文章编号:1002-2082 (2022) 03-0532-06

基于深度学习激光熔覆层树枝晶的形貌识别

郭士锐,王凯祥,崔陆军,李晓磊,郑 博,陈永骞
(中原工学院 机电学院,河南 郑州 450007)

摘 要: 在增材制造技术中,树枝晶的表征对于分析激光熔覆层的机械性能至关重要,但目前树枝晶的标记主要由人工完成,耗时长且容易引入人为误差,而深度学习可提高目标识别准确度。基于 U-net 网络提出了适于识别分割树枝晶的 BNC-Unet 网络,将串行注意力机制和 Batch Normalization 层有效地部署在上采样和下采样区域,调整图像特征的权重信息。选取交并比作为分割结果的评价指标,并对比了原 Unet 以及不同的改进方法在该指标下的结果。在测试集中,BNC-Unet 网络分割树枝晶准确率指标为 84.2%,比原 U-net 网络结果提升了 8.97%。该指标表明网络能准确地从激光熔覆层金相图中识别出树枝晶形貌,且识别树枝晶的准确率显著提高,便于在激光熔覆试验后评估熔覆层性能。

关键词: 激光熔覆;语义分割;树枝晶;深度学习;串行注意力
中图分类号: TN249; TP302.1 **文献标志码:** A **DOI:** 10.5768/JAO202243.0307001

Morphology identification of dendrites of laser cladding layer based on deep learning

GUO Shirui, WANG Kaixiang, CUI Lujun, LI Xiaolei, ZHENG Bo, CHEN Yongqian
(School of Mechanical and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China)

Abstract: In the manufacturing technology of additive materials, the characterization of dendrites is crucial for analyzing the mechanical properties of laser cladding layer. However, the labeling of the dendrites is mainly completed manually at present, which is time-consuming and easy to introduce artificial errors, while the deep learning can improve the accuracy of target recognition. Based on the U-net network, the BNC-Unet network which was suitable for the identification and segmentation of dendrites was proposed. The serial attention mechanism and the Batch Normalization layer were effectively deployed in the upsampling and downsampling regions to adjust the weight information of image features. The intersection over union (IoU) was selected as the evaluation index of the segmentation results, and the results of original U-net network and different improved methods under this index were compared. In the test set, the segmentation accuracy index of BNC-Unet network for dendrites is 84.2%, which is 8.97% higher than the results of original U-net network. The index shows that the BNC-Unet network can accurately identify the morphology of dendrites from metallographic diagrams of laser cladding layer, and the accuracy of dendrites identification is significantly improved, which is convenient for evaluating the properties of cladding layer after the laser cladding test.

Key words: laser cladding; semantic segmentation; dendrites; deep learning; serial attention

---

收稿日期: 2021-11-26; **修回日期:** 2022-02-22
基金项目: 中国纺织工业联合会科学技术指导项目 (2016085);河南省研究生教育改革与质量提升工程项目 (YJS2022AL057);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);中原工学院科研团队发展项目 (K2021TD002);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项目 (KSZN202002003)
作者简介: 郭士锐(1986—),男,博士,副教授,主要从事激光制造方面的研究工作。E-mail: laser@zut.edu.cn

引言

激光熔覆技术是将高功率密度的激光束的能量, 通过非接触方式使粉末材料加热并熔化, 熔化后的粉末材料快速冷却与微观结构的形成, 使得表面具有与基体材料本身不同的新性能[1-4]。在激光熔覆过程中, 急冷急热的能量变化过程会对微观组织的形成产生极大的影响, 而最终形成的微观组织与材料的性能有着紧密的联系[5-6]。在熔覆层凝固结晶时, 最常见的是以树枝状生长的枝晶, 且枝晶不同的形貌参数与其性能有着一定的对应关系[7]。因此, 实现自动识别枝晶形貌, 对研究熔覆层性能具有十分重要的价值。基于上述背景分析, 本文提出了一种从复杂的金相图中识别并分割出枝晶形貌的深度学习算法。

近些年, 基于深度学习的计算机视觉检测在不同工况均展现了其优异的能力[8-10]。特别是用于像素级别分类的语义分割更是在工业、医学等领域得到了大量的应用[11-12]。Li 等人利用收发声信号技术监测激光熔覆过程, 并通过深度学习神经网络提取特征向量, 从而分析熔覆状态以及是否存在产生裂纹的迹象[13], 该研究表明可以利用深度学习技术分析熔覆层状态及微观形貌。张海军等人基于遗传算法、二维最大类间方差提出双阈值分割法[14], 根据区域极大值初步确定阈值, 并对金相图像进行分割, 分割结果反馈调整阈值, 实现了对不同金相组织的分割。为本文实现在复杂背景中分割出枝晶形貌提供了一种新思路。熔覆层金相图中树枝晶与不规则晶、胞状晶等混杂在一起[15], 不利于对树枝晶形貌进行观察研究。将树枝晶形貌识别视为二分类问题, 通过语义分割的方法把树枝晶从复杂的金相图中分割出来, 便于后续研究枝晶形貌与熔覆层性能的具体联系。

针对复杂背景的二分类问题, 本文对原 Unet 网络加以改进, 引入串行放置的通道注意力机制、空间注意力机制和批量标准化(Batch Normalization, BN)层, 并部署在下采样和上采样过程, 建立了 BNC-Unet(BN+CBAM-Unet)(convolutional block attention module, CBAM)即枝晶形貌分割网络模型。本网络模型用于树枝晶形貌的识别分割, 在观测树枝晶形貌时排除其他晶粒的干扰, 有效地促进了树枝晶形貌对熔覆层性能影响的研究。

1　树枝晶形貌识别网络

语义分割是深度学习三大应用之一, 其特点是效率和精度高[16-17]。搭建适于识别并分割样本特征的深度学习模型, 建立训练集及测试集, 用以训练、测试模型, 并根据网络反馈优化参数, 最终可以快速、准确得到分割结果, 节省大量的时间。本研究以 U-net 网络模型为基础, U-net 网络本身是为了解决医学影响的细胞分割而提出, 便于辅助医护人员客观、准确地分析病理, 为做出准确的医学诊断提供支撑数据[18]。由于其出色完成了在复杂背景下分割样本目标而得到了众多领域专家的青睐[19]。针对样本背景复杂问题, 在原来的 U-net 网络基础上, 部署串行注意力模型和 BN 层, 并调整 2 种模型的部署位置, 使得本次改进的 BNC-Unet 模型能更好地胜任树枝晶形貌分割任务, 取得较为满意的分割结果。

1.1　总体网络架构

本次设计的树枝晶形貌识别网络主要分为 2 个部分, 其结构如图 1 所示。一是用于提取特征的下采样部分, 即编码区; 二是得到更加准确特征信息的上采样部分, 即解码区。树枝晶识别网络整体架构呈现字母 U 型, 接下来依次介绍 2 个区域。

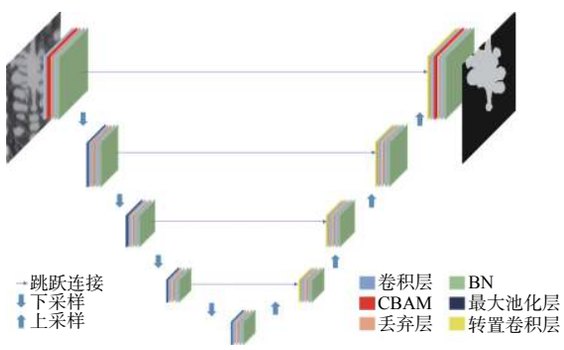


图 1　BNC-Unet 网络结构图

Fig. 1　Structure diagram of BNC-Unet network

第 1 部分是模型的编码部分, 用于提取金相特征, 该部分的各子过程大体包括了 2 个卷积层(其中卷积核为 3×3)、防止模型过拟合的丢弃层、用于对每层数据标准化处理的 BN 层、elu 激活函数层、最大池化层。在该部分的第一阶段引入 CBAM, 使得提取到的特征更加精准。将 3 通道 128 像素×128 像素的原始图像输入编码区, 经过卷积等操作生成 4 组通道数与大小不同的特征图, 其通道数和大小如表 1 所示。这些特征图用于与解码区中相同通道数、相同大小的特征图进行跳跃融合。经编码区最后一次处理得到 256 通道 8 像素×8 像素的特征图, 并传递至解码区。

表 1　特征图指标

Table 1　Feature map indexes

组别	通道数	大小 /(像素×像素)
1	16	128×128
2	32	64×64
3	64	32×32
4	128	16×16

第 2 部分是凝练特征信息的解码部分，该部分的每个阶段大体包括了用于扩大图像尺寸的 ConvTranspose2d 逆卷积层（卷积核为 2×2、步长为 2）、特征融合（解码区内的特征图与编码区内的特征图通道数、尺寸大小对应相等）、2 个卷积层（其中卷积核为 3×3）、丢弃层、BN 层、elu 激活函数层。在该部分的最后一个阶段引入 CBAM，细化特征，调整权重。由编码区传递的 256 通道 8 像素×8 像素的特征图经过上采样过程，最终生成单通道 128 像素×128 像素的灰度图像。

1.2　串行注意力机制原理

串行注意力机制，即将输入的特征图先经过通道注意力模块，然后将该模块的输出结果加权传递至空间注意模块，经过这 2 个模块的作用，最终将结果输出作为下一过程的输入。这种注意力机制模型在 2018 年由 Woo[20] 提出，并对比了单独使用这 2 个模块和空间注意模块加通道注意力模块的布局，结果表明先通道后空间得到的特征信息最为准确，其对目标特征的识别覆盖最广。串行注意力模型结构如图 2 所示，通道注意力模型与空间注意力模型串行放置，使得提取特征更为准确。

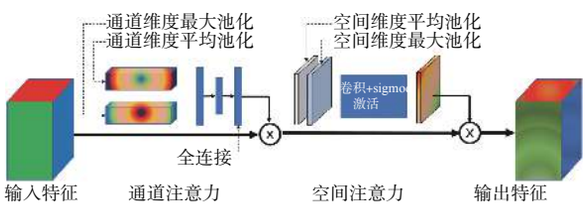


图 2　串行注意力模型结构图

Fig. 2　Structure diagram of serial attention model

在通道注意力模型内，焦点在于通道维度，特征图的每个通道即为各通道包含的特征。提取通道的池化参数，再经过同一个多层感知机（Multilayer Perceptron, MLP），输出值相加后经过归一化处理得出最终权重，得到的权重与特征图的乘积即为通道注意力模型输出的加权特征图。

在空间注意力模型内，焦点在于特征图中每个像素的权重。与通道注意力模型类似，在空间注意力模型中，提取不同通道的同一平面空间点的池化参数，经卷积和归一化处理得出最终权重，得到的权重与空间特征的乘积即为空间注意力模型输出的加权特征图。

特征图经过串行注意力模型后，特征权重得到进一步调整处理，有利于提高网络对样本特征学习、提取的能力，从而提高网络的性能。

1.3　网络参数的选取

网络参数的选择主要是 Batch Size 的确定、优化器的选取和损失函数的选取。

1.3.1　Batch Size 的确定和优化器的选取

Batch Size 的大小对网络的训练速度和梯度具有一定的影响，其优点在于可以有效地防止内存爆炸，加速收敛，避免局部最优。梯度方差如（1）式所示：

$$Var(g) = Var\left(\frac{1}{m}\sum_{i=1}^{m} g(x_i, y_i)\right) \tag{1}$$

式中：m 为 Batch Size。因为样本数量固定且随机，所以各样本方差相同，（1）式可简化为

$$Var(g) = \frac{1}{m} Var(g(x_i, y_i)) \tag{2}$$

由（2）式可以看出，梯度方差与批量大小成反比，一般 Batch Size 取值范围在 2～32 之间[21]，本次实验取 Batch Size 大小为 14，epoch 设置为 75。

学习率对于神经网络训练起到非常重要的作用，本次选用 Adam 优化算法。该算法整合了 AdaGrad 和 RMSProp 两种算法的优势[22]，与单独使用这 2 种算法相比，使用 Adam 算法收敛最快。该算法属于自适应学习率算法，其算法策略如下：

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t, \hat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{3}$$

式中：m_t为一阶动量项；$\hat{m}_t$为修正值；根据经验设定 β_1为 0.9。

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2, \hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{4}$$

式中：v_t为二阶动量项；$\hat{v}_t$为修正值；根据经验设定 β_2为 0.999。

$$W_{t+1} = W_t - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon} \hat{m}_t \tag{5}$$

式中：W_t为迭代模型参数；取ε值为 10^{-8}。Adam 优化器在收敛速度和准确率方面表现优秀，故本网络模型选择 Adam 优化器。

1.3.2　损失函数的选取

本次分割任务属于二分类问题, 故 *Loss* 损失函数选用二进制交叉熵。由于树枝晶特征与其背景中其余晶粒特征相差不大, 故本次设计的网络模型的输出层激活函数选用 Sigmoid 函数, 其函数原理如(6)式所示:

$$f(x)=\frac{1}{1+\exp(-x)} \tag{6}$$

从(6)式可以看出, 该函数可以把数值控制在 0～1 区间, 故能良好地应用于二分类问题。

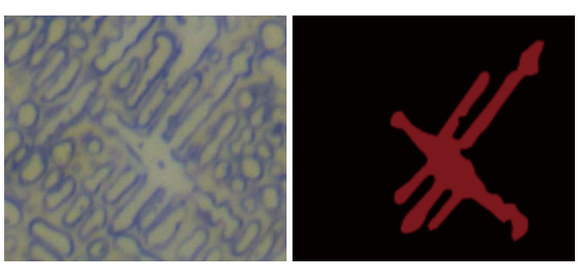
二元交叉熵损失函数如(7)式所示:

$$Loss=-\frac{1}{N}\sum_{i=1}^{N}\{y_i\cdot\log[p(y_i)]+(1-y_i)\cdot\log[1-p(y_i)]\} \tag{7}$$

式中: $p(y_i)$是标签y_i(0 或 1)出现的概率; N为样本数。故该文采用 binary cross entropy 作为损失函数, 且搭配 Sigmoid 使用。

2　实验结果与讨论

为了检验本次设计的基于语义分割树枝晶形貌识别网络的分割能力, 进行了激光熔覆实验, 并采集了训练网络所需的样本。本次的训练集和测试集均是由课题组激光熔覆样块在光学显微镜下捕获所得, 光学显微镜型号为 DM2700M(leica microsystems, GmbH)。使用型号为 LDP6000-60(laserline, GmbH)大功率光纤激光器进行熔覆实验, 将熔覆样块进行线切割、镶块处理, 经打磨抛光和王水腐蚀后在光学显微镜下捕获金相图像, 至此激光熔覆实验结束, 后续需要在计算机上进行下一阶段实验。利用 Labelme 软件为金相图制作标签后, 建立金相数据集, 其中包括 420 张训练图像和 20 张测试图像, 原金相图和标签如图 3 所示。



(a) 金相图　　(b) 标签

图 3　数据集

Fig. 3　Data set

本次研究使用的计算机操作系统为 Windows 10, CPU 为 Intel®Core™i5-10200H, RAM 为 8.00 GB, GPU 为 GTX1650Ti。网络搭建、运行、调试选用 Tensorflow 框架, 选择基于网页版编译器 Jupyter notebook 和 Python 语言编译开发, 该编译器可以直接在网页上调试代码, 程序结果直接在网页中显示, 便于直接观察。为了防止网络模型训练时出现过拟合现象, 本设计引入用于提前终止训练的 Early Stopping 命令, 并设置耐心值为 20 步, 即训练过程中 *Loss* 值出现连续 20 步无下降现象, 模型训练停止。

实验结果的评价指标选用交并比(intersection over union, IoU), 其表达了网络分割像素与标签像素总数上的比值。通常以混淆矩阵作为基础, 再计算评价指标数值。混淆矩阵如图 4 所示, A 为真实值, B 为预测值。预测与标记相同为 True; 相反为 False。预测值为树枝晶像素记 Positive; 反之为 Negative。

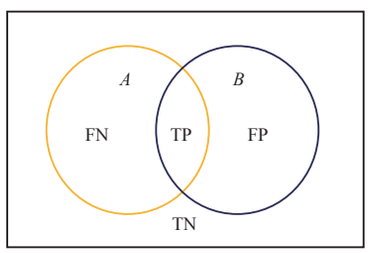


图 4　混淆矩阵示意图

Fig. 4　Schematic diagram of confusion matrix

IoU 计算方式为

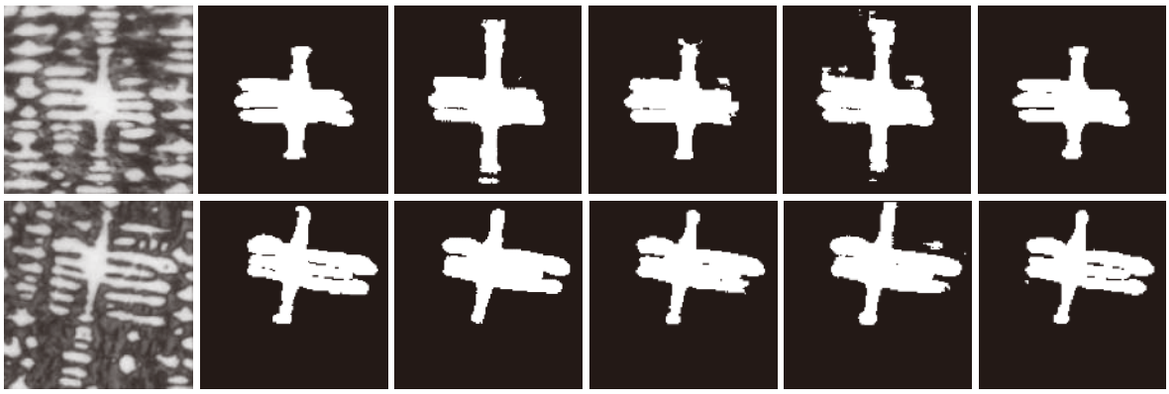
$$IoU=\frac{TP}{TP+FP+FN} \tag{8}$$

本网络共训练了 378 张图片, 验证了 42 张图片, 并做了 3 组对比试验, 其结果如表 2 所示。设定相同参数, 对比了 Unet、BN-Unet、CBAM-Unet、BNC-Unet 输出结果如图 5 所示。从表 2 可以看出本次针对树枝晶特征设计的 BNC-Unet 模型结果最好, 其 IoU 值为 84.20%, 比原 Unet 模型结果提高了 8.97%。从图 5 中的分割结果看, BNC-Unet 自动分割的树枝晶结果最为精确。

表 2　不同改进方式的网络结果 *Loss*、*IoU* 对比

Table 2　Comparison of Loss and IoU network results of different improved methods

网络架构	*Loss*	*IoU*/%
Unet	0.066 6	75.23
BN-Unet	0.065 8	81.05
CBAM-Unet	0.066 2	75.37
BNC-Unet	0.059 1	84.20



(a) 测试图像 (b) 标签 (c) 原Unet (d) BN-Unet (e) CBAM-Unet (f) BNC-Unet

图 5 各网络输出对比图

Fig. 5 Comparison graphics of each network output

3 结论

本实验基于 U-net 网络引入注意力模型和 BN 层，搭建了针对分割树枝晶的 BNC-Unet 网络，建立了用于训练网络和测试网络的数据集，将交并比 IoU 作为本次网络的评价指标，并对比了原 Unet 网络、BN-Unet 网络、CBAM-Unet 网络的分割结果。其对比结果表明 BNC-Unet 网络能够较为准确自动标记树枝晶形貌，BNC-Unet 网络分割准确率为 84.2%，比原 Unet 网络结果提高了 8.97%。注意力机制在浅层和深层对称部署一次可以在实现较好的特征提取结果的基础上而不增加网络训练时间，实现自动标记树枝晶形貌，便于在进行激光熔覆试验后，为自动分析熔覆层性能提供参考。

参考文献：

[1] 周圣丰，曾晓雁，胡乾午. 金属陶瓷复合涂层的激光熔覆与无裂纹的实现[J]. 应用光学，2008，29(1)：76-80.
ZHOU Shengfeng, ZENG Xiaoyan, HU Qianwu. Realization of laser cladding and crack-free ceramic-metal composite coatings[J]. Journal of Applied Optics, 2008, 29(1): 76-80.

[2] 张志强，杨凡，张天刚，等. 激光熔覆碳化钛增强钛基复合涂层研究进展[J]. 表面技术，2020，49(10)：138-151.
ZHANG Zhiqiang, YANG Fan, ZHANG Tiangang, et al. Research progress of laser cladding titanium carbide reinforced titanium-based composite coating[J]. Surface Technology, 2020, 49(10): 138-151.

[3] 张群莉，王梁，梅雪松，等. 激光表面改性技术发展研究[J]. 中国工程科学，2020，22(3)：71-77.
ZHANG Qunli, WANG Liang, MEI Xuesong, et al. Development of laser surface modification technology[J]. Strategic Study of CAE, 2020, 22(3): 71-77.

[4] 王强，李洋洋，杨洪波，等. 激光功率对17-4PH丝材激光熔覆组织及硬度的影响[J]. 表面技术，2021，50(3)：191-197.
WANG Qiang, LI Yangyang, YANG Hongbo, et al. Effect of laser power on laser cladding structure and hardness of 17-4PH wire[J]. Surface Technology, 2021, 50(3): 191-197.

[5] ZHANG Shitang, ZHOU Jiansong, GUO Baogang, et al. Friction and wear behavior of laser cladding Ni/hBN self-lubricating composite coating[J]. Materials Science and Engineering A, 2008, 491(1/2): 47-54.

[6] ASTA M, BECKERMANN C, KARMA A, et al. Solidification microstructures and solid-state parallels: recent developments, future directions[J]. Acta Materialia, 2009, 57(4): 941-971.

[7] 袁森，徐锦锋，时惠英，等. 灰铸铁中γ枝晶结晶形貌与机械性能的关系[J]. 热加工工艺，1996，25(4)：14-16.
YUAN Sen, XU Jinfeng, SHI Huiying, et al. Relationship between the crystal morphology of austenitic dendrites and mechanical property in gray lron[J]. Hot Working Technology, 1996, 25(4): 14-16.

[8] 范宏，侯云，李柏林，等. 基于区域特征的缺陷扣件视觉检测[J]. 铁道学报，2021，43(8)：132-138.
FAN Hong, HOU Yun, LI Bailin, et al. Visual inspection of defective fasteners based on regional features[J]. Journal of the China Railway Society, 2021, 43(8): 132-138.

[9] 周治国，刘开元，郑翼鹏，等. 一种基于深度学习的高速无人艇视觉检测实时算法[J]. 北京理工大学学报，2021，41(7)：758-764.
ZHOU Zhiguo, LIU Kaiyuan, ZHENG Yipeng, et al. A

real-time algorithm for visual detection of high-speed unmanned surface vehicle based on deep learning[J]. Transactions of Beijing Institute of Technology, 2021, 41(7): 758-764.

[10] 张炳力, 秦浩然, 江尚, 等. 基于RetinaNet及优化损失函数的夜间车辆检测方法[J]. 汽车工程, 2021, 43(8): 1195-1202.
ZHANG Bingli, QIN Haoran, JIANG Shang, et al. A method of vehicle detection at night based on RetinaNet and optimized loss functions[J]. Automotive Engineering, 2021, 43(8): 1195-1202.

[11] 宁芊, 胡诗雨, 雷印杰, 等. 基于多尺度特征和注意力机制的航空图像分割[J]. 控制理论与应用, 2020, 37(6): 1218-1224.
NING Qian, HU Shiyu, LEI Yinjie, et al. Segmentation of aerial image with multi-scale feature and attention model[J]. Control Theory & Applications, 2020, 37(6): 1218-1224.

[12] 何康辉, 肖志勇. LRUNet: 轻量级脑肿瘤快速语义分割网络[J]. 中国图象图形学报, 2021, 26(9): 2233-2242.
HE Kanghui, XIAO Zhiyong. LRUNet: a lightweight rapid semantic segmentation network for brain tumors[J]. Journal of Image and Graphics, 2021, 26(9): 2233-2242.

[13] LI Kaiqiang, LI Tao, MA Min, et al. Laser cladding state recognition and crack defect diagnosis by acoustic emission signal and neural network[J]. Optics & Laser Technology, 2021, 142: 107161.

[14] 张海军, 王春光, 翟改霞. 用小波分析进行基于遗传算法的2DMBV图像分割[J]. 自动化技术与应用, 2008, 27(4): 15-18.
ZHANG Haijun, WANG Chunguang, ZHAI Gaixia. 2DMBV image segmentation based on genetic algorithm with wavelet analysis[J]. Techniques of Automation and Applications, 2008, 27(4): 15-18.

[15] 徐一飞, 孙耀宁, 王国建, 等. 高速激光熔覆铁基合金涂层的组织及性能研究[J]. 中国激光, 2021, 48(10): 222-230.
XU Yifei, SUN Yaoning, WANG Guojian, et al. Microstructure and properties of iron-based alloys coatings prepared by high-speed laser cladding[J]. Chinese Journal of Lasers, 2021, 48(10): 222-230.

[16] 孟琭, 徐磊, 郭嘉阳. 一种基于改进的MobileNetV2网络语义分割算法[J]. 电子学报, 2020, 48(9): 1769-1776.
MENG Lu, XU Lei, GUO Jiayang. Semantic segmentation algorithm based on improved MobileNetV2[J]. Acta Electronica Sinica, 2020, 48(9): 1769-1776.

[17] 张哲晗, 方薇, 杜丽丽, 等. 基于编码-解码卷积神经网络的遥感图像语义分割[J]. 光学学报, 2020, 40(3): 46-55.
ZHANG Zhehan, FANG Wei, DU Lili, et al. Semantic segmentation of remote sensing image based on encoder-decoder convolutional neural network[J]. Acta Optica Sinica, 2020, 40(3): 46-55.

[18] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention(MICCAI 2015), Berlin, German: Springer, 2015: 234-241.

[19] 张文倩, 冯金超, 李哲, 等. 基于Unet的切连科夫激发的荧光扫描断层重建[J]. 中国激光, 2021, 48(17): 129-140.
ZHANG Wenqian, FENG Jinchao, LI Zhe, et al. Reconstruction for Cherenkov-excited luminescence scanned tomography based on unet network[J]. Chinese Journal of Lasers, 2021, 48(17): 129-140.

[20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision – ECCV, Berlin, German: Springer, 2018, 2018: 3-19.

[21] MASTERS D, LUSCHI C. Revisiting small batch training for deep neural networks[J]. arXiv, 2018, 1804: 07612.

[22] 饶竹一, 张云翔. 基于BiGRU和注意力机制的多标签文本分类模型[J]. 现代计算机, 2020(1): 31-35.
RAO Zhuyi, ZHANG Yunxiang. Multi-label text classification model based on BiGRU-attention neural network[J]. Modern Computer, 2020(1): 31-35.