融合双编码器注意力网络的钻孔结构面分割与参数可视化方法*

时乙入[1,2],段东[1,2],李杰[1,2],冯小静[1,2],何家雄[1,2],陆卫东[3,4],甘元平[3,4]

(1.太原理工大学 矿业工程学院, 山西 太原 030024;
2.太原理工大学 原位改性采矿教育部重点实验室, 山西 太原 030024;
3.新疆工程学院 安全科学与工程学院, 新疆 乌鲁木齐 830023;
4.新疆工程学院 新疆煤矿灾害智能防控与应急重点实验室, 新疆 乌鲁木齐 830023)

摘要:针对钻孔图像光照不均、背景复杂及结构面形态差异大等难题,提出了一种双编码器交互式注意力增强融合网络(DISA-FuseNet)。该网络基于UNet编码器-解码器架构,采用Swin Transformer和ResNet50双编码器结构,同时捕捉全局上下文信息与局部细节特征;通过统计聚合通道注意力模块(SACA)进行注意力加权,抑制岩体背景噪声和无关信息;底层特征经拼接后输入自适应层级增强模块(AHE),强化网络对多尺度结构面特征的捕捉能力。基于自建钻孔结构面数据集开展试验验证,将DISA-FuseNet与UNet、Deeplabv3+、Trans-UNet、DeepCrack及CrackFormer五种模型进行对比。结果表明,DISA-FuseNet在平均交并比、平均像素准确率、召回率、精确率、F1分数指标上均优于其他模型,各指标较基准模型UNet分别提升了8.44,9.51,8.52,10.60,9.62个百分点。进一步提出结构面参数提取与三维可视化方法,实现结构面参数的精准提取与交互式三维可视化模型构建。经实际工程案例验证,该方法具备良好的有效性和工程适用性。

关键词:结构面识别;钻孔图像;双编码器;交互式注意力增强融合网络;三维可视化

中图分类号:TD313;TP391 **文献标识码**:A **文章编号**:1005-2763(2026)02-0288-15

Borehole Structural Plane Segmentation and Parameter Visualization Method with Dual-Encoder Attention Network

SHI Yiru[1,2], *DUAN Dong*[1,2], *LI Jie*[1,2], *FENG Xiaojing*[1,2], *HE Jiaxiong*[1,2], *LU Weidong*[3,4], *GAN Yuanping*[3,4]

(1.College of Mining Engineering, Taiyuan University of Technology, Taiyuan, Shanxi 030024, China;
2.Key Laboratory of In-situ Property-improving Mining of Ministry of Education, Taiyuan University of Technology, Taiyuan, Shanxi, 030024, China;
3.College of Safety Science and Engineering, Xinjiang Institute of Engineering, Urumqi, Xinjiang 830023, China;
4.Xinjiang Key Laboratory of Coal Mine Disaster Intelligent Prevention and Emergency Response, Xinjiang Institute of Engineering, Urumqi, Xinjiang 830023, China)

Abstract: Aiming at the problems of uneven illumination, complex background and large difference in structural plane morphology of borehole images, a dual-encoder interactive attention enhanced fusion network (DISA-FuseNet) was proposed. Based on the UNet encoder-decoder architecture, the network adopted the dual-encoder structure of Swin Transformer and ResNet50 to capture global context information and local detail features simultaneously. The attention was weighted by the Statistical Aggregation Channel Attention (SACA) module to suppress the background noise and irrelevant information of the rock mass. The underlying features were inputted into the Adaptive Hierarchical Enhancement module (AHE) after splicing to enhance the network's ability to capture multi-scale structural plane features. Experimental verification was conducted based on the self-constructed borehole structural plane dataset, and DISA-FuseNet was compared with five models, namely UNet, Deeplabv3+, Trans-UNet, DeepCrack and CrackFormer. The results show that DISA-FuseNet outperforms all the other models in terms of mean intersection over union, mean pixel accuracy, recall, precision and F1-score, these indexes are

* **收稿日期**:2025-01-22
基金项目:国家自然科学基金项目(52274134,52474139,51604184);山西省基础研究计划项目(202403011242005)
作者简介:时乙入(2000—),女,山西运城人,硕士研究生,主要研究方向为图像分割、结构面特征参数数据分析。E-mail:Shiyiru1213@gmail.com
通信作者:段东(1979—),男,辽宁大连人,博士,副教授,主要研究方向为岩石力学。E-mail:duandong@tyut.edu.cn

increased by 8.44, 9.51, 8.52, 10.60 and 9.62 percentage points respectively compared with those of the baseline model UNet. Furthermore, a method for structural plane parameter extraction and 3D visualization is proposed, which enables the accurate extraction of structural plane parameters and the construction of an interactive 3D visualization model. The actual engineering case verifies that the method has good effectiveness and engineering applicability.

Key words: Structural plane identification, Borehole image, Dual-encoder, Interactive attention enhanced fusion network, 3D visualization

0 引言

精确捕捉结构面特征是分析岩体力学性质和评估工程岩体稳定性的首要工作[1]。结构面特征提取方法包括三维激光扫描技术[2]、岩石取芯[3]、钻孔摄像技术[4]等,其中钻孔摄像技术因能直观且清晰地揭示钻孔围岩内部的情况,在结构面探测方面具有无可替代的优势,目前已被广泛应用于多个领域,包括地质灾害评估、能源勘探、矿产勘查、水利工程和建筑行业等[5],同时广泛应用于煤矿井下,为岩体结构面特征分析提供了保障。目前,基于钻孔摄像技术的结构面识别与分类工作大多依赖于人工操作,存在工作量大、主观性强、易因人为因素导致识别遗漏或误差等问题。为有效减少操作错误并提高工作效率,众多专家学者运用计算机视觉技术进行结构面的智能识别。传统数字图像处理技术主要基于结构面的线性与灰度特征来实现识别和分割,葛云峰等[6]利用灰度共生矩阵技术定位钻孔图像中的结构面,借助 Canny 边缘检测算法,进一步实现了对结构面边缘的识别和提取。YANG 等[7]通过融合 Gabor 滤波器生成的纹理特征和钻孔图像颜色信息,采用模糊 c 均值聚类算法实现了结构面与完整区域的分离。LI 等[8]建立了基于 HSV 颜色空间的自适应岩土分层检测模型,并对模型所生成的二值化图像进行滤波处理,以抑制图像噪声,实现了对裂隙的分割。李清波等[9]通过浮点算法将钻孔图像转换为灰度图,提出了一种基于边缘信息的自适应阈值选择策略,以此实现了钻孔结构面的分割。上述研究为钻孔图像结构面智能化识别奠定了基础,但传统的数字图像处理技术对图像所呈现出的复杂纹理特征及结构面色彩的深浅差异考虑较少。

近年来,卷积神经网络(CNN)模型及基于自注意力机制的 Transformer 模型在图像处理领域均表现出明显的优势。在目标检测[10-11]、图像分类[12-13]和语义分割[14-15]中均得到广泛应用,一些学者也逐渐将深度学习算法应用于结构面提取中。DIAS 等[16]提出利用目标检测算法 Fast R-CNN 对钻孔图像中的结构面进行自动识别,以提高高对比度钻孔图像结构面自动检测的准确性。LIU 等[17]针对不同结构面间存在的类别不平衡问题,采用由两个 DeepLabV3+网络组成的双重网络框架来识别正弦垂直裂隙和破碎带,实现了不同类别结构面的分割。张野等[18]对 Unet、SegNet、DeepLabV3 和传统图像处理方法进行对比分析,证明了深度模型在应对含有多噪声的钻孔图像分割任务时的优越性。ZHANG 等[19]将预训练的 ResNet 网络与经典的 Unet 架构相融合,以实现钻孔结构面分割。李炜等[20]采用以 CSPDarkNet53 为特征提取网络的 YOLOv4 算法,对钻孔图像中的结构面进行识别定位。吴金等[21]提出了一种基于改进的 U-Net 网络模型的钻孔图像结构面识别方法,该模型加深了 Unet 网络层次,引入了通道注意力残差模块(SER)、卷积通道注意力模块(SEC)、多尺度空间注意力模块(RSAM),在处理复杂钻孔地层图像中的结构面时更占优势。

相较于传统的图像处理方法,深度学习技术在处理构造复杂的深层岩体结构面钻孔图像时展现出了更为优越的性能。但由于钻孔图像获取过程中光照条件较差,图像细节模糊,导致复杂的岩体背景中存在与结构面相似的高对比度纹理,使分割难度进一步加大。同时,结构面的深度不一,使其在不同压力条件下的形态各异,不同结构面之间在宽度与对比度上存在显著差异。目前,基于深度学习的多数研究仅限于对现有的网络模型进行参数调整,未能根据钻孔图像结构面特征设计出更加适配的网络结构。上述研究中应用的单一编码器架构在应对钻孔图像特有的复杂背景与多变结构面时,难以同时捕捉全局上下文信息与局部细节特征,导致对结构面的精确定位和边缘识别能力不足。同时,单一卷积操作因具有局部感受野和权重共享特性,易受噪声和高对比度纹理的干扰。传统注意力机制在处理多尺度特征和形态多样的结构面时缺乏足够的灵活性

和适应性,限制了模型在钻孔图像上的鲁棒性。针对上述问题,本文提出一种贴合结构面图像特性以及实际工程应用需求的双编码器分割模型 DISA-FuseNet(Dual-encoder Interactive SACA & AHE Enhanced Fusion Network),并设计算法进行结构面长度、宽度、粗糙度、倾向、倾角等参数提取和三维图像可视化,为岩体力学特征分析和工程设计提供了技术与可视化支持。

1 结构面分割

1.1 DISA-FuseNet 结构框架

DISA-FuseNet 整体结构如图 1 所示,提出了 Swin Transformer[22] 与 ResNet50[23] 双编码器融合结构,前者通过自注意力机制高效建模全局上下文信息,以适应长距离结构面形态,后者的残差连接则保障了梯度传播完整性,并提取了丰富的结构面边缘特征。针对岩体复杂背景问题,提出统计聚合通道注意力模块(Statistical Aggregation Channel Attention,SACA),进一步通过结合中值增强的空间通道注意力机制,动态调整特征图的通道与空间权重,抑制岩体背景噪声和无关信息影响。针对图像存在形态多样、宽度差异明显的结构面问题,提出自适应层级增强模块(Adaptive Hierarchical Enhancement,AHE),通过可变形空洞卷积动态适配不同尺度结构面几何特征,并结合注意力机制优化特征融合。

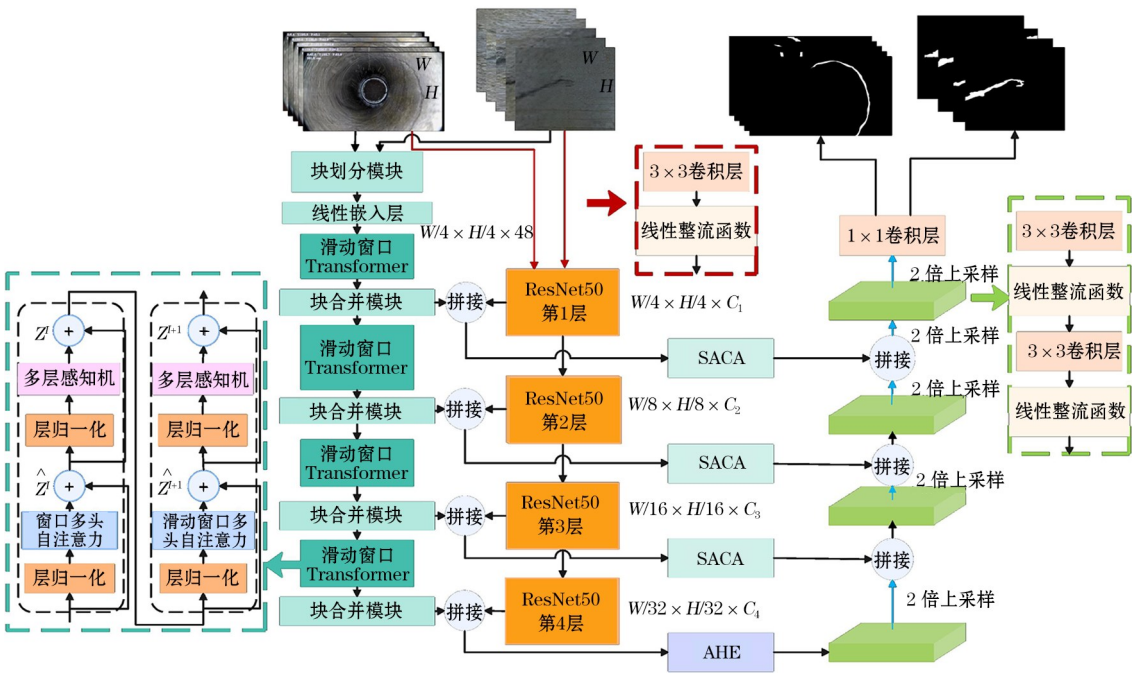


图 1 DISA-FuseNet 模型结构

Fig.1 DISA-FuseNet model structure

1.2 双编码器结构

沿用 U-Net[24] 的基础框架,基于编码器-解码器结构对骨干网络进行针对性设计。针对钻孔图像中结构面宽度跨量级变化与空间分布随机性,选用预训练的 ResNet50 和 Swin Transformer 进行相应层级间的逐层融合,同时捕捉微小结构面的边缘梯度特征和贯穿性结构面的拓扑连续性特征。拼接后的特征图通过统计聚合通道注意力模块 SACA 进行通道注意力加权,抑制岩体背景噪声干扰,底层特征拼接后输入自适应层级增强模块 AHE,通过并行的多个空洞卷积,同时捕捉不同尺度的结构面特征,充分利用各层有效特征,通过跳跃连接将融合后特征传递到解码器。解码器通过双卷积结构,应用转置卷积实现特征上采样。其中,ResNet50 由初始卷积层和 4 个残差层级(第 1 层至第 4 层)共五部分组成,网络结构如图 2 所示。网络中每层瓶颈层包含卷积块与恒等块(图中只展示出卷积块),卷积块通过 1×1 卷积进行降维操作,经过 3×3 卷积后使用再恢复维度,瓶颈结构在保证特征表达力的同时降低了计算复杂度,对处理高分辨率钻孔图像具有显著工程价值。跳跃连接部分同样使用 1×1 卷积调整维度(见图 2 中的虚线),恒等块维度不改变。

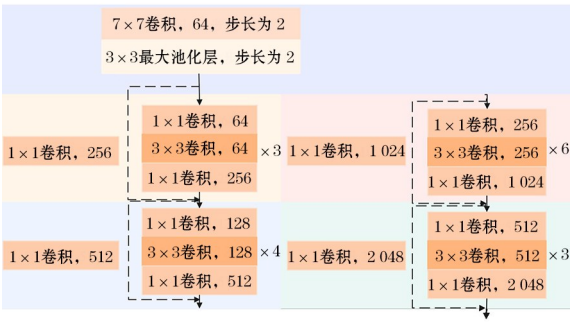


图 2　ResNet-50 网络结构

Fig.2　ResNet-50 network structure

Swin-Transformer 通过在连续的自注意力层之间移动窗口分区，从而替代 Transformer 中的多头自注意力模块（Multi-Head Self-Attention，MSA），其动态窗口划分策略与钻孔结构面无规律分布特征相适配。Swin-Transformer 模块主要由以下几个组件组成：层归一化（Layer Normalization，LN）在每个子层前进行标准化，确保输入特征具有零均值和单位方差，稳定训练过程；基于窗口的多头自注意力（Window-based Multi-Head Self-Attention，W-MSA）将输入特征划分为不重叠的窗口，在每个窗口内独立执行多头自注意力计算，降低处理钻孔图像时的内存消耗。计算公式见式(1)。

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{SoftMax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}+\boldsymbol{B}\right)\boldsymbol{V} \tag{1}$$

式中：$\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}\in\mathbb{R}^{M^2\times d}$ 分别为查询、键和值矩阵，$\mathbb{R}$ 为实数集；d 表示查询和键矩阵的维度；$\boldsymbol{B}\in\mathbb{R}^{(2M-1)\times(2M+1)}$ 表示偏置矩阵；M 为窗口尺寸；SoftMax( )为标准化函数。

移动窗口多头自注意力（Shifted Window Multi-Head Self-Attention，SW-MSA）将窗口划分进行平移，实现跨窗口信息交流，其跨窗口关联能力可捕捉岩体节理分隔的长距离结构面拓扑关系。连续的 Swin-Transformer 模块交替应用 W-MSA、SW-MSA，增强全局特征捕捉能力。多层感知机（Multi-Layer Perceptron，MLP）包含两层全连接层和一个 GELU 激活函数，用于对特征进行非线性变换，见式(2)。

$$MLP(\boldsymbol{X})=\mathrm{GELU}(\boldsymbol{X}\boldsymbol{W}_1+\boldsymbol{b}_1)\boldsymbol{W}_2+\boldsymbol{b}_2 \tag{2}$$

式中：$\boldsymbol{W}_1$、$\boldsymbol{W}_2$ 为权重矩阵；$\boldsymbol{b}_1$、$\boldsymbol{b}_2$ 为偏置向量；$\boldsymbol{X}$ 为输入特征向量；GELU( )为高斯误差线性单元激活函数。

每个子层后进行一次残差连接，缓解深层网络中的梯度消失问题。Swin-Transformer 完整计算过程见式(3)。

$$\begin{cases}\hat{z}_l=W\text{-}\mathrm{MSA}(\mathrm{LN}(z_{l-1}))+z_{l-1}\\ z_l=\mathrm{MLP}(\mathrm{LN}\hat{z}_l)+\hat{z}_l\\ \hat{z}_{l+1}=\mathrm{SW\text{-}MSA}(\mathrm{LN}(z_l))+z_l\\ z_{l+1}=\mathrm{MLP}(\mathrm{LN}(\hat{z}_{l+1}))+\hat{z}_{l+1}\end{cases} \tag{3}$$

式中，$\hat{z}_l$、z_l 分别表示第 l 个模块中的 W-MSA 和 SW-MSA 模块的输出。

1.3　统计聚合通道注意力模块

现有的通道注意力机制通常依赖于全局平均池化和全局最大池化来提取特征图的全局统计信息。然而，当存在复杂纹理与噪声时，此类方法进行特征提取易受其干扰，难以保证特征信息的有效性。针对钻孔图像中的离散型噪声，传统池化算子存在固有缺陷：平均池化会扩散噪声影响，最大池化则易受极端值干扰，故提出三模态统计聚合策略。边缘和纹理是图像中的重要特征，反映了物体的形状和细节。中值池化通过保留局部像素值的中位数，有效反映出图像具有连续性的边缘变化，且不因单个像素的变化而造成边缘模糊或断裂，以维护边缘的连续性和纹理的细腻度，其非线性特性能灵活适应不同类型图像的特征变化，在处理复杂特征时能更好地捕捉到多样化的特征模式。将中值池化与全局平均池化和全局最大池化结合应用于通道注意力机制，充分挖掘输入特征图的不同特性。全局平均池化可捕捉特征图的整体均值信息，全局最大池化强调特征图中的显著响应，而全局中值池化则在去除噪声的同时保留重要的特征信息。多统计聚合方法使通道注意力图能够更加全面、稳健地反映特征图的关键信息。SACA 模块如图 3 所示，其对输入特征图 $F\in\mathbb{R}^{C\times H\times W}$ 进行全局平均池化、全局最大池化和全局中值池化，分别得到 3 个池化结果 F_{avg}，F_{max}，$F_{\mathrm{median}}\in\mathbb{R}^{C\times1\times1}$，每个池化结果通过多层感知器 MLP 进行处理。MLP 由两个 1×1 卷积层和 ReLU 激活函数组成。第一个卷积层将特征维度从 C 降至 C/r，其中 r 为降维比例，第二个卷积层将特征维度恢复至 C，使用 Sigmoid 激活函数将输出值压缩至[0，1]范围内，分别得到 3 个注意力图 A_{avg}，A_{max}，$A_{\mathrm{median}}\in\mathbb{R}^{C\times1\times1}$，将 3 个注意力图进行元素级相加，得到最终的通道注意力图 A_{c}。

$$A_{\mathrm{c}}=\sigma(\mathrm{MLP}(F_{\mathrm{avg}}))+\sigma(\mathrm{MLP}(F_{\mathrm{max}}))+\sigma(\mathrm{MLP}(F_{\mathrm{median}})) \tag{4}$$

式中，σ 为激活函数。

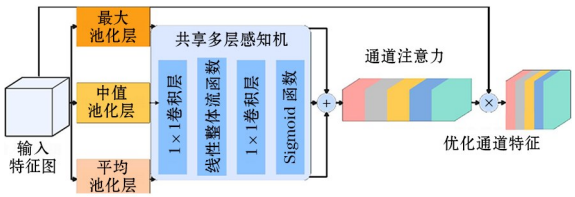


图 3 SACA 模块

Fig.3 SACA module

将通道注意力图 A_c 与原始输入特征图 F 进行元素级相乘,得到加权后的特征图 F':

$$F'=A_c\odot F \tag{5}$$

式中,⊙表示元素级相乘。

1.4 自适应层级增强模块

钻孔图像中结构面呈现直线型、弯曲型或分支形态,形态复杂多样,常含跨量级宽度差异的复合结构面。传统模块受固定感受野限制,难以兼顾裂隙宏观全局形态与局部细节特征,此外,全局特征与局部细节之间的关联未被充分建模,导致复杂结构面形态分割效果不佳。AHE 模块通过引入不同空洞率的多分支卷积,结合通道和空间注意力机制,可对多尺度特征进行校准,强化关键区域和重要通道的特征表达,在保留结构面局部细节信息的同时捕捉到裂隙的全局形态,模块结构如图 4 所示。

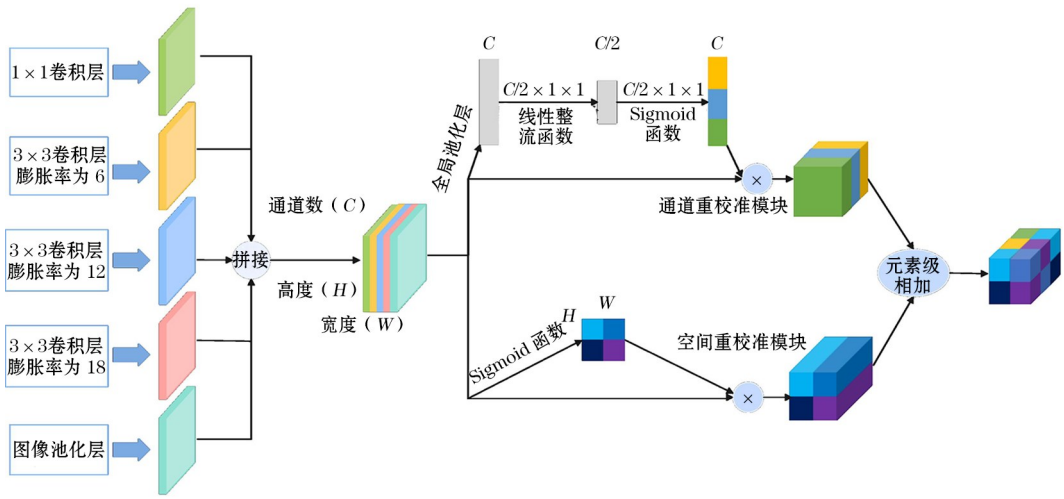


图 4 AHE 模块

Fig.4 AHE module

AHE 结合多尺度空洞卷积[25],并融合通道与空间注意力机制。多尺度空洞卷积部分包含 5 个分支,分支 1 采用 1×1 密集卷积核提取结构面边缘的高频信号;分支 2,3,4 采用 3×3 空洞卷积,空洞率逐层增加,其中中等空洞率卷积建立局部特征间的空间关联性,大空洞率分支利用扩展感受野建模结构面走向的连续性;分支 5 利用全局平均池化提取全局上下文信息。各分支输出经插值恢复至统一空间尺寸后,每个分支的输出通过通道维度拼接,形成综合特征图。多尺度特征整合后,通过通道和空间注意力机制对特征进行自适应调整,通道注意力对综合特征图进行全局平均池化,使用两层全连接层(FC)计算每个通道的重要性权重,通过 Sigmoid 激活函数进行归一化处理,将权重与原始特征逐通道相乘,完成通道加权。空间注意力在通道维度进行全局池化,使用 1×1 卷积生成每个空间位置的重要性权重,并与原始特征逐元素相乘,完成空间加权。随后,通道特征图与空间特征图通过逐元素加法融合,并进一步与原始特征图融合,以保留多尺度信息的完整性。最终通过 1×1 卷积实现特征降维,生成最终的输出特征图。

2 试验与分析

2.1 数据集

本试验数据源于 3 个钻孔深度为 30 m 的顶板岩层样本,采集其孔壁 360°全景展开图像作为原始数据集,所选岩层涵盖砂岩、泥岩等多种岩性,图 5 为钻孔摄像视频截图,图 6 为一小段连续的具有代表性的试验钻孔图像,图中展示了 3 种类型结构面,其中红色矩形框标选的结构面张开度较小,蓝色矩形框标选的结构面张开明显,黄色矩形框标选岩体破碎带(颜色标识见电子版)。各类结构面边界存在不同起伏特征,长度贯穿孔壁展开图的结构面居于多数,同时呈现如展开图左侧孔壁背景纹理较复杂

的现象。试验数据集裁剪图像大小为 380 px×320 px,横向和纵向裁剪步距分别为 120 px 和 100 px,共获取 2 000 张尺寸为 380 px×320 px 的图像。同时,考虑到实际工程中测定围岩物理力学性质时,普遍采用人工观看窥孔视频、通过视觉判断记录钻孔内部结构面发育状况的方法,因此选取图 5 中的直观反映围岩质量状态及结构面细节特征的钻孔摄像视频截图作为辅助数据集。此外,为降低过拟合风险,提高模型的泛化能力,本文对图像进行了数据增强操作,包括翻转、色彩扰动、错切形变、透视形变、弹性扭曲等 8 种方式,最终获取 14 856 张钻孔结构面图像,并按 8∶1∶1 的比例划分数据集,其中训练集 11 884 张,验证集 1 486 张,测试集 1 486 张。

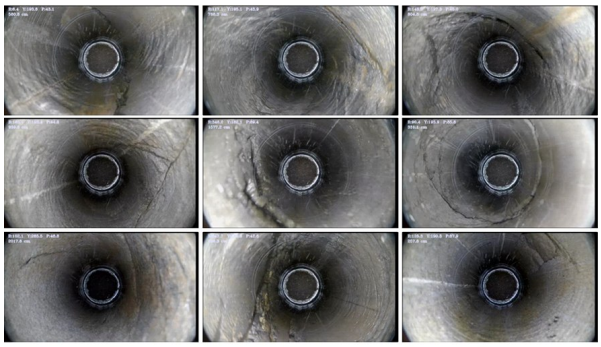


图 5 钻孔摄像视频截图

Fig.5 Borehole camera video screenshot

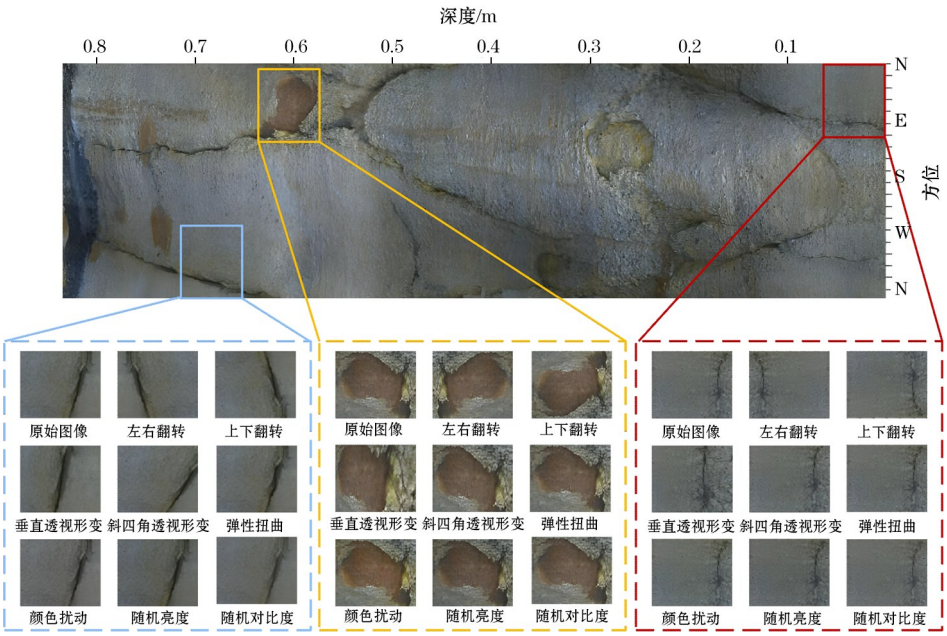


图 6 数据集示例

Fig.6 Dataset example

2.2 试验环境

本文试验均在 Ubuntu20.04 操作系统下展开,基于 GPU 硬件平台和 PyTorch 深度学习框架完成。具体试验环境配置见表 1。

表 1 试验环境配置

Table 1 Test environment configuration

名称	版本
操作系统	Ubuntu20.04
CPU	Intel Xeon Gold 6142 CPU
GPU	NVIDIA GeForce RTX 3090
Python	3.10.12
PyTorch	2.0.1+cu118

为保障模型训练过程中的稳定性,避免优化过程陷入局部最优解,试验采用随机梯度下降(SGD)算法作为模拟优化器。

初始学习率设定为 1×10^{-2},以提升训练初期的学习速度;动量设置为 0.9,用以在梯度方向上加速并抑制震荡;权重衰减系数设置为 5×10^{-4},用于正则化模型,以降低过拟合风险;试验中批次大小设定为 4。训练过程共进行了 300 轮次,以期达到更好的训练效果并减少过拟合风险。

2.3 试验评估指标

在探索和改进语义分割算法的过程中,评价模型的分割精度至关重要。为此,本文采用了多种评估指标来全面且客观地评价模型性能,具体包括平均交并比(J)、平均像素准确率(A)、召回率(R)、精确率(P)、F1 分数(F)。

基于上述指标建立混淆矩阵,用于统计分类模型归对类、归错类的样本个数的分类结果,见表2。

表2 混淆矩阵

Table 2 Confusion matrix

真实情况	预测结果	
	正例	反例
正例	TP(真正例)	FN(假反例)
反例	FP(假正例)	TN(真反例)

交并比(IoU)表示某类真实标签和预测值的交和并的比值,计算公式见式(6),平均交并比(J)为各类别 IoU 的平均值。

$$IoU=\frac{TP}{FP+FN+TP} \tag{6}$$

像素准确率(PA)表示图像预测结果中正确分类占总预测值的比例,计算公式见式(7),平均像素准确率(A)为各类别 PA 的平均值。

$$PA=\frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

精确度(P)表示模型预测为正类的样本中,实际上为正类的比例,计算公式见式(8)。

$$P=\frac{TP}{FP+TP} \tag{8}$$

召回率(R)表示所有实际为正类的样本中,模型能够正确预测为正类的比例,计算公式见式(9)。

$$R=\frac{TP}{FN+TP} \tag{9}$$

F1分数(F)可实现精确度和召回率的平衡考量,适用于类别不平衡的场景,对于每个类别 i,计算公式见式(10)。

$$F=2\times\frac{P_i\times R_i}{P_i+R_i} \tag{10}$$

2.4 损失函数

交叉熵用于度量两个概率分布间的差异性,交叉熵损失函数 L_{CE} 将图像中的全部像素作为一个整体,定义公式见式(11)。

$$L_{CE}=-\frac{1}{N}\sum_{i=1}^{N}[y_i\ln(p_i)+(1-y_i)\ln(1-p_i)] \tag{11}$$

式中:y_i 为第 i 个像素的真实标签,$y_i\in\{0,1\}$,0为背景,1为结构面;p_i 为第 i 个像素的预测概率值,$p_i\in[0,1]$;N 为图像总像素数。

在钻孔围岩结构面分割任务中,结构面目标较小,窥孔图像中结构面与背景像素的数量不平衡。语义分割任务中,L_{Dice} 损失函数常被用于处理正样本和负样本的不平衡分布情况。L_{Dice} 损失函数[19]不易受到正负样本不平衡的影响,在训练过程中更侧重于挖掘前景,但该函数对小目标分割任务不利,易导致梯度变化剧烈。L_{Dice} 损失函数的定义公式见式(12):

$$L_{Dice}=1-\frac{2\sum_{i=1}^{N}y_ip_i}{\sum_{i=1}^{N}y_i+\sum_{i=1}^{N}p_i} \tag{12}$$

针对数据集特征,本文采用 L_{CE} 与 L_{Dice} 相结合的方法作为语义分割损失函数,解决难以收敛和无法平衡正负样本的难题。L_r 为融合后的损失函数,计算公式见式(13):

$$L_r=(1-\lambda)\ln(L_{Dice})+\lambda L_{CE} \tag{13}$$

式中,λ 为交叉熵损失函数权重值。

2.5 不同模型试验结果分析

为揭示各模型在性能方面的差异,并进一步凸显 DISA-FuseNet 与钻孔结构面分割任务的适配性,选取经典图像分割方法 DeepLabV3+[25]、UNet、Trans-UNet[26],以及针对裂隙图像的分割方法 DeepCrack[27]、CrackFormer[28]五种语义分割模型进行对比分析,各模型的性能对比见表3。其中,UNet 模型由于其特有的对称结构和跨层连接,在医学图像分割等领域取得了显著成效。DeepLabV3+通过融合空洞空间金字塔池化(Atrous Spatial Pyramid Pooling,ASPP)与编码器-解码器结构,在处理多尺度特征图像方面具有显著优势。Trans-UNet 应用了针对序列处理的 Transformer 结构,以解决卷积网络内在局部性问题。DeepCrack 结合分层特征学习和多尺度融合结构,在公开图像数据集中实现了较高精度的裂隙分割。CrackFormer 针对裂隙特征设计全局注意力机制,可有效捕获长程依赖关系。

表3 不同分割模型对比试验结果

Table 3 Comparative test results of different segmentation models

模型	J/%	A/%	R/%	P/%	F/%
UNet	79.12	83.14	77.05	71.81	74.34
Trans-UNet	77.01	82.53	78.81	73.82	76.23
DeepLabV3+	81.22	88.14	78.71	77.85	78.28
DeepCrack	82.16	88.73	80.91	73.83	77.21
CrackFormer	81.55	89.30	78.02	76.27	77.14
DISA-FuseNet	87.56	92.65	85.57	82.41	83.96

由表3可知,DISA-FuseNet 模型展现出显著优势,分割精度指标 J、A、R、P、F 分别达到了

87.56%、92.65%、85.57%、82.41%、83.96%,较基准模型 UNet 分别提升了 8.44, 9.51, 8.52, 10.60, 9.62 个百分点。图 7 的分割可视化结果进一步证明了定量分析结论。由图 7(c)、图 7(d)可知,UNet、Trans-UNet 存在显著分割断裂现象,出现了大片漏检区域;由图 7(f)、图 7(g)可知,DeepCrack、CrackFormer 虽在完整性方面有所改善,但仍存在背景误检和局部细节丢失现象;由图 7(e)可知,DeepLabV3+虽误检率降低,但其分割边缘粗糙度较高,难以准确反映结构面形态特征;由图 7(h)可知,DISA-FuseNet 在保持结构面连续性的同时,有效抑制了背景干扰,呈现出更加清晰完整的边缘细节。

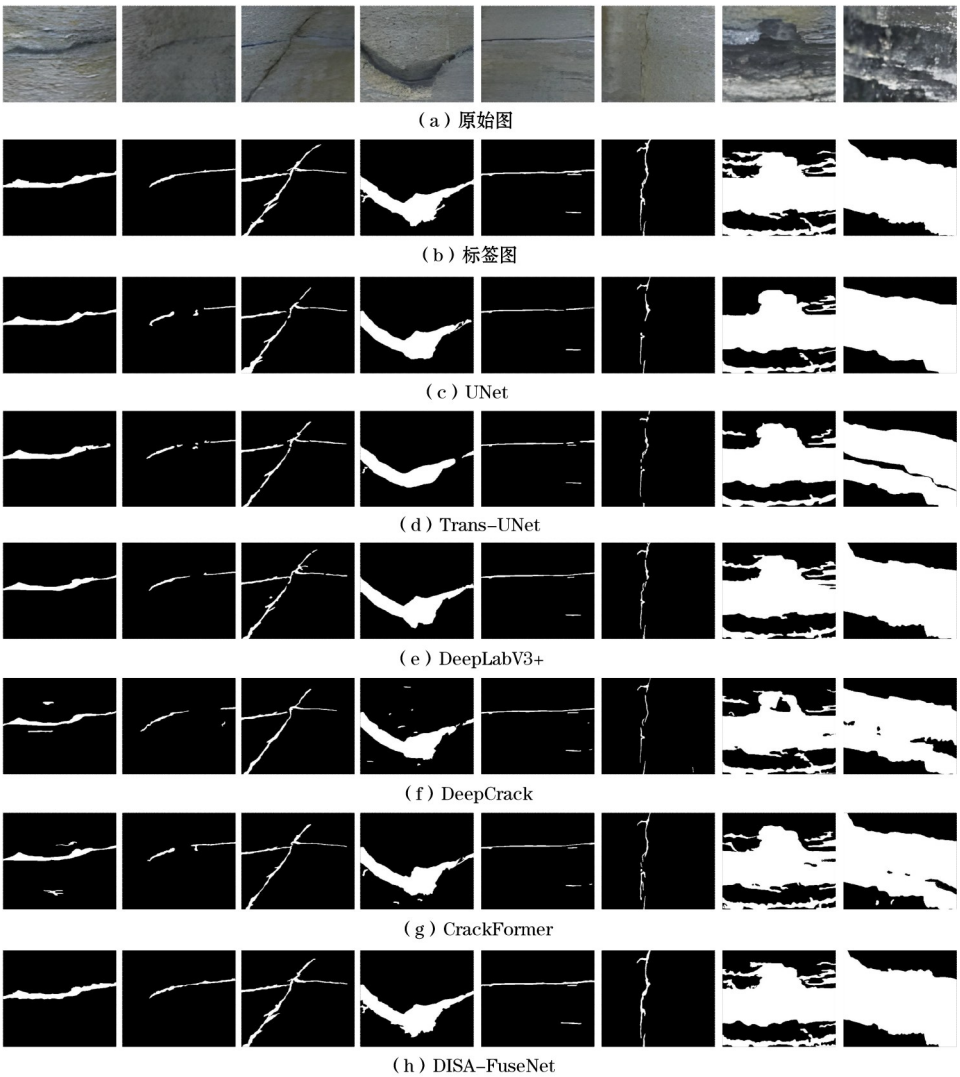


图 7 不同分割模型效果对比可视化

Fig.7 Visualization of effect comparison of different segmentation models

2.6 消融试验

为验证本文改进模型的有效性,本文设置了 3 组消融试验,模型 1(基准模型)采用原始 UNet 架构;模型 2 为双编码器 UNet 架构,引入 Transformer 与 ResNet50 作为双编码器;模型 3 在模型 2 的基础上继续引入统计聚合通道注意力模块(SACA);模型 4 则是在模型 3 的基础上进一步引入自适应层级增强模块(AHE)。4 种模型训练过程中平均交并比(J)的变化曲线如图 8 所示,消融试验结果见表 4,消融试验结果对比可视化见图 9。由表 4 可知,相较于模型 1,模型 2 的 J、A、R、P、F 分别提升了 3.78, 3.92, 3.92, 6.09, 5.04 个百分点,达到了 82.90%、87.06%、80.97%、77.90% 和 79.41%。图 9(d)展示了模型 2 的分割结果,相较于模型 1,模型 2 改善了结构面轮廓的几何连续性,验证了双编码器结构在局部细节保持与全局上下文建模方面的优势。

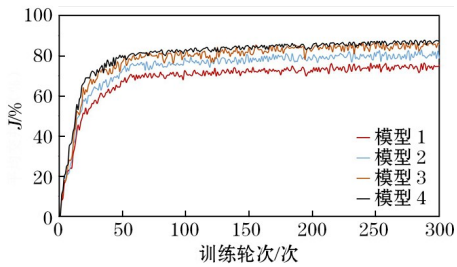


图 8 4 种模型的平均交并比变化曲线

Fig.8 The change curves of the average intersection ratio of the four models

模型 3 在模型 2 的基础上进一步引入了统计聚合通道注意力模块,使 J、A、R、P 和 F 分别提升了 3.30, 4.61, 2.18, 2.95 和 2.57 个百分点,达到了 86.20%、91.67%、83.15%、80.85%和 81.98%。由图 9(e)可知,模型 3 在裂隙边界的完整性和清晰度上有所改善,漏检现象明显减少,表明 SACA 在维护边缘连续性和起伏细腻度方面发挥了重要作用。模型 4 在模型 3 的基础上进一步引入了 AHE,使 J 和 A 分别提升了 1.36, 0.98 个百分点,达到了87.56%和 92.65%。R、P 和 F 也分别提升至85.57%、82.41%和 83.96%。由图 9(f)可知,模型 4 的分割结果与标签图最为相近,前景断裂问题得以缓解,进一步证明了 AHE 对结构面多尺度特征的捕捉能力。

表 4 消融试验结果

Table 4 Ablation test results

模型	双编码器	SACA	AHE	J/%	A/%	R%	P/%	F/%
1	×	×	×	79.12	83.14	77.05	71.81	74.37
2	√	×	×	82.90	87.06	80.97	77.90	79.41
3	√	√	×	86.20	91.67	83.15	80.85	81.98
4	√	√	√	87.56	92.65	85.57	82.41	83.96

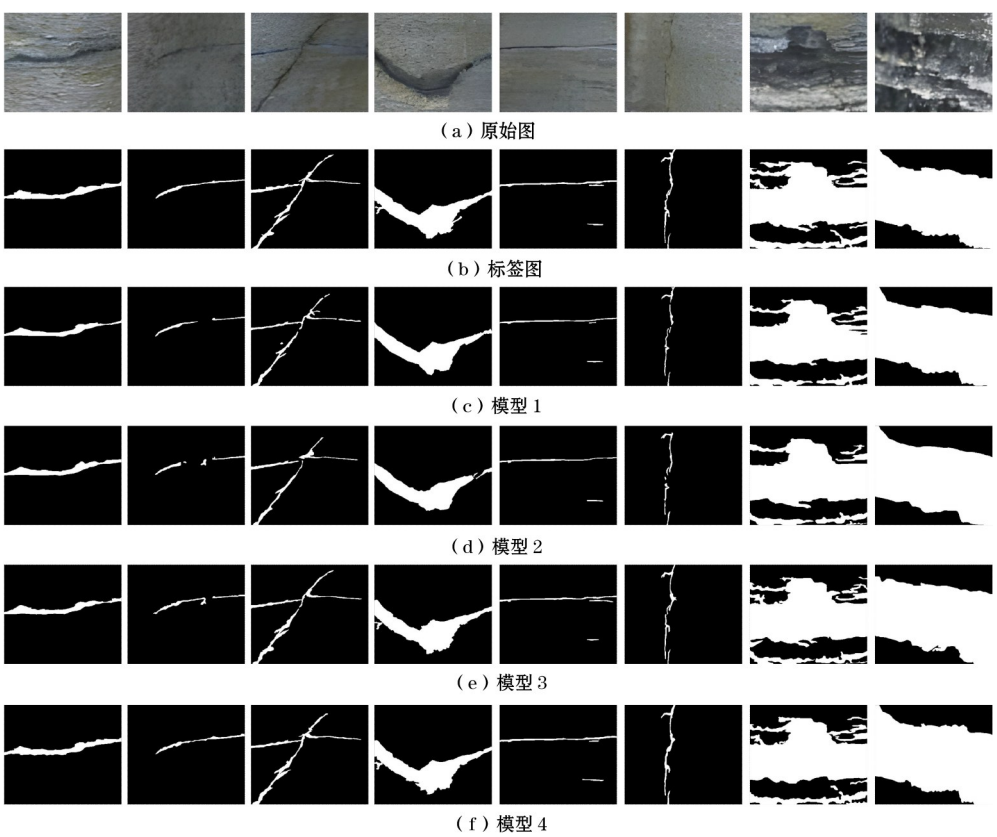


图 9 消融试验结果对比可视化

Fig.9 Visualization of comparison of ablation test results

3 结构面参数提取与可视化分析

3.1 结构面表征

3.1.1 长度

沿结构面骨架线对相邻节点采用欧氏距离 d,递进累加计算总长度:

$$d_{k,k+1}=\sqrt{(x_{k+1}-x_k)^2+(y_{k+1}-y_k)^2} \quad (14)$$

式中,(x_k,y_k) 和 (x_{k+1},y_{k+1}) 为骨架上相邻的两个像素点坐标。

将所有相邻像素点的距离累加,得到裂隙总长

度 L:

$$L=\sum_{k=1}^{N-1}d_{k,k+1} \tag{15}$$

式中,N 为骨架像素点的总数。

3.1.2 宽度

针对骨架线上各点 $P_k=(x_k,y_k)$,先确定该点的局部切线方向,再推导与其垂直的法线方向。

使用骨架中心线的邻近点 P_{k-1} 和 P_{k+1} 计算切线方向角 θ_k:

$$\theta_k=\arctan2(y_{k+1}-y_{k-1},x_{k+1}-x_{k-1}) \tag{16}$$

法线方向角 ϕ_k 计算公式为:

$$\phi_k=\theta_k+\frac{\pi}{2} \tag{17}$$

沿每个骨架点 P_k 的局部法线方向 ϕ_k 发射射线,检测射线与结构面边界的两个交点 $Q_k^{(1)}$ 和 $Q_k^{(2)}$。

射线可表示为:

$$Q_k(t)=P_k+t\cdot(\cos\varphi_k,\sin\varphi_k) \tag{18}$$

式中,t 为射线参数,$t\geqslant 0$。

使用直线扫描方法,沿射线方向逐步增加 t,直到射线首次与结构面边界相交,记录交点 $Q_k^{(1)}$,沿相反方向 $-\phi_k$ 重复此过程,记录交点 $Q_k^{(2)}$。在每个骨架点 P_k 处,通过计算交点 $Q_k^{(1)}$ 和 $Q_k^{(2)}$ 之间的距离,得到裂隙宽度 W_k:

$$W_k=\|Q_k^{(1)}-Q_k^{(2)}\| \tag{19}$$

式中,$\|\cdot\|$ 表示欧氏距离。

则裂隙的平均宽度 $\overline{W}$ 为:

$$\overline{W}=\frac{1}{N}\sum_{k=1}^{N}W_k \tag{20}$$

3.1.3 粗糙度

提取裂隙边界曲线,对相邻点的方向角差值进行标准差运算:

$$C=\mathrm{std}(\psi_{i+1}-\psi_i) \tag{21}$$

式中:C 为裂隙的粗糙度值;std( )为标准差函数;ψ_i 为边界曲线上第 i 个点的方向角。

裂隙边界曲线起伏越剧烈,表明粗糙度值越大。

3.1.4 空间参数

设结构面的二维曲线见式(22):

$$y=A\sin(\omega x+\varphi)+K \tag{22}$$

式中:(x,y) 为结构面二维曲线的坐标;A 为正弦曲线的振幅;ω 为角速度;φ 为相位偏移;K 为基线偏移量。通过正弦曲线的泰勒展开式对结构面二维曲线进行拟合,求解得到上述参数[29]。

设钻孔半径为 r,图像纵向比例为 S,则结构面的单位法向量 $\boldsymbol{N}_V=(N_{Vx},N_{Vy},N_{Vz})$ 为:

$$\begin{cases}N_{Vx}=\dfrac{AS\sin(\varphi)}{\sqrt{r^2+(AS)^2}}\\ N_{Vy}=\dfrac{AS\cos(\varphi)}{\sqrt{r^2+(AS)^2}}\\ N_{Vz}=\dfrac{r}{\sqrt{r^2+(AS)^2}}\end{cases} \tag{23}$$

倾向 α 由 $\boldsymbol{N}_V$ 投影在水平面上的向量(N_{Vx},N_{Vy})确定,经象限修正后为:

$$\alpha=\begin{cases}\arctan(\dfrac{N_{Vy}}{N_{Vx}}),N_{Vx}>0,N_{Vy}\geqslant 0\\ \arctan(\dfrac{N_{Vy}}{N_{Vx}})+180°,N_{Vx}<0\\ \arctan(\dfrac{N_{Vy}}{N_{Vx}})+360°,N_{Vx}>0,N_{Vy}<0\end{cases} \tag{24}$$

倾角 β 为结构面与水平面之间夹角,可由式(25)求得:

$$\beta=\arctan\left(\frac{\sqrt{N_{Vx}^2+N_{Vy}^2}}{N_{Vz}}\right) \tag{25}$$

3.2 三维可视化

3.2.1 三维映射

为在三维环境下展示结构面分布,建立二维展开图像坐标$(x_{\mathrm{img}},y_{\mathrm{img}})$到三维钻孔圆柱体$(X,Y,Z)$的几何映射关系:

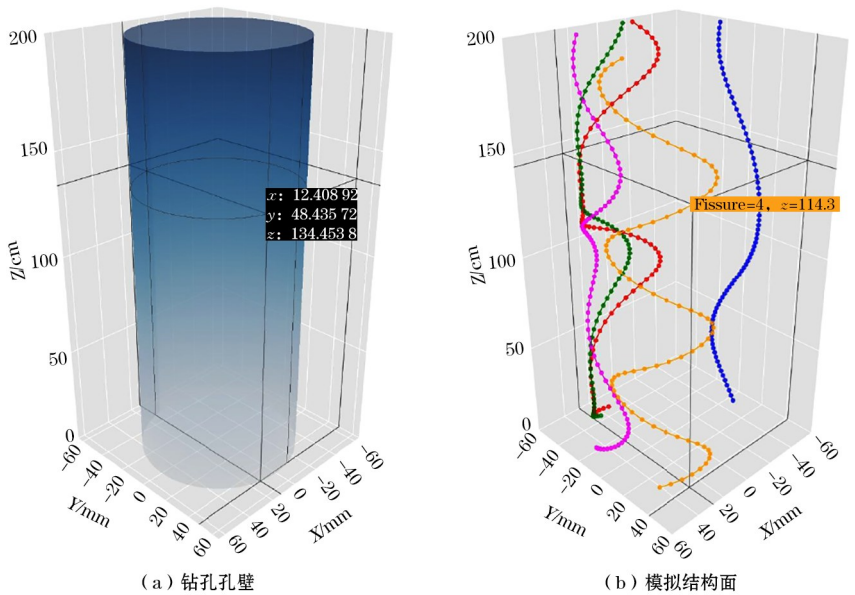
$$\begin{cases}X=R\cos\left(\dfrac{2\pi x_{\mathrm{img}}}{W}\right)\\ Y=R\sin\left(\dfrac{2\pi x_{\mathrm{img}}}{W}\right)\\ Z=\dfrac{y_{\mathrm{img}}}{S}\end{cases} \tag{26}$$

式中,W 为图像宽度。对裂隙像素进行逐点映射可在圆柱壁面重建裂隙形态。

3.2.2 可视化流程与示例

利用网格生成一组圆柱方位角 $\theta\in[0,2\pi]$ 与钻孔深度 $H'\in[0,H_{\max}]$,其中 $H_{\max}$ 为图像最大钻孔深度,基于该网络构建三维表面(X,Y,Z),完成钻孔孔壁的三维可视化,如图 10 所示。

逐像素提取边界后进行坐标映射,得到其在三维圆柱坐标系下的坐标分量$(x_{\mathrm{cyl}},y_{\mathrm{cyl}},z_{\mathrm{cyl}})$,再通过 Scatter3d 工具绘制该结构面的三维点线轨迹(见图 10(b)),当鼠标悬停至轨迹任意位置时,可实时显示该位置对应的结构面参数,包括长度、宽度、粗糙度、倾向 α 和倾角 β。



(a) 钻孔孔壁 (b) 模拟结构面

图 10 孔壁与模拟结构面可视化

Fig.10 Visualization of hole wall and simulated structural plane

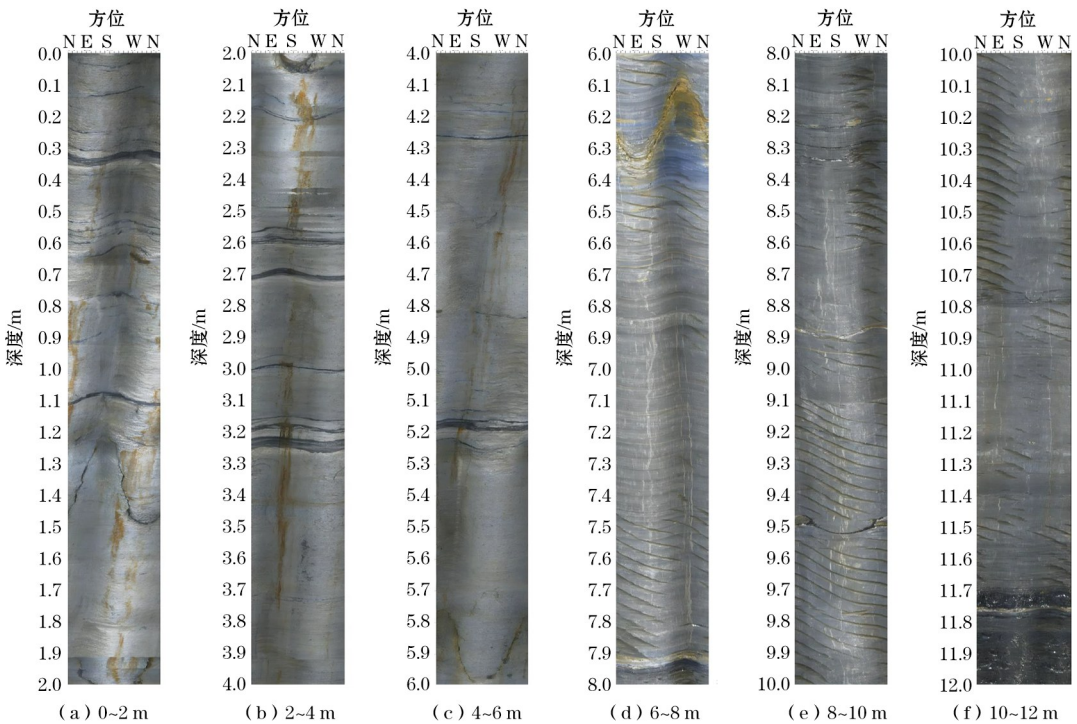
4 实例应用

4.1 工程概况

以山西某矿 13101 工作面为工程背景开展钻孔结构面表征工作。该工作面切眼煤厚 5.5～6.5 m，平均厚度为 6.0 m，煤层内生裂隙、外生裂隙发育，煤层结构复杂，含多层夹矸，夹矸为厚 0.1～0.8 m 的泥岩或砂质泥岩。煤层上方自下而上分别为 7.83 m 砂质泥岩、3.24 m 泥岩、0.48 m 细粒砂岩、1.94 m 砂质泥岩、0.96 m 中粒砂岩、4.06 m 砂质泥岩。

4.2 结构面识别与表征

图 11 为采用钻孔成像仪获得的部分结构面数据，图 12 为经 DISA-FuseNet 识别输出的二值掩码。



(a) 0~2 m (b) 2~4 m (c) 4~6 m (d) 6~8 m (e) 8~10 m (f) 10~12 m

图 11 钻孔展开柱状图

Fig.11 Borehole expansion histograms

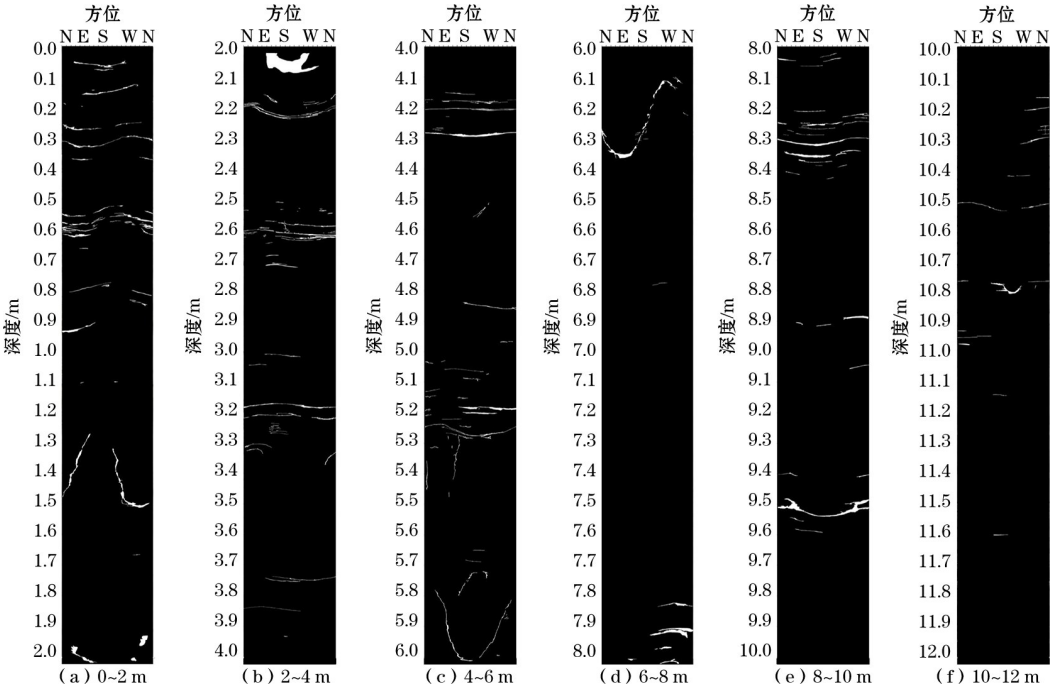


图 12　钻孔图像分割结果

Fig.12　Segmentation results of borehole images

对照原始展开图像与分割结果,二值图像中结构面边缘清晰、结构完整,表明 DIME-FuseNet 在钻孔图像特有的复杂背景中能够适应结构面多变的形态,具有抵抗噪声和高对比度纹理的能力。根据分割结果进行三维钻孔立体图像转换,以便更清晰、直观地观察岩体结构面形态(截取两段长 1.0 m 的钻孔图像,1# 段和 2# 段),如图 13 所示。表 5 为钻孔图像中提取的部分结构面参数。

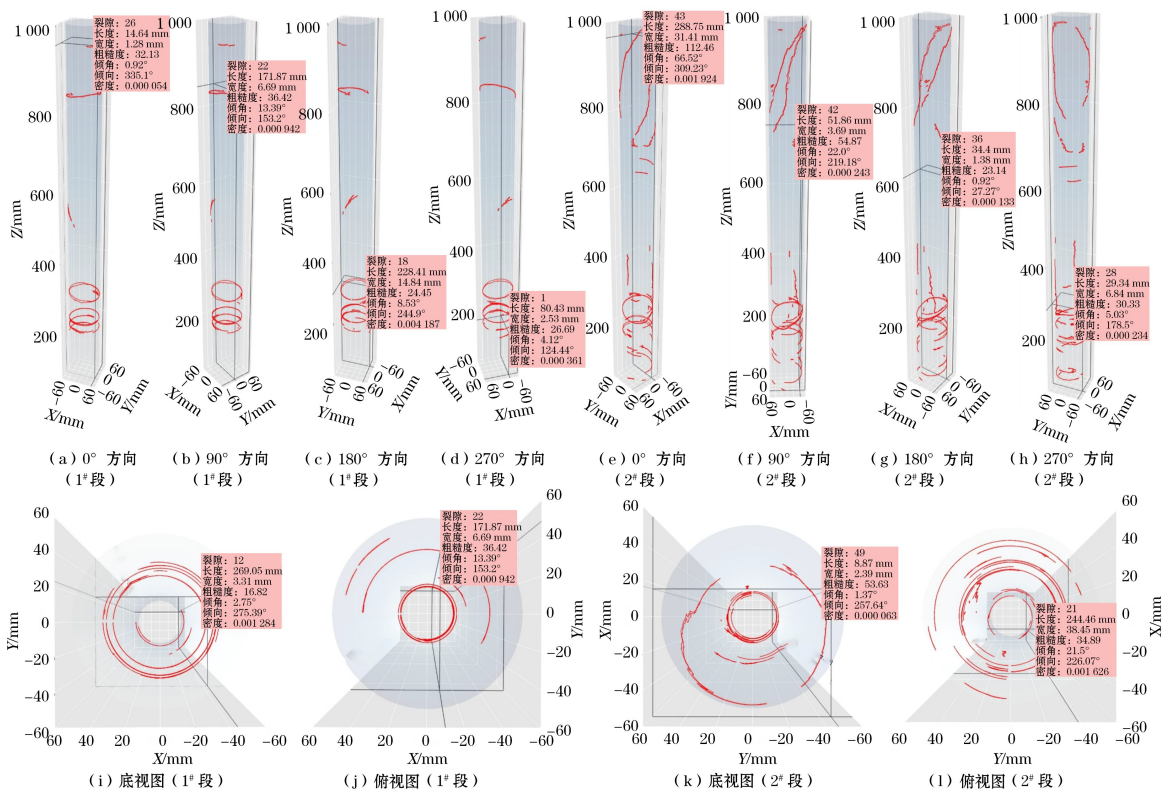


图 13　钻孔结构面三维展示

Fig.13　Three-dimensional display of borehole structural plane

表 5 钻孔图像中部分结构面参数

Table 5 Part structural plane parameters in borehole images

结构面编号	长度/mm	宽度/mm	粗糙度	倾角/(°)	倾向/(°)	密度
1	13.11	1.27	36.22	2.52	236.64	0.000 046
2	24.06	2.56	42.84	3.55	252.68	0.000 134
3	127.06	4.77	32.57	5.94	112.35	0.000 303
4	8.27	1.77	43.28	1.95	295.84	0.000 046
5	55.73	5.63	51.56	6.50	359.61	0.000 269
6	31.78	4.10	48.32	3.43	221.59	0.000 129
7	2.13	0.70	58.07	0.34	149.47	0.000 005
8	5.55	1.04	47.77	0.69	345.58	0.000 017
9	18.78	1.39	32.05	1.37	230.50	0.000 072
10	170.23	7.94	30.02	7.18	189.09	0.002 512
11	4.36	1.51	48.97	1.37	307.77	0.000 021
12	14.23	1.24	37.51	2.75	317.64	0.000 051
13	8.78	1.26	39.25	2.18	339.33	0.000 027
14	92.11	4.18	25.46	3.78	319.82	0.001 105
15	11.7	1.24	36.22	2.52	236.64	0.000 044
16	17.72	1.13	34.86	2.52	43.04	0.000 047
17	58.29	1.21	20.85	1.03	200.82	0.000 176
18	244.46	38.45	34.89	21.50	226.07	0.001 626
19	34.73	3.39	37.73	3.32	316.86	0.000 169
20	8.02	1.74	37.23	1.15	120.42	0.000 041
21	42.28	4.66	51.97	6.73	16.36	0.000 255
22	53.83	5.58	78.67	25.73	252.37	0.000 186
23	11.98	1.30	122.68	6.39	160.01	0.000 044
24	29.34	6.84	30.33	5.03	178.50	0.000 234
25	24.11	4.28	32.69	2.75	196.13	0.000 119

5 结论

本文针对钻孔围岩结构面分割任务中面临的光照不均,易造成图像细节模糊;岩体背景复杂,高对比度纹理易与结构面混淆;结构面深度不一导致其形态各异、宽度与对比度差别显著等问题,提出了一种双编码器交互式注意力增强融合网络(DIME-FuseNet)。基于自建钻孔结构面数据集开展试验验证,与 UNet、Deeplabv3+、Trans-UNet、DeepCrack 及 CrackFormer 五种模型进行对比。最后提出了一种结构面参数提取与三维可视化方法。得出以下主要研究结论。

(1) 基于编码器-解码器结构对骨干网络进行针对性设计,采用 Swin Transformer 和 ResNet50 双编码器结构同时捕捉全局上下文信息与局部细节特征;添加统计聚合通道注意力模块(SACA)进行全局平均池化、全局最大池化及全局中值池化以获取通道注意力,抑制岩体背景噪声和无关信息;设计自适应层级增强模块(AHE),融合并行多空洞率卷积与通道-空间注意力机制,针对结构面在宽度与对比度上的显著差异实现感受野的灵活适配,进而提升模型对多尺度结构面的精准分割性能。

(2) 不同分割模型对比试验结果,DISA-FuseNet 模型展现出显著优势,分割精度指标平均交并比、平均像素准确率、召回率、精确率、F1 分数分别达到了 87.56%、92.65%、85.57%、82.41%、83.96%,较基准模型 UNet 分别提升了 8.44,9.51,8.52,10.60,9.62 个百分点,在保持结构面连续性的同时,有效抑制了背景干扰,呈现出更加清晰完整的边缘细节,消融试验进一步验证了 DISA-FuseNet 模型的有效性。

(3) 提出并实现了一种结构面参数提取与三维可视化方法,实现了对结构面长度、宽度、粗糙度、密度及其倾向倾角参数的精确提取。利用 Plotly 等工具生成交互式的三维可视化图像,对照原始展开图像与分割结果,二值图像中结构面边缘清晰、结构完整,表明 DIME-FuseNet 模型在钻孔图像特有的复杂背景中能够适应结构面多变的形态,具有抵抗噪声和高对比度纹理的能力。研究结果可为后续岩体力学特征分析和工程设计提供了技术及可视化支持。

参考文献(References):

[1] 周维垣.岩体工程结构的稳定性[J].岩石力学与工程学报,2010,29(9):1729-1753.
ZHOU Weiyuan. Structure stability of rock mass engineering[J]. Chinese Journal of Rock Mechanics and Engineering,2010,29(9):1729-1753.

[2] 李杰林,白德威,杨承业,等.基于三维激光扫描点云数据的地下巷道岩体结构面识别及稳定性分析[J].黄金科学技术,2022,30(3):343-351.
LI Jielin, BAI Dewei, YANG Chengye, et al. Recognition and stability analysis of underground tunnel rock mass structural plane based on 3D laser scanning point cloud data [J]. Gold Science and Technology,2022,30(3):343-351.

[3] 康红普,司林坡.深部矿区煤岩体强度测试与分析[J].岩石力学与工程学报,2009,28(7):1312-1320.
KANG Hongpu, SI Linpo. Test and analysis of coal and rock mass strength in deep coal mine[J]. Chinese Journal of Rock Mechanics and Engineering,2009,28(7):1312-1320.

[4] WU Faquan, WU Jie, BAO Han, et al. Rapid intelligent evaluation method and technology for determining engineering rock mass quality[J]. Rock Mechanics Bulletin, 2023,2(2):100038.

[5] 季惠彬,宋琨,吕坤.地下水封洞库深部岩体结构面连通性研究[J].地质科技情报,2013,32(1):176-180.
JI Huibin, SONG Kun, LYU Kun. Connectivity of discontinuities of deep rock mass in water sealed underground storage cavern [J]. Bulletin of Geological Science and Technology,2013,32(1):176-180.

[6] 葛云峰,钟鹏,唐辉明,等.基于钻孔图像的岩体结构面几何信息智能测量[J].岩土力学,2019,40(11):4467-4476.
GE Yunfeng, ZHONG Peng, TANG Huiming, et al. Intelligent measurement on geometric information of rock discontinuities based on borehole image[J]. Rock and Soil Mechanics,2019,40(11):4467-4476.

[7] YANG Sen, LI Hongru, MA Li, et al. An automatic method for discontinuity recognition in coal-measure strata borehole images[J]. IEEE Access,2021,9:105072-105081.

[8] LI Li, YU Cui, HAN Zengqiang, et al. Automatic identification of the rock-soil interface and solution fissures from optical borehole images based on color features[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2019,12(10):3862-3873.

[9] 李清波,杜朋召.基于边缘阈值分割的钻孔图像 RQD 自动分析方法研究[J].岩土工程学报,2020,42(11):2153-2160.
LI Qingbo, DU Pengzhao. Automatic RQD analysis method based on information recognition of borehole images[J]. Chinese Journal of Geotechnical Engineering,2020,42(11):2153-2160.

[10] 余潇智,杨静,李兴超,等.基于卷积神经网络的目标检测算法综述[J].电子技术与软件工程,2022(24):171-176.
YU Xiaozhi, YANG Jing, LI Xingchao, et al. A review of object detection algorithms based on convolutional neural networks [J]. Electronic Technology and Software Engineering,2022(24):171-176.

[11] 刘宇晶.基于 Transformer 的目标检测研究综述[J].计算机时代,2023(5):6-10.
LIU Yujing. Summary of research on target detection based on Transformer[J]. Computer Era,2023(5):6-10.

[12] 彭斌,白静,李文静,等.面向图像分类的视觉 Transformer 研究进展[J].计算机科学与探索,2024,18(2):320-344.
PENG Bin, BAI Jing, LI Wenjing, et al. Survey on visual transformer for image classification[J]. Journal of Frontiers of Computer Science and Technology,2024,18(2):320-344.

[13] JANNAT F, WILLIS A. Improving classification of remotely sensed images with the swin transformer[C]// SoutheastCon 2022,New York:IEEE,2022:611-618.

[14] 吴玉超,林岚,王婧璇,等.基于卷积神经网络的语义分割在医学图像中的应用[J].生物医学工程学杂志,2020,37(3):533-540.
WU Yuchao, LIN Lan, WANG Jingxuan, et al. Application of semantic segmentation based on convolutional neural network in medical images [J]. Journal of Biomedical Engineering,2020,37(3):533-540.

[15] KIYA H, NAGAMORI T, IMAIZUMI S, et al. Privacy-preserving semantic segmentation using vision transformer [J]. Journal of Imaging,2022,8(9):233-247.

[16] DIAS L O, BOM C R, FARIA E L, et al. Automatic detection of fractures and breakouts patterns in acoustic borehole image logs using fast-region convolutional neural networks[J]. Journal of Petroleum Science and Engineering 2020,191:107099.

[17] LIU Cancan, KIM Jineon, SONG Jaejoon, et al. Intelligent recognition and identification of fracture types and parameters for borehole images based on developed convolutional neural networks and post-processing [J]. Engineering Fracture Mechanics,2023,292:109624.

[18] 张野,陈金桥,李炎隆.基于多深度模型的钻孔结构面智能识别与量化分析[J].地质科技通报,2023,42(6):31-41.
ZHANG Ye, CHEN Jinqiao, LI Yanlong. Intelligent recognition and quantitative analysis of borehole hydraulic geological images utilizing multiple deep learning models[J]. Bulletin of Geological Science and Technology,2023,42(6):31-41.

[19] ZHANG Ye, CHEN Jingqiao, LI Yanlong. Segmentation and quantitative analysis of geological fracture: a deep transfer learning approach based on borehole televiewer image[J]. Arabian Journal of Geosciences,2022,15(3):300.

[20] 李炜,刘耕,葛云峰,等.基于深度学习的钻孔图像岩体结构面识别[J].应用基础与工程科学学报,2024,32(3):702-720.
LI Wei, LIU Geng, GE Yunfeng, et al. Recognition of rock mass structural planes in borehole images based on deep learning[J]. Journal of Basic Science and Engineering,2024,32(3):702-720.

[21] 吴金,吴顺川,王焘,等.基于深度学习的复杂地层钻孔图像岩体结构面识别方法[J].清华大学学报(自然科学版),2024,64(7):1136-1146.
WU Jin, WU Shunchuan, WANG Tao, et al. Recognition method of rock mass structural planes in complex stratigraphic borehole images based on deep learning[J]. Journal of Tsinghua University (Science and Technology),2024,64(7):1136-1146.

[22] LIU Ze, LIN Yutong, CAO Yue, et al. Swin Transformer: hierarchical vision transformer using shifted windows[J]. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV),2021:10012-10022.

[23] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[J]. CORR, 2015, abs/1512.03385.

[24] 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. Heidelberg:Springer,2015:234-241.

[25] CHEN Liangchieh, ZHU Yukun, PAPANDREOU George, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the European conference on computer vision (ECCV 2018). Cham: Springer, 2018:801-818.

[26] CHEN Jieneng, LU Yongyi, YU Qihang, et al. TransUNet: transformers make strong encoders for medical image segmentation[J]. Computer Science,2021,arXiv:2102,04306.

[27] LIU Yahui, YAO Jian, LU Xiaohu, et al. DeepCrack: a deep hierarchical feature learning architecture for crack segmentation[J]. Neurocomputing,2019,338:139-153.

[28] LIU Huajun, MIAO Xiangyu, MERTZ Christoph, et al. CrackFormer: transformer network for fine-grained crack detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press,2021:3783-3792

[29] 张占旭,苏俊辉,吕光祖,等.数字钻孔图像岩体结构面自动化识别方法[J].重庆大学学报,2024,47(2):40-50.
ZHANG Zhanxu, SU Junhui, LYU Guangzu, et al. Automatic identification of rock structure surface based on digital borehole images [J]. Journal of Chongqing University,2024,47(2):40-50.