

文章编号: 2095-2163(2024)02-0156-06

中图分类号: TP391.41

文献标志码: A

基于双解码器网络的岩心 CT 图像分割

陈忠照¹, 滕奇志¹, 吴晓红¹, 何海波²

(1 四川大学 电子信息学院, 成都 610065; 2 成都西图科技有限公司, 成都 610065)

摘要: 在石油地质领域, 分析岩心内部微观孔隙结构的形态分布、特征参数, 对研究油气资源的渗流特性和储集性能具有重要意义。岩心 CT 图像具有噪点多、比度低、亮度不均匀的特点, 目前实际工程运用的孔隙提取方法, 仍然存在着需要大量人工交互且分割精度较低的问题。针对这些问题, 本文提出了基于双解码器网络的分割方法, 构建了岩心 CT 图像分割数据集, 用图像预处理网络分支, 辅助训练图像分割网络。实验结果表明, 本方法的模型参数量仅有 33.3 MB, 像素精度 PA 能达到 91.41%, 平均检测交并比 $MIoU$ 能达到 85.32%, 具有模型小、推理速度快、分割精度高的优点。

关键词: 岩心 CT 图像; 图像分割; 深度学习

Core CT image segmentation based on dual decoder network

CHEN Zhongzhao¹, TENG Qizhi¹, WU Xiaohong¹, HE Haibo²

(1 College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2 Chengdu Xitu Technology Co. Ltd, Chengdu 610065, China)

Abstract: In the field of petroleum geology, analyzing the morphological distribution and characteristic parameters of the microscopic pore structure inside the core is of great significance to study the seepage characteristics and reservoir performance of oil and gas resources. Core CT images have the characteristics of high noise, low ratio and uneven brightness. The pore extraction methods used in practical engineering still have the problems of requiring a lot of manual interaction and low segmentation accuracy. In order to solve these problems, this paper proposes a segmentation method based on dual decoder network which uses image preprocessing network branches to assist in training image segmentation network. The number of model parameters in this method is only 33.3 MB. Besides, we construct a core CT image segmentation dataset. Experimental results show that the pixel accuracy PA can reach 91.41%, and the average detection intersection and merge ratio $MIoU$ can reach 85.32%, which prove that our method has the advantages of small model, fast inference speed, and high segmentation accuracy.

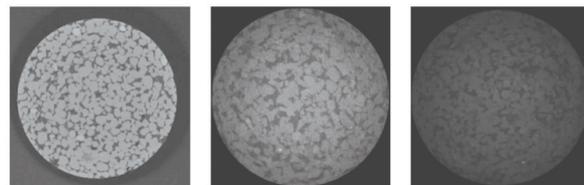
Key words: core CT image; image segmentation; deep learning

0 引言

研究储集层的孔隙空间结构对于指导油气资源的勘探和开发至关重要^[1], 对储集层孔隙结构进行量化分析, 能够为提高采收率和预测开采能力提供基础数据。

随着数字岩心的发展和计算机断层扫描 (Computer Tomography, CT) 技术^[2] 的普及, 岩心 CT 数字图像处理技术为研究人员提供了一种了解内部孔隙结构的新方法。通过分析岩心孔隙的二维图像并构建三维孔隙模型, 可以深入研究孔隙内部流体

的作用机理。因此, 首先需要对岩心 CT 图像孔隙部分进行准确的分割。如图 1 所示, 岩心 CT 孔隙图像具有噪点多、对比度低、亮度不均匀等分割难点。



(a) 正常孔隙图像 (b) 亮度不均孔隙图像 (c) 低对比度孔隙图像

图 1 不同情况的岩心 CT 孔隙图像

Fig. 1 CT pore images of different conditions in rock cores

基金项目: 国家自然科学基金(62071315)。

作者简介: 陈忠照(1998-), 男, 硕士研究生, 主要研究方向: 图像处理; 吴晓红(1964-), 女, 博士, 副教授, 主要研究方向: 图像处理、模式识别; 何海波(1968-), 男, 学士, 高级工程师, 主要研究方向: 图像处理、计算机应用。

通讯作者: 滕奇志(1961-), 女, 博士, 教授, 主要研究方向: 图像处理、模式识别。Email: qzteng@scu.edu.cn

收稿日期: 2023-10-17

目前,工程中分割岩心 CT 孔隙图像最常用的方法是基于阈值的最大类间方差法(OTSU)^[3],该方法通过选取最佳阈值,将图像分为目标和背景两部分,使得对应的类间方差值最大。对于亮度不均和对比度低的岩心 CT 孔隙图像,还需要先进行预处理矫正图像。直方图均衡化虽然可以增强图像的对比度,但同时也会放大岩心 CT 孔隙图像中的噪点;若采用对数变换或伽马变换矫正亮度不均图像,则需要同时根据图像情况手动设置变换参数。

近年来,随着深度学习的快速发展,相较于传统方法,基于深度学习的一些方法极大地改善了图像预处理和图像分割的质量。针对低照度图像增强问题,Tang 等^[4]提出了瓶颈引导通道注意(BGCA)模块,通过低光图像增强网络实现统一的通道信息,以实现更好的颜色恢复;Zhao 等^[5]提出了一种端到端颜色通道融合网络,设计了多尺度特征通道洗牌模块(MFCS)来组合不同尺度的图像特征,使得融合图像包含更丰富的信息,最后经过细节增强后生成输出图像。针对灰度 CT 图像分割问题,Liu 等^[6]提出了一种轻量级且精确的 U-Net 模型,将基准 U-Net 模型的特征图优化到最佳点,以在参数减少和分割性能之间取得平衡;Dabass 等^[7]提出了一种基于残差卷积和注意力学习的 U-Net 模型,采用残差卷积代替编码器和解码器中使用的传统卷积层,以及在传统 U-Net 架构的跳跃连接中实现的注意机制,从而增强了特征表示能力。但采用先图像增强再图像分割的两阶段处理过程,需要经过两个网络模型,不易于模型的训练和部署,因此本文提出了一种双解码器的岩心 CT 孔隙图像分割技术。

1 本文方法

本文提出的双解码器网络结构模型,将图像预处理网络和图像分割网络相结合,用图像预处理网络分支来辅助图像分割网络进行学习。在实际推理运用时,只需要部署图像分割网络,达到提高精度、简化模型的目的。编码阶段由 3 层 DenseBlock 模块来提取特征,解码阶段加入通道注意力模块融合对应层特征,并采用通道随机丢弃算法加强模型的学习能力。设计基于预处理图像的结构相似度损失和均方误差损失,联合基于分割图像的交叉熵损失 BCE Loss,通过设定不同的权重来共同约束编码过程。

1.1 双解码器网络结构

本文提出的双解码器网络包括一个编码器

Encoder 和两个解码器 DecoderSeg、DecoderPre,如图 2 所示。

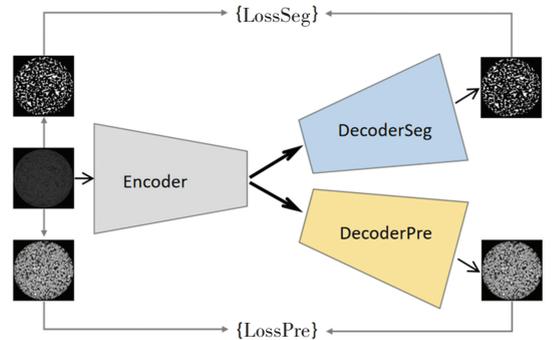


图 2 网络模型算法示意图

Fig. 2 Overview of the pipeline of our method

编码器 Encoder 接收输入孔隙图像,并提取特征输出编码信息,解码器 DecoderSeg、DecoderPre 分别将编码信息进行解码,经过各自的损失函数反向传播以后共同约束编码器的学习过程,总的约束损失函数由分割损失 LossSeg 和矫正损失 LossPre 加权组成,权重 λ_1 和 λ_2 由多次实验得出,见式(1):

$$\text{LossTotal} = \lambda_1 \text{LossSeg} + \lambda_2 \text{LossPre} \quad (1)$$

其中, LossPre 采用基于预处理图像的均方误差损失 MSE 和结构相似度损失 SSIM, LossSeg 为基于分割图像的交叉熵损失 BCELoss。

对于大小为 $m \times n$ 的图像 x 和 y ,均方误差损失 MSE 的计算见式(2)

$$\text{MSE}(x, y) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i, j) - y(i, j)]^2 \quad (2)$$

结构相似度损失 SSIM 的计算见式(3):

$$\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3)$$

其中, μ_x, μ_y 分别为 x, y 的平均值; σ_x^2, σ_y^2 分别为 x, y 的方差; σ_{xy} 是 x 和 y 的协方差; c_1, c_2 均为常数。

交叉熵损失 BCELoss 的计算见式(4):

$$\text{BCELoss}(x, y) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} -[y(i, j) \log x(i, j) + (1 - y(i, j)) \log(1 - x(i, j))] \quad (4)$$

1.2 编码器模块

编码器采用 DenseBlock 模块作为特征提取单元,与通过将输入和输出相加来构建残差块的 Residual (Res) 模块^[8]不同,在 DenseBlock 模块中,每一层都接收前面所有层的特征映射作为输入,并将自己的特征映射传递到后续的层中,从而形成一

种全互连接的结构^[9],每层所提取出来的特征映射均可供后续层使用,如图3所示。因此,DenseBlock具有更强的信息流动性,可以帮助模型更好地利用前面层的信息,逐级挖掘浅层细节信息,形成密集连接,增强特征传播,从而提高模型的性能。

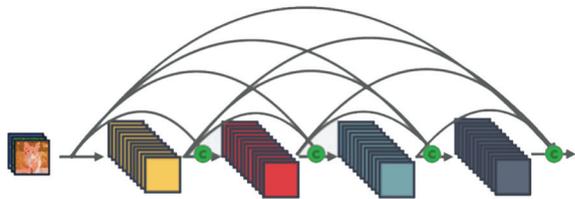


图3 DenseBlock 模块示意图

Fig. 3 Illustration on the DenseBlock module

1.3 解码器模块

解码器借鉴 Unet^[10] 的解码部分,Unet 的解码器部分能够实现精确定位,主要负责特征图的上采样。通过对特征图做上采样以及卷积操作后,同时融合了编码部分对应的特征图,增加了输出的分辨率,从而逐步恢复特征图像的像素信息及图像精度。Unet 将特征图从编码器的每一层传递到解码器的类似层,使得分类器可以考虑不同规模和复杂性的特征来做出决策,其网络结构如图4所示。

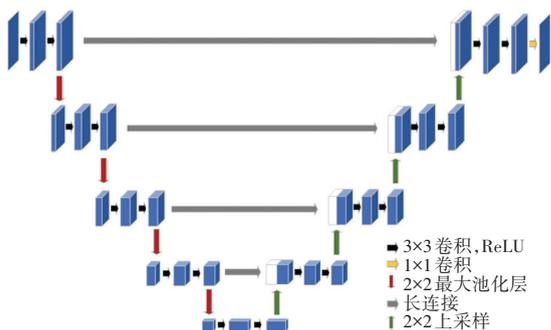


图4 UNet 网络结构示意图

Fig. 4 Structure of the UNet model

1.4 注意力机制

浅层特征经过 DenseBlock 模块后会得到密集特征层,对于图像分割来说并不是所有通道特征都能提供有效的分割信息。SE 通道注意力机制能够帮助网络自适应地调整各个通道之间权重关系,识别出最关键的特征信息,并根据其重要性自适应地调整通道特征的权重,以引导模型将更多的注意力集中在那些对分割结果影响最大的图像区域上,从而提高分割的准确性。SE 通道注意力机制示意图如图5所示。

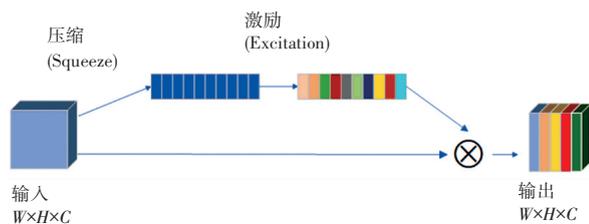


图5 SE 通道注意力机制示意图

Fig. 5 Illustration on the SE channel attention mechanism

SE 模块分为压缩 (squeeze) 和激励 (excitation) 两个阶段。如式(5)所示,压缩阶段使用全局平均池化(Global Average Pooling, GAP)计算出上一个特征图的全局语义信息^[11]。

$$z = \frac{1}{wh} \sum_{i=1}^h \sum_{j=1}^w t(i, j) \quad (5)$$

其中, z 表示经过 GAP 运算将通道中的特征图压缩成的一个实数; t 表示通道中的特征图; w 和 h 为特征图的宽和高。通过对每个通道进行压缩操作就可以得到一个含有全局语义特征的实数矩阵^[12]。

激励阶段主要完成对各通道压缩得来的特征信息进行权重分配,这一过程的表述见式(6):

$$s = \sigma(w_2 \delta(w_1 z)) \quad (6)$$

其中, z 为压缩阶段的输出并含有全局语义特征的实数矩阵; σ 和 δ 分别是 Sigmoid 和 ReLU 激活函数; w_1 和 w_2 是全连接层的权重矩阵。

1.5 通道随机丢弃算法

为了防止训练阶段的过拟合现象,通常会在卷积层后面加一层 Dropout,随机让特征图的部分像素点置零。对于图像数据而言,其是由许多像素点所组成,这些像素点通常都具有空间相关性,相邻像素点的像素值往往都是相近的。如果随机丢弃特征图上的像素点,依然可以通过相邻的像素点来传递特征信息,使得防止过拟合效果并不显著。

本文引入 Spatial Dropout^[13] 代替传统的 Dropout,不是随机将特征图像素点置零,而是将特征通道随机置零。这种做法有助于防止某些特征通道在经过 SE 模块后一直被赋予过高的权重,从而导致过度拟合,使得模型学习到更多的孔隙图像特征,提高模型的泛化能力。

1.6 整体网络模型

本文所用的整体网络的结构框架如图6所示,网络模型整体采用3层编解码模式,相比于 Unet 标准的5层编解码模式,减少了模型的参数量。编码阶段由3个编码模块 Encoder1-3 组成,采用 DenseBlock 模块来提取特征。具体地,将原始的岩

心 CT 孔隙图像作为输入, 并先通过一个卷积层进行粗精度的提取特征, 接着依次通过稠密连接模块和下采样模块, 构建 3 层图像特征金字塔; 解码阶段也由对应的 3 个解码模块 Decoder1-3 组成, 采用通道注意力机制对编码信息进行特征筛选, 并加入通

道随机丢弃机制, 增强模型的特征学习能力。解码模块 Decoder1-3 将上采样特征和编码模块 Encoder1-3 的编码特征进行通道拼接, 经过通道注意力模块融合特征信息, 并通过空洞卷积模块^[14], 进行上采样恢复为原始分辨率。

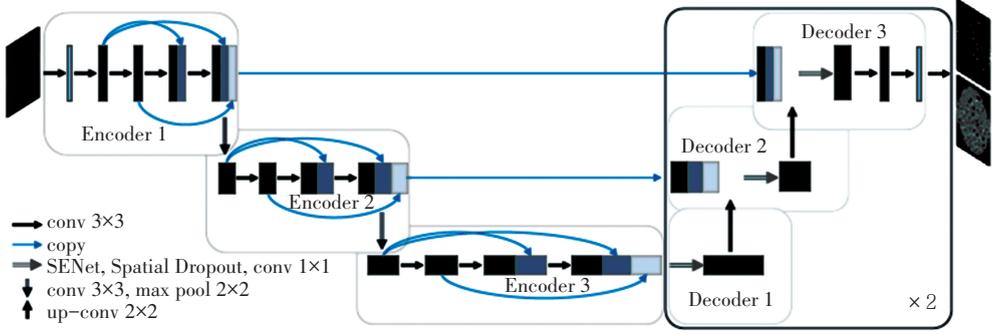


图 6 整体网络结构示意图

Fig. 6 Structure of the overall network

2 实验结果与分析

2.1 实验数据集

本文数据集标签对由一幅原图, 及其对应的矫正图和分割图组成, 如图 7 所示。使用对比度受限的自适应直方图均衡化(CLAHE)^[15] 手动调整适合参数, 并消除黑色背景区域对算法的影响, 得到对比度良好、亮度均匀的矫正图像, 再使用人工标记分割的方法得到孔隙分割图像。从 15 组岩心 CT 孔隙图像中各选取 800 幅图像, 将图像大小归一化为 960×960, 共制作了 12 000 个样本, 按 7:3 的比例划分训练集和验证集, 由于 CT 图像同一组之间存在层间相关性, 为保证测试集与训练集的差异性, 避免过拟合, 另选取 5 组岩心 CT 孔隙图像制作 4 000 张数据集作为测试集。

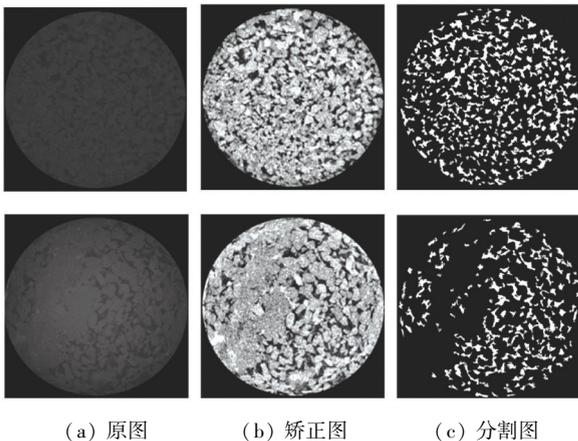


图 7 数据集示例

Fig. 7 Examples in the dataset

2.2 实验环境及训练参数

本文实验环境为 Ubuntu 18.04, CPU 为 Dual Intel Xeno E5-2600, GPU 为 NVIDIA GTX 1080Ti (Pascal)。在 PyTorch 1.9.0, cuda10.2 下完成模型的搭建和网络训练。

在训练阶段采用 SGD 作为优化器, 初始学习率为 1×10^{-4} , BatchSize 设置为 2, 共训练 100 个周期, 学习策略采用“burn-in”方法^[16], 在训练开始后一定数量的更新周期内逐渐增加学习率, 旨在执行一些随机或初始搜索, 有助于优化算法更好地收敛到全局最优解。计算公式如式(7):

$$lr = base_lr * (batch_num / burn_in)^{power} \quad (7)$$

其中, $base_lr$ 是基准学习率; $batch_num$ 是当前的批次; $burn_in$ 是指学习率从 0 逐渐上升到初始学习率所需的批次; $power$ 用于控制学习率变化速率。在超过 $burn_in$ 后的更新周期内, 学习速率将按照预设的策略逐渐减小。

为了在训练中加快网络收敛速度、防止过拟合现象发生, 本文使用 Batch Normalizing 标准化操作, 对样本进行标准化处理后, 该样本的输出结果不仅仅受到自身特征的影响, 还会被同属于一个 batch 中的其他样本所影响^[12]。这可以防止整个网络朝着同一个方向学习, 确保网络的学习过程更加稳定和可靠。

2.3 模型评价指标

为了客观评价预测图与标签图的相似程度, 实验采用像素精度 (Pixel Accuracy, PA) 以及平均交并比 (Mean Intersection over Union, MIoU) 作为模型

的客观评价指标。 PA 、 $MIoU$ 的表达式如下:

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

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

其中, TP 为真正值, TN 为真负值, FP 为假正值, FN 为假负值。像素精度 PA 是指正确分类的像素占所有像素的百分比, 从宏观角度描述了预测结果的准确率; $MIoU$ 是指预测图与标签图交集与并集的比例, 反应了预测图与标签图的重合情况, 从细节上描述了预测结果的准确率。

此外, 为了评价模型的大小和推理速度, 采用模型参数量 (Model Size) 和每秒推理请求数 IPS (Inferences Per Second) 作为评价指标。由于不同的硬件平台、模型大小和输入数据等因素都会影响 IPS 的结果, 本实验在同等硬件条件下, 计算各模型对 1 500 幅大小为 960×960 的岩心 CT 序列图的推理总时间, 并取平均值来获得更精确的结果。

2.4 结果分析

为了客观验证本文所用网络的分类性能, 使用同一数据集分别在 MobieNet^[17]、PSPNet^[18]、DeepLabv3^[19]、UCTransNet^[20]、UNet2023^[21] 等主流分割网络上进行了对比实验。不同网络模型在 PA 、 $MIoU$ 、Model Size 和 IPS 评价指标上的结果见表 1。

表 1 不同网络模型下孔隙图像分割结果评价指标

Table 1 Evaluation results for pore image segmentation of different models

网络模型	PA / %	$MIoU$ / %	Model Size/ MB	IPS / (次·s) ⁻¹
MobieNet	80.54	74.33	57.8	3.28
PSPNet	84.36	76.15	178	0.96
DeepLabv3	85.47	80.20	172.7	0.98
UCTransNet	92.64	85.72	278.8	0.54
UNet2023	88.03	82.75	136.5	1.12
本文方法	91.41	85.32	33.3	5.32

根据表 1 中数据可见, 由于本文提出的双解码器卷积神经网络模型利用图像预处理网络分支辅助分割, 将稠密连接层嵌入到编码阶段, 将通道注意力模块嵌入到解码阶段, 以达到更好的融合各通道的特征信息, 更精确的提取孔隙区域, 在 PA 和 $MIoU$ 指标上分别达到了 91.41% 和 85.32%, 仅比 UCTransNet 模型低 1.23% 和 0.4%, 明显优于其他网络模型。在 Model Size 和 IPS 指标上, 模型参数量 Model Size 为 33.3 MB, 比 MobieNet 网络模型还小

24.5 MB。实际工程运用中, 常需要一次性对成百上千幅岩心 CT 孔隙图进行分割, 在对 1 500 幅大小为 960×960 的岩心 CT 孔隙图进行推理预测时, UCTransNet 模型总用时为 46 min, 而本文算法模型总用时仅 5 min, 推理速度显著快于各主流网络模型。

最后, 为了直观地感受本文算法模型的分割效果, 图 8 给出了最大类间方差法、本文算法模型和 UCTransNet 模型的岩心 CT 孔隙图像分割结果对比, 并将分割结果图做伪彩色处理叠加到原图上。可以看出, 基于阈值分割的最大类间方差法会把岩石颗粒间的噪点过度分割出来, 而孔隙中间又分出来很多空洞, 破坏了孔隙的完整性。由深度学习模型自动化分割出的结果符合人眼视觉的主观感受, 孔隙边缘清晰, 孔隙内部完整, 噪点较少, 孔隙分割效果良好。在面对亮度不均匀的孔隙图像时, 最大类间方差法很难将孔隙分割出来, 而本文算法模型和 UCTransNet 模型具有良好的泛化性, 能够学习到孔隙图像的特征信息, 取得较为准确的分割结果。

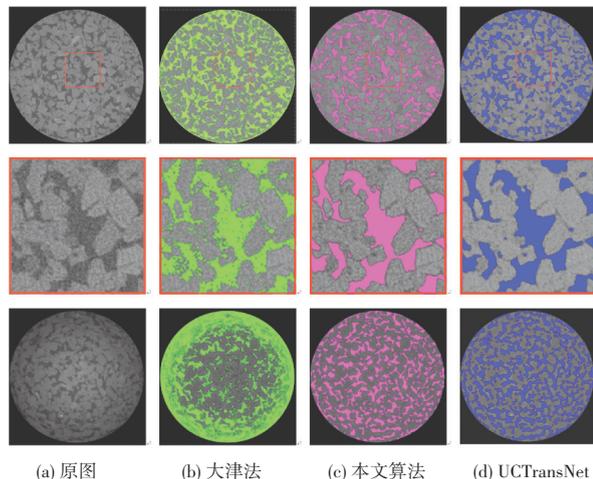


图 8 分割结果对比

Fig. 8 Comparison of segmentation results

3 结束语

本文针对传统岩心 CT 孔隙图像分割方法的准确度不高且需要大量人工交互的问题, 提出了一种双解码器网络的岩心 CT 孔隙图像分割方法, 在辅助分支的监督学习下, 网络模型训练完成后, 编码器已经能将低对比度或光照不均的孔隙图像编码为更利于分割的高维特征信息, 在实际推理应用中只需要部署编码器和分割解码器, 达到减小模型参数量的目的。对比实验表明, 分割结果不论是在主观效果上, 还是在客观指标上都有一定的优势。对多组

不同的岩心 CT 孔隙图像进行实验,均取得了较好的分割结果。实验结果表明,该方法模型参数量小,同时分割精度高,后期将继续优化网络模型,将算法应用于实际工程中。

参考文献

- [1] 王润涵,李兵,滕奇志.基于卷积神经网络的岩心 FIB-SEM 图像分割算法[J].计算机工程,2021,47(1):264-274.
- [2] HOUNSFIELD G N. Computerized transverse axial scanning (tomography): Part I. Description of system. 1973 [J]. The British Journal of Radiology, 1995,68(815):66-72.
- [3] OTSU N. A threshold selection method from gray level histograms [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979,9(1):62-66.
- [4] TANG P, GUO X, JU G, et al. Integration - and - diffusion network for low - light image enhancement [C] // Proceedings of 2021 IEEE International Conference on Image Processing (ICIP). IEEE, 2021: 1664-1668.
- [5] ZHAO L, GONG X, LIU K, et al. Color channel fusion network for low - light image enhancement [C] // Proceedings of 2021 IEEE International Conference on Image Processing (ICIP), IEEE, 2021: 1654-1658.
- [6] LIU Q, ZHOU S, LAI J. EdgeMedNet: lightweight and accurate U-Net for implementing efficient medical image segmentation on edge devices [J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2023,70(12):4329-4333.
- [7] DABASS M, CHANDALIA A, GUPTA H, et al. Lung segmentation in CT scans with residual convolutional and attention learning - based U-Net [C] // Proceedings of 2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON). IEEE, 2023: 240-245.
- [8] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 770-778.
- [9] 陈立潮,朝昕,潘理虎,等.基于部件关注 DenseNet 的细粒度车型识别 [J]. 智能系统学报, 2022,17(2):402-410.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C] // Proceedings of the 18th International Conference on Medical Image Computing and Computer - Assisted Intervention - MICCAI 2015. Munich: Springer International Publishing, 2015:234-241.
- [11] MIN Lin, CHEN Qiang, YAN Shuicheng. Network In Network [C] // Proceedings of the International Conference on Learning Representations (ICLR), 2013.
- [12] 倪黎,邹卫军.基于 SE 模块改进 Xception 的动物种类识别 [J]. 导航与控制, 2020,19(2):106-111.
- [13] ZHANG J, KONG X, LI Xueyi, et al. Fault diagnosis of bearings based on deep separable convolutional neural network and spatial dropout [J]. Chinese Journal of Aeronautics, 2022,35(10):301-312.
- [14] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolution [J]. arXiv preprint arXiv:1511.07122, 2015.
- [15] 梁爽,肖章林. CLAHE 融合低频 DCT 系数重变换的人脸识别算法 [J]. 计算机仿真, 2015, 32(8):255-259, 275.
- [16] Sutskever, Ilya, et al. On the importance of initialization and momentum in deep learning [C] // Proceedings of the 30th International Conference on Machine Learning (ICML). 2013: 1139-1147.
- [17] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv abs/1704.04861, 2017.
- [18] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 2881-2890.
- [19] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C] // Proceedings of the European Conference on Computer Vision (ECCV). IEEE, 2018:801-818.
- [20] WANG H, CAO P, WANG J, et al. UCTransNet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022,36(3):2441-2449.
- [21] D'ALESSANDRO V I, DE PLAMA L, ATTIVISSIMO F, et al. U-Net convolutional neural network for multisource heterogeneous iris segmentation [C] // Proceedings of 2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, 2023:1-5.