Manuscript title - LATEXtemplate for Computers & Geosciences

Author 1*a*, Author 2*b* and Author 3*c*

aAuthor 1 affiliation
bAuthor 2 affiliation
cAuthor 3 affiliation

ARTICLE INFO

Keywords:
Keyword 1
Keyword 2
Keyword 3
Keyword 4

ABSTRACT

Abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here, abstract text here.

CRediT authorship contribution statement

Author 1: Author 1 contribution . **Author 2:** Author 2 contribution . **Author 3:** Author 3 contribution.

1. Introduction

2. Related Work

3. FRAMEWORK

3.1. Framework Description

Figure X shows the overall framework of the MSA-CycleGAN proposed in this paper in realizing the task of converting Micro-CT to SEM images.

The original CycleGAN consists of two generators G_{L2H} and G_{H2L} two discriminators D_L and D_H. The overall network adopts a cyclic structure to ensure that the consistency of the image structure and the reliability of physical properties can be maintained in the absence of paired labels.

However, CycleGAN only focuses on pixel distribution matching and lacks explicit modeling of the microstructure in rock images. This may cause structural authenticity problems such as void ambiguity, boundary fractures, and crack strike deviation, which will affect downstream geotechnical modeling and physical simulation tasks.

Moreover, CycleGAN does not consider the physical indicators of rock images, such as porosity, connectivity, particle size distribution, etc., and the generated map may be visually realistic but the geological attributes are distorted. Moreover, CycleGAN may be unstable in training and excessively strong discriminators. The pattern crashes, that is, multiple inputs are mapped to approximately the same output, reducing the diversity of generation. This problem is more likely to occur in scenes with a limited number of rock image samples.

Table 1
Summary of Notations and Definitions.

$Symbol$	$Purpose\,and\,Description$
x	Image samples from source domain
y	Image samples from the target domain
p	A probability density function that describes the probability distribution of the random variable
$G(x), G(y)$	Represents generator output to source domain image x or target domain image y
G_{A2B}	Generator from source domain to target domain
G_{B2A}	Generator from target domain to source domain
D	discriminator
$\mathcal{L}_{\text{GAN}}$	adversarial Loss
$\mathcal{L}_{\text{cyc}}$	cycle consistency loss
$\mathcal{L}_{\text{id}}$	identity loss
$\lvert \cdot \rvert_1$	L1 norm, used to measure absolute differences between images

In order to solve the above problems, inspired by the existing research on multi-scale discriminant strategies and structure sensing mechanisms in reducing representation differences between domains, this paper systematically introduces multi-scale discriminators, structure retention loss and attention mechanisms into the CycleGAN framework to achieve multiple fidelity of image texture details and geological structure information. We named this new network model MSA-CycleGAN.

Specifically, in order to enhance the generator's ability to maintain microstructure, MSA-CycleGAN introduces a self-attention mechanism in the generator to improve the feature selectivity of key structural areas such as crack boundaries and pore distribution through information interaction between channels., thereby improving structural expression accuracy and edge clarity. At the same time, in order to improve the model's perception of texture and structural features at different scales, the discriminator is constructed as a multi-scale PatchGAN structure. Its dual discrimination path can evaluate the structural authenticity of the image at the original and down-sampled scales in parallel, thereby enhancing the model's consistent modeling capabilities for local details and global semantics. In order to further improve the response ability of the generator to key areas of the structure, a space-channel self-attention mechanism is introduced within the network, which effectively enhances the modeling selectivity of image boundaries and structural information, thereby improving the accuracy of structural expression and texture contrast.

As the training iteration deepens, the generator's characterization capabilities on structural domains at different scales continue to be enhanced, ultimately making the output image not only visually close to the real SEM image, but also highly consistent in physical properties such as porosity and connectivity. This method significantly improves the fidelity and interpretability of cross-domain transformation of geological images without relying on pixel paired data. We provide a summary of the symbols and their definitions in Table 1.

3.2. Generator with Multi-scale Feature Fusion

Due to the typical multi-scale characteristics of pores, cracks and texture structures in the image, the generated images face major challenges. Micro-scale pores usually appear as high-frequency, low-contrast edge textures, while millimeter-scale cracks show significant gray scale jumps and spatial extensibility. Due to the fixed receptive fields of traditional convolution operations, it is often difficult to simultaneously capture such structural information with significant scale differences, which may easily lead to structural dislocation and blurred boundaries in the generated image.

In order to improve the network's ability to model complex geological structures, multiple residual blocks are introduced in the encoding and decoding stages. Each residual block contains two 3×3 convoluted layers and ReLU activation functions, which directly transfers input features through residual connections, effectively alleviating the gradient disappearance problem and enhancing deep feature expression capabilities. This module can extract local textures and global structures at different scales, improving the consistency and detail integrity of the generated image.

Therefore, a multi-scale feature fusion mechanism is introduced into the generator design to improve the model's ability to model cross-scale spatial structures and enhance the structural expression and detail restoration performance of rock images. Specifically, a multi-scale generator architecture integrating residual blocks and attention modules is constructed. Through hierarchical feature extraction and contextual information fusion, effective modeling and high-fidelity reconstruction of multi-scale structural elements in rock images are achieved.

In order to enhance the model's ability to focus on key areas, an attention module is embedded in the intermediate feature extraction stage. This module dynamically adjusts the importance of each position in the feature map by calculating channel attention and spatial attention, thereby guiding the generator to pay attention to key structural areas in the image.

During the feature extraction stage, the model successively stacks multiple residual modules, and introduces a self-attention mechanism in each module to effectively improve the model's response to pore edges and crack boundaries by explicitly modeling the spatial dependencies within the feature map. Responsiveness to key areas such as crack boundaries.

In order to enhance the model's global modeling capabilities for key areas, during the feature extraction stage, the model successively stacks multiple residual modules, and introduces a self-attention mechanism in each module to improve modeling of long-distance dependencies and cross-area textures. Specifically, for the input feature map $X \in \mathbb{R}^{C \times H \times W}$, the Query $Q \in \mathbb{R}^{C' \times N}$, Key $K \in \mathbb{R}^{C' \times N}$ and Value $V \in \mathbb{R}^{C \times N}$ are obtained through 1×1 convolution mapping respectively. Where $C' = C/8, N$ is the number of pixels after the spatial dimension is expanded. Subsequently, the attention matrix $A \in \mathbb{R}^{N \times N}$ is obtained through matrix multiplication calculation,The (i, j) item represents the attention weight of the i position to the j position in the input feature map, which is used to measure

the similarity between the two in the feature space. Finally, each line is normalized through the *softmax* function to ensure that the sum of all attention weights is 1. A is obtained by Formula X.

$$A = Softmax(Q^T K) \tag{1}$$

Then, by applying this attention matrix A to the value matrix V, The weighted feature response is obtained, and the attention output feature map F_{att} is obtained from Formula X.

$$F_{att} = V A \tag{2}$$

Therefore, the final output feature at each location is the result of a weighted sum of the feature content of all locations, with the weights determined by the attention matrix A. That is, when the model generates new features for the current pixel, it will synthesize information from all other pixels in the reference map and determine the degree of influence based on their similarities, thereby improving the network's ability to model image structure and remote dependence.

Finally, it is reshaped back to the original feature map size, and residual fusion is performed with the input feature through a learnable scaling factor. The fused output feature map F' is shown in Formula X.

$$F' = \gamma \cdot F_{att} + F \tag{3}$$

Where γ is a learnable scaling factor used to control the intensity of the attention branch.

For each position in the input feature map, a set of attention weights is calculated by learning its correlation with all other positions, thereby dynamically adjusting the feature contributions of different regions. This mechanism can break through the limitations of the local receptive field of traditional convolution kernels and focus more on the expression of key areas while maintaining the original structural information.

3.3. Multi-scale Discriminator

In view of the fact that traditional single-scale discriminators usually use fixed-size image blocks as the discriminant unit, mainly focus on the authenticity judgment of local texture, and it is difficult to fully capture cross-region structural information, this paper introduces a multi-scale discriminant mechanism in the design of the discriminant. Considering that structural features such as pore connectivity and fracture networks in rock images have obvious multi-scale distribution characteristics, a single scale often cannot fully model global structural consistency. To this end, this

paper proposes a dual-resolution discrimination structure, in which the original resolution image and its down-sampled version are input simultaneously in the discrimination stage, and the authenticity judgment is independently carried out by two discriminators. Among them, the high-resolution branch focuses on the identification of detail textures, while the low-resolution branch focuses more on the overall rationality of the global structure. The discriminant outputs at the two scales are jointly used for the generator's counter loss calculation, thereby improving the model's comprehensive evaluation ability of image authenticity from different levels. This design effectively enhances the model's ability to model multi-scale structures in rock images, and improves structure restoration and artifact suppression performance.

The main branch discriminator receives raw image input with a resolution of 1024×1024 and uses the PatchGAN framework for local discrimination. Specifically, the main branch uses several convoluted layers to gradually down-sample the image and finally outputs a 64×64 discriminant feature map. The secondary branch aims at global structure perception. First, the input image is downsampled by average pooling at 2×2, reduced its size to 512×512, and finally outputs a discriminant map with a size of 32×32. This branch focuses on structural consistency in larger receptive fields, which helps make up for the main branch's limitations in macro-structure modeling. The final discrimination result is calculated from the discrimination maps output by the main and auxiliary branches respectively, and is subjected to equal weight weighted average to guide the training of the generator, so that it can take into account the local realism of the image and the overall structural consistency.

Let the input image be $x \in \mathbb{R}^{C \times H \times W}$, the main discriminator extracts the high-resolution texture from the original image as $D_{\text{main}}(x)$, uses two-dimensional average pooling to down-sample the input image to obtain $x_{\downarrow}$, and the auxiliary discriminator processes the down-sampled image to capture the global structure; if $D_{\text{down}}(x_{\downarrow})$, the fusion output $D_{\text{final}}(x)$ is obtained as the overall discriminator result is obtained from Formula X.

$$D_{\text{final}}(x) = \frac{1}{2}(D_{\text{main}}(x) + D_{\text{down}}(x_{\downarrow})) \tag{4}$$

The multi-scale discriminant structure not only enhances the perception of image global structure distortions, edge discontinuities and texture blurring, but also provides cross-scale and consistent semantic and detail feedback for the generator, thereby effectively improving the spatial structure of the generated image. Stability and quality. At the same time, while ensuring computing efficiency, this design significantly improves the discriminator's ability to distinguish complex texture and structural features in rock images, and provides a more effective training signal for generator optimization.

3.4. Upsampling module optimization

In the process of restoring images from low resolution to high resolution, traditional generation networks often use deconvolution for upsampling. However, previous studies have shown that deconvolution can easily produce checker-

board artifacts when stacked in multiple layers, especially when processing complex pore and boundary structures in rock images. This artifact can significantly affect the structural authenticity and discriminator stability of the image.

To this end, we optimized the structure of the upsampling module in the generator, and introduced an upsampling module design that combines interpolation fusion and residual connection into the generator structure. Specifically, each level of image upsampling operation uses nearest neighbor interpolation to double the feature map in spatial dimensions, followed by a convolution layer to re-extract details and compress channel dimensions. This design structurally avoids the uneven overlap problem inherent in deconvolution, while allowing smooth transition of features and retaining local texture.

In addition, in order to enhance the transfer of context information during the upsampling stage, this paper introduces a lightweight residual connection structure after upsampling at each layer, that is, an identity mapping is applied before and after convolution output, so that the original features can be directly transferred to subsequent layers without convolution transformation, thereby retaining mid-scale information to a certain extent and enhancing the continuity and edge clarity of high-resolution output images. Formula X represents the overall sampling process

$$\mathbf{y}_{i+1} = \text{Conv}\left(\text{Up}(\mathbf{y}_i)\right) + \mathbf{y}_i^{\uparrow} \tag{5}$$

where $\mathbf{y}_i$ denotes the input feature map at stage i, $\text{Up}(\cdot)$ is the nearest-neighbor interpolation, and $\mathbf{y}_i^{\uparrow}$ is the upsampled identity shortcut.

In the upsampling stage of the generator, the model is equipped with a total of four levels of upsampling modules, which are used to gradually restore the resolution of 64×64 in the self-coding stage of the feature map to the final 1024×1024. Each level of upsampling consists of two parts: First, the spatial size of the input feature map is expanded to twice the original through nearest neighbor interpolation, and then a layer of two-dimensional convolution is added to refine local features. This convolution operation uses a 3×3 convolution kernel, and the steps in the horizontal and vertical directions are 1, and the filling of pixels at the edges is 1, thus effectively enhancing the continuity and texture expression capabilities of context information without changing the space size. The number of output channels of the convolutional layer corresponding to the four-level upsampling is 6432168, and the number of channels is reduced step by step to compress feature dimensions and enhance semantic expression capabilities. This design not only ensures the continuous restoration of the image space size, but also effectively avoids Checkerboard artifacts common in deconvolution operations, ensuring the stability of the image structure and texture during the upsampling process.

Overall, this module design not only improves the stability of upsampling, but also effectively enhances the model's ability to restore image structure and edge information at different scales. It is especially suitable for high-fidelity

reconstruction tasks of microstructures and fractures in rock images.

3.5. Extended loss function for structural and edge fidelity

In order to achieve high-fidelity rock image conversion and structural reconstruction, this paper systematically expands and customizes the traditional CycleGAN loss function architecture, and constructs a multi-objective collaborative optimization strategy suitable for geological image super-resolution reconstruction tasks. Considering the high sensitivity of rock images to microscopic information such as pore connectivity, fracture structure and boundary topology in engineering applications, this paper further introduces identity mapping loss based on the resistance loss and cyclic consistency loss to improve the model's ability to express structure and maintain physical attributes.

First, generator training still focuses on the Least Squares Form of Adversarial Loss (LSGAN). The goal is to make the generated high-resolution SEM image indistinguishable from the real image before the discriminator, thus driving the generator to learn the target domain. Texture distribution and appearance features. Compared with traditional cross-entropy loss, LSGAN can provide smoother gradient feedback, enhance the stability of the model training process and mitigate pattern collapse. To measure the confrontational performance of generators and discriminators in the target domain, The adversarial loss function $\mathcal{L}_{GAN}(G, D)$ is defined as Equation X.

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{x \sim p_{data}(x)}[(D(x) - 1)^2] + \mathbb{E}_{y \sim p_{data}(y)}[(D(G(y)))^2] \tag{6}$$

Secondly, in order to maintain the geometric consistency of the image during cross-domain conversion, the model introduces a cyclic consistency loss term, which means that when the image is mapped from the source domain to the target domain and then mapped back to the original domain, the reconstruction result should be as consistent as the input image as possible,so we use cyclic consistency loss to encourage this behavior.This loss achieves closed-loop preservation of semantic content and structural boundaries through pixel-level L1 norm constraints. The cyclic consistency loss function $\mathcal{L}_{\text{cyc}}$ is defined as Equation X.

$$\mathcal{L}_{\text{cyc}}(G_{A2B}, G_{B2A}) = \mathbb{E}_{x \sim p_{\text{data}}(x)} \left[\|G_{B2A}(G_{A2B}(x)) - x\|_1 \right] + \mathbb{E}_{y \sim p_{\text{data}}(y)} \left[\|G_{A2B}(G_{B2A}(y)) - y\|_1 \right] \tag{7}$$

In addition, in order to suppress unnecessary transformation behavior of the generator on the target domain image, an identity mapping loss term is added to the model. When the generator receives an image of the target domain as input, its output should be approximate to the input image itself. This regular term helps stabilize the model learning process and avoid texture distortions and structural mismatches caused by excessive migration. The identity map loss

189    $\mathcal{L}_{id}$ is defined as equation X

$$\mathcal{L}_{id}(G_{A2B}, G_{B2A}) = \mathbb{E}_{y\sim p_{data}(y)}\left[\|G_{A2B}(y) - y\|_1\right] + \mathbb{E}_{x\sim p_{data}(x)}\left[\|G_{B2A}(x) - x\|_1\right] \tag{8}$$

190   To sum up, the constructed composite loss function not only inherits CycleGAN's unsupervised style migration

191 capabilities, but also integrates the structural priors and physical constraints required for geological image modeling,

192 providing a basis for realizing rock image reconstruction with engineering interpretation. Theoretical basis and training

193 guarantee.

4. EXPERIMENTAL RESULTS

4.1. Data sets and pretreatment

196   The experimental dataset used in this study was obtained from the publicly available Zenodo platform (DOI:

197 10.5281/zenodo.6420100), published by Liu et al. in 2022. It was designed to support research on multiscale dig-

198 ital rock image fusion using deep generative adversarial networks (GANs), and primarily contains high-resolution

199 grayscale digital rock images derived from micron-scale CT scans, encompassing a wide range of geological struc-

200 tures and scale features.

201   The dataset is composed of two main components: low-resolution (LR) images and learned style kernels (LSK),

202 along with a reference pretrained StyleGAN checkpoint file (network-snapshot-stylegan.pkl). The LR images reflect

203 the structural features of porous rocks at mesoscopic scales, suitable for coarse-scale analysis, while the LSK encodes

204 high-frequency textural features extracted from high-resolution images, providing a foundation for multiscale texture

205 modeling.

206   All images are stored in standard image formats and have been systematically organized to meet the requirements of

207 supervised learning tasks. The dataset is widely applicable to digital rock physics, geomechanical modeling, and deep

208 learning-based super-resolution or scale conversion tasks. In this study, a subset of the data was selected for training

209 and validating the proposed model for translating low-resolution micro-CT images into high-resolution SEM-style

210 images.

4.2. Comparison with state-of-the-art methods

212   In this study, we systematically evaluated the effectiveness of the proposed style-based data augmentation mecha-

213 nism in enhancing image generation quality and structural reconstruction performance. To ensure a fair comparison,

214 we reproduced and referenced several existing methods based on the same CycleGAN backbone, and conducted perfor-

215 mance comparisons on a unified dataset using consistent evaluation metrics, as shown in Tables IV and V. The results

216 demonstrate that, for the task of Micro-CT to SEM image translation, our method significantly outperforms baseline

Table 2
Performance comparison table of image generation and structure reconstruction methods.

	$ssim$	$PSNR$	$porosity error$
$baseline$	0.9304	22.61 dB	0.0973
$style500$	0.9912	36.58 dB	0.0047
$style400$	0.9879	30.26 dB	0.0067
$style300$	0.9893	34.34 dB	0.0109
$style200$	0.9861	31.81 dB	0.0057
$style200$	0.9875	33.36 dB	0.0066

217 models by incorporating StyleGAN-generated style samples into the training process.

218   Specifically, the model trained solely on original image data showed limitations in terms of SSIM, PSNR, and

219 porosity error. When StyleGAN-generated samples were integrated into the training dataset for augmentation, the

220 model's performance improved significantly. In particular, under the configuration with 500 style samples, the model

221 achieved optimal results, with SSIM reaching 0.9911, PSNR increasing to 36.58 dB, and porosity error decreasing to

222 0.0047substantially surpassing the results obtained without data augmentation. By contrast, the highest SSIM achieved

223 by baseline methods under the same setting was 0.985, while the standard CycleGAN model yielded an average SSIM

224 of only 0.974, confirming the advantage of the proposed style augmentation approach in enhancing visual quality and

225 structural fidelity.

226   Further analysis revealed that the introduction of style samples led to superior reconstruction of complex fracture

227 networks and multi-scale pore geometries, especially in terms of edge continuity and fine detail sharpness. Addition-

228 ally, we introduced an adaptive training strategy that progressively expands the coverage of the style space, enabling

229 the model to learn more generalized and diverse distributions. Overall, the synergy between the style augmentation

230 mechanism and structure-preserving objectives contributes to improved balance between visual realism and geological

231 interpretability, enhancing the models potential for practical geotechnical applications.

4.3. Ablation Study

233   Table X presents the results of ablation experiments conducted to evaluate the impact of key components in the

234 proposed model, including the discriminator architecture, attention mechanism, and different upsampling strategies.

235 The performance of image generation was assessed using three metrics: SSIM, PSNR, and porosity error, reflecting

236 the structural similarity, visual quality, and geological consistency of the generated images.

237   In the baseline model employing a single-scale discriminator, the SSIM reached 0.9731, PSNR was 32.11 dB, and

238 porosity error was 0.0098, indicating strong structural reconstruction capability. When the discriminator was replaced

239 with a multi-scale structure, a slight drop in SSIM to 0.9377 was observed, along with a notable decline in PSNR to

240 20.23 dB and an increase in porosity error to 0.038. These results suggest that multi-scale feature fusion, without the

241 support of auxiliary mechanisms, may interfere with image quality.

Table 3
Ablation Study Results.

	$ssim$	$PSNR$	$porosity error$
$single-scale discriminator$	0.9731	32.11 dB	0.0098
$no attentio$	0.9377	20.23 dB	0.038
$upsample deconv$	0.9784	35.88 dB	0.0066
$upsample nearest$	0.9723	26.67 dB	0.0115
$upsanple shuffle$	0.9839	34.87 dB	0.0071

Regarding the upsampling module, three strategies were compared: transposed convolution (deconv), nearest-neighbor interpolation (nearest), and shuffle upsampling. The results show that transposed convolution yields the highest image quality, achieving an SSIM of 0.978, PSNR of 35.88 dB, and the lowest porosity error of 0.0066, indicating its effectiveness in detail reconstruction. In contrast, the nearest-neighbor method performed relatively poorly, with SSIM dropping to 0.972, PSNR to 26.67 dB, and porosity error increasing to 0.0115, suggesting limitations in preserving fine details. The shuffle upsampling strategy achieved a more balanced performance across the three metrics, with SSIM reaching 0.984, PSNR at 34.87 dB, and porosity error at 0.0071, demonstrating superior structural restoration and edge clarity.

Overall, the analysis reveals that the choice of upsampling strategy significantly influences model performance. Transposed convolution and shuffle upsampling both demonstrate advantages in terms of visual quality and structural fidelity. In contrast, the design of the discriminator requires coordination with contextual mechanisms to avoid excessive noise introduced by multi-scale discrimination. Collectively, the selected module combinations in this study achieved a favorable balance across multiple evaluation metrics, highlighting the importance of each component in generating high-quality rock images.

5. Algorithm and implementation

References

Almahairi, A., Rajeshwar, S., Sordoni, A., et al., 2018. Augmented cyclegan: Learning many-to-many mappings from unpaired data, in: ICML, pp. 195–204.

Bermano, A.H., Gal, R., Alaluf, Y., et al., 2022. Stateoftheart in the architecture, methods and applications of stylegan. Computer Graphics Forum 41, 591–611.

Bizhani, M., Ardakani, O.H., Little, E., 2022. Reconstructing high fidelity digital rock images using deep convolutional neural networks. Scientific Reports 12, 4264.

Cao, D., Hou, Z., Liu, Q., et al., 2022. Reconstruction of three-dimension digital rock guided by prior information with a combination of infogan and style-based gan. Journal of Petroleum Science and Engineering 208, 109590.

Chen, G., Zhang, G., Yang, Z., et al., 2023. Multi-scale patch-gan with edge detection for image inpainting. Applied Intelligence 53, 3917–3932.

Chen, Y., Lin, H., Zhang, W., et al., 2024. Icycle-gan: Improved cycle generative adversarial networks for liver medical image generation. Biomedical

268   Signal Processing and Control 92, 106100.

269   Chi, P., Sun, J., Zhang, R., et al., 2024. Digital rock reconstruction enhanced by a novel gan-based 2d-3d image fusion framework. Advances in
270   Water Resources 193, 104813.

271   Jia, H., Cai, W., Huang, H., et al., 2022. Learning multi-scale synergic discriminative features for prostate image segmentation. Pattern Recognition
272   126, 108556.

273   Karras, T., Laine, S., Aittala, M., et al., 2020. Analyzing and improving the image quality of stylegan, in: CVPR, pp. 8110–8119.

274   Li, X., Li, B., Liu, F., et al., 2023. Advances in the application of deep learning methods to digital rock technology. Advances in Geo-Energy
275   Research 8, 5–18.

276   Qin, X., Xia, Y., Wu, J., et al., 2022. Influence of pore morphology on permeability through digital rock modeling: new insights from the euler
277   number and shape factor. Energy & Fuels 36, 7519–7530.

278   Reinhardt, M., Jacob, A., Sadeghnejad, S., et al., 2022. Benchmarking conventional and machine learning segmentation techniques for digital rock
279   physics analysis of fractured rocks. Environmental Earth Sciences 81, 71.

280   Yang, Y., Liu, F., Zhang, Q., et al., 2023. Recent advances in multiscale digital rock reconstruction, flow simulation, and experiments during shale
281   gas production. Energy & Fuels 37, 2475–2497.

282   Zhang, T., Ji, X., Lu, F., 2022. 3d reconstruction of porous media by combining scaling transformation and multi-scale discrimination using
283   generative adversarial networks. Journal of Petroleum Science and Engineering 209, 109815.

284   Zhao, J., Wang, F., Cai, J., 2021. 3d tight sandstone digital rock reconstruction with deep learning. Journal of Petroleum Science and Engineering
285   207, 109020.

286   Zhong, G., Ding, W., Chen, L., et al., 2023. Multi-scale attention generative adversarial network for medical image enhancement. IEEE Transactions
287   on Emerging Topics in Computational Intelligence 7, 1113–1125.