AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters

RESEARCH LETTER
10.1029/2025GL117039

Second-Order Degradation Modeling and Multiscale Feature Fusion for High-Fidelity Segmentation of Low-Quality Digital Rock Images

Lulu Xu[1], Zhenxue Dai[1,2], Yi Du[3], Xiaoying Zhang[1], Huichao Yin[4], Mohamad Reza Soltanian[5,6], Hung Vo Thanh[7], Meifeng Cai[8], and Kenneth C. Carroll[4]

[1]State Key Laboratory of Deep Earth Exploration and Imaging, College of Construction Engineering, Jilin University, Changchun, China, [2]School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao, China, [3]School of Mechanical and Electrical Engineering, Beijing Institute of Graphic Communication, Beijing, China, [4]Department of Plant and Environmental Sciences, New Mexico State University, Las Cruces, NM, USA, [5]Department of Geosciences, University of Cincinnati, Cincinnati, OH, USA, [6]Department of Environmental Engineering, University of Cincinnati, Cincinnati, OH, USA, [7]Waseda Research Institute for Science and Engineering, Waseda University, Tokyo, Japan, [8]School of Civil and Resource Engineering, University of Science and Technology Beijing, Beijing, China

Abstract Digital Rock Physics (DRP) is a critical tool for characterizing rock properties and modeling multiphase flow, but segmenting low-quality (LQ) rock images remains a key challenge due to partial volume blurring. In this study, we propose a method that leverages a second-order degradation model to generate physically meaningful synthetic LQ and high-quality (HQ) image pairs for training a full-scale connected UNet 3+, enabling accurate segmentation of LQ rock images with varying degradation levels. It captures the overall pore structure in LQ rock images while recovering fine details from HQ Scanning Electron Microscope data. We validate its effectiveness by benchmarking against the watershed-based segmentation method in terms of porosity, permeability, and pore size distribution. Our method delivers an efficient solution for LQ rock image segmentation, enhancing multiscale pore characterization and petrophysical predictions. This holds significant implications for advancing DRP workflows and deepening the understanding of subsurface rock systems.

Plain Language Summary Pore-scale imaging reveals fine details of rock microstructures, but quantitative analysis relies on precise segmentation, which denotes the process of partitioning images into distinct components such as pores and grains. While deep learning has advanced the field of rock image segmentation, it faces challenges with blurry low-quality (LQ) images, primarily caused by partial volume blurring. To address this, we employ a second-order degradation model to simulate real-world image degradation processes, generating synthetic LQ counterparts from sharp high-quality (HQ) scans. These LQ-HQ pairs are used to train UNet 3+, a deep neural network that fuses multiscale features to mitigate blurring effects and recover fine details. Our approach outperforms previous methods by generating more physically realistic training data and leveraging the superior architectural design of UNet 3+, enabling accurate segmentation—an essential step for reliably predicting rock properties and fluid flow in porous rock materials.

Key Points:

- A second-order degradation model simulates real-world image degradation, generating synthetic low-quality digital rock images for model training
- The full-scale connected UNet 3+ mitigates partial volume blurring, enhancing low-quality rock image segmentation via multiscale feature fusion
- Combining the second-order degradation model with UNet 3+ enables high-fidelity segmentation of low-quality digital rock images

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Z. Dai and X. Zhang,
dzx@jlu.edu.cn;
xiaoyingzh@jlu.edu.cn

Citation:

Xu, L., Dai, Z., Du, Y., Zhang, X., Yin, H., Soltanian, M. R., et al. (2025). Second-order degradation modeling and multiscale feature fusion for high-fidelity segmentation of low-quality digital rock images. *Geophysical Research Letters*, *52*, e2025GL117039. https://doi.org/10.1029/2025GL117039

Received 15 MAY 2025
Accepted 24 SEP 2025

Author Contributions:

Conceptualization: Lulu Xu
Formal analysis: Xiaoying Zhang
Methodology: Lulu Xu
Resources: Meifeng Cai
Software: Lulu Xu, Yi Du
Supervision: Meifeng Cai
Validation: Lulu Xu, Yi Du
Writing – original draft: Lulu Xu
Writing – review & editing: Lulu Xu, Xiaoying Zhang, Huichao Yin, Mohamad Reza Soltanian, Hung Vo Thanh, Kenneth C. Carroll

1. Introduction

Digital Rock Physics (DRP) serves as a key tool for characterizing complex physicochemical processes in subsurface systems (Blunt et al., 2013; Chen et al., 2022; K. Liu et al., 2025; Viswanathan et al., 2022). Over the past two decades, advances in high-quality (HQ) imaging and pore-scale modeling have significantly enhanced DRP's analytical capabilities (Jung et al., 2015; Niu, Wang, et al., 2020; You et al., 2021). These developments facilitate quantitative analysis of fluid transport and reactive processes (Iyare et al., 2025; Zhan et al., 2021, 2023), while the direct derivation of petrophysical parameters from Scanning Electron Microscopy (SEM) or X-ray micro-Computed Tomography (μCT) images highlights DRP's effectiveness in high-precision subsurface characterization (Carroll et al., 2015; C. Li et al., 2025; McDonald et al., 2016; Raeini et al., 2014). Image segmentation constitutes a critical step in the DRP workflow, laying the foundation for subsequent analyses including petrophysical parameter estimation and simulations of single- and multi-phase flow, as well as solute transport (Andra et al., 2013; S. Jia et al., 2023; Niu, Mostaghimi, et al., 2020). Traditional segmentation methods, including Otsu thresholding (Otsu, 1979), watershed-based methods (Beucher & Lantuéjoul, 1979), global and adaptive thresholding (Andra et al., 2013), and indicator kriging (Karimpouli & Tahmasebi, 2019), suffer from

three fundamental limitations: highly operator-dependent parameter tuning (Alqahtani et al., 2022; Cao et al., 2022; Qin et al., 2025); a labor-intensive segmentation threshold optimization process; and limited efficacy in segmenting low-contrast rock images (H. Wang et al., 2022).

In recent years, Deep Learning (DL) methods, particularly Convolutional Neural Networks (CNNs), have emerged as promising tools for rapid, robust segmentation of digital rock images (Purswani et al., 2020; Yu et al., 2020). Once trained, these models enable automated processing of large volumes of rock imagery, delivering consistent segmentation results for flow simulations and petrophysical parameter estimation without user subjectivity (Cherubini et al., 2019; Z. Huang et al., 2022; Z. Wang et al., 2025). Beyond CNNs, Transformer-based architectures have recently been applied to rock image segmentation, utilizing global context modeling to characterize complex pore structures (Y. Jia et al., 2024; M. Li et al., 2025; Xiong et al., 2023). However, a critical bottleneck for both approaches in DRP lies in the scarcity of large, accurately segmented ground truth data sets for model training (Y. Jia et al., 2024; Karimpouli & Tahmasebi, 2019), limiting their accuracy and practical utility, particularly for Transformer architectures, which rely on extensive data to fully leverage their architectural advantages (M. Li et al., 2025). To address this data scarcity, techniques have emerged, including mapping LQ to HQ rock images (Niu, Mostaghimi, et al., 2020; Varfolomeev et al., 2019), data augmentation (Karimpouli & Tahmasebi, 2019), and synthetic data generation (H. Wang et al., 2022). For instance, Karimpouli and Tahmasebi (2019) augmented 20 segmented Berea sandstone images using cross-correlation, then trained SegNet to achieve 96% test accuracy. H. Wang et al. (2022) generated HQ data via entropy-based masking Indicator Kriging (IK-EBM), then trained U-Net++ to mitigate partial volume blurring (PVB) in noisy LQ rock images. Training CNNs on HQ imaging data enhances their ability to segment LQ rock images: Varfolomeev et al. (2019) trained three CNNs on LQ-to-HQ data sets, finding U-Net superior to other architectures, though still challenged by fine details. Niu, Mostaghimi, et al. (2020) extended CNN training to SEM rock images, which in turn improved μCT image segmentation accuracy and consistency for petrophysical parameter estimation. Alqahtani et al. (2022) proposed a super-resolution segmentation framework for carbonates but faced challenges in accurately segmenting micropores due to PVB, resulting in reduced accuracy in specific regions.

CNNs drive significant advances in digital rock segmentation by eliminating user-induced bias and attaining superior accuracy over conventional methods when trained on sufficient HQ data. Despite this potential, their practical application remains limited by prevalent LQ rock images, characterized by blurring, noise, and artifacts stemming from imaging or data transmission processes. These degradations obscure fine-scale features (e.g., micropores) and impair the identification of phase boundaries (H. Wang et al., 2022). Current LQ segmentation workflows typically depend on expert-registered LQ-HQ training pairs to support CNN-based degradation recovery, but this approach faces inherent limitations: Rigid registration requirements constrain the quantity and diversity of paired data sets, leading to poor generalization to real-world LQ rock images with unseen degradation patterns (M. Liu & Mukerji, 2022). Consequently, two core barriers persist: (a) Insufficient quantity and diversity of training pairs restrict model robustness to out-of-distribution LQ rock images; (b) Inherent degradations in LQ rock images induce PVB, which impairs the detection of fine-scale features and the identification of phase boundaries.

To overcome these limitations, we propose an innovative DL-driven framework for high-fidelity segmentation of LQ digital rock images. Our framework integrates two synergistic core components: (a) A second-order degradation model that synthesizes physically meaningful, diverse LQ-HQ rock image pairs by emulating real-world cumulative degradation processes; (b) A full-scale connected UNet 3+ architecture trained on these synthetic pairs to extract and fuse multiscale pore structure features. By applying the degradation model to high-resolution SEM images, we generate extensive LQ-HQ pairs that cover a broad range of degradation variations. This dual strategy significantly boosts the training efficacy of UNet 3+, enabling high-fidelity segmentation of real-world degraded images while recovering fine-scale structural details from HQ references.

2. Methodology

This study leverages CNNs for segmenting real-world LQ digital rock images. The sandstone sample, sourced from the UK North Sea oil reservoir, exhibits secondary porosity from grain dissolution, complex pore geometries, and multi-modal pore size distributions (Scott et al., 2019). A high-resolution SEM image of 16,384 × 12,288 pixels with a resolution of 0.17 μm was downsampled to 8,192 × 6,144 pixels with a resolution of 0.34 μm via linear interpolation for CNN training, thereby enhancing computational efficiency. As illustrated

10.1029/2025GL117039

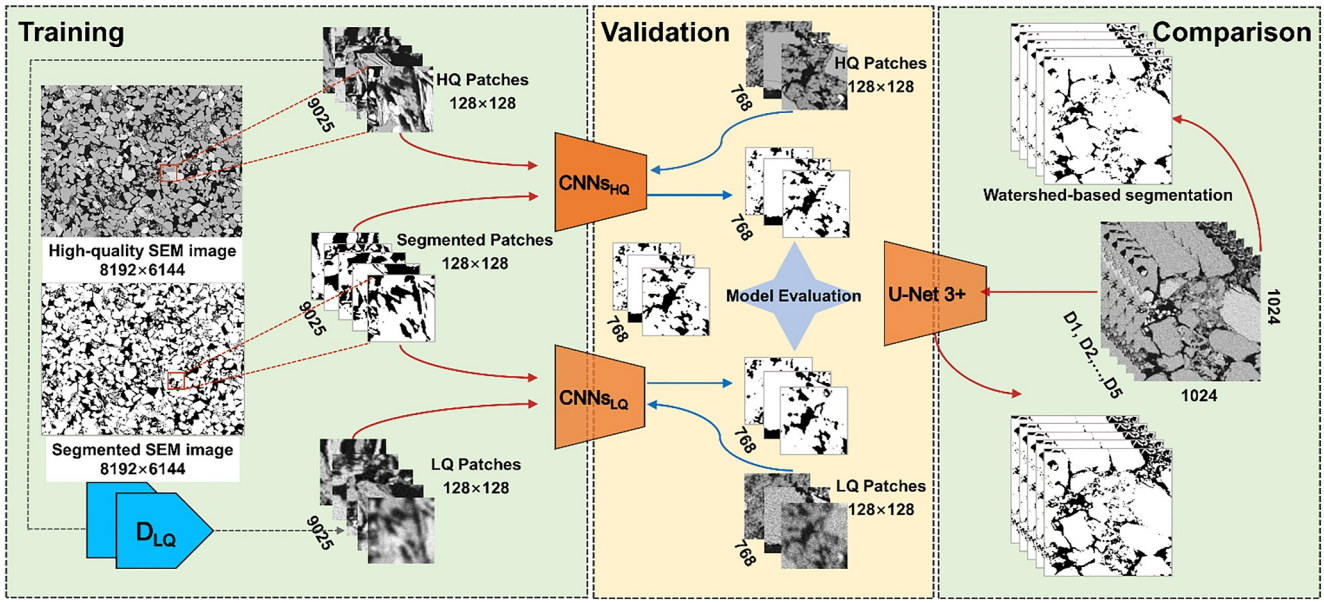

Figure 1. The framework for Convolutional Neural Network (CNN)-based segmentation of low-quality (LQ) digital rock images. The second-order degradation model (D_{LQ}) takes high-quality (HQ) Scanning Electron Microscope rock images as input to synthesize LQ rock images, while $CNNs_{HQ}$ and $CNNs_{LQ}$ refer to four distinct CNN architectures trained on HQ and LQ rock images, respectively.

in Figure 1, the framework integrates two synergistic components: (a) A physics-informed second-order degradation model that synthesizes realistic LQ rock images from HQ SEM inputs; (b) a full-scale connected UNet 3+ architecture that conducts hierarchical feature fusion to mitigate PVB effects. To maximize data diversity, the SEM image was partitioned into 128 × 128-pixel patches using a 64-pixel sliding stride, with each patch degraded individually, as detailed in Section 2.3. Validation involved five degradation levels (D1–D5, from mild to severe) on 1,024 × 1,024-pixel regions, where segmentation performance was quantified by evaluating petrophysical consistency against the watershed-based segmentation benchmark (Leu et al., 2014).

2.1. Second-Order Degradation Model

A critical challenge in DL-based segmentation of LQ digital rock images lies in the scarcity of training pairs that faithfully capture real-world degradation patterns. To address this, we propose a physics-informed second-order degradation model inspired by X. Wang et al. (2021)—one that synthesizes realistic LQ rock images by applying two sequential degradation processes to HQ digital rock SEM images. Unlike generative adversarial networks (GANs), this model directly mimics multi-stage physical degradation, avoids GAN-related training instability, and eliminates the need for large-scale annotated training data. Each stage follows an enhanced first-order degradation framework with distinctly tailored hyperparameters, integrating blur, additive noise, and JPEG compression. The base degradation process is defined as follows:

$$\mathrm{LQ} = \mathcal{D}(\mathrm{HQ}) = [(\mathrm{HQ} \circledast k) + n]_{\mathrm{JPEG}} \tag{1}$$

where $\mathcal{D}$ denotes the degradation process, $\circledast$ represents the convolution operation with blur kernel k, and n is additive noise. Detailed descriptions can be found in Supporting Information S1.

Real-world digital rock images undergo cumulative sequential degradation during acquisition and processing: scanning or internet downloads introduce device blur, sensor noise, sharpening artifacts, or JPEG compression; μCT imaging typically sacrifices resolution for a broader field of view, unlike high-resolution SEM imaging, which yields images with clear phase contrasts; real-time imaging prioritizes speed over quality; and post-imaging edits or data transmission introduce overshoot artifacts, unpredictable noise, or additional blur. Consequently, classical first-order degradation models cannot fully capture this multi-stage complexity, necessitating a higher-order degradation process for realistic simulation. To balance simplicity and effectiveness, we adopt a second-order degradation model consisting of two sequential steps (Figure S1 in Supporting Information S1), expressed as:

$$\mathrm{LQ} = \mathcal{D}^2(\mathrm{HQ}) = (\mathcal{D}_2 \, o \, \mathcal{D}_1)(\mathrm{HQ}) \quad (2)$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ denote the first and second sequential degradation steps, respectively. This structure better approximates real-world cumulative degradation than first-order models while remaining computationally feasible.

2.2. Full-Scale Connected UNet 3+ Architecture

The choice of CNN architecture critically impacts segmentation performance for LQ rock images affected by PVB. To identify the optimal architecture for this task, we evaluated four CNNs: fully convolutional network, UNet++, UNet 3+, and DeepLabV3+ architectures. Among these, UNet 3+ is a deeply supervised architecture designed to learn the mapping between LQ rock images and their HQ segmentation masks (H. Huang et al., 2020), enabling two-phase (pore and grain) segmentation. UNet 3+ incorporates two key improvements to enhance multiscale feature fusion (Figure S2 in Supporting Information S1): (a) Full-scale skip connections between the encoder and decoder, as well as across decoder sub-networks, which capture pore information across scales and refine the positions and boundaries of pores; (b) full-scale deep supervision, which effectively supports the learning of hierarchical representations from aggregated feature maps. Both mitigate PVB by preserving fine-grained structural details critical for pore-grain differentiation. Detailed descriptions of the other three architectures are provided in the Supporting Information. To further refine pore-boundary segmentation, we deploy a hybrid loss function integrating focal loss ($\mathcal{L}_{\mathrm{fl}}$), MS-SSIM loss ($\mathcal{L}_{\mathrm{ms\text{-}ssim}}$), and soft IoU loss ($\mathcal{L}_{\mathrm{soft\text{-}iou}}$):

$$\mathcal{L}_{\mathrm{seg}} = \mathcal{L}_{\mathrm{fl}} + \mathcal{L}_{\mathrm{ms\text{-}ssim}} + \mathcal{L}_{\mathrm{soft\text{-}iou}} \quad (3)$$

where $\mathcal{L}_{\mathrm{fl}}$ targets blurred pore-grain interfaces, $\mathcal{L}_{\mathrm{ms\text{-}ssim}}$ preserves multiscale structural integrity, and $\mathcal{L}_{\mathrm{soft\text{-}iou}}$ optimizes regional phase overlap, collectively alleviating PVB-induced segmentation errors.

2.3. Data Preparation and Training

The target HQ SEM image (8,192 × 6,144 pixels) was partitioned into two distinct regions: a left 6,144 × 6,144-pixel section designated for model training, and a right 2,048 × 6,144-pixel section reserved for model evaluation. The training region was cropped into 9,025 overlapping 128 × 128-pixel patches using a 64-pixel sliding stride, while the evaluation region was divided into 768 non-overlapping 128 × 128-pixel patches for consistency. Ground truth masks were generated by binarizing the original HQ SEM image, followed by identical patch cropping to align with the training and evaluation data sets.

Prior to training, all HQ patches were processed by our physics-informed parameterized degradation model to generate paired LQ data. Critically, two rounds of random degradation were applied to each patch to maximize diversity, boosting model generalizability for addressing the PVB problem. This model runs efficiently on an NVIDIA GeForce RTX 2060 (6 GB memory) thanks to its lightweight parameterized architecture, enabling rapid LQ data synthesis even on personal workstations. The paired HQ-LQ data sets were utilized to train the UNet 3+ model, which completed training in approximately 5 hr on a single NVIDIA V100 GPU. This efficiency originates from UNet 3+'s lightweight architecture tailored for multiscale feature fusion. After training, the model demonstrated precise segmentation capability on LQ rock images, accurately identifying submicron pores (<1 μm) that are imperceptible to the naked eye. The combined application of diverse synthetic degradations and the UNet 3+ network's efficient design ensured both training efficiency and robust performance on real-world degraded images.

3. Results

Figures 2a and 2b contrast HQ SEM-derived rock images against computationally degraded LQ counterparts generated via our second-order model. The original HQ rock images clearly resolve multiscale pore structures, including intergranular pores, soluble mineral pores, micropores, and microfractures, with sharp boundaries. Post-degradation, these sharp boundaries become blurred and noisy, with noticeable ringing and overshoot artifacts emerging at grain-pore edges. These artifacts originate from the *sinc* filter's spectral clipping of high-frequency edge information during phase transitions, mirroring distortions in DRP workflows. Notably, LQ rock images entirely lose detectable micropore signatures. More critically, the PVB effect poses significant

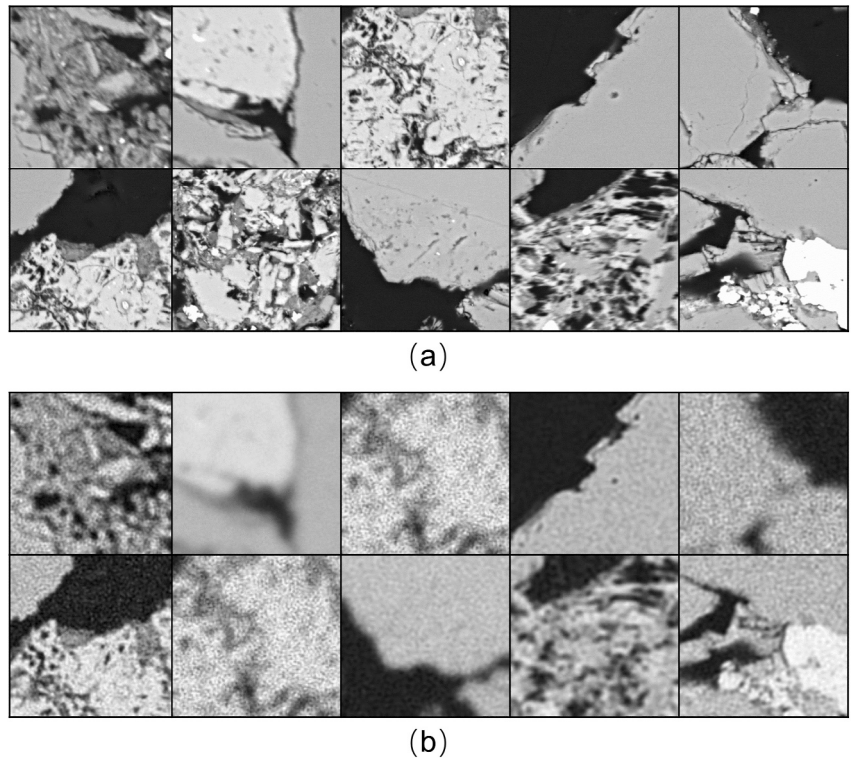

Figure 2. Comparison of (a) high-quality Scanning Electron Microscope rock images and (b) synthetic low-quality rock images generated by the second-order degradation model with distinct random processes. The image size is 128 × 128 pixels, and the resolution is 0.34 μm.

challenges to segmentation, as it causes blurred edges to resemble low-resolution imaging features. Additionally, degradation severity varies across LQ rock images, which is evident in inconsistent edge blurring and variable degrees of micropore loss, confirming the inherent randomness of real-world degradation processes. This diversity in LQ samples ultimately enhances the robustness of segmentation CNNs.

To assess the quality of synthetic LQ data, we quantify thresholding sensitivity in both HQ and degraded LQ rock images by analyzing porosity, permeability, and pore size distribution. Figure 3a demonstrates that a 10-unit grayscale increase causes greater porosity variation in LQ rock images than in HQ rock images. This confirms the higher threshold sensitivity of degraded LQ rock images, while the relatively uniform variation distribution highlights the diversity of data generated by the degradation model. Figure 3b presents permeability-threshold relationships for degraded LQ rock images and real HQ SEM images, where permeability is computed based on extracted pore structures using the Poiseuille-Darcy coupled model. The permeability of HQ data exhibits minimal fluctuation (±3% variation) over a ±10-unit grayscale deviation range. In contrast, degraded LQ data shows substantial permeability variation: a −5-unit grayscale shift results in a ∼40% change in permeability—aligning with the finding of Niu, Mostaghimi, et al. (2020) that HQ SEM-based permeability estimates are significantly less threshold-sensitive than μCT-derived permeability estimates. Pore size distribution was also analyzed at grayscale thresholds of 50 and 70 (Figure 3c). Degraded LQ rock images exhibit a higher number of apparent pores and larger variations in pore count compared to HQ SEM images. This originates from pore-grain boundary blurring in LQ rock images, which leads to partial false-positive labeling where grains are misidentified as pores. Thus, the generated LQ data exhibits higher sensitivity to thresholds, as small threshold changes trigger significant changes in segmentation results. Moreover, this synthetic LQ data provides robust diversity for CNN training, thereby improving the generalizability of the trained CNN model to unseen real-world LQ data.

Figure S3 in Supporting Information S1 shows that UNet 3+ achieves the best segmentation performance for LQ rock images. This superiority arises from its full-scale skip connections, which enable seamless fusion of low-level spatial details and high-level semantic features across scales, ultimately facilitating full capture of pore information. Consistent with this, H. Huang et al. (2020) reported that UNet 3+ can effectively mitigate PVB

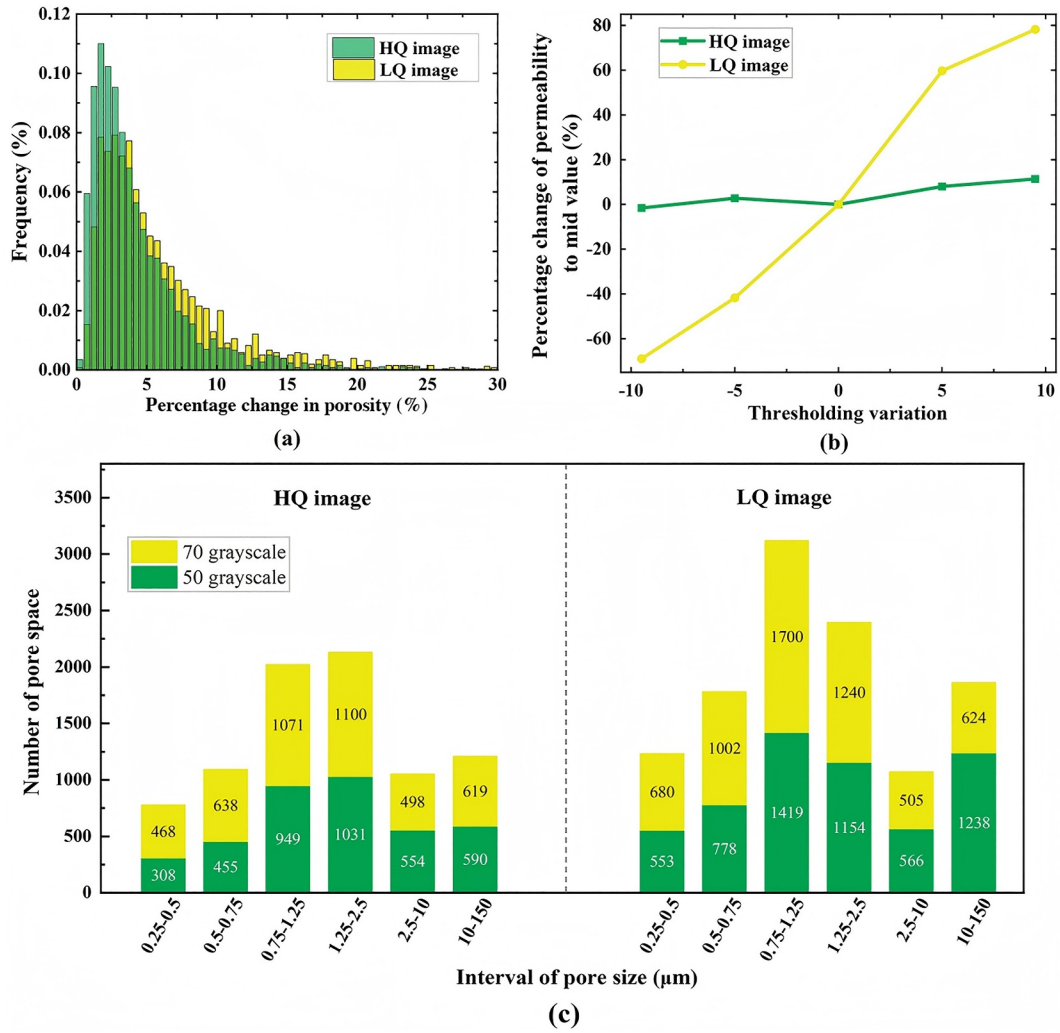

Figure 3. Thresholding sensitivity analysis of high-quality Scanning Electron Microscope (SEM) and degraded low-quality rock images: (a) porosity; (b) permeability; (c) pore size distribution. A total of 2,961 256 × 256-pixel patches were used for porosity analysis. For permeability and pore size distribution, one 4,000 × 4,000-pixel image, selected from the bottom-right corner of the SEM image, was employed.

while maintaining high computational efficiency. We further evaluated the trained UNet 3+ against the watershed method on degraded LQ rock images, with segmentation conducted patch-by-patch. Visually, Figure 4a shows both methods capture general pore structure features; however, local magnification reveals UNet 3+ better detects small pores and recovers finer details in the LQ rock image, consistent with the pore features of the HQ SEM image in Figure 4a. This is further validated by the pore size distribution in Figure 4b, where UNet 3+'s results align more closely with the HQ SEM-derived ground truth. Training UNet 3+ on our synthetic degraded data narrows the segmentation gap between LQ and HQ rock images, with particular effectiveness in recovering fine pore structures. To verify applicability, we segmented five 1,024 × 1,024 LQ rock images with varying degradation levels (D1 to D5, from mild to severe). As shown in Figure 4c, UNet 3+ exhibits minimal errors for porosity (−5.0% to −2.1%) and permeability (−8.8%–0.6%), whereas the watershed method yields significantly higher errors (>30% and >45%, respectively). These results underscore the critical role of recovered micropores in the accurate estimation of petrophysical properties, consistent with the findings of Xu et al. (2025), whose work experimentally demonstrated that micropores enhance pore connectivity to facilitate fluid transport, thereby improving the permeability of porous media. Overall, training UNet 3+ with the degradation model enhances the accuracy and stability of LQ rock image segmentation.

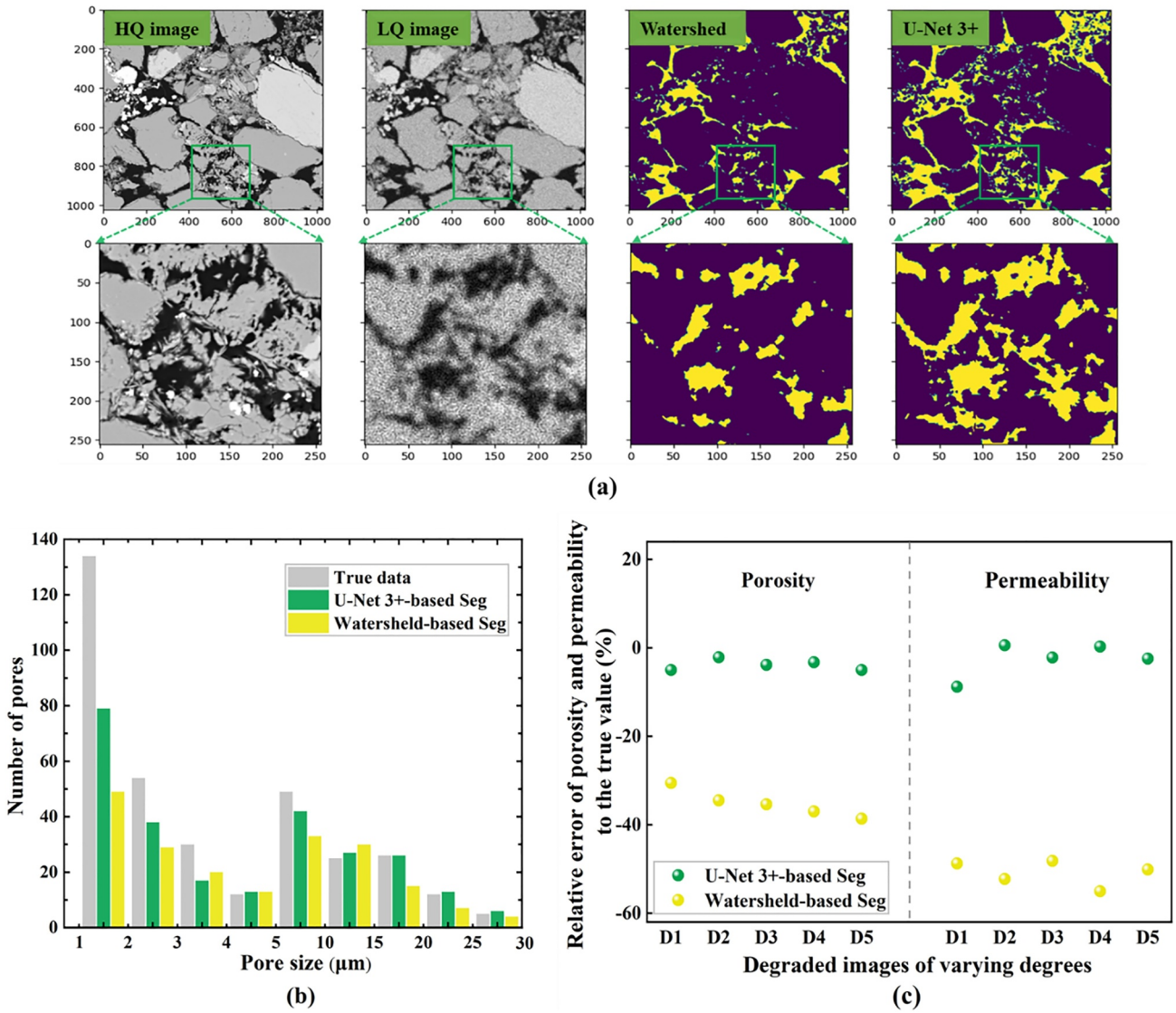

Figure 4. Comparison of segmentation performance between the watershed-based method and the UNet 3+ model: (a) segmented rock images; (b) pore size distribution; (c) porosity and permeability. The rock image size is 1,024 × 1,024-pixels, cropped from the evaluation region of the original Scanning Electron Microscope image.

4. Discussion

Recent advances in DL methods have significantly improved the segmentation of digital rock images, particularly by utilizing high-fidelity SEM imaging data to enhance the segmentation of LQ rock images. However, previous studies have primarily relied on the registration of LQ and HQ rock image data sets to achieve high-accuracy segmentation robust to threshold variations and mitigate the impact of user subjectivity. While effective in specific scenarios, this approach faces challenges with real-world LQ rock images due to the scarcity of sufficient paired training data. We address this challenge by developing a framework that generates training pairs for CNNs, specifically by simulating the degradation of HQ rock images while fully leveraging the high-detail characteristics of SEM imaging data. This framework is highly suitable for applications requiring wide-field HQ SEM imaging modes, especially for characterizing multiscale structures in diverse rock types, including tight sandstones (with microscale pores), carbonates (with vugular networks), and shales (with microfracture systems). These multiscale features are critical to fluid flow, as they directly regulate the transport behavior of fluids in porous media. Notably, capturing these varied scales requires a broad range of imaging resolutions, and the compatibility of the proposed framework with wide-field HQ SEM effectively meets this requirement.

The proposed method leverages the deep feature priors of CNNs trained on HQ SEM imaging data to reconstruct structural details that exceed what is achievable with LQ rock data. Given that μCT systems inherently operate at coarser resolutions than SEM's submicron imaging capability (Y. Wang et al., 2024), this physical constraint limits the method's effectiveness for μCT data sets. Currently limited to 2D images, the method cannot capture 3D pore connectivity, which is a critical factor governing fluid transport behaviors. Additionally, the degradation model relies on predefined parameters and fails to encompass all real-world imaging artifacts. Nevertheless, even modest improvements in segmentation accuracy and stability are critical for 3D DRP analysis, as micropores significantly impact petrophysical properties in 3D space. Notably, integrating Cycle-Consistent GANs (CycleGANs) with the degradation model enables accurate and robust μCT data segmentation on par with the segmentation results of SEM images (M. Liu & Mukerji, 2022). Looking ahead, critical future directions include refining the degradation model to more closely align with real-world imaging artifacts, extending the framework to 3D modalities, and leveraging transfer learning to enhance generalization capabilities (R. Li et al., 2025). These advancements will enhance the method's applicability across diverse imaging scenarios and rock analysis tasks.

5. Conclusions

We proposed an innovative framework for high-fidelity segmentation of LQ digital rock images, integrating a physics-informed second-order degradation model and a full-scale connected UNet 3+. The second-order degradation model effectively overcomes the bottleneck of paired training data scarcity by simulating real-world cumulative degradation processes of HQ SEM digital rock images, while the UNet 3+ mitigates the PVB effect through full-scale skip connections and ensures computational efficiency due to its optimized, reduced parameter count. Validation on LQ rock images spanning five degradation levels confirms that the proposed framework is a robust, effective approach for high-fidelity LQ rock image segmentation. The framework's segmentation results accurately capture submicron pores in LQ rock images and demonstrate strong consistency with HQ SEM reference results in terms of pore size distribution. Moreover, when benchmarked against the HQ SEM-derived ground truth, the framework's segmentation errors in porosity and permeability are substantially lower than those of the traditional watershed segmentation method. We conclude that the proposed framework offers an efficient, reliable solution for LQ rock image segmentation, which is vital for accurate characterization of multiscale pore structures and reliable petrophysical property estimation in DRP workflows. This solution thereby supports critical subsurface engineering applications, including hydrocarbon recovery and carbon capture and sequestration.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Corresponding SEM data for North Sea Sandstone are publicly available (Scott, 2020). The code for the second-order degradation model is based on the implementation provided by the following repository: https://github.com/xinntao/Real-ESRGAN.

Acknowledgments
This work has been funded by the National Natural Science Foundation of China (NSFC: U2267217, 42141011). A part of this work has been supported by the "Grant for Research" from the Japan Petroleum Institute (Grant BXRC02184001) and Waseda University Research Institute for Science and Engineering under the Encouraged Research Program (Grant BA070Z000500).

References

Alqahtani, N. J., Niu, Y., Wang, Y. D., Chung, T., Lanetc, Z., Zhuravljov, A., et al. (2022). Super-resolved segmentation of X-ray images of carbonate rocks using deep learning. *Transport in Porous Media*, *143*(2), 497–525. https://doi.org/10.1007/s11242-022-01781-9

Andra, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., et al. (2013). Digital rock physics benchmarks-part I: Imaging and segmentation. *Computers & Geosciences*, *50*, 25–32. https://doi.org/10.1016/j.cageo.2012.09.005

Beucher, S., & Lantuéjoul, C. (1979). Use of watersheds in contour detection. In *Paper presented at international workshop on image processing, real-time edge and motion detection*.

Blunt, M. J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., et al. (2013). Pore-scale imaging and modelling. *Advances in Water Resources*, *51*, 197–216. https://doi.org/10.1016/j.advwatres.2012.03.003

Cao, D., Ji, S., Cui, R., & Liu, Q. (2022). Multi-task learning for digital rock segmentation and characteristic parameters computation. *Journal of Petroleum Science and Engineering*, *208*, 109202. https://doi.org/10.1016/j.petrol.2021.109202

Carroll, K. C., McDonald, K., Marble, J., Russo, A. E., & Brusseau, M. L. (2015). The impact of transitions between two-fluid and three-fluid phases on fluid configuration and fluid-fluid interfacial area in porous media. *Water Resources Research*, *51*(9), 7189–7201. https://doi.org/10.1002/2015wr017490

Chen, L., He, A., Zhao, J., Kang, Q., Li, Z.-Y., Carmeliet, J., et al. (2022). Pore-scale modeling of complex transport phenomena in porous media. *Progress in Energy and Combustion Science*, *88*, 100968. https://doi.org/10.1016/j.pecs.2021.100968

Cherubini, A., Garcia, B., Cerepi, A., & Revil, A. (2019). Influence of CO_2 on the electrical conductivity and streaming potential of carbonate rocks. *Journal of Geophysical Research-Solid Earth*, *124*(10), 10056–10073. https://doi.org/10.1029/2018jb017057

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., et al. (2020). UNet 3+: A full-scale connected UNet for medical image segmentation. arXiv:2004.08790.

Huang, Z., Kurotori, T., Pini, R., Benson, S. M., & Zahasky, C. (2022). Three-dimensional permeability inversion using convolutional neural networks and positron emission tomography. *Water Resources Research*, *58*(3), e2021WR031554. https://doi.org/10.1029/2021wr031554

Iyare, U. C., Boampong, L. O., Li, W., Neil, C. W., Frash, L. P., Carey, J. W., et al. (2025). Interaction between dissolution and precipitation during olivine carbonation: Implications for CO_2 mineralization. *Chemical Geology*, *678*(5), 122645. https://doi.org/10.1016/j.chemgeo.2025.122645

Jia, S., Dai, Z., Zhou, Z., Ling, H., Yang, Z., Qi, L., et al. (2023). Upscaling dispersivity for conservative solute transport in naturally fractured media. *Water Research*, *235*, 119844. https://doi.org/10.1016/j.watres.2023.119844

Jia, Y., Wan, G., Li, W., Li, C., Liu, J., Cong, D., & Liu, L. (2024). EDR-TransUnet: Integrating enhanced dual relation-attention with transformer U-Net for multiscale rock segmentation on Mars. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 4601416. https://doi.org/10.1109/tgrs.2024.3432533

Jung, H. B., Carroll, K. C., Kabilan, S., Heldebrant, D. J., Hoyt, D., Zhong, L., et al. (2015). Stimuli-responsive/rheoreversible hydraulic fracturing fluids as a greener alternative to support geothermal and fossil energy production. *Green Chemistry*, *17*(5), 2799–2812. https://doi.org/10.1039/c4gc01917b

Karimpouli, S., & Tahmasebi, P. (2019). Segmentation of digital rock images using deep convolutional autoencoder networks. *Computers & Geosciences*, *126*, 142–150. https://doi.org/10.1016/j.cageo.2019.02.003

Leu, L., Berg, S., Enzmann, F., Armstrong, R. T., & Kersten, M. (2014). Fast X-ray micro-tomography of multiphase flow in Berea sandstone: A sensitivity study on image processing. *Transport in Porous Media*, *105*(2), 451–469. https://doi.org/10.1007/s11242-014-0378-4

Li, C., Liu, Y., Li, L., Wang, Z., & Li, H. (2025). Texture-based segmentation of SEM images of shale rocks and estimation of meso-scale elastic modulus by 2D FEM. *Rock Mechanics and Rock Engineering*, *58*(6), 6475–6491. https://doi.org/10.1007/s00603-025-04442-7

Li, M., Chen, M., Lu, W., Zhao, F., Yan, P., & Liu, J. (2025). RDT-FragNet: A DCN-transformer network for intelligent rock fragment recognition and particle size distribution acquisition. *Computers and Geotechnics*, *177*, 106809. https://doi.org/10.1016/j.compgeo.2024.106809

Li, R., Zhang, Q., Guo, S., Zhu, G., Li, H., Chen, X., & Yan, J. (2025). Intelligent identification and semantic segmentation of deep rock fracture based on deep ensemble learning and transfer learning. *Tunnelling and Underground Space Technology*, *157*, 106317. https://doi.org/10.1016/j.tust.2024.106317

Liu, K., Sun, J., Wu, H., Luo, X., & Sun, F. (2025). Shale sample permeability estimation using fractal parameters computed from TransUnet-based SEM image segmentation. *Computers & Geosciences*, *194*, 105745. https://doi.org/10.1016/j.cageo.2024.105745

Liu, M., & Mukerji, T. (2022). Multiscale fusion of digital rock images based on deep generative adversarial networks. *Geophysical Research Letters*, *49*(9), e2022GL098342. https://doi.org/10.1029/2022gl098342

McDonald, K., Carroll, K. C., & Brusseau, M. L. (2016). Comparison of fluid-fluid interfacial areas measured with X-ray microtomography and interfacial partitioning tracer tests for the same samples. *Water Resources Research*, *52*(7), 5393–5399. https://doi.org/10.1002/2016wr018775

Niu, Y., Mostaghimi, P., Shabaninejad, M., Swietojanski, P., & Armstrong, R. T. (2020). Digital rock segmentation for petrophysical analysis with reduced user bias using convolutional neural networks. *Water Resources Research*, *56*(2), e2019WR026597. https://doi.org/10.1029/2019wr026597

Niu, Y., Wang, Y. D., Mostaghimi, P., Swietojanski, P., & Armstrong, R. T. (2020). An innovative application of generative adversarial networks for physically accurate rock images with an unprecedented field of view. *Geophysical Research Letters*, *47*(23), e2020GL089029. https://doi.org/10.1029/2020gl089029

Otsu, N. (1979). Threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/tsmc.1979.4310076

Purswani, P., Karpyn, Z. T., Enab, K., Xue, Y., & Huang, X. (2020). Evaluation of image segmentation techniques for image-based rock property estimation. *Journal of Petroleum Science and Engineering*, *195*, 107890. https://doi.org/10.1016/j.petrol.2020.107890

Qin, S., Wang, Q., Zeng, Q., Ye, M., Fu, A., & Chen, G. (2025). Automatic recognition of debris rock lithology based on unsupervised semantic segmentation. *Computers & Geosciences*, *196*, 105790. https://doi.org/10.1016/j.cageo.2024.105790

Raeini, A. Q., Blunt, M. J., & Bijeljic, B. (2014). Direct simulations of two-phase flow on micro-CT images of porous media and upscaling of pore-scale forces. *Advances in Water Resources*, *74*, 116–126. https://doi.org/10.1016/j.advwatres.2014.08.012

Scott, G. (2020). North sea sandstone SEM images [Dataset]. *Digital Porous Media*. https://www.doi.org/10.17612/36F2-8Q45

Scott, G., Wu, K., & Zhou, Y. (2019). Multi-scale image-based pore space characterisation and pore network generation: Case Study of a North Sea sandstone reservoir. *Transport in Porous Media*, *129*(3), 855–884. https://doi.org/10.1007/s11242-019-01309-8

Varfolomeev, I., Yakimchuk, I., & Safonov, I. (2019). An application of deep neural networks for segmentation of microtomographic images of rock samples. *Computers*, *8*(4), 8040072. https://doi.org/10.3390/computers8040072

Viswanathan, H. S., Ajo-Franklin, J., Birkholzer, J. T., Carey, J. W., Guglielmi, Y., Hyman, J. D., et al. (2022). From fluid flow to coupled processes in fractured rock: Recent advances and new frontiers. *Reviews of Geophysics*, *60*(1), e2021RG000744. https://doi.org/10.1029/2021rg000744

Wang, H., Dalton, L., Fan, M., Guo, R., McClure, J., Crandall, D., & Chen, C. (2022). Deep-learning-based workflow for boundary and small target segmentation in digital rock images using UNet++ and IK-EBM. *Journal of Petroleum Science and Engineering*, *215*, 110596. https://doi.org/10.1016/j.petrol.2022.110596

Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data. *arXiv: 2107.10833*, 1905–1914. https://doi.org/10.1109/iccvw54120.2021.00217

Wang, Y., Qin, R., Wei, D., Li, X., Wang, P., & Ye, X. (2024). Super-resolution enhancement and segmentation for digital rock based on multi-task joint deep neural network. *Geoenergy Science and Engineering*, *243*, 213385. https://doi.org/10.1016/j.geoen.2024.213385

Wang, Z., Hou, Z., & Cao, D. (2025). Edge-guided segmentation of digital rock images: Integrating a pretrained edge aware path with the main segmentation path. *Computers & Geosciences*, *197*, 105884. https://doi.org/10.1016/j.cageo.2025.105884

Xiong, Y., Xiao, X., Yao, M., Liu, H., Yang, H., & Fu, Y. (2023). MarsFormer: Martian rock semantic segmentation with transformer. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 4600612. https://doi.org/10.1109/tgrs.2023.3302649

Xu, L., Cai, M., Dai, Z., Zheng, Z., Chen, W., Yin, S., et al. (2025). Fractal and NMR based characterizations of multi-scale pore structure alterations in tight sandstones due to $scCO_2$-water-rock interactions. *Chemical Engineering Journal*, *508*, 160898. https://doi.org/10.1016/j.cej.2025.160898

You, N., Li, Y. E., & Cheng, A. (2021). 3D carbonate Digital rock reconstruction using progressive growing GAN. *Journal of Geophysical Research: Solid Earth*, *126*(5), e2021JB021687. https://doi.org/10.1029/2021jb021687

Yu, Q., Xiong, Z., Du, C., Dai, Z., Soltanian, M. R., Soltanian, M., et al. (2020). Identification of rock pore structures and permeabilities using electron microscopy experiments and deep learning interpretations. *Fuel*, *268*, 117416. https://doi.org/10.1016/j.fuel.2020.117416

Zhan, C., Dai, Z., Soltanian, M. R., & Zhang, X. (2021). Stage-Wise stochastic deep learning inversion framework for subsurface sedimentary structure identification. *Geophysical Research Letters*, *49*(1), e2021GL095823. https://doi.org/10.1029/2021gl095823

Zhan, C., Dai, Z., Yang, Z., Zhang, X., Ma, Z., Thanh, H. V., & Soltanian, M. R. (2023). Subsurface sedimentary structure identification using deep learning: A review. *Earth-Science Reviews*, *239*, 104370. https://doi.org/10.1016/j.earscirev.2023.104370

References From the Supporting Information

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with atrous separable convolution for semantic image segmentation. *arXiv:1802.02611*, 833–851. https://doi.org/10.1007/978-3-030-01234-2_49

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *arXiv:1708.02002*.

Long, J., Shelhamer, E., & Darrell, T. (2014). Fully convolutional networks for semantic segmentation. *arXiv:1411.4038*.

Máttyus, G., Luo, W., & Urtasun, R. (2017). DeepRoadMapper: Extracting road topology from aerial images. In *Paper presented at 2017 IEEE International Conference on Computer Vision (ICCV)*.

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net Architecture for medical image segmentation. *arXiv: 1807.10165*, *11045*, 3–11. https://doi.org/10.1007/978-3-030-00889-5_1