Hybrid Deep Learning for 3D Reconstruction of Multi-Mineral Porous Media: Integrating U-Net and GAN for Enhanced Segmentation and Texture Preservation

Bahareh Keshavarz[a1], Mohsen Masihi[b2,] Mastaneh Hajipour Shirazi Fard[a3,] and Ebrahim Biniaz Delijani[a4]

[a] Department of Petroleum Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

[b] Department of Chemical & Petroleum Engineering, Sharif University of Technology Sharif University of Technology, Tehran, Iran

Abstract

This study proposes a hybrid deep learning approach combining U-Net and Generative Adversarial Network (GAN) architectures for the segmentation and texture-based reconstruction of 3D multi-mineral porous media images. The dataset consists of high-resolution 3D Leopard sandstone images, segmented into four key mineral classes: macro-pores, clay, quartz, and high-density minerals. Our approach leverages the feature extraction capabilities of a ResNet-18 backbone within U-Net, pre-trained specifically for multi-mineral segmentation, which then feeds these detailed features into a GAN framework for image reconstruction. The model effectively bridges segmentation and reconstruction, achieving superior image quality and structural fidelity compared to standalone GAN models by preserving intricate textures and maintaining macroscopic rock structures. Quantitative assessments reveal that the hybrid model yields porosity and absolute permeability values with minimal discrepancies (2.25% and 1.54% error, respectively) compared to actual data. These findings highlight the model's ability to replicate critical geophysical metrics and generate accurate 3D representations. Unlike traditional methods that either focused solely on segmentation or reconstruction, our model uniquely integrates segmentation-driven texture data for image reconstruction, offering a novel solution for geoscientific applications in hydrogeology, petroleum engineering, and environmental science.

Keywords: Deep Learning, Image Reconstruction, Multi Mineral Segmentation, GAN

(Generative Adversarial Network), UNet (U-shaped Convolutional Neural Network)

[1] bahareh.keshavarz@srbiau.ac.ir

[2] Corresponding Author: masihi@sharif.edu

[3] Corresponding Author: m.hajipour@srbiau.ac.ir

[4] biniaz@srbiau.ac.ir

1. Introduction

Characterization of porous media is vital in engineering applications. This needs the 3D structure of the medium which allows assessment of macroscopic properties [1].3D construction methods include physical scanning (e.g., X-ray [2], or CT [3]), which provids high-resolution 3D images without destructive sectioning [4]. However, these tools are often expensive, time-consuming, and not universally available; high-resolution FIB-SEM is limited to small samples, hindering representative analysis [5, 6].

Porous media images, as raw data, require analysis from texture identification [7] to modeling physical processes [8]. Segmentation simplifies these images for simulation [9], often using binary (solid/void) or multiphase approaches.

Convolutional neural networks (CNNs) have overcome segmentation challenges like inheritance errors [10], high sensitivity to physical modeling [11], and segmenting noisy images [12]. Most studies use SEM or higher-quality images, improving training, accuracy, and noise reduction compared to micro-CT. CNNs, combined with pooling, consider surrounding shape and texture, outperforming machine learning methods like Weka that rely on user-defined filters.

In CNN-based segmentation, networks like SegNet [13] and U-Net [14] (encoder-decoder based) convert input images to segmented ones. Sandstone samples are common due to clear mineral distinctions, allowing traditional segmentation. Here, machine learning's advantage diminishes, but it simplifies SEM imaging for higher accuracy [10] and outperforms traditional methods even for rock/fluid separation [15].

An early SegNet application used statistically generated datasets from manually segmented sandstone images, achieving 95% pixel accuracy, though lower (60%) for less frequent phases [4]. LeNet-5 was trained on grayscale micro-CT sandstone images (three phases), with physical accuracy assessed via permeability/porosity [10]. A study using QEMSCAN-segmented micro-CT sandstone found pixel accuracy did not correlate with physical accuracy (topology, flow simulation) for CNNs like SegNet, U-Net, and U-ResNet [11]. U-Net and U-ResNet achieved >99% pixel accuracies, with U-ResNet performing best in 3D physical and per-pixel accuracy.

Deep learning (DL) is increasingly used for porous media reconstruction [16] and feature evaluation [17, 18], automatically extracting image features and improving reconstruction.

Generative models, especially Generative Adversarial Networks (GANs) [19]—an unsupervised approach with a generator and a discriminator—produce realistic samples. Trained GANs quickly generate synthetic images consistent with training data.

GANs have been used for 3D porous media reconstruction from 3D training data [11, 20], where the generator creates images to deceive the discriminator. Valsecchi et al. [21] trained a GAN discriminator on 2D samples, requiring large memory and reconstructing small volumes. Feng et al. [22] used BicycleGAN (generator, discriminator, encoder), but it requires 3D training data, is inadequate for heterogeneous samples [23], and training three networks is difficult. Zhang et al. [24] developed an autoencoder-GAN model for 3D reconstruction from 2D samples, improving stability.

Conditional GANs (CGANs) address limited training data [14] by using conditional factors in generation [25], accelerating statistical reconstruction [25]. 3D reconstruction has used layer-by-layer stacking [25] or full 2D image reconstruction from limited information [26]. Shams [27] combined conditional/statistical GANs for reconstructing homogeneous/heterogeneous media from 2D images, feeding statistical information to the deep network. GANs and autoencoders were also used for multi-scale 3D reconstruction [19], with GANs for inter-grain and autoencoders for intra-grain porosity.

Stark et al. [28] introduced an open-source pipeline for reaction-diffusion models using level-set methods and GPUs. Siavashi et al. [29] used CNNs for upscaling multiphase flow simulations from low-resolution images. Yan et al. [30] developed a multiscale reconstruction algorithm using multiple dictionary learning. Yang et al. [31] presented a cGAN for reconstruction from low-resolution core images. Amiri et al. [32] used GANs for 2D to 3D heterogeneous reconstruction from 2D training images. Zhang et al. [33] proposed an RNN (3D-PMRNN) for efficient 3D reconstruction from 2D images. These DL advances improve digital rock modeling accuracy, efficiency, and scalability.

Percolation theory is able to quantify uncertainties using metrics like connected pore fraction [34]. New 3D correlations from 2D micro-CT images highlight the need for pore-scale structures. A multi-scale image reconstruction method integrates high/low-resolution micro-CT scans, maintaining features while reducing cost [35]. However, image segmentation remains a bottleneck;

While 3D reconstruction is well-researched, reconstruction using rock texture segmentation data needs further investigation. This work innovates by considering rock texture's impact on 3D reconstruction, using DL to reveal its influence on image quality and address heterogeneity. We propose a novel U-Net + GAN architecture: U-Net performs texture-based segmentation, and its encoder feeds texture features to the GAN for reconstruction, enabling texture-informed generation of multiple porous media reconstructions from an input image.

2. Materials and Methods

2.1. Data

The data used in this study consists of Leopard sandstone images along with their corresponding segmented images based on the rock texture of the sandstone. These images have dimensions of 800*800*800 and a voxel size of 2.15 um. The segmented images are classified into four classes, including macro-pores, clay, quartz, and high-density minerals. These data are publicly accessible on the digitalrocksportal.org website. For our 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). All data used had dimensions of 128*128*128. For training purposes, 1140 images with these dimensions were utilized, while 300 images were allocated for validation and 144 images were set aside for testing. Figure 1 presents a representative Leopard sandstone image alongside its corresponding segmentation, highlighting the classification into macro-pores, clay, quartz, and high-density minerals.

Low-quality inputs were not captured by a separate scanner; they were synthetically derived from the 800 × 800 × 800-voxel Leopard-sandstone micro-CT volumes available on the Digital Rocks Portal. Each high-resolution cube is first nearest-neighbour up-scaled to 896 × 896 × 896 solely so that every axis is an integer multiple of our 128-voxel patch size; this step adds voxels but does not add new information. The enlarged stack is then partitioned into non-overlapping 128 × 128 × 128 sub-volumes and linearly normalised to the [-1, 1] range. To mimic bench-top μCT artefacts, we apply an isotropic Gaussian blur ($\sigma = 1$ voxel) followed by 4 % additive Gaussian sensor noise.

2.2. Network Architecture

2.2.1. U-Net Architecture

The proposed U-Net model is designed for 3D multi-mineral segmentation, incorporating a ResNet-18 backbone for its encoder, which enhances feature extraction capabilities. The input to the model consists of 3D images with dimensions 128x128x128 and a single channel.

Encoder Path

✓ Input Layer and Initial Normalization:

- o The input layer accepts data of shape 128, 128, 128, 1.

- o Initial batch normalization is applied to stabilize the learning process.

✓ Convolutional Blocks:

- o The first convolutional block applies a 3D convolutional layer with 64 filters of size (3, 3, 3) followed by batch normalization and ReLU activation. The output is then down sampled using max pooling.

- o This block structure is repeated, with the number of filters doubling at each down sampling step (64, 128, 256, 512 filters respectively), reflecting a typical ResNet architecture.

✓ Residual Blocks:

- o Each residual block consists of two 3D convolutional layers with batch normalization and ReLU activation, followed by an additive identity shortcut connection, enhancing gradient flow during training.

✓ Final Encoding Layer:

- The final layer in the encoder path contains 512 filters, resulting in a highly abstract and feature-rich representation of the input data.

Decoder Path

✓ Up sampling and Concatenation:

- The decoder path mirrors the encoder but in reverse, utilizing up sampling layers to increase the spatial dimensions of the feature maps.

- At each up-sampling step, the feature maps are concatenated with the corresponding feature maps from the encoder (skip connections), ensuring the preservation of spatial context and finer details.

✓ Convolutional Blocks:

- Post-concatenation, 3D convolutional layers with batch normalization and ReLU activation are applied, progressively reducing the number of filters (512, 256, 128, 64, 32, 16).

✓ Final Convolution and Activation:

- The final layer uses a 3D convolutional layer with 4 filters, corresponding to the number of mineral classes, followed by a SoftMax activation function to produce the segmentation map.

Model Training and Performance

- Training: The model was trained for 100 epochs, each epoch taking approximately 15 minutes. This training duration ensured the model had sufficient time to learn and generalize the complex multi-mineral features from the input data.

- Optimizer and Loss Function: The Adam optimizer with a learning rate of 1×10^{-5} was used. The loss function employed was the categorical focal dice loss, which combines focal loss and dice coefficient, effectively handling class imbalance and improving segmentation accuracy.

Feature-Rich Output of the Encoder

The encoder's ability to produce feature-rich outputs is fundamental to the effectiveness of the U-Net architecture, especially in the context of multi-mineral segmentation. As the input image progresses through the layers of the encoder, each convolutional block extracts increasingly abstract features, capturing both local textures and global patterns crucial for distinguishing different minerals.

1. Initial Feature Extraction: The early convolutional layers in the encoder focus on extracting low-level features such as edges and textures..

2. Intermediate Layers: As the image data moves through the intermediate layers, the encoder captures more complex patterns, including shapes and the spatial relationships between different mineral components..

3. Deeper Layers: In the deeper layers of the encoder, the model learns high-level abstract features that represent the global context of the image..

The ResNet-18 backbone enhances this feature extraction capability through its residual connections. These connections allow for the effective training of deeper networks by mitigating the vanishing gradient problem. Residual connections enable the network to learn more complex features without degradation of the performance, thereby ensuring that even the deepest layers contribute valuable information to the encoding process.

In the context of multi-mineral segmentation, the feature-rich output of the encoder is particularly significant. Minerals often exhibit subtle differences in texture, shape, and spatial distribution, making their accurate identification challenging. The encoder's layered approach ensures that both local and global features are captured effectively, providing a comprehensive representation of the mineral content in the input images.

- Local Features: These include fine-grained details such as mineral grain boundaries, small-scale textures, and minute variations in composition, which are crucial for distinguishing between similar-looking minerals.

- Global Features: These encompass broader patterns such as the overall distribution of minerals, large-scale structures, and spatial relationships between different mineral deposits, which are essential for understanding the geological context.

Note that the combination of the U-Net's architectural strengths and the ResNet-18 backbone ensures that the encoder outputs are rich in features that capture both the intricate details and the overall context of the minerals present in the images. ==Table 1 provides a summary of the architecture of the proposed 3D U-Net model with the ResNet-18 backbone.==

2.2.2 GAN network

In this study, we integrate a pre-trained U-Net encoder with a Generative Adversarial Network (GAN) to create a powerful model capable of reconstructing images with multi-mineral segmentation information. This approach leverages the feature-rich output of the U-Net encoder, specifically trained on multimineral segmentation tasks, to enhance the quality of generated images by incorporating detailed mineral information.

The U-Net model used in this study has been pre-trained on a multimineral segmentation dataset. Its encoder consists of multiple convolutional and pooling layers that extract high-level features from input images..

The architecture of the U-Net encoder includes:

- Input Layer and Initial Normalization:

 - The input layer accepts data of shape 128x128x128x1.

 - Initial batch normalization is applied to stabilize the learning process.

- Convolutional Blocks:

 - The first convolutional block applies a 3D convolutional layer with 64 filters of size (3, 3, 3) followed by batch normalization and ReLU activation. The output is then downsampled using max pooling.

 - This block structure is repeated, with the number of filters doubling at each downsampling step (64, 128, 256, 512 filters respectively), reflecting a typical ResNet architecture.

- Residual Blocks:

 - Each residual block consists of two 3D convolutional layers with batch normalization and ReLU activation, followed by an additive identity shortcut connection, enhancing gradient flow during training.

- Final Encoding Layer:

   - The final layer in the encoder path contains 512 filters, resulting in a highly abstract and feature-rich representation of the input data.

The generator model aims to reconstruct high-resolution images using the features extracted by the U-Net encoder. It consists of several layers designed to upsample and refine the encoded features.

The architecture of the generator includes:

- Input Layer:

  - Accepts a 4x4x4x512 feature map from the U-Net encoder.

- Convolutional Transpose Layers:

  - Series of Conv3DTranspose layers to upsample the feature map gradually, doubling the spatial dimensions at each step.

  - These layers have 512, 256, 128, 64, and 32 filters respectively.

- Batch Normalization and LeakyReLU:

  - Applied after each Conv3DTranspose layer to stabilize training and introduce non-linearity.

- Gaussian Noise Layer:

  - Adds noise to the feature maps to improve robustness.

- Final Conv3DTranspose Layer:

  - Produces an output of size 128x128x128 with a single channel, followed by a tanh activation function to normalize the pixel values.

The detailed architecture of the Generator model is summarized in Table 2. A visual overview of the Generator's architecture is presented in Figure 2.

The discriminator evaluates the authenticity of generated images by distinguishing between real and fake samples. It follows a typical convolutional neural network architecture designed for 3D image data.

The architecture of the discriminator includes:

• Input Layer:

- Accepts images of size 128x128x128 with a single channel.

• Convolutional Layers:

- Series of Conv3D layers to downsample the input image, extracting hierarchical features.

- These layers have 64, 128, 256, and 512 filters respectively.

• Batch Normalization and LeakyReLU:

- Applied after each Conv3D layer to stabilize training and introduce non-linearity.

• Final Conv3D Layer:

- Reduces the spatial dimensions to 4x4x4 with a single channel.

• Activation Layer:

- Sigmoid activation to output a probability indicating the authenticity of the input image.

The Discriminator model is detailed in Table 3. A schematic representation of the Discriminator's structure is shown in Figure 3.

The GAN model is trained for 600 epochs, with each epoch taking approximately 15 minutes. The encoder's weights are frozen to preserve the learned mineral features. During training, the generator reconstructs images using the mineral-specific information provided by the encoder, while the discriminator learns to differentiate between real and generated images.

The novelty of this approach lies in the integration of a pre-trained U-Net encoder, specifically trained on multimineral segmentation tasks, with a GAN framework. The use of the ResNet-18 backbone in the U-Net encoder further enhances the feature extraction process, ensuring that detailed mineralogical information is retained in the generated images.

The generator and discriminator are optimized using the Adam optimizer with specific learning rates and momentum parameters to ensure stable training. The learning rate is 1e-4 for generator and 1e-4 for discriminator and momentum parameter is 0.5. The loss functions

involve binary cross-entropy, with the generator's loss calculated based on its ability to fool the discriminator, and the discriminator's loss calculated based on its ability to correctly distinguish real from generated images.

The use of a GAN framework enables the generation of high-quality, mineral-rich images. This approach ensures that the generated images retain detailed mineralogical features, enhancing the overall performance and applicability of the model in geoscientific tasks. This novel architecture demonstrates significant improvements in the accuracy and quality of image reconstructions,.

By feeding the feature tensor produced by the U-Net encoder directly into the GAN generator, the adversarial model is conditioned on segmentation-aware representations of mineral boundaries and pore textures. Stand-alone GANs explore the entire latent space in search of realistic images, but they do so blindly, guided only by the discriminator's real/fake signal; this often leads to blurry grain contacts, missing micro-fractures, or mode collapse. In contrast, the U-Net encoder has already distilled the input into high-level, geology-specific cues—such as the shape of macro-pores, clay rims, and quartz grain edges. When these cues seed the generator:

- Texture fidelity improves: the generator begins its up-sampling path with mineral-specific features rather than isotropic Gaussian noise, so fine rims, laminations, and facies contrasts are preserved instead of being hallucinated.

- Structural consistency is enforced: skip-connection indices from the encoder are injected at matching resolutions in the decoder. This cross-talk keeps large-scale pore topology and mineral contacts aligned, whereas unconditional GANs can drift away from the true spatial relationships.

- Training stabilizes: because the generator's initial activations already resemble plausible rock textures, the discriminator receives realistic candidates earlier in training. That narrows the adversarial gap, shortens convergence, and minimizes mode collapse—problems often reported for 3D-DCGAN and SliceGAN.

- Domain knowledge becomes part of the latent code: the frozen U-Net weights embed prior geological knowledge captured during supervised segmentation. Passing this

2.3 Loss Function

In this hybrid network, , we employ distinct loss functions for the generator and discriminator to optimize the training process.The generator loss is designed to encourage the generation of outputs that are indistinguishable from real data by the discriminator.

The loss is computed by summing two components: the loss when the discriminator correctly identifies real data (real_loss) and the loss when it correctly identifies generated data (generated_loss). In this formulation, the real_loss penalizes the discriminator for misclassifying real data as generated, while the generated_loss penalizes it for misclassifying generated data as real.

The loss functions are crucial for the stability and convergence of the hybrid network training. By strategically combining the U-Net with a GAN, the network benefits from the generative capabilities of the GAN while leveraging the structural advantages of the U-Net.

2.4. Evaluation Metrics

The evaluation of generated samples is conducted using the PoreSpy and OpenPNM packages, which are well-established tools for analyzing and simulating porous structures and properties, with applications across nanotechnology, petrochemical engineering, and materials science.

To assess static features of the generated samples, metrics such as porosity, the two-point correlation function $S_2(r)$, and the linear path function $L(r)$ are utilized. In 3D structures, this probability is calculated along three perpendicular directions, providing insights into the connectivity of local pores.

A graphical representation of the two-point correlation function $S_2(r)$ and linear path function $L(r)$, , is shown in Figure 4.

To investigate flow properties and dynamic characteristics, the relative permeability and capillary pressure of the reconstructed images are examined. This process involves several stages:

- Conversion of the heterogeneous environment images into binary images comprising solid and pore regions using the Otsu algorithm.

- Extraction of the existing network in the image using the Snow algorithm, which employs a watershed-based method to find an appropriate set of initial markers in the image.

- Subsequent extraction of geometric features of the network. These features, including pore and throat sizes, are of significance as they control the entire percolation and transport behavior of the network.

- Addressing any issues related to network clustering. Network clustering refers to the lack of connectivity of some pores with the main body of the network.

- Ensuring the network's integrity, which means that isolated pores, clustered non-connected pores, dead-end throats, repeated throats, and throats without an inlet should not be present in the network.

- Creating phases, where knowledge of the thermodynamic and physical properties of fluids and phases is required.

- Invasion Percolation, a simulation and analysis method used in porous media studies, is introduced. This method is widely employed in physics and engineering for modeling fluid flow in porous structures.

- Definition of the Invasion Percolation algorithm for simulation. Invasion Percolation simulates the process of fluid infiltration in a porous structure. It starts by applying pressure or allowing fluid infiltration from a specified starting point in the porous structure. The fluid then continuously advances step by step in the structure, forming new pores until reaching the exit surface.

- For calculating relative permeability, the Stokes flow algorithm is defined in the simulation. The Stokes flow algorithm is used in physics and engineering for modeling fluid flow in systems with low velocities and slow motion.

- Definition of the multi-phase flow model for relative permeability calculations.

This multi-stage approach allows a comprehensive assessment of both static and dynamic properties, enhancing our understanding of the generated samples' flow and connectivity characteristics within porous media.

3. Results and Discussion

This hybrid model leverages the feature extraction capabilities of U-Net's ResNet-18 backbone to capture intricate details of mineral textures, while the GAN framework enhances the realism of the reconstructed images. Through this unique combination, the proposed network demonstrates significant improvements in accurately segmenting and reconstructing images of porous media,. Figure 5 presents the schematic of the proposed network architecture. As shown, the output of encoder part of U-Net which is feature rich regarding the texture of reservoir, plus a random noise is the input of generative part of our hybrid model.

3.1. UNet

The first part was training UNet. In this study, the Unet model was trained using 3D images of Leopard sandstone, segmented into four distinct classes: macro-pores, clay, quartz, and high-density minerals. The original images were of size 800×800×800 with a voxel length of 2.15 µm, but they were subdivided into smaller sections of 128×128×128 for training purposes. The training process utilized an NVIDIA GeForce RTX 2070 with Max-Q Design GPU with 8 GB memory and was conducted over 80 epochs.

Figure 6 shows a representative input slice, its segmented ground-truth, and the masks produced by three different segmentation strategies: the proposed 3D-U-Net, Otsu's global thresholding, and k-means clustering. The U-Net prediction reproduces the complex pore network and mineral boundaries far more faithfully than the two unsupervised baselines, which both over smooth the fracture system and miss smaller clay pockets. This qualitative observation is mirrored by the quantitative scores: the U-Net attains an Intersection-over-Union (IoU) of 0.87 and an F1-score of 0.93, whereas Otsu and k-means achieve only 0.32 / 0.33 and 0.32 / 0.33, respectively. This figure therefore confirms—visually and numerically—the superiority of the learning-based approach for mineral-specific segmentation.

- Training and Validation Loss:

Figure 7 displays the training and validation loss over 80 epochs, providing insights into the UNet model's learning process. Initially, both losses decrease rapidly, indicating that the model is quickly grasping fundamental patterns within the data.

Around epoch 30, the validation loss stabilizes, while the training loss continues to decline gradually before leveling off at epoch 70. The close alignment between training and validation loss indicates minimal overfitting, suggesting that the model has learned to generalize well to unseen data. The steady convergence pattern highlights the model's stability and efficiency in learning complex features.

The minimal gap between the training and validation loss curves demonstrates that the UNet model is not only learning effectively but is also robust against overfitting, making it reliable for real-world data applications.

- Intersection over Union (IOU):

Figure 8 illustrates the IOU metric, a critical measure for segmentation tasks, as it captures the overlap between predicted and actual segmentation masks. Both the training and validation IOU scores exhibit consistent growth, ultimately reaching around 0.7.

The gradual increase in IOU reflects the model's improving ability to align its predictions with the ground truth. Although there is a slight gap between training and validation IOU, the curves remain closely aligned, indicating that the model generalizes well without overfitting. The steady climb in IOU across epochs also suggests the model is progressively refining its understanding.

With an IOU score of approximately 0.7, the UNet model effectively captures spatial relationships in segmentation, showing a high degree of accuracy. The minimal discrepancy between training and validation IOU further reinforces the model's strong generalization capabilities, making it reliable for precise segmentation.

- F1 Score:

Figure 9 tracks the F1 score, a metric combining precision and recall, particularly important for imbalanced segmentation tasks. Starting at around 0.45, the F1 score climbs steadily, reaching about 0.72 on the validation set by the final epochs.

The continuous rise in F1 score demonstrates that the model is successfully balancing precision and recall, accurately identifying relevant segments while minimizing false positives and false negatives. The close alignment between training and validation F1 scores suggests a well-generalized model that avoids overfitting and achieving a balanced performance across the dataset.

The high F1 score at the end of training confirms that the UNet model performs well in this segmentation task, handling the balance between precision and recall effectively. The minimal gap between training and validation scores further validates the model's ability to maintain consistent performance across different data samples.

The training and validation curves for loss, IOU, and F1 score collectively indicate that the UNet model is both accurate and well-generalized. The minimal gap between training and validation metrics across all curves points to a robust model that resists overfitting while achieving high segmentation accuracy. The final high IOU and F1 scores underscore the model's suitability for complex segmentation tasks, capturing intricate spatial patterns effectively.

- Data Preprocessing and Model Training:

The decision to divide the 800×800×800 3D images into smaller 128×128×128 patches proved advantageous for memory management during training, as it allowed efficient utilization of the GPU. However, this approach may introduce challenges related to boundary effects, which could potentially impact segmentation accuracy at the edges of patches. Future work could explore techniques such as overlap-tile strategies or larger patch sizes to mitigate these boundary issues. Additionally, further training or fine-tuning might enhance the model's performance, as indicated by the ongoing improvement in both IOU and F1 scores even after 80 epochs. Also in figure 10, the visual prediction masks are presented.

3.2. GAN-Based Image Reconstruction

Following the segmentation with U-Net, the next phase involved image reconstruction using a Generative Adversarial Network (GAN)..

Comparison of Reconstructed Images Across Epochs

To evaluate the performance of the proposed model, reconstructed images were generated and analyzed at different training epochs (120, 240, 360, and 480). As depicted in Figure 11, the quality

of the reconstructed images shows a marked improvement as training progresses. At 120 epochs, the reconstructions produced by the 3D-DCGAN model are highly noisy and lack discernible structures, reflecting its inability to capture the underlying mineral textures effectively. In contrast, the proposed hybrid model, even at an early training stage, begins to approximate the true textures, albeit with some blurring.

As training continues to 240 epochs, the superiority of the hybrid model becomes evident. The reconstructed images exhibit more defined boundaries between minerals and a more accurate representation of the porous structure. By 360 epochs, the hybrid model outperforms 3D-DCGAN, which still struggles with maintaining the integrity of mineral boundaries. At 480 epochs, the proposed model achieves reconstructions that are not only visually similar to the original but also structurally accurate, capturing fine details that are critical for accurate multi-mineral segmentation.

The comparison clearly demonstrates that the proposed hybrid model significantly outperforms the conventional 3D-DCGAN. While 3D-DCGAN fails to preserve the intricate details and textures, particularly at lower epochs, the hybrid model consistently produces high-quality reconstructions, with a progressive enhancement in both visual fidelity and structural accuracy.

To determine the information that is lost in a synthetic low-quality (Low-Q) scan—and how much of it is recovered by the hybrid U-Net + GAN—we compared a volume against the ground-truth high-resolution (Hi-Res) scan with three full-reference metrics: mean-squared error (MSE, lower = better), peak-signal-to-noise ratio (PSNR, higher = better) and structural similarity index (SSIM, higher = better). The results of the quantitative comparison between the Low-Q and Hi-Res scans, along with the improvements achieved by the proposed hybrid model, are presented in Table 4, highlighting significant enhancements in MSE, PSNR, and SSIM values.

A 4 dB PSNR gain and four-fold MSE drop indicate that most Gaussian blur and sensor noise have been suppressed. The SSIM rise from 0.70 to 0.85 shows that the reconstruction recovers crucial textural patterns.

Porosity and Absolute Permeability Analysis

The quantitative metrics such as porosity and absolute permeability were evaluated for the reconstructed images. Table 5 presents a comparison of these parameters between the original image and the reconstructed images averaged over 10 realizations from the hybrid model.

The porosity of the reconstructed images is slightly overestimated at 36.3%, compared to the actual porosity of 35.5%. The absolute permeability of the reconstructed images averages at 263 mD, closely aligning with the actual value of 259 mD.

These results indicate that the hybrid U-Net and GAN model is not only effective in reconstructing the geometric structure of porous media but also in accurately reproducing flow properties..

Capillary Pressure and Non-Wetting Phase Saturation

Figure 12 illustrates the relationship between capillary pressure and non-wetting phase saturation for both the original image and the GAN-reconstructed image. The capillary pressure increases steadily with the non-wetting phase saturation, indicating a strong correlation between the two variables. The comparison between the original and reconstructed images shows a slight deviation, with the GAN-reconstructed data slightly underestimating the non-wetting phase saturation at lower capillary pressures..

This result demonstrates the success of the GAN in preserving the macroscopic behavior of the system, as the reconstruction accurately reflects the non-linear capillary pressure response..

Two-Point Correlation Function

The two-point correlation function is a critical metric for assessing the spatial distribution of phases in porous media. Figure 13 compares the two-point correlation function of the original image and the GAN-reconstructed image. Both curves show a rapid decay, indicating that the spatial correlation diminishes as the distance between points increases.

Although the GAN-reconstructed image slightly deviates from the original in terms of the correlation at shorter distances, the overall trend is well-preserved. This consistency highlights the model's ability to replicate the fine-scale structural heterogeneity of the porous media. The minor discrepancies observed at short distances could be addressed by refining the GAN's loss function to emphasize small-scale structural features, potentially enhancing its reconstruction accuracy.

Relative Permeability

The relative permeability curves in Figure 14 provide a more detailed analysis of the fluid flow properties for both wetting and non-wetting phases in the x direction. The original and reconstructed images show similar trends in relative permeability for both phases,.

The GAN-reconstructed image successfully captures the transition points and the overall shape of the relative permeability curves. However, some divergence is observed in the intermediate saturation range, where the GAN reconstruction slightly overestimates the relative permeability of the non-wetting phase and underestimates that of the wetting phase. This deviation could be attributed to the complex nature of fluid interactions in porous media, which may require more advanced modeling techniques or a more refined GAN architecture to capture accurately.

The results of this study demonstrate the effectiveness of the hybrid U-Net and GAN architecture in both segmenting and reconstructing images of porous media. The integration of U-Net's feature extraction capabilities with the generative modeling of GAN allowed for the accurate recreation of both mineral textures and macroscopic fluid properties.

Additionally, increasing the patch size for U-Net segmentation or using a more advanced overlap-tile strategy could help mitigate boundary effects and further enhance segmentation performance.

Overall, the proposed model offers a promising approach for improving multi-mineral segmentation and texture-based image reconstruction in complex porous media, with potential applications in fields such as hydrogeology, petroleum engineering, and environmental science.

Computational considerations

- Memory footprint: The hybrid stacks a full 3-D U-Net encoder on top of a 3-D generator. Even with the encoder frozen, the model occupies noticeably more GPU memory than a stand-alone 3-D-DCGAN and roughly the same as two separate networks running back-to-back. Patch-wise training and mixed-precision arithmetic mitigate the overhead, but commodity laptop hardware remains insufficient.

- Training time: Although the hybrid converges in fewer epochs than an unconditional GAN, each epoch processes two networks in tandem. In wall-clock terms it therefore sits between a plain U-Net (fast) and a full-blown 3-D-DCGAN (slow). For projects where turnaround

time outweighs ultra-high texture fidelity, a lighter 2-D conditional GAN may still be preferable.

- Inference speed: During deployment the encoder executes once per volume and then hands off its feature tensor, so generation is only moderately slower than a vanilla GAN. Slice-based reconstructions such as SliceGAN remain faster in absolute terms, but they pay for that speed with reduced three-dimensional coherence.

- Parameter efficiency: Freezing the encoder keeps the count of trainable weights on par with a conventional GAN, yet the total parameter count is higher. Storage and checkpoint transfer therefore demand more disk and network bandwidth.

In workflows where geological realism and pore-scale connectivity are the dominant concerns—and ample GPU memory is available—the U-Net-conditioned GAN remains the most compelling choice. Conversely, when rare-phase preservation, minimal engineering effort, or strict hardware budgets top the priority list, earlier approaches such as classical conditioning, 2-D slice GANs, or even rule-based up-scaling can still be the better fit. Articulating these trade-offs clarifies how the proposed integration should be selected or adapted for real-world deployments rather than applied as a one-size-fits-all solution.

4. Conclusion

This study presents a hybrid U-Net and GAN model that significantly advances the segmentation and reconstruction of 3D multi-mineral textures in porous media images, specifically Leopard sandstone. Our approach combines the U-Net's robust feature extraction—powered by a ResNet-18 backbone pre-trained for mineral segmentation—with a GAN structure, allowing for the synthesis of high-resolution images that retain the original mineral textures and structural details. Quantitative evaluations confirm the model's accuracy, with reconstructed images achieving porosity estimates of 36.3% (compared to the actual 35.5%) and absolute permeability values of 263 mD (compared to the actual 259 mD), resulting in minor error margins of 2.25% for porosity and 1.54% for permeability. This accuracy indicates the model's capability in generating reliable 3D images essential for fluid dynamics simulations within porous media..

The innovation of this study lies in the integration of U-Net's segmentation-driven texture data with GAN's image synthesis capabilities, a combination previously unexplored for porous media

image reconstruction. This approach provides a more comprehensive solution by considering the impact of rock texture,. Unlike earlier models that separately handled segmentation and reconstruction, our model consolidates these processes, marking a substantial step toward addressing the complex structure and flow properties of rock formations. Future research will focus on further GAN optimizations to improve fluid interaction modeling and refining segmentation processes to enhance overall model accuracy and application in fields such as hydrogeology, reservoir simulation, and environmental engineering.

5. References

1. Tahmasebi, P., "Accurate modeling and evaluation of microstructures in complex materials", Phys. Rev. E, 97(2-1), (2018), doi: https://doi.org/10.1103/PhysRevE.97.023307.

2. Baruchel, J., et al., "Advances in synchrotron hard X-ray based imaging", C. R. Physique, 9(5–6), pp. 624–641 (2008), doi: https://doi.org/10.1016/j.crhy.2007.08.003.

3. Mostaghimi, P., Blunt, M.J., and Bijeljic, B., "Computations of absolute permeability on micro-CT images", Math. Geosci., 45(1), pp. 103–125 (2013), doi: https://doi.org/10.1007/s11004-012-9431-4.

4. Kamrava, S., Tahmasebi, P., and Sahimi, M., "Enhancing images of shale formations by a hybrid stochastic and deep learning algorithm", Neural Netw., 118, pp. 310–320 (2019), doi: https://doi.org/10.1016/j.neunet.2019.07.009.

5. Tahmasebi, P., "Nanoscale and multiresolution models for shale samples", Fuel, 217, pp. 218–225 (2018), doi: https://doi.org/10.1016/j.fuel.2017.12.107.

6. Javadpour, F., Fisher, D., and Unsworth, M., "Nanoscale gas flow in shale gas sediments", J. Can. Pet. Technol., 46(10), pp. 55–61 (2007), doi: https://doi.org/10.2118/07-10-06.

7. Kulikowski, J.L., "Computer-Aided Porous Materials Description Based on Morphological Spectra", Res. Dev. Mater. Sci., 11(4) (2019), doi: https://doi.org/10.31031/RDMS.2019.11.000769.

8. Yang, Q., Yao, J., Huang, Z., Zhu, G., Liu, L., and Song, W., "Pore-scale investigation of petro-physical fluid behaviors based on multiphase SPH method", J. Pet. Sci. Eng., 192, 107238 (2020), doi: https://doi.org/10.1016/j.petrol.2020.107238.

9. Rabbani, A., Jamshidi, S., and Salehi, S., "An automated simple algorithm for realistic pore network extraction from micro-tomography images", J. Pet. Sci. Eng., 123, pp. 164–171 (2014), doi: https://doi.org/10.1016/j.petrol.2014.08.020.

10. Niu, Y., Mostaghimi, P., Shabaninejad, M., et al., "Digital rock segmentation for petrophysical analysis with reduced user bias using convolutional neural networks", Water Resour. Res., 56 (2020), doi: https://doi.org/10.1029/2019WR026597.

11. Shams, R., Masihi, M., Boozarjomehry, R.B., et al., "Coupled generative adversarial and auto-encoder neural networks to reconstruct three-dimensional multi-scale porous media", J. Pet. Sci. Eng., 186, 106794 (2020), doi: https://doi.org/10.1016/j.petrol.2019.106794.

12. Varfolomeev, I., Yakimchuk, I., and Safonov, I., "An application of deep neural networks for segmentation of microtomographic images of rock samples", Computers, 8(4), p. 72 (2019), doi: https://doi.org/10.3390/computers8040072.

13. Badrinarayanan, V., Kendall, A., and Cipolla, R., "SegNet: A deep convolutional encoder-decoder architecture for image segmentation", IEEE Trans. Pattern Anal. Mach. Intell., 39(12), pp. 2481–2495 (2017), doi: https://doi.org/10.1109/TPAMI.2016.2644615.

14. Ronneberger, O., Fischer, P., and Brox, T., "U-Net: Convolutional networks for biomedical image segmentation", in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham: Springer, pp. 234–241 (2015), doi: https://doi.org/10.1007/978-3-319-24574-4_28.

15. Ar Rushood, I., Alqahtani, N., Wang, Y.D., et al., "Segmentation of X-ray images of rocks using deep learning", in SPE Annual Technical Conference and Exhibition, (2020), doi: https://doi.org/10.2118/201306-MS.

16. Feng, J., He, X., Teng, Q., et al., "Accurate and fast reconstruction of porous media from extremely limited information using conditional generative adversarial network", Phys. Rev. E, 100(3) (2019), doi: https://doi.org/10.1103/PhysRevE.100.033308.

17. Kamrava, S., Sahimi, M., and Tahmasebi, P., "Quantifying accuracy of stochastic methods of reconstructing complex materials by deep learning", Phys. Rev. E, 101(4), p. 043301 (2020), doi: https://doi.org/10.1103/PhysRevE.101.043301.

18. Kamrava, S., Tahmasebi, P., and Sahimi, M., "Linking morphology of porous media to their macroscopic permeability by deep learning", Transp. Porous Media, 131(2), pp. 427–448 (2020), doi: https://doi.org/10.1007/s11242-019-01352-5.

19. Goodfellow, I.J., et al., "Generative Adversarial Networks", (2014), [Online]. Available: https://arxiv.org/abs/1406.2661.

20. Wang, Y.D., Shabaninejad, M., Armstrong, R.T., and Mostaghimi, P., "Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images", Appl. Soft Comput., 104 (2021), doi: https://doi.org/10.1016/j.asoc.2021.107185.

21. Valsecchi, A., Damas, S., Tubilleja, C., and Arechalde, J., "Stochastic reconstruction of 3D porous media from 2D images using generative adversarial networks", Neurocomputing, 399, pp. 227–236 (2020), doi: https://doi.org/10.1016/j.neucom.2019.12.040.

22. Feng, J., Teng, Q., Li, B., et al., "An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning", Comput. Methods Appl. Mech. Eng., 368, 113043 (2020), doi: https://doi.org/10.1016/j.cma.2020.113043.

23. Zhang, T.F., Tilke, P., Dupont, E., et al., "Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks", Pet. Sci., 16(3), pp. 541–549 (2019), doi: https://doi.org/10.1007/s12182-019-0328-4.

24. Zhang, F., Teng, Q., Chen, H., He, X., and Dong, X., "Slice-to-voxel stochastic reconstructions on porous media with hybrid deep generative model", Comput. Mater. Sci., 186, 110018 (2021), doi: https://doi.org/10.1016/j.commatsci.2020.110018.

25. Feng, J., Teng, Q., He, X., et al., "Accelerating multi-point statistics reconstruction method for porous media via deep learning", Acta Mater., 159, pp. 296–308 (2018), doi: https://doi.org/10.1016/j.actamat.2018.08.026.

26. Feng, J., He, X., Teng, Q., Ren, C., Chen, H., and Li, Y., "Reconstruction of porous media from extremely limited information using conditional generative adversarial networks", Phys. Rev. E, 100(3), p. 033308 (2019), doi: https://doi.org/10.1103/PhysRevE.100.033308.

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

28. Stark, J., and Sbalzarini, I.F., "An open-source pipeline for solving continuous reaction–diffusion models in image-based geometries of porous media", J. Comput. Sci., 72, 102118 (2023), doi: https://doi.org/10.1016/j.jocs.2023.102118.

29. Siavashi, J., Najafi, A., Ebadi, M., and Sharifi, M., "A CNN-based approach for upscaling multiphase flow in digital sandstones", Fuel, 308, 122047 (2022), doi: https://doi.org/10.1016/j.fuel.2021.122047.

30. Yan, P., Teng, Q., He, X., Ma, Z., and Zhang, N., "Multiscale reconstruction of porous media based on multiple dictionaries learning", arXiv preprint (2022), doi: https://arxiv.org/abs/2205.08278.

31. Yang, Y., Liu, F., Yao, J., et al., "Multi-scale reconstruction of porous media from low-resolution core images using conditional generative adversarial networks", J. Nat. Gas Sci. Eng., 99, 104411 (2022), doi: https://doi.org/10.1016/j.jngse.2022.104411.

32. Amiri, H., Vogel, H., and Plümper, O., "New 2D to 3D reconstruction of heterogeneous porous media via deep generative adversarial networks (GANs)", J. Geophys. Res.: Mach. Learn. Comput., 1, e2024JH000178 (2024), doi: https://doi.org/10.1029/2024JH000178.

33. Zhang, F., He, X., Teng, Q., et al., "3D-PMRNN: Reconstructing three-dimensional porous media from the two-dimensional image with recurrent neural network", J. Pet. Sci. Eng., 208, 109652 (2022), doi: https://doi.org/10.1016/j.petrol.2021.109652.

34. Masihi, M., Shams, R., and King, P. R., "Pore level characterization of Micro-CT images using percolation theory", J. Pet. Sci. Eng., 211, 110113 (2022), doi: 10.1016/j.petrol.2022.110113.

35. Moslemipour, A., Sadeghnejad, S., Enzmann, F., et al., "Image-based multi-scale reconstruction of unresolved microporosity in 3D heterogeneous rock digital twins using cross-correlation simulation and watershed algorithm", Transp. Porous Media, 151, 2215–2240 (2024), doi: 10.1007/s11242-024-02111-x.

36. Reinhardt, M., Jacob, A., Sadeghnejad, S., et al., "Benchmarking conventional and machine learning segmentation techniques for digital rock physics analysis of fractured rocks", Environ. Earth Sci., 81, 71 (2022), doi: 10.1007/s12665-021-10133-7.

6. Biographies

- **Bahareh Keshavarz** is a Petroleum Engineering Ph.D. student at Islamic Azad University (Science and Research Branch), with M.Sc. (Tarbiat Modares Univ.) and B.Sc. (Amirkabir Univ. of Tech.) degrees. Her research focuses on deep learning for 3D image reconstruction in porous media, with expertise in Python, data science, AI, and asphaltene precipitation.

- **Dr. Mohsen Masihi**, Professor of Petroleum Engineering at Sharif University of Technology, holds a Ph.D. (Imperial College London, 2006). His research spans reservoir modeling, fractured reservoirs, and percolation theory, with over 160 publications and several books. He has held directorships at Sharif University and serves as Editor-in-Chief for the Journal of Petroleum Research.

- **Dr. Mastaneh Hajipour Shirazi Fard**, Assistant Professor in Petroleum Engineering at Islamic Azad University (IAU, Science and Research Branch), earned her Ph.D. in Chemical Engineering from Sharif University of Technology. Her research includes computational fluid dynamics (CFD), transport in porous media, nanofluids, and formation damage.

- **Dr. Ebrahim Biniaz Delijani**, Assistant Professor in Petroleum Engineering at Islamic Azad University (IAU, Science and Research Branch), holds a Ph.D. in Petroleum Engineering from Sharif University of Technology. His research focuses on transport phenomena in porous media, particularly enhancing drilling fluid properties with nanoadditives and understanding fluid flow.

7. Statements & Declarations

- **Funding**

- **Competing Interests**

The authors have no relevant financial or non-financial interests to disclose.

- **Author Contributions**

Bahareh Keshavarza conducted the research, performed data collection and analysis, and prepared the manuscript draft. Mohsen Masihi, Mastaneh Hajipour Shirazi Farda, and Ebrahim Biniaz Delijani provided guidance and critical revisions throughout the study. All authors read and approved the final manuscript.

- **Data Availability**

The datasets used during this study are publicly available on the Digital Rocks Portal and can be accessed at https://www.digitalrocksportal.org.
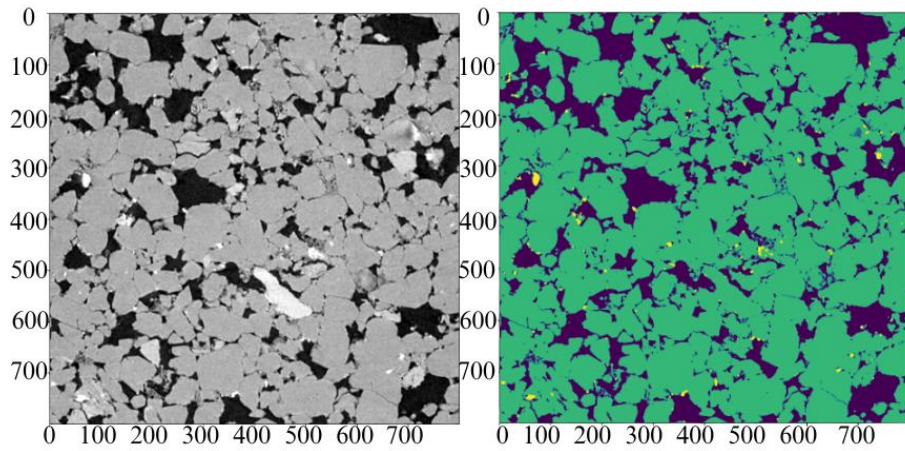
Figures:

Figure 1. Image of the Leopard sandstone reservoir and its corresponding segmentation image.
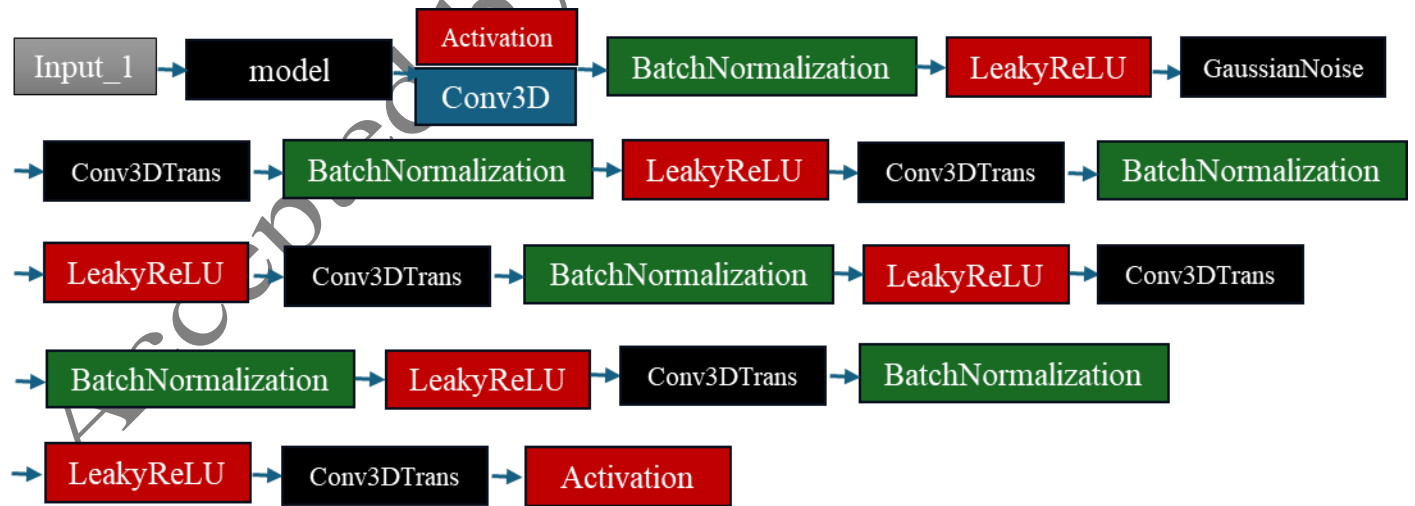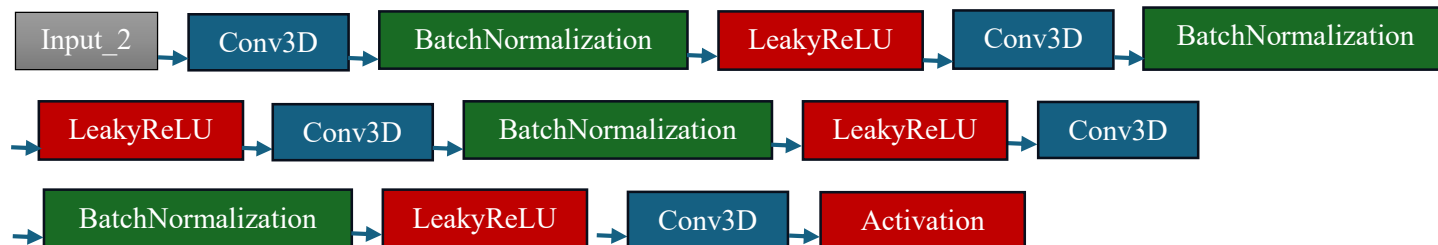
Figure 2. Architecture of the Generator model.

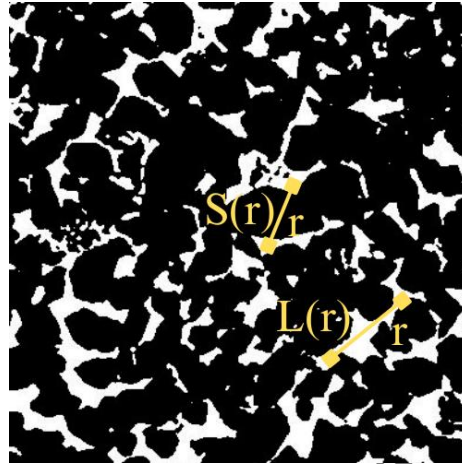Figure 3. Architecture of the Discriminator model.

Figure 4. Representation of porous structure features, showing the two-point correlation function $S_2(r)$ and linear path function $L(r)$.
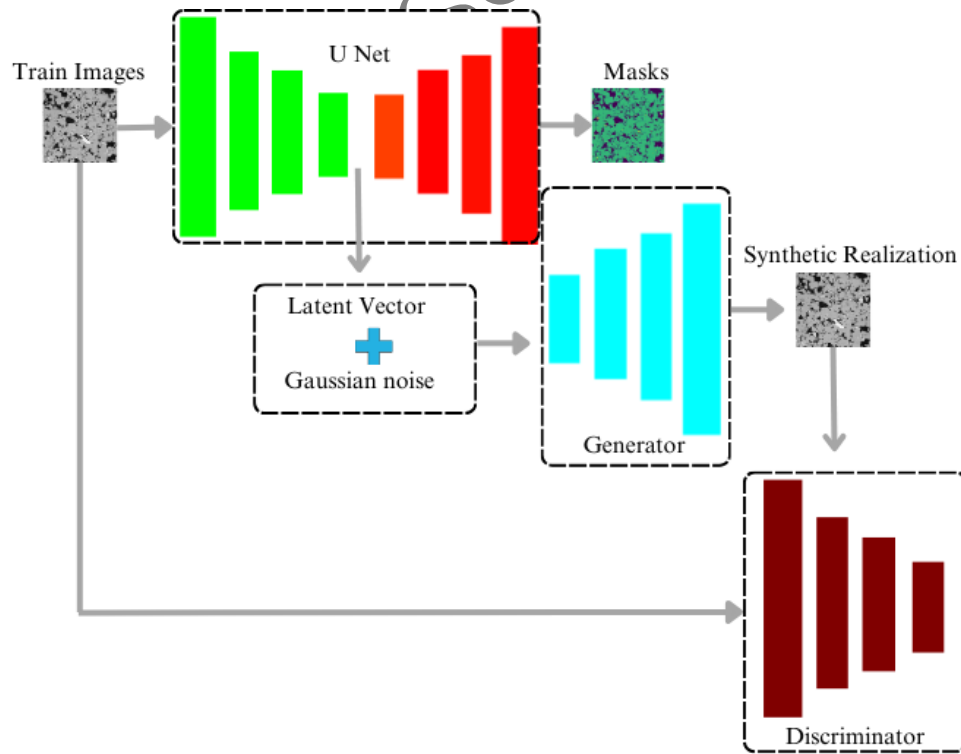
Figure 5. The schematic of the proposed network architecture.

Figure 6. Visual comparison of segmentation results on a Leopard-sandstone slice. (a) input, (b) ground-truth mask, (c) 3D-U-Net prediction (F1 = 0.93), (d) Otsu thresholding (F1 = 0.33), and (e) k-means clustering (F1 = 0.33).

Figure 7. Training and Validation Loss of UNet model.

Figure 8. IOU of UNet model.

Figure 9. F1 Score of UNet model.

Figure 10. Masks predicted by UNet model.

Figure 11. Comparison of 3D-DCGAN and the proposed hybrid model's performance in reconstructing mineral textures at different training epochs (120, 240, 360, and 480).

Figure 12. Capillary pressure vs. non-wetting phase saturation for original and GAN-reconstructed images

Figure 13. Two-point correlation function for original and GAN-reconstructed images.

Figure 14. Original and GAN-reconstructed relative permeability curves.

Figure 1. Image of the Leopard sandstone reservoir and its corresponding segmentation image.

Figure 2. Architecture of the Generator model.

Figure 3. Architecture of the Discriminator model.

Figure 4. Representation of porous structure features, showing the two-point correlation function $S_2(r)$ and linear path function $L(r)$.

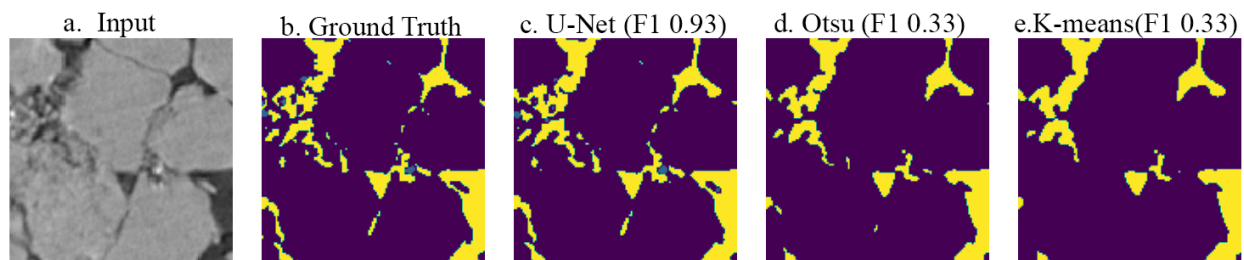Figure 5. The schematic of the proposed network architecture.

Figure 6. Visual comparison of segmentation results on a Leopard-sandstone slice. (a) input, (b) ground-truth mask, (c) 3D-U-Net prediction (F1 = 0.93), (d) Otsu thresholding (F1 = 0.33), and (e) k-means clustering (F1 = 0.33).
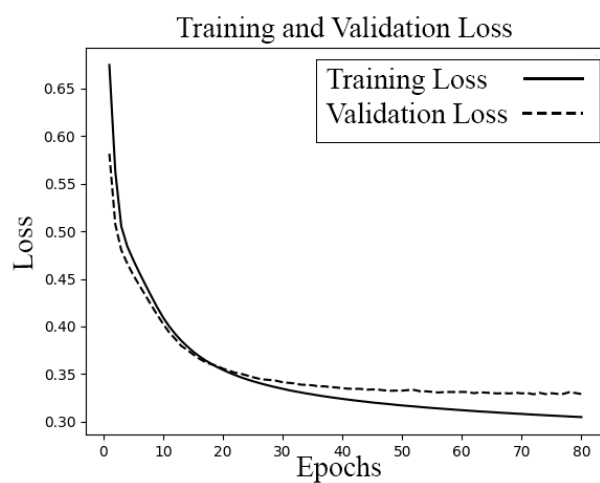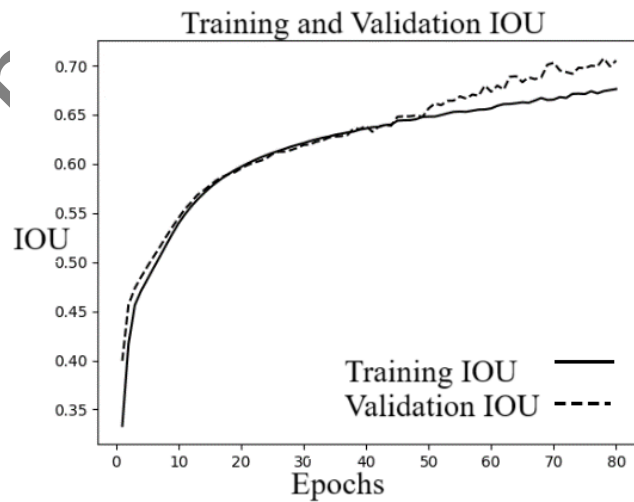

Figure 7. Training and Validation Loss of UNet model.

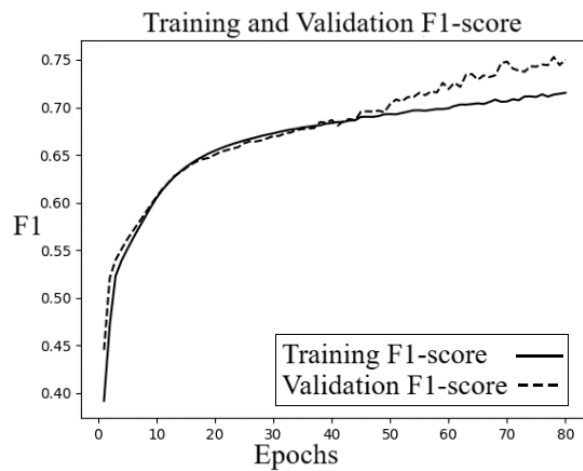Figure 8. IOU of UNet model.
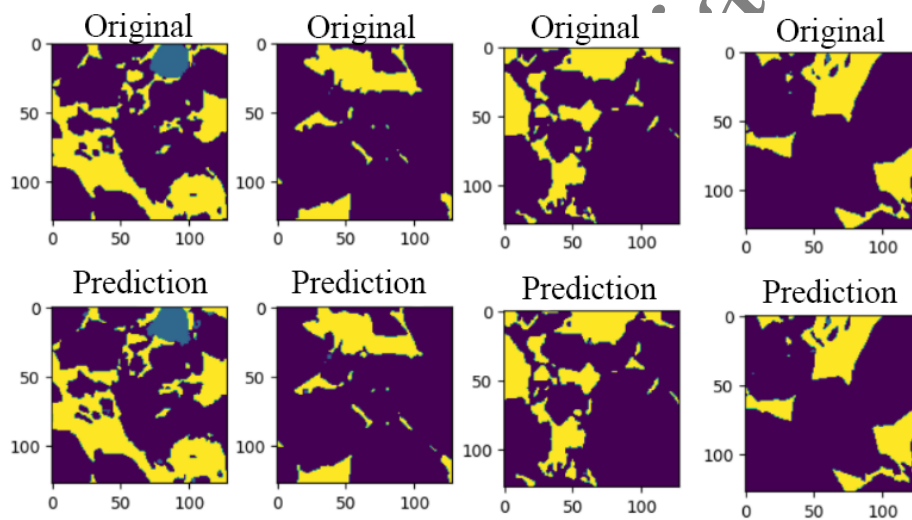
Figure 9. F1 Score of UNet model.

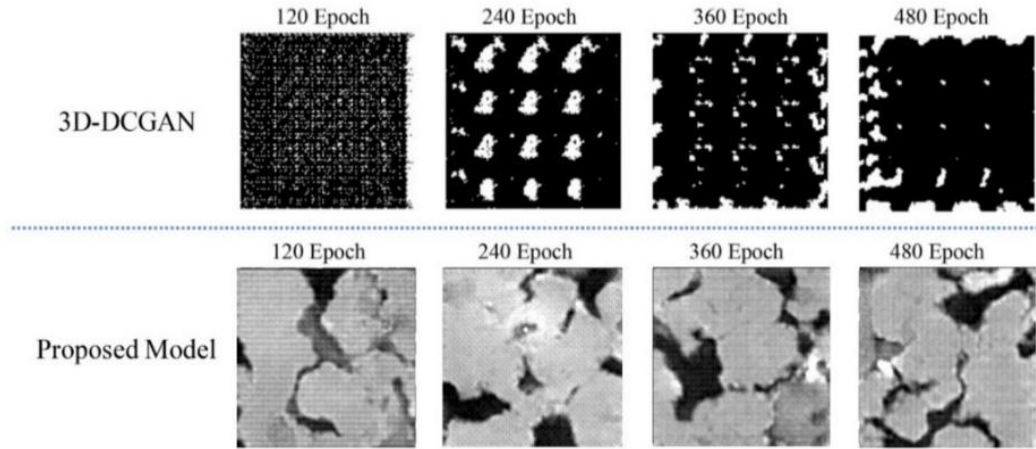Figure 10. Masks predicted by UNet model.

Figure 11. Comparison of 3D-DCGAN and the proposed hybrid model's performance in reconstructing mineral textures at different training epochs (120, 240, 360, and 480).
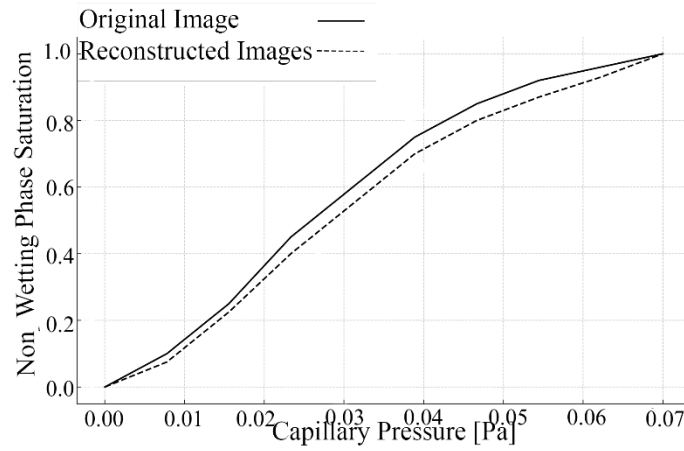

Figure 12. Capillary pressure vs. non-wetting phase saturation for original and GAN-reconstructed images
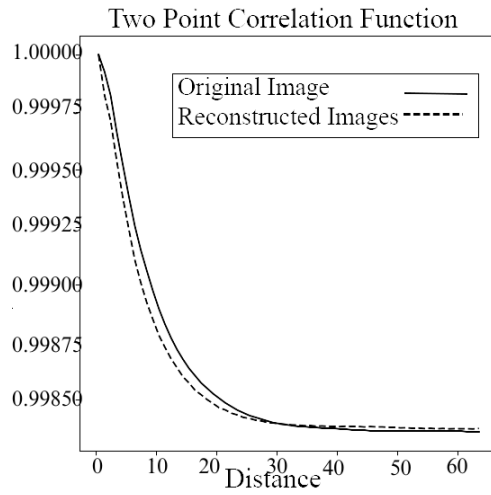
Figure 13. Two-point correlation function for original and GAN-reconstructed images.
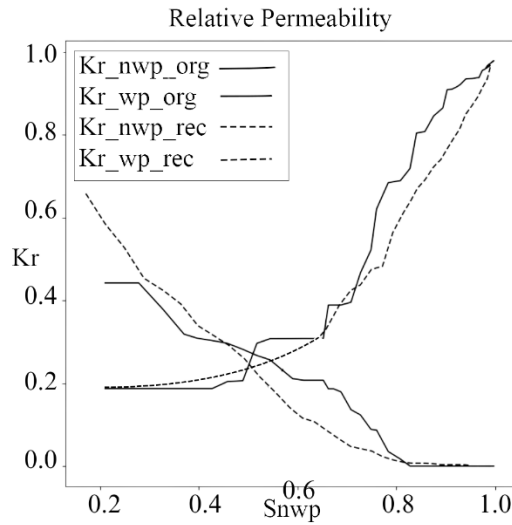

Figure 14. Original and GAN-reconstructed relative permeability curves.

Tables:

Table 1. Architecture of the 3D U-Net model

Table 2. Architecture of the Generator model.

Table 3. Architecture of the Discriminator model.

Table 4. Quantitative evaluation of information loss and recovery

Table 5. Comparison of actual and reconstructed images in terms of porosity and absolute permeability.

Table 1. Architecture of the 3D U-Net model

Layer	Type	Filters	Padding	Batch Normalization	Activation
1	Input	-	0	yes	-
2	Conv3D	64	0	yes	Relu
3	MaxPooling3D	-	0	yes	Relu
4	Conv3D	128	0	yes	Relu
5	Conv3D	256	0	yes	Relu
6	Conv3D	512	0	yes	Relu
7	Conv3D	256	0	yes	Relu
8	Conv3D	128	0	yes	Relu
9	Conv3D	64	0	yes	Relu
10	Conv3D	32	0	yes	Relu
11	Conv3D	16	0	yes	Relu
12	Conv3D	4	0	yes	Softmax
Total params	42,624,775	Trainable params	42,614,853	Non-trainable params	9,922

Table 2. Architecture of the Generator model.

Layer	Type	Filters	Padding	Batch Normalization	Activation
1	Input	-	0	no	-
2	Encoder	-	0	no	LeakyRelu
3	Conv3D	512	0	yes	LeakyRelu
4	GaussianNoise	-	0	no	LeakyRelu
5	Conv3DTrans	512	0	yes	LeakyRelu
6	Conv3DTrans	256	0	yes	LeakyRelu
7	Conv3DTrans	128	0	yes	LeakyRelu
8	Conv3DTrans	64	0	yes	LeakyRelu
9	Conv3DTrans	32	0	yes	LeakyRelu
10	Conv3DTrans	1	0	No	tanh
Total params	68,174,979	Trainable params	35,001,600	Non-trainable params	33,173,379

Table 3. Architecture of the Discriminator model.

Layer	Type	Filters	Padding	Batch Normalization	Activation
1	Input	-	0	no	-
2	Conv3D	64	0	yes	LeakyRelu
3	Conv3D	128	0	yes	LeakyRelu
4	Conv3D	256	0	yes	LeakyRelu
5	Conv3D	512	0	yes	LeakyRelu
6	Conv3D	1	0	no	Sigmoid
Total params	11,051,713	Trainable params	11,049,793	Non-trainable params	1,920

Table 4. Quantitative evaluation of information loss and recovery

Metric	Low-Q vs Hi-Res	Predicted vs Hi-Res	Improvement
MSE	$2*10^{-3}$	$5*10^{-4}$	$\downarrow \approx 4 \times$
PSNR	28	32	$\uparrow 4$
SSIM	0.7	0.85	$\uparrow 0.15$

Table 5. Comparison of actual and reconstructed images in terms of porosity and absolute permeability.

Parameter	Actual Image	Reconstructed Image (mean of 10 realizations)	Error percentage
Porosity (%)	35.5	36.3	2.25
Absolute Permeability (mD)	259	263	1.54