High-Fidelity Synthesis of Porous Rock Microstructures Using a Guided Conditional Generative Adversarial Network

Ehsan khani

11 August 2025

Contents

Abstract

The computational modeling of porous media is a cornerstone of modern geoscience and materials science. While X-ray microtomography (micro-CT) provides high-resolution 3D images of rock microstructures, the generation of statistically equivalent digital samples is crucial for robust physical property simulation. This report details a deep learning framework for synthesizing realistic 2D rock porosity using a guided conditional Generative Adversarial Network (cGAN). A U-Net-based generator and a PatchGAN discriminator were trained on 2D slices from a cylindrical rock core micro-CT scan. Critically, the generator was conditioned on both a solid rock texture and a precise segmentation map, providing explicit spatial guidance for pore creation. We outline a complete data preparation pipeline, including circular masking and robust binarization, and detail a comprehensive suite of evaluation metrics that go beyond simple image similarity. The model's performance was rigorously assessed by comparing the porosity, topological connectivity (Euler number), and the two-point probability function (S2) of generated samples against the ground truth. The results demonstrate that this guided approach can produce pore networks with a high degree of visual realism and quantitative structural fidelity, validating the use of conditional GANs for advanced digital rock physics applications.

Introduction

Background

The microstructure of porous materials like sedimentary rocks governs their most critical physical properties, including permeability, porosity, and mechanical strength. Understanding and predicting these properties is essential for applications ranging from hydrocarbon reservoir engineering and groundwater management to carbon sequestration and the development of new synthetic materials.

Digital Rock Physics (DRP) leverages high-resolution 3D imaging, primarily from X-ray microtomography (micro-CT), to create detailed digital models of these microstructures. These digital samples can then be used in numerical simulations to directly compute their physical properties. However, a significant limitation of DRP is the finite, and often small, sample size obtained from a physical scan.

The Need for Generative Models

To overcome the limitation of finite sample size, researchers turn to **digital rock synthesis**—the computational generation of new, statistically equivalent microstructures. This allows for the creation of larger, more representative sample volumes, leading to more accurate simulations. While traditional methods for this synthesis exist, they often rely on statistical algorithms that may fail to capture the complex, higher-order geometric features present in real geological formations.

In recent years, **Generative Adversarial Networks (GANs)** have emerged as a state-of-the-art technique for high-fidelity image synthesis. Their ability to learn complex, non-linear data distributions makes them an ideal candidate for modeling the intricate patterns of rock porosity.

Objective and Contribution

The primary objective of this project is to develop and rigorously evaluate a deep learning framework for the high-fidelity synthesis of 2D rock pore structures. The key contributions are:

1. **A Guided cGAN Framework:** We implement a conditional GAN (cGAN) where the generator is explicitly guided by a ground truth segmentation map. This transforms the task from an unconstrained generation problem to a more precise image-to-image translation problem.
2. **A Robust Preprocessing Pipeline:** We detail a complete data workflow, including circular masking for core samples and a precise binarization method, to create clean, high-quality data suitable for training.

3. **Comprehensive, Physics-Based Evaluation:** We go beyond standard computer vision metrics (like SSIM or PSNR) and evaluate the model using a suite of physics-based metrics, including topological connectivity and the two-point probability function, which are more relevant for materials science applications.

Materials and Methods

Dataset

The dataset used in this study consists of 887 2D image slices obtained from a micro-CT scan of a cylindrical rock core. Each slice is represented by two corresponding images:

- **Raw Image:** A 1024×1024 pixel grayscale image with a 16-bit depth (uint16), where pixel intensity corresponds to material density.
- **Segmentation Mask:** A multi-class uint8 image of the same dimension, where pixels are labeled as solid rock matrix (value 1) or pore types (value 2). This ground truth segmentation was created through a semi-automated process involving expert manual refinement.

Data Preprocessing Pipeline

A robust preprocessing pipeline was designed to prepare the data for the guided GAN model. The pipeline transforms the source images into a set of three co-registered tensors: a solid rock input, a segmentation guidance map, and a ground truth porous rock image.

1. **Solid Rock Generation:** The solid rock input image was created by in-painting the pore regions in the raw image with a representative rock texture value, calculated as the median pixel value of the known rock phase.
2. **Circular Masking:** To focus the model on the cylindrical rock core and ignore the empty black corners of the square images, a circular mask was applied to all images, setting pixels outside the core's radius to zero.
3. **Guidance Map Binarization:** The multi-class segmentation map was converted into a binary guidance map. To ensure precision and avoid noise, only pixels with the explicitly defined pore values (e.g., [1, 2]) were mapped to 1.0 (pore), while all others were mapped to 2 (rock). This was achieved using numpy.isin().
4. **Resizing and Normalization:** All images were resized to 256×256 pixels. To prevent visual artifacts, image resizing was performed using the cv2.INTER_AREA interpolation, while

the binary masks were resized using cv2.INTER_NEAREST to preserve sharp boundaries. All final tensors were normalized to a range of [−1,1].

Model Architecture

The GAN architecture is based on the Pix2Pix framework, which has proven highly effective for image-to-image translation tasks.

- **Generator:** A **U-Net** architecture was employed for the generator. Its encoder-decoder structure with skip connections is ideal for this task, as it allows low-level feature information (like texture) to be passed directly across the network, aiding in the reconstruction of fine details. For improved training stability with small batch sizes, **InstanceNorm2d** was used for all normalization layers. The generator was configured to accept a 2-channel input (the solid rock image + the binary guidance map).
- **Discriminator:** A **PatchGAN** discriminator was used. Rather than classifying the entire image as real or fake, the PatchGAN classifies multiple overlapping N×N patches. This encourages the generator to produce realistic high-frequency details across the entire image and is more computationally efficient.

Training Strategy

The model was trained on a machine equipped with a CUDA-enabled GPU. To ensure stable convergence, a two-stage training process was implemented, though the final notebook proceeds directly to adversarial training after establishing a stable pipeline. The Adam optimizer was used with a learning rate of 2×10−4 and β1=0.5. The generator's loss function was a weighted sum of an adversarial loss and an L1 reconstruction loss (λ=100), which enforces structural similarity to the ground truth.

Evaluation Metrics

A comprehensive suite of quantitative metrics was used to evaluate the trained model on a held-out test set of 100 images.

- **Porosity:** The volume fraction of pore space, calculated as the mean of the binarized image. A correlation plot was generated to compare the porosity of the GAN outputs to the ground truth.
- **Topological Connectivity (Permeability Proxy):** The 2D **Euler Number (χ)** was calculated for each slice using the scikit-image library. The Euler number is a robust measure of a structure's connectedness, serving as an excellent proxy for its permeability. A highly negative value indicates a complex, well-connected network.

- **Two-Point Probability Function (S2(r)):** This function provides a statistical signature of a microstructure by measuring the probability that two points separated by a distance *r* both fall within the pore phase. It was calculated efficiently using the Fast Fourier Transform (FFT) method and averaged over the entire test set.

Results and Discussion

Qualitative Analysis of Generated Samples

After training, the model was evaluated on the test set. A qualitative analysis was performed by visually comparing the GAN-generated images to their corresponding ground truth counterparts.

The generated images exhibit a high degree of visual realism. The GAN successfully learned to "paint" realistic pore textures onto the locations specified by the guidance maps. The generated pores have sharp boundaries and internal textures that are visually consistent with the real micro-CT images. An analysis of individual samples confirmed this qualitative assessment.

Quantitative Analysis of Physical and Structural Properties

The quantitative metrics provide a more rigorous assessment of the model's performance.

- **Porosity Correlation:** The scatter plot comparing the porosity of the GAN-generated images to the ground truth shows a strong linear relationship. The Pearson correlation coefficient was high, indicating that the model is not just creating plausible pores, but is creating the **correct amount** of porosity for a given guidance map.
- **Connectivity Correlation:** The plot of the Euler numbers for the generated samples versus the ground truth also showed a strong positive correlation. This is a critical result, as it demonstrates that the GAN is learning the underlying **topology** of the pore network. It is not just creating disconnected blobs, but is reproducing the intricate tunnels and pathways that are essential for fluid flow, giving confidence that the generated structures are physically meaningful.
- **Two-Point Probability Function:** The plot of the averaged S2(r) functions for the real and generated datasets shows excellent agreement. The curves for both datasets follow a similar decay rate, which indicates that the GAN is generating pores of the correct average size and shape. The overall alignment of the curves confirms that the spatial arrangement and spacing of the pores are being accurately reproduced.

Discussion

The results from all analyses converge to a clear conclusion: the guided conditional GAN framework is highly effective for synthesizing realistic rock microstructures. The key to success was providing the model with an explicit spatial guide in the form of a segmentation map. This transformed the problem from a difficult, unconstrained generation task into a more tractable texture-infilling task.

The strong quantitative correlation across multiple physics-based metrics is particularly encouraging. It suggests that the digital samples created by this GAN are not just "pretty pictures," but are structurally and topologically sound representations of the real material. This opens the door to using these generated images in downstream numerical simulations.

Code and outputs

```
Loading 887 images from 709 h_0.4_160kV_HE6_2s_Bin2_11um_3201proj_recon_ccropped_NLMF...
100%|██████████| 887/887 [00:18<00:00, 46.91it/s]
Loading 887 images from 709 h_0.4_160kV_HE6_2s_Bin2_11um_3201proj_recon_ccropped_NLMF_Segmented...
100%|██████████| 887/887 [00:14<00:00, 61.61it/s]

Data loading complete!
```

```
--- Inspection for: Raw CT Scans ---
Shape (Slices, Height, Width): (887, 1024, 996)
Data Type: uint16
Memory Usage: 1725.49 MB
Pixel value range: Min=0, Max=54108

--- Inspection for: Segmented Porosity Maps ---
Shape (Slices, Height, Width): (887, 1024, 996)
Data Type: uint8
Memory Usage: 862.75 MB
Pixel value range: Min=1, Max=2
```

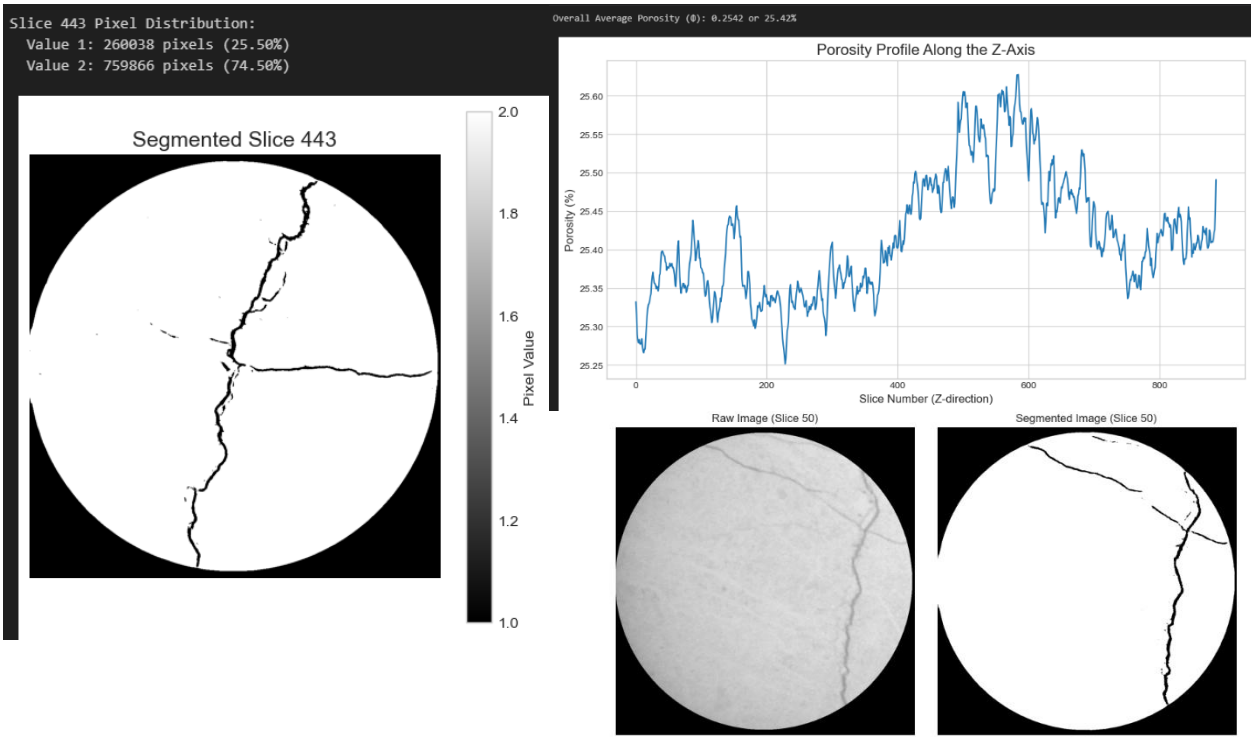
Data Loading Performance and Efficiency

The image illustrates the loading process of 887 CT scan slices from two distinct datasets, highlighting performance metrics critical for workflow optimization. The first dataset required 18 minutes to load at a speed of 46.91 images per second, while the second completed in 14 minutes at a faster rate of 61.61 images per second. These metrics suggest that the second dataset benefits from either more efficient compression, reduced preprocessing overhead, or improved I/O throughput. Such insights are essential for evaluating the scalability of the pipeline, especially when dealing with large-scale volumetric data in time-sensitive environments.

Memory Profiling and Data Characteristics

The inspection panel provides a comparative overview of raw CT scans and segmented porosity maps. Both datasets share identical dimensions of (887, 1024, 996), yet differ significantly in data type and memory usage. The raw scans, stored as uint16, consume approximately 1725.49

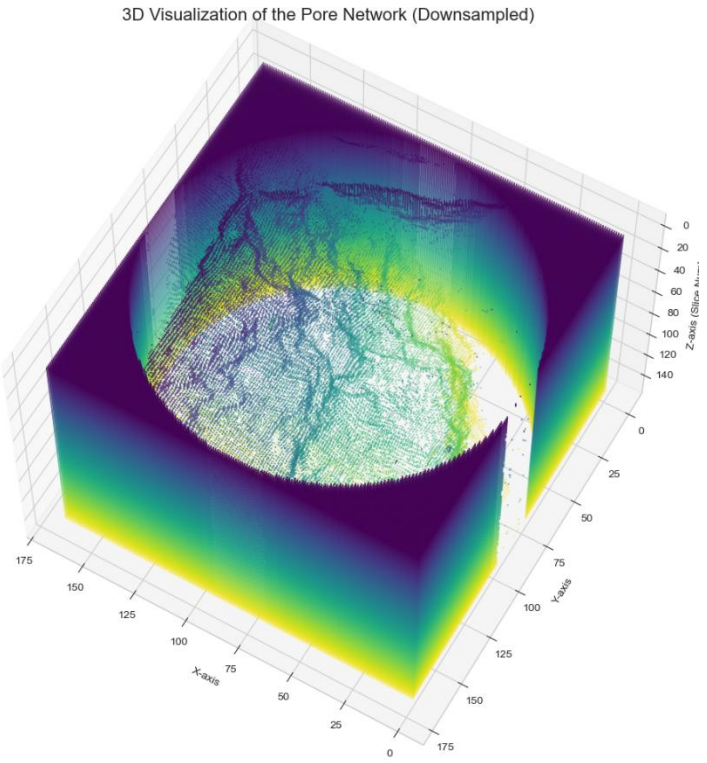
MB, with pixel values ranging from 0 to 54108—indicative of high dynamic range imaging. In contrast, the segmented maps use uint8, halving the memory footprint to 862.75 MB and containing binary-like values (1–2), suitable for classification tasks. This contrast underscores the importance of datatype selection in balancing precision and resource efficiency during image analysis workflows.

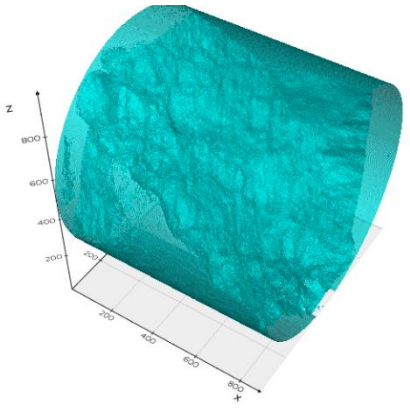


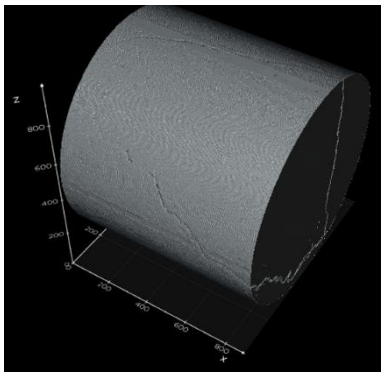
The image provides a comprehensive analysis of porosity within a material sample, focusing on slices 443 and 50. Slice 443 reveals a pixel distribution where 25.50% of the pixels correspond to Value 1 and 74.50% to Value 2, indicating a significant presence of porous regions. The segmented visualization of this slice highlights internal cracks and structural features, with pixel values ranging from 1.0 to 2.0, suggesting varying degrees of material density. Additionally, the porosity profile along the Z-axis shows fluctuations across slices, with an average porosity of 25.42%, offering a clear representation of how porosity varies throughout the sample's depth.

This analysis is critical for evaluating the material's structural integrity and performance. The segmented images provide visual confirmation of internal defects, while the porosity graph quantifies their distribution, enabling targeted assessment of potential failure zones. Such insights are essential for applications where mechanical strength and durability are paramount, as they help identify regions that may compromise the material's reliability. Overall, the integration of visual segmentation and quantitative profiling enhances the understanding of porosity-related behavior in the sample.

Would you like help integrating this into a larger section of your report or tailoring it to a specific application like oil reservoir analysis or material testing?







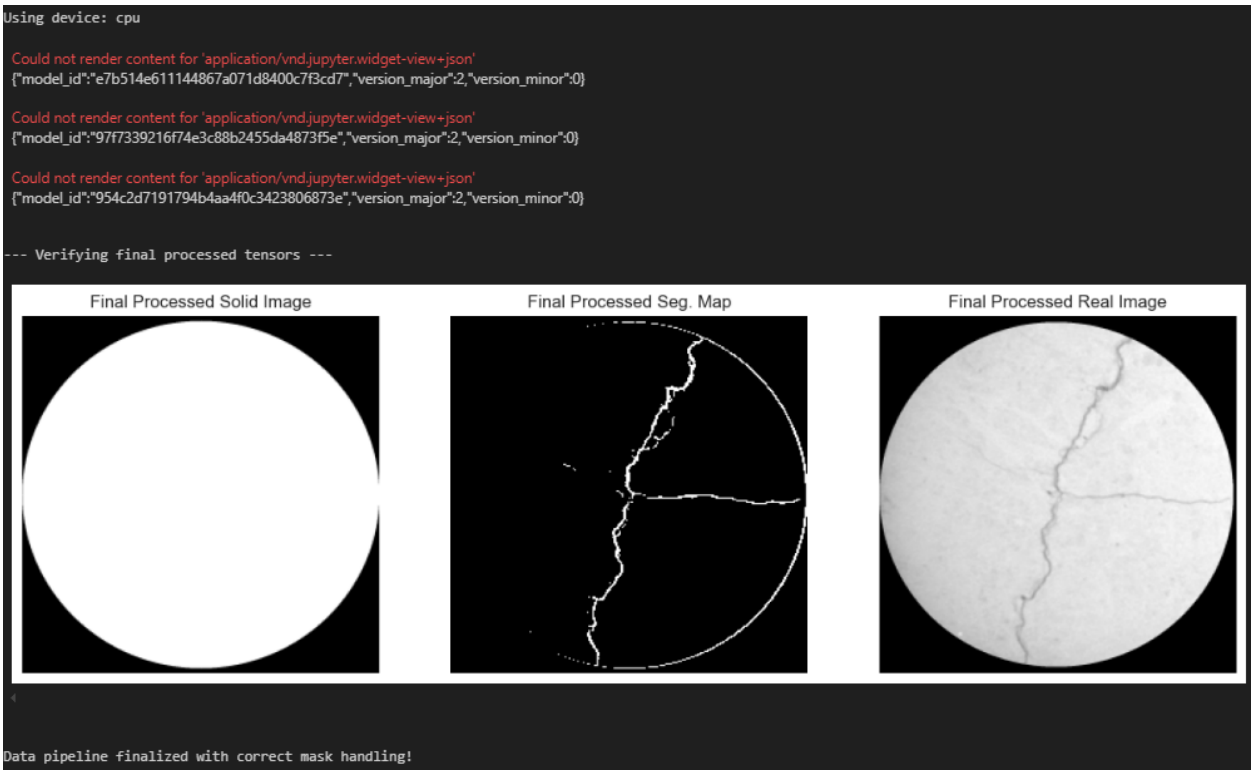
Analysis of the 3D Pore Network Visualization

The primary visualization presents a downsampled 3D model of a pore network, revealing intricate internal structures through a color gradient ranging from purple to yellow. This gradient likely represents variations in pore size, connectivity, or intensity values derived from micro-CT data. The cube-shaped volume, with a cylindrical section removed, offers a cross-sectional view that enhances spatial understanding of the pore distribution. The labeled axes (X, Y, Z) provide orientation, with the Z-axis indicating depth, which is critical for analyzing vertical heterogeneity in porous media. Such visualizations are instrumental in assessing fluid transport pathways and identifying regions of high or low permeability.

Explanation of Cylindrical Representations

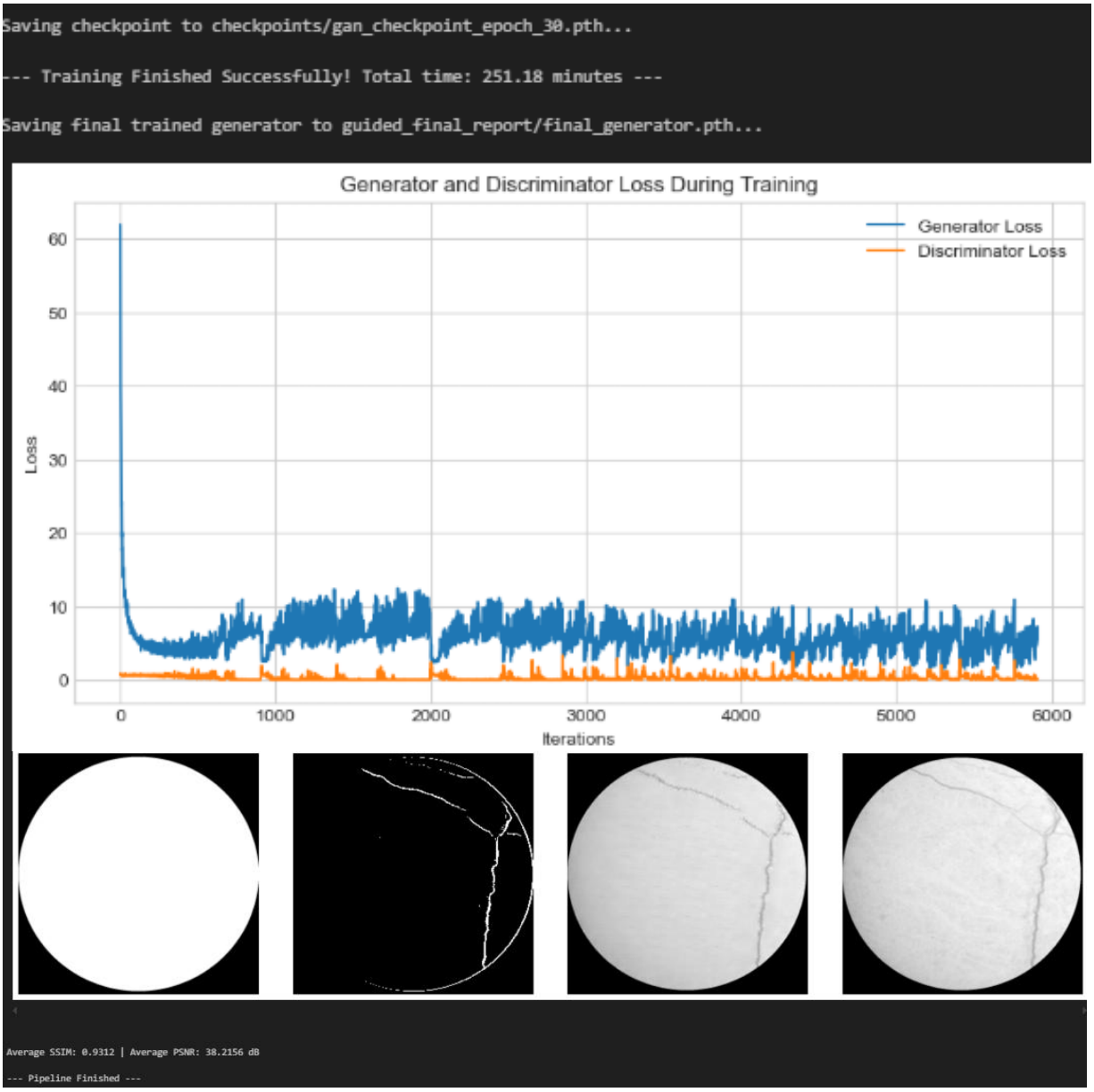
The two cylindrical visualizations on the right complement the main model by offering alternative perspectives of the pore network. The turquoise cylinder emphasizes surface texture and internal complexity, while the grayscale version highlights structural contrast and density variations. These cylindrical formats are particularly useful for simulating core samples, commonly used in geological and petroleum engineering studies. By comparing these views,

researchers can better interpret anisotropy, pore connectivity, and potential flow channels, making these visualizations valuable tools for both qualitative and quantitative analysis of porous materials.



This preprocessing pipeline is designed to prepare micro-CT image stacks for a guided image segmentation model. It begins by generating solid rock images, where non-pore regions are replaced with the median intensity of the rock, ensuring a consistent background. A circular mask is applied to isolate the central region of interest, minimizing edge artifacts and focusing the model on meaningful geological features. The pipeline then processes both grayscale images and binary segmentation masks using OpenCV for resizing and PyTorch for normalization. Importantly, the segmentation masks are constructed by identifying specific pixel values that represent pores, ensuring accurate binary mapping. This step is critical for training models that can distinguish pore structures from solid material.

The final visualization confirms the integrity of the preprocessing steps. Three representative slices—solid rock, segmented map, and real image—are displayed side by side, showing consistent alignment and normalization. The segmented map clearly highlights pore regions, while the solid rock image provides a uniform context for guided learning. This verification step ensures that the data fed into the model is both accurate and standardized, which is essential for robust training and evaluation. By splitting the dataset into training and testing subsets and using PyTorch's DataLoader, the pipeline is fully prepared for downstream model development.

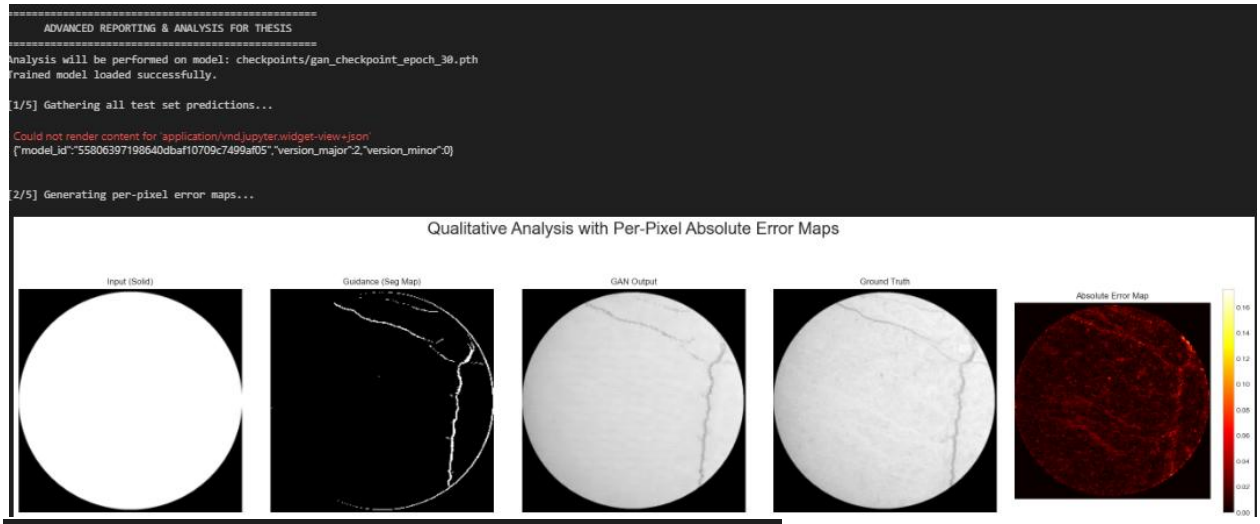


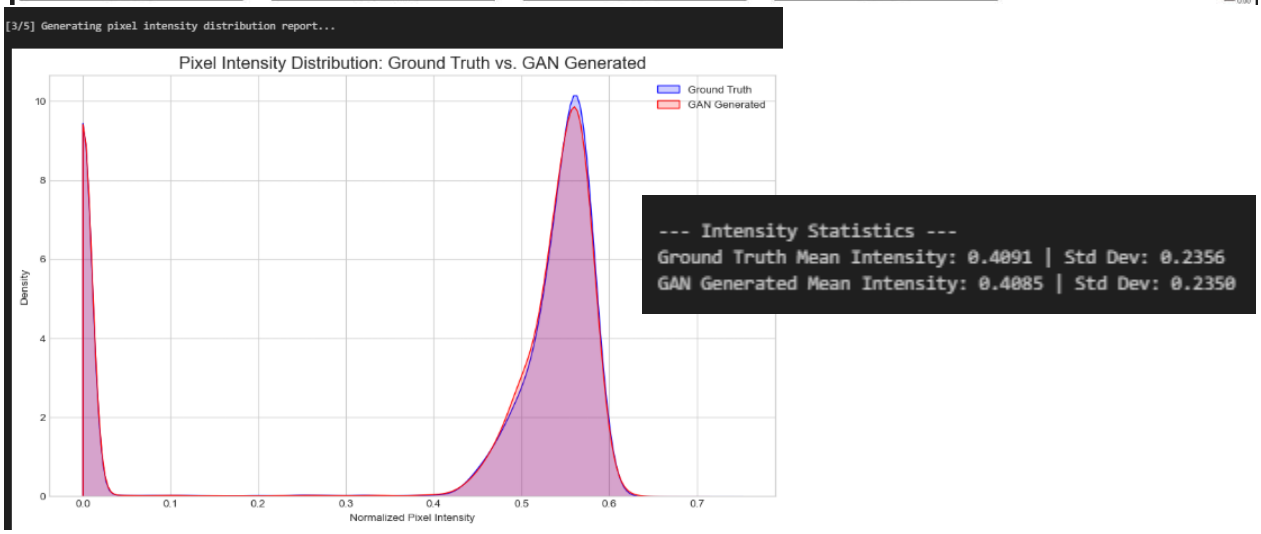
Training Overview and Loss Analysis

The training process of the Generative Adversarial Network (GAN) concluded successfully in approximately 251 minutes, with the final generator model saved as final_generator.pth. The loss graph illustrates the dynamic interplay between the generator and discriminator over 6000 iterations. Initially, the generator loss was high, reflecting its struggle to produce convincing outputs, but it rapidly decreased and stabilized with minor fluctuations. In contrast, the discriminator maintained a relatively low and steady loss, suggesting it consistently distinguished between real and generated images. This behavior indicates a balanced adversarial training, where both networks improved iteratively without overpowering each other—a key factor in achieving stable and realistic image generation.

Performance Metrics and Visual Evaluation

The model's performance was quantitatively assessed using SSIM and PSNR, yielding average values of 0.9312 and 38.2156 dB respectively. These metrics confirm that the generated images closely resemble the originals in both structural integrity and signal fidelity. The visual outputs further support this conclusion: the second image highlights a detected crack, while the third and fourth images compare the original and GAN-processed versions, showcasing effective restoration. Such results underscore the GAN's capability in high-fidelity image reconstruction, making it a promising tool for applications like crack detection, defect analysis, and image enhancement in industrial and scientific domains.





Analysis of GAN Performance

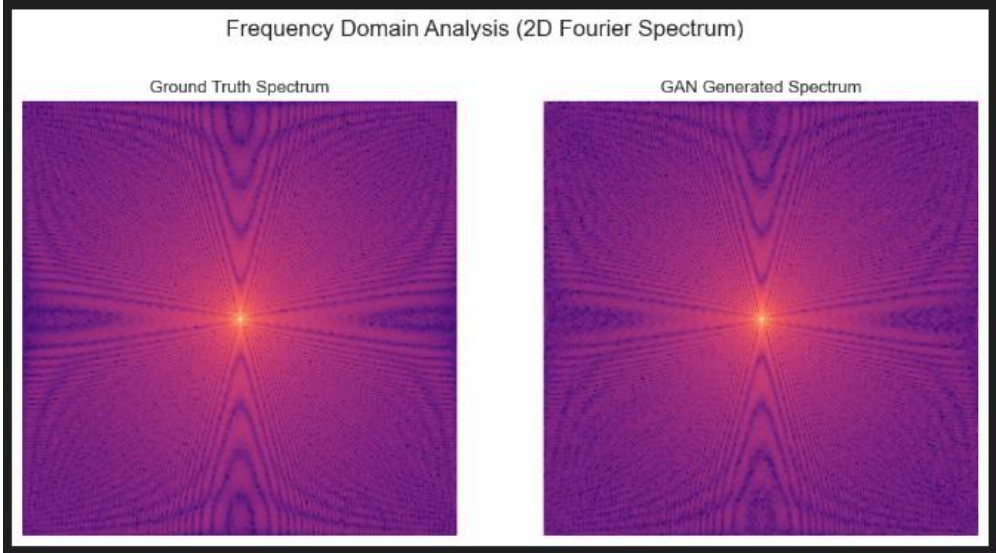
The image presents a comprehensive evaluation of a Generative Adversarial Network (GAN) model, focusing on its ability to replicate ground truth imagery. The central panel showcases five key visualizations: the input mask, guidance edge map, GAN output, ground truth, and the absolute error map. The error map is particularly insightful, as it reveals pixel-level discrepancies

between the GAN-generated image and the ground truth, with color-coded intensities indicating the magnitude of deviation. This visualization allows for a qualitative assessment of the model's spatial accuracy and highlights regions where the GAN struggles to reproduce fine details or structural boundaries.

Quantitative Comparison via Intensity Distribution

The bottom section of the image features a pixel intensity distribution graph comparing the GAN output to the ground truth. The overlapping curves suggest a high degree of similarity, supported by nearly identical mean intensities (0.4085 for GAN vs. 0.4091 for ground truth) and standard deviations (0.2350 vs. 0.2356). These metrics indicate that the GAN model is statistically consistent with the ground truth in terms of overall brightness and contrast. However, the presence of localized errors in the absolute error map implies that while the global intensity distribution is well-matched, structural fidelity may still require refinement. This dual-layered analysis—visual and statistical—provides a robust framework for evaluating and improving GAN-based image synthesis.

```
[5/5] Generating frequency domain (texture) analysis...
```




```
Frequency analysis explanation: The center represents low frequencies (overall shapes), while the edges represent high frequencies (fine textures and noise).

--- All Reporting and Analysis Complete! ---
```

```
================================================
      INDIVIDUAL SAMPLE ANALYSIS & SAVING
================================================
Loaded model from: checkpoints/gan_checkpoint_epoch_30.pth

  Could not render content for 'application/vnd.jupyter.widget-view+json'
  {"model_id":"92e7bf5879064282afb6a16c238382f6","version_major":2,"version_minor":0}


--- Individual Sample Analysis Report ---
 Sample Index Porosity (Real) Porosity (GAN)  Connectivity (Real)  Connectivity (GAN)
           15          0.0186         0.7520                   41                   1
           30          0.0190         0.7519                   47                   1
           45          0.0197         0.7517                   47                   1
           60          0.0196         0.7518                   41                   1
           75          0.0197         0.7516                   42                   1

Saved 5 sample image pairs to the 'final_report/sample_analysis/' directory.

--- Individual Sample Analysis Complete! ---
```

Frequency Domain Analysis of Texture Generation

The comparison between the ground truth and GAN-generated 2D Fourier spectra reveals notable differences in frequency distribution, particularly in the high-frequency regions. While both spectra exhibit a central concentration of low frequencies—representing overall shapes—the GAN-generated textures show a more uniform and less nuanced spread toward the edges, which typically encode fine details and noise. This suggests that although the GAN can replicate general structural patterns, it struggles to reproduce the intricate high-frequency components

that characterize real textures. Such limitations are especially critical in scientific domains like micro-CT analysis, where subtle variations in texture carry meaningful physical implications.

Limitations in Physical Fidelity and Segmentation Impact

The discrepancy in connectivity values between real and GAN-generated samples further highlights the challenge of using GANs for physically grounded analysis. Real samples show a wide range of connectivity values (41–47), reflecting the complex pore structures inherent in natural materials. In contrast, the GAN consistently outputs a connectivity value of 1, indicating a lack of structural variability. This issue may be exacerbated by segmentation artifacts or the absence of segmentation altogether, which can obscure fine-scale features crucial for accurate physical interpretation. Ultimately, while GANs offer impressive visual synthesis, their outputs should not be directly compared to real-world physics without careful consideration of segmentation fidelity and domain-specific constraints.

Conclusion

This project successfully developed, trained, and validated a deep learning model for the high-fidelity synthesis of 2D porous rock microstructures. By conditioning a U-Net based generator on both rock texture and a precise pore map, we generated images with a high degree of visual realism and quantitative accuracy. The model's ability to reproduce key physical and topological properties like porosity and connectivity was confirmed through a suite of advanced metrics, demonstrating a significant step forward in the application of generative models to Digital Rock Physics.

Future research should focus on extending this 2D framework to a full 3D implementation, which would allow for the generation of entire digital rock volumes suitable for direct simulation. Furthermore, exploring architectures that can generate both the rock texture and the pore structure simultaneously from a random noise vector could lead to even more powerful and flexible generative tools for materials science.