words image processing; cladding area crack; convolutional block attention module; semantic segmentation; U-net network

OCIS codes 100.2960; 100.3008; 150.1835

收稿日期: 2020-10-06; 修回日期: 2020-12-04; 录用日期: 2020-12-27

基金项目: 河南省高等学校重点科研项目(20A460033, 20A460031)、河南省科技攻关计划项目(202102210068)、河南省自然科学基金面上项目(212300410422)、安徽理工大学矿山智能装备与技术安徽省重点实验室开放基金(KSZN202002003)

通信作者: *laser@zut.edu.cn

1 引言

激光熔覆技术利用预铺粉末或同步送粉等方式使金属粉末在高能量激光束的辐照下熔凝,使零件表面形成一层冶金涂层,从而提高零件的硬度、耐磨性、抗氧化性及耐腐蚀性^[1]。激光熔覆技术已在特种构件的加工制造以及航空航天等领域得到了广泛应用,熔覆层质量制约着激光熔覆技术的应用范围^[2-3]。熔覆层及基体中的残余应力大于材料的抗拉强度时易在裂纹尖端脆硬相处产生应力集中,进而导致熔覆层开裂。微观裂纹指宏观形貌下不可见或比较模糊而无法准确判断的裂纹,识别微观裂纹并对其进行标注可在熔覆加工结束后及时反馈熔覆层的质量。分析熔覆层的质量需要快速准确的裂纹识别方法。深度学习算法可以实现对目标的批量标注,对一定激光加工参数下的裂纹进行批量识别可以得到特定工艺下的裂纹率^[4-5]。基于深度学习的裂纹识别方法加速了裂纹的自动化检测进程。

卷积神经网络使用不基于知识系统的训练模式来获取样本的多种特征,在处理多特征交互样本方面表现优异,是目前图像分割中常用的算法^[6-7]。Qi 等^[8]分析了图像灰度梯度特征(Cv)与金属微观裂纹尺寸之间的关系,提出了基于 Cv 的精确的定量检测方法。混凝土表面裂纹背景简单且裂纹尺度较大,文献[9]采用阈值法对混凝土表面的裂纹进行了识别。阈值法适用于对背景简单且对比度大的图像进行边界检测,对复杂的熔覆金属裂纹图像识别时

易将金属枝晶与划痕等错误地识别为裂纹。Yun 等^[10]针对金属裂纹缺陷的特性提出了卷积变分自编码器(CVAE),该编码器提高了基于卷积神经网络的缺陷识别系统的鲁棒性;同时,他们采用自编码的数据库训练网络,并采用网络对线扫描相机采集的金属宏观裂纹等缺陷进行了识别。在金属缺陷的识别应用方面,神经网络主要用于背景与目标边界较明显的情况,较少用于显微镜照片中激光熔覆裂纹的识别。

激光熔覆层裂纹目标较小,不便于分割,因此,本文以 500 倍显微镜下的裂纹图像构建训练集,以增加裂纹的显著度。本文提出了通道和空间信息增强的卷积神经网络,并将其用于语义分割,增加裂纹训练的权重。在构建的像素标注训练样本上训练网络参数,实现裂纹特征的提取与裂纹标注。将一个图像输入网络中,网络便可以输出对应的裂纹标注。

2 裂纹识别网络的理论基础

本文针对 U-net 网络结构进行改进,提出了熔覆区裂纹识别网络。由于在通常的图像分割任务中,待分割区域的像素占比一般较大且像素分布较为均匀,所以经典的 U-net 网络针对所有像素无差别地进行训练。针对裂纹等小目标特征,引入注意力模型(CBAM)层构建图层的加权特征,以提高训练效率。

2.1 网络的基本结构

基于 U-net 设计的网络结构如图 1 所示,网络

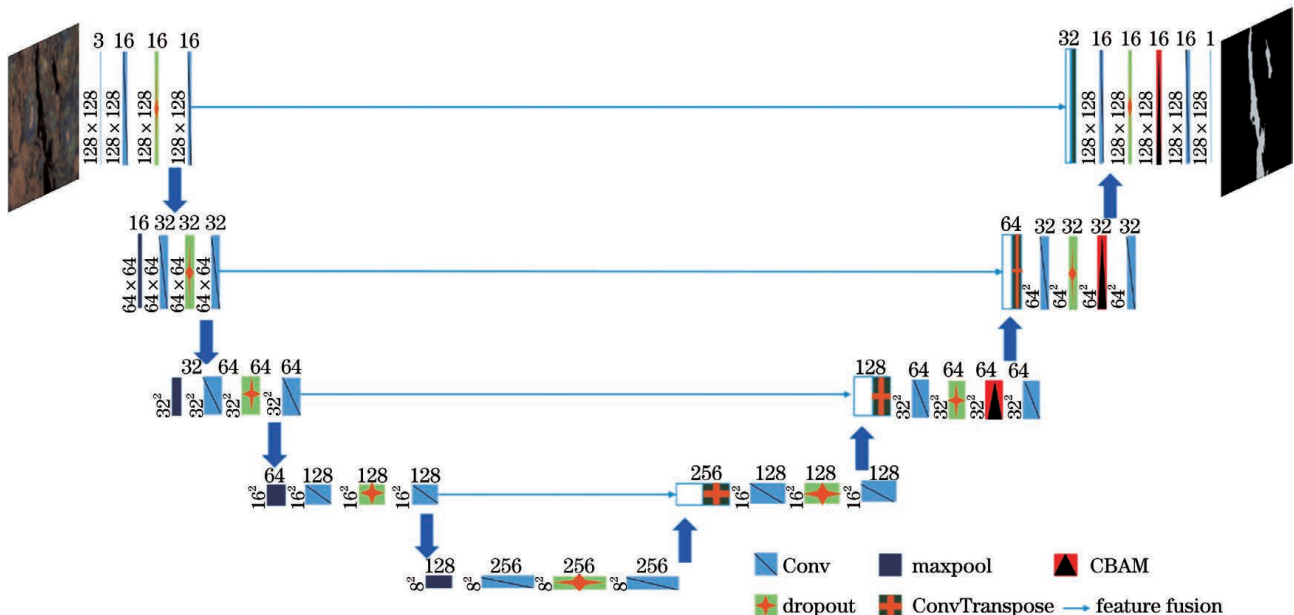


图 1 改进的 U-net 结构

Fig. 1 Improved U-net structure

分为下采样和上采样两部分。将图片 (128 pixel × 128 pixel) 输入网络, 在原图维度上进行两次卷积, 提取图像的基本特征。Dropout 层随机使一定比例参数为 0, 这样只需训练其余参数, 以便加快参数的训练速度, 增加网络的随机性, 使网络的泛化能力增强。特征信息存储在 16 层特征图上, 通过 maxpool 层去除提取的冗余信息, 特征图的大小减半, 层数不变。Maxpool 层后面的两个卷积层使特征图的层数翻倍。多次卷积后, 特征图尺寸为 8 × 8。两次卷积操作可使图像特征被提取至 256 层 8 × 8 的特征图上。上采样过程经过一个反卷积层使图层的尺寸翻倍, 然后与下采样的对应维度的图层信息进行一次特征融合。融合后的特征图经过两次卷积, 特征图层数减半, 最后在 128 × 128 × 1 的灰度图上展示裂纹标记。下采样与上采样一起组合成

为一个 U 型神经网络。与传统的 U-net 相比, 本文在图层的后三层加入了 CBAM 层, 以增加权重信息。两个融合层的尺寸相同, 不需要信息裁剪, 保留了上下图层的全部信息。

2.2 注意力模型层

增加网络的卷积层数和特征图厚度可以提高裂纹识别的准确率^[11], 但网络深度的增加会导致参数量和训练时间显著增加。CBAM 层通过计算特征图通道信息和空间信息参数分布得到通道层面和空间层面的权重信息。图 2 是 CBAM 层的结构图, 输入的特征图尺寸为 $H \times W \times C$, C 为通道数, H 为特征图高度, W 为特征图宽度, 在垂直于 C 轴方向和沿 C 轴方向分别进行通道运算和空间运算。将通道和空间注意力模型串联, 就可以输出具有权重的特征图。

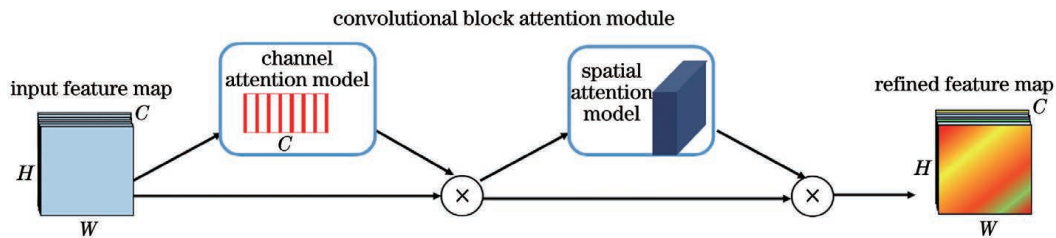


图 2 增加特征图训练权重信息的 CBAM 层

Fig. 2 CBAM layer for increasing feature map training weight information

通道注意力模型如图 3 所示。将特征图输入到模型中, 在每层特征图上分别作最大值池化和平均池化, 然后将池化结果输入到一个单隐层多层感知

器 (MLP) 中, 完成 channel 维度注意力特征的提取, 之后将两个输出量相加再通过激活函数就可以得到 $C \times 1 \times 1$ 的向量 (元素值代表相应通道的权重)。

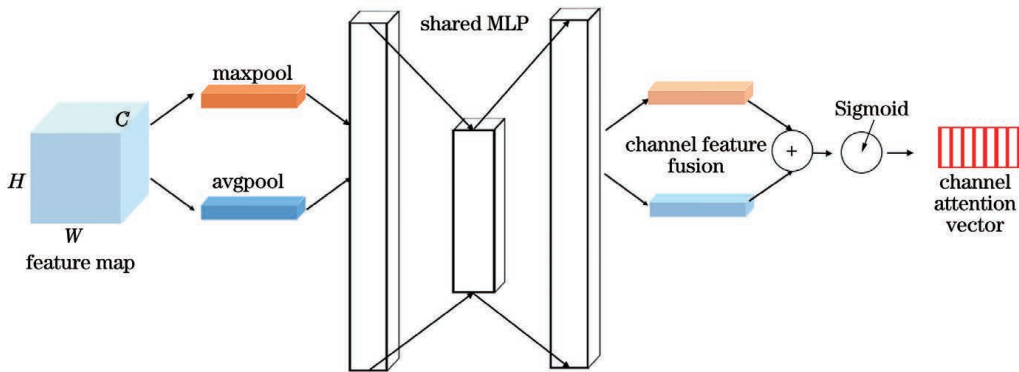


图 3 通道注意力模型结构

Fig. 3 Channel attention model structure

空间注意力模型如图 4 所示。将输入沿 channel 方向作最大值池化和平均池化, 然后将两层特征融合, 采用卷积神经网络提取特征图的空间注意力信息; 含有空间注意力信息的特征图层经激活函数归一化后, 就可以得到 $H \times W \times 1$ 的包含空间注意力信息的特征图。

将空间注意力模型与通道注意力模型串联, 可以得到每个像素都包含注意力信息的特征图层。特征图的输出维度与输入维度相同, 增加了特征图的权重信息, 从而可以提升模型的训练效率。

2.3 网络识别原理

确定了网络结构后, 待训练参数便随之确定。

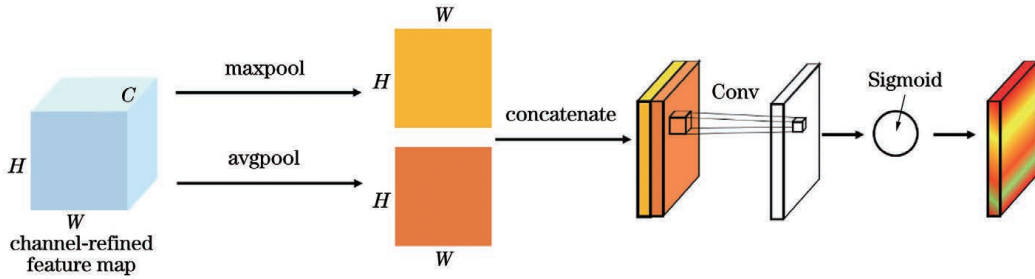


图 4 空间注意力模型结构

Fig. 4 Spatial attention model structure

网络训练的重点是损失函数和优化算法的选取。选用二元交叉熵 (BCELoss) 作为训练参数的损失函数。BCELoss 是衡量两个数学分布的交叉熵, 根据其值可以判断其为同分布的概率。 \hat{y}_n 为裂纹图片专家标注的序列, y_n 为对应的裂纹图神经网络自动标记后的序列。损失函数的表达式为

$$\text{BCELoss} = \hat{y}_n \cdot \exp(y_n) + (1 - \hat{y}_n) \times \exp(1 - y_n). \quad (1)$$

网络的训练过程可以表示为

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \{ \text{BCELoss}(\theta) \}, \quad (2)$$

式中: θ 是待训练的参数序列; θ^* 是使损失函数为全局最小值的参数序列。将裂纹图像的灰度矩阵输入到神经网络中, 以相应的裂纹专家标记序列作为输出期望, 通过损失函数 BCELoss 求解 θ^* 即可将图像分割问题转化为数学问题。

梯度下降算法选用 Adam 算法, 该算法的表达式为

$$\theta_t = \theta_{t-1} - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t, \quad (3)$$

式中: θ_{t-1} 为损失函数第 $t-1$ 次迭代后求得的网络参数; θ_t 为损失函数第 t 次迭代后求得的网络参数; η 为网络参数的初始学习率; \hat{v}_t 由前置参数更新幅度与当前参数的导数决定, 可对初始学习率进行动态调节; ϵ 为人为设定的常数值, 可确保参数的更新速度; $-\hat{m}_t$ 为包含所有网络参数的导数值的参数移动方向。

Adam 算法在回归参数时采用自适应学习率, 将前期网络参数求得的梯度作为当前网络参数迭代更新的一部分。将训练样本划分为训练集和验证集, 每次从训练集中批量抽取样本更新一次参数。

g_t 是每次更新参数时 BCELoss 对当前所有参数的导数向量, i 为总参数量。 g_t 的表达式为

$$g_t = \nabla \text{BCELoss}(\theta_t) = \begin{bmatrix} \partial \text{BCELoss}(\theta_1) / \partial \theta_1 \\ \vdots \\ \partial \text{BCELoss}(\theta_i) / \partial \theta_i \end{bmatrix}. \quad (4)$$

$-\hat{m}_t$ 是每次参数更新的方向。参数更新的方向由上一步参数移动方向和当前参数梯度共同决定。 \hat{m}_t 的表达式为

$$\hat{m}_t = \frac{\beta_1 \hat{m}_{t-1} + (1 - \beta_1) g_{t-1}}{1 - \beta_1^t}, \quad (5)$$

式中: β_1 依经验设置为 0.9。

\hat{v}_t 融合了前置梯度与当前梯度的平方和, 其表达式为

$$\hat{v}_t = \frac{\beta_2 \hat{v}_{t-1} + (1 - \beta_2) g_{t-1}^2}{1 - \beta_2^t}, \quad (6)$$

式中: β_2 依经验设置为 0.9。

\hat{v}_t 与初始学习率 η 结合形成动态学习率。可以依据多维参数的不同梯度动态调整学习速度, 以更快收敛到全局最优解。 ϵ 设置为 10^{-8} (以保持学习率始终为正)。

3 实验结果与讨论

激光熔覆裂纹识别系统主要由熔覆试样加工和裂纹识别两部分组成。完整的试样加工包括激光熔覆、线切割、样块镶嵌及金属抛光。完成试样加工后, 将样块放在显微镜下, 采集裂纹图片。实验中使用的光学显微镜的型号为 DM2700M, 其输出图片的像素为 1600×1200 。不同倍数显微镜镜头下的裂纹图片如图 5 所示。选择 500 倍下的裂纹图片作为神经网络的输入, 以增加小裂纹在图片中的显著度。

实验中使用的计算机操作系统为 Windows 10 教育版, CPU 型号为 Intel(R) Core(TM) i5-6500, ARM 内存为 8 GB, GPU 为 GTX 1660。显微镜下的裂纹识别网络基于 TensorFlow 框架和 Python 语言开发。TensorFlow 框架是 Google 公司开发的深度学习框架, 其在 GPU 上出色的分布式计算能力使其成为大数据运算的常用框架。编译器选用基于浏览器的 Jupyter notebook。本课题组基于激光

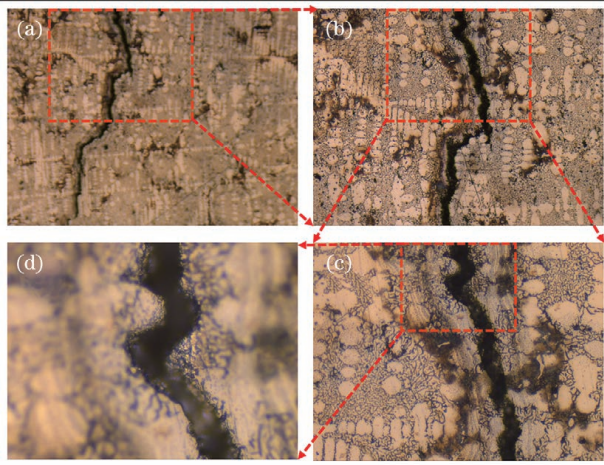


图 5 不同放大倍数下的裂纹形貌。(a)50 倍;
(b)100 倍;(c)200 倍;(d)500 倍

Fig. 5 Crack morphologies at different magnifications.
(a) 50 times; (b) 100 times; (c) 200 times;
(d) 500 times

熔覆层构建了包含 485 张熔覆区裂纹的数据库,裂纹图均由本课题组熔覆缺陷研究方向的研究生进行标注。随机选取 465 张图片作为训练样本,其余的 20 张图片用于测试。设置 batchsize 值为 16,每次批量从训练集中选取 16 个样本更新网络参数。批量化训练可以提高训练速度,同时会降低不良样本对参数的影响,批量化训练使得网络参数的更新更加稳定。隐层激活函数选择 Relu 激活函数,该函数汲取了批标准化操作(BN)的优点,缓解了梯度消失问题,降低了训练数据量。若网络在训练集上的准确率高,而在测试集上的准确率降低时,称为过拟合。采用 EarlyStopping 技术,求解经训练集训练后的函数在验证集上的 BCELoss 函数值,若 BCELoss 函数值在 15 步内不下降,则停止训练,以避免网络过拟合。

在图像语义分割领域,一般选用交并比(IoU)指标对分割结果进行评价,该指标表征网络自动分割图像与专家标记图像重合像素的个数与并集中像素总个数的比值。IoU 指标表征图像分割、定位的

精度^[12]。在混淆矩阵中,若某像素的预测值与专家标记相同,记为 True,否则记为 False。像素预测为裂纹像素的记为 Positive,否则记为 Negative。

表 1 混淆矩阵

Table 1 Confusion matrix

Expert mark	Prediction	
	True	False
True	True Positive (TP)	False Negative (FN)
False	False Positive (FP)	True Negative (TN)

IoU(α_{IoU})评价指标的表达式为

$$\alpha_{IoU} = \frac{\alpha_{TP}}{\alpha_{TP} + \alpha_{FN} + \alpha_{FP}}, \quad (7)$$

式中: α_{TP} 为将裂纹预测为裂纹的像素点数; α_{FN} 为将裂纹预测为非裂纹的像素点数; α_{FP} 为将非裂纹预测为裂纹的像素点数。

熔覆区裂纹图片中的裂纹为小目标,加入 CBAM 层可以增加网络对裂纹区域的训练权重。增加 CBAM 层不会显著增加网络的参数量,对训练速度的影响极小,可以提高模型的准确率。CBAM 层加入的位置会影响识别准确率:在浅层网络中添加 CBAM 层,裂纹识别的 IoU 准确率会下降;在深层网络中添加 CBAM 层,会大幅提升准确率。这与网络结构有关。在浅层网络中,被提取的图像特征较少,且是颜色、位置等浅层特征,加入 CBAM 层后会抑制部分特征,在后续提取中注意力效果会逐渐降低,最终导致准确率稍有下降。在深层网络中,特征图提取的特征增多,且包含更多的复杂特征,加入 CBAM 层可使重要特征得以加强,而次要特征被抑制。在网络后三层加入 CBAM 层后,网络的准确率上升了 2.7 个百分点,网络标记每张图片中裂纹的时间消耗为 23 ms。表 2 所示为在神经网络的不同位置(position)加入 CBAM 层后,在训练样本上训练神经网络时待训练的参数总量(parameters)以及训练结束后的测试准确率(accuracy)。

表 2 CBAM 加入的位置对训练参数量和准确率的影响

Table 2 Influence of the position of CBAM on the training parameter quantity and accuracy

Position	None	1	9	1,9	8,9	1,8,9	7,8,9	6,7,8,9
Parameters	1941105	1941368	1941368	1941631	1942051	1942314	1944342	1952921
Accuracy /%	77.1	76.9	78.7	79.2	79.7	79.1	79.8	79.8

训练时,设置训练步数为 70 次,每次从训练集中批量选取样本更新网络参数。每步训练完成后,

将验证集输入训练好的网络,分别求其 BCELoss 值和 IoU 值,通过验证集的 IoU 观察每次的参数更新

效果。再次将训练样本随机分割为训练集与验证集,并继续进行下次训练,直到训练停止,保存网络。将测试集导入到网络中验证网络的性能。网络在测试集上的 IoU 准确率为 79.8%。图像处理常用的自适应阈值法在相同测试集上的 IoU 准确率为 59.7%,测试集的实验结果如图 6 所示。采用自适应阈值法对裂纹进行标注时,标注结果易受背景信

息的干扰。与自适应阈值法标注相比,神经网络解读了裂纹的语义信息,并过滤了背景干扰,对金属图像中的裂纹分割得更精确。神经网络由数据驱动,依据不同图像裂纹分割任务中的裂纹尺度,可选择不同放大倍数下的裂纹图像作为训练样本。神经网络对裂纹的标注可以依据训练数据迁移至不同放大倍数激光熔覆裂纹的识别中。

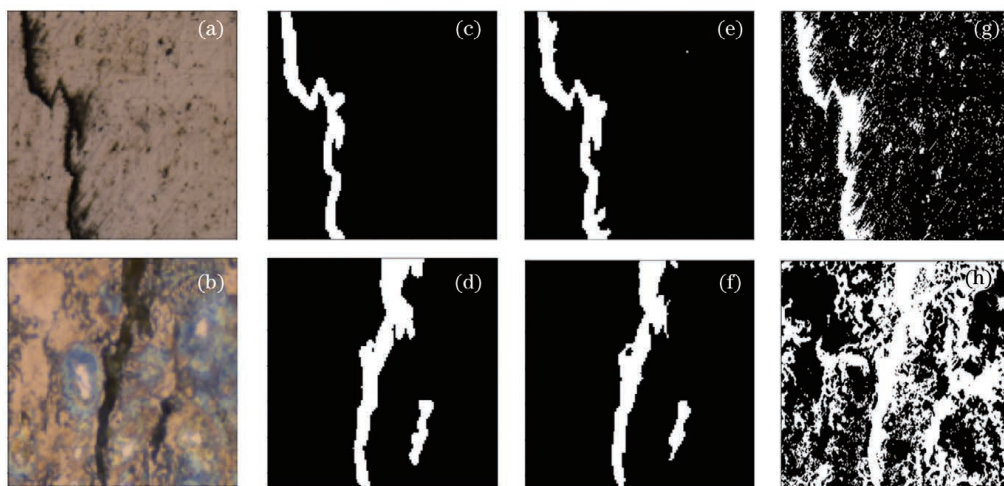


图 6 测试效果。(a)(b)待测图;(c)(d)人工标注;(e)(f)神经网络标注;(g)(h)自适应阈值法标注

Fig. 6 Test effect. (a)(b) Image to be tested; (c)(d) manual labeling; (e)(f) neural network labeling; (g)(h) adaptive threshold method labeling

4 结 论

本文基于 U-net 网络建立了激光熔覆裂纹自动标记神经网络。采集实验样本并将其分为训练集和测试集,在训练集中构建损失函数和回归算法,以训练网络参数;在测试集中采用 IoU 准确率评价网络对裂纹的自动识别精度。针对裂纹的小目标特性,引入注意力模型,以增加裂纹训练的权重。研究在网络的不同层引入注意力模型对网络性能提升的影响,结果表明:在深层网络添加注意力模型后,网络在熔覆区域裂纹测试集上的准确率为 79.8%(与未引入模型相比,准确率提升了 2.7 个百分点),标记每张图片耗时 23 ms。裂纹率与激光功率、送粉速率和扫描速度等工艺参数密切相关。通过该模型识别大量激光熔覆涂层图像得到不同工艺参数下的裂纹率,将其作为训练数据,然后将不同的激光加工工艺参数输入神经网络中,以工艺对应的裂纹率为输出,训练神经网络回归工艺参数与裂纹率之间的非线性关系,是下一步的研究工作。

参 考 文 献

- [1] Liu P, Chen Z K, Jin Q M, et al. Microstructure and corrosion behaviors of 316L coating fabricated by laser cladding[J]. Laser & Optoelectronics Progress, 2020, 57(3): 031402.
- [2] 刘朋, 陈志凯, 金全明, 等. 激光熔覆制备 316L 涂层组织及腐蚀行为研究[J]. 激光与光电子学进展, 2020, 57(3): 031402.
- [3] Fu F X, Zhang Y L, Chang G R, et al. Analysis on the physical mechanism of laser cladding crack and its influence factors[J]. Optik, 2012, 127(1): 200-202.
- [4] Zhou J Z, He W Y, Xu J L, et al. Strengthening mechanism and wear resistance of $Al_2O_3/Fe901$ composite coating prepared by laser cladding [J]. Acta Optica Sinica, 2019, 39(5): 0514001.
- [5] 周建忠, 何文渊, 徐家乐, 等. 激光熔覆 $Al_2O_3/Fe901$ 复合涂层的强化机制及耐磨性[J]. 光学学报, 2019, 39(5): 0514001.
- [6] Huang Y W. Research on large-scale photovoltaic panel detection technology based on UAV vision[D]. Nanning: Guangxi University, 2017: 2-4.
- [7] 黄钰雯. 基于无人机视觉的大规模光伏电池板检测技术研究[D]. 南宁: 广西大学, 2017: 2-4.
- [8] de Santis A, di Bartolomeo O, Iacoviello D, et al. Quantitative shape evaluation of graphite particles in ductile iron [J]. Journal of Materials Processing Technology, 2008, 196(1/2/3): 292-302.

[1] Liu P, Chen Z K, Jin Q M, et al. Microstructure and

- [6] Zhang Z H, Li B, Zhang W F, et al. Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks[J]. Journal of Manufacturing Systems, 2020, 54: 348-360.
- [7] Lu S, Tan Z W, Liu Y, et al. Realization of recognition for multi-mode optical fiber transmission speckle using neural network[J]. Acta Optica Sinica, 2020, 40(13): 1306001.
卢顺, 谭中伟, 刘艳, 等. 利用神经网络实现多模光纤传输散斑的识别[J]. 光学学报, 2020, 40(13): 1306001.
- [8] Qi Z C, Ni P J, Jiang W, et al. Quantitative detection of minor defects in metal materials based on variation coefficient of CT image[J]. Optik, 2020, 223: 165269.
- [9] Kim H, Ahn E, Cho S, et al. Comparative analysis of image binarization methods for crack identification in concrete structures [J]. Cement and Concrete Research, 2017, 99: 53-61.
- [10] Yun J P, Shin W C, Koo G, et al. Automated defect inspection system for metal surfaces based on deep learning and data augmentation [J]. Journal of Manufacturing Systems, 2020, 55: 317-324.
- [11] Wang S, Wu X, Zhang Y H, et al. Image crack detection with fully convolutional network based on deep learning [J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(5): 859-867.
王森, 伍星, 张印辉, 等. 基于深度学习的全卷积网络图像裂纹检测[J]. 计算机辅助设计与图形学学报, 2018, 30(5): 859-867.
- [12] Jiang B R, Luo R X, Mao J Y, et al. Acquisition of localization confidence for accurate object detection [M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11218: 816-832.