Advanced IntelliEngineering

Vol. 1, No. 1, p. 4-20, 2026

Invited review

# Deep learning in digital rock technology: A comprehensive review

Chengqian Liu[1], Liqun Shan[2], Jiuyu Zhao[3], Yuxuan Xia[1], Jianchao Cai[1]*

[1]*State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum, Beijing 102249, P. R. China*

[2]*Department of Information Technology, Georgia Southern University, Statesboro GA 30460, USA*

[3]*Exploration and Development Research Institute, Daqing Oilfield of CNPC, Daqing 163712, P. R. China*

**Keywords:**
Digital rock
deep learning
super-resolution reconstruction
image segmentation
physical property prediction
flow simulation

**Cited as:**
Liu, C., Shan, L., Zhao, J., Xia, Y., Cai, J. Deep learning in digital rock technology: A comprehensive review. Advanced IntelliEngineering, 2026, 1(1): 4-20.
https://doi.org/10.46690/aie.2026.01.02

**Abstract:**
Against the background of digital oil field development, deep learning has driven digital rock technology toward a more robust and data-driven characterization framework. This review comprehensively explores the integration of deep learning in pore structure analysis and micro-scale transport mechanisms. This paper systematically reviews the current status of four key areas, namely super-resolution reconstruction, image segmentation, physical property prediction, and flow simulation. Super-resolution reconstruction and image segmentation improve image quality and sharpen pore boundary identification, providing a solid basis for subsequent microscopic analysis. Physical property prediction further links microscopic structures to macroscopic properties, offering an effective approach for reservoir property evaluation. Flow simulation describes transport behavior in porous media and helps reveal flow patterns under different conditions. Despite significant progress, this review also identifies several challenges, including insufficient consideration of physical degradation during image acquisition, as well as limited availability of high-quality labeled data and three-dimensional spatial constraints. In addition, the generalization of physical property prediction and flow simulation models across scales and lithologies remains weak, and the incorporation of flow physics into neural architectures is still inadequate, which may affect physical consistency. This review also discusses emerging trends such as multi-scale modeling and physics-aware learning. Finally, future research directions are identified to support intelligent reservoir characterization and provide stronger technical support for oil and gas development decisions.

## 1. Introduction

Petroleum and natural gas serve as fundamental energy sources in modern industrial systems and provide essential feedstocks for the chemical industry. Supply capacity for these resources remains inextricably linked to economic performance, industrial development, and daily life[1–4]. Oil and gas consumption in China has risen steadily alongside industrialization and urbanization since the beginning of the 21st century[5]. While significant progress has been made in exploration theory, engineering equipment, and production technology, domestic output still struggles to meet growing demand. Consequently, China maintains a high dependence on imported energy resources[6]. External dependence reached 72.7% for crude oil and approximately 40% for natural gas in 2025. Improving the development efficiency of domestic resources and unlocking the potential of complex reservoirs are therefore critical tasks in the energy sector.

Reservoir development mechanisms dictate that the occurrence, migration, and production of hydrocarbons within rocks are controlled by the geometry, connectivity, and multiscale distribution of internal pore-throat systems[7, 8]. This control is particularly evident in unconventional reservoirs such as

Yandy Scientific Press

*Corresponding author.
*E-mail address*: liucq@student.cup.edu.cn (C. Liu); lshan@georgiasouthern.edu (L. Shan); zhaojiuyu@petrochina.com.cn (J. Zhao); xiayx@cup.edu.cn (Y. Xia); caijc@cup.edu.cn (J. Cai).

Received April 15, 2026; revised May 12, 2026; accepted May 27, 2026; available online May 31, 2026.

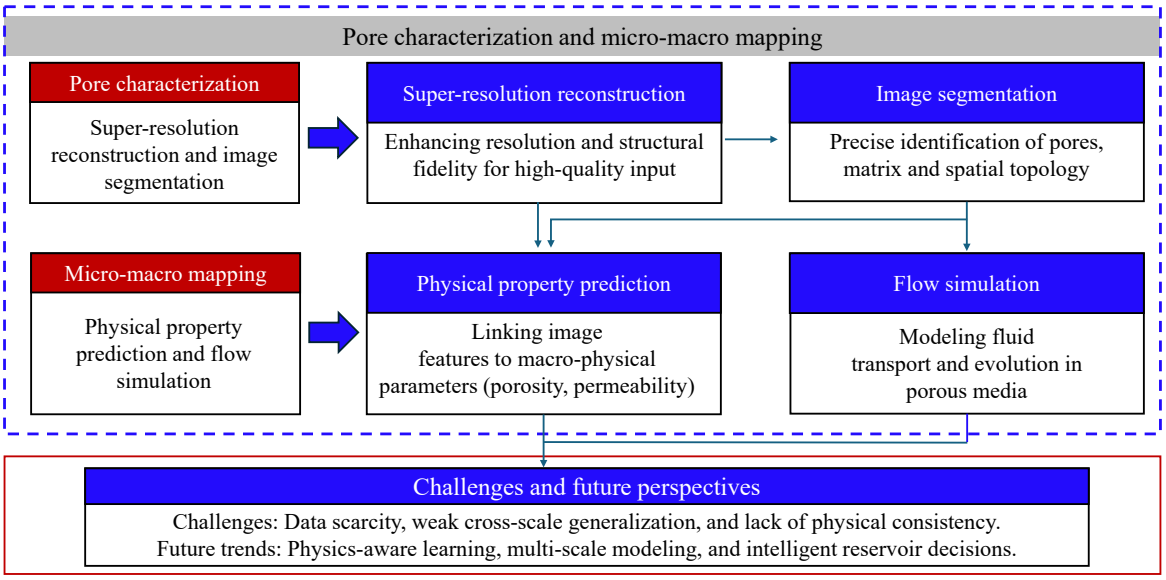


**Fig. 1**. Schematic overview of this review on super-resolution reconstruction, image segmentation, physical property prediction, and flow simulation.

tight rocks and shale. Micro- and nanoscale pores in these formations act as the primary storage spaces and the main channels for fluid flow and phase evolution[9, 10]. Accurate characterization of these structures is essential for understanding pore-scale flow behavior. Linking microscopic structures with macroscopic reservoir properties has become a central scientific challenge for efficient resource development.

Digital rock technology has emerged as a primary method for microscopic reservoir characterization[11, 12]. It uses high-resolution imaging to create three-dimensional (3D) models of pore-throat networks[13]. However, the technology faces several technical bottlenecks. A major issue involves the trade-off between imaging resolution and the field of view. High-resolution imaging captures fine details but covers a limited area. Large-field imaging covers more volume but fails to resolve small features. Additionally, physics-based numerical simulations involve high computational costs for large datasets. High-dimensional data also make manual feature extraction difficult.

Deep learning offers innovative strategies to overcome existing technical bottlenecks[14–17]. These computational methods have gained widespread adoption across the petroleum industry. Current applications include well-log reconstruction, seismic data processing, drilling optimization, and production prediction[18–20]. By processing massive datasets, deep learning architectures effectively extract features and fit complex nonlinear relationships.

The integration of digital rock technology and deep learning is shifting rock image analysis toward intelligent, multiscale, and high-precision workflows[21, 22]. This evolution occurs across three connected levels. At the image level, super-resolution reconstruction improves spatial resolution and preserves fine details under limited imaging conditions. The objective is to recover small pores and narrow throats while maintaining the field of view. At the microstructural level, image segmentation and recognition methods distinguish between pores, matrix, and mineral phases. These methods extract connectivity and topological features for quantitative characterization. Finally, property prediction models and flow simulations link these features to macroscopic parameters like porosity and permeability. This analytical chain extends from raw images to macroscopic behavior, supporting reservoir evaluation and flow-mechanism studies.

Unlike existing reviews on general artificial intelligence or intelligent oil and gas engineering, this study is confined to pore-scale technical workflows. While previous reviews on digital rocks or pore structure analysis emphasize experimental imaging hardware or traditional geometric and topological parameters, this work concentrates on a four-stage deep learning pipeline: Super-resolution reconstruction, image segmentation, physical property prediction, and multi-physics flow simulation. Based on this defined scope, the main contributions of this review are threefold: (1) Providing a structured overview of the four interconnected stages mentioned above; (2) systematically categorizing the challenges identified in the literature, including data limitations, weak cross-scale generalization, and physical inconsistencies; and (3) summarizing future trends, such as multi-scale modeling and physics-aware learning, based on the analyzed frameworks.

To elaborate on these components, the overall research framework of this review is illustrated in Fig. 1. Section 2 introduces the fundamental principles and challenges of digital rock technology. Section 3 focuses on super-resolution reconstruction and the recovery of fine details from low-resolution images. Section 4 reviews image segmentation methods for the identification of complex porous media. Section 5 summarizes deep learning models used for predicting physical properties. Section 6 discusses progress in flow simulation, focusing on pore-scale flow and multiphysics coupling. Section 7 provides a discussion on future research directions and prospects. Sec-

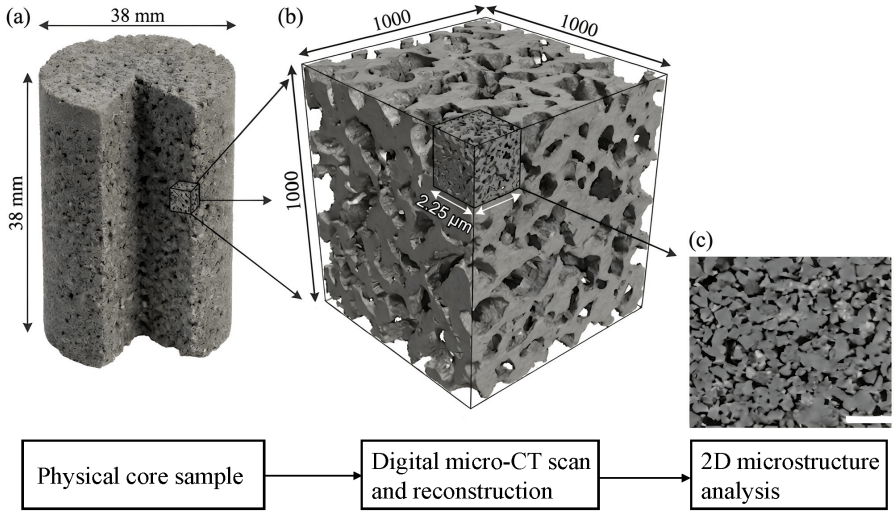


**Fig. 2**. Digital rock core analysis workflow: From sample to microstructure[23]. (a) Experimental-scale core sample and its geometric dimensions, (b) three-dimensional subvolume view of a digital rock cube obtained from X-ray microtomography. The cube contains 1,000 voxels along each edge, with a voxel size of 2.25 μm and an overall side length of 2.25 mm and (c) 2D cross-sectional image, showing the microscopic pore-structure. The white scale bar in the lower-right corner of each image corresponds to 500 μm.

tion 8 concludes the paper with a summary of major findings.

## 2. Digital rock technology

Digital rock technology utilizes high-resolution imaging methods including computed tomography (CT), nuclear magnetic resonance, scanning electron microscopy (SEM), and laser confocal microscopy. These techniques provide 3D grayscale data from rock samples. The resulting data undergo image filtering, threshold segmentation, and volumetric reconstruction to create realistic models of pore-throat networks[13]. During CT imaging, X-rays pass through the core sample while a detector records grayscale variations based on material density. Rotating the sample allows for the collection of multi-angle projection data used to reconstruct the 3D structure. Key parameters such as porosity, permeability, and pore-throat size distribution can then be extracted from these digital models.

Digital rock technology effectively converts complex reservoir microstructures into visual, measurable, and computationally usable digital objects. It provides a unified platform for detailed characterization and fluid behavior simulation. Representative multiscale characterization results obtained via X-ray microtomography illustrate the capabilities of this technology[23], as depicted in Fig. 2. It shows the process of extracting a digital rock volume from a physical core, reconstructing its 3D porous network, and analyzing its microstructure through two-dimensional (2D) cross-sections.

Digital rock technology is widely applied in pore-structure analysis and flow-mechanism studies, particularly for complex unconventional reservoirs[24, 25]. It enables the quantitative description of pore morphology, size, connectivity, and heterogeneity. These insights offer a reliable microscopic basis for reservoir evaluation and reserve estimation. Furthermore, reconstructed digital rocks serve as the basis for numerical simulations of fluid transport in porous media. This capability allows for the simulation of single-phase flow, multiphase flow, and reactive transport within a computational environment. Such simulations support development optimization and enhanced oil recovery. Digital rock technology thus bridges the gap between experimental observation, theoretical analysis, and numerical simulation[26].

Recent research into deeper and more heterogeneous reservoirs has highlighted the need for more efficient processing methods. Classical algorithm-based recognition often struggles with high-dimensional and large-scale data. The computational burden of physics-based models remains a significant barrier for large-scale applications. These challenges necessitate the introduction of deep learning to improve image processing efficiency and structural identification[27–29]. By linking pore structures with macroscopic properties through data-driven models, researchers can overcome the inherent limitations of traditional imaging, such as the persistent trade-off between spatial resolution and field of view, susceptibility to imaging artifacts, and the difficulty of accurately segmenting complex multiscale heterogeneities. This integration facilitates the transition from raw image data to actionable reservoir insights.

## 3. Super-resolution reconstruction

Digital rock technology utilizes micron- or submicron-scale imaging to convert complex pore spaces into computable 3D data volumes. These volumes serve as an essential microscopic foundation for pore-network extraction, flow simulation, and reservoir property prediction[30, 31]. Image resolution remains a critical bottleneck in this workflow. When the voxel size exceeds the pore-throat diameter, problems such as blurred pore boundaries, missing micropores, and distorted connectivity become pronounced. These degradations propagate to image segmentation and physical-field simulation, which ultimately undermines the reliability of evaluation re-

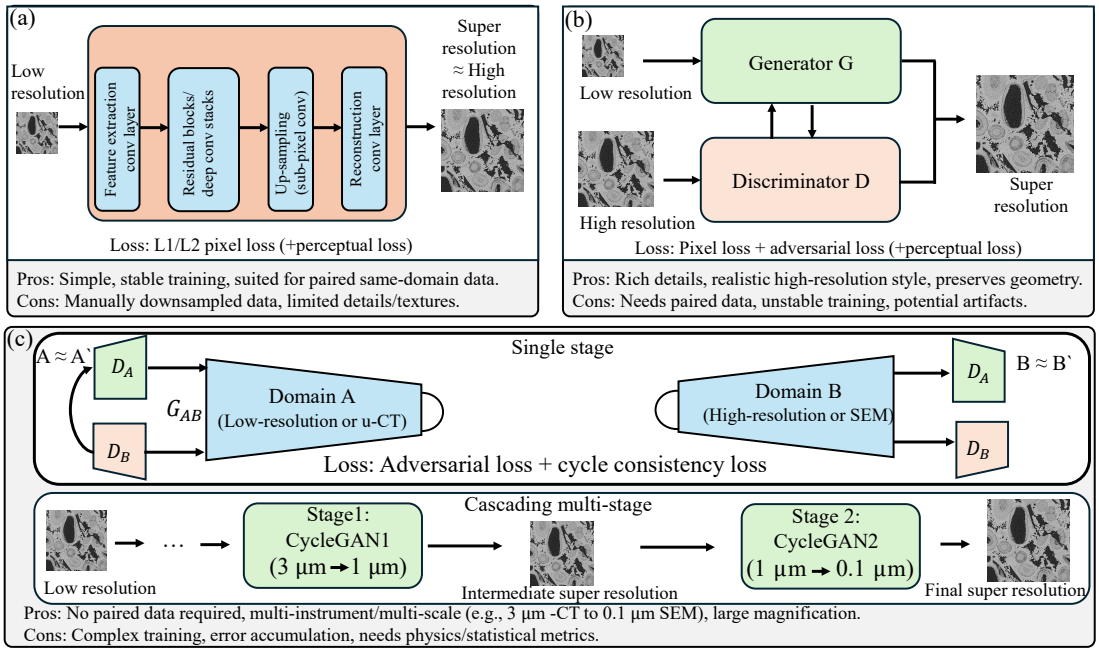


**Fig. 3**. Typical frameworks for digital rock super-resolution reconstruction: Comparison of paired CNN, GAN, and unpaired cascading CycleGAN approaches. (a) Paired convolutional neural network, (b) paired generative adversarial network and (c) unpaired cascading cycle-consistent generative adversarial network.

sults. This issue is particularly evident in multi-source imaging scenarios, including micro-CT and thin-section microscopy. Higher resolution is often accompanied by a smaller field of view, higher cost, and lower throughput. The conflict between field of view and resolution has therefore become a prominent challenge in digital rock research[32]. Improving structural resolution while preserving a representative volume is a key priority.

Conventional super-resolution methods often rely on resampling techniques such as bicubic interpolation. These algorithms are easy to implement and computationally inexpensive, but their performance is fundamentally limited. They perform smoothing on existing pixels without recovering missing high-frequency information[33, 34]. When applied to digital rock images, interpolation results in visual enlargement rather than the faithful reconstruction of pore-throat boundaries or complex mineral textures. Deep learning has transformed this situation. Super-resolution methods based on convolutional neural networks (CNNs), generative adversarial networks (GANs), and Transformers learn nonlinear mappings between low- and high-resolution images from large-scale samples. These models improve visual clarity and demonstrate advantages in pore-structure recovery, noise suppression, and physical-property preservation.

## 3.1 Evolution of super-resolution methods

The development of super-resolution reconstruction has evolved through several key architectural paradigms: Basic convolutional networks, residual and attention-enhanced models, lightweight architectures, and unpaired GANs.

The primary deep learning paradigms used in super-resolution reconstruction are summarized in Fig. 3. The frameworks illustrated in Figs. 3(a) and 3(b) rely on paired low- and high-resolution data. CNN-based frameworks utilize residual blocks and sub-pixel upsampling to achieve stable pixel-level reconstruction. These models are simple to train, though their ability to enhance fine textures is limited. Paired GAN-based methods employ adversarial learning between a generator and a discriminator. They recover high-resolution textures and local geometric features more effectively. However, their dependence on paired training data is a significant constraint, and the training process is often prone to instability. Fig. 3(c) illustrates the cascade cycle-consistent generative adversarial network (CycleGAN) strategy for unpaired conditions. This approach builds reversible mappings between domains through cycle-consistency constraints, which allows cross-resolution translation without paired data. A staged and progressive upscaling scheme is used to gradually reduce scale differences. This strategy enables cross-scale reconstruction from micrometer to submicrometer levels. Compared with paired methods, the unpaired cascade framework is more suitable for multi-instrument and multi-scale digital rock scenarios. Its training is more difficult, and the risk of error accumulation is higher. In practice, physical or statistical constraints are needed to regulate the generated results.

### 3.1.1 Basic convolutional network models

Early studies verified whether deep super-resolution methods could be applied effectively to digital rock images. A comparison of representative models, including super-resolution convolutional neural network, super-resolution residual network, enhanced deep residual network (EDSR), and wide activation deep super-resolution network, demonstrated substantial improvements in peak signal-to-noise ratio and structural

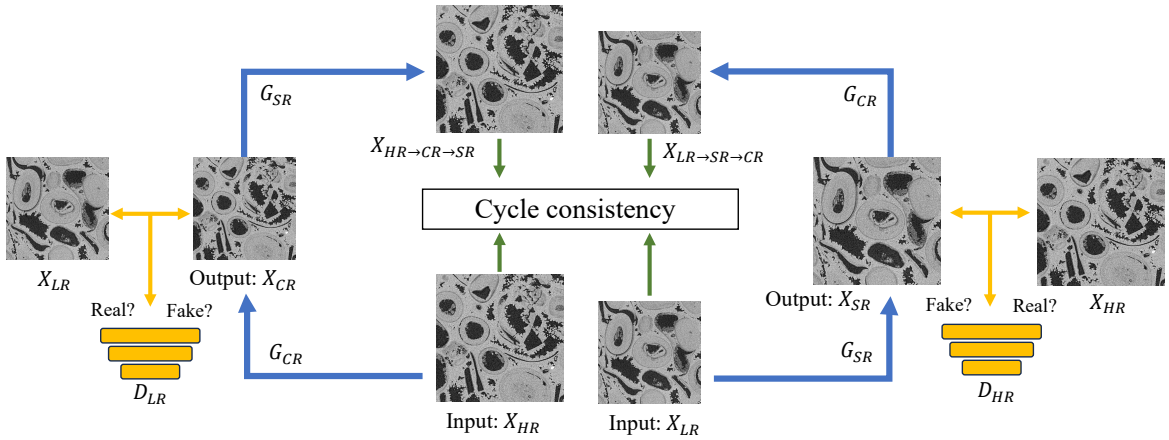


**Fig. 4**. Unpaired CycleGAN architecture utilized at each stage of the cascading super-resolution workflow[35]. The variables $X_{LR}$ and $X_{HR}$ represent the input low-resolution and high-resolution images, respectively. The generated super-resolution output is denoted by $X_{SR}$, while $X_{CR}$ represents the cycle-reconstructed image. The network employs two domain-mapping generators ($G_{SR}$, $G_{CR}$) and two discriminators ($D_{LR}$, $D_{HR}$), constrained by cycle consistency to preserve structural features.

similarity index over bicubic interpolation[36]. These models reconstructed pore edges more sharply and exhibited noise-suppression capabilities. Therefore, super-resolution reconstruction can serve as both an image enhancement tool and a preprocessing step for subsequent segmentation. Digital rock super-resolution reconstruction is a structural reconstruction process that supports pore recognition rather than a simple pixel enlargement. Even so, the training paradigm of early CNN-based methods generally relied on artificially downsampled data to build paired training sets. While this reliance is a limitation of the data preparation strategy rather than a fundamental flaw of the CNN architecture itself, their ability to handle real-world degradations caused by blur and instrument-related effects remained limited.

### 3.1.2 Residual learning and attention-enhanced models

Researchers introduced residual learning, attention mechanisms, and multi-scale feature fusion to improve the representation of complex rock structures. The CA-SRResNet model embeds residual connections and attention modules to strengthen adaptive responses to key structural regions in carbonate and sandstone images[37]. The HAMSR model adopts mixed attention and a multi-branch structure to prioritize the restoration of pores, fractures, and textures[38]. This approach improves edge sharpness and micropore recognition while controlling parameter counts. The RSMAN model introduces spatial multi-scale attention to handle rock images with complex spatial relationships[39]. These studies indicate that digital rock images require super-resolution models with efficient feature selection and multi-scale fusion due to high spatial heterogeneity. Model complexity, however, often results in higher computational costs. This trade-off must be considered for engineering applications.

### 3.1.3 Lightweight architectures

Lightweight modeling has become an important route toward the practical deployment of digital rock super-resolution reconstruction. For rock thin-section images, the OmniSR-M network reduces model complexity while maintaining reconstruction quality[40]. The LMDN model utilizes lightweight multi-scale feature distillation and gated large-kernel attention to recover microstructural features with a small parameter budget[41]. A basic trade-off still exists between lightweight design and high-frequency detail recovery. This limitation becomes more apparent at high magnifications where lightweight models may struggle to represent fine pore throats in complex rocks.

### 3.1.4 Unpaired generative adversarial networks

Obtaining strictly registered low- and high-resolution pairs is challenging. This difficulty has promoted the development of GAN-based unpaired super-resolution approaches. The NL-CycleGAN model introduces a non-local attention mechanism into the CycleGAN framework to enable unsupervised super-resolution reconstruction of rock CT images[42]. This method learns the mapping between domains under unpaired conditions, which is consistent with real digital rock acquisition workflows. Reconstructed images achieve a balance between global style preservation and local detail recovery. However, GAN-based methods may introduce pseudo-textures or false pores. Their validity must be verified through pore geometry and physical-property analysis.

A multi-stage cascading framework effectively performs unpaired cross-domain super-resolution reconstruction[35]. In this workflow, initial low-resolution micro-CT images are progressively mapped through intermediate scales to reach a target high-resolution SEM level. The core network at each stage of this progressive cascade utilizes an unpaired CycleGAN structure, as detailed in Fig. 4. The initial low-resolution image is enhanced step by step rather than undergoing a single drastic magnification. This staged strategy successfully prevents the structural distortions typically caused by large magnification jumps. Experimental results demonstrate that this cascading method outperforms single-stage ESRGAN models in micro-

pore recognition. The approach significantly improves small-pore visibility. However, researchers must carefully control potential error accumulation between consecutive enhancement stages.

The PatchSRGAN model emphasizes the role of local patch-level discrimination in pore-geometry recovery[43]. When the discriminator output patch size reaches a specific level, the fidelity of flow-related indicators such as porosity and pore-size distribution improve significantly. The model also remains robust under noisy conditions. Multi-scale GAN fusion methods further extend super-resolution reconstruction by integrating large-field 3D micro-CT geometries with fine-grained SEM surface details[44]. These methods produce synthetic digital rock images that combine a wide field of view with high precision. They show good agreement with real samples in porosity, two-point correlation functions, and effective permeability. However, these outputs are statistical reconstructions, and their one-to-one correspondence with real local structures should be interpreted with caution.

### 3.2 3D and physics-consistent models

As pore-network characterization moves into 3D space, 2D slice-based super-resolution reconstruction is no longer sufficient for many applications. Extending the VDSR model to 3D digital rock reconstruction demonstrated that at a 5× upscale factor, the permeability errors for the training and prediction sets of matrix shale are 0.21 and 0.22, respectively, while those for fractured shale are 0.11 and 0.15[31]. To optimize computational efficiency, the 3DSAFMN model combines spatial adaptive feature modulation and multi-scale information in a 3D framework; compared with existing advanced deep learning algorithms, this model reduces the parameter count by 45.5% and increases the reconstruction speed by 1.70 times[45]. The reliability of the super-resolved images was verified through geometric topology and direct flow simulation.

Transformer models provide an alternative technical route for long-range dependency modeling. The EAST model integrates self-attention and channel attention mechanisms, achieving a 1.85× speedup and a 78% reduction in parameters compared to existing advanced models[46]. Furthermore, EAST reduces the relative error of absolute permeability by 18.5% and 33% over the traditional tricubic interpolation method across two external samples, confirming the physical reliability of the output. The PAST model embeds physical accuracy directly into the design using a dual-path Transformer[47]. Compared to EAST, PAST improves the peak signal-to-noise ratio by 0.14 dB on coal, 0.37 dB on sandstone, and 0.13 dB on carbonate; notably, it yields a porosity error below 0.1% and a minimal permeability deviation of 0.15 millidarcies on coal samples. Additionally, the HLF-SASR model features a compact design with only 19.81 million parameters and utilizes cross-frequency collaborative attention on $200 \times 200$ and $400 \times 400$ image pairs for 2× super-resolution. This network successfully reduces porosity errors to 0.416% for carbonate and 0.114% for sandstone samples, restoring distinct pore-throat boundaries while suppressing pseudo-textures[48].

### 3.3 Summary and limitations

A systematic comparison of technical routes and application ranges is provided in Table 1. Research has evolved from early CNN-based reconstruction toward joint modeling for 3D structure and physical consistency. Basic CNN methods demonstrated the feasibility of deep learning for edge recovery. Models with attention mechanisms and multi-scale fusion significantly improved the restoration of complex pore structures. GAN-based unpaired methods addressed the difficulty of obtaining real high-resolution data. Transformers and 3D methods have pushed super-resolution results toward practical applications such as flow simulation.

Despite these significant advancements, super-resolution reconstruction in digital rock technology still faces inherent limitations. Models heavily reliant on GANs frequently hallucinate false pore connectivity, severely distorting downstream permeability evaluations. This reliability issue is compounded by the massive memory consumption of direct 3D volumetric processing, which restricts practical deployment on typical hardware. Meanwhile, conventional evaluations based on purely visual metrics or synthetic downsampling fail to capture the complex physical noise of real micro-CT scanners, leaving a persistent gap between visual enhancement and true petrophysical fidelity.

## 4. Image segmentation

Accurate characterization of microscopic pore structures is a central requirement as oil and gas exploration advances into deeper and more unconventional reservoirs. Digital rock technology has emerged as a vital tool by providing essential high-resolution pore information. Within the digital rock analysis workflow, segmentation accuracy directly determines the reliability of physical property prediction and flow simulation.

Traditional segmentation methods, such as Otsu's thresholding and watershed algorithms, are highly sensitive to noise and struggle to resolve complex multi-mineral coexistence and blurry pore-throat boundaries. Deep learning has revolutionized this step by transforming segmentation from grayscale thresholding into a high-dimensional semantic feature extraction process.

### 4.1 Lightweight CNNs and U-Net variants

As shown in Fig. 5, the U-shaped network (U-Net) serves as a foundational architecture in core image segmentation due to its symmetric encoder-decoder structure and skip connection mechanism. The encoder path captures the semantic information of complex pore geometries. The decoder path restores spatial resolution to delineate precise pore boundaries. Skip connections fuse high-resolution structural features, such as rock matrix edges, with deeper semantic information. This mechanism reduces data loss during training and improves the accuracy of pore network extraction.

Enhanced U-Net variants have been developed to address the inherent heterogeneity of core images. An enhanced U-Net model utilizing depthwise separable convolution and feature deep concatenation reduces model complexity while improving segmentation efficiency[49]. For blasted rock fragments, U-

**Table 1**. Comparison of major super-resolution methods for digital rock images.

| Model | Features | Advantages | Limitations |
|---|---|---|---|
| HAMSR[38] | Mixed attention; multi-branch network | Strong recovery of pores and fractures; fewer parameters | Mainly 2D; lack of 3D topological constraints |
| RSMAN[39] | Nested residual learning; multi-scale attention | Strong performance on complex spatial relationships | Redundant computation on simple images |
| OmniSR-M[40] | Lightweight multi-branch aggregation | Efficient and suitable for deployment | Inferior in very fine detail recovery |
| LMDN[41] | Feature distillation; gated large-kernel attention | Clear lightweight advantage; preserves microstructures | Performance limited by model capacity |
| NL-CycleGAN[42] | CycleGAN with non-local attention | No paired data required; suits real scenarios | Risk of pseudo-details; limited GAN stability |
| Cascading CycleGAN[35] | Cascaded unpaired label | Suitable for large magnification and cross-scale mapping | Complex pipeline; error may accumulate |
| PatchSRGAN[43] | Patch-based discriminator | Better local geometric fidelity; noise robust | Training is sensitive to sample quality |
| VDSR-3D[31] | Deep 3D convolution | Supports 3D permeability evaluation | Limited long-range dependency modeling |
| 3DSAFMN[45] | Spatial adaptive modulation; multi-scale fusion | High accuracy; better parameter efficiency | Needs verification for anisotropic samples |
| EAST[46] | Self-attention and channel attention fusion | Good quality; fast inference; robust | Cost increases with image size |
| PAST[47] | Dual-path Transformer; intensity normalization | Preserves CT values; joint segmentation output | Requires high-quality and large-scale data |

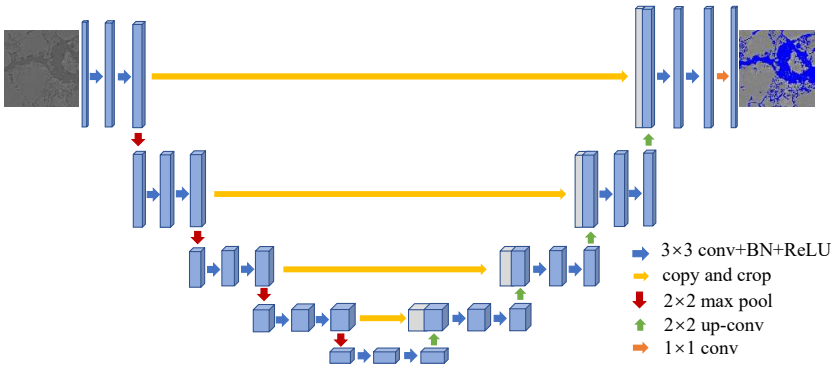


**Fig. 5**. Architecture of the U-shaped network for digital rock pore segmentation. The left side represents the encoder path for extracting features from raw grayscale rock images. The right side represents the decoder path for upsampling and generating the final blue pore segmentation mask. The horizontal arrows indicate skip connections. These connections fuse high-resolution spatial details, such as micro-fractures and pore boundaries, with deeper semantic information.

CARFnet incorporates attention mechanisms, residual learning modules, and focal loss functions to successfully mitigate severe class imbalance between narrow fractures and the background matrix[50]. This approach achieved a segmentation accuracy of 97.11% by reducing background interference. Furthermore, Res-VGG-UNet integrates VGG and ResNet elements within the U-Net framework to handle rock fractures and mineral coexistence. This hybrid architecture demonstrates superior geometric feature extraction capabilities compared to DeepLabv3+ in complex fracture recognition[51].

## *4.2 Large vision models and multimodal Transformers*

While CNNs perform well on local feature extraction, they often struggle with long-range dependencies and global contextual understanding. The emergence of large vision models has introduced a new paradigm for core image segmentation. CoreSAM serves as a lightweight fine-tuning method for fine-grained sandstone particle evaluation[52]. By adapting the segment anything model (SAM) with a rank-decomposition

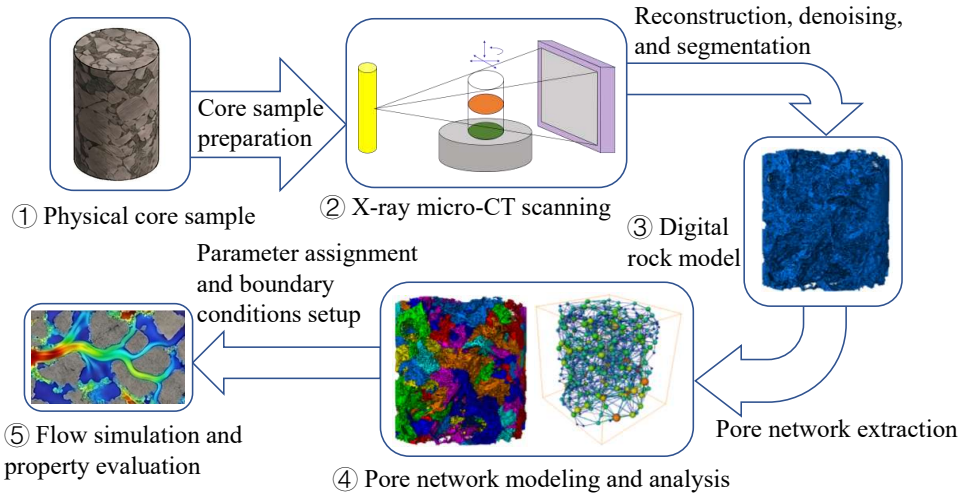


**Fig. 6**. Workflow for predicting physical properties in digital rock physics.

matrix adapter, this method enables rapid boundary extraction for multispectral rock particles, exhibiting strong zero-shot generalization capabilities. This system supports fully automated quantitative evaluation of grain size, sorting, and contact types.

Multimodal perception has further expanded the application scope of segmentation models. SAM-FuseNet combines SAM2 with cross-modal spatial attention and long-range context modeling to fuse RGB and thermal imaging data[53]. This framework improves rock recognition under complex lighting conditions or degraded imaging environments by leveraging complementary features from different sensors. Transformer-based architectures have also shown stronger global feature modeling capabilities than conventional CNNs for specific rock samples. A Multi-Channel Attention Transformer combined with Copy-Paste augmentation addresses the limited sample size of rock thin sections[54]. This approach integrates image sequence features acquired under different lighting angles to improve segmentation accuracy. Additionally, RDT-FragNet combines deformable convolutional networks with Transformer attention mechanisms[55]. The model maintains sensitivity to local edges while retaining awareness of global geometric structures, ensuring particle size distribution recognition remains close to the actual physical state.

### 4.3 Joint multi-task learning frameworks

Rather than treating super-resolution and segmentation as isolated steps, joint optimization of super-resolution methods and segmentation addresses the difficulty of identifying micron-scale pores in low-resolution images. A joint neural network framework integrates an EDSR module and a SegNet segmentation module in parallel[56]. This end-to-end design enables high-quality reconstruction and accurate segmentation of 3D digital rock. The resulting physical analysis remains highly consistent with high-resolution ground truth values. Similar strategies have been applied to unconventional reservoirs with complex microstructures. For shale characterization, a GAN-based integrated framework facilitates image enhancement, 8× super-resolution reconstruction, and automatic multi-mineral segmentation[57]. This method eliminates the need for paired samples and successfully reconstructs missing fine pores in low-resolution images. By sharing feature extraction layers, these multi-task models ensure that the segmented geometric models are physically consistent and ready for property evaluation.

### 4.4 Summary and limitations

To explicitly address the comparative differences among these evolving methodologies, Table 2 provides a systematic evaluation of representative segmentation models. In terms of data types, the field is transitioning from single-channel 2D grayscale images to multi-channel, true 3D, and multimodal inputs, which inherently increases data processing complexity. Regarding architectures, lightweight CNNs excel in local boundary precision, whereas large Transformers offer unparalleled generalization across different lithologies. Despite this progress, ongoing limitations persist. Researchers continually struggle to balance massive computational costs with inference accuracy while ensuring the strict physical consistency of the segmented multi-mineral topologies. Once the pore structure is accurately delineated through these methods, the resulting geometric models serve as the precise basis for quantitative evaluation, naturally driving the subsequent extraction and prediction of macroscopic physical properties.

## 5. Physical properties prediction

Reconstructed and segmented 3D pore structures provide high-quality representations of digital rock images. However, reservoir evaluation and development decisions ultimately rely on the quantitative prediction of physical properties, including porosity, permeability, and relative permeability. Deep learning has emerged as a vital direction in digital rock physics because it can learn the direct mapping between pore geometry and physical responses from digital rock images[58].

A standard workflow is generally followed in digital rock physics prediction, as illustrated in Fig. 6. Specifically, ①

**Table 2**. Comparative analysis models for digital rock image segmentation.

| Method | Input data | Advantages | Limitations |
|---|---|---|---|
| Enhanced U-Net[38] | Core images | Reduced model complexity; improved segmentation efficiency | Weak at modeling long-range global dependencies |
| U-CARFnet[50] | 2D grayscale core and fragment images | Effectively mitigates severe class imbalance in fracture networks | Limited global context; struggles with massive 3D volumetric data |
| Res-VGG-UNet[51] | 2D and 3D CT volumetric slices | Superior local geometric feature extraction for heterogeneous minerals | Higher computational overhead compared to standard U-Net |
| Multi-Channel Attention Transformer[54] | Rock thin section images | Strong global modeling; overcomes small sample size limits via Copy-Paste | Requires image sequences under different lighting angles |
| CoreSAM[52] | Multispectral 2D rock images | Rapid boundary extraction with extremely strong zero-shot generalization | High hardware requirements; requires task-specific adapters |
| RDT-FragNet[55] | Rock fragment images | Balances local edges sensitivity with global geometric structure awareness | High complexity from combining CNN and Transformer |
| SAM-FuseNet[53] | Multimodal data (RGB and Thermal) | Robust to noisy environments by fusing multi-source complementary data | Complex architecture; requires perfectly aligned multimodal inputs |
| EDSR-SegNet[56] | Low-resolution 3D digital rock images | End-to-end high-quality reconstruction and accurate segmentation | High memory consumption for joint 3D optimization |

a representative core sample is prepared to establish the physical basis for the digital model. Then, ② high-resolution raw data are acquired through X-ray micro-CT scanning. Following data acquisition, ③ the images undergo denoising, registration, and segmentation to be converted into pore-solid structural models. Subsequently, ④ the geometric and topological features are extracted to construct the pore network model for detailed analysis. Finally, ⑤ fluid simulations are performed on the model to achieve accurate prediction of macroscopic physical properties.

## *5.1 Porosity and permeability*

Current deep learning applications for porosity and permeability prediction in digital rock technology follow three primary paradigms: Multimodal attribute fusion, end-to-end direct image regression, and hybrid modular workflows. Multimodal methods integrate microstructural images with high-resolution physical attributes, while end-to-end frameworks map raw digital rock images directly to macroscopic properties. Complementing these approaches, hybrid modular pipelines combine deep learning segmentation with traditional geometric characterization. Together, these pathways establish rapid, data-driven mapping functions that bypass computationally prohibitive direct numerical fluid simulations, offering scalable solutions for digital petrophysical analysis.

Integrating physical logs with spatial data provides a powerful entry point for comprehensive characterization. Specifically, a dual-branch CNN workflow was developed to estimate both porosity and permeability by fusing X-CT images with high-resolution continuous physical attributes, such as bulk density and photoelectric factors[59]. This multimodal strategy integrates structural pore geometry with mineralogical information more effectively than purely binary image methods. Consequently, the trained network achieves excellent predictive performance, yielding a determination coefficient of 0.996 for the porosity model and a determination coefficient of 0.983 for the permeability model. This framework enables highly accurate property estimations over multiple rock intervals at a high vertical resolution of 0.625 mm. However, its transferability across different core intervals remains sensitive to data consistency and lithological variations.

To bypass the reliance on auxiliary physical attributes, researchers have increasingly pursued purely image-driven, end-to-end estimation pathways. A representative model utilizes a vision Transformer with soft convolutional inductive bias to predict permeability[60]. Convolutional kernels are incorporated into the architecture to capture local pore-throat details, while self-attention mechanisms describe global connectivity. Training data were generated using an improved quaternary stochastic growth method to produce multiple classes of pore structures, with the corresponding permeability labels computed directly on these same synthetic images via the lattice Boltzmann method (LBM). However, using LBM-generated labels for training potentially introduces a systematic bias, causing the deep learning model to inherit inherent numerical artifacts from fluid simulations, such as grid resolution limitations, boundary condition simplifications, or sub-resolution effects. Despite this, the model achieved a correlation coefficient of 0.985 on synthetic test data and maintained a high coefficient of 0.890 on a generalization dataset consisting of

200 natural rock samples via a two-stage transfer learning strategy. This class of methods is computationally efficient and has the potential to replace costly direct numerical simulations, though its robustness is still constrained by the mismatch between synthetic pore structures and natural rocks.

Enhancing the multi-lithology generalization of these image-driven methods remains a major focus as models transition to diverse natural formations. To address this, deep ConvNet architectures involving residual network and ResNeXt were deployed to learn the complex geometry of pore spaces for flow-based characterization[61]. For unseen sandstone samples, the network achieves an explained variance score of 0.87 and a mean absolute error of 0.040 darcies compared to simulated values, while executing permeability inference in less than 120 ms per sample.

While standard ConvNets excel at spatial feature extraction, alternative architectures aim to better capture topological relationships within the pore space. For example, a hybrid framework named Pore-GNN integrates CNNs with graph neural networks to predict absolute permeability[62]. A pre-trained CNN extracts multidimensional feature vectors from micro-CT images, which are subsequently embedded into graph structures for graph neural network processing. This architecture improves the coefficient of determination by up to 9% over baseline models and predicts permeability more than 500 times faster than numerical solvers.

In contrast to these single-stage end-to-end regressions, modular workflows offer a distinct advantage by combining deep learning segmentation with multi-stage property estimation for enhanced interpretability[63]. In this hybrid framework, a U-Net model achieves exceptional accuracy in segmenting raw CT images, allowing engineers to extract key geometric attributes like the hydraulic radius. These extracted features are evaluated using long short-term memory and random forest regressors for permeability estimation. Furthermore, a refined U-Net variant is deployed to directly predict the microscopic flow velocity field, reducing computational time by 98.59% compared to conventional LBM.

### 5.2 Relative permeability curves

Relative permeability curves are essential functional descriptors of multiphase flow behavior. These curves are controlled by the interaction of pore topology, wettability, and interfacial tension. Traditional estimation has relied primarily on pore network models or direct numerical simulation. To address the high dimensionality and strong nonlinearity of this problem, a deep learning framework has been developed to predict oil-water relative permeability curves directly from 3D images and fluid property parameters. In this design, 3D images, wettability, and interfacial tension served as inputs, while relative permeability curves calculated from an improved pore network model were used as labels. A hybrid model combining convolutional layers and ConvLSTM achieved high predictive accuracy for both sandstone and volcanic rock samples. Once trained, the model could rapidly evaluate how variations in wettability and interfacial tension affect curve shape[64]. The direct prediction of relative permeability curves extends digital rock analysis beyond scalar properties toward dynamic functions reflecting displacement efficiency. This approach offers significant value for enhanced oil recovery studies and displacement strategy optimization. Nevertheless, current methods depend largely on pore network simulation as the source of ground truth. The deviation between these simulated labels and systematic two-phase flow experiments has not yet been fully quantified. Furthermore, physical constraints such as endpoint consistency and monotonicity have not been systematically incorporated into the loss functions. Practical application therefore still requires calibration against a limited number of real displacement experiments.

### 5.3 Summary and limitations

To explicitly address the comparative differences among these evolving methodologies, Table 3 provides a systematic evaluation of representative physical property prediction models. The limited availability of high-quality digital rock samples with experimental labels remains a central bottleneck. Systematic two-phase flow datasets are exceptionally scarce. Consequently, most models still rely on numerical simulation results as supervisory signals[22, 65]. Synthetic pore structures can expand dataset sizes. However, they fail to capture the natural complexities introduced by multiscale pores, fractures, and diagenetic heterogeneity. Model generalization across different lithologies and basins requires much broader evaluation. Furthermore, physical property predictions lack a mature framework for enforcing rock-physics constraints. Crucial factors like endpoint conditions and porosity-permeability relationships are often ignored. Methods for upscaling microscopic pore-scale features to well-scale and seismic-scale parameters are also not well established. As a result, digital rock technology currently functions primarily as a local calibration tool in large-scale reservoir evaluation. Understanding these static properties is vital, yet practical engineering demands a deeper insight into dynamic behaviors. Therefore, extending these static predictions to complex dynamic flow simulations becomes an essential progression.

## 6. Flow simulation

Digital rock technology aims to reconstruct 3D pore structures from high-resolution images. This process establishes a quantitative link between microscopic pore morphology and macroscopic flow response. Predicting vital transport properties, such as permeability, relative permeability, and capillary pressure, relies heavily on this quantitative link. Traditional approaches often utilize direct numerical solvers, including Navier-Stokes equations, Stokes equations, or LBMs. These methods yield high-precision pore-scale flow fields; however, they incur prohibitive computational costs when applied to complex 3D geometries, high-resolution dynamic processes, or multi-phase interface tracking[66–68]. Deep learning has introduced a transformative alternative for flow simulation[69]. Rather than mapping structures exclusively to static scalar properties, advanced data-driven proxies are now utilized. These frameworks rapidly approximate flow fields[70, 71], dynamically reconstruct multidimensional physical fields, or

**Table 3**. Comparative analysis models for digital rock physical property prediction.

| Method | Input data | Advantages | Limitations |
|---|---|---|---|
| Dual-branch CNN[59] | X-CT images combined with bulk density and photoelectric factors | High resolution of 0.625 mm with excellent accuracy for both porosity and permeability | Sensitive to core data consistency and lithological variations |
| Vision Transformer[60] | Synthetic quaternary growth and natural rock images | Balances local details and global connectivity | Inherits numerical artifacts from LBM labels alongside synthetic-natural mismatch |
| Deep ConvNet[61] | Unseen natural sandstone micro-CT images | Extremely fast inference taking less than 120 ms per sample | Struggles to explicitly capture complex topological network relationships |
| Pore-GNN[62] | Micro-CT feature vectors mapped to graph networks | Directly models pore topology and operates 500 times faster than conventional numerical solvers | Heavily dependent on the quality of prior graph abstraction |
| Modular pipeline[63] | Raw CT images processed via U-Net and long short-term memory or random forest | Enhances physical interpretability and reduces velocity simulation time by 98.59% | Susceptible to multi-stage error accumulation across modules |
| Conv-ConvLSTM[64] | 3D images with fluid wettability and interfacial tension | Resolves strong nonlinearity and predicts full relative permeability curves rapidly | Relies on simulated labels with unquantified bias and lacks endpoint consistency |

accelerate traditional numerical solvers from different entry points within the simulation pipeline.

## 6.1 Direct fluid flow field reconstruction

Directly predicting pore-scale velocity and pressure fields offers a deeper understanding of microscopic flow laws without explicit numerical iteration. Unlike scalar regression, flow field reconstruction is a high-dimensional output problem that demands significant network capacity.

The PoreFlow-Net architecture mapped binary digital rock images to steady-state velocity fields using 3D CNNs[72]. Although the training phase primarily utilized synthetic geometries, the network successfully reproduced overall flow patterns and main channel distributions in real sandstone and carbonate samples. The computational speedup reached several orders of magnitude compared to traditional LBM calculations, demonstrating that deep convolutional networks can effectively approximate complex flow behaviors when provided with sufficiently diverse training data.

Relying solely on local geometric information can occasionally lead to errors in representing global pressure drops and long-range connectivity. To address this limitation, accuracy was improved by incorporating coarse-grid flow fields as an additional input to drive fine-grid reconstruction[73]. The coarse-grid data, derived from low-resolution fast simulations. These data carried essential physical information regarding the global pressure gradient and primary flow paths. This coarse physics-fine geometry strategy balanced global transport trends with local structural details, proving highly effective for heterogeneous samples containing large vugs or complex cavities.

## 6.2 Deep learning for numerical solver acceleration

Instead of completely replacing traditional physics-based models, another prominent strategy utilizes deep learning to accelerate heavy numerical calculation bottlenecks while retaining physical rigor. The primary motivation stems from the high computational toll and stability limitations associated with simulating multiphase systems. In conventional implementations, resolving three-dimensional transport physics inside intricate porous domains remains a grand challenge, especially when localized apertures and varying wettability conditions dictate complex two-phase flow dynamics and fluid segregation inside 3D fractures[75]. Capturing such intricate interfacial dynamics and transient saturation evolution demands massive computational grid resources. Furthermore, conventional phase-field LBMs face stringent limitations regarding computational efficiency, numerical stability, and volume conservation when executing multi-dimensional flow behaviors[76]. To bypass these substantial execution costs, deep neural frameworks can be seamlessly integrated into the early stages of the simulation pipeline to optimize solver initialization.

Optimizing solver initialization through intelligent convergence guessing represents a highly efficient strategy for runtime reduction. For instance, a CNN was deployed to predict an initial steady-state velocity field directly from the raw pore geometry, serving as a warm-start for subsequent traditional LBM iterations[77]. This hybrid prediction-correction mode significantly reduced the required convergence steps, shortening the total computation time by approximately one order of magnitude without losing the physical conservation guaranteed by the final numerical steps.

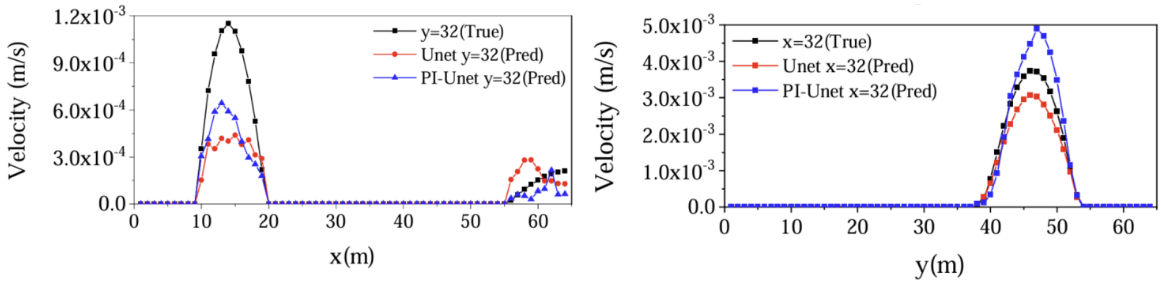


**Fig. 7**. Comparison of velocity profiles at middle cross-sections among LBM calculations, traditional U-Net, and PI-U-Net-BGK[74].

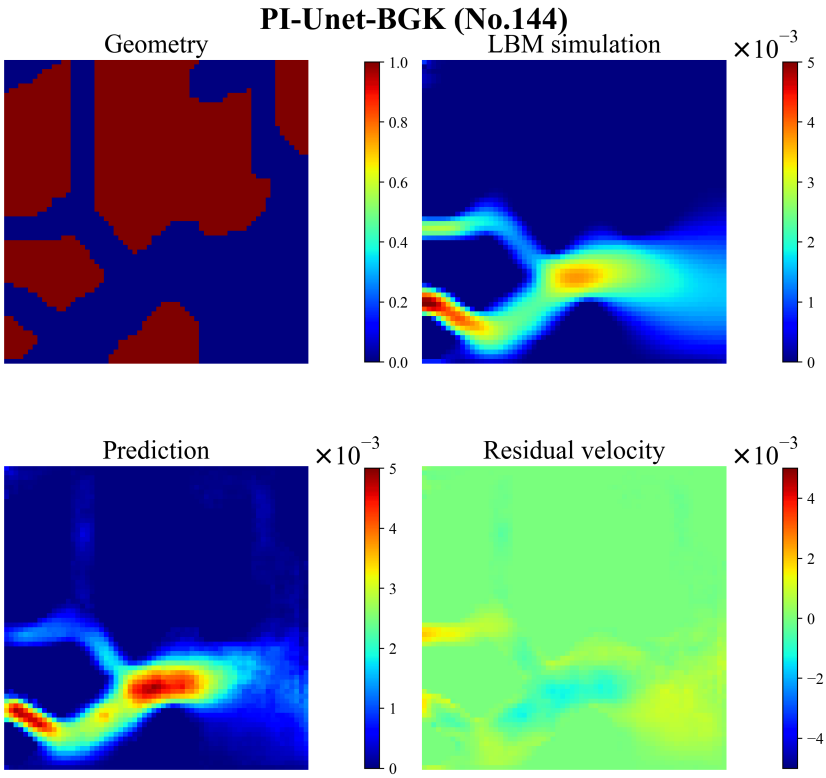


**Fig. 8**. Visual comparison of velocity distribution between LBM simulation and the PI-U-Net-BGK network[74].

### *6.3 Physics-aware learning and geometric constraints*

While the aforementioned hybrid modes rely on subsequent traditional steps to maintain physics, recent advanced frameworks embed physical mechanisms directly into the neural training process to prevent models from violating fundamental fluid mechanics. Specifically, a physics-informed U-Net network with the Bhatnagar-Gross-Krook collision model (PI-U-Net-BGK) was established[74]. This framework integrated the Bhatnagar-Gross-Krook collision operator as a physical constraint within the loss function, enabling the physics-constrained network to capture velocity peaks and sharp gradients more effectively than traditional U-Nets, especially in narrow pore throats.

A detailed comparison of velocity profiles at middle cross-sections highlighted the clear differences among LBM benchmarks, traditional U-Nets, and the PI-U-Net-BGK model Fig. 7. The PI-U-Net-BGK results aligned closely with LBM benchmarks regarding both peak shape and magnitude, whereas traditional U-Net architectures systematically underestimated velocity peaks. Visual comparisons of velocity distributions further illustrated the superior capability of physics-informed frameworks in characterizing complex pore-scale flows Fig. 8. Despite these advancements, physics-embedded models still face stability challenges when applied to transient flows or highly heterogeneous shale reservoirs.

Furthermore, convolutional networks are inherently restricted by regular voxel grids, which can limit the boundary accuracy of complex, curved pore spaces. Physics-informed neural networks paired with point cloud representations offer a viable solution to this grid limitation. A physics-informed PointNet framework effectively modeled pore-scale Stokes flow by using point clouds as input[78]. This method focused exclusively on the void space, thereby reducing memory over-

**Table 4**. Comparative analysis models for digital rock physical property prediction.

| References | Advantages | Limitations | Practical scenarios |
|---|---|---|---|
| Santos et al. [72] | Near-instantaneous inference; no boundary setup required | Low physical consistency; weak generalization to unseen rock types | Fast velocity and pressure field mapping for standard geometries |
| Zhou et al. [73] | Balances local details with global pressure gradients; improves accuracy in large vugs | Requires running a low-resolution numerical solver first | Flow reconstruction in heterogeneous samples with large vugs or cavities |
| Wang et al. [77] | Reduces iteration steps while retaining strict physical conservation laws | Speedup depends heavily on the baseline traditional numerical solver | Accelerating iterative solvers via warm-start field prediction |
| Zhao et al. [74] | Enhances gradient capturing; enforces strict local physical laws in throats | Faces convergence and stability risks in transient or shale flows | Flow field prediction in narrow channels and tight pore throats |
| Kashefi and Mukerji [78] | Eliminates solid matrix calculations; smooth boundary handling without grids | High training cost for complex or large-scale point clouds | Flow modeling along highly irregular and curved grain boundaries |
| Kamrava et al. [79] | Captures history-dependent transient flows without continuous simulations | High training complexity; scale scaling limitations | Tracking dynamic fluid evolution over time, such as multi-phase displacement |

head and eliminates redundant calculations on the solid matrix. Point clouds provided a smoother representation of complex boundaries and allowed for non-uniform spatial sampling, enabling the model to maintain high predictive accuracy in critical regions even when full-field supervision data remained unavailable.

In the temporal domain, the integration of physical constraints can be extended to model transient flow behaviors. By embedding mass conservation and the Navier-Stokes equations into deep networks, researchers can predict flow evolution over time[79]. This approach reduces the reliance on continuous, direct numerical simulation. It also holds significant potential for modeling dynamic, history-dependent processes like multi-phase displacement.

### 6.4 Summary and limitations

To provide a systematic evaluation of deep learning applications in digital rock physics, Table 4 comprehensively compares the input data, applications, advantages, limitations, and key references of the three mainstream paradigms. Current limitations of deep learning-based flow simulation stem heavily from the reliance on the black-box fitting of numerical data. This data dependency restricts model generalization across different lithologies and scales. Additionally, unconventional reservoirs contain broad pore size distributions, ranging from nanometers to centimeters. Single-resolution images are fundamentally insufficient to capture this structural span. Integrating robust physical mechanisms and multi-scale data remains a substantial hurdle for future digital rock research.

## 7. Future outlook

1) Foundation models and universal domain adaptation

Current segmentation and super-resolution models lack generalizability across different rock fabrics and scanner modalities. Future research should focus on domain-specific foundation models. These models can be pretrained on massive 3D rock datasets using self-supervised learning. This pretraining captures universal latent representations of complex pore topologies. Once established, the foundation models require minimal labeled samples for downstream fine-tuning. Additionally, advanced unsupervised domain adaptation methods are necessary. These methods help mitigate performance drops caused by variations in scanning resolutions and geological basins.

2) Architecture-level physics-consistent regularization

Purely data-driven models often generate physically non-compliant results in image segmentation and fluid simulation. Traditional penalty-based loss functions cannot strictly guarantee physical consistency. Future network architectures should directly hard-encode fundamental physical laws into network layers and activation functions. These governing equations include fluid-solid boundary conditions, mass conservation, and momentum conservation. Ensuring exact physical conservation at the architectural level is essential to generate reliable velocity and pressure fields.

3) Holistic multiscale and multimodal data fusion

Unconventional formations exhibit extreme diagenetic heterogeneity. Single-scale imaging cannot capture the full span of pore structures or support macroscopic upscaling. Future workflows should deploy multi-resolution neural networks to seamlessly fuse nested 3D volumetric data, such as nano-CT and micro-CT geometries. Furthermore, deep learning frameworks must integrate 3D structural images with non-imaging signals. Merging image matrices with well logs, mineral maps, and laboratory core measurements will effectively link pore-scale fluid behaviors to reservoir-scale engineering decisions.

4) Hybrid numerical and deep learning workflows for fluid simulation

Traditional solvers face high computational costs when tracking multiphase fluid interfaces. Purely data-driven models often violate fundamental conservation laws and fluid boundaries. Future research should focus on tight coupling mechanisms where neural networks accelerate specific numerical bottlenecks. For instance, deep learning can predict transient velocity and pressure fields to provide a warm-start initialization. The classical solver then performs the final iterations to ensure exact mass balance and boundary compliance. This coupled paradigm balances computational efficiency with strict physical rigor.

## 8. Conclusions

Deep learning is fundamentally reshaping the research paradigm of digital rock image analysis, evolving from traditional, experience-driven single tasks toward intelligent workflows characterized by multi-scale, high-precision, and multi-task collaboration. A comprehensive technical chain has emerged to support this transition, encompassing super-resolution reconstruction, image segmentation, physical property prediction, and flow simulation. Significant breakthroughs have been achieved in restoring microscopic pore-throat details, identifying complex medium boundaries, and accelerating parameter inversion. Moving forward, the field is transitioning from pure data-driven visual mapping to a deep synergy between physical mechanisms and neural networks. As algorithmic frameworks continue to mature, digital rock technology will successfully bridge the critical gap between pore-scale microscopic characterization and macro-scale reservoir engineering decisions.

## Acknowledgements

The authors are grateful to the National Natural Science Foundation of China (No. 42302143).

## Conflicts of interest

The authors declare no competing interest.



## References

[1] Alameedy, U., Abdul-Majeed, G., AlSaedi, H. N., et al. Role of artificial intelligence in enhancing efficiency and sustainability in middle eastern oil and gas: A review. Arabian Journal for Science and Engineering, 2026, 51(1): 1043-1066.

[2] Zhang, Q., Liu, J., Gao, Z., et al. Review on the challenges and strategies in oil and gas industry's transition towards carbon neutrality in china. Petroleum Science, 2023, 20(6): 3931-3944.

[3] Wang, F., Chen, D., Li, M., et al. A novel method for predicting shallow hydrocarbon accumulation based on source-fault-sand (SF-Sd) evaluation and ensemble neural network (ENN). Applied Energy, 2024, 359: 122684.

[4] Yang, D., Wang, P., Lu, J., et al. Leakage detection of oil and gas pipelines based on a multi-channel and multi-branch one-dimensional convolutional neural network with imbalanced samples. Computers in Industry, 2025a, 173: 104356.

[5] Yang, H., Feng, W., Diao, H., et al. Artificial intelligence in shale gas and oil: A comprehensive review of applications and challenges. Green and Smart Mining Engineering, 2025b, 2: 259-277.

[6] Tu, R., Liao, Q., Liu, Y., et al. Transformation of China's petroleum pipeline industry in the renewable energy era: A review. Renewable and Sustainable Energy Reviews, 2026, 233: 116863.

[7] Xiong, A., Zhou, Y., Shen, Z., et al. Total pore-throat size distribution characteristics and oiliness differences analysis of different oil-bearing tight sandstone reservoirs - A case study of Chang6 reservoir in Xiasiwan Oilfield, Ordos Basin. Fractal and Fractional, 2025, 9(11): 729.

[8] Davoodi, S., Makarov, N., Wood, D. A., et al. A review of artificial intelligence techniques applied to subsurface oil and gas reservoir simulation. Applied Soft Computing, 2026, 198: 115222.

[9] Liu, J., Li, J., Song, Z., et al. Flow characteristics of shale oil and its geological controls: A case study of the Lucaogou Formation in the Jimsar Sag, Junggar Basin. Petroleum Science, 2025, 22(10): 3900-3914.

[10] Wang, D., Ma, Y., Hu, Z., et al. Pore network modeling of gas-water two-phase flow in deformed multi-scale fracture-porous media. Petroleum Science, 2025, 22(5): 2096-2108.

[11] Zhou, Y., Nie, X. Identification and parameter characterization of pores and fractures in shales based on multi-scale digital core data. Advances in Geo-Energy Research, 2024, 13(2): 146-160.

[12] Xia, C., Shan, J., Li, Y., et al. Three-dimensional modeling of full-diameter micro-nano digital rock core based on CT scanning. Processes, 2026, 14(2): 337.

[13] Li, G., Zhao, J., Liu, Y., et al. Research on multi-scale digital core construction of carbonate rock with fractures and caves characteristics: A case study of dengying formation carbonate rocks in Sichuan Basin. Frontiers in Earth Science, 2025, 13: 1528829.

[14] Li, J., Teng, Q., Zhang, N., et al. Deep learning method of stochastic reconstruction of three-dimensional digital cores from a two-dimensional image. Physical Review E, 2023, 107(5): 055309.

[15] Luo, X., Sun, J., Zhang, R., et al. A multi-condition denoising diffusion probabilistic model controls the reconstruction of 3D digital rocks. Computers & Geosciences, 2024, 184: 105541.

[16] Golsanami, N., Guo, D., Fernando, S. G., et al. Tensile fracture propagation in deep coalbed methane layers under brazilian test: A quantitative analysis using digital rock technology and deep learning-based image segmentation. Rock Mechanics and Rock Engineering, 2025, 58(9): 10613-10641.

[17] Chen, Y., Chen, P., Dai, X., et al. Prediction of rate of penetration in oil and gas well drilling: A hybrid data-driven approach with multi-stage data preprocessing. Engineering Applications of Artificial Intelligence, 2026, 163: 113085.

[18] Maroufi, K., Zahmatkesh, I. Effect of lithological variations on the performance of artificial intelligence techniques for estimating total organic carbon through well logs. Journal of Petroleum Science and Engineering, 2023, 220: 111213.

[19] Davoodi, S., Al-Shargabi, M., Wood, D. A., et al. Advancement of artificial intelligence applications in hydrocarbon well drilling technology: A review. Applied Soft Computing, 2025, 176: 113129.

[20] da Luz, M. A. L., Vasconcelos, M. A. R. Hybrid seismic automation: Integrating artificial intelligence and deterministic methods in velocity analysis. Journal of Applied Geophysics, 2026, 250: 106259.

[21] Zhao, G., Wu, Y., Wei, X. A high-fidelity digital rock representation based on digital grinding combined with deep learning for four-dimensional lattice spring model. International Journal of Rock Mechanics and Mining Sciences, 2025, 186: 106004.

[22] Wu, P., Dong, Z., Liu, S., et al. Microstructure and physical properties of hydrate-bearing sediment: Digital core and deep learning application. Fuel, 2026, 404: 136283.

[23] Neumann, R. F., Barsi-Andreeta, M., Lucas-Oliveira, E., et al. High accuracy capillary network representation in digital rock reveals permeability scaling functions. Scientific Reports, 2021, 11: 11370.

[24] Du, Y., Yan, S., Fu, C., et al. Study on the influence mechanism of the supercritical $CO_2$-$H_2O$-coal interaction on the anisotropy of coal permeability via digital rock technology. Energy, 2025, 343: 139813.

[25] Zhang, H., Zhang, Z., Nie, X., et al. Multiscale digital core technology for analyzing the genesis of low-resistivity oil reservoirs in the guantao formation of bohai sea: Clay mineral-induced additional conductivity effect, pore structure regulation mechanism, and quantitative characterization. Geofluids, 2026, 2026(1): 9975345.

[26] Chi, P., Sun, J., Yan, W., et al. From digital rock to digital wellbore: Multiscale reconstruction and simulation. Advances in Geo-Energy Research, 2024, 13(1): 1-6.

[27] Jiang, F., Guo, Y., Tsuji, T., et al. Upscaling permeability using multiscale X-ray-CT images with digital rock modeling and deep learning techniques. Water Resources Research, 2023, 59(3): e2022WR033267.

[28] Fan, Y., Yan, W., Xing, H., et al. Investigating petrophysical properties of gas hydrate-bearing sediments using digital rock technology: A microscopic perspective. Petroleum Science, 2025, 22(5): 1889-1911.

[29] Wei, X., Qu, D., Zhou, Z., et al. Three-dimensional refined discrete element numerical modeling method and its application for reef limestone based on digital core technology. Computers and Geotechnics, 2025, 185: 107362.

[30] Murugesu, M. P., Krishnan, V., Kovscek, A. R. Enhancing prediction of fluid-saturated fracture characteristics using deep learning super resolution. Applied Computing and Geosciences, 2024, 24: 100208.

[31] You, S., Liao, Q., Yan, Z., et al. Super-resolution reconstruction of 3D digital rocks by deep neural networks. Geoenergy Science and Engineering, 2024, 237: 212781.

[32] Zohdparast, P., Kadkhodaie, A. High resolution estimation shear wave velocity and rock elastic properties using integration of the individual and hybrid machine learning models. Results in Engineering, 2026, 29: 109423.

[33] Lepcha, D. C., Goyal, B., Dogra, A., et al. Image super-resolution: A comprehensive review, recent trends, challenges and applications. Information Fusion, 2023, 91: 230-260.

[34] Su, H., Li, Y., Xu, Y., et al. A review of deep-learning-based super-resolution: From methods to applications. Pattern Recognition, 2025, 157: 110935.

[35] Tang, K., Wang, Y., Regaieg, M., et al. Unlocking sub-micrometer features in carbonate rocks: A cascading super-resolution approach for multiscale multi-instrument carbonate characterization. Water Resources Research, 2025, 61(8): e2025WR040865.

[36] Wang, Y., Armstrong, R. T., Mostaghimi, P. Enhancing resolution of digital rock images with super resolution convolutional neural networks. Journal of Petroleum Science and Engineering, 2019, 182: 106261.

[37] Shan, L., Bai, X., Liu, C., et al. Super-resolution reconstruction of digital rock CT images based on residual attention mechanism. Advances in Geo-Energy Research, 2022, 6(2): 157-168.

[38] Yang, Y., Wang, J., Liu, F., et al. Super-resolution of digital rock images with hybrid attention multi-branch neural network. Gas Science and Engineering, 2024, 128: 205395.

[39] Huo, F., Liu, K., Dong, H., et al. Rock core micrograph image super-resolution reconstruction using a novel deep residual network. Signal, Image and Video Processing, 2025, 19(5): 369.

[40] Liu, T., Xu, C., Tang, L., et al. OmniSR-M: A rock sheet with a multi-branch structure image super-resolution lightweight method. Applied Sciences, 2024, 14(7): 2779.

[41] Zhang, Y., Bi, J., Xu, L., et al. Lightweight multi-scale attention feature distillation network for super-resolution reconstruction of digital rock images. Geoenergy Science and Engineering, 2025, 246: 213628.

[42] Liu, C., Liu, Y., Shan, L., et al. Enhancing unsupervised rock CT image super-resolution with non-local attention. Geoenergy Science and Engineering, 2024, 238: 212912.

[43] Nwankwo, I., Male, F., Karpyn, Z. Patch-based super-resolution generative adversarial network for high-resolution image reconstruction in digital core analysis. Transport in Porous Media, 2026, 153(4): 50.

[44] Liu, M., Mukerji, T. Multiscale fusion of digital rock images based on deep generative adversarial networks. Geophysical Research Letters, 2022, 49(9): e2022GL098342.

[45] Wang, J., Yang, Y., Liu, F., et al. Digital rock super-resolution reconstruction with efficient 3D spatial-

adaptive feature modulation network. Geoenergy Science and Engineering, 2025, 248: 213748.
[46] Xing, Z., Yao, J., Liu, L., et al. Efficiently reconstructing high-quality details of 3D digital rocks with super-resolution transformer. Energy, 2024, 300: 131499.
[47] Niu, Z., Zhu, C., Zhang, W., et al. Physically-accurate super-resolution transformer for 3D digital rock CT reconstruction. Earth Science Informatics, 2026, 19(4): 49.
[48] Gao, T., Zhu, C., Niu, Z., et al. HLF-SASR: A high-low frequency guided structure-aware super-resolution network for digital rock images. Mathematical Geosciences, 2026, https://doi.org/10.1007/s11004-026-10287-9. (in Press)
[49] Yang, Z., Wu, H., Ding, H., et al. Enhanced U-Net model for rock pile segmentation and particle size analysis. Minerals Engineering, 2023, 203: 108352.
[50] Jin, C., Liang, J., Fan, C., et al. Study on segmentation of blasting fragment images from open-pit mine based on U-CARFnet. PLoS One, 2023, 18(9): e0291115.
[51] He, C., Sadeghpour, H., Shi, Y., et al. Mapping distribution of fractures and minerals in rock samples using Res-VGG-UNet and threshold segmentation methods. Computers and Geotechnics, 2024, 175: 106675.
[52] Ren, Y., Zeng, C., Li, X., et al. Intelligent evaluation of sandstone rock structure based on a visual large model. Petroleum Exploration and Development, 2025, 52(2): 548-558.
[53] Zhu, C., Wang, J., Zhang, L., et al. Sam-fusenet: Segment anything guided multimodal fusion for rgb-thermal aerial robotic perception. IEEE Transactions on Geoscience and Remote Sensing, 2025, 64: 1-12.
[54] Ren, Y., Li, X., Bi, J., et al. Multi-channel attention transformer for rock thin-section image segmentation. Journal of Engineering Research, 2025, 13(2): 730-740.
[55] Li, M., Chen, M., Lu, W., et al. RDT-FragNet: A DCN-Transformer network for intelligent rock fragment recognition and particle size distribution acquisition. Computers and Geotechnics, 2025, 177: 106809.
[56] Wang, Y., Qin, R., Wei, D., et al. Super-resolution enhancement and segmentation for digital rock based on multi-task joint deep neural network. Geoenergy Science and Engineering, 2024, 243: 213385.
[57] Liu, F., Yang, Y., Yang, H., et al. Pore structure properties characterization of shale using generative adversarial network: Image augmentation, super-resolution reconstruction, and multi-mineral auto-segmentation. Petroleum Exploration and Development, 2025, 52(5): 1262-1274.
[58] Menke, H. P., Maes, J., Geiger, S. Upscaling the porosity-permeability relationship of a microporous carbonate for darcy-scale flow with machine learning. Scientific Reports, 2021, 11(1): 2625.
[59] Herrera, E., Oms, O., Remacha, E. Porosity and permeability estimations from X-Ray tomography images and data using a deep learning approach. Applied Sciences, 2026, 16(3): 1613.
[60] Tang, Z., Qiu, S., Wang, J., et al. A hybrid model for permeability prediction of porous media based on convit and transfer learning. International Journal of Heat and Mass Transfer, 2026, 256: 128083.
[61] Alqahtani, N. J., Chung, T., Wang, Y., et al. Flow-based characterization of digital rock images using deep learning. SPE Journal, 2021, 26(4): 1800-1811.
[62] Alzahrani, M. K., Shapoval, A., Chen, Z., et al. Pore-GNN: A graph neural network-based framework for predicting flow properties of porous media from micro-CT images. Advances in Geo-Energy Research, 2023, 10(1): 39-55.
[63] Wang, F., Zai, Y. Image segmentation and flow prediction of digital rock with U-Net network. Advances in Water Resources, 2023, 172: 104384.
[64] Xie, C., Zhu, J., Wang, J., et al. Direct prediction of relative permeability curve from 3D digital rock images based on deep learning approaches. International Journal of Rock Mechanics and Mining Sciences, 2023, 170: 105544.
[65] Ali, A., Mahmoud, A. A., Al Ramadan, M., et al. Assessing the influence of olive waste on the filtration properties, elasticity, porosity, and strength of oil well cement. Scientific Reports, 2025, 15(1): 6897.
[66] Wang, Y., Lin, G. Efficient deep learning techniques for multiphase flow simulation in heterogeneous porousc media. Journal of Computational Physics, 2020, 401: 108968.
[67] Balashov, V., Savenkov, E., Khlyupin, A., et al. Two-phase regularized phase-field density gradient navier-stokes based flow model: Tuning for microfluidic and digital core applications. Journal of Computational Physics, 2025, 521: 113554.
[68] Wang, Q., Wang, H., Wang, X., et al. Numerical simulation study on microseepage and displacement mechanism of coal based on digital core technology. Physics of Fluids, 2025, 37(9): 096610.
[69] Asadolahpour, S. R., Jiang, Z., Lewis, H., et al. Deep learning for pore-scale two-phase flow: Modelling drainage in realistic porous media. Petroleum Exploration and Development, 2024, 51(5): 1301-1315.
[70] Chung, J., Marcato, A., Guiltinan, E. J., et al. Accelerating multiphase simulations with denoising diffusion model driven initializations. Journal of Geophysical Research: Machine Learning and Computation, 2024, 1(4): e2024JH000293.
[71] Oommen, V., Shukla, K., Desai, S., et al. Rethinking materials simulations: Blending direct numerical simulations with neural operators. npj Computational Materials, 2024, 10(1): 145.
[72] Santos, J. E., Xu, D., Jo, H., et al. PoreFlow-Net: A 3D convolutional neural network to predict fluid flow through porous media. Advances in Water Resources, 2020, 138: 103539.
[73] Zhou, X., McClure, J. E., Chen, C., et al. Neural network-based pore flow field prediction in porous media using super resolution. Physical Review Fluids, 2022, 7(7): 074302.
[74] Zhao, J., Wu, J., Wang, H., et al. Single phase flow simulation in porous media by physical-informed Unet network based on lattice Boltzmann method. Journal of

Hydrology, 2024, 639: 131501.

[75] Munarin, F., Gouze, P., Nepomuceno Filho, F. Two-phase flow dynamics in 3D fractures: Influence of aperture, wettability, and fluid properties from lattice boltzmann simulations. Advances in Water Resources, 2025, 206: 105133.

[76] Li, X., Peng, C., Zhang, C., et al. Comparative study and critical assessment of phase-field lattice boltzmann models for laminar and turbulent two-phase flow simulations. ArXiv Preprint ArXiv: 2511.11360, 2025.

[77] Wang, Y., Chung, T., Armstrong, R. T., et al. ML-LBM: Predicting and accelerating steady state flow simulation in porous media with convolutional neural networks. Transport in Porous Media, 2021, 138(1): 49-75.

[78] Kashefi, A., Mukerji, T. Prediction of fluid flow in porous media by sparse observations and physics-informed pointnet. Neural Networks, 2023, 167: 80-91.

[79] Kamrava, S., Sahimi, M., Tahmasebi, P. Simulating fluid flow in complex porous materials by integrating the governing equations with deep-layered machines. npj Computational Materials, 2021, 7(1): 127.