

基于 μ CT 和深度学习的煤裂隙智能提取方法

胡咤咤¹, 张寻¹, 金毅^{2,*}, 巩林贤², 黄文辉³, 任建吉⁴, Norbert Klitzsch⁵

(1. 河南理工大学 能源科学与工程学院, 深井岩层控制与瓦斯抽采技术应急管理部科技研发平台, 河南 焦作 454003; 2. 河南理工大学 资源与环境学院, 河南 焦作 454003; 3. 中国地质大学(北京) 能源学院, 北京 100083; 4. 河南理工大学 计算机科学与技术学院, 河南 焦作 454003; 5. 亚琛工业大学 应用地球物理与地热能研究所, 德国 亚琛 52074)

摘要: 煤储层裂隙的精细描述对于煤层气资源的勘探开发具有重要意义, 裂隙的尺寸、走向、分布密度等直接影响煤层的渗透性, 准确识别和提取煤层中的裂隙信息是揭示体积压裂过程中裂缝网络形成与扩展机理的关键。传统的裂隙提取方法往往依赖人工标注和基于一定图像处理技术的特征提取, 这些方法在精度和效率上存在明显不足。提出一种基于 Trans-UNet 网络和 μ CT 图像的煤裂隙提取方法, Trans-UNet 结合了 Transformer 模块和卷积神经网络(CNN)的优点, 不仅具备全局特征提取能力, 还能够捕捉图像中的局部细节特征, 大幅提高了分割精度和网络的鲁棒性。首先对煤样 μ CT 图像进行预处理, 包括使用差值法提高图像质量、使用数据增强技术扩大样本数量等。随后, 利用 Trans-UNet 网络对处理后的图像进行分割, 提取裂隙特征, 并比较不同神经网络模型的分割结果。结果表明, 提出的方法在数据集上表现出优越性能, Trans-UNet 模型在煤裂隙提取上的准确性(Accuracy)、精确度(Precision)、F1 分数(F1-Score)和交并比(IoU)分别达到 91.3%、89.5%、89.8% 和 84.0%, 相较于 U-Net、U-Net++ 等其他多种智能模型有显著提升。结合 μ CT 图像的细粒度特征, 将 Trans-UNet 网络应用于煤裂隙提取任务, 是一种高效且准确的解决方案, 为煤层气勘探开发领域的相关图像处理任务提供了新的思路。

关 键 词: Trans-UNet; μ CT 图像; 煤裂隙; 图像分割; 深度学习

中图分类号: P31

文献标志码: A

Intelligent coal fracture extraction method using μ CT and deep learning

HU Zhazha¹, ZHANG Xun¹, JIN Yi^{2,*}, GONG Linxian², HUANG Wenhui³, REN Jianji⁴, Norbert Klitzsch⁵

(1. *Science and Technology R&D Platform of Emergency Management Ministry for Deep Well Ground Control and Gas Extraction Technology, School of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo 454003, China*; 2. *School of Resources and Environment, Henan Polytechnic University, Jiaozuo 454003, China*; 3. *School of Energy, China University of Geosciences (Beijing), Beijing 100083, China*; 4. *School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China*; 5. *Institute of Applied Geophysics and Geothermal Energy, RWTH Aachen University, Aachen 52074, Germany*)

Abstract: The detailed description of coal fractures is of great significance for the exploration and development of coalbed methane resources. The size, orientation, and distribution density of fractures directly affect the permeability of the coal seam. Accurate identification and extraction of fracture information in coal seams are crucial for revealing the mechanisms of fracture network formation and propagation during hydraulic fracturing. Traditional fracture extraction methods often rely on manual labeling and feature extraction based on certain image processing techniques, which have significant limitations in terms of accuracy and efficiency. This paper proposes a coal fracture extraction method based on Trans-UNet and μ CT images. Trans-UNet combines the advantages of Transformer modules and Convolutional Neural Network (CNN), possessing both global feature extraction capability and the ability to capture local details in images, significantly improving segmentation accuracy and network robustness. First, the μ CT images of coal samples are pre-processed, including using interpolation methods to improve image quality and data augmentation techniques to increase the number of samples. Subsequently, the processed images are segmented using the Trans-UNet network to extract fracture features, and the segmentation results of different neural network models are compared. The results show

收稿日期: 2024-09-06; 接收日期: 2025-02-10

基金项目: 国家自然科学基金青年项目(42402184); 河南省科技攻关项目(232102320200); 河南省高等学校重点科研项目(23A170006)

第一作者: 胡咤咤, 1990 年生, 男, 湖北黄冈人, 博士, 副教授。E-mail: zhazha.hu@hpu.edu.cn

通信作者: 金毅, 1979 年生, 男, 湖北鄂州人, 博士, 教授, 博士生导师。E-mail: jinyi2005@hpu.edu.cn

© Editorial Office of Coal Geology & Exploration. OA under CC BY-NC-ND

that the proposed method outperforms other models on the dataset. The Trans-UNet model achieves an Accuracy of 91.3%, Precision of 89.5%, F1-score of 89.8%, and Intersection over Union (IoU) of 84.0%, significantly outperforming other intelligent models such as U-Net and U-Net++. Combining the fine-grained features of μ CT images, the application of the Trans-UNet network to coal fracture extraction tasks is an efficient and accurate solution, providing new insights for image processing tasks in related fields.

Keywords: Trans-UNet; μ CT images; coal fractures; image segmentation; deep learning

煤层气作为一种重要的清洁能源，在全球范围内的能源开发中扮演着越来越重要的角色^{[1][4]}。裂隙作为煤层的重要储层特征，是煤层气运移的主要通道^{[5][6]}。在煤层气开发过程中，准确识别和提取煤层中裂隙信息是揭示体积压裂过程中裂缝网络形成与扩展机理的关键。煤中裂隙的分布特征对煤矿开采安全性也会产生影响^[9]，它的存在可能导致煤层稳定性降低，增加发生地质灾害的风险。

传统的煤裂隙特征研究主要通过井下煤壁或钻孔岩心观测其宏观特征，或利用光学显微镜、扫描电镜查明其微观特征^[10]。近年来，一些图像处理技术开始应用于裂隙的定量表征，例如基于图像灰度值的阈值分割方法^[11]、基于区域生长法的分割方法^[12]，还有混合迭代局部阈值的分割方法(TILT)等^[13]。虽然这些方法在一定程度上能解决一些实际问题，但仍存在精度低、效率差等不足。上述人工观察方法需要丰富的经验，耗时量大，并且肉眼还难以识别一些微小裂隙。基于灰度的阈值分割法等精度不高，无法准确区分与裂隙灰度值相差不大的显微组分和矿物。因此，目前迫切需要一种更加高效且精确的裂隙识别方法。

微计算机断层扫描(micro computed tomography, μ CT)是一种先进的无损三维成像技术^[14]错误!未找到引用源。，其能够对岩石样本的内部结构和成分进行高分辨率成像错误!未找到引用源。。 μ CT 获取的煤样内部高分辨率图像，可以详细展现裂隙的尺寸、方位、连通性及其在煤层中的空间分布，该技术已被证明是定量研究煤裂隙网络的有效工具错误!未找到引用源。错误!未找到引用源。。

此外，随着计算机视觉和深度学习技术的发展，卷积神经网络(convolutional neural network, CNN)在图像分割领域的应用也有着显著的进步错误!未找到引用源。，基于深度学习的自动化分割技术在医学图像处理等领域表现出了优越的性能^[23]。近年来，针对煤裂隙的智能提取研究也取得了显著进展。例如，王登科等^[24]提出的MCSN模型基于U-Net网络，通过引入迁移学习和深度可分离空洞卷积模块(DCAC)，显著提升了煤裂隙的识别精度和鲁棒性。郝天轩等^[25]使用基于U-Net网络的改进模型，通过直方图均衡化、高斯双边滤波和拉普拉斯算子等预处理方法，有效提高了

煤裂隙的提取能力。郑江韬等^[26]采用Waifu2x卷积神经网络模型和U-Net网络，显著提高了煤岩CT图像的分辨率和裂隙信息提取精度。此外，J. YU等^[27]提出的DeepLab V3+模型通过优化损失函数和网络结构，智能地提取煤样CT图像中的裂隙，有效避免了传统阈值分割方法的局限性。F. LU等^[28]提出的FPMF-U-Net模型通过并行多尺度特征融合和平均池化操作，进一步提升了复杂煤裂隙的分割效果。S. Karimpouli等^[29]通过使用卷积神经网络方法对均质和非均质的煤样裂隙进行分割，发现3D CNN在提取3D空间特征方面比2D CNN更为稳健和准确，特别是在复杂的非均质样本中，3D CNN实现了96.7%的准确率，显著优于传统分割方法。这些研究成果为煤裂隙的自动提取提供了重要的技术支撑^{[24][29]}。Transformer模型^[30]的引入为处理复杂图像提供了全新的思路，其通过自注意力机制，能够有效捕获长距离依赖关系和全局信息，提升图像的理解能力。CNN通过多层卷积操作和池化，可以自动学习图像中的局部特征，而Transformer则能够捕捉图像中的全局关系，两者结合有望更准确地提取复杂的煤裂隙结构。

笔者结合 μ CT技术，融合CNN、Transformer架构的Trans-UNet网络，提出一种高效、精确的煤裂隙提取方法。通过结合 μ CT图像的丰富结构信息和Trans-UNet的强大特征提取能力，旨在提升煤裂隙特征提取的精度和效率，为煤层气资源的勘探与开发提供可靠的技术支持。

1 μ CT图像数据集

1.1 样品的数据采集

研究采用来自澳大利亚Bowen盆地上二叠统中挥发分烟煤 μ CT图像的公开数据集(<https://www.digitalrockspotal.org/projects/86>)，图像分辨率为1505像素×1497像素，原柱塞煤样直径25 mm、长35 mm。该数据集包含同一煤样在干燥和湿润状态下的各80张原始灰度图像，图像数据集的获取过程可参考L. H. Ramandi等(2016)错误!未找到引用源。。首先使用 μ CT将烘干的煤样进行扫描得到干燥样品的图像，然后使用碘化钠和氯化钾的混合溶液(物质的量比值3:2)对

干燥样品进行饱和处理，采用相同分辨率进行二次扫描后获得湿润样品的图像。如图 1 所示，该煤样发育一定数量的裂隙。

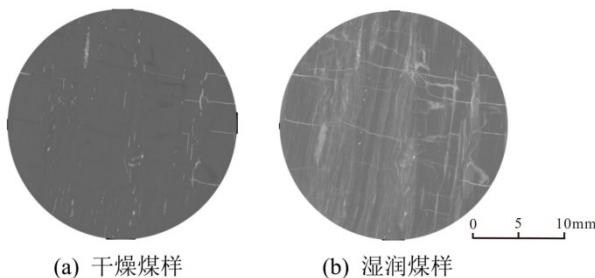


图 1 干燥/湿润状态下煤样的原始灰度图像
Fig.1 Raw grayscale images of the dry or wet coal sample

1.2 图像预处理

从图 1 可以看出，干燥和湿润样品的原始灰度图像的对比度和清晰度都不高，因此，在进行图像分割前需要对图像进行预处理。通过计算干样与湿样 μ CT 图像的灰度差值(亮度差值)来突出显示图像中的某些组分，以增强图像中不同组间的对比度。公式如下：

$$D(x,y) = I_1(x,y) - I_2(x,y) \quad (1)$$

如果两张图像中的某些区域在亮度或颜色上有所不同，这种差异在差值图像中会变得更加明显，以突出显示感兴趣的区域，如煤裂隙的边缘部分。首先使用 Matlab 软件将湿润样品图像与干燥样品图像进行配准^{错误!未找到引用源。}，为尽量减少成像条件的影响，以干燥样品图像为基准对湿润样品图像进行校正，一些不受对比剂影响的组分(如矿物)在两组图像中可获得相似的灰度值。然后将干燥样品的图像与湿润样品的图像相减，生成差值图像(图 2)。与图 1 相比，差值图像更清晰地显现出煤样的裂隙网络。

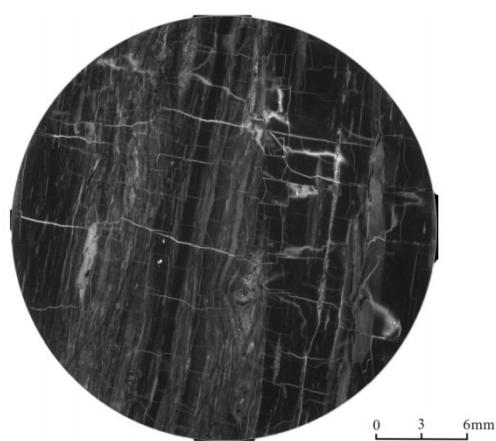


图 2 图像预处理后的差值图像(白色是孔隙和裂隙，灰色是

无机矿物，黑色是有机显微组分)
Fig.2 Difference image formed after image preprocessing (white indicates pores and cleats, gray represents inorganic minerals, black denotes organic macerals)

1.3 数据增强

数据增强的目的是通过生成更多的训练样本以提高模型的泛化能力，防止过拟合，并改善模型在测试数据上的性能^{[33][35]}。在煤裂隙提取任务中，由于数据集的图像数量有限，数据增强可以帮助扩充数据集，从而提高模型对裂隙特征的识别能力。因此，为了提高基于 Trans-UNet 的煤裂隙分割提取模型在 μ CT 图像上的表现，采用 python 编程语言中的 Albumentations 库对数据集进行系统的增强处理，将原本 80 张的差值图像数据集扩充到了 1 000 张。

数据增强方法具体步骤如下。

(1) 随机角度旋转：将 80 张原始 μ CT 图像中的每一张图像经过随机角度旋转($-25^\circ\sim25^\circ$)，得到不同视角下的旋转图像。这个步骤循环进行 10 次最终得到 800 张具有各种视角的数据，用以模拟不同采集角度下的图像格式，有助于提高模型对不同视角下煤裂隙的识别能力。

(2) 随机水平和垂直翻转：将 800 张旋转好的每一张图像以 50% 的概率进行翻转(水平翻转、垂直翻转，或两者都翻转)，这一步可以增加样本的多样性和对称性，以模拟不同方向的样本，提高模型对方向变化的鲁棒性(鲁棒性指的是一个系统、算法或模型在面对不确定性、噪声干扰或其他异常情况下的稳定性和可靠性)。

(3) 随机缩放：将每一张翻转过后的图像进行随机缩放(0.8~1.2 倍)，得到大小(分辨率)不同的 800 张的图像数据集，以模拟不同距离下的视图变化，增强模型对尺度变化的鲁棒性。

(4) 随机裁剪：从原始的 80 张 μ CT 图像中随机裁剪固定大小的区域(例如 256 像素×256 像素)以生成局部视角样本，得到 200 张额外的图像，增加训练样本的多样性，保留图像的局部细节，提高模型对局部特征的敏感度，最终得到共 1 000 张不同分辨率的数据集。

(5) 随机亮度调整：对每一张图像进行随机的亮度调整，范围为 0.7~1.3 倍，用以模拟不同的图像亮度，增强模型对亮度变化的鲁棒性。

(6) 随机对比度调整：调整每一张图像的对比度(0.7~1.3 倍)，用以模拟不同图像质量，提高模型在不同对比度下的表现。

(7) 随机添加高斯噪声: 向图像中添加高斯噪声(均值为 0, 标准差为 0.01), 用以模拟实际成像过程中可能出现的噪声, 提高模型的噪声处理能力。

经过上述一系列数据增强步骤后得到的部分图像如图 3 所示。扩充后的数据集规模显著提升, 每张图像的分辨率、角度互有区别, 图像之间的亮度与对比度也不尽相同, 提升了数据的多样性, 使得模型能够学习到更加丰富的裂隙特征, 提高对不同裂隙形态的识别能力。

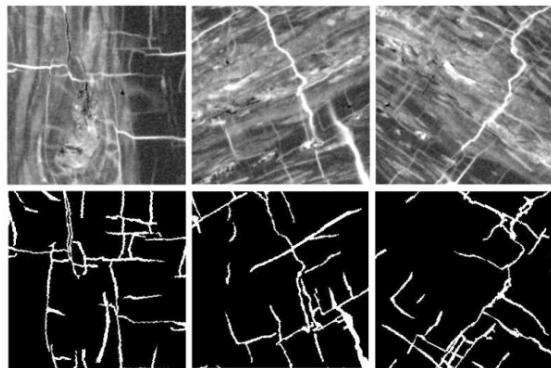


图 3 数据增强后的部分图像(煤样 CT 数据集+标签图像)
Fig.3 Selected augmented images (CT image dataset and labeled images of the coal sample)

1.4 图像重采样

为了便于图像数据输入 Trans-UNet 模型, 要将所有图像的分辨率进行统一。利用图像重采样技术将不同分辨率的图像统一成相同的空间尺寸。

上述经过数据增强后的 1 000 张图像数据集中的每张图片都拥有不同的分辨率, 故使用双线性插值法(Bilinear Interpolation)进行图像重采样, 将所有图像重采样至 512 像素×512 像素。这种方法通过考虑原始图像中相邻像素的值来计算新图像中像素的值, 从而在调整图像尺寸时尽量保留原始图像的细节和特征。具体步骤如下。

- (1) 读取经过数据增强后的图像。
- (2) 使用双线性插值法计算每个目标像素值。用原始图像中相邻 4 个像素点的加权平均值来确定新图像中对应像素的值, 该方法保证了图像的平滑性和连贯性。
- (3) 将计算出的像素值应用到目标尺寸的图像中, 获得符合要求的统一尺寸图像。

此外, 使用边缘填充(Padding)技术处理原始图像边界。如果图像的尺寸小于 512 像素×512 像素, 就在图像周围添加填充区域, 以确保最终输出图像的尺寸一致。填充区域的像素值设置为 0(即黑色), 以最小化

对图像内容的影响。

图像重采样后进行质量检查, 确保图像的细节和特征在重采样过程中得到有效保留, 主要检查图像的清晰度、边缘保留情况以及重采样前后的尺寸一致性。最终将完成的图像统一存储为 JPEG 文件格式, 在保证图像质量的同时确保文件的读取速度, 以便后续数据加载和模型训练。

2 煤裂隙分割提取模型

2.1 Trans-UNet 网络架构

传统的 U-Net 模型在医学图像分割等领域表现出色, 但其主要通过卷积操作来提取局部特征, 对于全局信息的捕捉能力有限。Transformer 模型则擅长处理全局信息, 但在图像分割任务中, 直接使用 Transformer 会导致空间细节的丢失。Trans-UNet 结合二者优势, 实现既能够捕捉全局信息, 又能够保留空间细节的高效图像分割。本文采用 Trans-UNet 作为煤裂隙分割的主要模型, 适用于处理复杂图像的分割任务, 特别是当图像中包含复杂的结构和纹理。

Trans-UNet 的网络架构如图 4 所示。Trans-UNet 先将输入图像分割成小块(patches), 并通过全连接层将小块嵌入到高维特征空间中。随后, 使用多个 Transformer 层处理这些特征, 每层包含多头自注意力机制(MHSA)和前馈神经网络(FFN), 并通过层归一化和残差连接提高模型的稳定性和训练效率。然后, U-Net 解码器从 Transformer 编码器的中间层提取特征并通过跳跃连接传递到解码器相应层, 保留空间细节。解码器通过上采样操作逐步恢复图像分辨率, 并在每个步骤后应用卷积操作增强空间信息。将 Transformer 编码器的特征与解码器特征融合, 逐步恢复图像分辨率, 生成与输入图像尺寸相同的分割结果。最后, 通过 1×1 卷积层和 softmax 激活函数生成精确的分割结果, 实现高效且准确的图像分割。

其中, Transformer 模块捕获长距离依赖的核心是多头自注意力机制。在传统的卷积神经网络中, 感受野受限于卷积核的大小, 难以有效处理全局信息。而自注意力机制通过计算每个特征与其他所有特征的相关性, 能够直接建模长距离依赖关系, 使得 Transformer 能够同时关注图像中的多个全局特征, 从而全面捕捉图像的全局信息。

理论上 Trans-UNet 适合复杂的煤裂隙分割任务。首先煤样裂隙图像中往往存在类别不平衡问题(背景区域占主导, 而裂隙区域较小), 传统方法在这种场景

下表现不佳。Trans-UNet 通过将 Transformer 与 U-Net 相结合，实现了局部特征和全局特征的有效融合，实现全面捕捉图像的复杂结构，从而提高分割精度。其次煤裂隙图像具有复杂的结构和纹理特征，这些特征通常具有长距离依赖关系(例如，裂隙的走向和分布可

能跨越较大范围甚至于全图)。传统 CNN 方法由于感受野的限制，难以有效捕捉这些长距离依赖，而 Transformer 的注意力机制能够直接建模全局信息，适合处理这类复杂任务。

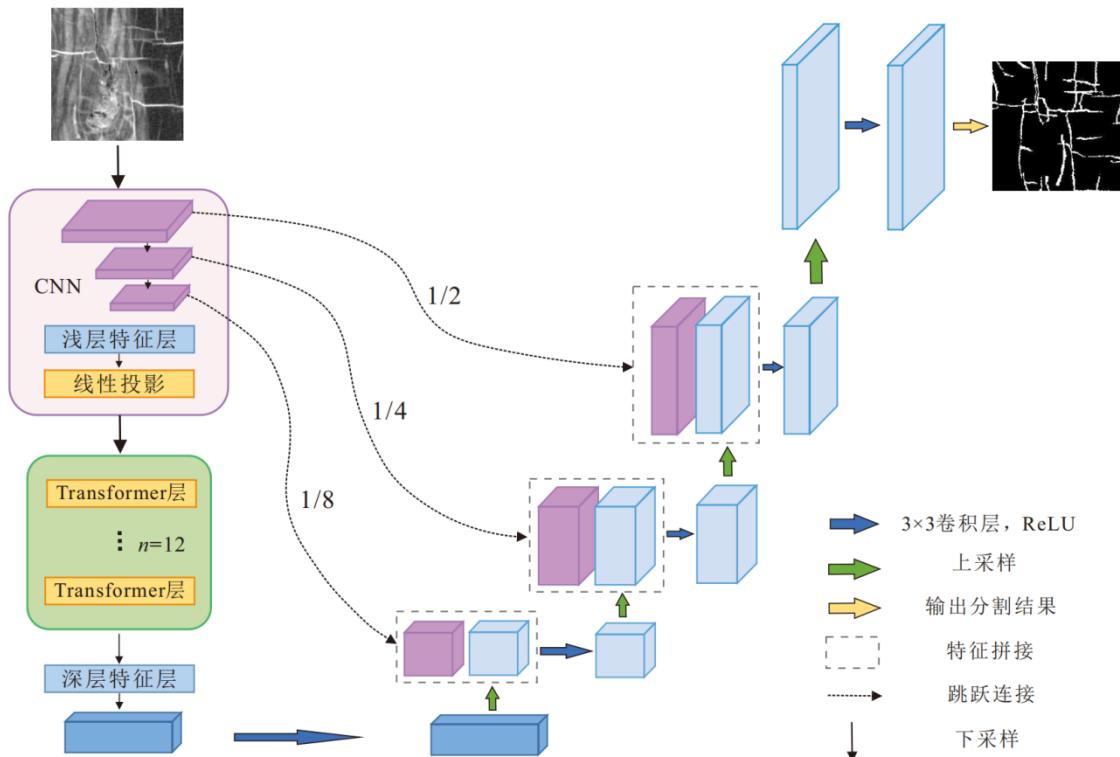


图 4 Trans-UNet 网络架构
Fig.4 Trans-UNet network architecture

2.2 多头自注意力机制

多头自注意力机制(Multi-Head Self-Attention, MHSA)是 Transformer 模型中的核心组件之一，用于捕捉输入序列中的长距离依赖关系。在 Trans-UNet 中，MHSA 被用于捕捉图像中的全局信息，具体实现步骤如图 5 所示。

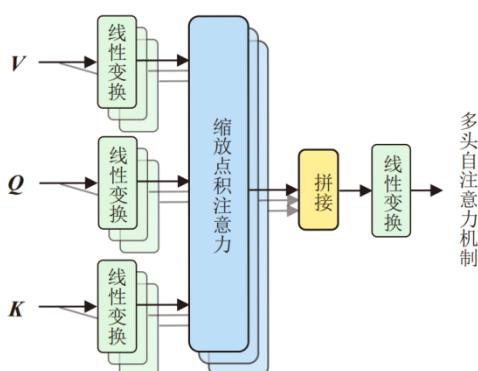


图 5 多头自注意力机制
Fig.5 Multi-Head Self-Attention Mechanism

将由图像小块构成的特征矩阵输入多头自注意力机制模块中。每个小块经过嵌入后被映射到一个高维特征空间，形成一个特征向量，这些特征向量依次构成输入序列。对于每个注意力头，输入序列经过 3 个线性变换，生成查询(Q)、键(K)和值(V)矩阵，通过计算查询矩阵和键矩阵的点积得到注意力分数。然后再通过一个缩放因子进行缩放，使内积不会过大。随后，使用 softmax 函数将这些分数转换为概率分布，表示每个位置的重要程度。通过注意力分数对矩阵值进行加权求和，得到每个注意力头的输出特征矩阵，再进行拼接，并通过一个线性变换生成最终的输出特征矩阵。拼接和线性变换的操作使得模型能够同时关注不同的位置和特征。

多头自注意力机制允许模型在不同子空间中捕捉不同的依赖关系，增强了模型的表示能力。通过并行计算多个注意力头，模型能够同时关注不同的位置和特征，提取更丰富的全局信息。

2.3 网络训练与优化

Trans-UNet 模型训练通常遵循以下步骤：

(1) 预处理：对输入数据进行归一化、增强和其他预处理操作，提高模型的泛化能力。将处理后的数据集划分为训练集和验证集，以适应分割任务。

将包含 1 000 张预处理 μ CT 图像数据集(分辨率 512×512 像素)按 8 : 2 划分为训练集和验证集。其中，训练集包含 800 张随机挑选的预处理图像，用于模型训练；验证集包含验证集一(开发集)和验证集二(测试集)。验证集一中含有 100 张随机挑选的图像，用于调整超参数和监控模型性能；验证集二包含剩余的 100 张图像，用于最终性能评估。这种划分避免了测试集的“泄露”，确保模型性能评估的可靠性，真实反映模型在之前未见数据上的表现。

(2) 构建 Trans-UNet 网络架构：编码器部分采用卷积神经网络提取局部特征，生成嵌入特征序列。将嵌入序列输入 Transformer 模块，通过自注意力机制捕获全局信息，增强模型对长距离依赖的建模能力。解码器部分利用跳跃连接结合多尺度特征信息，从而实现对目标区域的精准分割。

(3) 设置损失函数：在煤裂隙特征提取任务中，采用 Dice 损失函数和交叉熵损失函数的加权组合^{错误!未找到引用源。}，可以更好地处理类别不平衡问题，以便更准确地度量模型预测结果与真实分割结果之间的差异，提升分割性能，具体形式为：

$$\text{Loss} = \alpha \cdot \text{CrossEntropyLoss} + \beta \cdot \text{DiceLoss} \quad (2)$$

交叉熵损失函数(CrossEntropyLoss)和 Dice 损失函数(DiceLoss)的数学表达式分别为(交叉熵损失函数公式中的 log 明确是以自然常数 e 为底，相当于自然对数 ln)：

$$\text{CrossEntropyLoss} = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i)) \quad (3)$$

$$\text{DiceLoss} = 1 - \frac{2 \times \sum_{i=1}^N y_i p_i + \epsilon}{\sum_{i=1}^N y_i + \sum_{i=1}^N p_i + \epsilon} \quad (4)$$

(4) 输入图像进行训练：训练时， μ CT 图像首先被输入到 Trans-UNet 模型的编码器中。编码器通过多个卷积层逐步提取图像的局部特征，生成多尺度的特征图。这些特征图随后被展平并转换为序列，输入到 Transformer 模块中。Transformer 模块利用自注意力机制，捕获特征图中的全局信息，增强对长距离依赖的建模能力。处理后的特征序列再通过解码器，利用跳

跃连接将多尺度特征信息重新组合，逐步恢复空间细节，最后生成分割掩码。训练过程中，使用 Adam 优化器和随机梯度下降(SGD)算法进行训练迭代，并设置以下超参数：

批量大小(指在一次训练迭代中输入到模型中的样本数量)：在实验设备(内存 256 GB、显存 16 GB，下文将详细描述)的限制下选择 8，平衡内存占用和梯度更新效率。

学习率(指在每次迭代中模型参数更新的步长)：初始学习率设为 0.001，经多次实验验证，此值既能避免局部最小值停滞，又能加快训练速度。

(5) 评估与调整：在每个训练周期结束后，在验证集上评估模型性能。主要监控指标包括损失值、分割精度、Dice 系数等，以选择最佳模型并调整超参数。通过引入早期停止策略^{错误!未找到引用源。}，避免过拟合。

(6) 模型保存和测试：在训练完成后，保存性能最佳的模型权重及配置文件，用于后续应用和优化。在测试集上评估最终模型性能。

此外，为提高 Trans-UNet 模型在煤裂隙提取任务中的性能，训练过程中还加入了以下优化策略：

(1) Dropout 应用：在 Trans-UNet 的 Transformer 模块中，采用 Dropout 随机丢弃神经元，丢弃比例为 0.5。在训练时随机丢弃神经元输出(设为 0)，而在测试时根据丢弃率缩放神经元输出，从而减少对特定神经元的依赖，有效降低过拟合现象，提高模型的泛化能力。

(2) L2 正则化：在上述步骤(3)损失函数中加入 L2 权重正则化项，权重强度(λ)设置为 0.000 1。通过对大权重施加惩罚，限制模型复杂度，避免过拟合，并平滑损失函数的变化，提升训练稳定性。

(3) 学习率调度：采用学习率下降计划(LR-Scheduling)，每 10 个 epoch 将学习率减小为原来的 0.1。此策略结合反向传播算法^{错误!未找到引用源。}，能够提升训练效率，加速收敛，同时避免震荡和过拟合。

3 实验与结果分析

3.1 实验设置

本实验在以下计算环境中进行：

(1) 单个中央处理器(CPU)：Intel Xeon Silver 4214(12 核 24 线程)，提供强大的计算能力，支持大规模数据处理和模型训练。

(2) 内存(RAM)：256 GB，能够满足大规模数据处理和训练过程中高内存消耗的需求。

(3) 显卡(GPU): NVIDIA RTX A4000 显卡(16 GB), 支持高效的深度学习模型训练和推理。

(4) 操作系统: Windows 10, 确保系统的兼容性和稳定性。

(5) 编程语言: Python (3.8 版本), 用于实现数据处理、模型构建和训练。

(6) 开发环境(编译器): PyCharm (2021.1 版本), 用于代码编写、调试和项目管理等集成开发。

(7) 深度学习框架: PyTorch (1.11 版本), 提供灵活的深度学习模型构建和训练功能。

3.2 评价指标

为了全面评估基于 Trans-UNet 的煤裂隙网络提取模型的性能, 选择以下主要评价指标: 准确性(Accuracy)、精确度(Precision)、召回率(Recall)、F1 分数(F1-Score)、交并比(IoU)及 Dice 系数(Dice Coefficient)^[39]。准确性是衡量模型在所有预测中正确预测的比例; 精确度衡量模型预测为正类的样本中真正为正类的比例; 召回率衡量实际正类样本中被正确预测为正类的比例; F1 分数是精确度和召回率的调和平均数, 综合考虑了精确度和召回率的表现; 交并比用于评价分割结果的精确度, 计算预测区域与真实区域的交集与并集的比率; Dice 系数类似于 F1 分数, 但专门用于二分类任务的评价, 主要衡量预测结果与真实标签的重叠情况。各评价指标数值越大, 说明图像分割结果越好, 越接近真实的裂隙网络。公式如下:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (5)$$

$$P = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \times 100\% \quad (8)$$

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} \quad (9)$$

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (10)$$

3.3 Trans-UNet 训练过程分析

分析训练过程可以帮助评估模型性能, 实时监控损失值和交并比, 检查模型是否有效学习。若训练效果不佳, 可以及时调整超参数(学习率、批量大小等), 以提高训练效果。

设置以下超参数: 学习率 0.001, 批量大小 8, 训练迭代轮次 100, 学习率每 10 轮衰减为初始值的 0.1 倍。分别设置迭代次数为 80、100、150 与 200, 发现迭代次数为 100 时训练效果最佳, 如图 6 所示。

优化器负责在训练过程中根据损失函数的梯度更新模型参数。本文选择 Adam 优化器, 其具备自适应学习率调整的能力, 有助于加快训练的收敛速度。损失函数方面, 采用交叉熵损失函数和 Dice 损失函数的加权组合, 其中交叉熵损失用于优化分类准确性, 而 Dice 损失专注于解决类别不平衡问题, 进一步提升分割结果的细节表现。

在训练过程中, 首先进行前向传播以计算预测结果, 随后基于损失函数计算损失值, 并通过反向传播更新模型参数。训练过程中实时监控训练损失并验证损失的变化, 确保模型能够稳定收敛。

训练过程如图 6 所示。从图中可以看出, Trans-UNet 在训练过程中表现出良好的收敛性, 损失值逐渐减小, 同时在验证集上的性能指标(交并比)稳定提升。这表明 Trans-UNet 能够有效捕获 μCT 图像中煤裂隙的局部和全局特征, 提升分割性能。

为防止过拟合, 引入早停策略, 即在验证损失不再改善时提前终止训练, 避免模型在验证集上的性能下降, 同时提高模型的泛化能力。

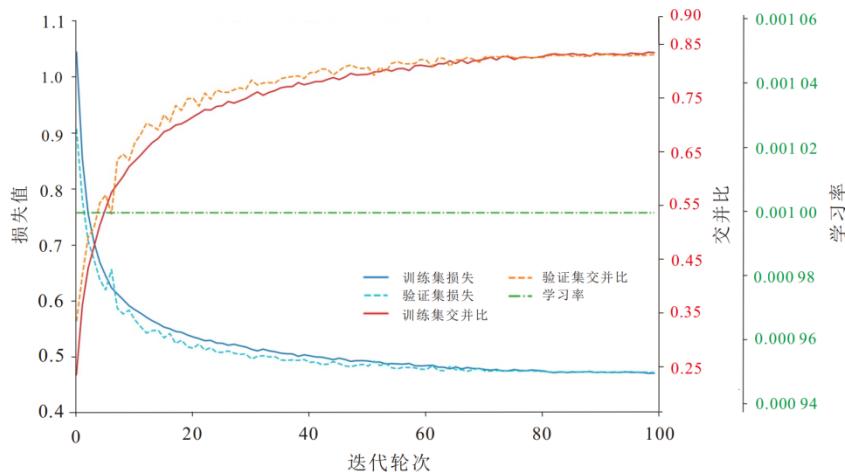


图 6 Trans-UNet 模型的训练过程
Fig.6 Training process by using the Trans-UNet model

3.4 Trans-UNet 模型与多种无监督机器学习算法

对比

为了验证 Trans-UNet 模型在煤裂隙提取中的性能, 使用同一数据集, 将其与 3 种常见的无监督机器

学习算法 (K-means 聚类、高斯混合模型和基于密度的聚类算法) 进行对比。图 7 中展示了 5 组煤 μ CT 图像的分割结果, 包括人工标注的裂隙标签(图 7b)和 4 种方法的分割结果(图 7c-图 7f)。

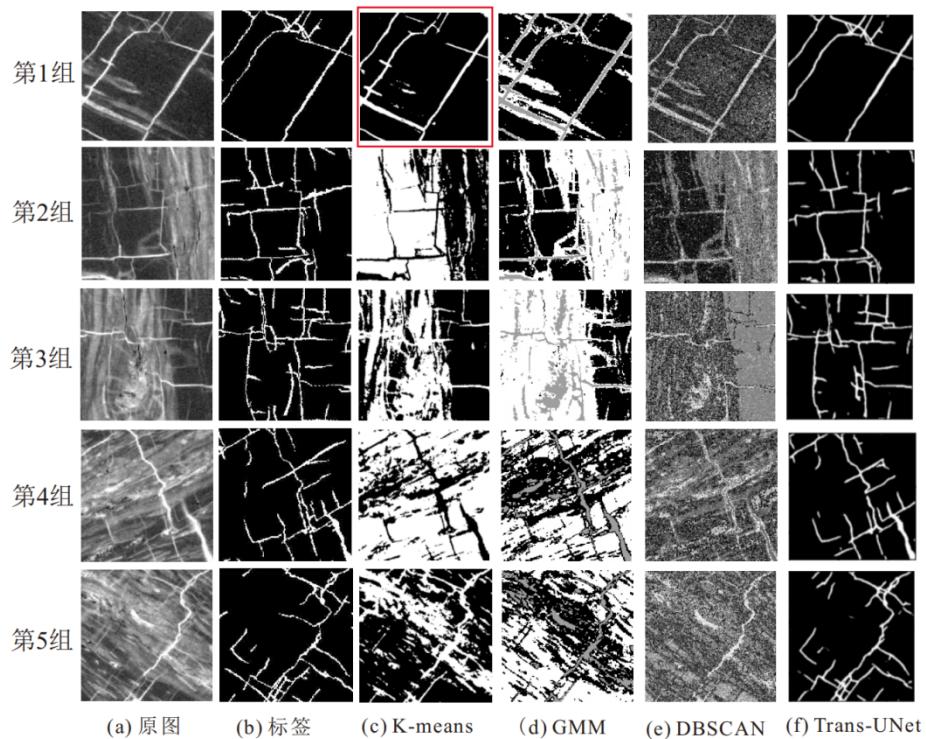


图 7 Trans-UNet 模型与三种无监督机器学习算法在 μ CT 图像上分割结果的可视化
Fig.7 Visualization of segmentation results on μ CT images by Trans-UNet model and three unsupervised machine learning algorithms

从分割精度、特征提取能力、噪声处理及局限性等方面对 4 种方法进行对比:

K-means 是一种基于欧几里得距离的无监督聚类算法, 应用于图像分割时, 能够通过像素灰度值的不同将图像分为若干类。从分割结果上看(图 7c), 它能

够在一定程度上识别煤中的裂隙(第 1 组), 但对于复杂背景下的裂隙识别效果较差, 容易受噪声影响, 产生大量伪裂隙(第 2、4 组)。

高斯混合模型算法(GMM)假设数据符合多个高斯分布, 通过将像素归属到概率最大的类别来实现分

割。分割结果表明(图 7d), GMM 在背景较简单的区域(第 1 组)能够分割出主裂隙, 然而在裂隙宽度较窄或背景纹理复杂的区域(第 3、4、5 组), 出现了较多的错分。

DBSCAN 是一种基于密度的聚类算法, 能够根据点的密度分布发现聚类结构并进行分割。DBSCAN 能识别出连通性较好的裂隙(第 1、2 组)。然而, 在裂隙交错复杂的区域(第 4、5 组), 过度分割和漏分现象较为明显。

相比之下, Trans-UNet 能够清晰、准确地提取裂隙形态, 尤其是在裂隙分布复杂或背景纹理明显的区域(第 3、5 组), 能够识别一些开度极小的裂隙, 对裂隙宽度、方向、连通性等细节有出色的捕捉能力, 显著减少了伪裂隙。

综上所述, Trans-UNet 在煤样 μ CT 图像的裂隙提取任务中展现出显著优势, 其性能明显优于 K-means、GMM 和 DBSCAN 等传统无监督算法, 能够更精准地提取裂隙特征。

3.5 Trans-UNet 模型与其他多种智能模型对比

为了进一步评估 Trans-UNet 模型在提高分割精度和处理复杂图像特征方面的有效性, 检验其在处理煤 μ CT 图像时是否具有优势, 将其与 5 种智能分割模型进行对比, 这 5 种模型分别为 U-Net、U-Net++、Mask R-CNN、DeepLab v3+ 和 SegFormer, 对比使用同一数据集, 同一超参数设置。

图 8 详细展示了 5 组煤 μ CT 图像的分割结果, 包括人工标注的裂隙标签(图 8b)和六种智能图像分割模型的分割结果(图 8c—图 8h)。

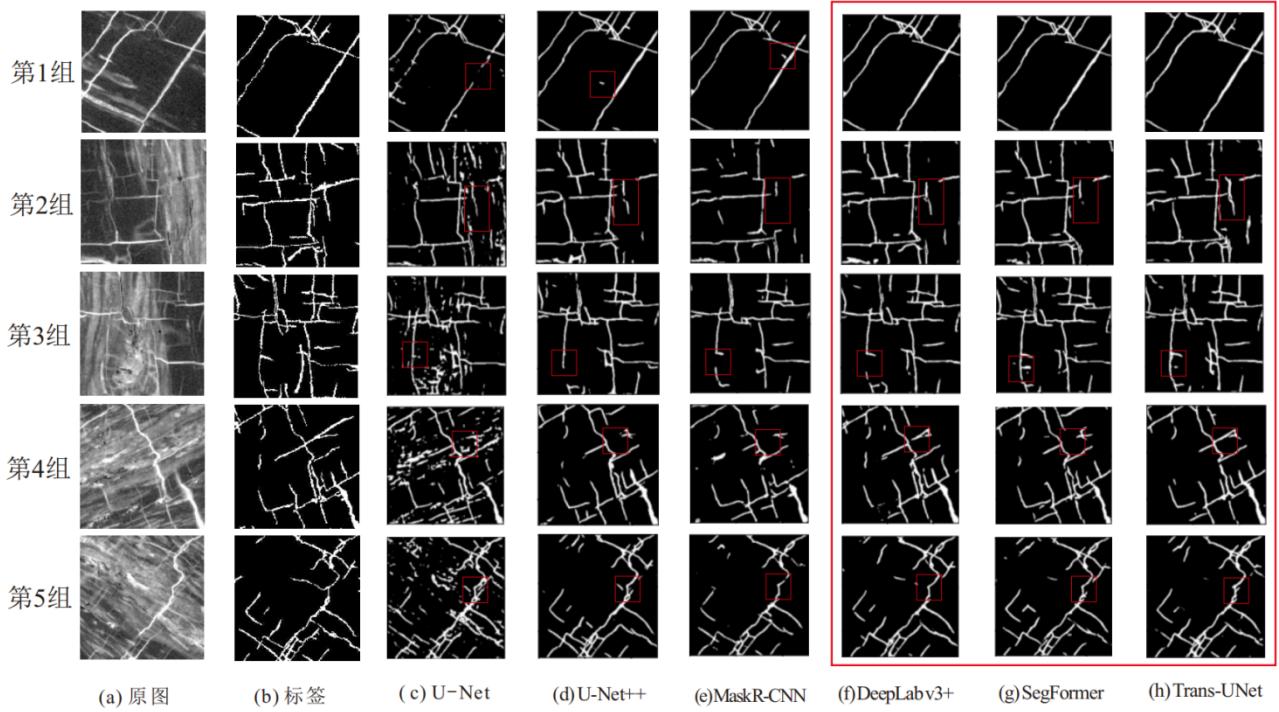


图 8 不同网络模型在 μ CT 图像上的分割结果可视化

Fig.8 Visualization of segmentation results of different network models on μ CT images

(1) U-Net 是经典的深度学习分割模型, 能够较好地提取裂隙区域(图 8c), 但在背景复杂时易出现漏检和伪裂隙(第 4、5 组), 而 Trans-UNet 在这些地方能够更准确地捕捉裂隙特征, 提升分割效果。

(2) U-Net++通过跳跃连接增强了特征传递能力, 在处理相对规律的裂隙时表现良好(图 8d), 但在复杂背景中仍存在错分问题(第 3、5 组), 而 Trans-UNet 在这类噪声较多的区域表现出更强的鲁棒性, 降低了误分率。

(3) Mask R-CNN 结合了目标检测和像素级分割功能, 能够较好处理复杂裂隙结构(图 8e), 但在狭窄或交错的裂隙区域易出现误分(第 1、4 组), 且对背景噪声处理较弱。相比之下, Trans-UNet 在这些区域展现了卓越的分割性能。

(4) DeepLab v3+通过多尺度特征提取和空洞卷积展现了较强分割能力(图 8f), 简单背景时(第 1、2 组)能够有效提取主要裂隙, 但在复杂背景下(第 3、4 组)仍存在错分和漏分问题, 而 Trans-UNet 在这些情况下的分割准确性更高。

(5) SegFormer 利用 Transformer 架构进行特征处理, 对连通性和细节的捕获能力较强(图 8g), 但在裂隙边界模糊或细小区域(第 3 组)存在误分, 而 Trans-UNet 在细节捕捉方面表现更优, 实现了更高的分割精度。

综上, Trans-UNet 在各组图像中表现出色, 尤其在裂隙分布复杂或背景纹理丰富的区域(第 3、5 组), 能够准确提取裂隙形态, 显著减少了伪裂隙。尽管在边界模糊或灰度过渡区域(第 2 组)可能出现轻微漏分, 但整体上, Trans-UNet 在细小裂隙、连通区域和边界识别方面的表现均明显优于 U-Net、U-Net++、Mask R-CNN、DeepLab v3+ 和 SegFormer 等模型, 后者往往存在错分或漏分问题。

除了定性评价外, 本研究还对这六种智能模型的分割结果进行定量分析。表 1 展示了各模型在多个性能指标上的表现。

表 1 六种网络模型评价指标对比
Table 1 Comparison of evaluation metrics of six network models

方法	准确性/%	精确度/%	召回率/%	F1 分数/%	IoU	Dice 系数
U-Net	80.3	77.1	79.5	78.3	0.65	0.78
U-Net++	87.7	84.4	86.3	85.3	0.77	0.87
Mask R-CNN	84.6	82.2	83.4	82.8	0.73	0.84
DeepLab v3+	88.4	86.5	87.1	86.8	0.81	0.90
SegFormer	90.2	88.3	88.8	88.5	0.82	0.90
Trans-UNet	91.3	89.5	90.1	89.8	0.84	0.91

(1) 准确性: 在五种模型中, Trans-UNet 在裂隙识别上的准确性最高, 达到 91.3%。U-Net(80.3%)受限于其相对简单的编码器-解码器架构, 难以在捕捉全局和细节特征之间做好平衡, 在复杂的裂隙交错区域容易出错。U-Net++(87.7%)尽管通过密集跳跃连接增强了特征融合能力, 但其局部感知能力不足, 导致某些细节区域分割错误。Mask R-CNN(84.6%)基于目标检测的两阶段方法更适合目标实例分割, 无法很好地处理裂隙这种复杂的线性结构。DeepLab v3+(88.4%)和 SegFormer(90.2%)这两者在大范围区域的分割表现较好, 但对细节(如裂隙边缘)捕捉不足, 尤其在细小裂隙区域容易导致误分。

(2) 精确度: Trans-UNet 在精确度上达到了 89.5%, 显著优于其他模型。U-Net(77.1%)和 U-Net++(84.4%)对边缘特征的提取较为模糊, 容易将背景误识别为裂隙, 导致误检较多。Mask R-CNN(82.2%)的检测机制在密集裂隙区域难以避免裂隙粘连, 进一步降低了精确度。DeepLab v3+(86.5%)和 SegFormer(88.3%)由于

对小尺度裂隙特征的捕捉能力较弱, 偶尔出现细小裂隙漏检的情况, 导致分割边界不够清晰。

(3) 召回率: Trans-UNet 在召回率上达到了 90.1%, 说明其对目标区域的捕捉能力最强。U-Net(79.5%)和 U-Net++(86.3%)在分割复杂形态的裂隙区域时, 较难捕捉到细小或模糊的裂隙, 容易导致部分裂隙被遗漏。Mask R-CNN(83.4%), 其两阶段方法的定位精度欠佳, 导致对细节区域的召回不足。DeepLab v3+(87.1%)尽管具有较强的全局感知能力, 但在边缘区域的处理稍显粗糙。SegFormer(88.8%)对整体裂隙区域识别较好, 但对极窄裂隙的敏感度稍低。

(4) F1 分数和 Dice 系数: Trans-UNet 分别达到 89.8% 和 0.91, 是综合分割性能的最佳体现。U-Net 和 U-Net++识别结果中分割边界的不完整性和细节的遗漏使得整体分割质量下降。Mask R-CNN 在特征提取过程中, 过分依赖局部信息, 导致对细微特征的表现不佳。DeepLab v3+ 和 SegFormer 在裂隙边缘处理上有所欠缺, 容易造成边界模糊或不连续。

(5) IoU(交并比): Trans-UNet 的 IoU 为 0.84, 领先于其他模型, 表明其预测区域与真实区域重合度更高。U-Net(0.65)和 U-Net++(0.77)由于对复杂裂隙特征的刻画能力有限, 造成分割结果的预测区域与真实区域偏差较大。Mask R-CNN(0.73)的两阶段框架在稀疏裂隙区域表现尚可, 但在裂隙密集区难以精确拟合目标形状。DeepLab v3+(0.81)和 SegFormer(0.82)虽然具有较好的全局信息提取能力, 但在边缘细节上表现略差, 导致区域覆盖精度下降。

总体来说, Trans-UNet 通过结合 Transformer 模块和卷积网络的特点, 有效增强了全局信息的捕捉能力, 特别是在包含复杂裂隙特征的分割任务中, 表现出了如下显著优势: (1) Transformer 能够捕捉长距离依赖, 提升裂隙整体连贯性; (2) 编码器-解码器结构有效结合了全局与局部特征, 使得裂隙区域分割更准确; (3) 多尺度特征融合能力增强, 使其对细小裂隙和复杂边缘的识别能力显著提升。这些优势使得 Trans-UNet 不仅在定量指标上全面优于其他模型, 还能在实际分割任务中提供更可靠的结果, 为煤裂隙的精细描述提供了坚实的技术支持。

3.6 消融实验

为了评估 Trans-UNet 模型中不同模块对煤裂隙提取任务的影响, 设计消融实验, 分别移除 Transformer 编码器和卷积编码器, 并与原始 U-Net 和 Trans-UNet 进行对比。U-Net 作为基线模型, 仅使用

卷积编码器和解码器。Trans-UNet 作为完整模型，包含卷积编码器、Transformer 编码器和解码器。Trans-UNet(移除 Transformer 编码器)仅使用卷积编码器和解码器，去掉了 Transformer 编码器。Trans-UNet(移除卷积编码器)仅使用 Transformer 编码器和解码器，去掉了卷积编码器。

实验中，所有方法的训练数据、超参数设置和评估指标保持一致。实验结果见表 2。

表 2 消融实验结果
Table 2 Ablation experiment results

方法	准确性/%	精确度/%	召回率/%	F1 分数/%	IoU	Dice 系数
U-Net	80.3	77.1	79.5	78.3	0.65	0.78
Trans-UNet(移除 Transformer 编码器)	85.2	83.4	84.7	84.0	0.72	0.84
Trans-UNet(移除卷积编码器)	87.6	86.1	87.2	86.6	0.78	0.88
Trans-UNet	91.3	89.5	90.1	89.8	0.84	0.91

通过表 2，可以直观地看出：

(1) Transformer 编码器在 Trans-UNet 中起到了关键作用，显著提升了模型对煤裂隙的提取能力。移除 Transformer 编码器后，模型性能有所下降，但仍优于 U-Net。准确性、精确度、召回率和 F1 分数分别下降了 6.1%、6.1%、5.4% 和 5.8%，IoU 和 Dice 系数分别下降了 0.12 和 0.07。这表明 Transformer 编码器对模型性能的提升起到了关键作用，但卷积编码器仍然能够提供一定的局部特征提取能力。

(2) 卷积编码器在提取局部特征方面也具有重要作用，但其影响相对较小。移除卷积编码器后，模型性能也出现下降，但下降幅度小于移除 Transformer 编码器。准确性、精确度、召回率和 F1 分数分别下降了 3.7%、3.4%、2.9% 和 3.2%，IoU 和 Dice 系数分别下降了 0.06 和 0.03。这表明卷积编码器在提取局部特征方面具有重要作用，但 Transformer 编码器能够部分弥补其缺失。

(3) 综合使用卷积编码器和 Transformer 编码器的 Trans-UNet 能够充分利用局部和全局信息，从而在煤裂隙提取任务中取得最佳性能。Trans-UNet 在所有指标上均显著优于其他 3 种方法，尤其与 U-Net 进行对比，在 IoU 和 Dice 系数上分别提高了 0.19 和 0.13。这表明 Transformer 编码器在捕捉全局信息和长距离依赖关系方面具有显著优势，能够更好地提取煤裂隙的特征。

4 结论

(1) 提出了一种基于 Trans-UNet 的煤裂隙网络提取方法，并在 μ CT 扫描得到的煤图像数据集上进行了实验验证。结果表明，Trans-UNet 有效提高了煤裂隙的分割精度，解决了传统方法在处理复杂裂隙网络图像中的不足。该方法不仅优于多种无监督机器学习的图像分割方法，在与其他智能模型进行对比的多项评价指标上，Trans-UNet 也表现优异。具体而言，该模型在准确性(91.3%)、精确度(89.5%) 和召回率(90.1%) 指标上显著优于原始 U-Net 等基准模型。此外，Trans-UNet 在 IoU(0.84) 和 Dice 系数(0.91) 指标上的良好表现，证明了其在处理煤裂隙任务中的有效性和可靠性。

(2) 创新性地引入 Trans-UNet 模型，结合 Transformer 模块的全局特征提取能力和卷积网络的局部细节捕捉能力，优化了模型在煤裂隙提取中的表现。研究表明，Trans-UNet 能够在复杂的煤 μ CT 图像中更精确地分割目标区域，特别是在细小裂隙和复杂边缘区域的分割效果尤为显著，该方法有望在煤层气勘探开发中得到推广应用，为煤储层裂隙的精细描述提供更加准确可靠的技术支持。

(3) 本次所用的煤数据集规模相对有限，可能影响模型的泛化能力，因此未来需要扩展和丰富数据集以进一步提升模型的分割精度和适应性。另外，尽管 Trans-UNet 展现出出色的分割性能，但其较高的计算复杂度和相对较长的训练时间却可能成为实际应用场景中实时性和高效率目标的潜在障碍。

(4) 未来相关研究应聚焦在以下几个核心方向：首先是模型优化，发掘更为高效的网络架构(采用轻量化的 Transformer 模块)以及创新性的训练策略，削减计算复杂度、缩短训练时间，从而显著提升模型在实际部署中的运行效率。其次是数据扩展，通过获取更加多样化的煤 μ CT 图像数据，使得数据涵盖更广泛的地质背景，以增强模型的泛化能力和稳定性。最后是实际应用验证，它是检验研究成果的关键环节，需进一步推进研究成果在煤层气勘探开发现场实时检测上的应用，通过实践验证模型的可靠性、实用性和适应性，确保其在真实环境中能够发挥预期效能。

符号注释：

A 为准确度，%； $D(x,y)$ 为两幅图像在位置 (x,y) 的像素差值；F1 为 F1 分数，%；Loss 为总损失值；CrossEntropyLoss 为交叉熵损失；DiceLoss 为 Dice 损失； N 为图像的数量，个； $I_1(x,y)$ 和 $I_2(x,y)$ 分别为两张图像在 (x,y) 处的像素值； p_i 为预测分割图的像元和； P

为精确度, %; R 为召回率, %; IoU 为交并比; TP 为真正例, 个; TN 为真反例, 个; FP 为假正例, 个; FN 为假反例, 个; X 为模型分割出的预测区域; y_i 为真实分割图的像元和; Y 为真实区域; α 、 β 分别为交叉熵损失的权重系数和 Dice 损失的权重系数, 用于平衡两种损失的影响; ϵ 为很小的正数, 用于防止数值不稳定。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

参考文献(References)

- [1] JIN Yi, ZHENG Junling, DONG Jiabin, et al. Fractal topography and complexity assembly in multifractals[J]. *Fractals*, 2022, 30(3): 2250052.
- [2] 李国永, 姚艳斌, 王辉, 等. 鄂尔多斯盆地神木-佳县区块深部煤层气地质特征及勘探开发潜力[J]. 煤田地质与勘探, 2024, 52(2): 70–80.
LI Guoyong, YAO Yanbin, WANG Hui, et al. Deep coalbed methane resources in the Shenmu-Jiaxian Block, Ordos Basin, China: Geological characteristics and potential for exploration and exploitation[J]. *Coal Geology & Exploration*, 2024, 52(2): 70–80.
- [3] 李斌, 杨帆, 张红杰, 等. 神府区块深部煤层气高效开发技术研究[J]. 煤田地质与勘探, 2024, 52(8): 57–68.
LI Bin, YANG Fan, ZHANG Hongjie, et al. Technology for efficient production of deep coalbed methane in the Shengfu Block[J]. *Coal Geology & Exploration*, 2024, 52(8): 57–68.
- [4] MOORE T A. Coalbed methane: A review[J]. *International Journal of Coal Geology*, 2012, 101: 36–81.
- [5] 叶桢妮, 侯恩科, 段中会, 等. 不同煤体结构煤的孔隙-裂隙分形特征及其对渗透性的影响[J]. 煤田地质与勘探, 2019, 47(5): 70–78.
YE Zhenni, HOU Enke, DUAN Zhonghui, et al. Fractal characteristics of pores and microfractures of coals with different structure and their effect on permeability[J]. *Coal Geology & Exploration*, 2019, 47(5): 70–78.
- [6] 施雷庭, 赵启明, 任镇宇, 等. 煤岩裂隙形态对渗流能力影响数值模拟研究[J]. 油气藏评价与开发, 2023, 13(4): 424–432.
SHI Leiting, ZHAO Qiming, REN Zhenyu, et al. Numerical simulation study on the influence of coal rock fracture morphology on seepage capacity[J]. *Petroleum Reservoir Evaluation and Development*, 2023, 13(4): 424–432.
- [7] 李学博, 刘春春, 张武昌, 等. 高煤阶煤裂隙发育特征对煤层气开发的影响: 以沁水盆地南部郑庄区块为例[J]. 中国煤层气, 2022, 19(3): 12–15.
LI Xuebo, LIU Chunchun, ZHANG Wuchang, et al. Influence of development characteristics of high-rank coal fractures on coalbed methane development: Taking Zhengzhuang Block in South Qinshui Basin as an example[J]. *China Coalbed Methane*, 2022, 19(3): 12–15.
- [8] 王昱, 宋晓夏, 胡咤咤, 等. 西山煤田屯兰区块煤层气低产井的增产改造措施及效果分析[J/OL]. 河南理工大学学报(自然科 学版), 2024 : 1–13 [2025-02-13]. <http://kns.cnki.net/kcms/detail/41.1384.N.20240821.1707.002.html>.
- [9] 王跃鹏, 孙正财, 刘向君, 等. 煤层割理结构及其对井壁稳定的影响研究[J]. 油气藏评价与开发, 2020, 10(4): 45–52.
WANG Yuepeng, SUN Zhengcai, LIU Xiangjun, et al. Study on cleat structure and its influence on wellbore stability in coal seams[J]. *Reservoir Evaluation and Development*, 2020, 10(4): 45–52.
- [10] 胡秋嘉, 刘世奇, 毛崇昊, 等. 基于 X-ray CT 与 FIB-SEM 的无烟煤孔隙发育特征[J]. 煤矿安全, 2021, 52(9): 10–15.
HU Qiujiā, LIU Shiqī, MAO Chonghao, et al. Characteristics of pores and fractures in anthracite coal based on X-ray CT and FIB-SEM[J]. *Safety in Coal Mines*, 2021, 52(9): 10–15.
- [11] KETCHAM R A. Computational methods for quantitative analysis of three-dimensional features in geological specimens[J]. *Geosphere*, 2005, 1(1): 32–41.
- [12] IASSONOV P, GEBRENEGUS T, TULLER M. Segmentation of X-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures[J]. *Water Resources Research*, 2009, 45(9): W09415.
- [13] DENG Hang, FITTS J P, PETERS C A. Quantifying fracture geometry with X-ray tomography: Technique of Iterative Local Thresholding (TILT) for 3D image segmentation[J]. *Computational Geosciences*, 2016, 20(1): 231–244.
- [14] GOLAB A, WARD C R, PERMANA A, et al. High-resolution three-dimensional imaging of coal using microfocus X-ray computed tomography, with special reference to modes of mineral occurrence[J]. *International Journal of Coal Geology*, 2013, 113: 97–108.
- [15] WILDENSCHILD D, SHEPPARD A P. X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems[J]. *Advances in Water Resources*, 2013, 51: 217–246.
- [16] RAMANDI H L, IRTZA S, SIROJAN T, et al. FracDetect: A novel algorithm for 3D fracture detection in digital fractured rocks[J]. *Journal of Hydrology*, 2022, 607: 127482.
- [17] HU Zhazha, LU Shuangfang, KLAVER J, et al. An integrated imaging study of the pore structure of the Cobourg limestone: A potential nuclear waste host rock in Canada[J]. *Minerals*, 2021, 11(10): 1042.
- [18] MATHEWS J P, CAMPBELL Q P, XU Hao, et al. A review of the application of X-ray computed tomography to the study of coal[J]. *Fuel*, 2017, 209: 10–24.
- [19] TAN Jianquan, ZHOU Wenrui, LIN Ling, et al. A review of semantic medical image segmentation based on different paradigms[J]. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 2024, 20(1): 1–25.
- [20] KNACKSTEDT M A, LATHAM S, MADADI M, et al. Digital rock physics: 3D imaging of core material and correlations to acoustic and flow properties[J]. *The Leading Edge*, 2009, 28(1): 28–33.
- [21] CNUDDE V, BOONE M N. High-resolution X-ray computed

- tomography in geosciences: A review of the current technology and applications[J]. Earth-Science Reviews, 2013, 123: 1–17.
- [22] LEI Lian, YANG Qiliang, YANG Ling, et al. Deep learning implementation of image segmentation in agricultural applications: A comprehensive review[J]. Artificial Intelligence Review, 2024, 57(6): 149.
- [23] 尹艺晓, 马金刚, 张文凯, 等. 从 U-Net 到 Transformer: 混合模型在医学图像分割中的应用进展[J/OL]. 激光与光电子学进展, 2024: 1–38 [2025-02-13]. <http://kns.cnki.net/kcms/detail/31.1690.TN.20240612.0920.042.html>.
YIN Yixiao, MA Jingang, ZHANG Wenkai, et al. From U-Net to Transformer: Progress in the application of hybrid models in medical image segmentation[J/OL]. Laser & Optoelectronics Progress, 2024 : 1–38 [2025-02-13]. <http://kns.cnki.net/kcms/detail/31.1690.TN.20240612.0920.042.html>.
- [24] 王登科, 房禹, 魏建平, 等. 基于深度学习的煤岩 Micro-CT 裂隙智能提取与应用[J]. 煤炭学报, 2024, 49(8): 3439–3452.
WANG Dengke, FANG Yu, WEI Jianping, et al. Intelligent extraction of Micro-CT fissures in coal based on deep learning and its application[J]. Journal of China Coal Society, 2024, 49(8): 3439–3452.
- [25] 郝天轩, 徐新革, 赵立桢. 煤岩裂隙图像识别方法研究[J]. 工矿自动化, 2023, 49(10): 68–74.
HAO Tianxuan, XU Xinge, ZHAO Lizhen. Research on image recognition methods for coal rock fractures[J]. Journal of Mine Automation, 2023, 49(10): 68–74.
- [26] 郑江韬, 齐子豪, 刘佳存, 等. 基于卷积神经网络的煤岩微裂隙提取方法[J]. 矿业科学学报, 2022, 7(6): 680–688.
ZHENG Jiangtao, QI Zihao, LIU Jiacun, et al. Segmentation of micro-cracks in fractured coal based on convolutional neural network[J]. Journal of Mining Science and Technology, 2022, 7(6): 680–688.
- [27] YU Jinxia, WU Chengyi, LI Yingying, et al. Intelligent identification of coal crack in CT images based on deep learning[J]. Computational Intelligence and Neuroscience, 2022, 2022(1): 7092436.
- [28] LU Fengli, FU Chengcai, ZHANG Guoying, et al. Convolution neural network based on fusion parallel multiscale features for segmenting fractures in coal-rock images[J]. Journal of Electronic Imaging, 2020, 29(2): 23008.
- [29] KARIMPOULI S, TAHMASEBI P, SAENGER E H. Coal cleat/fracture segmentation using convolutional neural networks[J]. Natural Resources Research, 2020, 29(3): 1675–1685.
- [30] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000–6010.
- [31] RAMANDI H L, ARMSTRONG R T, MOSTAGHIMI P. Micro-CT image calibration to improve fracture aperture measurement[J]. Case Studies in Nondestructive Testing and Evaluation, 2016, 6: 4–13.
- [32] SUN Haoran. A review of 3D–2D registration methods and applications based on medical images[J]. Highlights in Science, Engineering and Technology, 2023, 35: 200–224.
- [33] 孙书魁, 范菁, 孙中强, 等. 基于深度学习的图像数据增强研究综述[J]. 计算机科学, 2024, 51(1): 150–167.
SUN Shukui, FAN Jing, SUN Zhongqiang, et al. Survey of image data augmentation techniques based on deep learning[J]. Computer Science, 2024, 51(1): 150–167.
- [34] 章展熠, 张宝荃, 王周立, 等. 多茶类 CNN 图像识别的数据增强优化及类激活映射量化评价[J]. 茶叶科学, 2023, 43(3): 411–423.
ZHANG Zhanyi, ZHANG Baoquan, WANG Zhouli, et al. Data enhancement optimization and class activation mapping quantitative evaluation for CNN image recognition of multiple tea categories[J]. Journal of Tea Science, 2023, 43(3): 411–423.
- [35] 王气洪, 贾洪杰, 黄龙霞, 等. 联合数据增强的语义对比聚类[J]. 计算机研究与发展, 2024, 61(6): 1511–1524.
WANG Qihong, JIA Hongjie, HUANG Longxia, et al. Semantic contrastive clustering with federated data augmentation[J]. Journal of Computer Research and Development, 2024, 61(6): 1511–1524.
- [36] SU Run, ZHANG Deyun, LIU Jinhuai, et al. MSU-Net: Multi-scale U-Net for 2D medical image segmentation[J]. Frontiers in Genetics, 2021, 12: 639930.
- [37] SHAN Liang, HU Bin, CHEN Long, et al. Detecting COVID-19 on CT images with impulsive-backpropagation neural networks[C]//2022 34th Chinese Control and Decision Conference (CCDC), 2022: 2797–2803.
- [38] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Delving deep into rectifiers: Surpassing human-level performance on image net classification[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 1026–1034.
- [39] 张成业, 李飞跃, 李军, 等. 基于 DeepLabv3+ 与 GF-2 高分辨率影像的露天煤矿区土地利用分类[J]. 煤田地质与勘探, 2022, 50(6): 94–103.
ZHANG Chengye, LI Feiyue, LI Jun, et al. Recognition of land use on open-pit coal mining area based on DeepLabv3+ and GF-2 high-resolution images[J]. Coal Geology & Exploration, 2022, 50(6): 94–103.

(责任编辑 范章群)