Generation of Heterogeneous Pore-Space Images Using Improved Pyramid Wasserstein Generative Adversarial Networks

Linqi Zhu[1,2,*]**, *Branko Bijeljic***[1]**, *Martin J. Blunt***[1]

[1]Department of Earth Science and Engineering, Imperial College London, SW7 2BP, London, United Kingdom
[2]Key Laboratory of Marine Geophysics and Georesource, Institute of Deep-sea Science and Engineering, Chinese Academy of Sciences, 572000, Sanya, China

Key Points:

- The improved GAN generates random heterogeneous pore-space images of arbitrary size.
- The method reproduces geometric and flow properties of images of five heterogeneous sandstones and carbonates.
- The method captures pore-space features over a range of length scales through a pyramid generator structure, among other improvements.

Author Contributions:

Linqi Zhu: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Software, Validation, Writing - Original Draft, Writing - Review & Editing.

Branko Bijeljic: Project Administration, Supervision, Data Curation, Writing - Review & Editing.

Martin Blunt: Conceptualization, Data Curation, Project Administration, Supervision, Resources, Writing - Review & Editing.

ORCID:

Linqi Zhu - 0000-0001-5725-2805

Branko Bijeljic - 0000-0003-0079-4624

Martin Blunt - 0000-0002-8725-0250

Corresponding author: Linqi Zhu, `linqi.zhu@imperial.ac.uk, zhulq@idsse.ac.cn`

Abstract

We use Wasserstein Generative Adversarial Networks to learn and integrate multi-scale features in segmented three-dimensional images of porous materials, enabling the dependable generation of large-scale representations of complex porous media. A Laplacian pyramid generator is introduced which creates pore-space features with a hierarchy of spatial scales. Feature statistics mixing regularization enhances the ability of the model generation to reliably maintain multi-scale pore-space features of images by increasing diversity. The method is tested on a variety of X-ray images of porous rocks. The generated images can be of any size – cm-scale ten-billion-cell images are generated to demonstrate the power of the approach – which have two-point correlation functions, porosity, permeability, Euler characteristic, curvature, and specific surface area close to the training datasets. The images demonstrate a considerable improvement over previously-published studies using generative adversarial networks.

Plain Language Summary

High-resolution three-dimensional X-ray microscopy can be used to image the pore space of porous materials. Machine learning algorithms trained on billion-voxel datasets can generate a statistical ensemble of representative images of arbitrary size, for rock characterization, modelling and analysis. However, current methods cannot easily capture the features at different spatial scales seen in many heterogeneous rocks. We develop an improved multi-scale neural network that automatically reproduces features of the pore space of different size. By testing on five sandstone and carbonate samples we show that the generated images accurately capture geometric and flow properties and are superior to those produced using existing machine learning techniques. Finally we demonstrate the power of the method by generating ten-billion voxel images, representing a cm-sized sample with almost micron resolution.

1 Introduction

Three-dimensional imaging, typically using X-rays, has revolutionized our understanding of flow in porous media, since both the pore structure and the fluids within can be seen in otherwise opaque materials. This technology has been applied to study water resources, fuel cells, carbon dioxide sequestration, hydrogen storage, and the recovery of oil and natural gas (Chai et al., 2022; Wang et al., 2022; Jiang et al., 2023; Wang et al., 2023; Li et al., 2023). In imaging there is an inevitable trade-off between resolution which has to be sufficiently high to resolve the pore space accurately, and field of view to capture a representative volume of the medium on which macroscopic properties can be observed or calculated (Saxena et al., 2018; Wang et al., 2021; Sadeghnejad et al., 2023).

The statistical generation of pore-space images through various numerical methods, can potentially achieve both high resolution and the required field of view (FoV) (Liu & Mukerji, 2022). Traditional solutions capture spatial statistical information, or simulate sedimentary processes to create random porous media. Until recently the state-of-the-art was multiple point statistics that had been successfully applied to generate three-dimensional pore-space images (Okabe & Blunt, 2004).

In recent years, the rapid development of machine learning, especially generative models, combined with the wide availability of good-quality three-dimensional data for training, has transformed our ability to generate images. Mosser et al. (2017, 2018) were the first to reconstruct random porous media, including a sandstone, a carbonate and beadpacks, based on Generative Adversarial Networks (GANs). Zha et al. (2020) employed Wasserstein GANs with a gradient penalty to generate images of shale. Liu et al. (2019) combined data augmentation techniques with Deep Convolution GANs (DC-

GANs). Shams et al. (2020) integrated GANs with autoencoders to generate sandstones with intra-granular pores. Zhao et al. (2021) applied a DCGAN for the three-dimensional reconstruction of tight sandstones. Zhang et al. (2022) incorporated convolutional neural networks (CNNs) into a traditional GAN, forming a method to quickly and accurately produce large-scale three-dimensional porous media. However, adversarial learning methods such as DCGAN find it challenging to train networks to generate complex random porous media with features over a range of length scales.

Current challenges in pore-space generation are: ① how to ensure that multi-scale heterogeneous pore features are learnt during the training phase; and ② after obtaining a plethora of multi-scale features, how to enhance control over the training process. These two aspects determine whether the model can provide reliable multi-scale generation. In this letter, we propose a new method based on an improved generative adversarial network to address these issues. In terms of multi-scale feature learning, it introduces a pyramid structure into the generator. During the training process, it captures different scale features through cooperative training between layers of different sizes, generating images in a coarse-to-fine manner. Subsequently, it incorporates feature statistics mixing regularization (FSMR) to strengthen the constraints of the discriminator, preventing the generative model from being sensitive to specific pore features, thereby enhancing the model's generative diversity. Furthermore, to better integrate multi-scale features and improve model training stability to obtain a reliable model, the Wasserstein GAN- Lipschitz Penalty (LP) is added to the model. The new model is termed the Improved Pyramid Wasserstein GAN (IPWGAN). Leveraging these advantages, we offer a superior solution for the rapid and reliable generation of large-scale complex porous media, be it heterogeneous sandstones or carbonates.

2 Methodology

A GAN consists of two networks: a generator and a discriminator. The generator learns how to produce samples, while the discriminator aids the training of the generator by determining the authenticity of the generated images. For the generator, we have incorporated a simplified pyramid structure to the conventional model, facilitating the extraction of multi-scale pore features and the connectivity between different types and sizes of pores. For the discriminator, we have employed the feature statistics mixing regularization method to encourage diversity in the generated samples. This prevents the model from focusing solely on textures, which could lead to a singular pattern that fails to replicate the diverse and complex pore spaces found in real multi-scale heterogeneous porous media. Additionally, we opted to use the WGAN-LP loss function in place of the traditional GAN loss function to enhance training stability, diversify the images generated, and prevent gradient vanishing and explosion issues. The proposed workflow is illustrated in Figure 1.

2.1 Pyramid generator structure

In GANs, the generator's performance directly determines the quality of image generation. For image generation tasks, traditional generators typically upscale a random noise vector through multiple deconvolution operations until it matches the size of the training samples. Throughout the training process, the generator is trained as a continuous, integrated network. In contrast, the pyramid-style generator adopts a hierarchical structure. In this design, each layer of the network is connected in series, with each layer specifically handling data of a certain resolution. For instance, the bottom layer mainly processes high-resolution data, while the top layer deals with low-resolution data. This layered approach allows the pyramid GAN to capture the multi-scale characteristics of the data more accurately, resulting in higher quality image generation. Currently, there are various pyramid-style generators available, such as the Laplacian pyramid (Denton et al., 2015) and the dual pyramid (Li et al., 2022). In this study, we chose the Lapla-

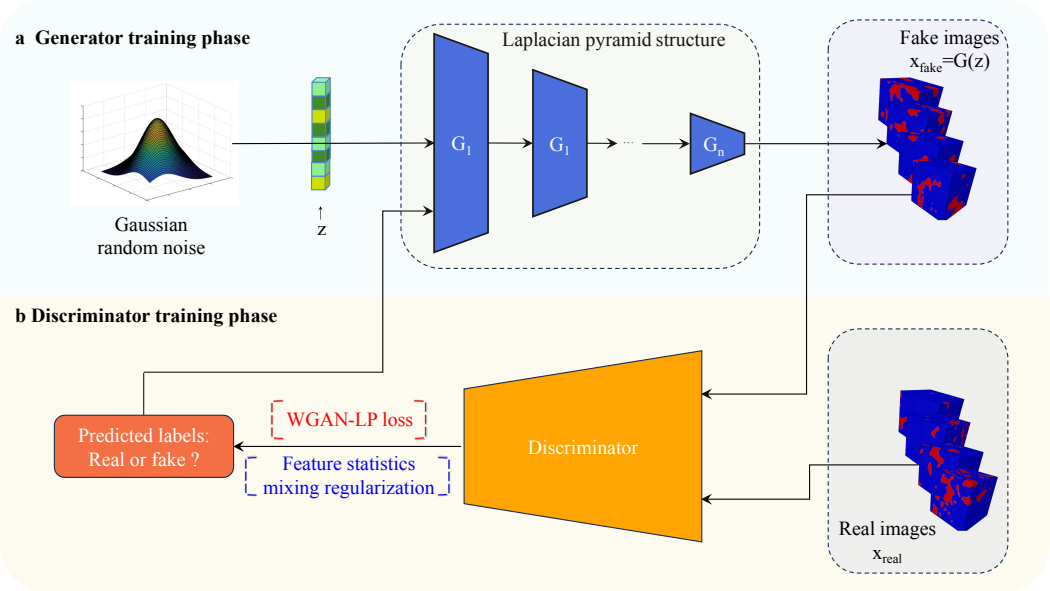


Figure 1. Proposed workflow for generating artificial images of porous media, which includes the incorporation of a Laplacian pyramid structure in the generator. a. Architecture and training method of the generator. The pyramid generator is generated layer by layer and can be regarded as multiple independent modules. It extracts from the latent high-dimensional vector z, composed of Gaussian random noise, and approaches the features of real three-dimensional digital rock data through layer-by-layer training to obtain the fake images. b. Architecture and training method of the discriminator with WGAN-LP loss and feature statistics mixing regularization. An equal number of fake images and real images will be labelled and input into the discriminator. Through training, it tries to distinguish between fake and real images. In addition to the error of the WGAN that needs to be calculated, the results of the LP penalty term used to control training smoothness and convergence, and the FSMR regularization term used to control sample diversity are added to the total error. Based on the error, the difference between fake and real is judged, and the error is fed back to the generator to guide the error backpropagation training. The generation phase is similar to the generator training phase, but it does not perform error backpropagation. It only takes vectors as input to obtain realistic fake images. The size of the fake images is determined by the size of the input vector z.

cian pyramid as the foundation and simplified it by omitting the step of calculating losses separately for each layer to accelerate training. This simplification ensures the model's lightweight nature, allowing it to be trained on a single GPU.

The Laplacian pyramid is composed of a set of generator networks. Each generator, when constructed, not only utilizes random noise but also receives feedback from the downsampled output of the previous generator. This approach is used to stack and learn multi-scale features. Assuming there is a set of generator models $G_0, \ldots, G_K$, the input for any generator model $(1<k\leq K)$ is:

$$\tilde{I}_k = d(\tilde{O}_{k-1}) + \tilde{z}_k = d(\tilde{G}_{k-1}(\tilde{I}_{k-1})) + \tilde{z}_k \tag{1}$$

In the formula, $\tilde{I}_k$ represents the input for the k^{th} generator model G_k. The function $d(\cdot)$ refers to the downsampling operation. $\tilde{O}_{k-1}$ is the output of the $k{-}1^{\text{th}}$ gen-

erator model $\tilde{G}_{k-1}$, and $\tilde{z}_k$ is the corresponding random noise vector for $\tilde{G}_k$. Therefore, the Laplacian pyramid generator consists of networks that are trained independently but share information with adjacent layers. This approach ensures the targeted acquisition of multi-scale information from images. Changes in the generator do not affect the training process of the discriminator. In other words, the discriminator still uses the fake images output by the last generator as its input.

2.2 Mixing and regularization of feature statistics

Unlike the generator, considering memory usage during training and ensuring the diversity of generated samples (which is crucial for complex porous media, as strong heterogeneity leads to diverse pore structure features with multiple scales), we chose the feature statistics mixing regularization (FSMR) discriminator (Kim et al., 2022). Typically, discriminators are highly sensitive to the texture of images (for digital rock core images, this refers to features such as pore shapes), which to some extent reduces the diversity of image generation. FSMR encourages the discriminator's predictions to remain invariant to the style of the input image. This is achieved by using the mean and standard deviation of intermediate feature maps mixed from another training sample within the discriminator and penalizing the difference between the original and mixed outputs.

The specific operation involves generating mixed features of the original image and the reference image in the discriminator's feature space. Feature statistics mixing (FSM) is defined for feature maps x with respect to feature maps y as AdaIN followed by linear interpolation:

$$\mathrm{FSM}(x, y) = \alpha x + (1 - \alpha)\mathrm{AdaIN}(x, y) = \alpha x + \frac{\sigma(y)(x - \mu(x))}{\sigma(x)} + \mu(y) \tag{2}$$

where $\alpha \sim \mathrm{Uniform}(0, 1)$ controls the intensity of feature perturbation. AdaIN is adaptive instance normalization, $x, y \in \mathbb{R}^{C\times D\times H\times W}$ are features obtained by a pretrained encoder, and $\mu(\cdot)$ and $\sigma(\cdot)$ denote their mean and standard deviation, calculated for each channel, respectively. Assuming the i^{th} layer of the discriminator is f_i, the content feature image is represented as c, and the style feature image is represented as s (randomly selected from the current training batch samples), the mixed feature maps $\tilde{x}$ and $\tilde{y}$ are:

$$\begin{cases} \tilde{x}_1 = f_1(c) \\ \tilde{y}_1 = f_1(s) \\ \tilde{x}_{i+1} = f_{i+1}(\mathrm{FSM}(\tilde{x}_i, \tilde{y}_i)) \\ \tilde{y}_{i+1} = f_{i+1}(\mathrm{FSM}(\tilde{y}_i, \tilde{x}_i)) \end{cases} \tag{3}$$

The calculation method for FSM is presented in Kim et al. (2022). The corresponding penalty term is computed based on the loss:

$$L_{\mathrm{FSMR}} = E_{c,s\sim p_{\mathrm{data}}}\left[(D(c) - D_{\mathrm{FSM}}(c, s))^2\right] = E_{c,s\sim p_{\mathrm{data}}}\left[(D(c) - \mathrm{Linear}(\tilde{x}_n))^2\right] \tag{4}$$

where n refers to the n^{th} layer of the discriminator, original output is represented by $D(c)$, and mixed output is represented by $D_{\mathrm{FSM}}(c, s)$.

2.3 WGAN-LP

The use of the JS divergence in the original objective function of DCGAN can lead to various problems, such as a vanishing gradient, and poses very high requirements for control during the training process (Mosser et al., 2017). WGAN was proposed to address the instability in the training of the original GAN. It offers a solution from the perspective of optimizing the objective function. However, the training of WGAN requires the function to be within the space of 1-Lipschitz functions. There are many schemes to achieve this condition, such as weight clipping, spectral normalization, gradient penalty, L_p normalization (L_p-norm), and so on. Due to the small and relatively stable penalty

weight required by (L_p-norm) normalization, it is used in this paper, resulting in WGAN-LP (Zhou et al., 2018).

The combination of WGAN-LP and FSMR loss forms the final loss function:

$$L = \min_G \left\{ \max_{\|D\|_L \leq 1} \left[E_{x\sim p(x)}[D(x)] - E_{z\sim q(z)}[D(G(z))] \right] + \lambda_1 E_{\hat{x}\sim p(\hat{x})} \left[\left(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1 \right)^2 \right] + \lambda_2 L_{\mathrm{FSMR}} \right\} \tag{5}$$

$E_{x\sim p(x)}[D(x)]$ represents the expected output of the real sample x on the discriminator $E_{z\sim q(z)}[D(G(z))]$ represents the expected output of the fake sample produced by the generator G on the discriminator D. λ_1 is the weight coefficient for L_p-norm normalization, which is used to control the importance of the penalty term. In this study, it is recommended to be set to 5. $\hat{x}$ is an interpolation between the real and the generated sample to ensure that the discriminator is Lipschitz continuous throughout the entire data space. λ_2 is the weight coefficient for FSMR. In this study, it is recommended to be set to 10. A loss function with multiple constraints can achieve more reliable results, making it suitable for adapting to the complexity of the pore space in heterogeneous rock images faced in the reconstruction task. In addition, some other means to improve the training effect are used, such as the generator being updated once after 5 discriminator updates to ensure that the discriminator is effective each time the generator is updated.

3 Data and results

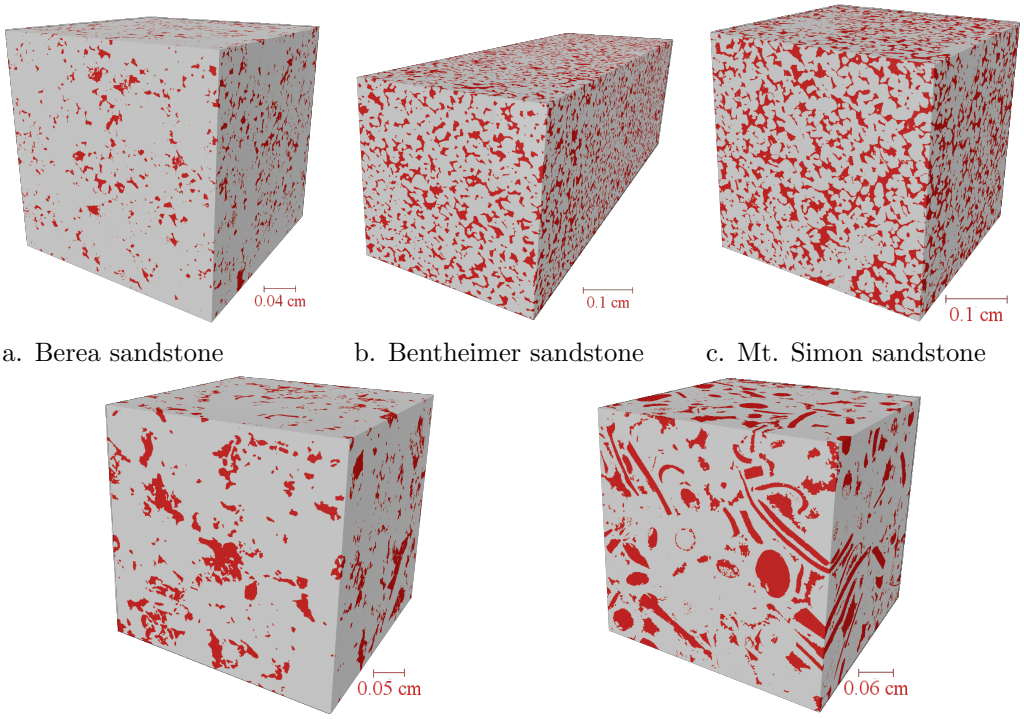
3.1 Data preparation and training

We selected dry images of three sandstones and two heterogeneous carbonate rocks as test data, Figure 2. The sandstones are Berea sandstone (1200^3 voxels, voxel size: 2.0 μm) (Garfi et al., 2020), Bentheimer sandstone ($1000^2\times3000$ voxels, voxel size: 3.58 μm) (Lin et al., 2018), and Mt. Simon sandstone (1200^3 voxels, voxel size: 2.8 μm) (Fan et al., 2020). The carbonates are Estaillades (1000^3 voxels, voxel size: 3.58 μm) (Muljadi et al., 2016) and Savonnières (1000^3 voxels, voxel size: 3.8 μm) (Bultreys et al., 2016).

We conducted the training on a workstation equipped with a single NVIDIA Quadro RTX 5000 graphics card, a dual Intel(R) Xeon(R) Gold 5222 @ 3.80GHz CPU, and 64GB of memory. Prior to training, the rock samples need to be cut into sub-samples for use. Due to memory limitations, we recommend a size of 64^3 for the training data as a first choice. However, for highly heterogeneous rocks, larger sizes are needed for the algorithm to learn multi-scale features. The two-point correlation function was used to determine size of the training data (further detail is provided in the Supplementary Information). For the 64^3 size training samples, the learning rate is 0.0002, the maximum number of training iterations is 800, and size of the random noise vector z input into the generator is 100. For the 128^3 size training samples, the learning rate is 0.0003, the maximum number of training iterations is 800, and size of the random noise vector z input into the generator is 512. For the choice of the network structure for the discriminator and generator, we followed the scheme proposed by Mosser et al. (2017) which can automatically select the appropriate network structure according to the size of the training samples. For sizes 64^3 and 128^3, the generator network structure consists of 14 and 16 layers respectively, and the discriminator network structure consists of 16 and 18 layers respectively. The specific feature parameters, training parameters, and network hyperparameters for the generator and discriminator for the five samples are shown in Table 1.

3.2 Results

Figure 3 displays two-dimesnional cross-sections of the original heterogeneous pore images, the images reconstructed using the IPWGAN method proposed in this paper, and the images reconstructed using the commonly used DCGAN method (Salimans et al., 2016; Mosser et al., 2017). Examining the generated results at a size larger than the



a. Berea sandstone b. Bentheimer sandstone c. Mt. Simon sandstone

d. Estaillades carbonate e. Savonnières carbonate

Figure 2. Visualization of the five heterogeneous rock samples studied in this paper. The pore space is shown in red and the solid in grey. a. The heterogeneity of Berea sandstone primarily originates from secondary processes, with its pore spaces being affected by clay and calcite cementation. This also means that the porosity of the selected Berea sandstone sample is very low, not exceeding 10%. The heterogeneity of b. Bentheimer sandstone and c. Mt. Simon sandstone both arise from the varying sizes and non-uniform distribution of grains. d. Estaillades carbonate exhibits extreme heterogeneity in pore size distribution, which greatly impacts pore connectivity. e. The pore heterogeneity in Savonnières carbonate is due to multiple pore types: intergranular and microconnected macroporosity (hollow ooids), next to intergranular, and intraooidic microporosity. The hollow ooids are only connected to the other macropores by microporosity in the ooid shells. The generation of Savonnières carbonate is the most challenging of these five samples.

Table 1. Properties of the training data and parameters during training. ϕ refers to the connected porosity, K refers to the absolute permeability, Euler denotes the Euler characteristic per unit volume of the void space. Curv refers to the average curvature of the solid/void interface, TSS refers to the training sample size, NTS refers to the number of training samples, LRVS refers to the latent random vector size, TGS refers to the time to generate a sample of 300^3 with a single NVIDIA Quadro RTX 5000 graphics card, SSA stands for specific surface area.

Parameters	Berea sandstone	Bentheimer sandstone	Mt. Simon sandstone	Estaillades carbonate	Savonnières carbonate
Volume	1200^3	$1000^2 \times 3000$	1200^3	1000^3	1000^3
Voxel size (µm)	2.0	3.58	2.8	3.58	3.8
ϕ (-)	0.07	0.22	0.27	0.10	0.14
Euler (mm^{-3})	-3.02×10^4	-1.98×10^4	-2.04×10^4	-2.26×10^3	-8.75×10^3
Curv (mm^{-1})	5.2×10^{-5}	1.29×10^{-5}	8.2×10^{-6}	1.33×10^{-5}	1.22×10^{-5}
SSA (mm^{-1})	19.4	71.0	78.2	70.1	51.5
TSS	128^3	64^3	64^3	128^3	128^3
NTS	7.29×10^2	1.035×10^4	5.832×10^3	1.372×10^3	1.0×10^3
LRVS	512	100	100	512	512
TGS (s)	5.34	2.79	2.47	5.62	5.59

training sample size is critical for heterogeneous porous rocks, and also tests the model's upscaling capability. IPWGAN captures pore shapes and connectivity more similar to the real images. Visually, the proportion of pores of different size is captured, and the pore shapes match the real images closely, indicating that diverse multi-scale visual details are captured.

In contrast, the results from DCGAN seem quite uniform, and even after multiple training attempts, it still cannot achieve good results in reconstructing Savonnières carbonate. 600^3 images have also been created to further assess the model's upscaling capability, and the results are shown in Figure S3. Details about the hyperparameter selection are also provided in the Supplementary Information.

To quantitatively evaluate the quality of the generated images, we calculated porosity, absolute permeability (computed by Porefoam1f, Raeini et al. (2012); Bijeljic et al. (2013)), curvature, specific surface area, and Euler characteristic for generated 300^3 and 600^3 images. Additionally, we analyzed their respective two-point correlation functions. The results for the artificial 300^3 images of Berea sandstone, Estaillades carbonate, and Savonnières carbonate are shown in Figure 4.

The DCGAN does not consider targeted learning and the maintenance of multi-scale features during the training process. Besides the functional value variations with size differing from the base-case images, it can be observed that at larger sizes, it cannot maintain pore features, giving poor predictons in many cases. In contrast, the IPWGAN can maintain functional regularity well within the displayed size range. This is because IPWGAN captures both global and local image features. It is evident that it provides good predictions of statistical parameters, especially for Berea and Estaillades. Savonnières is more complex, and the IPWGAN does not fully learn all features of the images, resulting in a parameter range smaller than real images, even though we have used pyramid structures and FSMR. However, the results are still significantly better than DCGAN.

To prove the cross-scale upscaling capability of the proposed model, we generated cm-scale $2{,}200^3$ (ten-billion-cell) images (Figure 5). This represents the size on which

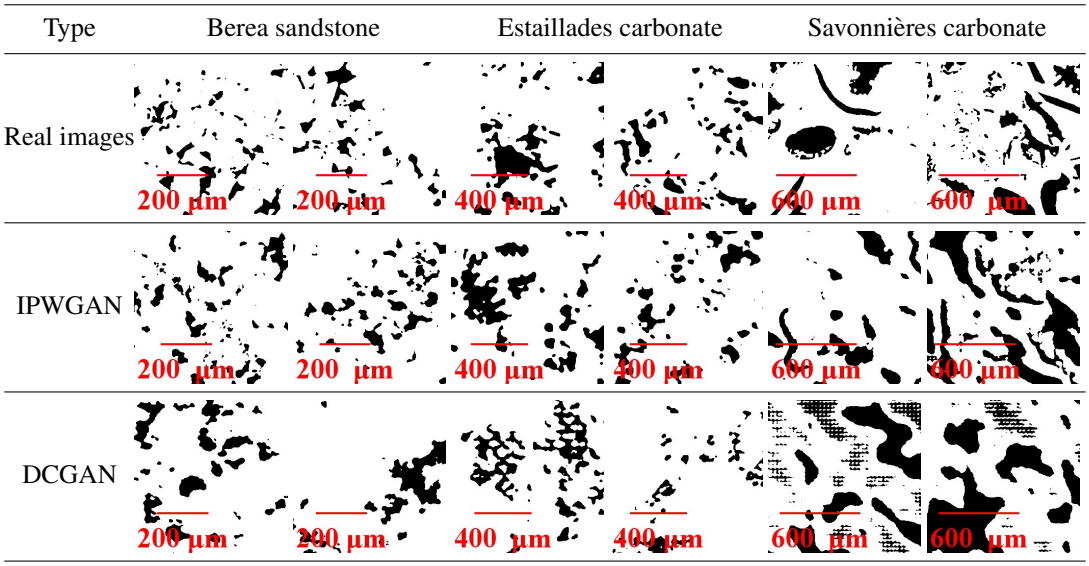


Figure 3. Two-dimensional cross-sections of three-dimensional images of Berea, Estaillades, and Savonnières, which have pronounced heterogeneity. Each type of three-dimensional image is presented with two distinct two-dimensional cross-sections images to show different characteristics. Visualizations for Bentheimer and Mt. Simon are shown in the Supplementary Information.

traditonal measurements of average flow properties are acquired. We can generate images of this size while maintaining high resolution. The reconstructed images show stability in both visualization with porosity, permeability, surface area, curvature and Euler characteristic values that fall within the range of the training data. This confirms that IPWGAN can truly achieve cross-scale upscaling, which is crucial for petrophysics and porous media flow studies.

Table 2 lists the error statistics for all results, further revealing the difference in results between IPWGAN and the representative machine learning reconstruction method DCGAN. Particularly for parameters such as Euler characteristic, that quantifies connectivity, IPWGAN showcases better results.

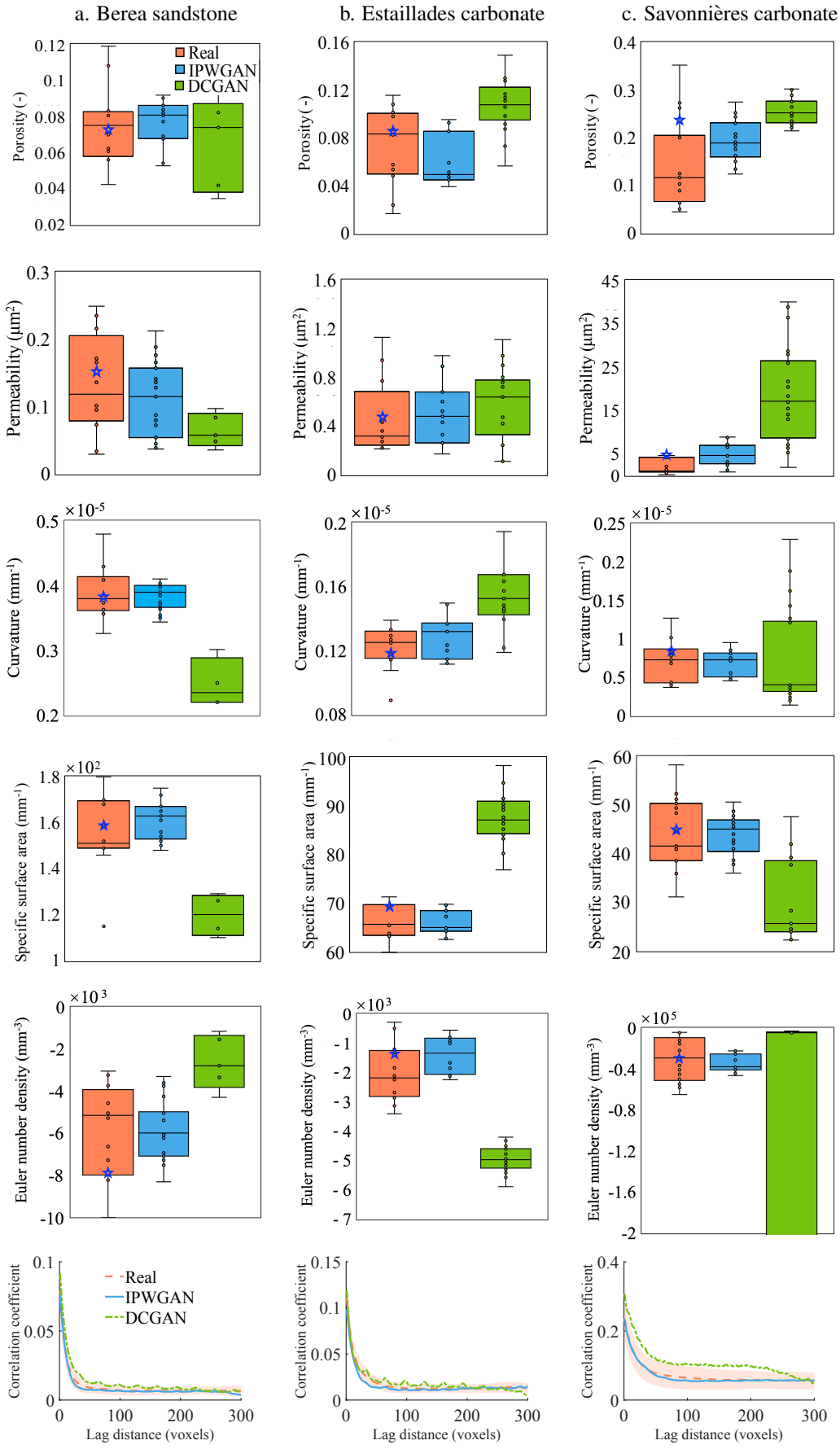
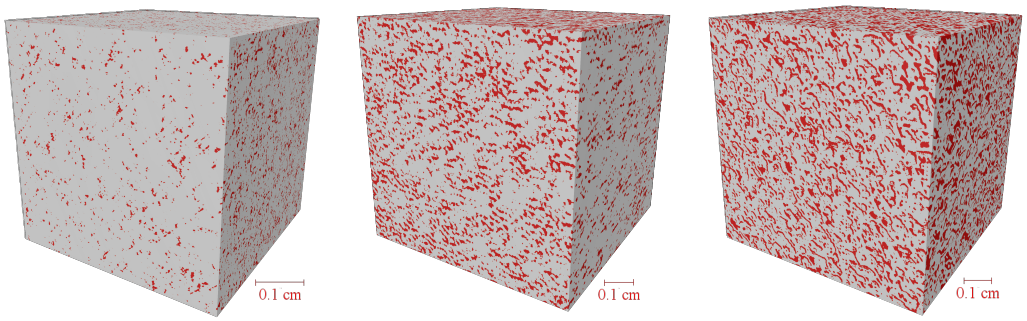

a. Berea sandstone
b. Estaillades carbonate
c. Savonnières carbonate
Real
IPWGAN
DCGAN
Porosity (-)
Permeability (μm²)
Curvature (mm⁻¹)
×10⁻⁵
Specific surface area (mm⁻¹)
×10²
Euler number density (mm⁻³)
×10³
×10⁵
Correlation coefficient
Lag distance (voxels)

Figure 4. The upper plots display the porosity, permeability, average curvature range, specific surface area, and Euler characteristic per unit volume for the actual images (represented in red) in comparison to predictions made using IPWGAN (in blue) and the previously-published DCGAN (in green) (Mosser et al., 2017) for Berea, Estaillades, and Savonnières.This box plot provides a detailed representation of the statistical distribution of the dataset. The bottom of the box (the lower quartile) and the top of the box (the upper quartile) delineate the range for the middle 50% of the data, while the line within the box represents the median, indicating the central point of the data. The "whiskers" of the box plot extend to the minimum and maximum values, illustrating the range of the data excluding outliers. Additionally, the plot includes outlier and inner points. The statistics are based on 21 randomly selected 300^3 samples from the entire image, with corresponding IPWGAN and DCGAN each randomly generating the same size and number of synthetic images. The lower plots illustrate the two-point correlation function for the three aforementioned samples, derived from the statistics of 21 images. The red dashed line represents the results from the real images, the blue solid line represents the results from IPWGAN-generated images, and the green dashed line represents the results from DCGAN. The shaded areas show the range of the real images plus or minus one standard deviation from the mean. The blue stars represent the average of the results from three 300^3 randomly-selected sections of the $2,200^3$ generated images.



a. Berea $2{,}200^3$ b. Estaillades $2{,}200^3$ c. Savonnières $2{,}200^3$

Figure 5. Generated images for Berea sandstone, Estaillades carbonate, and Savonnières carbonate of size $2{,}200^3$ at the centimetre scale, where red indicates porosity and grey represents the solid. The generated images retain the visual characteristics of the original images and exhibit distinct heterogeneity, consistent with that of the original samples.

Table 2. Parameters and two-point correlation function results of Berea, Estaillades, and Savonnières with 300^3 voxel statistics: mean relative error, MRE, Eqs. (S2), ϕ refers to the connected porosity, K refers to the absolute permeability, Curv refers to the curvature, Euler denotes the Euler characteristic per unit volume, SSA stands for specific surface area, and TE indicates the average relative difference in the two-point correlation function between real and fake images for each two-point distance, Eq. (S3). The improvement of IPWGAN over DCGAN is significant, with the mean relative errors for all items mostly below 0.1.

Samples	Error of parameters and function	MRE	
		IPWGAN	DCGAN
Berea sandstone	ϕ	0.02	0.16
	K	0.23	1.08
	Euler	0.11	0.66
	Curv	0.08	0.57
	SSA	0.01	0.36
	TE	10.0	17.3
Estaillades Carbonate	ϕ	0.22	0.32
	K	0.24	0.37
	Euler	0.45	2.10
	Curv	0.12	0.21
	SSA	0.002	0.25
	TE	8.49	13.2
Savonnières carbonate	ϕ	0.23	0.41
	K	0.61	0.90
	Euler	0.01	3.08
	Curv	0.01	0.03
	SSA	0.001	0.43
	TE	8.84	20.2

4 Discussion

Previous work on three-dimensional image generation mainly focused on generating small-scale samples of simple porous rocks, without giving sufficient attention to effective training of multi-scale pore structure information (Yan et al., 2023). The method we propose integrates pyramid structure generators, FSMR, and other enhancement techniques. This allows the generated images to maintain both local and global features and can retain various properties of heterogeneous porous rocks even when expanding to nearly 10 times the training sample size. Our method truly overcomes the trade-off between resolution and field of view for heterogeneous digital rock data. Images with greater heterogeneity and more complex pore spaces are harder to train perfectly, due to the insufficient sample volume and low model complexity.

The results using a single GPU are already satisfactory, but some more memory-intensive direct enhancement methods could be considered, such as the self-attention module, which focuses on long-range pore connectivity (Zhang et al., 2019). Furthermore, during modelling, it is essential to reinforce control of the topology. Savonnières is a good example; its topological features differ significantly from the other images studied. During training, it was challenging to capture good topological characteristics. Approaches including TopoGan that add quantitative topological feature loss constraints are an excellent direction for further improvement of the method (Wang et al., 2020). Specific convolutions, such as dynamic serpentine convolution, are helpful in capturing the topological features of porous media. Additionally, multi-modal learning combined with graph neural networks is a potential solution, as graph neural networks are sensitive to topol-

ogy (Holzinger et al., 2021). Moreover, the main constraint in generating $10{,}000^3$ images (trillion-cells) or larger is neither the model (as the statistical features of images generated are stable) nor the generation time (as it takes 5 hours using high performance computing with 2 cores and also 4TB memory for a single $2{,}200^3$ image), the main limitation is memory.

5 Conclusions

We have proposed an innovative method to generate heterogeneous three-dimensional pore space images from training samples by effectively capturing multi-scale features while preserving the inherent physical properties. Within the method, the pyramid generator structure integrates multi-scale digital rock images, effectively revealing the characteristics and relationship between large and small pores. Feature statistics mixing regularization enhances image diversity and improves the discernment of genuine local features, a common challenge in practice.

The method was applied to five heterogeneous rock images: three sandstones and two carbonates. The IPWGAN produced good predictions of porosity, absolute permeability, specific surface area, curvature, Euler characteristic and the two-point corelation function with better results than obtained using the previously-published DCGAN.

Finally, the IPWGAN was able to generate stable multi-billion-cell images, representing cm-sized samples at micron resolution. This provides benchmark reconstructions for modeling and analysis of porous structures at different length scales.

6 Open Research

The data set in this paper can be accessed through figshare data (analysis): 10.6084/m9.figshare.24516712 and github (code): https://github.com/ImperialCollegeLondon/porescale/tree/master/codes/IPWGAN.

Acknowledgments

This project was funded by National Natural Science Foundation of China (No. 42106213); International Postdoctoral Exchange Fellowship Program (No. PC2022059) ; National Key Research and Development Program of China (No. 2021YFC3100601). We thank Ryan Arnstrong, Sajjad Foroughi, Min Li, and Zhuangzhuang Ma for valuable discussions during the course of our research.

References

Bijeljic, B., Raeini, A., Mostaghimi, P., & Blunt, M. J. (2013). Predictions of non-fickian solute transport in different classes of porous media using direct simulation on pore-scale images. *Physical Review E*, *87*(1), 013011.

Bultreys, T., Stappen, J. V., Kock, T. D., Boever, W. D., Boone, M. A., Hoorebeke, L. V., & Cnudde, V. (2016). Investigating the relative permeability behavior of microporosity-rich carbonates and tight sandstones with multiscale pore network models. *Journal of Geophysical Research: Solid Earth*, *121*(11), 7929–7945.

Denton, E. L., Chintala, S., Fergus, R., et al. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. *Advances in neural information processing systems*, *28*.

Fan, M., McClure, J. E., Armstrong, R. T., Shabaninejad, M., Dalton, L. E., Crandall, D., & Chen, C. (2020). Influence of clay wettability alteration on relative permeability. *Geophysical Research Letters*, *47*(18), e2020GL088545.

Garfi, G., John, C. M., Lin, Q., Berg, S., & Krevor, S. (2020). Fluid surface coverage showing the controls of rock mineralogy on the wetting state. *Geophysical Research Letters*, *47*(8), e2019GL086380.

Holzinger, A., Malle, B., Saranti, A., & Pfeifer, B. (2021). Towards multi-modal causability with graph neural networks enabling information fusion for explainable ai. *Information Fusion*, *71*, 28–37.

Kim, J., Choi, Y., & Uh, Y. (2022). Feature statistics mixing regularization for generative adversarial networks. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 11294–11303).

Lin, Q., Bijeljic, B., Pini, R., Blunt, M. J., & Krevor, S. (2018). Imaging and measurement of pore-scale interfacial curvature to determine capillary pressure simultaneously with relative permeability. *Water Resources Research*, *54*(9), 7046–7060.

Liu, S., Zhong, Z., Takbiri-Borujeni, A., Kazemi, M., Fu, Q., & Yang, Y. (2019). A case study on homogeneous and heterogeneous reservoir porous media reconstruction by using generative adversarial networks. *Energy Procedia*, *158*, 6164–6169.

Mosser, L., Dubrule, O., & Blunt, M. J. (2017). Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Physical Review E*, *96*(4), 043309.

Mosser, L., Dubrule, O., & Blunt, M. J. (2018). Stochastic reconstruction of an oolitic limestone by generative adversarial networks. *Transport in Porous Media*, *125*(1), 81–103.

Muljadi, B. P., Blunt, M. J., Raeini, A. Q., & Bijeljic, B. (2016). The impact of porous media heterogeneity on non-darcy flow behaviour from pore-scale simulation. *Advances in water resources*, *95*, 329–340.

Okabe, H., & Blunt, M. J. (2004). Prediction of permeability for porous media reconstructed using multiple-point statistics. *Physical Review E*, *70*(6), 066135.

Raeini, A. Q., Blunt, M. J., & Bijeljic, B. (2012). Modelling two-phase flow in porous media at the pore scale using the volume-of-fluid method. *Journal of Computational Physics*, *231*(17), 5653–5668.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training gans. *Advances in neural information processing systems*, *29*.

Shams, R., Masihi, M., Boozarjomehry, R. B., & Blunt, M. J. (2020). Coupled generative adversarial and auto-encoder neural networks to reconstruct three-dimensional multi-scale porous media. *Journal of Petroleum Science and Engineering*, *186*, 106794.

Wang, F., Liu, H., Samaras, D., & Chen, C. (2020). Topogan: A topology-aware generative adversarial network. In *Computer vision–eccv 2020: 16th european conference, glasgow, uk, august 23–28, 2020, proceedings, part iii 16* (pp. 118–136).

Yan, W., Chi, P., Golsanami, N., Sun, J., Xing, H., Li, S., & Dong, H. (2023). Analysis of reconstructed multisource and multiscale 3-d digital rocks based on the cycle-consistent generative adversarial network method. *Geophysical Journal International*, *235*(1), 736–749.

Zha, W., Li, X., Xing, Y., He, L., & Li, D. (2020). Reconstruction of shale image based on wasserstein generative adversarial networks with gradient penalty. *Advances in Geo-Energy Research*, *4*(1), 107–114.

Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2019). Self-attention generative adversarial networks. In *International conference on machine learning* (pp. 7354–7363).

Zhang, H., Yu, H., Meng, S., Huang, M., Micheal, M., Su, J., ... Wu, H. (2022). Fast and accurate reconstruction of large-scale 3d porous media using deep learning. *Journal of Petroleum Science and Engineering*, *217*, 110937.

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

Zhou, C., Zhang, J., & Liu, J. (2018). Lp-wgan: Using lp-norm normalization to stabilize wasserstein generative adversarial networks. *Knowledge-Based Systems*, *161*, 415–424.