JOURNAL OF PETROLEUM GEOMECHANICS
(JPG)

Original article

Three-Dimensional Reconstruction of Porous Media Images Using the Vox2Vox Model in Presence of Multimineral Segmentation Information

Bahareh Keshavarz[1], Mohsen Masihi[2*], Mastaneh Hajipour Shirazi Fard[1], Ebrahim Biniaz Delijani[1]

1- Faculty of Petroleum and Chemical Engineering, Islamic Azad University Science and Research Branch, Tehran, Iran
2- Chemical & Petroleum Engineering Department, Sharif University of Technology Sharif University of Technology, Tehran, Iran

Received: 27 January 2024; Accepted: 22 October 2024

DOI: 10.22107/jpg.2025.436814.1226

Keywords

Deep Learning, Image Reconstruction, Multimineral Segmentation, Vox2Vox model.

Abstract

The evaluation of morphological, topological, statistical, and flow properties of porous media needs high-resolution images of porous media at the pore scale. However, direct access to high-quality tomographic images can be costly and impractical. Image reconstruction techniques offer a viable solution for obtaining visually realistic image data. Among these methods, deep learning-based approaches have gained significant attention from researchers. However, the utilization of petrographic information from porous media in training such models is rare. In this study, we investigate the impact of incorporating multimineral segmentation information in the training of the VOX2VOX model. The data used in this model consists of 3D images of a sandstone reservoir along with segmented images into 5 classes, including macro-pores, clay, quartz, feldspar, and high-density minerals. Additionally, the Vox2Vox model has been trained with binary segmented images containing pores and solid phases to compare the effect of multimineral information on model and reconstructed images. Incorporating multi-mineral segmentation from the porous media significantly enhances the model's image reconstruction capabilities, as observed through comparisons of various dynamic and static features. Incorporating a five-class dataset has led the model to exhibit lower error at the outset of training, stabilizing after approximately 30 epochs, whereas this point for the model based on the two-class dataset is around 60 epochs. Furthermore, the comparison parameters for image quality, specifically SSIM Score: 0.95, MSE Score: 0.00013, and PSNR Score: 38.93, are observed for the first model, while for the second model, they are SSIM Score: 0.89, MSE Score: 0.00112, and PSNR Score: 29.49. The original image has a porosity of 0.229, which increases to 0.233 and 0.244 in the reconstructed images from the models based on five-class and two-class datasets, respectively. Additionally, the graphs of other parameters also demonstrate the superiority of the model based on the five-class dataset.

1. Introduction

The evaluation of heterogeneous mediums, such as porous spaces, plays a fundamental role in many engineering applications [1]. By accessing the three-dimensional structure of porous media, many macroscopic features, such as transport, permeability, and electro-magnetic properties, can be effectively assessed [2]. A precise understanding of the three-dimensional structure of porous media is crucial for comprehending and developing their physical properties [3]. Essentially, methods for constructing the three-dimensional structure of porous media can be categorized into two groups: the first group

* **Corresponding Author:** e-mail: masihi@sharif.edu

involves physical scanning methods that directly employ advanced tools for creating digital three-dimensional structures of porous media [4]. X-ray imaging [5] and CT scanning [6] eliminate the need for destructive sample sectioning and provide high-resolution three-dimensional images of porous media [7]. Although physical scanning techniques enable the direct acquisition of the three-dimensional structure of porous media, they still have limitations that are challenging to overcome. These tools are typically expensive and may not be available everywhere, and their processes are usually time-consuming. Moreover, the high resolution of FIB-SEM can only be applied to small samples, making it difficult to obtain the requirements for providing a representative spatial distribution and other sample characteristics [8], [9].

The images of porous media are considered as raw data which needs later analysis, ranging from rock texture identification [10] to modeling physical processes [11]. Segmentation is one of the preliminary processes employed to simplify or comprehend the images of porous media for simulation techniques [12]. Porous media images may consist of two or more phases that need to be identified and separated to gain a clear understanding of the microscopic structure. Often, binary segmentation is used which aims to detect two regions in the image, solid and void space. The alternative is to use the multiphase segmentation by which various solid rock components can be identified.

With the advancement in computational capabilities and the availability of vast amounts of data, deep learning techniques have seen significant development. Currently, an increasing number of deep learning methods are being employed for reconstructing porous media [13] and evaluating their features [14], [15]. Deep learning methods can automatically extract image features and improve the reconstruction process.

Deep learning is implemented through various methods, among which generative models are commonly used to produce realistic samples similar to real images. Generative models are responsible for creating samples based on the input training dataset. Among these models, Generative Adversarial Networks (GANs) are the most widely used [16]. GAN is an unsupervised deep learning approach that consists of two convolutional networks: the discriminator and the generator. The role of the generator is to create artificial samples that appear real to the discriminator, while the discriminator identifies the realism of the generated images in comparison to real data. When the deep model is trained, it is expected to quickly generate various synthetic images with dynamic and statistical features that are reasonably consistent with the training images.

Isola et al. introduced the Pix2Pix model, which investigates the image-to-image translation task using conditional generative adversarial networks (cGANs). They presented a framework consisting of a generator network and a discriminator network that learns the mapping between input and output images based on a large sample set [17].

The Pix2Pix generator network utilizes a U-Net architecture, consisting of an encoder to extract features from the input image and a decoder to generate the output image. This U-Net architecture incorporates skip connections between the encoder and decoder paths, facilitating the extraction of fine-grained details during the process.

The training process involves a min-max optimization, where the generator network aims to produce output images that are indistinguishable from real images, while the discriminator network tries to correctly discriminate between the generated and real images. Through training, the generator network acquires the ability to improve the generation of realistic and coherent output images. The Pix2Pix model has been successfully applied in various image-to-image translation tasks. It has demonstrated good performance in these tasks. The innovative combination of conditional generative adversarial learning and the U-Net architecture in the Pix2Pix model has provided better capabilities for image translation [17].

Pan et al. [18] present a stochastic Pix2pix model for 2D fluvial reservoir modeling, incorporating field data and diverse facies geometries. By constraining model realizations to available geophysical and petrophysical interpretations, the approach accelerates multiphysics inversion, requiring only production history matching. Despite its 2D focus, an extension to 3D is easily achieved through zone-by-zone modeling. Addressing a common machine learning challenge, the method extends conditional generative adversarial networks with a penalty term to generate diverse realizations respecting

conditional data. Trained on realistic spatial models from a rule-based fluvial reservoir simulator, the method accurately reproduces patterns, even with varying conditioning data.

The study of Buono et al. [19] explores the underutilization of digital rock physics in investigating microporous volcanic rocks, known for their complex microstructures. They propose a framework for optimizing 3D/4D imaging of these rocks, based on a 3D multiscale study of a tuff using X-ray microtomography. High-resolution scans are found necessary for accurate characterization of microstructure and petrophysical properties. Notably, 2D U-Net and pix2pix networks trained on paired data demonstrate effectiveness in facilitating high-resolution imaging of large microporous volcanic rocks, marking an initial application of deep learning-based super-resolution to unconventional non-sedimentary digital rocks and real scans.

Anderson et al. [20] proposed the utilization of deep learning image translation models to forecast high-contrast focused ion beam-scanning electron microscopy (FIB-SEM) image volumes based on transmission X-ray microscopy (TXM) images, particularly in scenarios where only 2D paired training data is accessible. They introduced a regularization technique aimed at enhancing the generation of 3D volumes from 2D-to-2D deep learning image models. The most effective GAN models were found to be the WGAN models, with a particular emphasis on the pix2pix WGAN and SRGAN 2x WGAN models, which exhibited superior performance.

Recently, GANs have been used for reconstructing porous media based on three-dimensional training data [21], [22]. The first step in training a GAN model for porous media image reconstruction is preparing the dataset. Therefore, different parts of the three-dimensional original image volume are extracted. During training, the generator plays the role of generating unrealistic images that can deceive the discriminator. On the other hand, the discriminator must learn how to distinguish between real and fake images.

In another study, Zhang et al. [23] developed a fully connected deep model for reconstructing three-dimensional images from two-dimensional samples. This model includes an automatic encoder and a GAN, offering a solution to the stability and challenges of convolutional GANs, making notable progress in this field.

Limited training data is a significant concern that was addressed through conditional GANs [24]. CGAN, as a generative model, has tackled some of the challenges faced by conventional GANs [25]. Unlike regular GANs, the generator networks in CGANs create new samples based on both noise and conditional data factors. However, the role of the discriminator function remains similar to that of a standard GAN. Given CGAN's performance in studies related to porous media, it has recently gained attention for accelerating the statistical reconstruction algorithm [25]. While two-dimensional reconstruction methods were relatively straightforward, layer-by-layer stacking in the Z direction has been utilized for three-dimensional reconstruction [25]. Later, they introduced a CGAN framework for reconstructing complete two-dimensional images from limited information, employing an intermediate information map and the complete image [26].

Shams [27] utilized the combination of conditional and statistical GAN networks for reconstructing homogeneous and heterogeneous porous media from a two-dimensional image. Information from the statistical approach was fed into the deep network with conditional data, and the reconstruction was performed based on a deep generator method. This integration enabled the reconstruction of non-homogeneous samples, which is a key challenge in this field.

Moreover, GAN networks and an autoencoder were employed for reconstructing three-dimensional samples of porous media [21]. Multi-scale reconstruction was conducted for sandstone and carbonate rocks from samples taken from an Iranian oilfield. The GAN network predicted inter-grain porosity, while the autoencoder provided intra-grain porosity information. This approach provides a reliable method for the multi-scale reconstruction of porous media samples [21].

Due to the necessity of reconstructing three-dimensional images of porous media, extensive research and studies have been conducted in this area to produce high-quality three-dimensional samples. However, the utilization of information obtained from the segmentation of rock textures for three-dimensional image reconstruction needs further investigation.

In this work, it is aimed to enhance the quality of generated images by providing the deep learning model with features extracted from segmented images of porous media. To achieve this, the

Pix2Pix model, which has proven applications in image-to-image translation, is utilized enabling the reflection of rock texture features into the generated images of porous media. The Pix2Pix model generates one type of image based on its paired counterpart of a different type. By ranking the segmented rock texture images of porous media, our proposed approach focuses on generating three-dimensional images of these mediums, marking a starting point for utilizing rock texture information in image reconstruction of porous media. Noteworthy is the presented network as it needs less memory to be trained. Training deep learning model is a real challenge specially when it comes to 3D images. This superiority makes it simple to train the presented model for other reservoirs.

2. Materials and Methods

2.1. Data

The data used in this study consists of Bentheimer sandstone images and segmented images based on the rock texture of those samples. The images have dimensions of 800*800*800 and a voxel size of 2.15 um. The segmented images have been classified into five classes, including macro-pores, clay, quartz, feldspar, and high-density minerals. In this segmentation, 25.38% corresponds to the first class, 1.42% corresponds to the second class, 72.12% corresponds to the third class, 0.99% corresponds to the fourth class, and 0.09% corresponds to the fifth class. These data are publicly accessible on the digitalrocksportal.org website. In this study, we divided the data into three sets: 128 initial slices for validation data (dimensions: 128*800*800), 128 subsequent slices for test data (dimensions: 128*800*800), and the remaining 544 slices for training data (dimensions: 544*800*800). The patch-size used to train model is 64*64*64. A sample image is shown in Fig. 1.

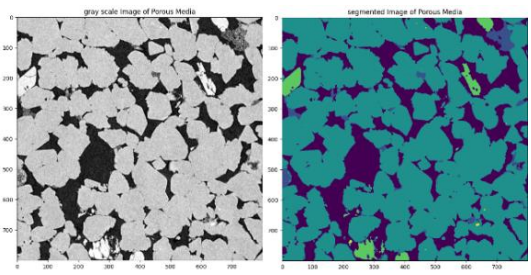

Figure 1. An image of the Bentheimer sandstone reservoir and its corresponding segmentation image.

2.2. Network Architecture

To generate three-dimensional images for the representation of porous media, we utilized a Vox2Vox model, inspired by the Pix2Pix model.
The Vox2Vox model, akin to the Pix2Pix model, consists of a generator network and a discriminator network. The generator network is constructed using a combination of U-Net and Res-Net architectures [28], while the discriminator network is built based on the PatchGAN architecture [29]. The network architecture used in this work is inspired by Cirillo et al. [30], who proposed it for brain tumor segmentation using the data provided by the BraTS Challenge 2020. Some modifications have been applied to the architecture to transform its application from image segmentation to image reconstruction. The details of the network architecture are as follows.
The generator network includes the following components:

- A 3D image of a porous media with one channel as input.
- Four downsampling blocks, each consisting of 3D convolutions with a kernel size of 4, stride 2, and same padding, followed by batch normalization [31] and Leaky ReLU activation function. The number of filters used in the first 3D convolution is 64, and this number doubles in each down-sampling block.
- Four central blocks, comprising 3D convolutions with a kernel size of 4, stride 1, and same padding, followed by instance normalization and Leaky ReLU activation function.
- Three upsampling blocks, each consisting of 3D transposed convolutions with a kernel size of 4 and stride 2, followed by instance normalization and ReLU activation function. Each 3D convolution from the input is connected to the corresponding decoder output.
- A single 3D transposed convolution with one filter, a kernel size of 4, and stride 2, followed by the tanh activation function. The output has dimensions of $1 \times 64 \times 64 \times 64$, and it represents the reconstructed image based on the segmentation of the porous media.

On the other hand, the discriminator network includes the following components:

- A segmented image based on the geological tissue and a 3D image with one channel, which can be either generated from real images or

synthetic images.
- Convolutional layers similar to those used in the generator network.
- A 3D convolution with one filter, a kernel size of 4, stride 1, and the same padding, similar to the previous layers. The output has dimensions of $1 \times 6 \times 6 \times 6$, and it represents the quality of the images generated by the generator network.

All 3D convolutional and 3D transposed convolutional layers utilize a kernel size of 4, He et al.'s weight initialization method [32], and employ the same padding. To reduce GPU memory consumption, a single convolutional layer is used in the encoder and decoder blocks, whereas a standard U-Net typically employs two convolutional layers in each block. The use of 2 strides in these convolutions implies that the encoder/decoder directly reduces/increases the input volume by a factor of 2. The model was trained using an NVIDIA GeForce RTX 2070 with a Max-Q Design graphics card, which has a memory capacity of 8 gigabytes and was trained for 150 epochs.

2.3. Loss Function

For two models, the generator and the discriminator, in the Vox2Vox architecture, two loss functions are utilized.

The discriminator's loss function calculates two distinct types of errors, one for classifying real samples and the other for classifying samples generated by the generator. The objective of this function is to encourage the discriminator network to accurately distinguish between real and generated samples with high precisions.

Initially, the function assigns a value of one to labels for real samples. It then computes the loss function for classifying real samples using these labels and the discriminator's output, measuring the disparity between the discriminator's output and the true labels. Subsequently, another function calculates the loss for classifying samples generated by the generator, employing labels with a value of zero. This loss reflects the difference between the discriminator's output and the true labels for the generator's images. The final discriminator loss is the sum of these two losses.

The generator loss function is a combination of two components: GAN loss and mean absolute error (L_1 loss). The mean absolute error is designed to evaluate the performance of the generator in the Vox2Vox architecture. The objective of the generator loss is to encourage the generator to create high-quality and realistic images that closely resemble the target images.

GAN loss, using binary classification labels, measures the discrepancy between the discriminator's output for the generated samples and the target labels, which are set to all ones. The GAN loss motivates the generator to generate images that effectively deceive the discriminator and are recognized as real samples.

The L_1 loss, also known as mean absolute error, computes the absolute pixel-wise differences between the target images and the generated images. It calculates the average absolute difference between corresponding pixels in the two images. The L_1 loss penalizes the generator for producing images that deviate significantly from the target images in terms of pixel values.

The total generator loss is obtained by combining both the GAN loss and the L_1 loss. The importance of the L_1 loss is controlled by a hyperparameter that determines the trade-off between the GAN loss and the L_1 loss. A higher value of this hyperparameter gives more significance to the L_1 loss, while a lower value emphasizes the GAN loss more.

2.4. Data Augmentation

Data augmentation is an important and widely used technique in the field of machine learning, applied to enhance the performance and accuracy of machine learning models. In this approach, various modifications and transformations are applied to the training data to increase the diversity and generalization capability of the model.

The primary objective of performing data augmentation is to increase the number of training samples and create multiple variations of the data. By applying operations such as rotation, translation, resizing, cropping, adding noise, and more, new instances of data are generated, which appear distinct from the original data.

In this study, we employed rotation, vertical flipping, and horizontal flipping as data augmentation techniques, which were applied to the training data at the beginning of each epoch. These techniques augment the dataset by generating variations of the original data, thereby enriching the diversity of training samples and enhancing the model's ability to recognize and adapt to different scenarios and features present in the data.

2.5. Evaluation Metrics

The evaluation of generated samples will be performed using the Porespy and OpenPNM packages. PoreSpy and OpenPNM are two popular and powerful libraries in the field of analysis and simulation of porous structures and properties. These libraries find extensive applications in nanotechnology research, petrochemical engineering, materials science, and other related fields.

For evaluating the performance of the model, porosity, two-point correlation function, and linear path function are utilized to examine the static features of the generated samples. The two-point correlation function denoted as $S_2(r)$, represents the probability of finding two arbitrary points with a distance of r from each other within the same phase, i.e., either the solid rock or pore phase.

The linear path function characterizes the connectivity of local pores. It calculates the probability of a straight-line segment l_r with length being entirely situated within the pore phase. In reality, this probability is only computed in the three perpendicular directions of a three-dimensional structure.

To investigate flow properties and dynamic characteristics of the reconstructed images, the relative permeability and capillary pressure have also been employed. These measures are utilized to assess the flow properties and dynamic attributes of the reconstructed images. Simulation of dynamic features of the generated samples, including relative permeability and capillary pressure, involves several stages including extraction of the existing network in the image, subsequent extraction of geometric features of the network, addressing any issues related to network clustering, creating phases, definition of the Invasion Percolation algorithm, definition of the Stokes flow algorithm, definition of the multi-phase flow model for relative permeability calculations.

3. Results and Discussion

In this section, we will discuss and examine the results obtained from training the models, followed by a comparison between the reconstructed images and the original test data in terms of dynamic and static features of the porous media.

For this study, we have developed two generative models, each trained with 150 iterations. These two models have been created using the same dataset but in different ways. In the first model, a segmented dataset containing five distinct classes, including macro-pores, clay, quartz, feldspar, and high-density minerals, has been utilized. However, in the second model, the segmented dataset consists of only two classes: solid and pores. These two models have been designed to enable a comparison between the two reconstructed samples and determine whether improvements in the reconstructed images can be achieved by considering segmentation based on rock texture in the porous media.

The generated models will be evaluated and compared based on their ability to recreate the dynamic and static features of the original test data. The evaluation will focus on the accuracy of the reconstructed images, particularly concerning the different classes and structures present in the porous media. The comparison will provide insights into the effectiveness of each model and its capability to capture the essential characteristics of the porous media. The discussion will revolve around the strengths and limitations of the generative models, as well as potential areas for further improvement and refinement.

3.1. Loss Function Plot

In evaluating the performance of the VOX2VOX model for 3D image reconstruction of porous media using paired grayscale and segmented images, a comprehensive analysis of the loss function plots is essential. The loss function serves as a crucial metric for assessing the accuracy of the model's predictions, offering insights into its performance and avenues for improvement.

Figures 2 and 3 present the loss function plots for the model trained with a segmented dataset containing five classes and subsequently for the model trained with a segmented dataset consisting of two classes. The inclusion of geological texture information in the 5-class segmented dataset appears to facilitate improved and more rapid learning of features in the reconstructed images, as compared to the 2-class segmented dataset. A discernible distinction between Figure 2 and Figure 3 is evident, with the 5-class segmented dataset exhibiting lower initial error levels, fluctuating between 10 to 20, in contrast to the 2-class segmented dataset where the error varies

between 10 to 25 during the initial epochs. Furthermore, the loss function plot for the 2-class segmented dataset indicates higher oscillations.

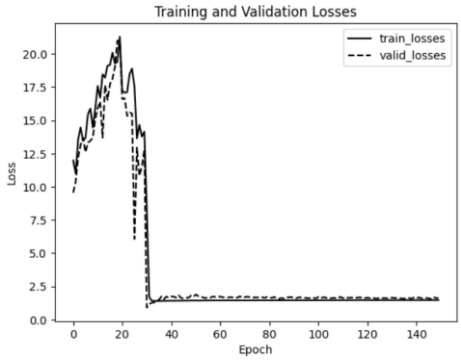

Figure 2. Loss Function Plot for the Model with a Segmented Dataset of Five Classes.

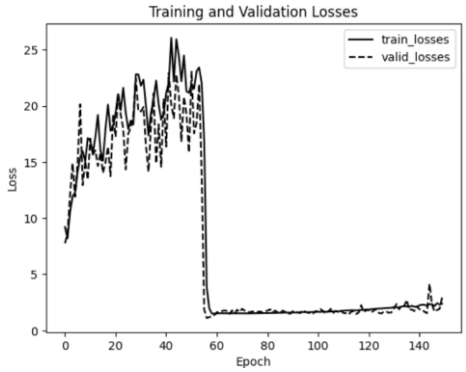

Figure 3. Loss Function Plot for the Model with a Segmented Dataset of Two Classes.

A noteworthy observation is the accelerated learning rate of the model when incorporating multimineral segmentation. The loss function of the model trained with the 5-class segmented dataset attains a relatively stable state around epoch 30, whereas the same state is achieved for the 2-class segmented dataset around epoch 60. These findings suggest that the model significantly benefits from datasets involving multimineral segmentation, resulting in improved and expedited training. This aspect holds particular significance when applying the model to new reservoirs, positively influencing both efficiency and outcomes.

In both loss function plots, post reaching a stable state, the error remains consistent within a limited range, and there is no discernible gap between the training and validation errors. This observation suggests that both models adequately fit the data without experiencing overfitting or underfitting issues. It implies that the complexity of the VOX2VOX network and the volume of data used in the training process are appropriate for the given task.

By providing a detailed examination of the loss function plots, our analysis substantiates the effectiveness of incorporating geological texture information through multimineral segmentation in enhancing the VOX2VOX model's learning capabilities for 3D image reconstruction of porous media. These insights contribute to advancing the field by shedding light on the nuanced impact of segmentation information on model performance.

3.2. Parametric Comparison of Images

In Figure 4, the original test data sample, the reconstructed sample generated by the model based on a 5-class dataset, and the reconstructed sample produced by the model based on the binary dataset are depicted. Notably, the inclusion of multimineral segmentation information in the 5-class dataset significantly enhances the reconstruction quality and perceptual clarity of the images, particularly in capturing intricate details of the rock texture within the porous media. Conversely, the reconstructed images from the model trained on the binary dataset exhibit a comparatively lower level of quality and clarity.

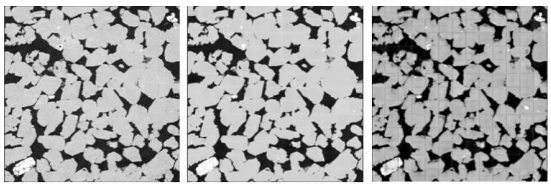

Figure 4. a: Original Image, b: Reconstructed Image considering the information of the rock texture in the porous media, and c: Reconstructed Image considering the pores and solid phase.

To provide a more comprehensive understanding of the implications of our findings, we further analyze the results using three commonly employed similarity evaluation metrics: Structural Similarity Index Measure (SSIM), Mean Squared Error (MSE), and Peak Signal-to-Noise Ratio (PSNR). These metrics, widely utilized in computer vision and image processing, offer valuable insights into the quality of reconstructed images and facilitate meaningful comparisons with ground truth images.

The results of these three metrics are summarized in Table 1. It is evident from the data that the

reconstructed images, incorporating information on rock texture through multimineral segmentation, display a higher degree of similarity to the original samples when compared to those reconstructed by the model based on the binary dataset. The enhanced performance is reflected in superior SSIM values, lower MSE, and higher PSNR scores, all indicative of improved accuracy and correlation with the ground truth images.

In conclusion, our findings underscore the significance of incorporating multimineral segmentation information for achieving superior image reconstruction in porous media. The improved quality and clarity observed in the reconstructed images align with the enhanced performance metrics, emphasizing the efficacy of our approach. This nuanced analysis provides a clearer and more detailed discussion of the implications of our results, contributing to the advancement of the field by highlighting the importance of considering rock texture information in 3D image reconstruction of porous media.

Table 1. Three metrics for the reconstructed images.

	SSIM Score	MSE Score	PSNR Score
5-class dataset	0.95	0.00013	38.93
Binary dataset	0.89	0.00112	29.49

3.3. Static Features of Reconstructed Images

Porosity:

The findings from our study reveal important insights into the reconstruction of 3D images of porous media using the VOX2VOX model. We observed a porosity of 0.229 in the original image, and upon employing our model trained with 5-class segmentation data, the reconstructed image demonstrated a porosity of 0.233. In contrast, when using the model trained with 2-class segmentation data, the porosity in the reconstructed image was measured at 0.244. The porosity error for the model trained with 5-class segmentation data was calculated at 1.78%, while the model trained with 2-class segmentation data exhibited a porosity error of 6.29%.

It is evident that the inclusion of multimineral segmentation information significantly influences the quality of reconstructed images, as evidenced by the improved resemblance to the original sample in terms of porosity. Our results suggest that leveraging rock texture information, particularly through the incorporation of multimineral segmentation, contributes to a more accurate representation of the porous media. This enhanced fidelity in reconstruction is crucial for applications where precision in porosity measurements is paramount.

- The Two-Point Correlation Function:

In Figure 5, the Percent Change in Two-Point Correlation Function between the Original Image and Reconstructed Images is presented. Notably, the measure indicates a closer resemblance to the image reconstructed from the model based on rock texture.

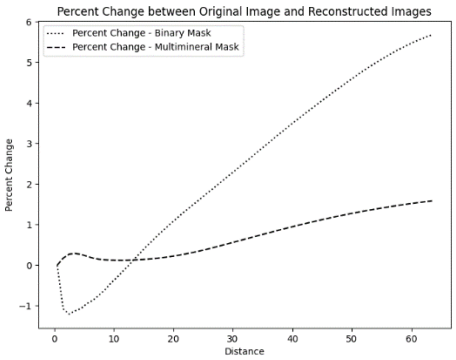

Figure 5. Percent Change in Two-Point Correlation Function: Original Image vs Reconstructed Images

The analysis reveals that the two-point correlation function closely aligns with the rock texture when utilizing the 5-class segmented dataset for image reconstruction. The percentage change between the two-point correlation function of the reconstructed images based on the 5-class segmented dataset model and the original image ranges from 0 to less than 2 percent. In contrast, the difference for reconstructed images based on the 2-class dataset model and the original image is observed to be between 0 to 6 percent.

This quantitative parameter analysis reinforces the significance of employing segmentation based on rock texture in porous media for enhancing the quality of reconstructed images. The close resemblance between the reconstructed image and the original image, particularly in the context of the two-point correlation function, underscores the efficacy of the 5-class segmentation approach.

This outcome implies a more faithful representation of the intricate details of the porous media, contributing to a nuanced understanding of its structural characteristics.

These findings offer valuable insights into the role of segmentation information, specifically the multi-mineral segmentation approach, in influencing the reconstruction process. The lower percentage change observed in the 5-class segmentation model suggests a more accurate reconstruction, emphasizing the importance of incorporating detailed mineralogical information in porous media imaging. This nuanced exploration provides a foundation for advancing the field, shedding light on the nuanced relationship between segmentation strategies and the fidelity of reconstructed images in porous media research.

- The Linear Path Function

Figure 6 shows the linear path function plot for the original image, the image reconstructed based on the 5-class segmentation, and the image reconstructed based on the 2-class segmentation.

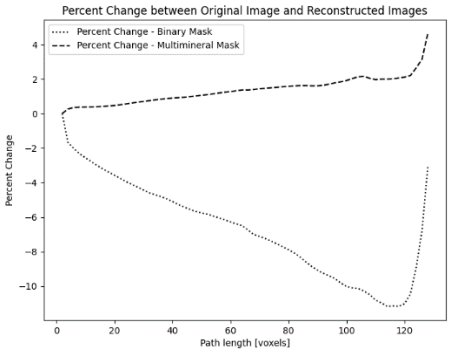

Figure 6. Percent Change in Linear Path Function: Original Image vs Reconstructed Images

Similarly, to the two-point correlation function's plot, in this graph, the values for the original sample and the sample reconstructed from the model overlap. However, the sample reconstructed from the second model shows slight differences. This indicates that although the reconstructed images from the two-class segmentation model also yield acceptable results, the images based on the other model have higher quality and exhibit closer similarity to the original sample.

3.4. Dynamic Features of Reconstructed Images

Absolute Permeability

The absolute permeability related to the original data was 73.97 mD. This value was 75.29 mD with an error of 1.78% for the reconstructed image based on the 5-class data model, compared to the absolute permeability of the original data. For the generated image from the 2-class data model, it was 98.76 mD with an error of 33.51% relative to the absolute permeability of the original data. This indicates that the first model has produced images much closer to the original data compared to the second model.

- Relative Permeability curves:

In Figure 7, the relative permeability curves derived from three examined images are presented.

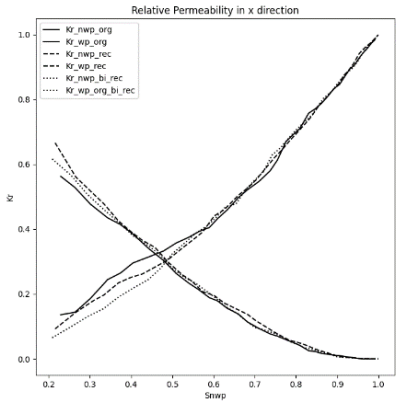

Figure 7. Relative Permeability Plot.

Although this chart does not exhibit a particular superiority of the model based on rock texture, it should be noted that significant simplifications have been applied to simulating the flow in the target sample, as explained in the previous sections. Therefore, the relative permeability chart can be considered approximate, and accurately determining the proximity of the results of the model based on rock texture to the original sample is challenging. However, the results indicate the proximity of the relative permeability of the images generated by both models to the original sample.

Table 2 presents information about the networks extracted from the three investigated images using the Snow algorithm, along with the errors relative to the original image values.

Table 2. Some parameters of pores and throats in the network.

	Original image	5 phase mask	5 phase mask error	2 phase mask	2 phase Mask error
number of pores	5971	5500	-7.88%	10016	67.74%
number of throats	9851	9681	-1.72%	15627	58.63%
pore.equivalent_diameter (μm)	23.28	24.94	7.13%	20.61	-11.45%
pore.volume (μm3)	35337.4	39094.	10.63%	22390.8	-36.64%
pore.surface_area (μm2)	3266.96	3558.28	8.92%	2381.92	-27.09%
throat.equivalent_diameter (μm)	16.87	18.01	6.75%	16.63	-1.42%

As observed, the network parameters extracted from the images of the first model are much closer to the original image compared to the parameters extracted from the images of the second model. This network serves as the basis for simulating relative permeability, and these factors along with subsequent simplifications could be the reason for the lack of significant differences in relative permeability among the samples.

- Capillary Pressure Plot:

Figure 8 shows the capillary pressure curves for three samples: the original image, the image reconstructed from the model based on rock texture information, and the image reconstructed from the model based on the two-class mask.

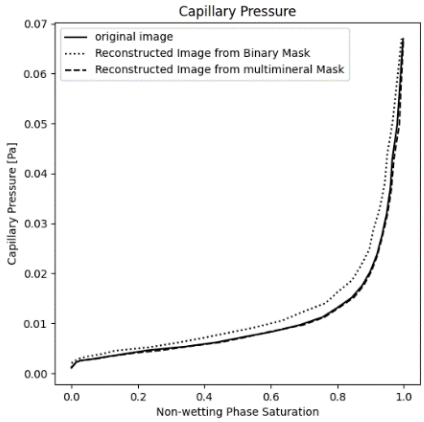

Figure 8. Capillary Pressure Plot.

The examination of this parameter also demonstrates the superiority of the model that had access to the geological texture information of the porous media. As evident in the graph, the capillary pressure in the reconstructed images from this model exhibits a much closer resemblance to the original sample. The investigation of this parameter provides further evidence to support the claim that utilizing geological texture information leads to higher-quality generated images.

4. Conclusions and recommendations

Three-dimensional reconstruction of porous media using deep learning methods is an intermediate solution that can provide researchers with realistic samples much more easily, cost-effectively, and rapidly compared to direct imaging methods. In this context, Generative Adversarial Networks (GANs) have received significant attention and efficient approaches have been proposed. However, the incorporation of geological texture information of porous media has not been extensively investigated in these methods. This information can be provided to the model through image segmentation based on geological textures. In this work, we focused on examining the impact of incorporating this information into the conditional GAN model called Vox2Vox, and the results indicated that the geological texture information can lead to the enhancement of three-dimensional reconstructed images of porous media and expedite the model's learning process. By utilizing two different data types, where one represents the solid part as a whole and the other segregates solids based on minerals, the model acquires valuable insights into mineral information. This investigation highlights the influence of petrographic information on the model's performance, ultimately leading to enhanced 3D image reconstruction.

The Vox2Vox deep learning model, which belongs to the family of conditional GANs, takes the porous media image along with its corresponding mask as an input. The dataset utilized in this work consisted of three-dimensional images of Bentheimer sandstone samples, along with their 5-class masks representing macro-pores, clay, quartz, feldspar,

and high-density minerals, with dimensions of 800*800*800. In this segmentation, 25.38% corresponds to the first class, 1.42% corresponds to the second class, 72.12% corresponds to the third class, 0.99% corresponds to the fourth class, and 0.09% corresponds to the fifth class. The model was trained first with 5-class segmentation images and second with 2-class segmentation images, comprising pores and solids, to compare the results and determine the impact of geological texture information in the model's training process.

A comparison of the error function graphs in the two models demonstrates that the presence of geological texture information significantly aids in the simpler and faster training of the model. Due to the initial error fluctuations, it was observed that in the 5-class data-based model, the error was significantly lower than in the 2-class data-based model. Furthermore, in the first model, the error converged to stability after approximately 30 epochs, while for the second model, this convergence occurred at nearly twice as many epochs, around 60.

The structural similarity index measure, mean squared error, and signal-to-noise ratio for the 5-class data-based model were 0.95, 0.00013, and 38.93, respectively. For the 2-class data-based model, these values were 0.89, 0.00112, and 29.49, respectively. The comparison in all parameters indicates that the model with access to geological texture information exhibits a much higher similarity to the original image.

Moreover, static and dynamic features of the medium such as porosity, two-point correlation function, linear path function, relative permeability, and capillary pressure of the reconstructed images from both models were compared. The porosity error in the images reconstructed from the model with geological texture information is 1.78% compared to the original image, whereas this number is 6.29% for the second model. Additionally, the absolute permeability differed by 1.78% for the reconstructed images from the geological texture-based model compared to the original image, whereas, for the 2-class data-based model, this difference was 33.51%. Furthermore, the similarity between the network extracted from the reconstructed images of the 5-class data-based model and the network from the original image was of paramount importance and notably conspicuous.

It should be noted that, firstly, simplifications were applied in simulating, and secondly, approximately 97.5% of the data pertains to only two classes, with these differences emerging from the remaining three classes, constituting 2.5% of the total data. Therefore, in reservoirs with greater geological heterogeneity, significant improvements in the reconstructed images will be evident in these parameters.

Ultimately, the results indicate that multimineral segmentation information of porous media not only leads to higher quality and more realistic reconstructed images but also simplifies and expedite the model training process. Additionally, the presented network requires less memory for training, which is particularly advantageous, especially when it comes to training 3D models for reconstruction. This simplifies the training process with images from various reservoirs. Considering the high efficiency of deep learning methods, it is recommended to design novel networks that can better utilize geological texture information in reconstructing images of porous media, enabling us to achieve higher-quality generated images. Furthermore, it is recommended to test this model in reservoirs exhibiting greater heterogeneity among rock textures to determine whether more pronounced improvements will be observed in the parameters measured. As a final suggestion for future work, sensitivity analysis concerning the number of classes in image segmentation is proposed.

5. References

[1] S. Torquato, “Random heterogeneous materials: microstructure and macroscopic properties,” p. 701, 2002.

[2] P. Tahmasebi, “Accurate modeling and evaluation of microstructures in complex materials,” *Phys Rev E*, vol. 97, no. 2–1, Feb. 2018, doi: 10.1103/PHYSREVE.97.023307.

[3] G. T. Herman, “Fundamentals of Computerized Tomography,” 2009, doi: 10.1007/978-1-84628-723-7.

[4] K. J. Lange, H. Carlsson, I. Stewart, P. C. Sui, R. Herring, and N. Djilali, “PEM fuel cell CL characterization using a standalone FIB and SEM: Experiments and simulation,” *Electrochim Acta*, vol. 85, pp. 322–331, Dec. 2012, doi: 10.1016/J.ELECTACTA.2012.08.082.

[5] J. Baruchel *et al.*, "Advances in synchrotron hard X-ray based imaging," *C R Phys*, vol. 9, no. 5–6, pp. 624–641, Jun. 2008, doi: 10.1016/J.CRHY.2007.08.003.

[6] P. Mostaghimi, M. J. Blunt, and B. Bijeljic, "Computations of Absolute Permeability on Micro-CT Images," *Math Geosci*, vol. 45, no. 1, pp. 103–125, Jan. 2013, doi: 10.1007/S11004-012-9431-4/METRICS.

[7] S. Kamrava, P. Tahmasebi, and M. Sahimi, "Enhancing images of shale formations by a hybrid stochastic and deep learning algorithm," *Neural Networks*, vol. 118, pp. 310–320, Oct. 2019, doi: 10.1016/J.NEUNET.2019.07.009.

[8] P. Tahmasebi, "Nanoscale and multiresolution models for shale samples," *Fuel*, vol. 217, pp. 218–225, Apr. 2018, doi: 10.1016/J.FUEL.2017.12.107.

[9] F. Javadpour, D. Fisher, and M. Unsworth, "Nanoscale gas flow in shale gas sediments," *Journal of Canadian Petroleum Technology*, vol. 46, no. 10, pp. 55–61, Oct. 2007, doi: 10.2118/07-10-06/180817/NANOSCALE-GAS-FLOW-IN-SHALE-GAS-SEDIMENTS.

[10] J. L. Kulikowski, "Computer-Aided Porous Materials Description Based on Morphological Spectra," *Research & Development in Material Science*, vol. 11, no. 4, Sep. 2019, doi: 10.31031/RDMS.2019.11.000769.

[11] Q. Yang, J. Yao, Z. Huang, G. Zhu, L. Liu, and W. Song, "Pore-scale investigation of petro-physical fluid behaviors based on multiphase SPH method," *J Pet Sci Eng*, vol. 192, p. 107238, Sep. 2020, doi: 10.1016/J.PETROL.2020.107238.

[12] A. Rabbani, S. Jamshidi, and S. Salehi, "An automated simple algorithm for realistic pore network extraction from micro-tomography images," *J Pet Sci Eng*, vol. 123, pp. 164–171, Nov. 2014, doi: 10.1016/J.PETROL.2014.08.020.

[13] J. Feng, X. He, Q. Teng, C. Ren, H. Chen, and Y. Li, "Accurate and Fast reconstruction of Porous Media from Extremely Limited Information Using Conditional Generative Adversarial Network," *Phys Rev E*, vol. 100, no. 3, Apr. 2019, doi: 10.1103/PhysRevE.100.033308.

[14] S. Kamrava, M. Sahimi, and P. Tahmasebi, "Quantifying accuracy of stochastic methods of reconstructing complex materials by deep learning," *Phys Rev E*, vol. 101, no. 4, p. 043301, Apr. 2020, doi: 10.1103/PHYSREVE.101.043301/FIGURES/9/MEDIUM.

[15] S. Kamrava, P. Tahmasebi, and M. Sahimi, "Linking Morphology of Porous Media to Their Macroscopic Permeability by Deep Learning," *Transp Porous Media*, vol. 131, no. 2, pp. 427–448, Jan. 2020, doi: 10.1007/S11242-019-01352-5/METRICS.

[16] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," Jun. 2014, [Online]. Available: http://arxiv.org/abs/1406.2661.

[17] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 5967–5976, Nov. 2016, doi: 10.1109/CVPR.2017.632.

[18] W. Pan, C. Torres-Verdin, and M. J. Pyrcz, "Stochastic Pix2pix: A New Machine Learning Method for Geophysical and Well Conditioning of Rule-Based Channel Reservoir Models," Nat Resour Res 30, 1319–1345 (2021). https://doi.org/10.1007/s11053-020-09778-1.

[19] G. Buono, S. Caliro, G. Macedonio *et al.* Exploring microstructure and petrophysical properties of microporous volcanic rocks through 3D multiscale and super-resolution imaging. *Sci Rep* 13, 6651 (2023). https://doi.org/10.1038/s41598-023-33687-x.

[20] T.I. Anderson, B.Vega, J. McKinzie *et al.* 2D-to-3D image translation of complex nanoporous volumes using generative networks. *Sci Rep* 11, 20768 (2021). https://doi.org/10.1038/s41598-021-00080-5

[21] R. Shams, M. Masihi, R. B. Boozarjomehry, and M. J. Blunt, "Coupled generative adversarial and auto-encoder neural networks to reconstruct three-dimensional multi-scale porous media," *J Pet Sci Eng*, vol. 186, p. 106794, Mar. 2020, doi: 10.1016/J.PETROL.2019.106794.

[22] Y. Da Wang, M. Shabaninejad, R. T. Armstrong, and P. Mostaghimi, "Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images," *Appl Soft Comput*, vol. 104,

Jun. 2021, doi: 10.1016/j.asoc.2021.107185.

[23] F. Zhang, Q. Teng, H. Chen, X. He, and X. Dong, "Slice-to-voxel stochastic reconstructions on porous media with hybrid deep generative model," *Comput Mater Sci*, vol. 186, p. 110018, Jan. 2021, doi: 10.1016/J.COMMATSCI.2020.110018.

[25] J. Feng, Q. Teng, X. He, and X. Wu, "Accelerating multi-point statistics reconstruction method for porous media via deep learning," *Acta Mater*, vol. 159, pp. 296–308, Oct. 2018, doi: 10.1016/J.ACTAMAT.2018.08.026.

[26] J. Feng, X. He, Q. Teng, C. Ren, H. Chen, and Y. Li, "Reconstruction of porous media from extremely limited information using conditional generative adversarial networks," *Phys Rev E*, vol. 100, no. 3, p. 033308, Sep. 2019, doi: 10.1103/PHYSREVE.100.033308/FIGURES/22/MEDIUM.

[27] R. Shams, M. Masihi, R. B. Boozarjomehry, and M. J. Blunt, "A hybrid of statistical and conditional generative adversarial neural network approaches for reconstruction of 3D porous media (ST-CGAN)," *Adv Water Resour*, vol. 158, p. 104064, Dec. 2021, doi: 10.1016/J.ADVWATRES.2021.104064.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 770–778, Dec. 2015, doi: 10.1109/CVPR.2016.90.

[29] M. D. Cirillo, D. Abramian, and A. Eklund, "Vox2Vox: 3D-GAN for Brain Tumour Segmentation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12658 LNCS, pp. 274–284, Mar. 2020, doi: 10.1007/978-3-030-72084-1_25.

[30] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance Normalization: The Missing Ingredient for Fast Stylization," Jul. 2016, Accessed: Aug. 06, 2023. [Online]. Available: https://arxiv.org/abs/1607.08022v3

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," Feb. 2015, [Online]. Available: http://arxiv.org/a