

基于改进 UNet 网络的烧结矿气孔分割^①



周思雨, 储岳中, 张学锋

(安徽工业大学 计算机科学与技术学院, 马鞍山 243032)

通信作者: 周思雨, E-mail: 1535141658@qq.com

摘 要: 在烧结矿生产过程中, 烧结矿形成的气孔是烧结矿的质量评估的重要参数. 由于烧结矿的气孔形状不一、气孔边缘模糊等问题, 导致分割出的气孔误差率较大. 为了能更准确地分割出气孔, 先对烧结矿图像进行 OpenCV 图像预处理. 对比传统的图像分割算法, 本文提出一种基于改进 UNet 网络对预处理后的烧结矿气孔图像进行分割的算法. 在 UNet 网络编码中引入残差和拼接连接结合思想的改进模块, 以获得更多的气孔特征信息. 实验结果表明, 改进的算法在 MIoU 和 Dice 指标均优于传统 UNet 网络 and 传统图像分割.

关键词: 深度学习; 图像处理; 烧结矿图像; 烧结矿气孔; 语义分割; UNet 网络

引用格式: 周思雨, 储岳中, 张学锋. 基于改进 UNet 网络的烧结矿气孔分割. 计算机系统应用, 2022, 31(7): 278–284. <http://www.c-s-a.org.cn/1003-3254/8559.html>

Segmentation of Sinter Pores Based on Improved UNet Network

ZHOU Si-Yu, CHU Yue-Zhong, ZHANG Xue-Feng

(School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China)

Abstract: In the process of sinter production, the pores formed in the sinter are an important parameter for sinter quality evaluation. Due to the different shapes of sinter pores and their fuzzy edge, the error rates of pores segmented are large. In order that the pores can be segmented more accurately, the sinter image is preprocessed by OpenCV. Through a comparative analysis of the traditional image segmentation algorithm, this study proposes an algorithm based on an improved UNet network to segment the preprocessed sinter pore image. An improved module integrating residual and splicing is introduced into UNet network coding to obtain more information on pore features. The experimental results show that the improved algorithm is better than both the traditional UNet network and the traditional image segmentation algorithm in mean intersection over union (MIoU) and the Dice index.

Key words: deep learning; image processing; sinter image; sinter pores; semantic segmentation; UNet network

在烧结矿生产中, 烧结矿的 FeO 含量是烧结矿强度的一个关键因素. 当 FeO 含量较高时, 会导致烧结矿还原性能变差, 从而影响烧结矿的强度. 一个烧结矿的气孔数量和面积可以判断其 FeO 含量, 因此检测烧结矿气孔数量和面积是很重要的研究方向. 传统的检测气孔数量和面积的方法是通过人工进行的, 由于烧结矿数量多, 导致工作量大, 使得检测困难. 因此, 很多小型企业只是对气孔进行粗略的观察, 并没有进行正规

检测和记录.

随着科学技术的发展, 自动化以及信息化等技术逐步应用到工业领域中. 各种图像分割技术应用到工业中, 文献 [1,2] 都是利用图像分割, 对矿石进行粒度检测. 文献 [3] 利用图像分割技术进行钢铁的缺陷检测. 针对传统图像分割, 常用的方法有阈值分割^[4]、边缘检测^[5]、形态学分水岭算法^[6]、Hough 变换算法和轮廓查找算法^[7] 等. 由于烧结矿气孔的图像存在气孔

^① 基金项目: 安徽省重点实验室研究项目 (TZJQR002-2021); 安徽高校自然科学研究项目 (KJ2019A0063)

收稿时间: 2021-09-30; 修改时间: 2021-10-25; 采用时间: 2021-11-08; csa 在线出版时间: 2022-05-17

形状不规则、轮廓不清晰、气孔相互粘连,增加了气孔的分割难度,从而出现了很多改进传统的分割算法的研究,如针对粘连图像进行分水岭算法的改进^[8].近年来,运用卷积神经网络对图像进行分割,具有很大的优势并取得了较大的进展^[9,10]. Shelhamer等^[11]提出的全卷积网络 (FCN) 使得语义分割得到了快速发展,该网络相比传统的 CNN 分割的网络更加高效,减少了像素块带来的冗余存储和计算. Ronneberger等^[12]提出的 UNet 网络主要用于生物医学图像的分割,使用小数据集训练得到不错的结果,使得很多研究者关注. Zhou等^[13]提出了 UNet++使用密集跳跃连接,通过叠加不同层次的特征,减小了编码特征和解码特征之间的语义鸿沟. 全真等^[14]提出结合残差网络的方式,形成 Res-UNet 网络进行图像轮廓优化. 文献 [15] 结合特征层密集连接的方式,提出 Dense-UNet 模型用以改善分割性能. 这些网络都是在 UNet 网络基础上进行改进,存在着在上采样时冗余和丢失部分特征,不适用于烧结矿气孔细节特征分割.

为了获得更高的准确度,本文提出先对原始烧结矿气孔图像进行 OpenCV 图像处理^[16-18] 后进行图像分割. 为减少烧结矿图像分割气孔信息的丢失,在原始的 UNet 上结合残差连接和拼接连接的思想提出改进的编码器,加强了气孔轮廓边缘特征的提取.

1 图像预处理和传统的图像分割

1.1 图像预处理

由于受到现场的光线、灰尘等因素干扰,需要对图像进行去噪处理. 为减少干扰,只进行单个烧结矿的采集和处理. 图像预处理的过程如图 1 所示.

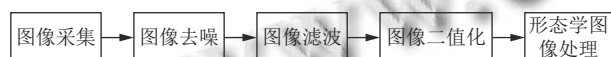


图 1 图像预处理的过程

通过对烧结矿图像进行像素相加处理,如图 2(a),消除部分噪声,提高像素亮度. 使用非线性滤波器中值滤波处理,如图 2(b),中值滤波对于消除图像中的随机噪声和脉冲噪声具有较好的效果,又能很好的保护细节信息^[19]. 之后将图像二值化,如图 2(c),增强了烧结矿图像中气孔和其他地方的对比度,看出烧结矿图像上仍存在噪声信息. 最后通过运用形态滤波,开运算和闭运算联合,有效的去除二值图像中存在的前景和背景噪声,并在一定程度上平滑了边界,如图 2(d).

1.2 传统图像分割

本实验利用传统的边缘检测算法和分水岭分割算法对烧结矿图像进行分割,基于边缘检测的图像分割算法是通过检测包含不同区域的边缘来解决分割问题,是最先想到也是研究最多的方法之一. 分水岭算法是一种基于拓扑理论的数学形态学的分割方法. 通过实验,边缘检测算法对图像中的噪声进行了分割,将部分背景也分割成气孔. 分水岭算法存在着过度分割图像的现象. 利用形态学的处理,结合开运算和闭运算的操作,减少噪声再进行分水岭分割,仍然存在烧结矿气孔像素信息相近的背景进行了分割. 开运算是先对图像先进行腐蚀然后膨胀,闭运算是先对图像先进行膨胀后进行腐蚀. 式 (1) 是开运算的定义. 式 (2) 是闭运算的定义.

$$X \circ Y = (X \odot Y) \oplus Y \quad (1)$$

$$X \cdot Y = (X \oplus Y) \odot Y \quad (2)$$

其中, X 代表原图, Y 代表结构元, $\odot$ 是腐蚀操作, $\oplus$ 是膨胀操作.

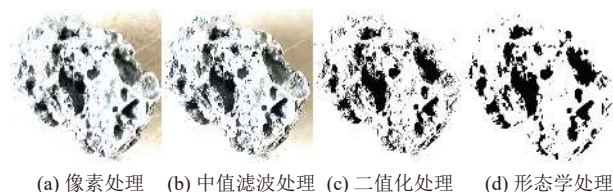


图 2 图像预处理

如图 3 所示. 由于传统的分割方法需要找到最佳阈值,当分割不同烧结矿图像时,都要改变阈值才能得到更好的效果,这个过程极其复杂. 传统分割算法不具有灵活性,导致分割出来的气孔误差率较大. 所以不能很好地适用于烧结矿气孔的分割.

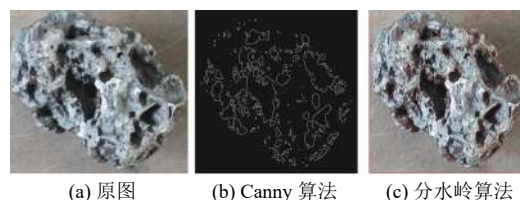


图 3 传统图像分割

2 网络结构与算法原理

2.1 UNet 网络

UNet 网络最早是应用在医学影像分割中,相比于另外几种经典的语义分割网络具有训练数据量少并且

精度高的特点,符合烧结矿气孔分割的要求.图4为UNet网络结构.网络由编码器和解码器两个部分组成.编码部分是通过Conv3×3卷积、ReLU激活函数和2×2最大池化对输入图片进行降维和特征提取,使得特征图的尺寸逐渐减小、特征维度逐渐增加.解码部分通过对特征图进行2×2转置卷积以及1×1卷积,使特

征图片尺寸增加,维度增加,加强了气孔特征的细节提取.解码器与编码器之间通过跳跃连接进行特征图融合,使得不同层气孔特征融合连接到解码相对应层中,其中浅层特征用来分割气孔,深层特征用来定位气孔的位置,二者的有效结合有助于增加气孔特征的多样性.因此,UNet网络非常适用于分割烧结矿气孔图像.

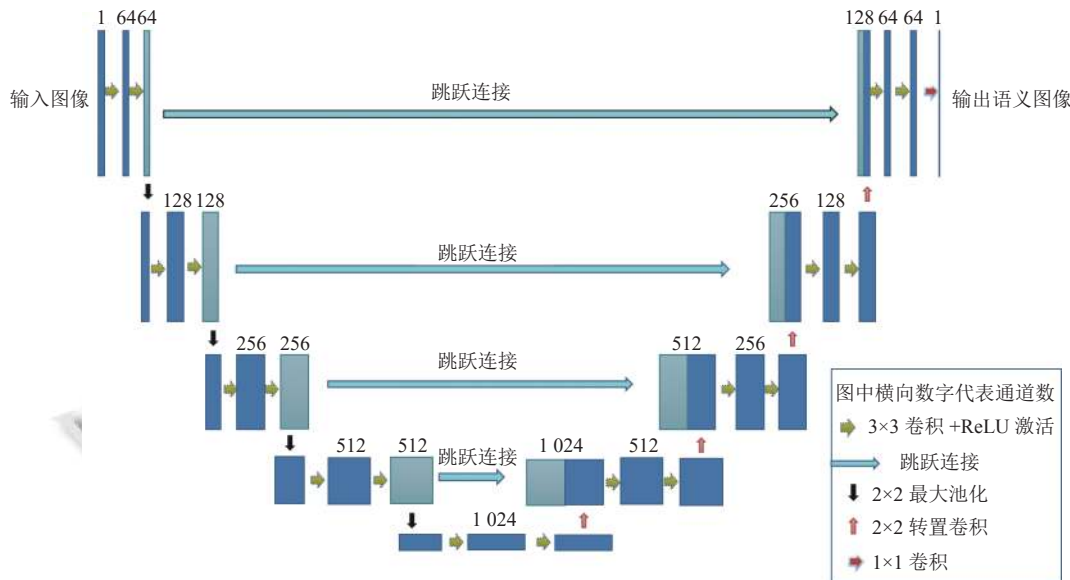


图4 UNet网络结构

2.2 残差结构

为获取更多的烧结矿气孔的信息和特征,可以通过加深网络来实现.用一种“短路”(shortcut connections)的模型来帮助多层的非线性的神经网络的拟合^[20].结合残差结构思想解决网络加深所带来网络退化的问题,使网络可以加深到更深的层次,利用到更多的信息.

$$x_{l+1} = x_l + F(x_l) \quad (3)$$

当输入为 x_l 时,学习残差是 $F(x_l)$, x_{l+1} 表示下一层输入.当残差为0时,此时堆积层仅仅做了恒等映射,此时把该处的深度忽略掉,这样保证网络性能至少不会下降.但是实际上残差不会为0,这也会使得堆积层在输入特征基础上能学习到新的特征,从而拥有更好的性能.如果残差映射 $F(x_l)$ 的结果的维度与跳跃连接 x_l 的维度不同,对 x_l 进行升维操作,两个维度相同,然后进行相加计算.如图5所示.

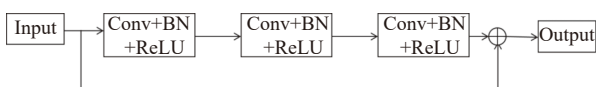


图5 残差结构

残差结构可以认为是烧结矿气孔图像的信息,且图像的信息可以直接进行传递,所以在卷积层上学到的东西就比较少,只是学习特殊的表达特征,而全部的信息是可以用残差的另一个恒等式保存下来,一定程度上解决网络退化问题,提升了梯度的反向传播,使得网络更容易训练.所以残差结构使得网络需要学习的烧结矿气孔特征变少,容易学习.

2.3 扩展模块结构

由于残差连接可以通过前后层特征相加实现特征重用,使训练过程中梯度的反向传播更容易.为使烧结矿气孔特征不容易丢失,可以采用拼接连接,当提取到新的特征的时候,拼接前面层特征.所以,为了使烧结矿气孔能更加准确地分割,本文提出了一种结合残差拼接连接思想的模块,在原始UNet网络的解码器的基础上,在进行一次卷积过后,将其结果直接与下一次卷积结果进行拼接,使气孔特征更少地丢失,提取到更多的气孔特征.加入特征映射,将提取的原始气孔特征和新的气孔特征进行多尺度融合,使用1×1的卷积进行相加操作.最后将特征融合学习得到新特征进行输出.

$$x_{l+1} = F(C) + x_l \tag{4}$$

$$C = C(X_1 + X_2) \tag{5}$$

其中, X_1 和 X_2 是通过 BN (batch normalization)、ReLU、 3×3 卷积操作得到的结果, $C(\cdot)$ 是拼接函数. 通过将 X_1 和 X_2 特征进行拼接, 能得到更多的烧结矿气孔特征, 且 X_1 直接与输出连接, 在反向传播过程中, 使这部分传递的更容易. 输入通过 1×1 卷积操作, 使与前面输出维度相同, 然后进行相加计算, 最后输出结果, 如图 6.

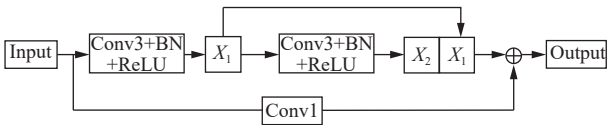


图 6 扩展连接结构

2.4 改进的 UNet 网络架构

UNet 网络能够高效地利用不同层的特征, 但是作为独立模型使用时还存在着很多不足, 每一层提取特征使用了两次 $\text{Conv}3\times 3$ 卷积, 很多气孔的细节信息丢

失, 分割效果较差, 网络层数还较浅, 提取气孔特征较少, 使得特征学习次数也有限, 表达性不强. 本文使用改进的残差拼接连接结构作为特征提取模块, 结合 UNet 网络的设计思路, 网络结构如图 7 所示. 采用改进的残差拼接结构 Block 移植到编码器部分, 进行烧结矿气孔特征的提取. 操作过程如下: 在特征提取部分, 将得到的尺寸为 512×512 烧结矿气孔图像输入到改进的 UNet 网络特征提取模块中, 采用不同 Block 模块数量, 加深网络, 形成多层特征融合, 学习到更多的新特征, 提取了更高级的烧结矿特征. 利用上采样对烧结矿气孔图像进行还原图像分辨率, 使用特征复制拼接操作将特征提取部分和上采样部分相同尺度的特征进行拼接融合. 上采样部分的终层卷积使用 Sigmoid 分类器目标分类函数, 以提高学习效率, 保证分类的准确率. 本文方法可以有效地解决 UNet 烧结矿气孔特征遗失的问题, 减少网络参数并且很大程度上防止梯度消失和爆炸.

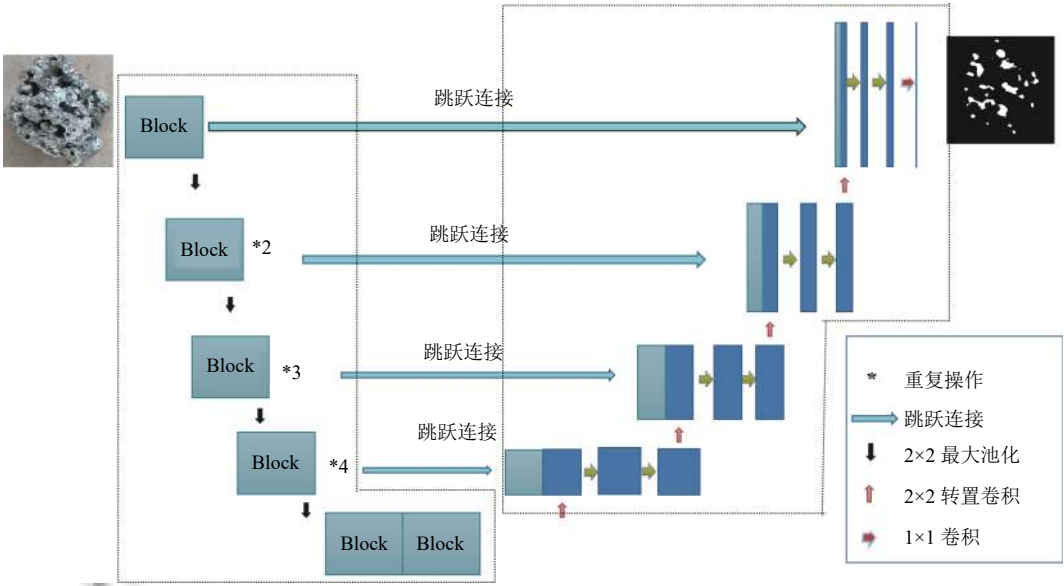


图 7 改进网络架构

3 实验结果与分析

3.1 实验设置

本文利用工业相机采集了尺寸为 $3024 \text{ 像素}\times 4032 \text{ 像素}$ 的烧结矿图像 200 张, 并挑选出 100 张作为数据样本. 为增加训练速度和效果, 将烧结矿图像裁剪成为 $512 \text{ 像素}\times 512 \text{ 像素}$ 尺寸. 将数据集 (100 张图片) 划分为训练集、验证集和测试集, 训练集用以训练模型, 验

证集用以对模型进行调优工作, 测试集用以测试模型的性能. 利用 labelme 对数据样本进行标注, 制作标签集, 样本集与对应的标签集即构成了新的数据集. 图 8(a) 是原图, 图 8(b) 是通过人工分割气孔, 进行标注的标签图像.

原始的烧结矿图像样本是进行预处理过的数据集. 为扩充样本的数量, 对训练集和验证集进行数据增强.

通过随机旋转、翻转和镜像等处理将数据集扩充到 200 张。测试集不作处理。用生成的样本数据训练网络模型,模型初始设置中网络模型权重为随机值,批量大小设置为 2,采用 Adam 作为模型的优化器,学习率设置为 0.000 01,训练批次设置为 100,损失函数使用经典的二分类交叉熵损失函数,如式 (6) 所示。 y_i 表示样本 i 的 label,正类为 1,负类为 0, p_i 表示样本 i 预测为正类的概率根据训练集和验证集的损失,来判断网络模型是否收敛。

$$L = \frac{1}{N} \sum_i -[y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (6)$$

实验环境为 Intel Xeon E5 v3 处理器, 32 GB 内存, NVIDIA GeForce RTX 2080 Ti 11 GB 显卡,采用 CUDA 10.0 架构,使用 PyTorch 深度学习框架, Python 3.6 语言编写。

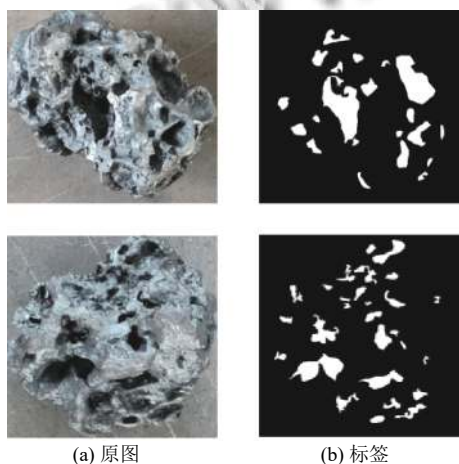


图 8 数据与标签

3.2 评价指标

本文为了对烧结矿气孔分割网络性能进行评估,采用准确率 (Accuracy) 和 Dice 系数 (Dice)、平均交并比 (MIoU) 来量化分析本文算法分割结果,计算公式如式 (7)–式 (9):

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

$$Dice = \frac{2|R_{pred} \cap R_{gt}|}{|R_{pred}| + |R_{gt}|} = \frac{2TP}{2TP + FP + FN} \quad (8)$$

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (9)$$

式 (7) 和式 (8) 中, R_{gt} 表示 ground truth 的气孔区域面积; R_{pred} 表示本文算法预测的分割结果; TP 表示算法图和标准图中各个气孔像素均为 1 的像素点数量; TN 为结果图和标准图中像素值均为 0 的像素点数量; FP 表示测试结果图中像素值为 1 而标准图像素值为 0 的像素点数量; FN 表示结果图像素值为 0 且标准图像素值为 1 的像素点数量。式 (9) 中 p_{ij} 表示真实值为 i , 被预测为 j 的数量, $k+1$ 是类别个数 (包含空类)。 p_{ii} 是真正的数量。 p_{ij} 和 p_{ji} 则分别表示假正和假负。

3.3 实验结果分析

由图 9 可知,改进的 UNet 网络 (UNetnext) 相较原始 UNet 网络,损失函数收敛的更快,并且收敛于更小的值,因此改进后的 UNet 网络具有更高的训练效率。

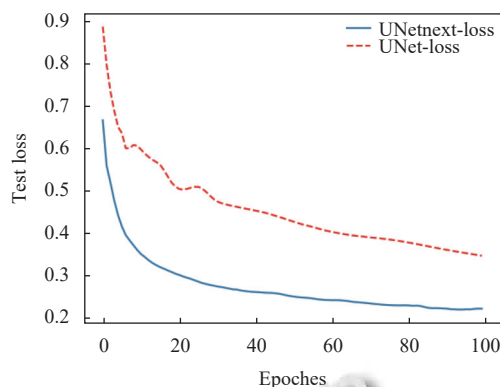


图 9 网络训练过程的损失函数

为了验证本算法的分割能力,本文选取传统的 UNet 和具有密集连接的 UNet++ 进行对比实验,其中各网络的数据集和参数设置与本文算法相同,计算得到各模型网络的精度评价指标。

从表 1 评价指标可知,UNet 网络准确率达到 95.55%, UNet++ 网络准确率达到 95.73%,改进网络准确率达到 95.56%,可以看出,各网络的准确率都很高,可能是由于前景气孔面积过小,导致分割的准确率都高,并不能说明气孔分割的准确。UNet++ 网络的 MIoU 和 Dice 系数高于原始的 UNet 网络,说明密集连接提取到了更多的烧结矿的特征。改进的网络模型的 MIoU 和 Dice 系数也优于原始的 UNet 网络,改进的残差拼接连接提取到了更多的特征。

为了测试本文改进的模型网络应用效果更好,进行可视化效果显示,结果见图 10 所示。每一行是一个烧结矿的样本,图 10(a) 为原始烧结矿图像,图 10(b) 为真

值图像,图 10(c)–图 10(e) 分别为 UNet、UNet++和本文网络的分割结果。

从图 10 中虚线方框中可以看出, UNet 网络可以大致分割出烧结矿的气孔,但是由于气孔轮廓模糊,很多气孔都没有分割出来,存在着欠分割现象, UNet++网络由于存在密集连接使模型能在一定程度上提高了对细节部分的检测能力,对轮廓模糊的气孔进行了分割,但是对气孔轮廓分割不够精细,有些气孔还存在着过分割的现象,相比,改进的网络结合了残差和拼接连接,更准确地分割烧结矿气孔。

将前面的轮廓图二值化. 利用 OpenCV 的相关算法获取各个模型算法轮廓图的参数, 利用 findContours 函数算法, 检测出每个图像中的轮廓的个数, 由于烧结矿气孔大小不一, 统计没有规律, 设置一定的周长或面积的限制, 排除特小的气孔轮廓, 再进行气孔个数统计。

表 1 网络模型的评价指标 (%)

模型	Accuracy	MIoU	Dice
UNet	95.55	72.67	66.48
UNet++	95.73	73.74	70.02
改进网络模型	95.56	73.97	68.89

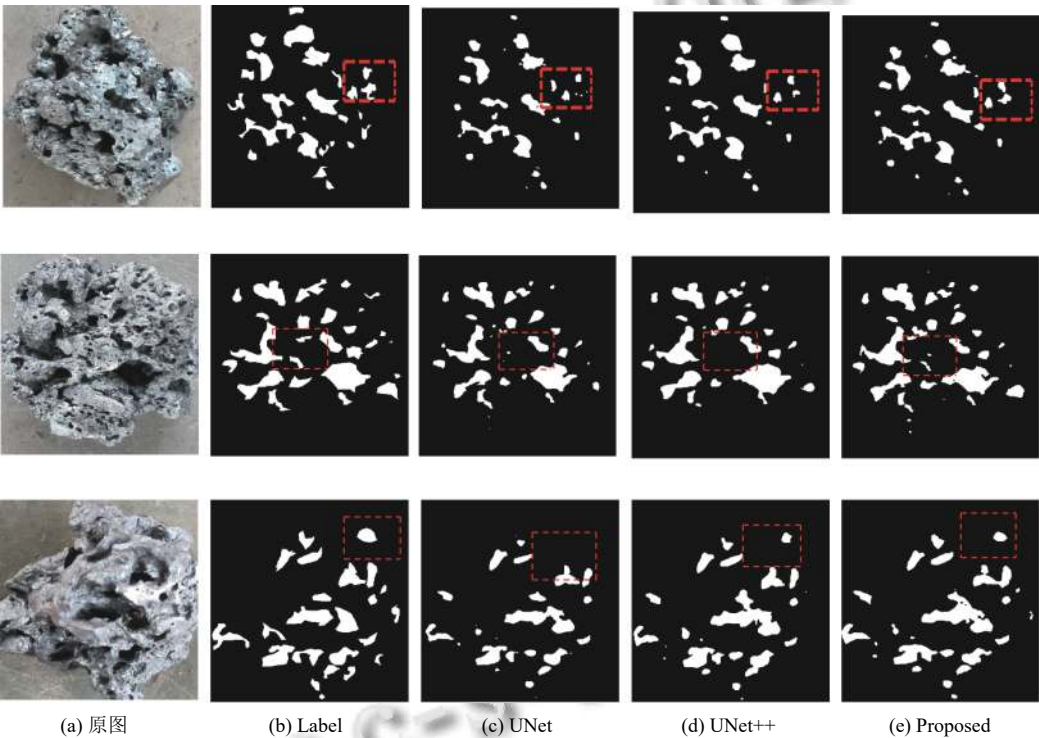


图 10 网络分割效果

如表 2 所示, 通过与原始的图像气孔数量进行比较, 改进网络分割出来的烧结矿气孔数量比 OpenCV 改进的分水岭算法检测出的气孔数更加准确, 比 UNet 和 UNet++网络分割出来的气孔数量误差率较小。

4 结论与展望

本文通过传统图像分割算法分割烧结矿气孔效果较差的不足, 应用图像语义分割技术, 提出 UNet 网络, 结合残差和拼接连接的思想, 分割烧结矿气孔图像. 利

用扩展连接模块将 UNet 的每个子模块替换, 通过不断增加提取模块, 进行多层特征融合, 使得网络在保持原本空间信息的情况下, 获取更多气孔的局部特征. 使用残差连接, 解决了网络退化的问题. 使用拼接连接, 保证了更多的特征信息. 该思想能够有效分割出气孔区域, 有利于更好判断气孔的占比和数量, 对于烧结矿的烧结质量具有重要价值. 然而, 由于烧结矿气孔图像样本是人工标注, 气孔图像不清晰等诸多因素的影响, 改进后的网络仍然存在气孔边缘轮廓不精确, 错把背景当成气孔检测出来等问题. 因此, 后续模型的改进方向

仍然是提升模型的分割精度, 并且对气孔实现更加精确的语义标注。

表2 各算法分割气孔的个数

图像编号	label	分水岭算法	UNet	UNet++	改进的模型
1	30	26	28	23	30
2	27	30	22	28	28
3	20	25	15	16	25
4	27	28	26	27	29
5	31	40	35	25	43
6	25	28	26	22	26
7	23	20	22	20	23
8	22	25	22	19	23
9	23	21	22	23	23
10	21	20	17	21	21

参考文献

- 董珂, 蒋大林. 基于改进分水岭变换的矿石图像分割算法. 计算机工程与设计, 2013, 34(3): 899–903. [doi: 10.3969/j.issn.1000-7024.2013.03.030]
- 柳小波, 张育维. 基于 U-Net 和 Res_UNet 模型的传送带矿石图像分割方法. 东北大学学报 (自然科学版), 2019, 40(11): 1623–1629. [doi: 10.12068/j.issn.1005-3026.2019.11.019]
- 师伟婕, 黄静静, 王茂发. 基于两种 U 型网络的钢铁图像缺陷检测方法对比. 北京信息科技大学学报, 2021, 36(1): 63–68.
- 卢才武, 齐凡, 阮顺颂. 基于深度图像分析的细粒度矿石分级测定方法. 应用科学学报, 2019, 37(4): 490–500. [doi: 10.3969/j.issn.0255-8297.2019.04.006]
- 罗朝阳, 张鹏超, 姚晋晋, 等. 一种基于形态学的边缘检测算法. 计算机应用与软件, 2020, 37(2): 176–181, 247.
- 蔡改贫, 刘占, 汪龙, 等. 基于形态学优化处理的标记符分水岭矿石图像分割. 科学技术与工程, 2020, 20(23): 9497–9502. [doi: 10.3969/j.issn.1671-1815.2020.23.036]
- 刘慧英, 王小波. 基于 OpenCV 的车辆轮廓检测. 科学技术与工程, 2010, 10(12): 2987–2991. [doi: 10.3969/j.issn.1671-1815.2010.12.043]
- 张文飞, 韩建海, 郭冰菁, 等. 改进的分水岭算法在粘连图像分割中的应用. 计算机应用与软件, 2021, 38(6): 243–248. [doi: 10.3969/j.issn.1000-386x.2021.06.039]
- 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展. 仪器仪表学报, 2020, 41(4): 167–182.
- 李梦怡, 朱定局. 基于全卷积网络的图像语义分割方法综述. 计算机系统应用, 2021, 30(9): 41–52. [doi: 10.15888/j.cnki.csa.008078]
- Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640–651. [doi: 10.1109/TPAMI.2016.2572683]
- Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015. 234–241.
- Zhou ZW, Siddiquee MMR, Tajbakhsh N, et al. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Transactions on Medical Imaging, 2020, 39(6): 1856–1867. [doi: 10.1109/TMI.2019.2959609]
- 全真, 徐爱俊. 基于改进 ResNet-UNet 的立木图像分割方法. 中南林业科技大学学报, 2021, 41(1): 132–139.
- Guan S, Khan AA, Sikdar S, et al. Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal. IEEE Journal of Biomedical and Health Informatics, 2020, 24(2): 568–576. [doi: 10.1109/JBHI.2019.2912935]
- 刘培军, 马明栋, 王得玉. 基于 OpenCV 图像处理系统的开发与实现. 计算机技术与发展, 2019, 29(3): 127–131. [doi: 10.3969/j.issn.1673-629X.2019.03.027]
- 宋冬梅. 数学形态学在数字图像处理中的应用研究. 计算技术与自动化, 2021, 40(2): 136–139.
- 张艳楠. 图像处理中的中值滤波及其改进. 中国新通信, 2018, 20(2): 230–231. [doi: 10.3969/j.issn.1673-4866.2018.02.190]
- 燕红文, 邓雪峰. 中值滤波在数字图像去噪中的应用. 计算机时代, 2020, (2): 47–49.
- Wu ZF, Shen CH, van den Hengel A. Wider or deeper: Revisiting the ResNet model for visual recognition. Pattern Recognition, 2019, 90: 119–133. [doi: 10.1016/j.patcog.2019.01.006]

(校对责编: 孙君艳)