

改进 UNet-VAE 网络的土壤多类型孔隙三维分割方法

韩巧玲^{1,3,4,5}, 宋美慧^{1,3,5}, 席本野², 赵 玥^{1,3,4,5*}, 赵燕东^{1,3,5}

(1. 北京林业大学工学院, 北京 100083; 2. 北京林业大学林学院, 北京 100083; 3. 北京市教育委员会城乡生态环境北京实验室, 北京 100083; 4. 国家林业局林业装备与自动化国家重点实验室, 北京 100083; 5. 北京林业大学智慧林业研究中心, 北京 100083)

摘要: 不同类型土壤孔隙结构会随生物活动和非生物作用发生形变, 从而影响其生态功能, 准确分割不同类型的孔隙结构对于研究孔隙结构与生态功能演变关系具有重要意义。针对单个类别孔隙的分割方法分割精度低、分类标准单一、鲁棒性差, 无法准确分割和判别生物孔隙、裂隙等相交部分孔隙结构的问题。该研究针对不同类型孔隙尺度差距大的特点, 提出了一种改进 UNet-VAE 网络模型, 实现土壤多类型孔隙分割。改进 UNet-VAE 网络引入多尺度特征融合注意力模块, 以实现多尺度信息融合和冗余信息筛选。结合变分自动编码器生成网络 (variational autoencoder, VAE), 引入噪声和辅助损失函数, 以增强网络的泛化能力和鲁棒性。试验结果表明: 改进 UNet-VAE 方法在土壤多类型孔隙 (裂隙、生物孔、不规则孔隙和球状孔隙) 三维分割中达到了 93.83% 的平均准确率, 与次优 VNet 方法相比, 平均准确率、精确率、召回率和 F1 值分别提升了 3.32, 5.06, 8.97 和 8.63 个百分点, 特别是对于不规则孔隙 4 项指标分别提升了 4.88, 15.46, 15.70 和 15.50 个百分点, 表明改进 UNet-VAE 法可准确分割多类型孔隙, 也验证了深度学习技术在多类型孔隙判别的可行性, 可为揭示土壤孔隙结构与演化研究提供有效工具。

关键词: 土壤; 孔隙; 卷积神经网络; 图像分割; 3D UNet; 注意力模块

doi: 10.11975/j.issn.1002-6819.202311075

中图分类号: S152; TP391.41

文献标志码: A

文章编号: 1002-6819(2024)-18-0081-09

韩巧玲, 宋美慧, 席本野, 等. 改进 UNet-VAE 网络的土壤多类型孔隙三维分割方法[J]. 农业工程学报, 2024, 40(18): 81-89. doi: 10.11975/j.issn.1002-6819.202311075 <http://www.tcsae.org>

HAN Qiaoling, SONG Meihui, XI Benye, et al. Three-dimensional segmentation method of soil multi-category pores based on improved UNet-VAE network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(18): 81-89. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202311075 <http://www.tcsae.org>

0 引言

土壤孔隙是容纳水分和空气的空间, 它在促进一系列土壤生态功能的重要过程中发挥着重要作用, 如为微型和中型动物提供栖息地, 支持水、气、营养物质的存储和运移, 以及其他关键过程^[1-4]。孔隙的大小和形状对其生态功能起着决定型作用^[5-6]。因此, 根据形态对孔隙进行准确分类具有重要的意义。

近年来, 基于计算机断层扫描 (computed tomography, CT) 进行孔隙分类、以及土壤孔隙结构和功能关系的研究越来越广泛^[5-8]。前人基于土壤 CT 图像, 对孔隙进行了大量的研究, 如 CAPOWIEZ 等^[9] 基于体积阈值对孔隙开展了分类; ZHANG 等^[10] 进一步引入形状因子消除了大体积的非生物孔隙。然而, 有研究表明, 海绵状裂隙和管状生物孔隙之间存在紧密连接^[11], 难以通过传统图像分割方法进行分割, 如 ROONEY 等^[12] 分割后仍发

现与生物孔垂直相交的片状孔隙; 由于外力活动生物孔隙破碎生成的小体积孔隙, 具有连通性差等特征, 在水气运移方面与非生物孔的功能更相似^[13]。由此可见, 有必要结合孔隙形态与形成原因分类, 对不同的孔隙功能进行研究。

生物孔能提高透水性、溶质运移和渗透率^[14-15]。耕作、冻融或干湿等非生物活动形成的非生物孔形状各异^[16]: 形成的呈片状的裂隙可作为优先流路径, 增加农田水分和养分的流失以及地下水污染的风险^[17]; 呈不规则形态的孔隙, 有利于储水储气^[18]; 球状孔隙由土壤干燥过程中空气截留形成^[19]。基于孔隙形态将土壤孔隙分为生物孔、裂隙、不规则孔隙和球状孔隙, 可以为研究孔隙结构与功能演变提供数据基础。

孔隙分割技术发展较快且取得了一定成效, 但也面临着若干技术层面的挑战: 1) 仅使用特征阈值进行多类别分割难以区分不同类别的相交孔隙, 尤其是常出现相交的裂隙和生物孔^[11-12,20]; 2) 受孔隙多尺度特征影响, 目前分割方法对于单个类别孔隙分割精度低; 3) 现有三维分割方法自动化程度低, 需要手动多次调试每个 CT 图像, 选择适当阈值以确保准确分割的重要性^[21]。近年来, 深度学习在土壤科学领域中, 逐渐得到了广泛的应用^[22-23], 为解决这些挑战提供了一种有效的思路和方法。

收稿日期: 2023-11-09 修订日期: 2024-03-26

基金项目: 国家自然科学基金面上项目 (32071838); 国家自然科学基金青年项目 (32101590)

作者简介: 韩巧玲, 副教授, 研究方向为生态信息智能检测、图像处理与模式识别等。Email: hanqiaoling0@163.com

*通信作者: 赵玥, 博士, 教授, 研究方向为图像处理与模式识别、机器学习、人脸识别、大数据分析等。Email: zhaoyue0609@126.com

深度学习能够实现体素化分割, 解决相交孔隙难以界定的问题; 通过训练模型学习多尺度特征, 从而提高分割精度; 并且实现自动特征提取和分割, 无需手动调试和选择阈值。在经典卷积网络方面, 韩巧玲等^[24]基于简化卷积网络进行孔隙分割。FU 等^[25]结合半监督训练模型、多尺度感受野结构提高孔隙多尺度特征信息提取能力。虽然深度学习从多个方向不断提升分割精度, 但均是针对整体孔隙进行分割, 无法实现多类型孔隙高精度分割, 且仍存在多尺度孔隙特征难以学习及分割方法泛化性差等问题。

为解决阈值法相交孔隙难以分割、精度和自动化程度较低问题, 实现土壤多类型孔隙精确分割, 本文提出了一种改进 UNet-VAE 网络模型, 以分割生物孔、裂隙、不规则孔隙和球状孔隙 4 类土壤孔隙。该模型在 3DUNet 网络基础上, 提出多尺度融合注意力模块, 融合多尺度信息、筛选冗余信息, 并在训练中通过变分自编码器 (variational autoencoders, VAE) 分支引入噪声和辅助损失函数, 以增强网络泛化能力、提高网络鲁棒性, 以期土壤孔隙结构精细量化表征提供数据基础, 为揭示土壤孔隙演化在生态系统中的作用提供科学依据。

1 材料与方法

1.1 图像采集与预处理

1.1.1 数据获取

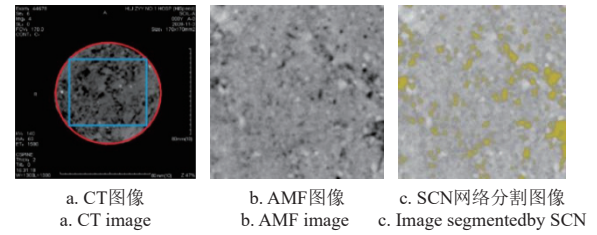
土壤样品采集于中国黑龙江省克山农场, 样品内径为 10 cm, 高度为 10 cm。于土层深为 0~40 cm 处分别修整黑土剖面。通过机械分层法, 以 10 cm 的间距为标准使用环刀 (100 cm³) 逐层进行原状黑土取样。采集的黑土样本质地为黏壤土, 有机质含量为 65 g/kg, 大孔隙率为 54.95%, 平均体积含水率为 31.71%。随机采取 3 个土样进行冻融处理, 分别进行 1、3、5 次冰冻和融化^[26]。采用黑龙江中医药大学第一附属医院 CT 扫描中心的螺旋 CT 扫描仪对样品进行扫描并获取高分辨图像 (每像素点实际长度为 0.236 mm)。每个土壤样本的每次处理产生 220 个横截面图像, 共产生 21 组图像, 详细处理方法见文献^[27], 将数据集按照 8:1:1 划分为训练集、验证集和测试集。

1.1.2 数据预处理

土壤的原始 CT 图像、矫正图像和初始分割图像如图 1 所示。土筒和 PVC 玻璃管在运输过程中的接触振动, 会影响土壤边界孔隙结构, 因此使用内切法对土壤 CT 图像进行裁剪以去除干扰区域, 保留图 1a 中的蓝框区域。通过自适应中值滤波方法去除孤立的噪声点, 增强孔隙的边界特征, 确保 CT 图像的可用性, 如图 1b 所示。滤波后的图像集为后续土壤孔隙分割和标记任务奠定数据基础。

为了获得精确的土壤孔隙结构, 使用 SCN 方法^[24]进行土壤孔隙分割并建立孔隙分割数据集。SCN 方法的土壤孔隙分割平均准确率为 99.61%, 分割图像如图 1c

所示。



注: 蓝框代表裁剪范围; 红圈代表实际扫描范围。AMF 为自适应中值滤波处理。SCN 为简化卷积网络。

Note: Blue frame represents the cropping range; Red circle represents the actual scanning range. AMF is adaptive median filtering. SCN is simplified convolutional network.

图 1 土壤 CT 孔隙分割数据集过程示意图

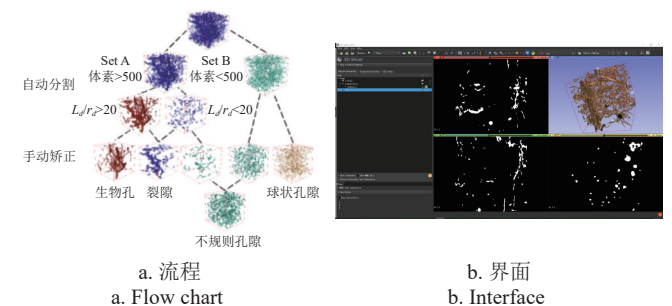
Fig.1 Diagram of soil CT pore segmentation dataset process

1.1.3 孔隙数据集标注

基于分割后的孔隙数据集, 采用自动分割和手动校正相结合的方法进行 4 类土壤孔隙结构的标定。裂隙是弯曲的、不同宽度的片状结构^[16]; 生物孔包括非分支、分支和网状孔结构^[16]; 不规则孔隙呈现随机的形状和大小, 外壁可能是有角度的或平滑弯曲的^[28]; 球状孔隙尺寸较小^[19]。

标定流程如图 2a 所示, 首先, 采用 Python 软件, 结合三维特征, 体素、形状因子 (孔长 L_d 与等效孔半径 r_d 的比值) 进行自动分割^[13]。然后, 使用图 2b 所示三维标定软件 3D Slicer, 手动校正标定相交的裂隙、生物孔及不同特征的非生物孔隙, 建立了裂隙、生物孔、不规则孔隙和球状孔隙的真值数据集。其中裂隙、生物孔、不规则孔隙和球状孔隙的体素分别标记标签 1、2、3 和 4, 其余体素被视为背景 (标记为空)。为了可视化 4 类孔隙空间分布及特征, 对 4 类孔隙分配不同颜色: 蓝色代表裂隙, 红色代表生物孔, 绿色代表不规则孔隙, 黄色代表球状孔隙。

每个多类型孔隙三维标定数据都由 5 名具有土壤物理学知识背景人员进行重复标定, 以消除主观性对标定精度的影响。



注: L_d 为孔长; r_d 为等效孔半径。

Note: L_d is pore length; r_d is equivalent pore radius.

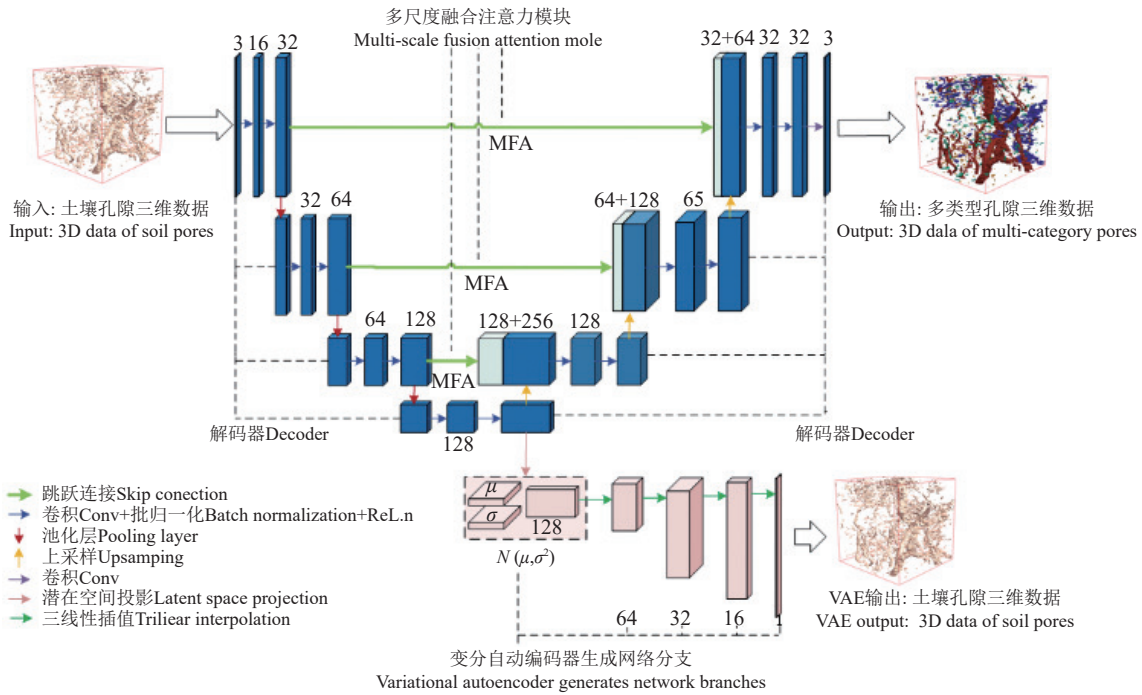
图 2 土壤孔隙标定流程和 3D Slicer 标定界面演示图

Fig.2 Soil pore calibration process and 3D Slicer calibration interface demonstration diagram

1.2 改进 UNet-VAE 土壤孔隙三维多类别分割模型构建
将 3D UNet 中 UNet 二维卷积替换为三维卷积, 实

现多感受野俘获整体空间信息进行三维分割。由于使用跳跃连接直接融合编码器与解码器输出的语义信息，吸取了大量冗余信息，导致其对多尺度土壤孔隙分割能力效果较差，尤其对于不规则孔隙难以辨别。为解决上述问题，本文基于 3D UNet 进行改进，提出改进 UNet-VAE 土壤多类型孔隙三维分割网络，网络结构如图 3 所示。编码器通过下采样卷积进行特征提取、解码器通过上采样卷积恢复图像并进行类别分割。提出了多尺度融合注

意力模块来筛选由于卷积学习产生的冗余信息；通过局部注意力学习小尺度孔隙（不规则孔隙和球状孔隙）空间特征；通过全局注意力提取大尺度孔隙（裂隙和生物孔）特征信息，从而融合不同类型孔隙多尺度特征，以提高不同类型孔隙的分割精度。同时，在训练中通过 VAE 分支引入噪声从而避免过拟合；引入 VAE 辅助损失函数，通过对潜在空间中的向量进行约束，使得网络学得不同类型孔隙的特征分布，从而提高网络的泛化能力。



注：MFA 为多尺度融合注意力模块；图中数字表示空间维度数； $N(\mu, \sigma^2)$ 是均值为 μ ，方差为 σ^2 的多元正态分布。

Note: MFA is the multi-scale fusion attention module; the number represents the spatial dimension; $N(\mu, \sigma^2)$ is a multivariate normal distribution with a mean of μ and a variance of σ^2 .

图 3 改进 UNet-VAE 网络结构

Fig.3 Structure of improved UNet-VAE network

1.2.1 多尺度融合注意力模块

为融合多尺度特征信息、减少冗余信息，本研究提出多尺度融合注意力模块（multi-scale fusion attention, MFA），其结构如图 4 所示。该模块在跳跃连接中加入基于注意力门的多尺度特征融合注意力（attentional feature fusion, AFF）。

MFA 解码器侧输入特征图 G 经上采样后尺度与编码器侧输入特征图 X 一致，皆为 $C_X \times H_X \times W_X \times L_X$ （ C_X 为编码器侧输入特征图 X 的通道数， H_X 为特征图 X 的高度， W_X 为特征图 X 的宽度， L_X 为特征图 X 的深度）。通过特征相加，实现初始特征融合，并通过 ReLU 激活函数和逐点卷积得到筛选冗余信息后的注意力特征图 X_A 。计算式如下：

$$X_A = \text{PWConv}_1(\delta(\text{UPConv}(G) + X)) \quad (1)$$

式中 UPConv 上卷积使 C_g 通道数减半， H_g 、 W_g 、 L_g 分别扩大 2 倍（ C_g 为解码器侧输入特征图 G 的通道数， H_g 为特征图 G 的高度， W_g 为特征图 G 的宽度， L_g 为特征图 G 的深度）；PWConv₁1×1 点卷积将 C_X 通道数缩减

为 1，即 $X_A \in 1 \times H_X \times W_X \times L_X$ ； δ 表示 ReLU 激活函数。

注意力特征图 X_A 与 X 相乘，得到筛选冗余信息后的特征图 X' 。

$$X' = X \otimes X_A \quad (2)$$

特征图 X' 在通道注意力模块中分别使用局部注意力学习细节信息和全局注意力吸取上下文特征； $L(X')$ 计算式如式 (3) 所示：

$$L(X') = B(\text{PWConv}_3(\delta(B(\text{PWConv}_2(X'))))) \quad (3)$$

式中 PWConv₂1×1 点卷积将 X' 通道数减少为原先的 $1/\tau$ ； B 表示 BatchNorm 层；PWConv₃1×1 的卷积将通道数目恢复成与原输入通道数目相同； τ 为通道缩放比。

相比局部特征通道注意力，全局特征通道注意力需要对 X' 先进行全局平均池化操作，其计算式 $G(X')$ 如下：

$$G(X') = B(\text{PWConv}_3(\delta(B(\text{PWConv}_2(\text{Pool}(X')))))) \quad (4)$$

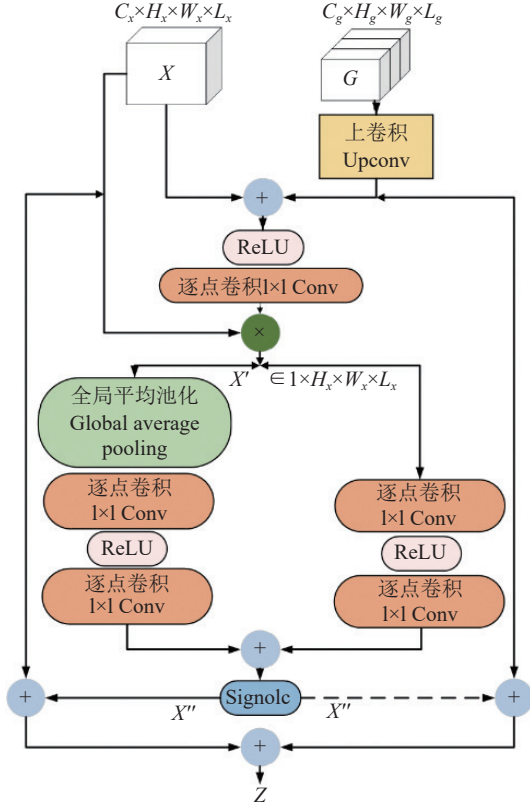
式中 Pool 表示全局池化。

得到 $G(X')$ 之后为 X 和 G 分配注意力权重，得到输出 Z 公式如下：

$$X'' = \text{Sigmoid}(L(X') + G(X')) \quad (5)$$

$$Z = X'' \otimes X + (1 - X'') \otimes G \quad (6)$$

式中 Sigmoid 表示 Sigmoid 激活函数； X'' 为权重注意力特征图。



注： X 是尺度为 $C_x \times H_x \times W_x \times L_x$ 的编码器侧输入特征图， G 是尺度为 $C_g \times H_g \times W_g \times L_g$ 的 MFA 解码器侧输入特征图， X' 是尺度为 $1 \times H_x \times W_x \times L_x$ 的筛减冗余信息后的特征图， X'' 是权重注意力特征图， Z 是最终输出特征图，ReLU 是 ReLU 激活函数，Sigmoid 是 Sigmoid 激活函数。
Note: X is the encoder side input feature map with a scale of $C_x \times H_x \times W_x \times L_x$, G is the MFA decoder side input feature map with a scale of $C_g \times H_g \times W_g \times L_g$, X' is the feature map after filtering redundant information with a scale of $1 \times H_x \times W_x \times L_x$, X'' is the weighted attention feature map, Z is the final output feature map, ReLU is the ReLU activation function, and Sigmoid is the Sigmoid activation function.

图 4 多尺度融合注意力模块结构图

Fig.4 Structure diagram of multi-scale fusion attention module

1.2.2 变分自动编码器生成网络分支

VAE 是一个带有编码器和解码器的生成模型，可将输入数据 x 编码为潜在变量 z ，并将潜在变量解码为输出数据 \bar{x} 。具体来说，VAE 假设潜在变量来自高斯分布，VAE 的编码器学习输入数据的分布参数：均值 (μ) 和方差 (σ^2)。然后从分布参数中采样得到潜在变量 z 并解码为输出数据 \bar{x} 。

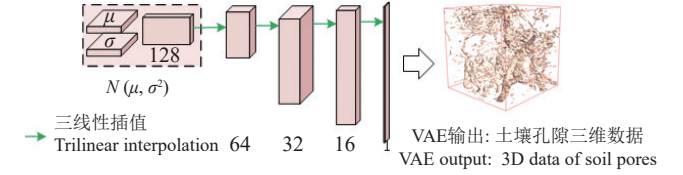
即 VAE 模型通过编码过程 $Q(z|x)$ 将样本映射为潜在变量 z ，并假设潜在变量服从多元正态分布 $P(x) \sim N(0, I)$ ，解码器 $P(x|z)$ 从隐藏变量 z 中抽取样本，生成指定图像 \bar{x} 。通过最大化变分下界共同训练近似后验模型和生成模型。变分下界 $E(q)$ 的表达式为

$$E(q) = E_q[\lg p(x|z)] - D_{\text{KL}}[q(z|x)||p(z)] \quad (7)$$

式中 $\lg p(x|z)$ 是一个条件对数似然，表示在给定潜

在变量 z 的条件下，观测数据 x 的对数概率。 $D_{\text{KL}}[q(z|x)||p(z)]$ 是使用 KL 散度衡量两个概率分布 $q(z|x)$ 和 $p(z)$ 之间差异。 $q(z|x)$ 是给定观测数据 x ，编码器输出的潜在变量 z 的后验分布。 $p(z)$ 是潜在变量 z 的先验分布。

VAE 分支结构如图 5 所示，其主要作用是避免过拟合问题并提高网络的泛化能力。由图 5 可知，编码器端输出数据被减少到 128 的低维潜在空间，其中 64 维度用于表示平均值，64 维度用于表示标准差。从具有给定均值和标准差的高斯分布中提取样本，然后按照与解码器相同的架构重建为输入图像维度。



注：数字为空间维度数。

Note: The number represents the spatial dimension.

图 5 变分自动编码器生成网络解码器结构

Fig.5 Structure of variational automatic encoder generates network decoder

1.2.3 损失函数

损失函数 L 由 Dice 损失和 VAE 损失组合构成：

$$L = L_{\text{Dice}} + 0.1(L_{\text{Rec}} + L_{\text{KL}}) \quad (8)$$

式中 L_{Dice} 为 Dice 损失， L_{Rec} 为计算重建图像和输入图像之间的差距。 L_{Rec} 越小，生成图像与输入图像越相似。 L_{KL} 是估计的正态分布和先验分布之间的 Kullback-Leibler 散度。

选择 0.1 的超参数（正则化因子权重）以在 Dice 损失和 VAE 损失之间提供良好的平衡^[29]。

1) Dice 损失

设 y 和 \hat{y} 分别是分割和模型预测的真值，为了避免训练数据没有标签，如 $y = \hat{y} = 0$ ，将 ε （用于提高函数稳定性的小正数常量）添加到分子和分母中。Dice 损失定义如下：

$$L_{\text{Dice}}(y, \hat{y}) = 1 - \frac{2y\hat{y} + \varepsilon}{y + \hat{y} + \varepsilon} \quad (9)$$

通过 Dice 损失计算分割预测和真值之间的差距，Dice 损失越小分割越准确。

2) VAE 损失

VAE 损失是 VAE 分支上的重建损失 L_{Rec} 和标准 VAE 惩罚项 L_{KL} 的总损失。在本研究中， L_{Rec} 是每个体素的均方误差。

$$L_{\text{VAE}} = L_{\text{Rec}} + L_{\text{KL}} \quad (10)$$

$$L_{\text{Rec}} = \|x_{\text{rebuilding}} - x\|_2^2 \quad (11)$$

式中 $x_{\text{rebuilding}}$ 和 x 分别表示重建图像和输入图像； L_{VAE} 是 VAE 损失； L_{Rec} 是每个体素的均方误差，计算重建图像和输入图像之间的差距。 L_{Rec} 越小，生成图像与输入图像越相似。

由于重构是在潜在空间中采样进行生成。在网络反向传播中，生成器编码器趋势是潜在空间方差逐渐趋近于 0。为了保证潜在空间的随机性，引入 L_{KL} 。

L_{KL} 是估计的正态分布和先验分布之间的 Kullback-Leibler 散度。

$$L_{KL} = \frac{1}{N_{\text{totalvoxels}}} \sum (\mu^2 + \sigma^2 - \lg \sigma^2 - 1) \quad (12)$$

式中 $N_{\text{totalvoxels}}$ 是图像体素的总数； μ 为图像体素的均值； σ^2 为图像体素的方差； $\lg \sigma^2$ 对方差取对数减少对小方差的惩罚；减 1 平衡损失项。 L_{KL} 越小，解码器输出的和潜在空间采样分布越接近，即越接近正态分布，从而防止潜在空间方差为 0。

1.2.4 评价指标

为了更好地评价网络对于多类型土壤孔隙分割性能。本研究以识别准确率 A 、精确率 P 、召回率 R 、F1 分数 (F_1) 以及 4 项指标的平均值进行指标评价，指标数值越大效果越好^[30]。各指标计算式如下：

$$A = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \quad (13)$$

$$P = \frac{T_P}{T_P + F_P} \quad (14)$$

$$R = \frac{T_P}{T_P + F_N} \quad (15)$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (16)$$

式中 T_P 为正确分类孔隙； F_N 为正确分类孔隙分割为其他类型孔隙及背景； F_P 为其他类型孔隙及背景分割为正确分类孔隙； T_N 为其他类型孔隙及背景分割为其他类型孔隙及背景。

1.2.5 试验环境以及参数设置

本研究所有网络模型均在 Featurize 服务器租赁平台上 linux 环境下使用 Pytorch 进行训练。硬件配置如下：内存大小 60.9 GB，CPU 为 30 核 AMD EPYC 7742，显卡型号为 RTX A6000，显存为 51.0 GB，模型训练超参数设置如下：batch size 设置为 1，训练本文所有方法优化器选用 AdamW，学习率为 0.001，权重衰减为 0.001，所有网络模型训练时最大迭代次数为 200。

2 结果与分析

2.1 消融试验结果对比分析

为验证各改进模块在多类型孔隙分割上的有效性，对各改进模块进行消融试验。对比原始 3D UNet，增加 MFA、VAE 和结合 2 个改进点的改进 UNet-VAE 网络模型分割性能，分别统计不同模型识别分类的土壤孔隙（裂隙、生物孔、不规则孔隙和球状孔隙）的准确率、精确率、召回率和 F1 分数。结果如表 1 所示。

从表 1 可以看出，在相同的训练环境下，在加入 MFA 后，4 类孔隙分割都有了明显提升效果。4 类孔隙准确率分别提升了 13.53、15.17、17.04 和 15.94 个百分

点；F1 分数分别提高了 14.18、10.96、24.87 和 27.47 个百分点。对体积较小、形状不规则的不规则孔隙性能提升效果最为明显，F1 分数和准确率分别提升了 24.87、17.04 个百分点。小体积的球状孔隙的 F1 分数和准确率分别提升了 27.47、15.94 个百分点。因此，使用 MFA 可以提升 4 类孔隙特征学习能力，对于小尺寸孔隙结构具有良好的分割效果。

与原始 3D UNet 网络相比，增加 VAE 分支后的网络模型 4 类孔隙均提升了分割性能，4 类孔隙准确率分别提升了 4.96、6.00、8.24 和 7.33 个百分点；F1 分数分别提高了 3.44、4.13、14.51 和 16.53 个百分点。说明通过引入噪声和辅助损失函数，VAE 有效地提高了网络模型泛化性，从而提高网络分割性能。

与次优指标表现的改进 UNet 网络相比，改进 UNet-VAE 网络（模型 2）对于 4 类孔隙准确率分别提升了 0.52、0.50、0.77 和 0.40 个百分点；F1 分数分别提高了 0.27、0.84、2.79 和 1.25 个百分点，具有最佳的分割效果。

为了直观体现不同模块对于改进网络的影响，使用 4 类孔隙平均分割指标，在表 1 平均值中可以看出 MFA 和 VAE 都对 4 类孔隙分割效果有明显提升。

表 1 不同改进网络的多类型孔隙分割效果

类型 Categories	指标 Index	模型 1 Model 1	模型 2 Model 2	模型 3 Model 3	模型 4 Model 4
裂隙 Cracks	准确率	78.24±6.42	91.77±4.23	83.20±5.21	92.29±3.71
	精确率	72.70±9.89	82.81±12.42	74.92±11.93	83.24±11.63
	召回率	66.31±10.67	84.77±8.98	70.98±10.39	84.23±9.20
	F1 分数	68.95±8.71	83.13±8.43	72.39±9.42	83.40±9.11
生物孔 Biological pores	准确率	75.57±6.35	90.74±4.34	81.57±5.11	91.24±3.80
	精确率	76.75±9.86	89.31±7.68	79.38±10.59	88.53±8.22
	召回率	77.38±10.82	87.01±10.61	83.06±9.23	89.08±8.81
	F1 分数	76.92±9.38	87.88±8.30	81.05±9.39	88.72±8.14
不规则孔隙 Irregular pores	准确率	76.63±6.87	93.67±2.90	84.87±4.31	94.44±2.39
	精确率	45.65±11.35	71.34±13.46	63.27±12.59	75.41±11.80
	召回率	49.71±6.31	73.40±10.22	60.95±7.54	74.93±9.79
	F1 分数	47.16±8.59	72.03±11.21	61.67±9.22	74.82±9.34
球状孔隙 Spherical pores	准确率	81.04±5.33	96.98±1.25	88.37±3.11	97.38±1.10
	精确率	64.51±5.12	90.66±4.98	82.09±4.12	91.84±4.53
	召回率	61.39±6.02	90.03±3.82	76.79±5.02	91.27±3.32
	F1 分数	62.73±4.31	90.20±2.74	79.26±3.75	91.45±2.49
平均值 Mean value	准确率	77.87±6.24	93.29±3.18	84.50±4.43	93.83±2.75
	精确率	64.90±9.05	83.83±9.63	83.16±9.80	84.75±9.04
	召回率	63.70±8.45	83.97±8.40	82.78±8.04	84.87±7.78
	F1 分数	63.94±7.74	83.66±7.67	82.66±7.94	84.59±7.27

注：模型 1：原始 3D UNet 网络；模型 2：结合了 MFA 的改进 UNet 网络；模型 3：引入了 VAE 的 UNet-VAE 网络；模型 4：结合了 MFA 和 VAE 的改进 UNet-VAE 网络。下同。

Note: Model 1: the original 3D UNet network; Model 2: an improved UNet network that combines MFA; Model 3: a UNet-VAE network with the introduction of VAE; Model 4: an improved UNet-VAE network that combines MFA and VAE. Same as below.

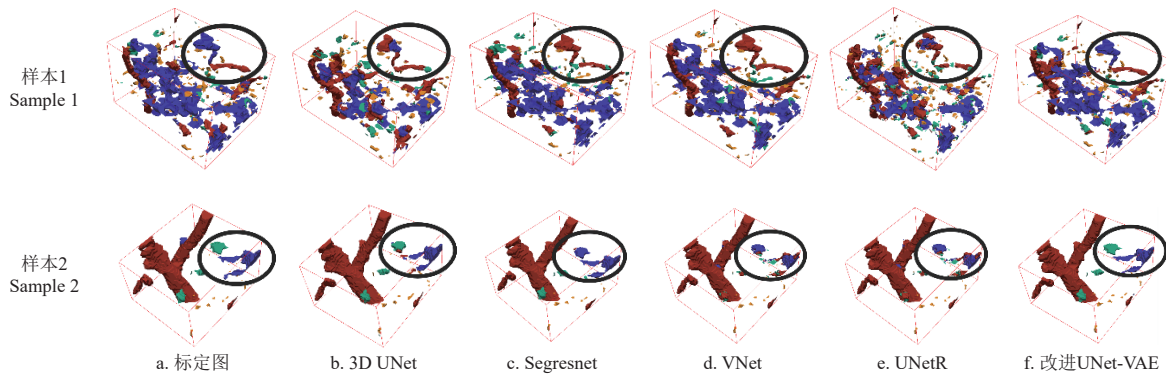
结合 MFA 的改进 UNet 网络分别在平均准确率、精确率、召回率和 F1 分数上提升了 15.42、18.93、20.27 和 19.72 个百分点。引入 VAE 的 UNet-VAE 网络分别在平均准确率、精确率、召回率和 F1 分数上提升了 6.63、18.26、19.08 和 18.72 个百分点。本文提出的同时结合 MFA 和 VAE 分支的方法对于 4 类孔隙分割具有最优性能，平均准确率、精确率、召回率和 F1 分数分别达到

93.83、84.75、84.87 和 84.59 个百分点, 说明了 2 个改进点的有效性。

2.2 不同三维分割模型对比分析

3D UNet、Segresnet、VNet 和 UNetR 作为三维分割中的经典模型, 近些年来受到广泛认可和使用。因此, 在孔隙多类型分割效果试验中, 选取以上网络与本研究提出的改进 UNet-VAE 网络模型进行分割效果对比。随机选取土壤图像样本进行结果展示, 5 种网络对于多类

别孔隙的分割结果如图 6 所示和表 2 所示。对于大体积的裂隙和生物孔, 如样本 1, 改进 UNet-VAE 网络可以准确分割类型及范围。3D UNet、Segresnet、VNet 和 UNetR 对于裂隙(蓝色)和生物孔(红色)难以分辨, 其中 UNetR 由于 Transformer 对数据集数量有着较高要求, 难以学习特征, 各类孔隙欠分割现象明显。3D UNet、Segresnet 和 VNet 将平面特征明显的裂隙分类为生物孔, 说明了卷积网络对于全局信息和大尺寸特征学习的欠缺。



注: 样本 1 圈中是代表性裂隙和生物孔, 样本 2 圈中是代表性裂隙和不规则孔隙。红色为生物孔, 蓝色为裂隙, 绿色为不规则孔隙, 黄色为球状孔隙。
Note: In sample 1, circles include representative fractures and biogenic pores, while in sample 2, circles include representative fractures and irregular pores. The red represents biological pores, the blue represents cracks, the green represents irregular pores, and the yellow represents spherical pores.

图 6 不同网络分割效果对比
Fig.6 Comparison of different network segmentation effects

表 2 不同网络的多类型分割效果

网络 Networks	类型 Categories	准确率 Accuracy	精确率 Precision	召回率 Recall	F1 分数 F1-score
3D UNet	裂隙	78.24±6.42	72.70±9.89	66.31±10.67	68.95±8.71
	生物孔	75.57±6.35	76.75±9.86	77.38±10.82	76.92±9.38
	不规则孔隙	76.63±6.87	45.65±11.35	49.71±6.31	47.16±8.59
	球状孔隙	81.04±5.33	64.51±5.12	61.39±6.02	62.73±4.31
	平均	77.87±6.24	64.90±9.05	63.70±8.45	63.94±7.7
Segresnet	裂隙	87.96±3.44	71.39±17.14	77.88±5.33	73.60±20.73
	生物孔	82.16±4.00	78.10±10.93	80.82±7.04	78.96±7.59
	不规则孔隙	83.49±3.09	38.23±8.98	42.25±5.55	39.67±6.94
	球状孔隙	91.59±1.61	82.23±4.52	65.60±3.91	72.81±2.20
	平均	86.30±3.03	67.49±10.39	66.64±5.4	66.26±9.36
VNet	裂隙	90.06±3.60	81.53±11.07	70.45±16.80	74.29±12.91
	生物孔	87.44±3.35	83.60±7.18	85.31±9.28	84.32±7.74
	不规则孔隙	89.56±2.83	59.95±8.53	59.23±4.39	59.32±5.59
	球状孔隙	94.99±1.52	83.68±5.35	88.66±1.91	85.96±2.04
	平均	90.51±2.82	79.69±8.03	75.91±8.09	75.97±7.07
UNetR	裂隙	84.56±4.28	69.23±13.62	71.80±22.99	64.11±7.30
	生物孔	79.06±2.72	74.02±12.96	81.11±16.24	81.73±11.46
	不规则孔隙	82.90±3.16	37.31±9.62	59.17±6.60	60.02±6.58
	球状孔隙	93.76±1.22	82.63±4.82	87.44±2.55	87.77±1.87
	平均	85.07±2.84	65.80±10.25	74.88±12.09	65.35±6.80
改进 UNet-VAE	裂隙	92.29±3.71	83.24±11.63	84.23±9.20	83.40±9.11
	生物孔	91.24±3.80	88.53±8.22	89.08±8.81	88.72±8.14
	不规则孔隙	94.44±2.39	75.41±11.80	74.93±9.79	74.82±9.34
	球状孔隙	97.38±1.10	91.84±4.53	91.27±3.32	91.45±2.49
	平均	93.83±2.75	84.75±9.04	84.88±7.78	84.60±7.27

对于小体积的不规则孔隙, 如样本 2, Segresnet、

VNet 和 UNetR 均将其误分为裂隙。且其余网络均出现产生欠分割现象, 导致连通的裂隙被误分为不规则孔隙及球状孔隙, 而改进 UNet-VAE 网络可以准确将断裂形成的不规则孔隙进行分类。说明了改进对于小尺寸孔隙特征学习及网络泛化能力的有效性。

由表 2 可知, 3D UNet 的分割精度最低, Segresnet 比起 3D UNet 有一定提升, 平均准确率、精确率、召回率和 F1 分数分别提升了 8.43、2.59、2.94 和 2.32 个百分点。VNet 对于 4 类孔隙特征指标均有较大提升, 平均准确率、精确率、召回率和 F1 分数分别提升了 12.64、14.79、12.21 和 12.03 个百分点。UNetR 作为典型 Transformer 网络, 与 VNet 相比平均准确率、精确率、召回率和 F1 分数分别降低了 5.44、13.89、1.03 和 10.62 个百分点。改进 UNet-VAE 在 4 类孔隙中均达到了所有方法中的最佳指标, 平均准确率、精确率、召回率和 F1 值分别达到了 93.83、84.75、84.88 和 84.60 个百分点。与次优 VNet 方法相比, 平均准确率、精确率、召回率和 F1 值分别提升了 3.32、5.06、8.97 和 8.63 个百分点, 特别是对于不规则孔隙准确率、精确率、召回率和 F1 值分别提升了 4.88、15.46、15.70 和 15.50 个百分点。综上, 改进 UNet-VAE 实现了高精度多类型孔隙三维分割, 对 4 类孔隙均有良好的特征学习能力, 与标定图最为接近, 方法泛化能力强。

3 结论

为进一步了解孔隙形态、形成原因和孔隙功能的联系, 本研究依据孔隙形态分类、结合多阈值分割及人工

矫正标注建立 4 类孔隙(裂隙、生物孔、不规则孔隙和球状孔隙)真值数据集。针对不同类型孔隙演变中产生相交的现象,引入深度学习进行土壤多类型孔隙三维分割。针对不同类型孔隙分割精度低、分割方法自动化程度低、鲁棒性差等问题,提出了一种改进 UNet-VAE 土壤多类型孔隙三维分割方法,通过融合多尺度特征并结合生成网络提高模型泛化性。通过消融试验和定性定量比较分析证明:

1) 改进 UNet-VAE 网络模型可以精确分割不同类型孔隙,针对 4 类孔隙分割准确率分别达到 92.29%、91.24%、94.44% 和 97.38%,平均准确率达到 93.83%。实现了高精度土壤 4 类孔隙分割,可为研究土壤物理结构提供先进的技术手段。

2) 基于提出的 MFA 的局部和全局注意力机制,改进 UNet-VAE 网络提高了多尺度特征信息的融合能力,并筛减冗余信息。通过 VAE 引入噪声信息和辅助损失函数,改进 UNet-VAE 网络通过学习不同类型孔隙特征,自动分割多类型孔隙,有效提高了分割方法的鲁棒性,增强了网络的泛化能力。

本文方法实现了对黑土 4 类孔隙的精准分割,而对于其他土壤不同功能孔隙的分类与分割和不同孔隙的结构演变研究,将是进一步研究的重点。

[参 考 文 献]

- [1] KRAVCHENKO A N, GUBER A K. Soil pores and their contributions to soil carbon processes[J]. *Geoderma*, 2017, 287: 31-39.
- [2] RABOT E, WIESMEIER M, SCHLÜTER S, et al. Soil structure as an indicator of soil functions: A review[J]. *Geoderma*, 2018, 314: 122-137.
- [3] VOGEL H J, EBERHARDT E, FRANKO U, et al. Quantitative evaluation of soil functions: Potential and state[J]. *Frontiers in Environmental Science*, 2019, 7: 164.
- [4] 丁天宇, 郭自春, 钱泳其, 等. 秸秆还田方式对砂姜黑土有机碳组分和孔隙结构的影响[J]. *农业工程学报*, 2023, 39(16): 71-78.
DING Tianyu, GUO Zichun, QIAN Yongqi, et al. Effects of straw return methods on the soil organic carbon fractions and pore structure characteristics of Shajiang black soil (Vertisol)[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2023, 39(16): 71-78. (in Chinese with English abstract)
- [5] PFEIFER J, KIRCHGESSNER N, COLOMBI T, et al. Rapid phenotyping of crop root systems in undisturbed field soils using X-ray computed tomography[J]. *Plant Methods*, 2015, 11: 1-8.
- [6] 彭珏, 陈家赢, 王军光, 等. 中国典型地带性土壤团聚体稳定性与孔隙特征的定量关系[J]. *农业工程学报*, 2022, 38(18): 113-121.
PENG Jue, CHEN Jiaying, WANG Junguang, et al. Linking aggregate stability to the characteristics of pore structure in different soil types along a climatic gradient in China[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2022, 38(18): 113-121. (in Chinese with English abstract)
- [7] 邱琛, 韩晓增, 陈旭, 等. CT 扫描技术研究有机物料还田深度对黑土孔隙结构影响[J]. *农业工程学报*, 2021, 37(14): 98-107.
QIU Chen, HAN Xiaozeng, CHEN Xu, et al. Effects of organic amendment depths on black soil pore structure using CT scanning technology[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2021, 37(14): 98-107. (in Chinese with English abstract)
- [8] 徐洋洋, 张兴, 左西宇, 等. 再生水灌溉对土壤表层大孔隙的影响[J]. *农业工程学报*, 2023, 39(23): 113-122.
XU Yangyang, ZHANG Xing, ZUO Xiyu, et al. Effects of reclaimed water irrigation with different water quality on surface soil macro-pores[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2023, 39(23): 113-122. (in Chinese with English abstract)
- [9] CAPOWIEZ Y, SAMMARTINO S, MICHEL E. Using X-ray tomography to quantify earthworm bioturbation non-destructively in repacked soil cores[J]. *Geoderma*, 2011, 162(1/2): 124-131.
- [10] ZHANG Z B, ZHOU H, ZHAO Q G, et al. Characteristics of cracks in two paddy soils and their impacts on preferential flow[J]. *Geoderma*, 2014, 228: 114-121.
- [11] GARBOU A, MUNKHOLM L J, HANSEN S B. Tillage effects on topsoil structural quality assessed using X-ray CT, soil cores and visual soil evaluation[J]. *Soil and Tillage Research*, 2013, 128: 104-109.
- [12] ROONEY E C, BAILEY V L, PATEL K F, et al. Soil pore network response to freeze-thaw cycles in permafrost aggregates[J]. *Geoderma*, 2022, 411: 115674.
- [13] ZHANG Z, LIU K, ZHOU H, et al. Three dimensional characteristics of biopores and non-biopores in the subsoil respond differently to land use and fertilization[J]. *Plant and Soil*, 2018, 428: 453-467.
- [14] BOTTINELLI N, ZHOU H, CAPOWIEZ Y, et al. Earthworm burrowing activity of two non-Lumbricidae earthworm species incubated in soils with contrasting organic carbon content (Vertisol vs. Ultisol)[J]. *Biology and Fertility of Soils*, 2017, 53: 951-955.
- [15] KOESTEL J, LARSBO M. Imaging and quantification of preferential solute transport in soil macropores[J]. *Water Resources Research*, 2014, 50(5): 4357-4378.
- [16] RINGROSE-VOASE A J. Measurement of soil macropore

- geometry by image analysis of sections through impregnated soil[J]. *Plant and Soil*, 1996, 183: 27-47.
- [17] 张中彬, 彭新华. 土壤裂隙及其优先流研究进展[J]. 土壤学报, 2015, 52(3): 477-488.
ZHANG Zhongbin, PENG Xinhua. A review of researches on soil cracks and their impacts on preferential flow[J]. *Journal of Soil Science*, 2015, 52(3): 477-488. (in Chinese with English abstract)
- [18] CHEN M, LI Y, JIANG X, et al. Study on soil physical structure after the bioremediation of Pb pollution using Microbial-induced carbonate precipitation methodology[J]. *Journal of Hazardous Materials*, 2021, 411: 125103.
- [19] PAGLIAI M. Pore Morphology and soil function [M]// GLINSKI J, HORABIK J, LIPIEC J. *Encyclopedia of Agrophysics. Encyclopedia of Earth Sciences Series*. Dordrecht: Springer, 2011: 640-645.
- [20] 周云艳, 徐琨, 陈建平, 等. 基于 CT 扫描与细观力学的植物侧根固土机理分析[J]. 农业工程学报, 2014, 30(1): 1-9.
ZHOU Yunyan, XU Kun, CHEN Jianping, et al. Mechanism of plant lateral root reinforcing soil based on CT scan and mesomechanics analysis[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2014, 30(1): 1-9. (in Chinese with English abstract)
- [21] WANG H, QIAN H, GAO Y. Characterization of macropore structure of remolded loess and analysis of hydraulic conductivity anisotropy using X-ray computed tomography technology[J]. *Environmental Earth Sciences*, 2021, 80: 1-15.
- [22] EBRAHIMI M K V, LEE H, WON J, et al. Estimation of soil texture by fusion of near-infrared spectroscopy and image data based on convolutional neural network[J]. *Computers and Electronics in Agriculture*, 2023, 212: 108117.
- [23] MENG C, YANG W, BAI Y, et al. Research of soil surface image occlusion removal and inpainting based on GAN used for estimation of farmland soil moisture content[J]. *Computers and Electronics in Agriculture*, 2023, 212: 108155.
- [24] 韩巧玲, 赵玥, 赵燕东, 等. 基于全卷积网络的土壤断层扫描图像中孔隙分割[J]. 农业工程学报, 2019, 35(2): 128-133.
HAN Qiao, ZHAO Yue, ZHAO Yandong, et al. Pore segmentation in soil tomography images based on fully convolutional network[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2019, 35(2): 128-133. (in Chinese with English abstract)
- [25] FU Y, ZHAO Y, ZHAO Y, et al. Semi-supervised segmentation of multi-scale soil pores based on a novel receptive field structure[J]. *Computers and Electronics in Agriculture*, 2023, 212: 108071.
- [26] 韩巧玲. 基于 CT 图像的黑土大孔隙精细分割与重构方法研究[D]. 北京: 北京林业大学, 2020.
HAN Qiaoling. *Fine Segmentation and Reconstruction of Black Soil Macropore Based on CT Image*[D]. Beijing: Beijing Forestry University, 2020. (in Chinese with English abstract)
- [27] 韩巧玲, 周希博, 宋润泽, 等. 基于序列信息的土壤 CT 图像超分辨率重建[J]. 农业工程学报, 2021, 37(17): 90-96.
HAN Qiaoling, ZHOU Xibo, SONG Runze, et al. Super-resolution reconstruction of soil CT images using sequence information[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2021, 37(17): 90-96. (in Chinese with English abstract)
- [28] 高朝侠, 徐学选, 赵娇娜, 等. 土壤大孔隙流研究现状与发展趋势[J]. 生态学报, 2014, 34(11): 2801-2811.
GAO Chaoxia, XU Xuexuan, ZHAO Jiaona, et al. Review on macropore flow in soil[J]. *Acta Ecologica Sinica*, 2014, 34(11): 2801-2811. (in Chinese with English abstract)
- [29] MYRONENKO A. 3D MRI brain tumor segmentation using autoencoder regularization[C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4. Berlin: Springer International Publishing, 2019: 311-320.
- [30] 韩巧玲, 柏浩, 赵玥, 等. 采用染色示踪技术的土壤优先流自动分割与量化系统[J]. 农业工程学报, 2021, 37(6): 127-134.
HAN Qiaoling, BAI Hao, ZHAO Yue, et al. A soil priority flow automatic segmentation and quantification system using dye tracing technology[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2021, 37(6): 127-134. (in Chinese with English abstract)

Three-dimensional segmentation method of soil multi-category pores based on improved UNet-VAE network

HAN Qiaoling^{1,3,4,5}, SONG Meihui^{1,3,5}, XI Benye², ZHAO Yue^{1,3,4,5*}, ZHAO Yandong^{1,3,5}

(1. School of Technology, Beijing Forestry University, Beijing 100083, China; 2. School of Forestry, Beijing Forestry University, Beijing 100083, China; 3. Beijing Laboratory of Urban and Rural Ecological Environment, Beijing Municipal Education Commission, Beijing 100083, China; 4. Key Lab of State Forestry Administration for Forestry Equipment and Automation, Beijing 100083, China; 5. Research Center for Intelligent Forestry, Beijing Forestry University, Beijing 100083, China)

Abstract: Soil pores can play a significant role in the crucial process related to soil ecological functions. However, it is still challenging on the relationship between pore structure and functional evolution, due to the lack of non-destructive and non-intrusive systems for the spatial structure of multiple types of pores. Among them, accurate segmentation of pore types and ranges was fundamental to the research. In this study, an improved UNet-VAE network was proposed to segment the multiple pores in soil for the first time. Taking the typical black soil as the research object, a Simplified Convolutional Network (SCN) was used to segment the three-dimensional data of soil pores. According to the segmented pore dataset, a combination of automatic segmentation and manual correction was used to obtain four types of soil pore structures as ground truth. A multi-scale fusion attention module was proposed to filter out the redundant information generated by convolutional learning using a 3D UNet network. Local attention was used to learn the spatial features of small-scale pores (irregular and spherical pores). Global attention was used to extract the feature information of large-scale pores (cracks and biological pores), in order to fuse the multi-scale features for high segmentation accuracy in the different categories of pores. Meanwhile, the commonly used segmentation networks (such as 3D UNet, Segresnet, VNet, and UNetR network) were used to compare the segmentation of the multiple pores. The experimental results showed that the improved UNet-VAE network accurately determined the range and the category of pores. Specifically, the UNetR network was difficult to learn the features, due to the high requirement of a Transformer for the number of datasets in the large-scale cracks and biological pores. Convolutional networks (such as 3D UNet, Segresnet, and VNet) failed to learn the global and large-scale features, where the cracks with obvious planar features were classified as biological pores. Furthermore, the Segresnet, VNet, and UNetR network misclassified the small-scale irregular pores as cracks. By contrast, the improved UNet VAE network achieved the best performance in the four categories of pores, with the average accuracy, precision, recall, and F1 values reaching 93.83%, 84.75%, 84.88%, and 84.60%, respectively. Compared with the suboptimal VNet, the average accuracy, precision, recall, and F1 value increased by 3.32, 5.06, 8.97, and 8.63 percentage points, respectively. Especially for irregular pores, the accuracy, recall, and F1 values increased by 4.88, 15.46, 15.70, and 15.50 percentage points, respectively. In summary, the improved UNet-VAE network was achieved in the high-precision and three-dimensional segmentation of multiple categories of pores, indicating better feature learning for all four categories of pores. The three-dimensional segmentation was achieved in the better classification of intersecting pores, high segmentation accuracy of single pores, and high automation level. Technical support was offered to explore the relationship between pore structure and ecological function evolution. This finding can also provide a data basis for the precise quantitative characterization of soil pore structure, in order to reveal the role of soil pore evolution in ecosystems.

Keywords: soils; pores; convolutional neural network; image segmentation; 3D UNet; attention module