

文章编号: 2095-2163(2021)11-0068-08

中图分类号: TP391.41

文献标志码: A

基于 U-net 的铸体岩石薄片图像孔隙自动提取

刘凯文, 熊淑华, 滕奇志

(四川大学 电子信息学院图像信息研究所, 成都 610065)

摘要: 使用岩石铸体薄片图像对岩石孔隙特征进行分析已经成为国内外石油地质部门常用方法之一, 自动精确地分割铸体岩石薄片中的孔隙区域是定量计算孔隙参数的前提。目前传统 RGB 阈值分割方法精度不高, 需要大量人工交互, 而一些主流图像分割的深度学习网络泛化性能差, 难以运用到实际中。针对这些问题, 本文在 U-net 网络的基础上, 提出了一种融合注意力机制和循环残差网络的模型。引入循环残差模块扩展网络深度, 又融合了注意力机制模块, 增加特征信息的学习权重。采用油田实验室常见的多种铸体薄片进行了实验, 均取得了较好的分割结果, 验证了本文方法的有效性和泛化性。

关键词: 铸体岩石薄片; 图像分割; 深度学习; 注意力机制; U-net

Automatic extraction of pores in slice image of rock castings based on U-net

LIU Kaiwen, XIONG Shuhua, TENG Qizhi

(Institute of Image Information, College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

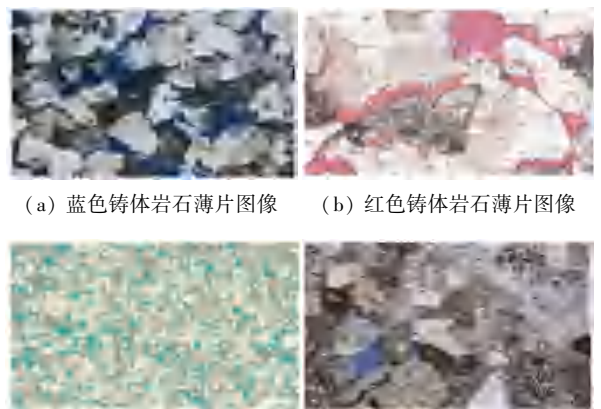
[Abstract] It has become one of the common methods for petroleum geology departments at home and abroad to analyze the characteristics of rock pores by using the image of cast rock thin section. Automatic and accurate segmentation of the pore area in cast rock thin section is the premise of quantitative calculation of pore parameters. At present, the traditional RGB threshold segmentation method is not accurate and needs a lot of human interaction. However, the generalization performance of some mainstream image segmentation deep learning networks is poor, which is difficult to apply to practice. To solve these problems, this paper proposes a model based on u-net, which combines attention mechanism and cyclic residual network. The cyclic residual module is introduced to expand the network depth, and the attention mechanism module is integrated to increase the learning weight of feature information. Experiments are carried out on a variety of common cast thin sections in the oilfield laboratory, and good segmentation results are obtained, which verify the effectiveness and generalization of the proposed method.

[Key words] cast rock slice; image segmentation; deep learning; attention mechanism; U-net

0 引言

在岩石薄片分析中, 孔隙的特征参数可以表征岩层微观结构, 如何精准地分割出铸体薄片中的孔隙区域是后续分析工作的关键前提。对于岩石铸体薄片而言, 铸体的颜色通常分为蓝色和红色两种, 如图 1(a)、(b) 所示。一方面由于各地使用的颜料和用量以及加工手法差异, 不同地区的铸体岩石薄片会有颜色深浅的区别, 导致不同铸体薄片的孔隙区域色差较大, 如图 1(a)、(c) 所示; 另一方面由于不同区域地质结构的多样性, 部分薄片的孔隙分布不均匀, 某些视域范围被整个颗粒面积所占据, 导致有些铸体薄片面孔率较低, 如图 1(d) 所示。这些因素都对孔隙提取工作造成了不利影响。因此, 本文提出一种高效率、高精度的自动分割算法, 有助于研

究人员更有效率地完成分析工作。



(a) 蓝色铸体岩石薄片图像 (b) 红色铸体岩石薄片图像

(c) 浅色铸体岩石薄片图像 (d) 低孔隙度铸体岩石薄片图像

图 1 铸体岩石薄片图像

Fig. 1 Rock cast thin sheet image

之前用于岩石铸体图像的分割方法以孔隙区域

基金项目: 国家自然科学基金(62071315)。

作者简介: 刘凯文(1995-), 男, 硕士研究生, 主研究方向: 数字图像处理; 熊淑华(1969-), 女, 博士, 副教授, 硕士生导师, 主要研究方向: 数字图像处理、图像通信; 滕奇志(1961-), 女, 博士, 教授, 博士生导师, 主要研究方向: 数字图像处理、模式识别。

通讯作者: 熊淑华 xiongsh@scu.edu.cn

收稿日期: 2021-08-23

的彩色空间特征分割的传统方法为主^[1]。张婷等采用形态学分水岭算法响应孔隙边缘, 获得连续封闭孔隙边缘信息, 该方法缺点是太易受噪声和灰度值变化的影响^[2]; 利用区域种子点生长的方法, 对岩石薄片图像中的颗粒与颗粒以及孔隙与颗粒进行了划分^[3]; B. Obara 根据不同岩石图像的特点, 在不同彩色空间中对图像进行分割^[4]; Gorsevski 对岩石图像中的边界像素点进行标记, 再以边界像素点为界分割岩石图像, 但是会出现伪边缘, 造成分割不准确的情况^[5]。

近些年随着机器视觉领域的迅猛发展, 也出现了基于神经网络的深度学习分割方法。虽然学习方法在一定程度上优于传统方法, 但是也存在一些问题, 如: 李周等通过深度学习的图像分割网络 FCN 对铸体图像进行孔隙提取, 但是由于 FCN 网络在池化阶段的信息损失, 导致提取结果不佳, 且模型受客观因素影响, 实用性不强^[6]; 蔡宇恒等提出以 U-net 网络为骨架, 在编码器部分使用残差块与空洞卷积的结合, 提升了网络深度与广度, 但是在解码部分只是简单的使用短链接进行特征融合, 对于上文中提到的某些低孔隙度的铸体薄片, 以及某些较低分辨率的铸体薄片的分割效果差, 网络泛化性能有待提高^[7]。为了解决上述问题, 本文在 U-net 的基础上, 进行了优化改进, 在输入端编码阶段, 使用循环残差结构, 进一步提升网络深度, 提取出更深层次的语义; 在解码部分使用了注意力机制, 抑制不相关区域的学习权重, 对卷积后的图像进行校准, 从而取得理想的分割效果。实验证明, 对低孔隙度和颜色较浅的铸体薄片也有较好的分割效果和更好的泛化性能。

1 相关技术

1.1 U-net

U-net 是 Ronnberger 等人 2015 年提出的一种主要由编码器与解码器构成的 U 型对称网络结构即 (Encoder-Decoder) 结构, 如图 2 所示^[8]。编码器主要负责图像的下采样, 通过两个 3×3 的卷积层、一个 2×2 的最大池化层和 $ReLU$ 函数进行 4 次下采样, 对每层的输入进行 2 次 3×3 卷积操作之后, 进行一次 2×2 的最大池化, 使特征图尺寸下降, 卷积核数目翻倍, 从而可以更好地捕捉和传播语义信息, 完成对特征像素点的提取。解码器则主要负责特征图的上采样, 通过 2×2 反卷积层和编解码器之间的跳跃连接进行 4 次上采样, 并且每次上采样过程都结合下采样各层信息和上采样输入信息扩大特征图

尺寸, 同时卷积核减半, 从而逐步恢复特征图像的像素信息 & 图像精度。

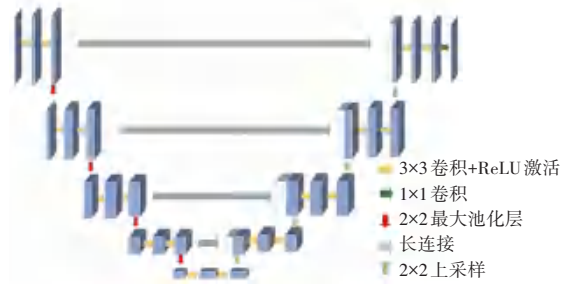
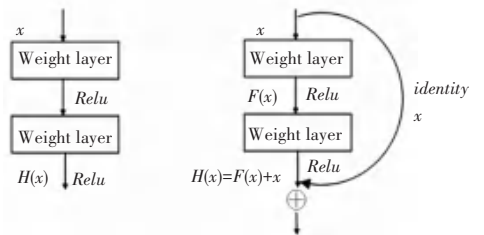


图 2 U-net 网络结构

Fig. 2 U-net network structure

1.2 残差网络 ResNet

残差网络 (ResNet) 是 He 等人 2016 年提出的一种引入跳跃连接的神经网络模型^[9]。残差网络由上层输出 x 和求和前的网络映射 $F(x)$ 组成, $H(x)$ 是输入到求和后的网络映射, 与正常深度神经网络比较, 不难发现正常深度神经网络的做法是寻找 x 到 $H(x)$ 的恒等映射函数, 如图 3(a) 所示; 而引入残差后是寻找 $H(x) = x + F(x)$ 的映射, 即转换为一个残差函数 $F(x) = H(x) - x$, 只要 $F(x) = 0$, 就构成了一个恒等映射 $H(x) = x$, 如图 3(b) 所示。明显比直接寻找输入 x 到输出 $H(x)$ 的映射要容易很多, 即在一定程度上减少了模块中需要计算的参数, 并且后者明显对输出的变化更敏感, 去掉相同的主体部分 x , 从而更能突出微小的变化。



(a) 正常神经网络模块 (b) 残差模块

图 3 网络模块

Fig. 3 Network module

1.3 注意力机制

近年来, 随着深度学习在机器视觉领域的飞速发展, 注意力机制在图像分类、语义分割和自然语言处理等研究中得到广泛应用。卷积运算作为卷积神经网络使用最为频繁的运算模块, 其原理是从空间维度以及通道维度将特征图像中局部感受野上的特征进行无差别融合, 从而获取多维度多尺度的空间信息。但对于不同的图像分类任务, 不同通道所包含信息的重要程度也是不同的, 如果仅仅是将各个通道信息进行无差别融合, 在一定程度上会丢失许多重要信息。

因此, Hu 等人 2018 年提出了 SE (squeeze and excitation) 模块, 该模块建立了通道之间的关系, 引入了注意力机制, 提高重要信息的学习权重, 抑制对网络性能影响不大的无用信息, 从而达到改善卷积神经网络通道特征的效果^[10], 如图 4 所示。

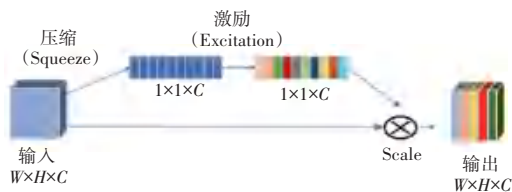


图 4 SE 模块结构

Fig. 4 SE module structure

SE 模块主要由压缩 (squeeze) 和激励 (excitation) 两种结构组成, 如图 5 所示。压缩结构采用平均池化来聚合全局信息, 即输入特征图为 X , 其维度为 $R^{H \times W \times C}$, H 和 W 分别表示特征图空间信息的高和宽两个维度, C 表示特征图的通道数。沿着空间维度 $R^{H \times W}$ 将通道中的每一个二维特征图转换成一个实数, 通道数 C 保持不变, 得到一个 $R^{1 \times 1 \times C}$ 维度的通道向量, 该向量在一定程度上可以表示特征图的全局特征; 而激励结构使用一个全连接神经网络, 对压缩之后的结果做一个非线性变换。该结构由全连接层—*ReLU*—全连接层—*Sigmoid* 组成, 来捕获各个通道之间的关系, 从而确定每个通道重要程度。在第一个全连接层将通道数 C 降维为 c/r , 其中 r 为缩放系数, 经过 *ReLU* 函数后再经过一个全连接层, 通道数由 c/r 扩大为 C , 紧接着是一个 *Sigmoid* 函数, 至此生成一个 $R^{1 \times 1 \times C}$ 维度的通道向量。该向量为各个通道的学习权重, 最后将激励后得到的权重向量与输入图像 X 逐通道相乘, 依据权重将重要特征信息放大, 同时抑制无关信息的影响。

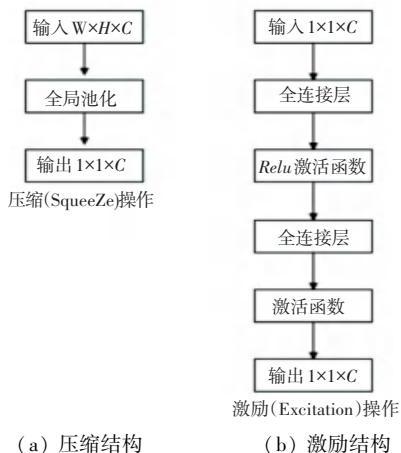


图 5 压缩结构与激励结构

Fig. 5 Compression structure and excitation structure

2 改进后的网络模型

借鉴 U-net 的设计思想, 本文设计的网络结构仍然采用 (Encoder-Decoder) 编码器-解码器结构, 实现端到端的像素分类。但是传统 U-net 的编码器部分只是卷积层以及 *ReLU* 函数的简单级联进行 4 次下采样, 这样的降采样针对铸体岩石薄片这类具有复杂细微的边界特征的图像, 提取能力有限, 对最终分割性能贡献不大。

为了在分割铸体岩石图像时获取更为细微的特征像素, 本文在参考了 Md Zahangir Alom 等人 2018 年提出的一种基于残差块结构的递归残差卷积神经网络的思想^[11], 首先是把残差网络 (ResNet) 和循环卷积网络 (RCNN) 的结构整合到了 Encoder-Decoder 结构里, 代替 U-net 原有的 2 个 3×3 的卷积层, 同时本文在整合过程中对残差模块进行了优化。相较于常规两层残差学习单元, 本文采用三层的残差学习单元。常规的两层残差的结构中, 由于残差等于目标输出减去输入, 即输入、输出维度需保持一致, 因此常规的二层残差结构包含两个相同的输入、输出通道数的 3×3 卷积, 而本文采用的三层残差模块, 如图 6 所示。在中间的 3×3 的卷积前后分别使用了 1×1 卷积, 1×1 卷积将输入通道数进行了先降维再升维的操作, 降低了参数计算的复杂度。例如: 输入和输出都是 256 维的参数量, 在常规残差模块中就是 2 个 3×3 的卷积, 参数量为 $3 \times 3 \times 256 \times 256 \times 2$ 共计 1 179 648。而本文使用的三层残差模块的参数量 $1 \times 1 \times 256 \times 64 + 3 \times 3 \times 64 \times 64 + 1 \times 1 \times 64 \times 256$ 共计 69 632, 减少到了 1/16, 从而避免了在深层网络中进行过多的参数计算, 减轻了网络负担。另外如果有输入、输出维度不同的情况, 该三层残差模型可以对输入 x 做一个线性映射变换, 再连接到后面的层即可。

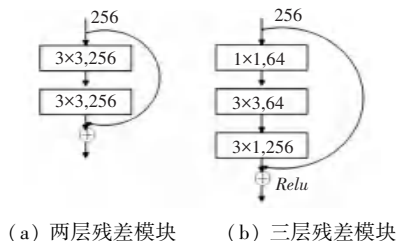


图 6 残差模块

Fig. 6 Residual module

本文提出的融合注意力机制和循环残差的网络模型如图 7 所示, 其中循环残差模块 (recurrent residual block), 不是简单的 conv+Relu 层, 而是在对

残差采取两个时间步长 ($timestep = 0 \sim 2$) 循环两次, 做两个循环卷积操作, 即当前的输出被添加到前一段时间步长的输出, 当前时间步长的输出作用于下一步的输入, 如图 8 所示, 这样可以有效地增加网络深度。



图 7 融合注意力机制和循环残差的网络模型

Fig. 7 A network model integrating attention mechanism and cyclic residual

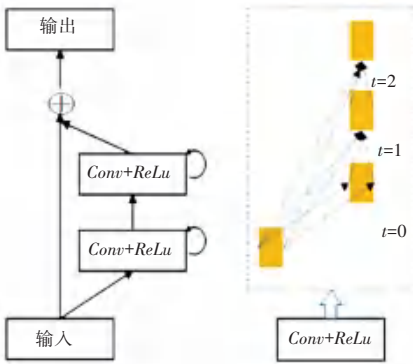


图 8 循环卷积操作

Fig. 8 Cyclic convolution structure

此处循环卷积的操作是根据离散时间步长执行的。例如, 考虑循环卷积模块中的第 l 层中的 x_l 输入样本, 以及来自循环卷积层 (RCL) 中第 k 个特征映射中的输入样本的 (i, j) 处的像素, 另外假设网络 O_{ijk}^l 是在时间步长 t 时的输出, 可以表示为式 (1):

$$O_{ijk}^l(t) = (w_k^f) * x_l^{f(i,j)}(t) (w_k^r)^T * x_l^{r(i,j)}(t-1) + b_k \quad (1)$$

这里, $x_l^{f(i,j)}(t)$ 和 $x_l^{r(i,j)}(t-1)$ 分别是标准卷积层和第 l 层的 RCL 的输入; w_k^f 和 w_k^r 的值分别是标准卷积层和第 k 个特征映射的 RLC 的权重; b_k 是偏差。经过标准线性整体单元 (ReLU) 激活函数, 可得 y , 式 (2):

$$y = f(O_{ijk}^l(t)) = \max(0, O_{ijk}^l(t)) \quad (2)$$

将循环卷积神经网络单元的最终输出定义为式 (3):

$$F(x_l, w_l) = y_{1 \times 1}(x) + y_{3 \times 3}(x) + y_{1 \times 1}(x) \quad (3)$$

循环残差的残差操作可以用等式 (4) 表示。

$$x_{l+1} = x_l + F(x_l, w_l) \quad (4)$$

采用不同时间步长的特征融合的方式能够获取更具有表达性的特征, 也有助于提取较低层次的特征。

最后, 考虑到某些低孔隙度的铸体岩石薄片的孔隙分布不均匀, 某些视域范围被整个颗粒面积所占据, 即整个特征图中主要以无关信息为主, 重要的特征信息占比较小, 因此本文在对编码后的每个像素特征与解码中对应特征进行拼接之前, 使用一个注意力门, 重新调整编码的输出特征, 从而达到抑制特征图中的无关信息。

注意力门结构如图 9 所示。下采样层同层的特征图 $g = g_i(F_g \times H_g \times W_g \times D_g)$, 进行 $1 \times 1 \times 1$ 卷积运算得到 $W_g^T g_i$, 上采样层的上一层的特征图 $x^l = x_i^l(F_l \times H_x \times W_x \times D_x)$, 进行 $1 \times 1 \times 1$ 卷积运算得到 $W_x^T x_i^l$, 将上两步得到的特征图 $W_g^T g_i$ 和 $W_x^T x_i^l$ 相加后再进行 ReLU, 得到 $\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)$, 即调整了 g 和 x_i^l 的权重, σ_1 为 ReLU 激活函数, 然后再使用 $1 \times 1 \times 1$ 卷积运算, 得到 $q_{att}^l = \psi^T(\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)) + b_\psi$, 接着对 q_{att}^l 进行 sigmoid 激活函数, 得到最终的注意力系数 (α_i^l), 最后, 让 x^l 与注意力系数相乘, 就能把注意力放到目标区域上。

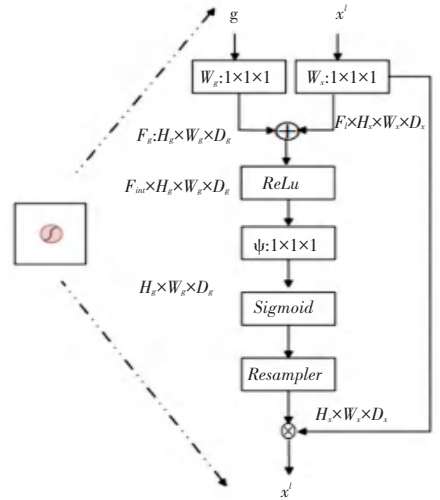


图 9 注意力门结构

Fig. 9 Attention gate structure

在网络结构中引入注意力门, 将跳跃连接中不相关信息以及噪声的激活响应抑制, 只凸显相关信息的激活值。在中间层做跳跃连接时使用注意力门, 这样确保不同尺度的注意力单元能够获取更大范围内的相关性信息, 而不仅仅依赖于局部信息进行预测。

3 结果分析

3.1 数据集

本文所用数据集由铸体图像以及提取孔隙区域后的二值图像标签构成, 数据集里的铸体图像均是

由高清相机在显微镜下拍摄采集,接着将采集到的图像裁剪成 300×300 分辨率的图像用作数据集中的原始图像,而数据集中的标签图则是将裁剪后的标签图采用人工手动标记的方式提取该图中孔隙区域,本文所用数据集中原图和标签图总计各 2 500 张,如图 10 所示。

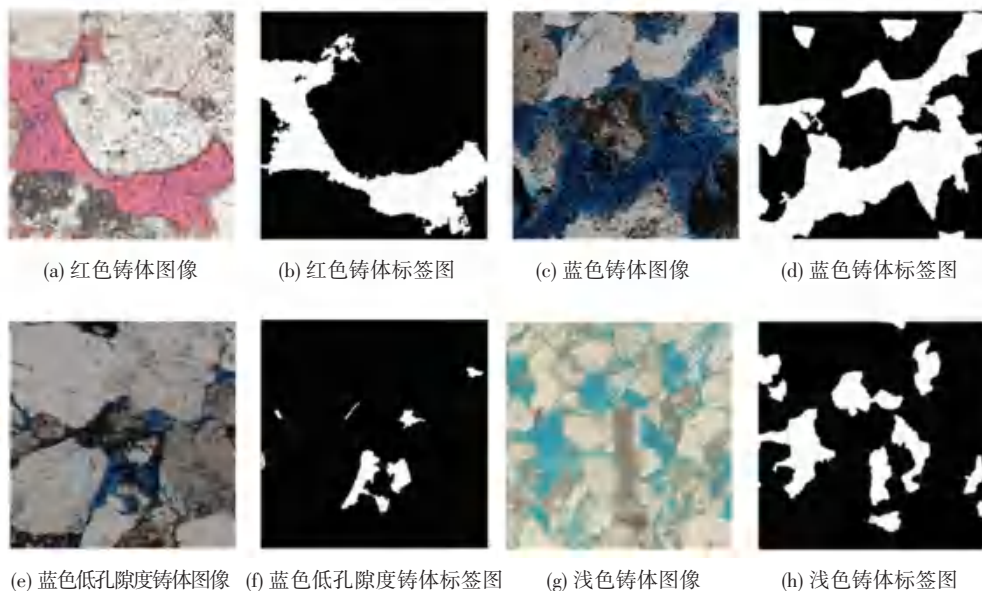


图 10 数据集图示

Fig. 10 Dataset images

3.2 实验环境

本次实验的软件及硬件环境操作系统: windows 10,以 tensorflow-gpu-1.4.0 为深度学习框架,使用 keras-1.0.8 神经网络库作为程序接口进行深度学习模型的设计、调试、与评估,使用 CUDA-v8.0 作为 GPU 加速运算平台,显卡型号为 NVIDIA GeForce GTX 1060 6 GB。

3.3 实验参数设置

由于 Adam 算法在训练时自适应的调节学习率,且收敛速度更快,故选择 Adam 作为优化器,初始学习率设置为 1×10^{-4} , BatchSize 设置为 3,共训练 50 轮,每轮训练为 2 500 张训练图,见表 1。

表 1 具体实验参数

Tab. 1 Experimental parameters

参数名称	参数取值
优化器	Adam
BatchSize	3
学习率	1×10^{-4}
迭代次数	50
每轮训练集数据量	2 500

卷积层的激活函数使用的是修正线性单元

在实际应用中,一张铸体岩石薄片图像大小普遍都远大于 300×300 ,因此在预测时会先将原图裁剪为 512×512 的图像,再送入训练好的模型中分割提取,最后将分割好的结果进行拼接操作,从而得到最终完整的孔隙分割提取结果图。

(ReLU),该函数是从底部开始修正的一种函数,数学公式为(5):

$$f(x) = \max(0, x) \quad (5)$$

其中, x 表示的是激活层的输入,该激活函数相较于逻辑回归函数(sigmoid),能使网络更快速地收敛。

对铸体岩石薄片仅是针对其孔隙区域进行提取,其余区域当作背景处理,这就是典型的二分类问题,所以本文使用二进制交叉熵作为损失函数,式(6):

$$C = -\frac{1}{N} \sum_{i=1}^N y_i \ln a_i + (1 - y_i) \ln (1 - a_i) \quad (6)$$

其中, y_i 为期望的输出, a_i 为神经元实际输出。该损失函数与传统的方差代价函数相比,克服了权重更新过慢的缺陷。

使用 sigmoid 激活注意力门 AG 中的最后输出,起到特征组合的作用,式(7):

$$S(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

3.4 评价指标

本实验采用的评价指标为准确率(accuracy)和平均交并比(mean intersection over union, mIOU),上

述指标为图像分割领域常用的评价分割指标,在一定程度上能够评价出训练模型的综合性能。 $Accuracy$ 和 $mIOU$ 的定义如式(8)和式(9):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$mIOU = \frac{1}{k + 1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (9)$$

其中, TP 为真正值,即预测值为 1,真实值为 1; TN 为真负值,即预测值为 0,真实值为 0; FP 为假正值,即预测值为 1,真实值为 0; FN 为假负值,即预测值为 0,真实值为 1; k 表示 $k + 1$ 类(其中包含一个空类或背景),其关系如图 11 所示。本实验针对岩石铸体薄片的孔隙提取,属于二分类问题,因此共分为两类,一类为孔隙区域,一类为背景区域。

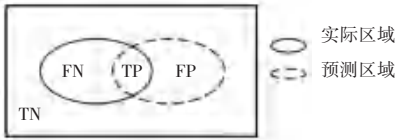


图 11 图像分割性能指标评直观示意图

Fig. 11 Intuitive diagram of image segmentation performance indicators

准确率 (Accuracy) 即标记正确的像素 (真正值与真负值) 占总像素的百分比; 均交并比 (mIOU) 为语义分割的标准度量, 计算两个集合的交集和并集之比, 语义分割的问题中, 这两个集合为真实值 (ground truth) 和预测值 (predicted segmentation)。这个比例可以变形为正真数 (intersection) 比真正、假负、假正 (并集) 之和。在每个类上计算交并比 IOU, 之后求和平均。

3.5 结果分析

通过统计每轮训练的准确率以及训练损失可以发现, 损失函数随着训练迭代次数的增加而逐渐收敛到一个极小值, 准确率随着迭代次数的增加而逐渐增高, 如图 12 和图 13 所示。

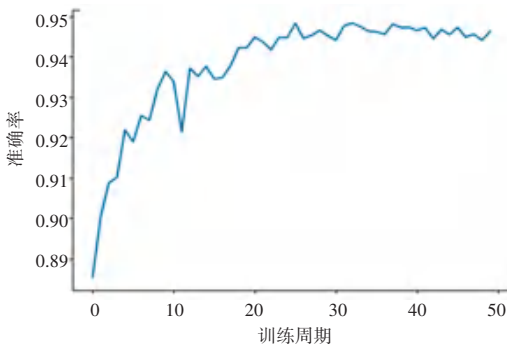


图 12 训练准确率曲线

Fig. 12 Training accuracy curve

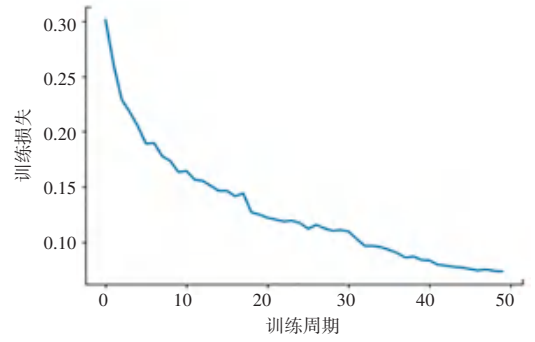


图 13 训练损失函数曲线

Fig. 13 Training loss function curve

为了验证本文所提出的融合注意力机制的循环残差网络分割方法的准确性和泛化性, 使用相同的数据集分别对 U-net, U-net++, R2U-Net 等主流神经网络训练后的模型与本文所改进的新模型进行多组对比实验, 并且为了证明用 U-net 作为基础网络对铸体岩石薄片这类边缘情况复杂的图像具有较好的分割效果, 本文还使用了 Segnet, Fcn 主流图像分割网络进行对比实验。测试数据集包含红、蓝铸体岩石薄片, 共计 300 张图像, 将这些图像统称为铸体效果明显的图像, 评价结果见表 2。

表 2 铸体效果明显的图像不同网络下的测试结果

Tab. 2 Test result of different networks on image with good effect

网络类型	像素精度 PA/%	平均交并比/%
U-net	92.87	82.10
U-net++	91.23	85.36
R2U-net	94.14	83.14
U-net+Attention	93.29	81.66
Segnet	92.11	76.23
Fcn8s	91.89	78.56
本文方法	94.83	88.04

为了说明该模型具有更好的泛化性能, 本文又将部分低孔隙度铸体岩石薄片以及部分颜色过深或过浅的铸体图像, 组成一组测试集, 用来测试在不同方法的泛化性能, 将这些图像统称为铸体效果不明显的图像, 评价结果见表 3。

表 3 铸体效果不明显的图像不同网络下的测试结果

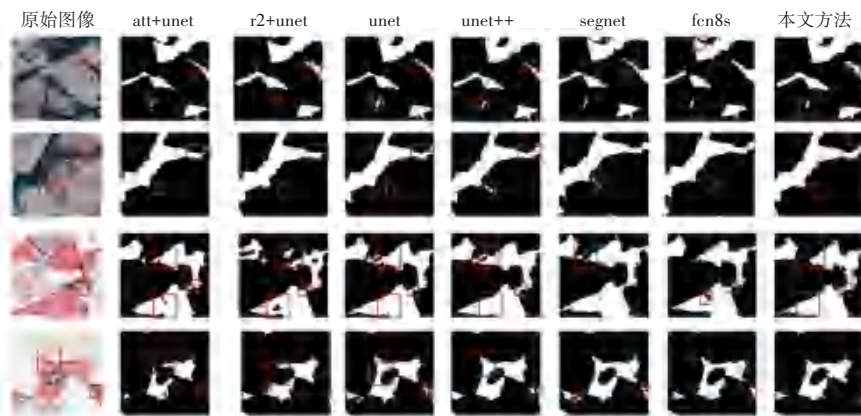
Tab. 3 Test result of different networks on image with bad effect

网络类型	像素精度 PA/%	平均交并比/%
U-net	86.71	75.23
U-net++	87.23	74.12
R2U-net	88.45	76.21
U-net+Attention	88.96	77.76
Segnet	81.41	70.56
Fcn8s	87.36	76.45
本文方法	90.12	71.66

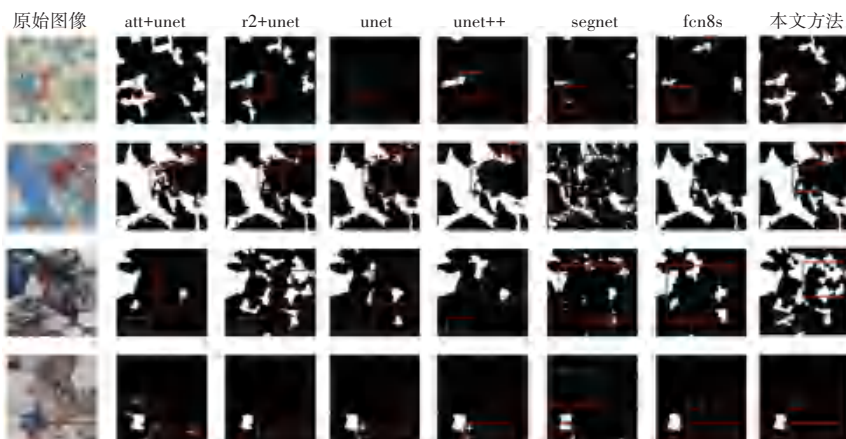
铸体效果明显的图像在不同网络下的分割结果, 如图 14(a) 所示; 铸体效果不明显的图像在不同

网络下的分割结果,如图 14(b)所示。通过分析图 14(a)组的结果对比图,可以看出与 Segnet 和 Fcn 这些经典图像分割网络相比以 U-net 为基础网络具有更好的分割效果,边缘更加平滑;其次,通过 att-unet 网络提取的孔隙结构与 U-net, Unet++相比,常出现一些欠分割的现象,其主要原因在于网络深度较浅,缺少某些细节信息,在计算注意力权重时将某些与背景相似的像素点误当作背景处理,即训练时该部分学习权重较小,导致了分割结果出现欠分割的现象;而 U-net 网络缺少一些细节处的提取,存在一定程度的过分割,且分割出的结果图的边缘区域存在明显毛刺突起的现象,不够平滑;R2U-Net 也出现了欠分割现象可能是因为网络深度加深之后,梯度不稳定导致某些浅层的学习能力下降。本文网络为了解决随着模型深度加深导致浅层学习能力的下降的问题,将 R2U-Net 与注意力机制融合,既能保证网络具有足够的深度来提取细节信息,又引用注意力机制,加强浅层特征信息的学习权重,避免随着深度加深导致浅层特征学习能力下降的问题。

为了进一步验证加入的注意力机制模块能够有效地加强重要信息的学习权重,抑制对网络性能影响不大的无用信息。图 14(b)组选取了铸体效果不明显的岩石薄片,比如颜色过浅,和低孔隙度的铸体岩石薄片进行对比实验,可以看出 Fcn 网络由于采用了反卷积进行上采样,其边缘存在很明显的锯齿效应,Segnet 网络由于提取特征方式过于单一,所以对于这些特殊铸体薄片分割出现明显的噪点;其次可以发现传统 U-net 网络和 Unet++网络在那些铸体颜色过浅使得目标与背景颜色较为相似的薄片图像中出现严重欠分割现象,甚至出现完全没有分割出目标的情况,如图 14(b)第一组所示。而 R2U-Net 因为网络具有足够深度,可以提取到某些细节特征,其分割结果优于传统 U-net 网络和 Unet++网络,但仍存在欠分割的情况;融合注意力机制后,通过加强特征信息学习权重,抑制无关信息的影响,能够使网络在无关信息居多的低孔隙度铸体薄片,和铸体颜色过浅,目标特征不明显的浅色铸体薄中有较好的分割效果,加强了网络的泛化性能。



(a) 铸体效果明显的岩石图像其不同网络下的分割结果



(b) 铸体效果不明显的图像其不同网络下的分割结果

图 14 结果图分析

Fig. 14 The results of segmentation

4 结束语

本文针对传统的岩石铸体薄片图像孔隙提取方法存在的分割精度不高,实用性不强以及部分应用于岩石铸体图像孔隙提取的卷积神经网络泛化性差等问题,提出了一种新的基于深度学习的岩石铸体图像孔隙提取方法。该方法在 U-net 网络为骨架的基础上融合注意力机制的循环残差网络,U-net 为骨架保证了端到端的自动分割;循环残差网络保证具有足够的网络深度,从而获取更多的深层细节特征信息;融合注意力机制,加强特征信息的学习权重,抑制无关背景信息影响,解决了低孔隙度的铸体薄片和颜色过浅的铸体薄片对孔隙分割的影响,使得本网络具有更好的泛化性能。采用油田实验室常见的多种铸体薄片进行了实验,均取得了较好的分割结果,验证了本文方法的有效性和泛化性。由于模型深度的加深,不可避免地增加了整个网络的复杂度和参数量,在一定程度上存在训练时间长,梯度不稳定等问题。在接下来的研究中,将对该模型进行进一步的改良,在保证具有泛化性和有效性的前提下,优化网络结构,减少计算参数,使该模型更适合实际工程的应用。

参考文献

[1] ALIGHOLI S, LASHKARIPOUR G R, KHAJAVI R, et al. Automatic mineral identification using color tracking[J]. Pattern Recognition,

2017, 65:164-174.

- [2] 张婷,徐守余,王子敏. 储层微观孔喉网络图形识别方法[J]. 吉林大学学报(地球科学版), 2011, 41(5):1646-1650.
- [3] A P A, B O C, ANDREAS GÜNTHER C, et al. Semi-automatic segmentation of petrographic thin section images using a "seeded-region growing algorithm" with an application to characterize weathered subarkose sandstone[J]. Computers & Geosciences, 2015, 83:89-99.
- [4] OBARA B. A new algorithm using image colour system transformation for rock grain segmentation[J]. Mineralogy & Petrology, 2007, 91(3-4):271-285.
- [5] GORSEVSKI P V, et al. Detecting grain boundaries in deformed rocks using a cellular automata approach[J]. Computers & Geosciences, 2012,42:136-142.
- [6] 李周. 岩石铸体薄片图像的孔隙提取及颗粒分割研究[D]. 成都: 四川大学, 2018.
- [7] 蔡宇恒,滕奇志,涂秉宇. 基于深度学习的岩石铸体薄片图像孔隙自动提取[J]. 科学技术与工程, 2020, 533(28):300-307.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
- [9] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016:770-778.
- [10] JIE H, LI S, GANG S, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [11] ALOM M Z, HASAN M, YAKOPCIC C, et al. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation[J]. 2018, 5:716-723.

(上接第 67 页)

参考文献

[1] 陈华伟,吴禄慎,袁小翠. 基于链码改进算法的钢轨表面缺陷识别[J]. 计算机工程与设计, 2015, 36(11):3097-3101.

[2] 陈华伟,吴禄慎,袁小翠. 2DPCA 算法设计及其在扣件识别中的应用[J]. 计算机工程与设计, 2016, 37(2):424-428.

[3] 欧阳,罗建桥,李柏林,等. 嵌入标记信息的铁路扣件状态检测主题模型[J]. 计算机工程, 2018, 44(6):200-206.

[4] 张雨,王志晖,李柏林. 基于截线法的铁路扣件骨架提取算法[J]. 铁道标准设计, 2019, 63(4):73-78.

[5] 张辉,李平, Q.M.Jonathan,等. 基于改进马尔可夫随机场的钢轨缺陷分割[J]. 计算机工程与设计, 2020, 41(4):160-169.

[6] 林菲,杨子明,李永光,等. 钢轨扣件检测技术研究[J]. 中国铁路, 2019(6).

[7] 张强,李朝奎,李俊晓,等. 一种改进的基于法矢方向调整的平面点云分割方法[J]. 地理与地理信息科学, 2015, 31(1):45-48.

[8] 王雅男,王挺峰,田玉珍,等. 基于改进的局部表面凸性算法三维点云分割[J]. 中国光学, 2017, 10(3):348-354.

[9] 李仁忠,刘阳阳,杨曼,等. 基于改进的区域生长三维点云分割[J]. 激光与光电子学进展, 2018, 55(5):051502.

[10] 高天一,韩慧妍,韩燮. 基于骨架的三维点云模型分割[J]. 计算机工程与设计, 2019, 40(5):1418-1423.

[11] 王希,陈晓波,刁俊通. 人体散乱点云数据的区域分割算法[J].

机械设计与研究,2020,36(1):1-6.

- [12] 王冲,李殿能,邓君裕,等. 改进法线方向的点云实时分割提取平面方法研究[J]. 计算机测量与控制, 2018, 26(5):210-213.
- [13] Mohammed Merzougui, Ahmad El Allaoui. Region growing segmentation optimized by evolutionary approach and Maximum Entropy[J]. Procedia Computer Science, 2019, 151:1046-1051.
- [14] Anh-Vu Vo, Linh Truong-Hong. Octree-based region growing for point cloud segmentation[J]. Isprs Journal of Photogrammetry & Remote Sensing, 2015, 104:88-100.
- [15] NURUNNABI A, BELTON D, WEST G. Robust Segmentation for Large Volumes of Laser Scanning Three-Dimensional Point Cloud Data[J]. IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(8):4790-4805.
- [16] HONG Z. Lung Lobe Segmentation Based on Lung Fissure Surface Classification Using a Point Cloud Region Growing Approach[J]. Algorithms, 2020, 13(10).
- [17] DONG Y, YANG W, WANG J, et al. An improved supervoxel 3D region growing method based on PET/CT multimodal data for segmentation and reconstruction of GGNs[J]. Multimedia Tools and Applications, 2020, 79(11).
- [18] HAN X, WANG X, LENG Y, et al. A Plane Extraction Approach in Inverse Depth Images Based on Region-Growing[J]. Sensors, 2021, 21(4):1141.