ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 41,No. 1,Jan. 2026,pp. 160－173
DOI:10. 16337/j. 1004-9037. 2026. 01. 011

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

基于改进GLU-Net的岩石薄片显微图像拼接

向文亮[1]，熊淑华[1]，何海波[2]，滕奇志[1]，何小海[1]

(1. 四川大学电子信息学院，成都 610065;2. 成都西图科技有限公司，成都 610024)

摘　要：岩石薄片显微图像常常呈现出局部纹理复杂、模糊以及高噪声的特征，导致传统的特征提取和匹配算法在高分辨率岩石薄片显微图像拼接中容易出现找不到有效特征点而难以实现全景拼接的问题，并且处理速度较慢。针对上述问题，提出了一种基于改进GLU-Net的岩石薄片显微图像拼接方法。该方法通过结合改进的相关性计算模块增强全局与局部对应关系，使用特征金字塔网络实现多尺度特征融合，设计添加了自适应卷积注意力机制优化关键区域注意力，并使用全局与局部解码器获取光流，最后对图像进行单应性变换实现拼接，构建了一种新的图像拼接网络模型。实验结果表明，与传统图像拼接算法和其他经典图像拼接网络模型相比，提出的网络表现出更好的拼接效果，在自制数据集拼接测试中的拼接准确率达到了86.75%，每组平均配准耗时为0.394 s，在提高拼接准确率的同时有效平衡了处理效率。

关键词：岩石薄片图像；特征融合；卷积注意力机制；光流估计；图像拼接

中图分类号：TN391　　**文献标志码：**A

引用格式：向文亮，熊淑华，何海波，等. 基于改进GLU-Net的岩石薄片显微图像拼接[J]. 数据采集与处理，2026，41(1)：160-173. XIANG Wenliang，XIONG Shuhua，HE Haibo，et al. Image stitching of rock thin sections microscopic images based on improved GLU-Net[J]. Journal of Data Acquisition and Processing，2026，41(1)：160-173.

引　　言

在地质学研究中，岩石矿物的偏光特性因晶体结构差异而在各类光环境中呈现出独特的光学表现[1]。地质学家传统上使用偏光显微镜[2-3]对制备的岩石薄片[4-6]进行显微图像[7]采集，这些图像在矿物微观结构研究[8]、矿物颗粒分类识别[9-11]中发挥了重要作用。然而，利用偏光显微镜在不同偏光模式和倍率下的成像，获取的单视域图像虽能揭示局部微观细节，但其局限性在于无法展示岩石薄片的整体特性。大视域成像虽能呈现岩石薄片的全局信息，却难以捕捉细微结构特征。因此，需要将多张单视域图像拼接成全幅面高分辨率图像，以实现岩石薄片宏观特征与微观细节的有机统一与综合分析。

图像配准[12]是图像拼接的关键环节。传统配准方法主要基于特征提取[13]和特征匹配[14]算法。SIFT算法对尺度、旋转和仿射变换以及光照、噪声具有高度鲁棒性，但需要处理的数据量较大，配准速度较慢。针对此问题，Wu等[15]对SIFT算法进行了优化，跳过初始尺度空间构建步骤，结合GLOH描述符增强鲁棒性，并引入基于曼哈顿距离的特征筛选及快速样本一致性算法进行误匹配剔除，该方法在多个桉树林航拍数据集上显著提升了匹配精度与运行效率。陈丽娟等[16]针对太空遥感图像特征稀疏的问题，结合Hessian矩阵与尺度空间提取匹配点，利用Harris角点构建特征描述子，并通过RANSAC筛选匹配对，估算单应矩阵实现图像拼接。Deb等[17]针对图像特征匹配中存在的匹配精度不足与计算资源受限等问题，基于AKAZE特征检测框架提出改进方法，将FREAK描述子与优化的KNN-RANSAC匹配策略相结合，实现了在尺度变换、模糊及旋转等干扰下的稳健匹配。

基金项目：国家自然科学基金(62071315)；四川省国际科技创新合作项目(2024YFHZ0289)。
收稿日期：2025-07-18；**修订日期：**2025-10-10

近年来,随着深度学习技术的发展,基于深度学习的图像拼接网络模型逐渐涌现。Sarlin等[18]在SuperPoint等局部特征基础上,引入图神经网络与自注意力机制,通过对源图像与目标图像的特征图进行图结构建模和信息交互,实现端到端关键点匹配,该方法在匹配精度、稳定性以及遮挡和视角变化适应性方面均显著优于传统算法。Truong等[19]提出的GLU-Net网络架构,构建了从低分辨到高分辨的多级稠密匹配框架,具备较高的匹配准确率与鲁棒性,适用于图像拼接、位姿估计和三维重建等多类场景。Sun等[20]提出了LoFTR网络,该方法在不使用显式关键点的前提下,借助Transformer架构直接在密集特征图上建立图像之间的空间对应关系,采用由粗到细的匹配机制,实现了较高精度和鲁棒性的图像匹配。Truong等[21]提出的PDC-Net+在密集光流估计中融合概率建模与置信度预测,提升了匹配精度与鲁棒性,并具备不确定性估计能力。Lindenberger等[22]提出的LightGlue网络,在SuperGlue架构基础上引入了轻量注意力模块与自适应推理机制,可根据图像对匹配难度动态调节推理深度与宽度,从而实现高精度、高效率的局部特征匹配。Potge等[23]提出了XFeat网络,该方法在保持匹配精度的同时,显著降低计算开销,通过极简CNN架构实现关键点检测、局部描述与匹配细化,并支持稀疏与半稠密两种模式。

与自然场景图像拼接不同,岩石薄片偏光显微图像具有独特的成像特点,在拼接过程中面临以下挑战:(1)岩石矿物纹理结构复杂且具有高度重复性,易导致局部特征匹配错误;(2)大量弱纹理或模糊区域缺乏明显的角点与边缘,导致提取的特征点稀疏且不稳定;(3)不同区域整体外观高度相似,易造成全局误匹配;(4)高倍显微镜采集产生的大规模高分辨率图像序列对拼接速度提出了高要求。因此,通用的图像拼接算法在岩石薄片场景下往往表现不佳,需要针对性设计与改进。

本文针对上述挑战,基于深度学习框架提出了一种基于改进GLU-Net的岩石薄片显微图像拼接网络,通过多尺度特征金字塔融合网络提升匹配精度,结合改进的相关性计算模块提升全局与局部匹配能力,引入自适应卷积注意力机制增强关键区域的特征表达,并进行全局与局部解码估计光流,最后对图像进行单应性变换实现拼接。实验验证表明,该方法在配准精度和速度上均表现较好,为岩石薄片显微图像的全景图拼接提供了有效的解决方案。

1 图像采集与数据集制作

1.1 岩石薄片序列图像采集

通过如图1所示的岩石薄片图像采集系统采集岩石薄片显微图像。岩石薄片图像采集系统由带有电动载物台的偏光显微镜、载物台控制器和工业相机组成。

岩石薄片在显微镜下的相机拍摄视域远小于整张薄片的面积,因此通过载物台控制器控制载物台移动来连续采集多张单视域序列图像,直至采集完整张薄片。通过控制载物台按“Z”字形路径沿X和Y方向移动来采集岩石薄片不同区域的显微图像。采集图像时,载物台通过Z方向移动对岩石薄片进行聚焦,以获得清晰的图像。由于需要将这些序列图像拼接成完整薄片图像,因此采集的每个相邻序列图像之间都需要存在一定的重叠区域,从而对相邻图像进行光流估计实现拼接。通过设定载物台沿X和Y方向的移动步长来确定重叠区域比例。

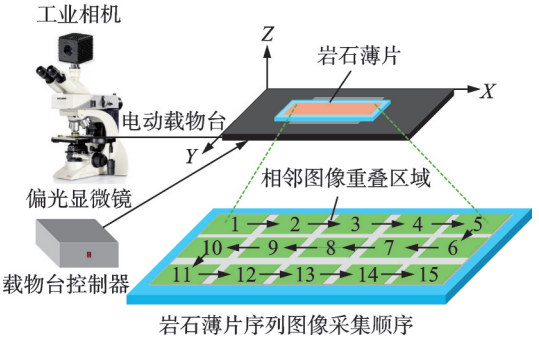


图1 岩石薄片图像采集系统

Fig.1 Rock thin-section image acquisition system

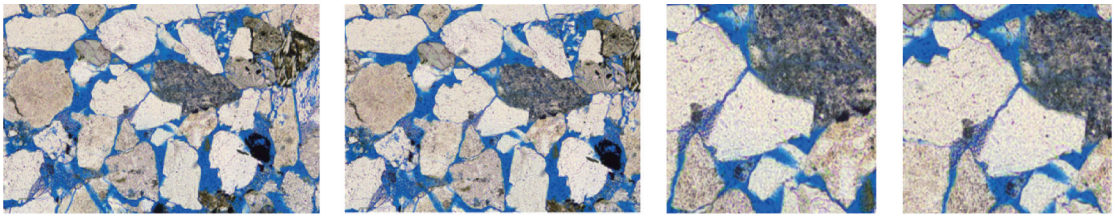
1.2 数据集制作

为了训练本文的深度学习模型,需要制作用于岩石薄片图像拼接的数据集。由于真实光流标签难以通

过人工标注获取,因此采用自监督方法生成光流标签以制作数据集,分为4个步骤:源图像的获取、图像单应性变换、图像裁剪和光流标签的生成。

第一步,源图像的获取。使用岩石薄片图像采集系统对不同岩石薄片进行图像采集,共选取3 000张分辨率为2 736×1 824像素的岩石薄片图像作为原始数据集图像。

第二步,图像单应性变换。为每张源图像随机生成一个3×3单应性矩阵。使用生成的单应性矩阵对源图像进行透视变换,生成对应的目标图像。变换采用双线性插值,确保目标图像分辨率保持2 736×1 824像素,避免边界丢失。图2展示了1张源图像(图2(a))及其单应性变换生成的目标图像(图2(b))。



(a) 源图像 (a) Source image　(b) 目标图像 (b) Target image　(c) 裁剪后的源图像 (c) Cropped source image　(d) 裁剪后的目标图像 (d) Cropped target image

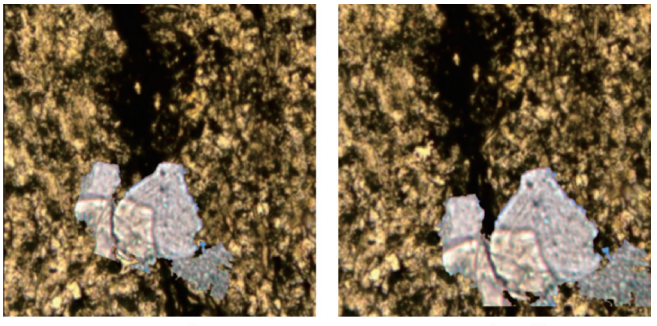
图2 源图像对及其裁剪后的图像对

Fig.2 Source image pair and the corresponding cropped image pair

第三步,图像裁剪。原始数据集图像分辨率较高,将其作为数据集图像数据进行训练运算量过大,为降低内存需求、减小网络训练数据量从而提升训练效率,对源图像和目标图像进行裁剪,减小图像尺寸。为使裁剪区域包含代表性的岩石薄片矿物颗粒和纹理,采用中心区域裁剪的方式,裁剪区域为520×520像素。图2(c)是对图2(a)的裁剪结果,图2(d)是对图2(b)的裁剪结果。

第四步,光流标签的生成。光流是图像序列中像素运动的二维矢量场,表示源图像中每个像素在目标图像中的位移。光流矢量由两个分量组成:x方向偏移和y方向偏移,反映了图像间的空间变换。通过每个源图像和目标图像对应的单应性矩阵计算光流,从而得到光流标签。具体过程是:首先,创建520×520像素的网格坐标(x,y),表示源图像像素位置;然后,应用单应性矩阵变换坐标,得到目标图像像素位置(x',y');最后,计算光流flow$=(x-x',y-y')$,并调整为张量格式。

此外,模型训练时还使用了COCO数据集[24],通过将自制的岩石COCO数据集中的矿物成分目标物融入自制岩石薄片图像,得到最终的数据集,模拟岩石薄片图像中可能出现的杂质或外部干扰,增强模型的泛化能力和推理的鲁棒性。从自制COCO数据集中随机选取图像并缩放为520×520像素分辨率,提取每个图像的边界框和分割掩码,使用掩码对目标物进行像素级分割,生成前景图像。背景区域置为透明。如图3所示,将提取的COCO目标物叠加到岩石薄片数据集图像上,使用α通道进行透明融合,融合公式如下



(a) 源图像 (a) Source image　(b) 目标图像 (b) Target image

图3 训练时融合COCO目标物的自制数据集图像

Fig.3 Custom dataset image with COCO objects fused during training

$$I_{\text{fused}}(x,y)=\alpha\cdot I_{\text{object}}(x,y)+(1-\alpha)I_{\text{back}}(x,y) \tag{1}$$

式中:I_{fused}为合成数据集图像,I_{object}为目标物图像,I_{back}为岩石薄片数据集图像,α为α通道值,范围从0到1。

2 改进GLU-Net的岩石薄片图像拼接网络

基于改进GLU-Net的岩石薄片显微图像拼接网络模型整体结构如图4所示,主要包括6个部分:特征提取模块(Visual geometry group, VGG)、多尺度特征融合模块(Feature pyramid network for multi-scale feature fusion, FPN)、相关性计算模块(Globally-optimized correspondence volume, GOCor)、多尺度卷积注意力机制模块(Convolutional attention mechanism module, CBAM)、解码器与光流估计模块以及单应性变换与拼接模块。

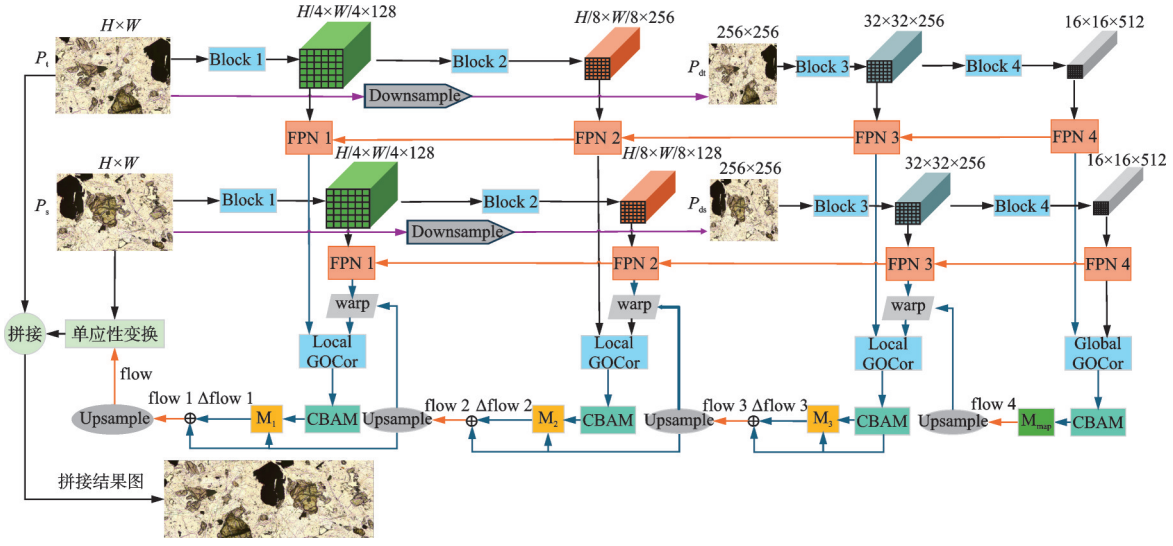


图4 网络整体结构图

Fig.4 Overall network architecture diagram

岩石薄片显微图像拼接的目的是通过光流估计建立两张图像间的精确对应关系以实现对齐。首先,将两张待拼接的岩石薄片显微图像及其降采样后的两张图像依次输入到VGG特征提取模块中,以获得每张图像的多尺度特征图;然后,将这些特征图输入FPN模块进行融合,生成通道数一致的特征金字塔;随后,通过相关性计算模块GOCor计算目标图像和源图像之间的全局和局部相关性,并通过CBAM模块进一步增强特征表示;接着,通过解码器预测多分辨率光流并进行精炼和上采样;最后,得到最终的光流用于匹配点提取,用于对输入图像进行单应性变换实现拼接。改进的GLU-Net网络能够灵活处理不同分辨率的图像对,适用于多种岩石薄片显微图像的拼接任务。

2.1 多尺度特征融合金字塔网络

岩石薄片图像通常呈现出复杂纹理、局部非均匀和高噪声的特征。使用VGG金字塔从图像中提取多尺度特征图,这些特征图能够捕捉从低级纹理到高级语义的多样信息。

VGG金字塔的结构如图5所示。VGG金字塔采用分层架构,每个块由多个3×3卷积层组成,后接ReLU激活函数,并在块末尾设置单一的最大池化层,从而逐步降低空间维度并增加通道深度。通过较小的3×3卷积核替代较大核,降低了模型参数量。输入图像经过VGG特征提取,分别得到1/4分辨率特征图C_1和1/8分辨率特征图C_2,通道数分别是128和256。将输入图像降采样为256×256像素分辨

率后进行特征提取，分别得到32×32像素分辨率特征图C_3和16×16像素分辨率特征图C_4，通道数分别是256和512。

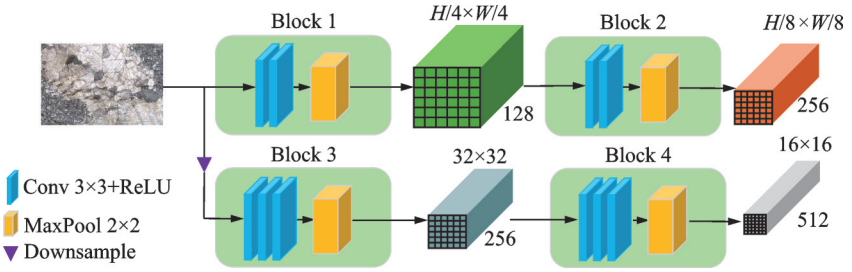


图5 VGG结构图

Fig.5 VGG structure diagram

为了充分整合多尺度信息以及应对偏光显微镜采集图像过程中出现的噪声和光照不均，添加多尺度特征融合模块，对提取的各层特征图进行融合，增强特征的表达能力。多尺度特征融合模块如图6所示。对提取的4层特征图分别使用1×1大小的卷积核进行卷积，将通道数统一调整为256通道。对第4层特征图C_4使用3×3卷积进行平滑得到特征图F_4；对特征图F_4进行双线性插值上采样后与特征图C_3相加进行融合，再使用3×3卷积进行平滑得到特征图F_3，以此类推可得到特征图F_2和F_1。通过自上而下与自下而上的路径结合特征，将高分辨率的空间细节与低分辨率的高级语义进行融合，生成适合多分辨率处理的特征图，从而提升匹配精度。

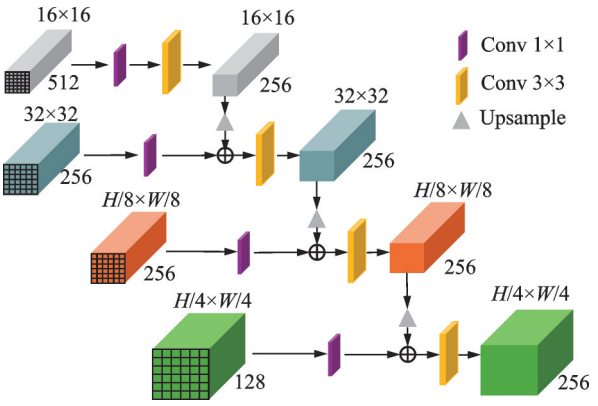


图6 多尺度特征融合模块结构图

Fig.6 Multi-scale feature fusion module structure diagram

2.2 相关性计算模块

得到源图像和目标图像不同分辨率的特征图后，需要计算源图像和目标图像特征之间的相关性，为后续光流估计提供初始对应关系。改进的特征相关性计算模块GOCor引入多目标滤波器预测器，通过结合全局和局部相关性分析，克服传统静态相关性方法的局限性，从而提升匹配精度和鲁棒性。

源图像的特征图F_s和目标图像的特征图F_t的尺寸为$H\times W\times C$，其中$C=256$。首先将特征图F_t输入到滤波器预测器中，计算初始滤波器ω_0，具体公式如下

$$\omega_0(u,v)=\frac{P_{\text{init}}(F_s,F_t)(u,v)}{\|P_{\text{init}}(F_s,F_t)\|_2} \tag{2}$$

式中：P_{init}为初始化子网络，(u,v)为窗口内的空间偏移，$\|\cdot\|_2$表示二维欧几里得距离的L2范数。

对ω_0进行迭代优化，获取优化滤波器ω^*，更新规则为

$$\omega_{t+1}=\omega_t-\eta\cdot\nabla_\omega L(\omega_t,F_s,F_t) \tag{3}$$

式中：η为学习率，∇表示求梯度，t为迭代索引，$L(\omega_t,F_s,F_t)$为损失函数。该损失函数定义为

$$L(\omega_t,F_s,F_t)=-\tau(i,j)+\lambda\|\omega_t\|_2^2 \tag{4}$$

式中:λ为自适应正则化系数;$\tau(i,j)$为初始相关性,定义为

$$\tau(i,j)=\sum_{u,v}\omega_t(u,v)\cdot F_t(i+u,j+v) \tag{5}$$

将优化滤波器ω^*与F_t进行卷积,计算增强的相关性R,生成对应关系体,计算公式如下

$$R(i,j)=\sum_{u,v}\omega^*(u,v)\cdot F_t(i+u,j+v) \tag{6}$$

式中:$R\in\mathbf{R}^{H'\times W'\times(2m+1)\times(2m+1)}$为对应关系体,$m$定义搜索范围。

全局GOCor专注于计算图像间的全局变换,如平移、旋转或尺度变化,基于较高层特征,以捕捉整体上下文信息。使用全局GOCor计算源图像的第4层特征图与目标图像的第4层特征图之间的全局相关性。局部GOCor侧重于局部区域的细粒度匹配,基于较低层特征,以细化局部边界或纹理对齐。使用局部GOCor计算源图像与目标图像第1、2、3层特征图之间的局部相关性。

2.3 自适应卷积注意力模块

岩石薄片显微图像在拼接过程中可能出现岩石颗粒关键区域和无关背景无法区分的情况,因此设计添加了自适应卷积注意力模块(CBAM),通过对通道和空间两个维度施加自适应的注意力权重,增强网络对重要特征的关注,同时抑制不相关的背景噪声。自适应卷积注意力模块结构如图7所示。

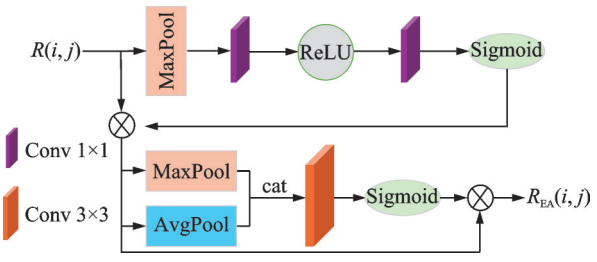


图7 自适应卷积注意力模块结构图

Fig.7 Adaptive convolutional attention module structure diagram

首先对输入的对应关系体$R(i,j)$进行自适应步长全局平均池化得到$R^{1\times1\times B\times C}$,其中$B$为批次大小,$C$为通道数,$H$和$W$分别为高度和宽度。然后通过1×1卷积和ReLU函数激活降维至$C/16$,再使用1×1卷积和Sigmoid函数激活重塑为$R^{1\times1\times B\times C}$,并与$R(i,j)$逐元素相乘得到$R_{SE}(i,j)$。对$R_{SE}(i,j)$进行最大池化得到$R_M(i,j)$,进行平均池化得到$R_A(i,j)$,$R_M(i,j)$和$R_M(i,j)$的尺寸都为$H\times W\times B\times1$,沿通道维度进行拼接得到$R^{H\times W\times B\times2}$;再通过3×3卷积压缩为$R^{H\times W\times B\times1}$,并施加Sigmoid激活后与$R_{SE}(i,j)$逐元素相乘,得到最终的对应关系体$R_{EA}(i,j)$。

2.4 解码器与光流估计

映射解码器由5个卷积块和1个输出层组成,是全局解码器。将第4层特征图的对应关系体R_4输入到映射解码器中,R_4依次通过5个3×3卷积块,通道数分别为128、128、96、64和32,逐层提取特征并应用ReLU激活;最后,将最终的32通道特征映射到2通道,再反归一化并减去网格坐标转换为初始光流flow_4,用于后续多尺度光流估计的起点。上述解码过程为

$$m_4=M_{\text{map}}(R_4) \tag{7}$$

$$\text{flow}_4(\boldsymbol{d})=m_4(\boldsymbol{d})-\boldsymbol{d} \tag{8}$$

式中:M_{map}为映射解码器,m_4为每个像素对应的归一化匹配坐标;$\boldsymbol{d}$为像素在特征图中的归一化坐标,$m_4(\boldsymbol{d})$为该点在另一张特征图中的匹配位置坐标,$\text{flow}_4(\boldsymbol{d})$为输出的光流,表示从当前位置指向其匹配点的偏移。

第1、2、3层特征图使用光流解码器进行解码,光流解码器由5个卷积块和1个光流预测层组成,是局部解码器。使用了动态步长调整策略,即根据前一层光流预测误差自适应调整本层卷积步长,使多

尺度光流估计更加稳定。将对应关系体$R_l(l=3,2,1)$输入到光流解码器中，光流解码器首先接收对应关系体$R_l(l=3,2,1)$和上一层光流，将两者沿通道维度拼接生成初始输入特征；随后，该特征依次通过5个3×3卷积块，每层卷积后与输入特征通过DenseNet连接沿通道维度拼接，累积特征信息并应用LeakyReLU函数激活；最后，将32通道特征映射到2通道，生成光流增量Δflow_l，该增量与上采样的上一层光流Δflow_{l+1}相加得到当前层光流flow_l，用于多尺度光流估计的逐步精炼。上述解码过程为

$$\Delta\text{flow}_l = M_l(R_l, \text{up}(\text{flow}_{l+1})) \quad l=3,2,1 \tag{9}$$

$$\text{flow}_l = \Delta\text{flow}_l + \text{up}(\text{flow}_{l+1}) \quad l=3,2,1 \tag{10}$$

式中：M_l为光流解码器，up(•)表示上采样。

得到的第1层特征图对应的光流flow_1经过上采样后即获得最终预测的光流flow。从光流场中选取具有高置信度的对应点作为匹配点对集合。

2.5 图像单应性变换与拼接

为实现图像的几何对齐，通过单应性变换完成图像的配准与拼接。单应性矩阵$\boldsymbol{H}$为

$$\boldsymbol{H}=\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \tag{11}$$

式中：h_{11}、h_{12}、h_{21}、h_{22}控制旋转、缩放、剪切等线性变换，h_{13}、h_{23}控制平移变换，h_{31}、h_{32}控制透视畸变。

由于岩石薄片显微序列图像是通过载物台水平面移动采集得到，故图像间可以忽略旋转、缩放等变形，只考虑平移变换，从而减小计算量。单应性矩阵可简化为

$$\boldsymbol{H}_1=\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

式中：t_x和t_y分别表示X方向和Y方向的平移量。

为估计单应性矩阵，采用随机抽样一致性算法(Random sample consensus, RANSAC)对其进行优化。从匹配点对集合中随机选取至少4对点，计算初始单应性矩阵。对所有匹配点计算内点，定义内点为满足重投影误差小于阈值(5像素)的点。重复上述过程若干次，选择对应内点最多的单应性矩阵作为最终的单应性矩阵。使用估计的单应性矩阵对图像进行单应性变换实现拼接。

为了保证全薄片拼接图的完整性，对于拼接失败的单视域图像，使用平均位移补偿法进行补偿拼接。平均位移补偿法定义为

$$\begin{cases} \bar{x}=\dfrac{1}{i}\displaystyle\sum_{k=1}^{i-1}x_k & i\in[1,m] \\ \bar{y}=\dfrac{1}{j}\displaystyle\sum_{l=1}^{j-1}y_l & j\in[1,n] \end{cases} \tag{13}$$

式中：$\bar{x}$表示序列图像X方向匹配位移均值，$\bar{y}$表示序列图像Y方向匹配位移均值，i表示当前拼接的序列图像所在列数，j表示当前拼接的序列图像所在行数，m和n分别表示全视域图像包含的单视域图像的行数和列数。如果当前序列图拼接失败，则使用$\bar{x}$和$\bar{y}$作为单应性矩阵中的平移参数对该图像进行拼接。

3 实验结果分析与讨论

3.1 实验环境

本文的模型训练和图像拼接实验在Linux操作系统下进行。服务器的CPU为Intel(R) Core(TM) i7－9700CPU@3.00 GHz×8，内存为32 GB；GPU为NVIDIA GeForce RTX 2080 Ti，显存为11 GB。

3.2 评价指标

为了有效评估不同算法和模型对岩石薄片图像的拼接效果,采用以下评价指标:

(1) 平均端点误差(Average end-point error,AEPE)

AEPE[25]是评估光流估计或图像对应关系质量的一个指标,表示预测点位置与真实点位置之间欧几里得距离的平均值。在密集匹配和光流估计任务中,AEPE反映了整体预测偏差大小,其值越小,光流估计越准确,表明算法(模型)对图像的拼接效果越好;反之亦然。AEPE定义为

$$\mathrm{AEPE}=\frac{1}{N}\sum_{k=1}^{N}\|\boldsymbol{X}_k'-\boldsymbol{X}_k\|_2 \tag{14}$$

式中:N表示参与计算的有效像素数,$\boldsymbol{X}_k=(x_k,y_k)$表示目标图像中的第$k$个像素位置,$X_k'=(x_k',y_k')$表示模型预测的该像素在源图像中的对应位置,$\|\cdot\|_2$表示二维欧几里得距离的L2范数。

(2) 正确关键点百分比(Percentage of correct keypoints, PCK)

PCK[25]是评估预测点质量的指标,表示估计预测点位置与真实点位置之间的欧几里得距离小于某一阈值δ的百分比。PCK是一个基于误差范围的精度指标,用于衡量预测关键点的正确率,其值越大,预测正确的关键点越多,表明算法(模型)对图像的拼接效果越好;反之亦然。其定义为

$$\mathrm{PCK}(\delta)=\frac{1}{N}\sum_{k=1}^{N}\Phi(\|\boldsymbol{X}_k'-\boldsymbol{X}_k\|_2\leqslant\delta) \tag{15}$$

式中:$\Phi(\cdot)$为指示函数,成立为1,不成立为0;δ为误差容忍阈值,比如1、3和5像素。

(3) 准确率(Accuracy)

准确率[26]是指拼接成功的图像组数占拼接图像组总数的百分比,其值越高,表明拼接成功的图像越多,算法(模型)对图像的拼接效果越好;反之亦然。判断图像是否拼接成功是通过结合客观参数评估和主观观测评估的方法来判定的。首先使用客观参数评估:从图像匹配对的光流中获取匹配点,并去除外点,若图像匹配对的匹配点数量不少于4,则进行图像单应性变换和拼接;反之则认为该图像匹配对拼接失败。为避免误匹配情况,对客观参数评估后的拼接结果图进行主观观测,即通过视觉观察拼接图,如果未出现错位和拼接缝,则判定为拼接成功;反之则判定为拼接失败。

(4) 单组平均配准耗时(Average registration time per group, ARTG)

ARTG[27]表示算法(模型)在每组图像配准所需的平均时间,单位为s,由所有测试图像组的配准时间之和除以图像组数计算得到。该指标数值越小,代表配准效率越高。

3.3 实验结果与分析

3.3.1 对比实验

为了测试本文搭建的模型的性能,分别从12张岩石薄片采集序列图像制作数据集,分辨率均为1 368×912像素,并随机选取2 000组作为测试集进行拼接测试,每组图像包含1张基准图像和1张待拼接图像。

表1中的实验组别1~4展示了近年来一些具有代表性的深度学习图像拼接网络的拼接效果,可以看到,PDC-Net网络的各项性能指标表现较好,其中准确率达到了82.15%,单组平均配准耗时为0.407 s。表1中的实验组别5~9展示了5个经典传统图像配准算法的拼接效果,可以看到,SIFT算法准确率较高,达到了80.30%,但耗时较长,单组平均配准耗时为0.619 s。相较于PDC-Net,本文网络的AEPE降低了约24.37%,PCK-1px和PCK-5px相较PDC-Net分别提升了约8.92%和5.09%。本文网络相对SIFT算法AEPE降低了约44.06%,PCK-1px和PCK-5px分别提升了17.75%和10.84%。本文网络的准确率相比PDC-Net网络提升了4.60%,相比SIFT算法提升了6.45%,单组平均配准耗时为0.394 s,配准速度比PDC-Net稍快,相比SIFT算法配准速度得到显著提升;而ORB算法的配准速度虽然最快,

但其准确率最低。从表1可以看出,本文提出的网络模型拼接准确率最高,同时兼顾了配准时间,相较于其他算法和模型具有更佳的图像拼接效果。

表1 本文网络与其他图像拼接网络和算法的性能指标对比

Table 1 Comparison of performance metrics between the proposed network and other image stitching networks and algorithms

实验组别	算法/网络	AEPE ↓	PCK-1px/% ↑	PCK-5px/% ↑	Accuracy/% ↑	ARTG/s ↓
1	LoFTR	20.83	49.72	68.53	74.10	0.316
2	PDC-Net	11.08	66.97	82.94	82.15	0.407
3	LightGlue	11.90	64.22	81.68	81.50	0.392
4	XFeat	13.41	61.16	79.76	80.40	0.323
5	SIFT	14.98	58.14	77.19	80.30	0.619
6	SURF	17.29	51.35	75.38	78.25	0.590
7	BRISK	30.36	26.23	58.33	62.70	1.614
8	AKAZE	26.83	31.82	62.21	70.05	0.492
9	ORB	51.04	8.73	32.30	38.45	**0.240**
10	Ours	**8.38**	**75.89**	**88.03**	**86.75**	0.394

图8分别展示了GLU-Net网络、PDC-Net网络、SIFT算法以及本文网络对2组岩石薄片显微图像的配准结果和拼接结果,其中每组图像对应的左图为配准结果,右图为拼接结果。可以看到,本文网

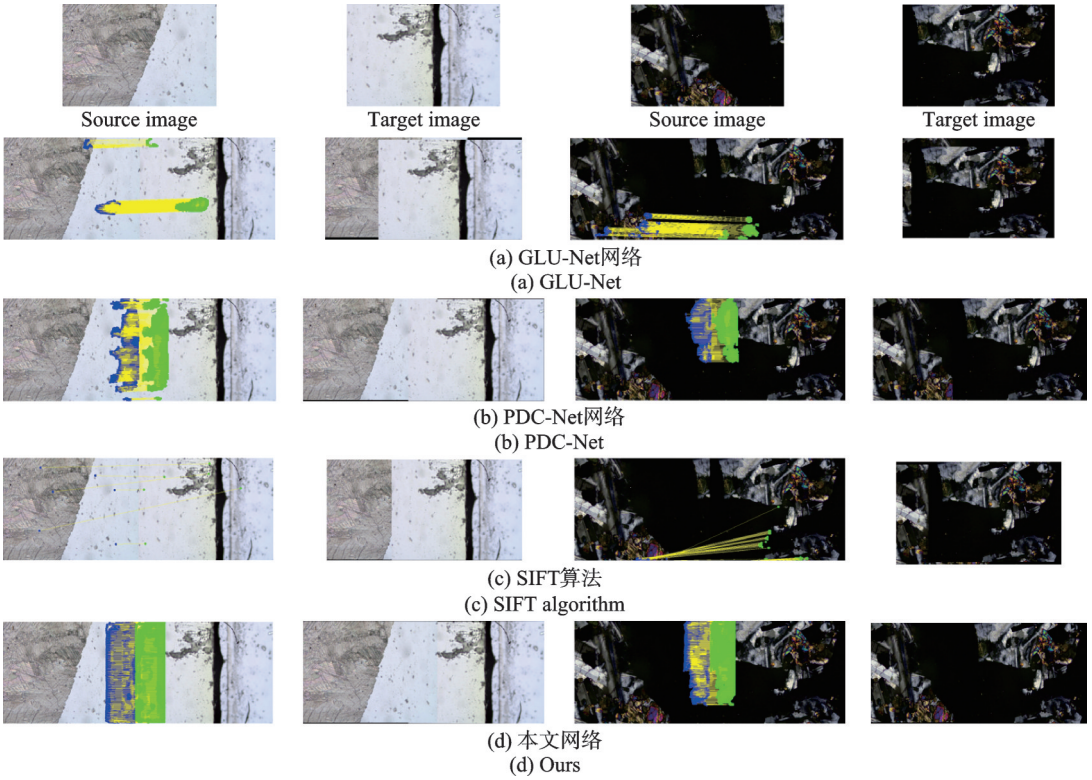


图8 不同算法和网络模型对源图像和目标图像的配准结果和拼接结果对比

Fig.8 Comparison of registration and stitching results for source and target images using different algorithms and network models

络的配准结果图具有更多正确的匹配点连线,而且拼接结果图没有出现拼接错位现象,拼接效果更好。

3.3.2 消融实验

消融实验的结果如表2所示。在基础模型GLU-Net上添加GOCor,然后在此基础上分别添加FPN和CBAM,同时添加FPN和CBAM则构成了本文网络。可以看到,随着模型的逐步改进,虽然ARTG在缓慢增长,但是其他指标在逐步优化。相较于GLU-Net基础网络,本文网络的AEPE从23.95下降至8.38,表明显著提升了模型光流估计的准确性;本文网络的PCK-1px和PCK-5px分别从初始的47.03%和67.20%提升至75.89%和88.03%,表明提取的正确匹配点占比显著提升。同时,本文网络的准确率在逐步提高,达到了86.75%,与初始GLU-Net网络相比提高了13.45%,模型的图像拼接性能得到有效提升。

表2 消融实验性能指标对比

Table 2 Comparison of performance metrics in ablation experiments

实验组别	算法/网络	AEPE↓	PCK-1px/%↑	PCK-5px/%↑	Accuracy/%↑	ARTG/s↓
1	GLU-Net	23.95	47.03	67.20	73.30	**0.301**
2	GLU-Net+GOCor	18.46	51.72	71.76	76.15	0.320
3	GLU-Net+GOCor+FPN	13.73	63.04	78.39	79.80	0.372
4	GLU-Net+GOCor+CBAM	11.36	65.13	81.28	80.55	0.387
5	Ours	**8.38**	**75.89**	**88.03**	**86.75**	0.394

3.3.3 泛化性实验

在显微镜对岩石薄片进行图像采集的过程中,由于机械抖动、光照变化等原因可能会导致采集的图像出现微小的旋转和尺度变换。为了验证本文网络的泛化能力,使用本文网络及原单应性矩阵对复杂几何变换下的岩石薄片显微图像和跨场景非岩石薄片图像进行拼接测试。

(1) 复杂几何变换下的岩石薄片显微图像拼接

制作500组具有旋转、缩放和投影变换的自制数据集图像对本文网络进行拼接测试。

表3给出了本文网络对复杂几何变换下岩石薄片显微图像拼接的评估结果,可以看到,拼接准确率达到81.80%,表明本文网络能够较好地适应具有不同几何变换的岩石薄片图像的拼接任务。仅当图像存在较大幅度的复杂几何变换且大面积区域纹理高度相似时,才可能出现拼接失败。图9展示了2组岩石薄片显微图像成功配准和拼接的结果。

表3 复杂几何变换下岩石薄片显微图像的拼接效果评估

Table 3 Evaluation of stitching performance on thin-section microscopic images under complex geometric transformations

实验组别	网络	AEPE↓	PCK-1px/%↑	PCK-5px/%↑	Accuracy/%↑	ARTG/s↓
1	Ours	11.54	65.72	82.62	81.80	0.419

(2) 跨场景图像拼接

在数据集MegaDepth中随机选取500组图像作为测试集对本文网络进行拼接测试,评估结果如表4所示,可以看到,拼接准确率达到了80.40%,表明本文网络对具有复杂几何变换的非岩石薄片图像拼接任务具有较好的适应能力。图10展示了2组不同场景的非岩石薄片图像成功配准和拼接的结果。

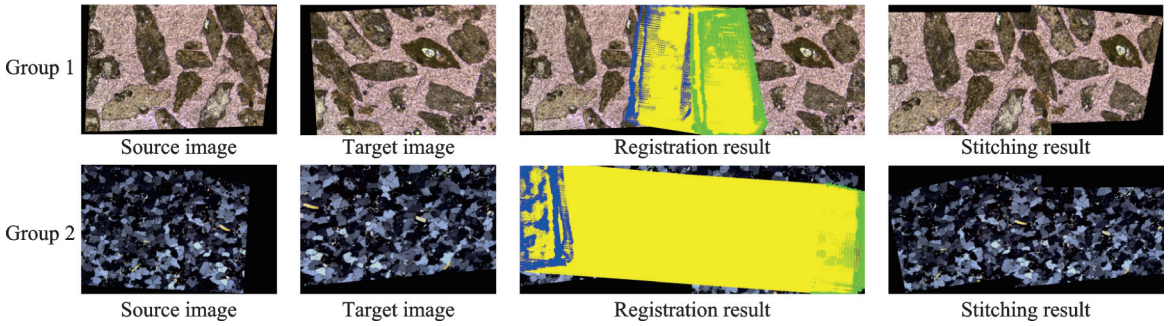


图 9 本文网络对复杂几何变换下的岩石薄片显微图像的配准和拼接结果

Fig.9 Registration and stitching results of rock thin-section microscopic images under complex geometric transformations by the proposed network

表 4 跨场景图像的拼接效果评估

Table 4 Evaluation of cross-scene image stitching performance

实验组别	网络	AEPE ↓	PCK-1px/% ↑	PCK-5px/% ↑	Accuracy/% ↑	ARTG/s ↓
1	Ours	13.79	59.64	78.03	80.40	0.414

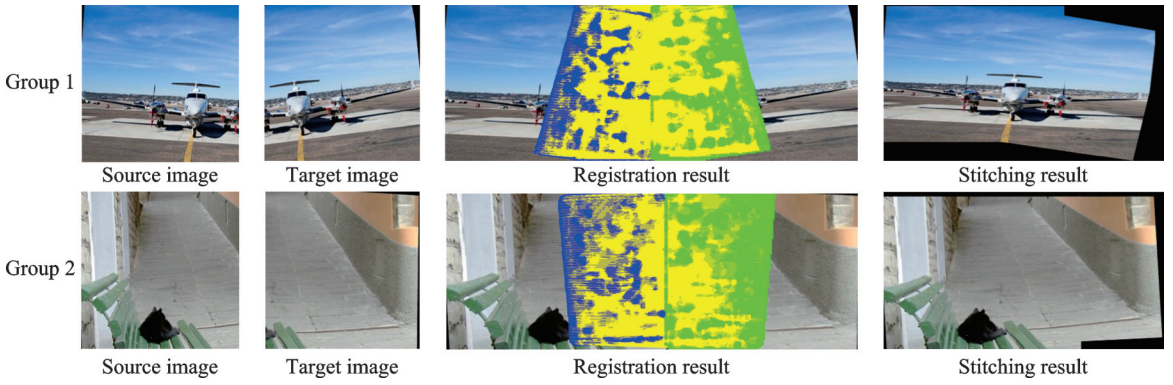
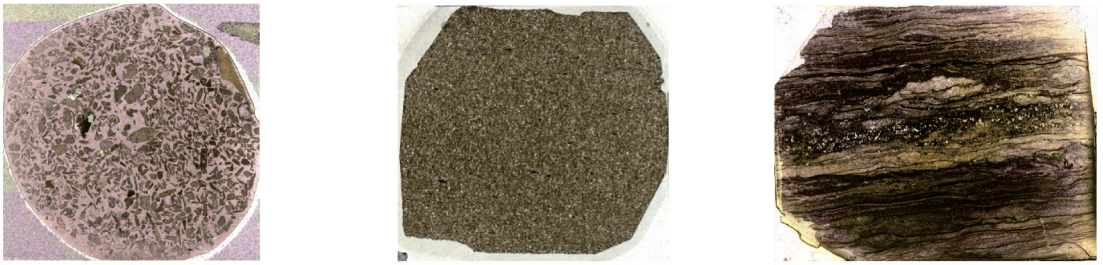


图 10 本文网络对跨场景非岩石薄片图像的配准和拼接结果

Fig.10 Registration and stitching results of non-rock thin-section images across different scenes by the proposed network

3.3.4 全薄片图像拼接

将 1.1 节中如图 1 所示的岩石薄片序列图像采集顺序作为序列图像拼接顺序，使用本文网络对采集的 12 张完整岩石薄片的序列图像进行拼接。其中的 3 张岩石薄片全幅面拼接图如图 11 所示，



(a) 完整的薄片拼接图 1
(a) Complete thin-section stitching image 1

(b) 完整的薄片拼接图 2
(b) Complete thin-section stitching image 2

(c) 完整的薄片拼接图 3
(c) Complete thin-section stitching image 3

图 11 完整岩石薄片拼接图像

Fig.11 Stitched image of complete rock thin-section

分别由432、330、528张序列图像拼接而成,分辨率分别为24 562×24 620、12 147×10 732和30 324×24 250像素。

4 结束语

针对传统特征提取与匹配算法在高分辨率岩石薄片偏光显微图像拼接中存在局部纹理复杂、全局相似度高、模糊区域无特征点导致全薄片图像拼接失败以及配准速度慢的问题,提出了一种基于改进GLU-Net的岩石薄片显微图像拼接方法。该方法在使用改进GOCor模块增强全局和局部对应关系计算的基础上,引入特征金字塔结构实现多尺度特征的融合,设计添加了自适应卷积注意力机制优化关键区域特征的注意力分配,然后通过全局映射解码器和局部光流解码器预测光流,最后基于获取的光流对图像进行单应性变换实现拼接,并且使用平均位移补偿法确保拼接出完整的岩石薄片图像,构建了改进的岩石薄片显微图像拼接网络架构。实验结果表明,与传统配准算法以及近年一些具有代表性的深度学习图像拼接网络相比,本文提出的网络具有拼接准确率高的优势,并且实现了配准速度与精度的平衡,具有在地质学研究中实时拼接的应用价值。本文网络的局限性在于,当图像存在复杂几何变换且大面积区域纹理浅弱并高度相似时,可能会出现拼接错位。后续研究将扩充包含复杂几何变换与复杂纹理特征的数据集,以进一步提升模型的拼接准确率。此外,可进一步优化特征提取与匹配策略,结合多类型图像特征并在多尺度层次上进行训练,从而进一步增强模型的拼接性能与泛化能力。

参考文献:

[1] CHEN Z, LI X, CHEN H, et al. The characteristics of lithofacies and depositional model of fine-grained sedimentary rocks in the ordos basin, China[J]. Energies, 2023, 16(5): 2390.

[2] FITZGERALD D J, BARRIOS C, LIU T, et al. A Novel polarized light microscope for the examination of birefringent crystals in synovial fluid[J]. Gout, Urate, and Crystal Deposition Disease, 2024, 2(4): 315-324.

[3] ALAVI R, CHANCY O, TRUDEL B, et al. Quantitative polarization microscopy as a potential tool for quantification of mechanical stresses within 3D matrices[J]. Acta Biomaterialia, 2025, 200: 236-250.

[4] LIU T, LI C S, LIU Z B, et al. Research on image identification method of rock thin slices in tight oil reservoirs based on mask R-CNN[J]. Energies, 2022, 15(16): 5818.

[5] KONG X P, QIU B H, LI Y J, et al. Application of thin slice technology of loose cuttings sample in early stage of experimental geology[J]. Journal of Physics: Conference Series, 2024, 2834(1): 012129.

[6] GUO X, CHEN Y, HE S, et al. Multi-scale channel enhanced transformer for rock thin sections identification and sequence consistency optimization[J]. Computational Geosciences, 2025, 29(3): 19.

[7] YU C, WU W, ZHENG J, et al. A multi-stage adaptive Otsu thresholding algorithm for pore segmentation in rock thin-section images[J]. Earth Science Informatics, 2025, 18(2): 239.

[8] YAN S, HAN L, MENG L, et al. Application of clay minerals in the synthesis of cretaceous mudstones and microstructural analysis of their deterioration mechanism[J]. Scientific Reports, 2024, 14(1): 29414.

[9] SHU J, HE X, TENG Q, et al. MMGC-Net: Deep neural network for classification of mineral grains using multi-modal polarization images[J]. Journal of Rock Mechanics and Geotechnical Engineering, 2025, 17(6): 3894-3909.

[10] SHU J, HE X, SU G, et al. Identification of plagioclase extinction-angle features from polarized images using deep neural network[J]. Physical Review E, 2024, 110: 045303.

[11] 先永利, 陈学健, 彭真明, 等. 基于深度学习的岩石钻孔全景图像识别[J]. 数据采集与处理, 2025, 40(3): 675-685.
XIAN Yongli, CHEN Xuejian, PENG Zhenming, et al. Rock drilling panoramic image recognition based on deep learning[J]. Journal of Data Acquisition and Processing, 2025, 40(3): 675-685.

[12] CHANG Y, LI Z, XU W. CGNet: A correlation-guided registration network for unsupervised deformable image registration [J]. IEEE Transactions on Medical Imaging, 2025, 44(3): 1468-1479.

[13] 赵帅, 张春堂, 樊春玲. 基于深度特征提取的深海序列图像拼接网络[J]. 电子测量技术, 2025, 48(3): 180-187.

ZHAO Shuai, ZHANG Chuntang, FAN Chunling. Deep-sea sequence image stitching network based on deep feature extraction[J]. Electronic Measurement Technology, 2025, 48(3): 180-187.

[14] 张琳娜, 陈建强, 吴妍, 等. 基于双匹配配准算法的多重复纹理图像拼接[J]. 数据采集与处理, 2021, 36(2): 334-345.
ZHANG Linna, CHEN Jianqiang, WU Yan, et al. Multi-repetitive texture image stitching based on dual matching registration algorithm[J]. Journal of Data Acquisition and Processing, 2021, 36(2): 334-345.

[15] WU P, LIU H, YI X, et al. Research on the registration of aerial images of cyclobalanopsis natural forest based on optimized fast sample consensus point matching with SIFT features[J]. Forests, 2024, 15(11): 1908.

[16] 陈丽娟, 虞业泺, 刘晨龙, 等. 卫星视觉导航图像拼接方法研究[J]. 电信科学, 2024, 40(12): 86-92.
CHEN Lijuan, YU Yeluo, LIU Chenlong, et al. Research on satellite vision navigation image stitching method[J]. Telecommunications Science, 2024, 40(12): 86-92.

[17] DEB P, KAR N, HASSAN L K, et al. Advanced copy-move forgery detection: Utilizing AKAZE in conjunction with SIFT algorithm for image forensics[J]. Microsystem Technologies, 2024, 31(3): 1-9.

[18] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: Learning feature matching with graph neural networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 4938-4947.

[19] TRUONG P, DANELLJAN M, TIMOFTE R. GLU-Net: Global-local universal network for dense flow and correspondences [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 6258-6268.

[20] SUN J, SHEN Z, WANG Y, et al. LoFTR: Detector-free local feature matching with transformers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 8922-8931.

[21] TRUONG P, DANELLJAN M, TIMOFTE R, et al. PDC-Net+: Enhanced probabilistic dense correspondence network[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(8): 10247-10266.

[22] LINDENBERGER P, SARLIN P E, POLLEFEYS M. LightGlue: Local feature matching at light speed[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2023: 17627-17638.

[23] POTGE G, CADAR F, ARAUJO A, et al. XFeat: Accelerated features for lightweight image matching[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2024: 2682-2691.

[24] LEE Y, LAMA B, JOO S, et al. Enhancing human key point identification: A comparative study of the high-resolution VICON dataset and COCO dataset using BPNET[J]. Applied Sciences, 2024, 14(11): 4351.

[25] MELEKHOV I, TIULPIN A, SATTLER T, et al. DGC-Net: Dense geometric correspondence network[C]//Proceedings of 2019 IEEE Winter Conference on Applications of Computer Vision. [S.l.]: IEEE, 2019: 1034-1042.

[26] LIU J, LI Y, WANG W, et al. Automatic, high-accuracy image registration in confocal microscopy[J]. Applied Optics, 2017, 56(32): 8924-8930.

[27] BOROVEC J, MUNOZ-BARRUTIA A, KYBIC J. Benchmarking of image registration methods for differently stained histological slides[C]//Proceedings of 2018 25th IEEE International Conference on Image Processing. [S.l.]: IEEE, 2018: 3368-3372.

作者简介:

向文亮(2000-),男,硕士研究生,研究方向:图像处理与识别,E-mail:xwlscue@163.com。

熊淑华(1969-),女,副教授,硕士研究生导师,研究方向:多媒体通信,E-mail:xiongsh@scu.edu.cn。

何海波(1968-),男,高级工程师,研究方向:图像处理、计算机应用,E-mail:351813868@qq.com。

滕奇志(1962-),女,教授,博士生导师,研究方向:数字图像处理、模式识别、三维图像重建及分析、计算机应用等,E-mail:qzteng@scu.edu.cn。

何小海(1964-),**通信作者**,男,教授,博士生导师,研究方向:图像处理与网络通信、人工智能与大数据分析、图像通信等,E-mail:hxh@scu.edu.cn。

(编辑:夏道家)

Image Stitching of Rock Thin Sections Microscopic Images Based on Improved GLU-Net

XIANG Wenliang[1], XIONG Shuhua[1], HE Haibo[2], TENG Qizhi[1], HE Xiaohai[1*]

(1. College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China; 2. Chengdu Xitu Technology Co. Ltd., Chengdu 610024, China)

Abstract: Rock thin-section microscopic images frequently exhibit complex local textures, blurriness, and high noise levels, posing significant challenges for traditional feature extraction and matching algorithms. These methods often fail to identify effective feature points in high-resolution rock thin-section images, hindering the realization of panoramic stitching while also resulting in slow processing speeds. To address the aforementioned issues, a rock thin- section microscopic image stitching method based on an improved GLU-Net has been proposed. This method integrates an enhanced correlation computation module to improve global and local correspondence, employs a feature pyramid network for multi-scale feature fusion, incorporates a designed adaptive convolutional attention mechanism to optimize attention to key regions, utilizes global and local decoders to obtain optical flow, and applies homography transformation for image stitching, thereby constructing a novel image stitching network model. Experimental results demonstrate that, compared to traditional image stitching algorithms and other classical image stitching network models, the proposed network achieves superior stitching performance. In stitching tests on a self-constructed dataset, a stitching accuracy of 86.75% has been attained with an average registration time of 0.394 s per pair, effectively balancing enhanced accuracy with processing efficiency.

Highlights:

1. A fusion module, feature pyramid network for multi-scale feature fusion, and an improved correlation module, globally-optimized correspondence volume, are introduced to enhance cross-scale representation and global-local correspondence modeling for weak and repetitive textures.
2. A multi-scale convolutional attention mechanism module is designed and embedded to adaptively modulate spatial and channel attention, suppress noise-dominated responses, and highlight discriminative structures.
3. A coarse-to-fine optical-flow estimation and stitching framework is built with a global mapping decoder and local optical-flow decoders. Alignment is completed via homography-based warping, and an average-displacement compensation strategy handles locally mismatched tiles to ensure continuity and completeness of full thin-section mosaicking.
4. Experiments show the proposed method improves stitching accuracy and efficiency on a self-built dataset and under complex geometric transformations and cross-dataset evaluations, outperforming traditional stitching algorithms and representative deep learning stitching models.

Key words: rock thin-section image; feature fusion; convolutional attention mechanism; optical flow estimation;image stitching

---

Foundation items: National Natural Science Foundation of China (No.62071315); Sichuan Provincial International Science and Technology Innovation Cooperation Project (No.2024YFHZ0289).

Received: 2025-07-18; **Revised:** 2025-10-10

***Corresponding author**, **E-mail:** hxh@scu.edu.cn.