

基于扩散模型的岩石薄片图像超分辨率重建^①



杜睿山^{1,2}, 穆文轩¹, 孟令东²

¹(东北石油大学 计算机与信息技术学院, 大庆 163318)

²(东北石油大学 油气藏及地下储库完整性评价黑龙江省重点实验室, 大庆 163318)

通信作者: 杜睿山, E-mail: durs918@163.com

摘要: 针对岩石薄片图像超分辨率重建过程中因纹理复杂导致现有重建方法效果不理想的问题, 提出面向岩石薄片图像的超分辨率网络模型 (super-resolution denoising diffusion probability model of rock slice, rsDDPMSR). 针对传统上采样方法往往会导致伪影和低分辨率图像先验信息利用不充分的问题提出分层特征增强网络 (layered feature enhancement network, LFE-Net), 利用双通路网络对平稳小波变换分解后的高频与低频分量进行分层特征增强. 为引导扩散模型的生成方向并提供丰富先验信息, 将经过 LFE-Net 增强后的低分辨率特征与目标高分辨率加噪图像特征通道拼接作为扩散模型的条件输入. 在 U-Net 的基础上设计了双编码器多尺度噪声预测网络 (ACA-U-Net) 有效处理岩石薄片多尺度信息并在跳跃连接中引入时间感知的自适应交叉注意力机制适配扩散模型不同去噪阶段的特征分布变化增强模型对关键区域的关注程度, 有效提升图像重建细节. 实验结果表明, rsDDPMSR 在 2×、4×、8×放大倍数下, 峰值信噪比 (PSNR) 和结构相似度 (SSIM) 相比于 CAMixerSR、SDFlow、IDM 和 SR3 等主流重建方法具有更优的重建效果.

关键词: 岩石薄片; 超分辨率重建; 小波变换; 扩散模型; 多尺度特征

引用格式: 杜睿山, 穆文轩, 孟令东. 基于扩散模型的岩石薄片图像超分辨率重建. 计算机系统应用, 2026, 35(2): 132-140. <http://www.c-s-a.org.cn/1003-3254/10059.html>

Super-resolution Reconstruction of Rock Thin Slice Image Based on Diffusion Model

DU Rui-Shan^{1,2}, MU Wen-Xuan¹, MENG Ling-Dong²

¹(College of Computer & Information Technology, Northeast Petroleum University, Daqing 163318, China)

²(Heilongjiang Provincial Key Laboratory of Oil and Gas Reservoir and Underground Gas Storage Integrity Evaluations, Northeast Petroleum University, Daqing 163318, China)

Abstract: In view of the problem that the existing reconstruction methods have no ideal effect due to the texture complexity during the super-resolution reconstruction of rock thin slice images, this study proposes a super-resolution denoising diffusion probability model of rock slice (rsDDPMSR). To solve the problem that the traditional upsampling methods tend to cause artifacts and insufficient utilization of prior information of low-resolution images, it puts forward a layered feature enhancement network (LFE-Net). Meanwhile, a dual-path network is employed to conduct layered feature enhancement on the high-frequency and low-frequency components decomposed by the stationary wavelet transform. The low-resolution features enhanced by LFE-Net are combined with the feature channels of the target high-resolution noisy image as the conditional input of the diffusion model to guide the generation direction of the diffusion model and provide rich prior information. Based on U-Net, a double-encoder multi-scale noise prediction network (ACA-U-Net) is designed to effectively process multi-scale information of rock slices, and a time-aware adaptive cross-attention mechanism is introduced into the skip connections to match the feature distribution changes in different denoising stages of the diffusion model to enhance the model's attention to key areas and effectively improve image reconstruction details. The

① 基金项目: 黑龙江省科技创新基地项目 (JD24A009)

收稿时间: 2025-07-11; 修改时间: 2025-08-01; 采用时间: 2025-08-19; csa 在线出版时间: 2025-12-19

CNKI 网络首发时间: 2025-12-22

experimental results show that rsDDPMSR has a better reconstruction effect at 2×, 4×, and 8× magnification than other mainstream reconstruction methods, such as CAMixerSR, SDFlow, IDM, and SR3.

Key words: rock thin slice; super-resolution reconstruction; wavelet transform; diffusion model; multi-scale feature

1 引言

岩石薄片图像的定量分析在地质勘探等领域^[1-3]具有重要作用. 当前利用深度学习技术对岩石薄片图像进行自动化定量分析因其能够得到快速、准确的分析结果而被广泛应用^[4-6].

在模型训练过程中往往仅通过对原始图像的旋转、切割等简单手段进行预处理. 但岩石薄片图像相比于自然场景中的物体(如人脸、建筑物等)通常具备更复杂的细节纹理和微观结构且部分类别间差距较小, 因此需要提供更精细的特征信息达到精确识别效果. 图像超分辨率(super-resolution, SR)重建^[7,8]技术旨在从低分辨率(low-resolution, LR)图像中恢复清晰的高分辨率(high-resolution, HR)图像, 有效恢复原始图像中的微观结构和纹理细节, 提升模型对岩石薄片图像的分析能力.

当前主流的SR重建算法主要分为基于插值的重建方法^[9]、基于重构的重建方法^[10]和基于深度学习的重建方法^[11]. 基于插值和重构的重建方法相比于基于深度学习的重建方法效果较差的原因在于: 基于插值的SR重建方法仅根据已有像素信息进行估计, 无法真正增加图像信息, 在插值过程中会产生不同程度的伪影; 基于重构的SR重建方法往往难以充分利用先验知识进行精准建模导致重建结果不够准确.

由于岩石薄片图像纹理复杂等特点, 基于深度学习的SR重建算法在岩石薄片图像上面临挑战. 基于回归式模型的SR重建算法^[12]由于缺乏对抗学习机制难以生成逼真纹理, 图像重建效果较差; 基于生成对抗网络(generative adversarial network, GAN)的SR重建算法^[13]往往训练过程复杂, 易出现梯度消失、模式崩溃等问题且难以理解复杂纹理图像的数据分布造成生成结果与原始目标相差过大; 基于扩散模型的SR重建算法^[14]相较上述算法具备更稳定的训练过程, 减少了模式崩溃风险, 更适合处理特征复杂的数据.

为应对岩石薄片图像超分辨率重建任务的挑战, 提出分层特征增强网络(layered feature enhancement network, LFE-Net). 对平稳小波变换(stationary wavelet transform, SWT)后的低频与高频子带图像通过特定网

络将LR子带图像映射到HR空间以增强子带图像细节并减少传统上采样方法伪影的产生.

针对岩石薄片图像具备多尺度特征信息的特点, 提出一种融合全局信息和局部信息的双编码器多尺度噪声预测网络(adaptive cross attention U-Net, ACA-U-Net), 有效增强网络图像重建效果.

针对岩石薄片图像特征复杂导致模型训练过程不稳定的问题, 提出将小波域内的LR子带图像与目标HR加噪子带图像进行通道拼接作为DDPM(denoising diffusion probabilistic model)逆向去噪过程中的条件输入, 以提供丰富的特征信息并引导模型图像生成方向.

2 相关工作

2.1 基于扩散模型超分辨率重建算法

随着生成式模型的发展, 条件概率扩散模型因其独特的生成机制以及强大的复杂数据分布捕捉能力在SR领域表现出巨大的潜力.

文献[15]提出了一种基于DDPM的逐步细化SR算法, 通过将LR图像与HR图像进行通道拼接作为DDPM的条件输入以引导生成方向. 文献[16]提出了一种二值化DDPM, 通过优化U-Net结构提高二值化模块的表达可靠性. 文献[17]通过构建LR编码网络对LR图像进行数据增强并通过双三次插值上采样将重建结果与目标HR图像对齐作为条件输入引导DDPM推理方向得到SR图像. 文献[18]提出了基于DDPM的遥感图像SR算法通过融合残差连接和自注意力机制的深度噪声训练网络以获得更优的超分辨率效果. 文献[19]提出了残差超分辨率DDPM, 通过在高分辨率和低分辨率图像之间移动残差来实现图像间的转移有效减少扩散步骤数量. 文献[20]通过DDPM结合多层自注意力机制和基于小波变换的LR编码器增强DDPM对超声图像超分辨率任务的适用性.

2.2 去噪扩散概率模型

DDPM通过设定时间步 $t \in \{1, 2, \dots, T-1, T\}$ 进行前向扩散学习噪声分布然后通过逆向过程从噪声中逐步恢复清晰图像其中前向扩散马尔可夫过程, 如式(1)所示, X_{t-1} 和 X_t 分别对应时间步 $t-1$ 和 t 的数据, $\alpha_t \in (0, 1)$

为预定义时间步长参数,用于控制 t 时刻步长加入的噪声程度, I 表示单位矩阵, $N(x;\mu,\sigma^2)$ 表示均值为 μ ,方差为 σ^2 的正态分布.

$$q(X_t|X_{t-1}) = N(X_t; \sqrt{\alpha_t}X_{t-1}, (1 - \alpha_t)I) \quad (1)$$

逆向去噪过程如式(2)所示:

$$X_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(X_t - \frac{\beta_t}{\sqrt{1 - \alpha_t}} \varepsilon_\theta(X_t, t) \right) + \sigma_t Z \quad (2)$$

其中, $\beta_t = 1 - \alpha_t$ 表示噪声程度, $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$ 表示到 t 步为止所有步骤中保留的原始信息比例, Z 是标准正态分布 $N(0, I)$ 中抽取的随机变量, σ_t 是方差项, ε_θ 为经过特

定训练的神经网络对去除噪声的数据分布估计.

DDPM 通过逐步学习如何从噪声中恢复原始清晰图像,适合处理多尺度的图像特征,能够同时考虑全局特征与局部特征,为超分辨率重建任务提供坚实的理论基础.

3 岩石薄片超分辨率重建网络模型

本文提出的岩石薄片超分辨率重建网络模型 (super-resolution denoising diffusion probability model of rock slice, rsDDPMSR),如图1所示,由LFE-Net和结合ACA-U-Net的DDPM共同组成.

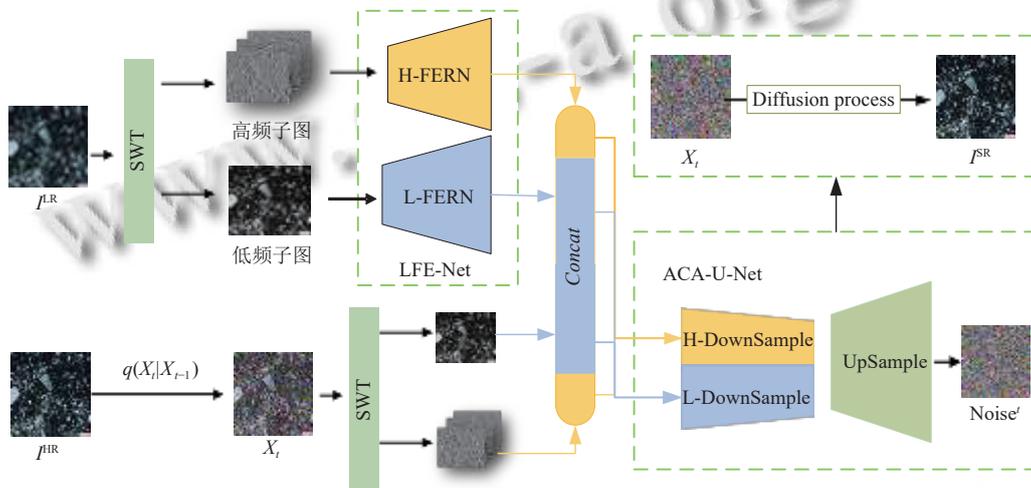


图1 rsDDPMSR 总体结构图

通过 LFE-Net 对 LR 图像经平稳小波变换分解后的低频和高频子图进行针对性特征增强,以解决重建过程中的细节信息丢失问题,并为 DDPM 提供多尺度特征输入,从而有助于引导扩散模型生成更加真实准确的重建结果.在 DDPM 的逆向去噪过程中,将平稳小波变换分解后的目标 HR 子图与增强后的 LR 子图通道拼接,作为 ACA-U-Net 噪声预测网络的初始条件引导预测噪声分布,通过迭代细化最终生成高分辨率岩石薄片图像.

3.1 分层特征增强网络

LFE-Net 分层特征增强网络如图2所示,包含针对高频子图的高频增强残差网络 (high-frequency enhanced residual network, H-FERN) 和针对低频子图的低频增强残差网络 (low-frequency enhanced residual network, L-FERN).对 SWT 分解得到的 LR 高频和低频子图进行分层处理,通过特征融合的方式增强图像结构.输出特征作为 DDPM 的条件向量,为后续扩散

过程提供多尺度地质先验约束.

(1) 高频增强残差网络

针对高频分量对噪声敏感,直接放大易放大噪声,从而降低图像质量这一问题,设计 H-FERN 以精准恢复高频细节.具体结构如图2中 H-FERN 部分所示,通过深度卷积对高频分量的拼接图的每个通道进行单独特征提取,利用可变形卷积 (deformable convolution, DCN) 动态适配岩石薄片的不规则几何结构,并利用结合了 CBAM 注意力机制和可变形卷积的变形注意力增强模块 (deformable attention enhancement block, DAEB),实现对多尺度上下文信息捕捉以及局部细节信息的保留.

DAEB 结构如图3所示,其中可变形卷积设置偏移量范围为 $[-25, 25]$,以允许卷积核在特征图上进行较大范围的移动,从而更好地适应不同的局部结构.在 DAEB 模块连接中加入残差连接增强特征传递能力的同时有效解决梯度消失问题.

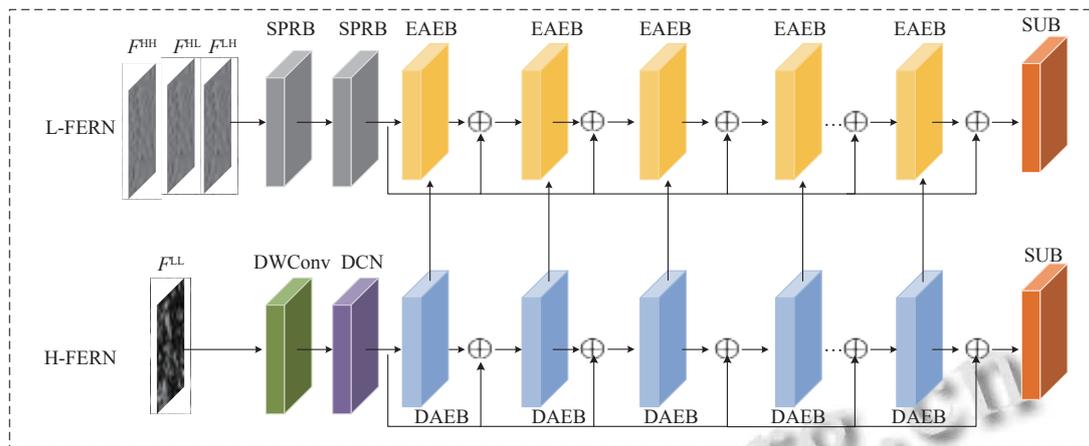


图2 LFE-Net 总体结构图

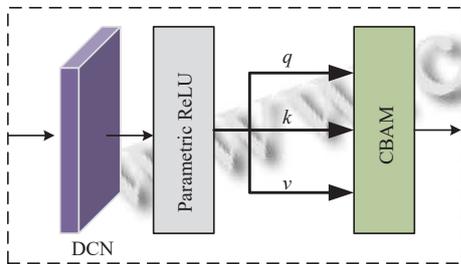


图3 DAEB 结构图

(2) 低频增强残差网络

针对岩石薄片图像低频分量内容复杂导致难以捕捉有效特征问题设计 L-FERN 以增强低频子图细节. 具体结构如图 2 中 L-FERN 部分所示. 初始低频子图 $F^{LL} \in R^{c \times h \times w}$ 通过结构保持残差模块 (structure-preserving residual block, SPRB) 增强图像边缘信息. 将增强后的低频子图 F_S 与当前第 t 个 DAEB 模块的高频增强子图 f_t^{HR} 作为 EAEB 模块的初始输入得到低频特征增强图 $F_t \in R^{c \times h \times w}$.

SPRB 结构如图 4 所示, 通过深度卷积进行特征提取, 利用点卷积扩展通道数, 再将通道数压缩回 C , 同时对输入使用固定的拉普拉斯卷积核进行卷积操作增强边缘信息. 将得到的结果相加得到最终增强输出 Y .

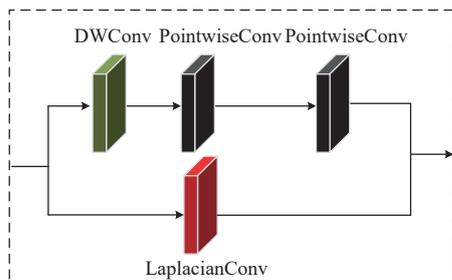


图4 SPRB 结构图

EAEB 结构如图 5 所示, 通过指数增长空洞率的空洞卷积根据输入特征动态捕捉多尺度信息, 并通过可变形卷积适配不规则的几何结构. 将特征提取结果与 DAEB 增强后的特征信息通过 CBAM 注意力机制进行特征融合实现低频子图特征增强.

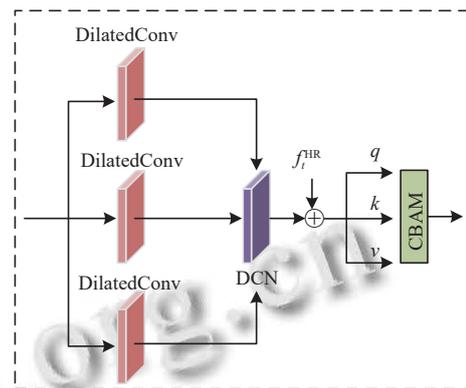


图5 EAEB 结构图

(3) 亚像素上采样模块

亚像素上采样模块 (subpixel upsampling block, SUB) 将增强后的子图特征 f 经过深度卷积层, 对特征进行初步提取得到提取特征 f' , 如式 (3) 所示, 其中 D 为深度卷积层.

$$f' = D(f) \tag{3}$$

通过亚像素卷积像素排列组合得到目标高分辨率增强子图上采样结果 F , 如式 (4) 所示, 其中 PS 为亚像素卷积层, s 为上采样因子.

$$F = PS(f', s) \tag{4}$$

3.2 扩散模型 rsDDPM

为引导扩散模型生成方向并提供多尺度特征信息,

本文提出将经过 LFE-Net 增强后的低频、高频特征子图与 SWT 分解的目标 HR 低频、高频加噪子图进行通道拼接作为条件输入, 如图 6 所示, 从而定义了一个新的条件化机制, 如式 (5) 所示:

$$q(X_t|X_{t-1}, C_L, C_H) = N(X_t; \mu_t, \sigma_t^2 I) \quad (5)$$

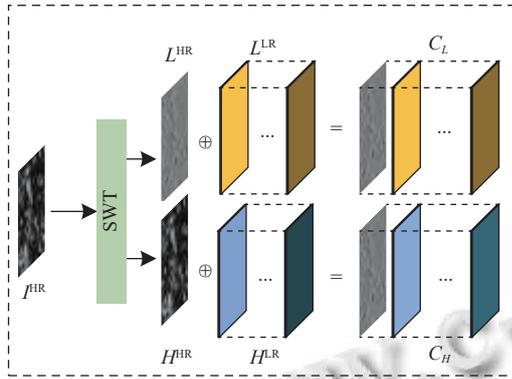


图 6 数据拼接方式

μ_t 计算如式 (6) 所示, 其中 $C_L = \text{Concat}(L^{LR}, L^{HR})$, f 为神经网络预测的函数, 根据条件输入 C_L 和 C_H 调整均值和方差。

$$\mu_t = \sqrt{\alpha_t} X_{t-1} + \beta_t f(C_L, C_H) \quad (6)$$

σ_t^2 计算如式 (7) 所示, 其中 $C_H = \text{Concat}(H^{LR}, H^{HR})$, g 为神经网络预测的函数根据条件输入 C_L 和 C_H 调整均值和方差。

$$\sigma_t^2 = (1 - \alpha_t) g(C_L, C_H) \quad (7)$$

同时在 U-Net 的基础上设计 ACA-U-Net 噪声预测网络如图 7 所示, 通过双编码网络对高频和低频信息分别进行特征信息处理. 对输入的高频和低频特征子图通过编码网络提取特征并通过逆转换的方式对高频和低频特征进行合并. 为更好适应不同时间步长的噪声去除过程以及过滤无关信息, 在跳跃连接中引入时间感知的自适应交叉注意力机制适配不同去噪阶段的特征分布变化。

时间感知自适应交叉注意力机制 (cross attention with time embedding, CATE) 通过一个嵌入层将时间步长 t 映射到高维空间, 生成查询向量 Q 来显式建模时间步与特征空间的关系. 令 d 为嵌入维度, i 是嵌入向量索引, W_{emb} 和 b_{emb} 分别是嵌入层的权重矩阵和偏置项. 对于原始输入特征 F_{in} , 通过残差连接来保留原始特征信息, 防止梯度消失. 通过权重矩阵 W_K 和 W_V 生成 K 和 V , 如式 (8)–式 (10) 所示:

$$Q = W_{\text{emb}} \times \sin\left(\frac{t}{10000} \frac{2i}{d}\right) + b_{\text{emb}} \quad (8)$$

$$K = W_K(F_{\text{in}}) \quad (9)$$

$$V = W_V(F_{\text{in}}) \quad (10)$$

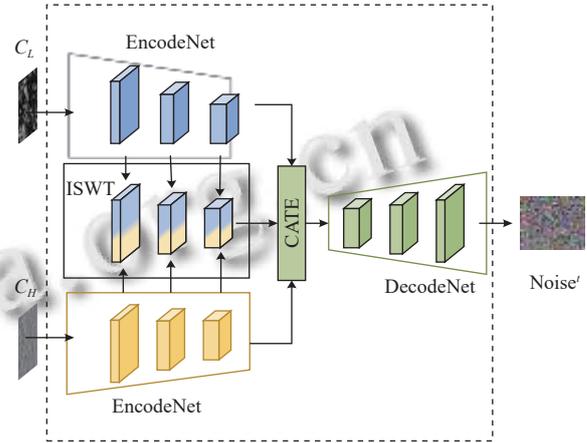


图 7 ACA-U-Net 总体结构图

通过拉普拉斯算子分别提取小波域内 LR 和 HR 加噪图像的边缘特征, 得到边缘特征图 E_{LR} 和 E_{HR} . 通过归一化生成空间注意力掩码 M_{spatial}^{LR} 和 M_{spatial}^{HR} , 并根据时间步的不同对二者进行加权融合, 形成最终的空间注意力掩码 M_{spatial} . 帮助模型聚焦关键区域, 忽略背景噪声与无关区域. 注意力权重计算公式如式 (11) 所示:

$$\text{Attn} = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}} \odot M_{\text{spatial}}\right) \quad (11)$$

CATE 的时空联合建模的方式能够突出显示重要信息并减少不相关信息的影响, 从而生成更加精细、准确的重建结果。

4 实验结果与分析

4.1 数据集

本文采用南京大学岩石教学薄片显微图像数据集训练模型. 该数据集涵盖 3 大类岩石中 90% 以上的常见岩石类型, 共 108 种岩石 324 个岩石薄片的 2634 张偏光显微图像. 遵循样本选取的随机性、代表性、平衡性原则选取每种岩石薄片对应图片共 972 张作为训练集以及 324 张图片作为测试集. 为验证 rsDDPMSR 算法重建性能, 本文在放大因子为 2×、4× 和 8× 的条件下进行超分辨率重建实验, 并对实验结果进行分析。

4.2 实验环境及参数设置

本文基于 PyTorch 深度学习框架构建相应网络模

型, 实验在处理器 Intel(R) Core(TM) i9-14900KF, 显卡 NVIDIA GeForce RTX 4090 24 GB, Python 3.9, CUDA 11.8 环境下进行. LFE-Net 中高频与低频网络的 DAEB 和 EAEB 块使用数量均为 5 个, 在 rsDDPM 中, 令 $t = 2000$ 作为扩散时间步长. 在训练过程中采用 L1 损失函数, 网络迭代过程中采用线性预热策略在前 1×10^5 步逐渐增加学习率之后保持迭代降低学习率到 $2E-5$. 以 1×10^4 次迭代作为指标, 使用 Adam 优化器优化网络参数, 其中超参数 $\gamma_1 = 0.9$, $\gamma_2 = 0.999$, $\alpha = 1E-4$, 图像训练批量为 16, 并以 PSNR、SSIM 和 FID (Frechét inception distance) 作为评价算法的评价指标.

4.3 消融实验

为验证 LFE-Net 中各模块的有效性设计 5 组实验, 以双三次上采样模块作为 Base 组, 其他为仅添加 SUB 上采样模块、添加 SUB 和 L-FERN、添加 SUB 和 H-FERN 以及添加所有模块的 4 种情况, 对比结果如表 1 所示. 对照项 1 表明 SUB 模块相比于双三次上采样模块效果较好. 对照项 2、3 在其基础上分别添加 L-FERN 和 H-FERN 网络模块.

表 1 LFE-Net 消融实验结果

| 对照项 | SUB | L-FERN | H-FERN | PSNR (dB)↑ | SSIM↑ | FID↓ |
|------|-----|--------|--------|------------|--------|-------|
| Base | — | — | — | 16.70 | 0.7855 | 61.34 |
| 1 | √ | — | — | 17.02 | 0.7932 | 57.33 |
| 2 | √ | √ | — | 17.34 | 0.8057 | 57.12 |
| 3 | √ | — | √ | 17.26 | 0.7989 | 54.52 |
| 4 | √ | √ | √ | 17.69 | 0.8334 | 52.35 |

实验结果表明, 在仅添加 SUB 模块的情况下 3 个指标均有提升, 表明其有助于提高生成图像与真实图像之间的相似性. 当进一步引入 L-FERN 后, 相较于仅加入 SUB 的情况, L-FERN 不仅能够提升图像质量, 同时还保持了生成图像的多样性. 当使用 H-FERN 替代 L-FERN 后, 虽然 PSNR 和 SSIM 略低于 SUB+L-FERN 的组合, 但 FID 的大幅改善显示 H-FERN 有效减少了生成图像与实际图像分布差距, 从而提高模型的整体表现.

最终整合所有模块时, 模型达到最佳水平, 即 PSNR 为 17.69 dB, SSIM 为 0.8334, FID 为 52.35. 表明 3 个模块的协同不仅在图像质量上实现了提升, 同时也在生成图像的真实性方面表现出色.

为验证 LFE-Net 在模型收敛以及超分辨率重建任务上的有效性, 本文对比了使用 LFE-Net 增强后的特征子图作为输入和不经增强仅通过平稳小波变换的原始子图作为输入的两种情况的迭代次数与 PSNR

的关系, 结果如图 8 所示.

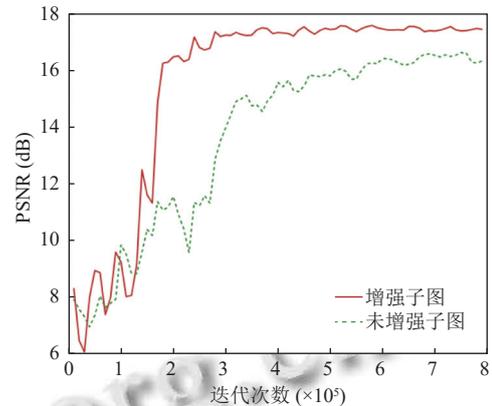


图 8 迭代次数和 PSNR 关系

可以看出, 随着迭代次数的增加, 使用增强子图作为输入的 PSNR 数值整体优于未增强子图作为输入的情况, 并且在使用增强子图作为输入时算法在 4.7×10^5 次迭代中达到了最高 PSNR 数值, 而未使用则在 7.6×10^5 次迭代中达到最优. 增强子图由于更有效地提取了低分辨率图像中的特征, 丰富了输入信息, 从而辅助扩散模型能够更好地收敛, 以达到更好的超分辨率效果.

为验证 ACA-U-Net 改进方法的有效性设计 4 组实验, 以原始 U-Net 作为 Base 组, 其他为仅使用双编码器 U-Net 和仅增加 ACA-Attention 的 U-Net 以及结合两种方式的 ACA-U-Net 共 3 种情况. 对比结果如表 2 所示.

表 2 ACA-U-Net 消融实验结果

| 对照项 | 双编码网络 | CATE | PSNR (dB)↑ | SSIM↑ | FID↓ |
|------|-------|------|------------|--------|-------|
| Base | — | — | 17.46 | 0.8081 | 59.12 |
| 1 | √ | — | 17.57 | 0.8148 | 56.84 |
| 2 | — | √ | 17.54 | 0.8123 | 55.61 |
| 3 | √ | √ | 17.69 | 0.8334 | 52.35 |

根据对照项 1 在仅添加双分支网络结构后 PSNR 和 SSIM 相较于 Base 组分别提升了 0.11 dB 和 0.0067, 且 FID 下降了 2.28 的结果表明: 通过独立处理高频和低频特征的编码方式能够使网络更好地提取多尺度特征, 有效增强模型对于图像细节和主体结构的分离表征能力. 对照项 2 在单独添加 CATE 时, PSNR 和 SSIM 的提升不如双分支网络显著但 FID 改善更为明显表明. CATE 在减少生成图像与实际图像分布之间的差距方面表现优异, 能够使模型在不同时间步中关注相应关键区域. 对照项 3 相较于 1、2 提升最为明显, 体现出双编码器提供的多频特征为注意力机制提供了更丰富的调控对象. 二者相互协调证明了 ACA-U-Net 的

有效性.

4.4 对比实验

为验证本文算法有效性,与双三次插值算法 Bicubic^[21]、生成对抗网络的 ESRGAN^[22]、基于 Transformer 的 SwinIR^[23]和基于扩散模型的 SR3^[19]算法以及当前主流的 SDFlow^[24]、IDM^[25]、CAMixerSR^[26]算法在放大因子为 2×、4×和 8×环境下进行对比实验分析.

根据表 3–表 5 结果显示,随着放大倍数增加,所有算法的 PSNR 和 SSIM 普遍上升,且 FID 普遍下降.这表明更大的放大倍数需要生成更多的细节信息,增加了任务的难度. Bicubic 在所有放大倍数下均为最差,反映了传统插值方法的局限性.

表 3 2×SR 下各算法客观指标对比结果

| 对照项 | PSNR (dB)↑ | SSIM↑ | FID↓ |
|-----------|--------------|---------------|--------------|
| Bicubic | 20.11 | 0.7585 | 86.14 |
| ESRGAN | 20.89 | 0.8116 | 35.42 |
| SwinIR | 21.80 | 0.7730 | 37.06 |
| SDFlow | 22.43 | 0.8241 | 39.21 |
| IDM | 20.34 | 0.7933 | 31.64 |
| CAMixerSR | 23.12 | 0.8021 | 37.22 |
| SR3 | 25.78 | 0.8654 | 35.77 |
| rsDDPMSR | 25.97 | 0.8799 | 30.93 |

表 4 4×SR 下各算法客观指标对比结果

| 对照项 | PSNR (dB)↑ | SSIM↑ | FID↓ |
|-----------|--------------|---------------|--------------|
| Bicubic | 17.08 | 0.5053 | 96.29 |
| ESRGAN | 17.09 | 0.5383 | 49.30 |
| SwinIR | 19.47 | 0.6687 | 60.81 |
| SDFlow | 19.32 | 0.6314 | 48.34 |
| IDM | 17.66 | 0.5242 | 54.94 |
| CAMixerSR | 19.13 | 0.6825 | 60.81 |
| SR3 | 19.44 | 0.7062 | 45.34 |
| rsDDPMSR | 19.57 | 0.8556 | 45.11 |

表 5 8×SR 下各算法客观指标对比结果

| 对照项 | PSNR (dB)↑ | SSIM↑ | FID↓ |
|-----------|--------------|---------------|--------------|
| Bicubic | 16.03 | 0.3362 | 124.25 |
| ESRGAN | 16.33 | 0.4337 | 60.94 |
| SwinIR | 17.00 | 0.3581 | 78.11 |
| SDFlow | 16.14 | 0.4662 | 56.13 |
| IDM | 15.08 | 0.3386 | 66.85 |
| CAMixerSR | 16.09 | 0.4639 | 65.74 |
| SR3 | 17.34 | 0.6812 | 54.22 |
| rsDDPMSR | 17.69 | 0.8334 | 52.35 |

SR3 和本文所提方法在多个指标上优于传统基于生成对抗网络方法,表现出在高倍放大时,扩散模型具有更强的稳定性和生成质量,同时本文所提方法在 4×和 8×的高难度任务中优势更加明显,SSIM 和 FID 指标上均领先于其他方法,表明本文所提方法生成的

图像分布更接近于真实图像,具备更好的多样性和真实性,能够更好地保留图像的整体结构和语义信息.

本文对不同纹理的岩石薄片图像进行主观实验对比,并采用 Grad-CAM 方法对 CATE 注意力与传统注意力机制进行可视化对比分析.

不同时间步的注意力热力图如图 9 所示,左为不同时间步原图,中为 CATE 注意力热力图,右为多头注意力热力图.对比时间步分别为 150、300 和 500 的不同岩石薄片图像可以看到:使用 CATE 注意力机制时模型在早期噪声相对较大时更倾向于覆盖整个图像范围以更多地关注粗略结构的重建,而在后期则更加集中于局部细节特征.

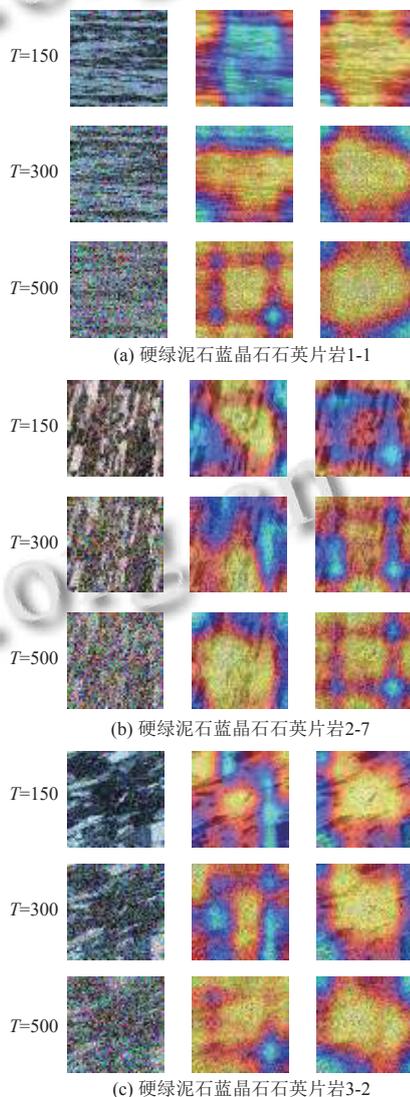


图 9 不同岩石薄片注意力可视化

传统注意力机制无论在哪一个时间步进行推理,始终都会以相同的方式计算注意力权重,导致无法根

据去噪的不同阶段来调整模型的关注点。

不同算法在 $8\times$ 环境下, 对不同岩片图像的实验结果如图 10-图 12 所示. 本文算法对图像纹理细节重建更为真实, 能够生成更为自然的边缘过渡. SR3 算法与本文具有相近结果, 但因其多尺度特征利用较差, 缺乏对细粒度图像信息处理能力导致重建的图像保真度相

对较弱, 对细节恢复存在不足. 本文所提出的重建算法在主观上效果最佳, 具备更好的细节纹理恢复和高保真效果. 通过 LFE-Net 充分融合利用图像的高频和低频信息, 双编码噪声预测网络能够提升模型理解能力, 自适应通道注意力机制调整模型在不同时间步的关注方向. 使重建图像观感相较其他算法有很大提升.

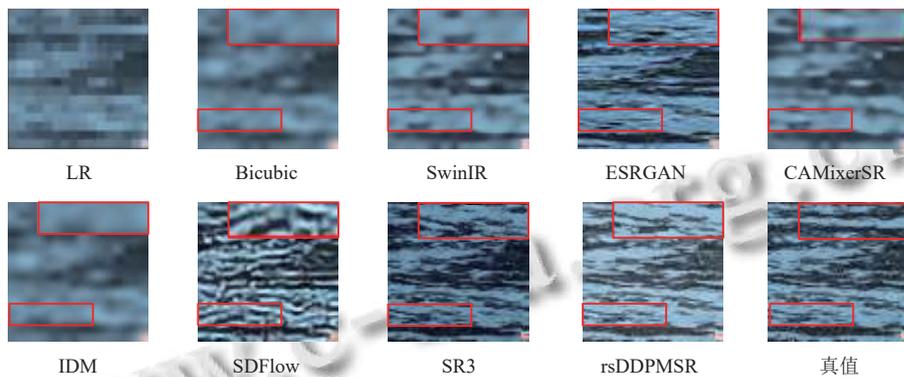


图 10 硬绿泥石蓝晶石石英片岩 1-1 在 $8\times$ 环境下的结果

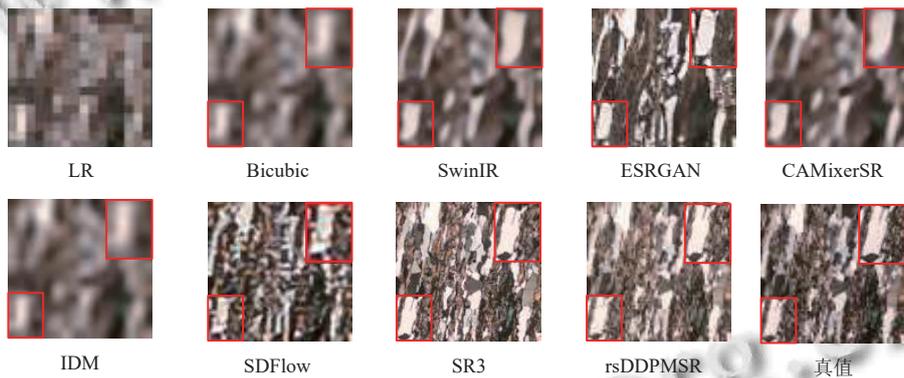


图 11 硬绿泥石蓝晶石石英片岩 2-7 在 $8\times$ 环境下的结果

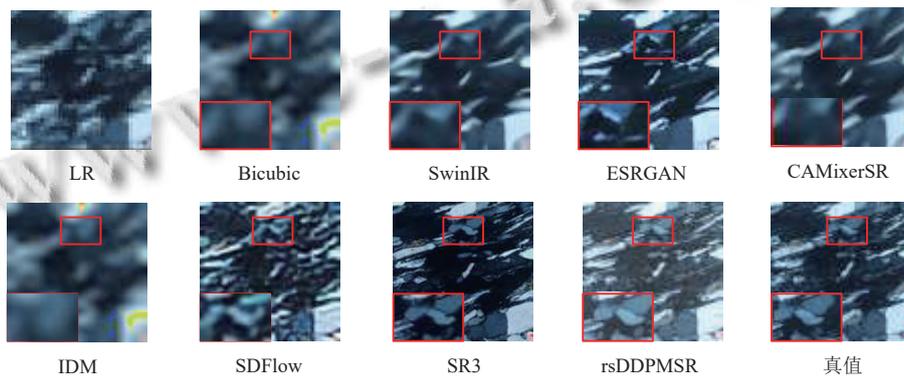


图 12 硬绿泥石蓝晶石石英片岩 3-2 在 $8\times$ 环境下的结果

5 结论

本文提出了一种融合小波域特征和扩散模型的针对岩石薄片图像的超分辨率重建网络模型. 通过分层增强网络融合图像多尺度特征丰富图像信息, 利用

SWT 分解后的子图与目标 HR 加噪子图进行通道拼接作为 DDPM 的条件输入实现岩石薄片图像的超分辨率重建, 通过与当前主流超分辨率重建算法对比证明, 本文提出的 rsDDPMSR 网络模型可以实现更有效的

岩石薄片图像超分辨率重建.

参考文献

- 1 Zhang Y, Zhang XL, Li ZL, *et al.* Hierarchical intelligent lithology recognition for thin section images using enhanced DenseNet. *Earth Science Informatics*, 2025, 18(1): 114. [doi: [10.1007/s12145-024-01663-2](https://doi.org/10.1007/s12145-024-01663-2)]
- 2 Chen FZ, Xu J, Tang LG, *et al.* Microscopic investigation of chemical and physical alteration behaviors in sandstone minerals induced by CO₂-H₂O-rock interactions during CO₂ saline aquifer storage. *Science of the Total Environment*, 2025, 958: 177897. [doi: [10.1016/j.scitotenv.2024.177897](https://doi.org/10.1016/j.scitotenv.2024.177897)]
- 3 Remitti F, Festa A, Nirta G, *et al.* Role of folding-related deformation in the seismicity of shallow accretionary prisms. *Nature Geoscience*, 2024, 17(7): 600–607. [doi: [10.1038/s41561-024-01474-6](https://doi.org/10.1038/s41561-024-01474-6)]
- 4 任书杰, 胡勇, 何文祥, 等. 基于深度学习的砂岩组分显微图像识别. *科学技术与工程*, 2024, 24(9): 3727–3736. [doi: [10.12404/j.issn.1671-1815.2302321](https://doi.org/10.12404/j.issn.1671-1815.2302321)]
- 5 李春生, 刘涛, 刘宗堡, 等. 基于空间注意力机制的 Mask R-CNN 致密储层岩石薄片图像鉴定. *中国石油大学学报(自然科学版)*, 2024, 48(4): 24–32.
- 6 Zhao L, Zhang HL, Sun XD, *et al.* Application of ResUNet-CBAM in thin-section image segmentation of rocks. *Information*, 2024, 15(12): 788. [doi: [10.3390/info15120788](https://doi.org/10.3390/info15120788)]
- 7 Maiseli B, Abdalla AT. Seven decades of image super-resolution: Achievements, challenges, and opportunities. *EURASIP Journal on Advances in Signal Processing*, 2024, 2024(1): 78. [doi: [10.1186/s13634-024-01170-y](https://doi.org/10.1186/s13634-024-01170-y)]
- 8 Yu M, Shi JC, Xue CH, *et al.* A review of single image super-resolution reconstruction based on deep learning. *Multimedia Tools and Applications*, 2024, 83(18): 55921–55962.
- 9 刘宇豪, 吴福培, 吴树壮, 等. 基于插值超分辨的双目三维重建方法. *中国光学(中英文)*, 2024, 17(4): 862–874.
- 10 Dan J, Jin T, Chi H, *et al.* PIRN: Phase invariant reconstruction network for infrared image super-resolution. *Neurocomputing*, 2024, 599: 128221. [doi: [10.1016/j.neucom.2024.128221](https://doi.org/10.1016/j.neucom.2024.128221)]
- 11 Chen H, Li H, Yao CX, *et al.* Image super-resolution based on improved ESRGAN and its application in camera calibration. *Measurement*, 2025, 242: 115899.
- 12 Gendy G, Hou JC, Sabor N, *et al.* Lightweight image super-resolution network based on dynamic graph message passing and convolution mixer. *Expert Systems with Applications*, 2025, 263: 125683. [doi: [10.1016/j.eswa.2024.125683](https://doi.org/10.1016/j.eswa.2024.125683)]
- 13 Park SW, Jung SH, Sim CB. Correction: NeXtSRGAN: Enhancing super-resolution GAN with ConvNeXt discriminator for superior realism. *The Visual Computer*, 2025, 41(10): 8119. [doi: [10.1007/s00371-025-03834-8](https://doi.org/10.1007/s00371-025-03834-8)]
- 14 Wang R, Zhou NN. Improved conditional diffusion model for image super-resolution. *IET Image Processing*, 2025, 19(1): e13313.
- 15 Saharia C, Ho J, Chan W, *et al.* Image super-resolution via iterative refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(4): 4713–4726.
- 16 Chen Z, Qin HT, Guo Y, *et al.* Binarized diffusion model for image super-resolution. *Proceedings of the 38th International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2024. 30651–30669.
- 17 许晓阳, 张梦飞. 融合 LR 编码网络和扩散模型的遥感图像超分辨率算法. *计算机工程与应用*, 2024, 60(22): 271–281. [doi: [10.3778/j.issn.1002-8331.2311-0166](https://doi.org/10.3778/j.issn.1002-8331.2311-0166)]
- 18 付奕博, 谢东海, 王志博, 等. 基于条件控制扩散模型的遥感图像超分辨率增强算法. *地球信息科学学报*, 2024, 26(10): 2384–2393. [doi: [10.12082/dqxxkx.2024.240315](https://doi.org/10.12082/dqxxkx.2024.240315)]
- 19 Liu TY, Han S, Xie LR, *et al.* Super-resolution reconstruction of ultrasound image using a modified diffusion model. *Physics in Medicine & Biology*, 2024, 69(12): 125026.
- 20 Song XL, Liu X, Luo ZX, *et al.* High-resolution iterative reconstruction at extremely low sampling rate for Fourier single-pixel imaging via diffusion model. *Optics Express*, 2024, 32(3): 3138–3156. [doi: [10.1364/OE.510692](https://doi.org/10.1364/OE.510692)]
- 21 Keys R. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981, 29(6): 1153–1160. [doi: [10.1109/TASSP.1981.1163711](https://doi.org/10.1109/TASSP.1981.1163711)]
- 22 Wang XT, Yu K, Wu SX, *et al.* ESRGAN: Enhanced super-resolution generative adversarial networks. *Proceedings of the 2018 European Conference on Computer Vision*. Munich: Springer, 2019. 63–79.
- 23 Liang JY, Cao JZ, Sun GL, *et al.* SwinIR: Image restoration using Swin Transformer. *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021. 1833–1844.
- 24 Sun WJ, Chen ZZ. Learning many-to-many mapping for unpaired real-world image super-resolution and downscaling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(12): 9874–9889. [doi: [10.1109/TPAMI.2024.3428546](https://doi.org/10.1109/TPAMI.2024.3428546)]
- 25 Chen B, Zhang ZY, Li WQ, *et al.* Invertible diffusion models for compressed sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025, 47(5): 3992–4006. [doi: [10.1109/TPAMI.2025.3538896](https://doi.org/10.1109/TPAMI.2025.3538896)]
- 26 Wang Y, Liu Y, Zhao SJ, *et al.* CAMixerSR: Only details need more “attention”. *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2024. 25837–25846.

(校对责编: 张重毅)