Petroleum Science 23 (2026) 3682–3697

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

TransUNet framework improved computed tomography image segmentation for core pore evolution

Shao-Hua Zhou [a,b], Tian-Bao Liu [a], Yu Zhao [a], Shao-Hao Yin [b], Zhi-Yu Liu [b], Ji-Jun Liu [c], Ling-Wei Du [d], Yue-Tong Zhao [d], Wei-Guang Shi [a,b,e,*]

[a] *College of Chemistry and Chemical Engineering, Northeast Petroleum University, Daqing, 163318, Heilongjiang, China*
[b] *Department of Mathematics and Statistics, Northeast Petroleum University, Daqing, 163318, Heilongjiang, China*
[c] *Daqing Oilfield No.5 Production Plant, Daqing, 163002, Heilongjiang, China*
[d] *School of Earth Sciences, Northeast Petroleum University, Daqing, 163318, Heilongjiang, China*
[e] *National Key Laboratory of Continental Shale Oil, Institute of Unconventional Oil & Gas, Northeast Petroleum University, Daqing, 163318, Heilongjiang, China*

A R T I C L E I N F O

Article history:
Received 20 October 2025
Received in revised form
11 February 2026
Accepted 7 April 2026
Available online 10 April 2026

Edited by Min Li

Keywords:
Synchrotron radiation CT
Image segmentation
TransUNet
Argillaceous microporous network
Alkaline flooding porosity correction

A B S T R A C T

The development of oil and gas is constrained by difficulties in dynamically characterizing pore structures. Traditional methods inadequately represent the complex interactions between mineral dissolution, precipitation, and fluid flow. This study addresses these gaps by introducing a Transformer U-Neural Network (TransUNet) for computed tomography (CT) image segmentation. The integrated workflow combines conventional CT (Resolution of 5.4 μm) and synchrotron radiation CT (Resolution of 0.8 μm) for dynamic flooding, imaging, segmentation, and precise 3D pore network extraction, overcoming resolution limits. TransUNet's strong global attention and feature extraction reduce overfitting and deliver high-accuracy segmentation of minerals, pores, and argillaceous microporous networks (AMN), achieving 74.92% intersection over union (IoU) for AMN. A porosity correction method improves conventional CT porosity accuracy to 94% of gas-measured values. Alkaline flooding experiments reveal: (1) initial clay swelling reduces small pore size by ~50% as alkaline ions destabilize clay; (2) mineral dissolution, such as dolomite, creates secondary pores, increasing 80 μm pores by 1.8 times; (3) silicate dissolution increases porosity and leads to a 93.7% rise in permeability. Clay reorganization enhances the AMN by 46.1%. The pore size distribution shifts to log-normal at steady state, and throat connectivity improves flow capacity. This work pioneers Transformer-based CT image segmentation, introduces cross-resolution prediction, and clarifies pore regulation by mineral phase changes, establishing a new paradigm for chemical flooding in sandstone reservoirs.

1. Introduction

The efficient development of oil and gas resources is fundamentally constrained by the precise characterization of reservoir pore structures and their dynamic evolution mechanisms (Lai et al., 2023; Xu et al., 2022). Different flooding technologies affect pore structures through distinct mechanisms (Wang et al., 2024a). For example, supercritical CO_2 reduces oil viscosity and promotes mineral dissolution, and its miscible flooding enhances oil mobilization in large pores; however, in low-permeability reservoirs, migration of fine particles generated by mineral dissolution may obstruct pores and throats, leading to permeability reduction (Cao et al., 2024). Similarly, water flooding can enlarge pores and throats in high-permeability reservoirs, while in low-permeability ones, it often induces clay swelling and fine migration, exacerbating throat blockage (Wang et al., 2025). Chemical flooding techniques, such as ternary composite flooding, can increase pore-throat radius and coordination number, but silicate dissolution and clay recombination may increase the proportion of micropores (Liang et al., 2021). Despite extensive research, existing experimental and analytical approaches remain inadequate for capturing dynamic pore evolution and the competitive processes mineral dissolution and precipitation during flooding (Wang et al., 2024b). Numerical simulations typically

* Corresponding author.
E-mail address: shiweiguang@nepu.edu.cn (W.-G. Shi).

https://doi.org/10.1016/j.petsci.2026.04.008
1995-8226/

rely on idealized assumptions and struggle to represent strong reservoir heterogeneity, resulting in persistent knowledge gaps (Yang et al., 2025). In particular, the quantitative impact of mineral dissolution-precipitation competition and the dynamic reaction pathways between flooding fluids and minerals, and the mechanisms by which clay hydration and swelling regulate pore space and fluid migration remain poorly understood.

Computed tomography (CT) imaging provides a powerful non-destructive approach for three-dimensional visualization of porous media (Hu et al., 2022). Through digital core analysis, CT enables quantitative tracking of pore network evolution at different flooding stages. However, pores in unconventional reservoirs exhibit pronounced multi-scale heterogeneity (Zhang et al., 2024a). The resolution of conventional CT (>5 μm) is insufficient to accurately identify sub-micron pores and mineral dissolution boundaries, and transient structural changes during dynamic flooding are often missed due to spatiotemporal resolution constraints (Zhao et al., 2023). These limitations hinder a deeper understanding of pore evolution mechanisms. Synchrotron radiation CT presents a transformative solution, achieving resolutions down to 500 nm and enabling precise characterization of micro-pore structures, mineral distributions, and dissolution features. It has become an effective analytical core tool for three-dimensional pore networks (Zhao et al., 2020). However, this technique introduces new challenges: the extraction of meaningful features from massive CT datasets has become a critical research frontier. Furthermore, current applications focus predominantly on static structural analysis rather than in-situ dynamic processes (Liu et al., 2023a). Real-time monitoring of transient pore evolution and mineral phase changes during dynamic flooding therefore remains technically challenging, making it even impossible to further explore its mechanisms.

Recent advances in deep learning, particularly U-neural network (U-Net), have demonstrated strong potential for CT image segmentation (Wang and Chen, 2021), as these models enable effective integration and analysis of multi-scale and multi-resolution CT images. Unlike traditional image processing methods, neural networks automatically learn hierarchical features from data through convolutional operations and nonlinear activation functions, and adapt well to low contrast and complex textures in CT images (Gu et al., 2023). The U-Net architecture adopts an encoder-decoder structure, in which the encoder extracts high-level semantic features through down-sampling, while the decoder restores spatial resolution through up-sampling. Skip connections can bridge corresponding encoder and decoder layers, facilitating the fusion of spatial details with high-level semantic information and enabling end-to-end, pixel-level segmentation (Song et al., 2022). For instance, Res-U-Net achieved an Intersection over Union (IoU) of 62.13% in fracture detection (Kalinaki et al., 2023). Despite these advantages, convolutional neural network (CNN)-based models inherently suffer from limited receptive fields, which restrict their ability to capture long-range spatial dependencies and complex multi-scale mineral-pore interactions. Moreover, most existing CNN-based frameworks are not explicitly designed for multi-target segmentation or dynamic spatiotemporal analysis. To address these limitations, recent studies have introduced attention mechanisms, multi-scale feature fusion strategies, and optimized loss functions. Among these approaches, Transformer U-Neural Network (TransUNet) has emerged as a representative improved model that combines the local feature extraction of CNNs with the global context modeling of Transformers (Chen et al., 2024). Skip connections enable efficient fusion of details and semantics, which significantly improves boundary segmentation accuracy for complex targets. Nevertheless, applying such advanced models to the analysis of dynamic pore evolution during core flooding remains challenging, and segmentation frameworks that give real-time performance with high accuracy are still urgently needed to advance the understanding of pore evolution mechanisms.

To address these challenges, this study aims to overcome the limitations of conventional experimental and analytical approaches in characterizing dynamic pore processes and cross-scale mineral reactions. We propose an integrated framework that combines multi-scale CT imaging with deep learning, establishing a comprehensive "dynamic flooding-imaging-segmentation -analysis" workflow. Within this framework, a TransUNet architecture is employed to explicitly model long-range spatial dependencies in synchrotron radiation CT images, while a cross-resolution prediction strategy is developed to synergistically utilize both high- and low-resolution CT data. In addition, a porosity correction formula applicable to the sandstone cores in this study is introduced to compensate for resolution-induced errors, thereby alleviating the submicron resolution limitations inherent to conventional CT imaging. The proposed TransUNet-based segmentation model achieves high-precision identification of minerals, pores, and clay-related microporous networks from low-resolution CT images, with an average IoU of 0.82 on the validation dataset. Experimentally, core samples from the Daqing Oilfield are subjected to alkaline flooding, and their pore structure evolution is monitored through the combined use of conventional CT and synchrotron radiation CT. This multi-scale observation system enables systematic elucidation of the dynamic pore balance mechanism driven by mineral phase transformations during alkaline flooding and provides insights into mitigating the risks associated with secondary precipitation. Ultimately, this research transcends the limitations of static idealization and dynamic black-box modeling, offering a systematic methodological framework and practical guidance for enhancing the development efficiency of oil and gas resources.

2. Experiment and image segmentation method

2.1. Alkali flooding experimental materials and equipment

Experimental materials: the core samples were obtained from the Sanan No. 5 block of the Daqing Oilfield (No.2 Production Plant), located within the central depression of the Songliao Basin on the southern flank of the Saertu Anticline. The formation dip averages 18.5° on the western flank and 3.7° on the eastern flank. The target oil layer occurs at a burial depth of approximately 1062 m, with an in-situ temperature of 49.4 °C. The original formation pressure is 10.99 MPa, the saturation pressure is 7.96 MPa, and the pressure coefficient is 1.04. The reservoir consists of grayish-green sandstone and mudstone interbeds of the Lower Cretaceous, representing inner and outer delta-front facies of a fluvial-delta system. Lithology is dominated by fine sandstone and siltstone, with fine sand content of 42.4%, silt content of 33.8%, and a median grain size of 0.124 mm. Cementation is primarily loose and argillaceous, with a clay content of 10.7%. In terms of mineral composition, the oil layer contains 25%–35% quartz, 30%–40% feldspar, 20%–40% rock fragments, and 5%–12% clay minerals. Within the clay fraction, illite is dominant (58%–67%), followed by kaolinite (5%–28%), with illite-smectite and chlorite-smectite mixed-layer minerals accounting for 7%–18%. And the core parameters are shown in Table 1.

The experimental system for multi-scale core analysis under realistic reservoir conditions integrated multiple advanced characterization tools, as detailed in Section 2.4. Reservoir fluid injection was simulated using a ZJHW-1 multi-function high-temperature and high-pressure flooding apparatus (Fig. S1). Pore

Table 1
Core parameter.

Parameter	Value	Parameter	Value
Core length, cm	1	Core diameter, cm	0.4
Porosity, %	18.2	Permeability, mD	28.7
NaOH concentration, wt%	2.0	Injection velocity, mL·h^{-1}	3

structure was characterized using both conventional CT for 3D analysis of rock structure and connectivity, and synchrotron radiation CT (SR-CT) (at the SSRF-BL13HB beamline). The latter technique, leveraging its high brightness and coherence for nanoscale resolution, was employed to dynamically monitor pore-scale fluid transport and mineral reactions. The experimental workflow is illustrated in Fig. S2, with detailed parameters for the characterization techniques provided in Sections 2.5 and 2.6.

2.2. Permeability measurement

Permeability was measured using a Gasperm gas permeameter, which operates under constant temperature and pressure conditions based on the steady-state flow method. The measurement range of the instrument spans from 0.01 mD to 10 D (Kudasik et al., 2023). Gas permeability (K, mD) was calculated according to Darcy's law:

$$K = \frac{2Q_0 P_0 \mu L}{A\left(P_1^2 - P_2^2\right)} \tag{1}$$

where Q_0 is the volumetric flow rate of the gas at atmospheric pressure P_0 (cm^3/s), μ is the gas viscosity at the measurement temperature (mPa·s), L and A are the length and cross-sectional area of the core sample (cm, cm^2), respectively, and P_1 and P_2 are the inlet and outlet pressures (MPa), respectively. The reported K is the absolute (intrinsic) permeability, a property of the rock itself independent of the fluid.

Experimental procedure: the core samples were cleaned, dried, and cooled to room temperature prior to measurement. Their dimensions were precisely measured to calculate the cross-sectional area. Each sample was then placed in a core holder, and a confining pressure was applied to ensure a proper seal. A system leak check was performed before testing. During measurement, nitrogen (or air) was steadily introduced at the core inlet. The system was allowed to stabilize until steady-state flow was achieved. The inlet pressure (P_1), outlet pressure (P_2, typically atmospheric), and the stabilized gas flow rate (Q_0) were recorded. For low-permeability samples, a soap film flowmeter was used to enhance measurement accuracy. The absolute permeability (K) was finally calculated by substituting all measured parameters into Eq. (1).

2.3. Porosity measurement

Core porosity was determined using the helium expansion method, a classic technique known for its high accuracy (Lyu et al., 2021). The porosity (ϕ) is calculated as:

$$\phi = \left(1 - V_g/V_b\right) \times 100\% \tag{2}$$

where V_b is the bulk volume of the core sample, and V_g is the grain volume.

Prior to measurement, core samples were processed into regular cylinders with ground, flat end faces. They were then washed with oil to remove residual fluids and dried at 105 °C until a constant weight was achieved. The dried samples were cooled and stored in a desiccator for subsequent use. The bulk volume (V_b) was calculated from precise diameter and length measurements obtained using a vernier caliper. The measurement was performed using a helium porosimeter. The system was connected, and the helium source pressure was adjusted to 0.6–0.7 MPa. A system leak check was conducted to ensure integrity. A pressure-volume standard curve was first established using calibrated steel blocks. The dried core sample was then placed into the core holder, and any remaining void space was filled with additional steel blocks if necessary. After the system reached equilibrium, the stabilized pressure was recorded. The corresponding total volume of the sample and filler blocks was determined from the pre-established standard curve. The grain volume of the core (V_g) was obtained by subtracting the known volume of the filler blocks from this total volume. Finally, the porosity (ϕ) was calculated using Eq. (2).

2.4. Experimental steps for alkali flooding

Alkaline flooding was performed using a 2.0 wt% NaOH solution. Strong alkali (NaOH) was selected to establish the high-pH environment requisite for the process, with the specific concentration optimized through preliminary tests to balance effective displacement performance against potential core damage (Taheri et al., 2025). The solution was prepared with deionized water, and its initial pH was measured using a pH meter calibrated with standard buffer solutions. Core samples, characterized by their pre-measured porosity and permeability, were first subjected to CT to acquire pre-flooding baseline images. The flooding experiment was then conducted at 45 °C to simulate reservoir conditions, using a high-temperature, high-pressure core flooding apparatus. During the injection, the inlet-outlet pressure difference was monitored in real-time, and the pH of the effluent was continuously recorded. Following the flooding experiment, the core's porosity and permeability were re-measured. Post-flooding structural alterations and fluid distribution were analyzed through comparative CT imaging, utilizing both conventional and high-resolution synchrotron radiation CT. The resulting images are correspondingly labeled as the pre-flooding, during-flooding, and post-flooding states. The experimental setup and detailed workflow are schematically illustrated in Fig. S1 and S2.

2.5. Conventional CT tests

The core samples subjected to strong alkaline flooding were prepared through a systematic pretreatment process. First, they were dried at 80 °C for 24 h to remove moisture. Subsequently, vacuum saturation was performed using simulated formation water. Pore volume was determined via gravimetric method using the formula: (saturated weight − dry weight)/fluid density. CT scanning was conducted using an X-ray digital core analysis system (GE, Germany). The scanning parameters were configured as follows: acceleration voltage 120 kV, current 80 μA, single projection exposure time 0.65 s, and voxel size (resolution) 5.4 μm (Lin et al., 2025; Zhang et al., 2024). Prior to scanning, the X-ray source and detector were calibrated to accommodate the sample dimensions. The entire length of the core was scanned to ensure full coverage and high image quality. A standard reconstruction algorithm (STD) was applied, resulting in a three-dimensional data volume with dimensions of 1700 × 1700 × 1440 voxels.

2.6. Synchrotron radiation CT (SR-CT) test

The experiment was conducted at the BL13HB X-ray Imagin Application Beamline of the Shanghai Synchrotron Radiation Facility (SSRF) (Ji et al., 2023). This beamline provides a spatial resolution of 0.8 μm and a temporal resolution of 0.5 ms for CT imaging. Prior to sample placement, X-ray fluctuation correction and background normalization were performed. This process involved acquiring multiple flat-field images (captured under X-ray illumination without the sample in the field of view) and dark-field images (captured without X-ray exposure to assess detector background). A cylindrical sample was securely fixed at the center of the rotation stage, with its long axis positioned vertically and aligned to the center of the field of view. The sample was then rotated through a total angle of 180° in increments of 0.1°. Imaging was performed using a monochromatic X-ray energy of 18 keV (corresponding to a wavelength of 0.689 Å). The resulting visible light was projected onto a CCD detector via a 4× objective lens. With an effective pixel size of 6.5 μm and a 4× magnification, each original projection generated a two-dimensional X-ray attenuation map with a resolution of 2048 pixels. The final reconstructed CT data volume had dimensions of 2048 × 2048 × 1780 voxels.

2.7. Image segmentation model based on the TransUNet

The image segmentation in this study was performed using a model based on the TransUNet architecture (Chen et al., 2024). This hybrid design integrates the local feature extraction capability of CNN with the global contextual modeling strength of the Transformer, enhanced by skip connections that preserve detailed spatial information for precise pixel-level localization (Xu and Zhang, 2025; Hoyer et al., 2024). This combination makes it particularly effective for complex segmentation tasks such as analyzing core sample images.

The model architecture, illustrated in Fig. 1, adopts a symmetric encoder-decoder structure. The encoder first utilizes a CNN backbone (e.g., ResNet-50) with 3 × 3 convolutional kernels and a stride of 2 to perform four levels of down sampling, extracting hierarchical features. The highest-resolution feature map from the final CNN stage is then divided into 16 × 16 patches. These patches are linearly projected into a 768-dimensional embedding space, augmented with positional encodings, and processed by a Transformer module comprising 12 layers, each with 12 self-attention heads. The decoder reconstructs the segmentation map through successive upsampling stages using transposed convolution (3 × 3 kernels, stride 2). It incorporates enhanced skip connections from the encoder, gated by squeeze-and-excitation (SE) attention blocks, to effectively fuse multi-scale features. A final 1 × 1 convolutional layer generates the pixel-wise output. Additional architectural details are provided in Figs. S3 and S4. The selection of this model is supported by its demonstrated performance. On the synapse multi-organ segmentation dataset, it achieves a dice coefficient of 82.3%, outperforming a conventional U-Net by 3.2 percentage points (Fig. S5). This performance gain is attributed to the Transformer's expansive global receptive field, the hybrid encoding strategy that leverages both local and global features, and the refined feature fusion mechanism facilitated by the enhanced skip connections.

3. Results and discussion

3.1. SR-CT images and low-resolution CT images of cores

The grayscale response of minerals in core CT imaging originates from the material-dependent X-ray absorption characteristics, which are influenced by density and atomic number (Hoyer et al., 2024; Yeh et al., 2017). Minerals with high density and high atomic number-such as pyrite and dolomite—exhibit strong X-ray attenuation and thus appear bright white in CT images. In contrast, mid-to low-density silicates (e.g., quartz and plagioclase, with densities around 2.65–2.76 g·cm^{-3}) demonstrate moderate attenuation and are depicted in light to medium gray. Clay

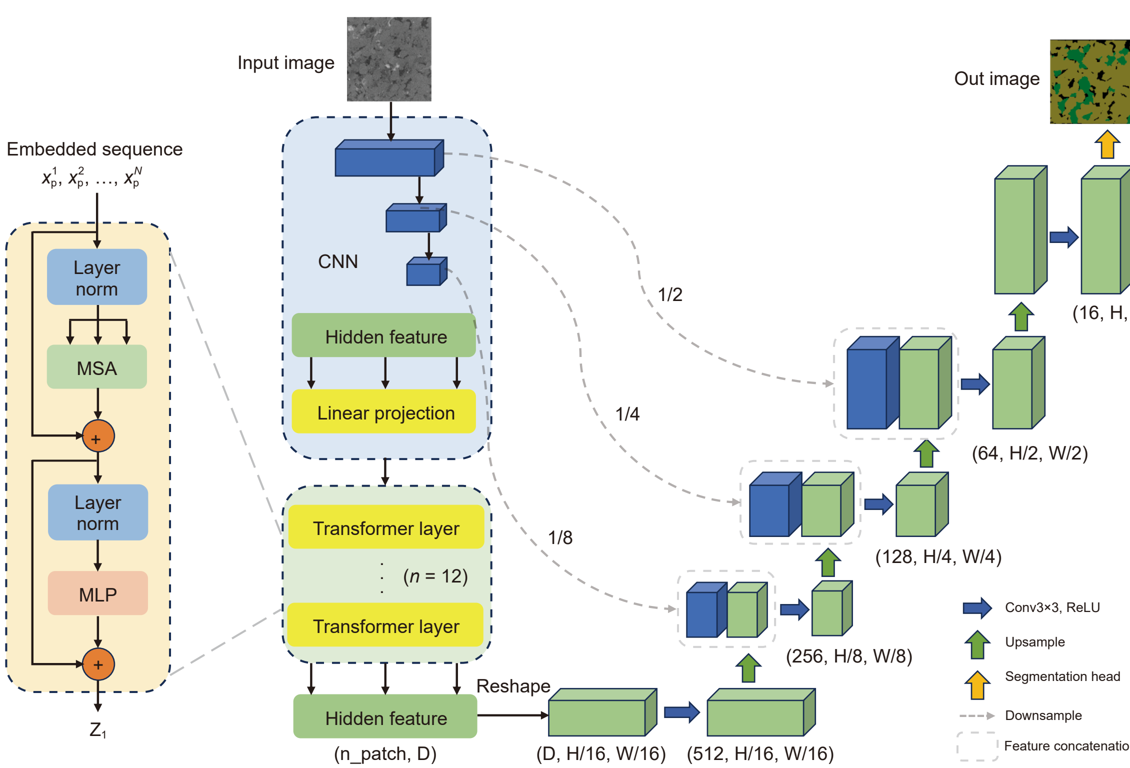

Fig. 1. Schematic diagram of the TransUNet architecture. The network comprises an input image, an encoder (composed of CNN and Transformer components), a decoder, and an output layer.

minerals, often containing microporosity, have even lower effective density (<2.6 g·cm^{-3}) and thus appear dark gray to nearly black. Pores and fractures, with density close to zero, are clearly delineated as pure black regions against the mineral matrix. Despite the widespread use of conventional CT systems with resolutions around 5.4 μm (as shown in Fig. 2(a)), inherent limitations become apparent under higher magnification. For example, the edges of major minerals such as quartz appear blurred with indistinguishable corners (Fig. 2(b) and (c)). Even the boundaries between high-attenuation minerals and quartz are noticeably diffuse, which hinders accurate quantification of chemical interactions and fluid transport phenomena along interfaces.

Additionally, the representation of pores systematically deviates from ideal models: instead of uniform black areas, many pore regions display nebulous gray signals. This inaccuracy stems from two factors: the inability to resolve submicron pores due to limited resolution, and the presence of low-density secondary minerals (e.g., clays) within pores, which alter local attenuation properties. It should be noted that limitations in spatial resolution give rise to two principal artifacts. First, blurred interfaces between mineral phases hamper the accurate delineation of dissolution–precipitation fronts and the quantification of fluid flow pathways. Second, the ambiguity of grayscale values within pore spaces complicates the reliable segmentation of pores from fine-grained or low-density minerals. This ambiguity introduces errors in porosity calculations and permeability predictions. Therefore, although conventional CT at 5.4 μm resolution remains a widely used tool in core analysis, it suffers from fundamental constraints in delineating mineral boundaries and characterizing micropores (Vásárhelyi et al., 2020). To accurately investigate mineral interactions and construct reliable pore-network models, it is necessary to apply high-resolution CT techniques with resolutions below 1 μm, complemented by multi-scale validation.

Conventional CT has a resolution of about 5.4 μm. This resolution poses inherent limitations for characterizing mineral interfaces and identifying micropores. Therefore, we employed SR-CT, which achieves a resolution of up to 0.8 μm. This technique performed high-precision three-dimensional imaging of the core samples. At this higher resolution, SR-CT clearly revealed the spatial distribution of high-attenuation metallic minerals, such as dolomite. It also revealed mid-to-low grayscale silicate minerals, like plagioclase and quartz, which form the bedrock matrix and appear as dark gray regions. More importantly, the technique significantly enhanced the sharpness of mineral boundaries. The outlines of bedrock and mineral interfaces became distinctly identifiable, contrasting sharply with the blurred boundaries observed in conventional CT images. Further comparison demonstrated that nano-CT (with resolution <1 μm) also captured fine mineral details. Conventional CT images showed blurry areas where clay minerals and pores overlapped. SR-CT clearly resolved these areas as well-defined pore networks. The distribution of these pores closely matched the spatial arrangement of the clay minerals. Given the complex interweaving of clay minerals and micropores at the submicron scale, these features were

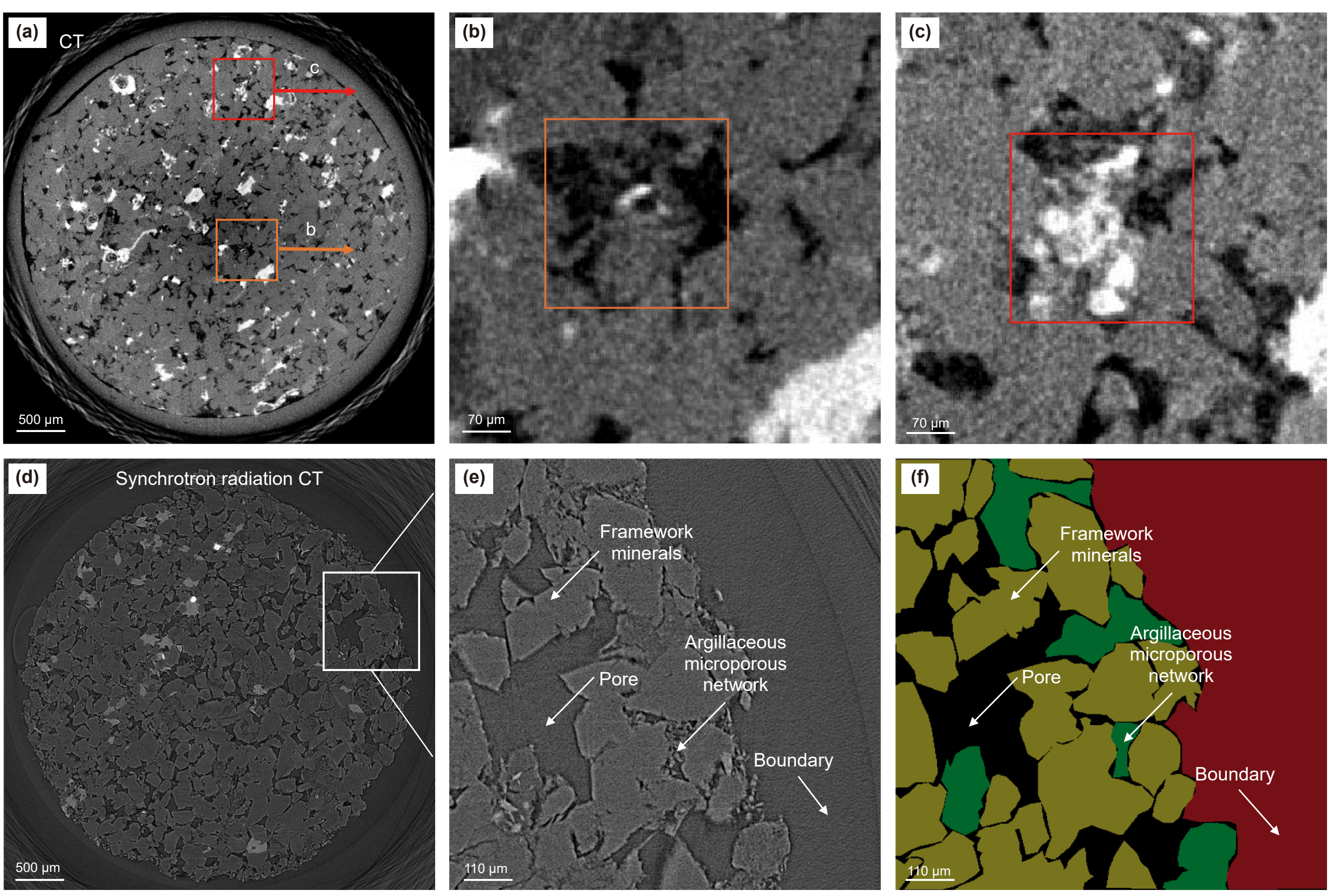

Fig. 2. CT image sequence and annotation results of the core sample. (**a**) CT image of the core sample; (**b**, **c**) Magnified views of local regions in (a), showing microscopic details such as pore structures; (**d**) Synchrotron radiation computed tomography (SR-CT) image of the same core sample, demonstrating significantly higher resolution than conventional CT; (**e**) Magnified view of a selected area in (d), revealing clearer microstructural features; (**f**) Manually annotated result (label image) corresponding to the region in (e), used for training deep learning-based image segmentation models.

collectively categorized as a mixed network. The core images were manually annotated into four components: boundary (red), framework minerals (yellow), pore (black), and argillaceous microporous network (green, AMN) (Fig. 2(e) and (f); Table 2 and Table S1). The high-resolution and boundary-sharpening capability of SR-CT, as confirmed by our analysis, enabled the accurate delineation of topological relationships between clay minerals and micropores (e.g., connectivity and spatial coupling). These high-fidelity annotated images served as the training dataset for the TransUNet model. This approach ensured an accurate representation of microstructures and established a reliable basis for quantifying pore-mineral interaction mechanisms.

3.2. The TransUNet model for segmentation of SR-CT core images

Fig. 3 illustrates the image segmentation workflow. Core flooding experiments provided CT images, with SR-CT (2048 × 2048 resolution, ~1780 slices/core) and routine CT (1700 × 1700 resolution, ~1440 slices/core) employed for high- and standard-resolution imaging, respectively. Ten representative high-resolution images were selected for model development. To address the challenges of processing large images and manual annotation, each original image was subdivided into 16 uniform patches, yielding 160 total patches for preprocessing. Preprocessing steps included format conversion, cropping, manual annotation, data augmentation, and masking, after which the data were partitioned into training, validation, and test sets. An advanced deep learning model (TransUNet, incorporating an attention mechanism) was trained and evaluated on this dataset. The final trained model was then applied to segment the full set of CT images, enabling subsequent quantitative analysis of rock pore structure. The remaining images (beyond the ten used for training) were utilized for segmentation prediction and core-scale quantitative analysis.

These sub-images were then pixel-level annotated under expert guidance into four categories, a process critical for the accuracy of subsequent supervised learning. To improve the model's generalization capability, the annotated dataset was expanded using data augmentation techniques. These techniques included translation, scaling, rotation, and elastic transformation. After augmentation, the total number of images reached 800. The preprocessed images subsequently underwent mask processing to generate label matrices that were strictly spatially aligned with the original data, resulting in a structured dataset suitable for deep learning. The dataset was divided into training and validation sets using an 8:2 ratio. The detailed data processing procedure is provided in the supporting information, specifically in Sections S3.1 to S3.2 and Figs. S6–S12. Regarding model selection, the traditional U-Net was deemed insufficient for modeling long-range spatial relationships. Consequently, the TransUNet architecture, which incorporates a Transformer self-attention mechanism, was adopted (configuration details are provided in Table 1). Its key advantages include: (1) multi-scale feature fusion between the encoder and decoder via n_skip = 3, and (2) the ability of Transformer blocks with vit_patches_size = 16 to effectively capture global spatial dependencies. During training, the model accepted input slices of size 512 × 512. Training proceeded for 100 epochs using the RMSprop optimizer with a learning rate of 0.001 and a batch size of 5 to ensure efficient GPU memory utilization. To address class imbalance, a weighted categorical cross-entropy loss function was employed (category_weight = 4 was set to emphasize the recognition of pores and argillaceous micro-porosity networks).

As shown in Fig. 4, both the training and validation loss curves converge rapidly to approximately 0.1, with a corresponding error rate below 10%. The model achieves a MIoU of 0.95 on the training set and 0.82 on the validation set. This dual convergence of metrics confirms that the model effectively learned distinguishing features of minerals and pores without overfitting, as evidenced by its

Table 2
Training parameters for the rock CT image segmentation model based on the TransUNet.

Parameter	Training parameters	Parameter	Training parameters
Train_rate, %	80	Learning rate	10^{-4}
Test_rate, %	20	Batch size, number/time	5
Category weight, weight coefficient	4	n_skip, number of pieces/layers	3
Loss	Categorical_ crossentropy	Epochs, rounds	100
Evaluating indicator	PA, MPA, CPA, IoU, MIoU	Vit_patches_size, pixels	16
Optimizer	RMSprop	Img_size, pixels	512

Note: PA (Pixel Accuracy), MPA (Mean Pixel Accuracy), CPA (Class Pixel Accuracy), IoU (Intersection over Union), MIoU (Mean Intersection over Union).

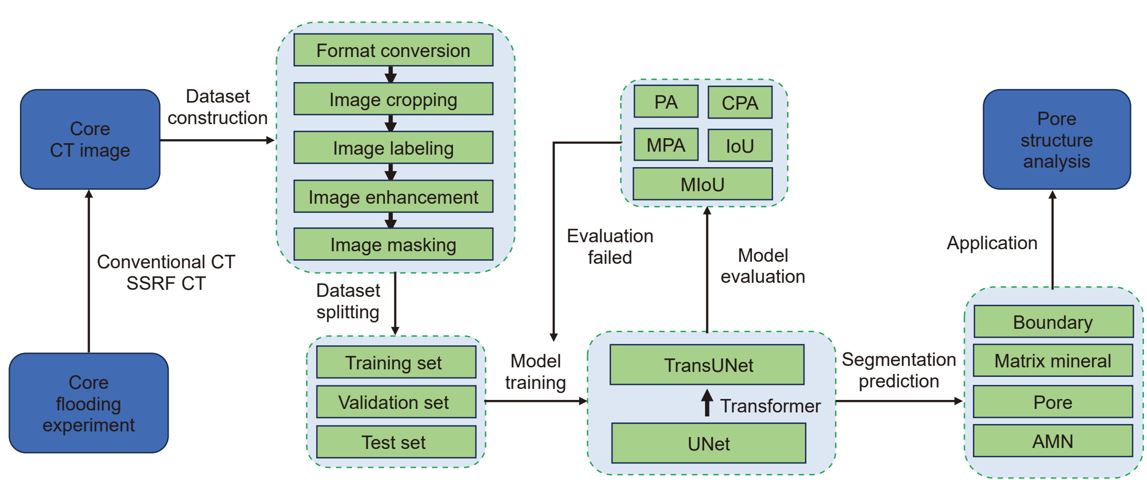

Fig. 3. Workflow of core image segmentation and pore structure analysis. 1. Core flooding experiment → 2. CT scanning → 3. Image preprocessing → 4. Dataset division → 5. U-Net/TransUNet model training → 6. Model evaluation → 7. Pore segmentation prediction → 8. Quantitative pore structure analysis. The parameters for model evaluation are listed in Table 2.

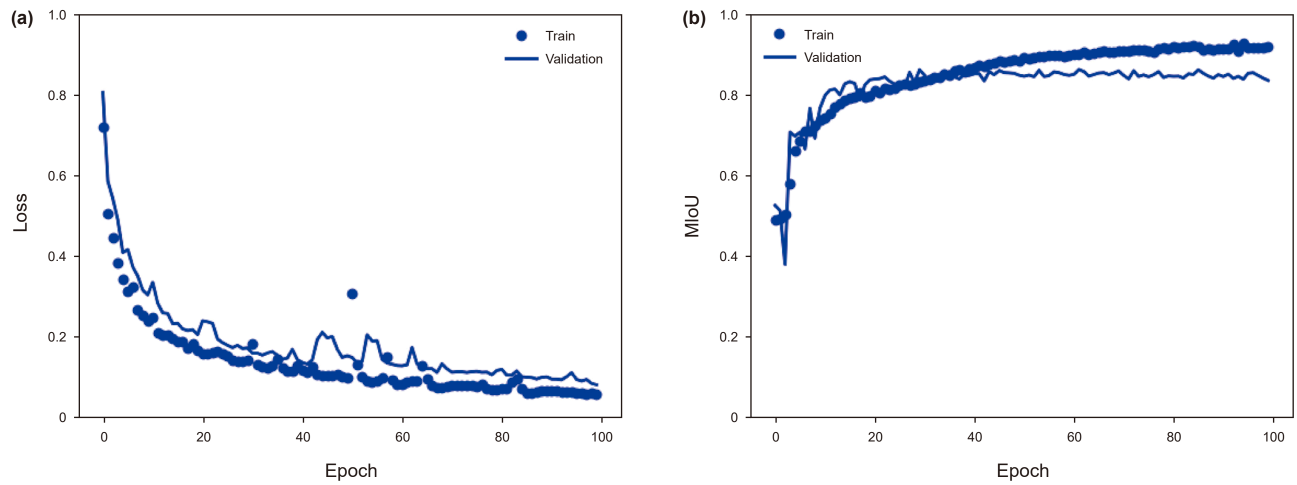

Fig. 4. Performance curves of the TransUNet model: (**a**) the training and validation loss across epochs, reflecting the model's convergence behavior; (**b**) the MIoU during training, demonstrating the segmentation accuracy improvement over iterations.

consistent performance on new, unseen core samples. Segmentation performance was rigorously quantified on the validation set using a five-metric system: Pixel Accuracy (PA) for overall quality; Class Pixel Accuracy (CPA) and Mean PA (MPA) for per-category (e.g., pore, mineral) accuracy; and IoU with its mean (MIoU) for boundary alignment precision. As summarized in Table 3, the TransUNet model attained a PA of 0.8720 and an MIoU of 0.8208.

The TransUNet model demonstrated superior performance, particularly in segmenting challenging components. Most notably, it achieved a significant 6% increase in recognition accuracy for the AMN category (0.7634 vs. U-Net's 0.7191), with its IoU for AMN also substantially higher (0.7492 vs. 0.6983). While maintaining consistently high CPA values (steadily above 0.82) for framework minerals and pores—performance on par with U-Net—the key advancement lies in its accurate delineation of complex clay-micro pore composite structures, which are often misidentified by conventional methods. This capability is attributed to the boundary-focused mechanism inherent in the TransUNet architecture. The resulting high-precision segmentation provides a reliable structural foundation for subsequent pore network modeling and fluid flow analysis.

A comparative analysis of CT images (Fig. 5(a)–(c) series) visually demonstrates the advancements offered by the TransUNet model. The U-Net model accurately segments framework minerals and macro-pores, as shown in Fig. 5(b1)–(b3). However, its receptive field is limited. This limitation is inherent to convolutional operations. It leads to misclassification within the AMN. This problem is particularly evident in the areas highlighted by the boxes. In contrast, TransUNet leverages the global contextual modeling capability of Transformer's self-attention mechanism. As shown in Fig. 5(c1)–(c3), it maintains high segmentation accuracy for minerals and pores while more precisely reconstructing the topological structure of the AMN, thereby correcting missegmented regions from U-Net. This results in a 7% improvement in IoU for the AMN, with segmentation results aligning closely with manual annotations. The successful implementation of TransUNet enables accurate submicron-level segmentation of four components: boundaries, framework minerals, pores, and the AMN.

By integrating the Transformer's global perception with U-Net's local feature extraction, TransUNet overcomes the limitations of conventional methods in characterizing complex pore-mineral microstructures. This is evidenced by its stable training convergence (Fig. 4), superior quantitative metrics (Table 3), and enhanced visual delineation (Fig. 5). Notably, while maintaining high accuracy for framework minerals and pores, the model achieves a significant 6% increase in CPA and a 7% rise in IoU for the AMN. These technical advancements translate into two key application values. First, the precise depiction of AMN connectivity (mean MIoU >0.82, Fig. S12) avoids the permeability overestimation common in traditional methods that confuse micropores with clay minerals, thereby providing a more reliable structural basis for seepage simulation. Second, the sharply defined mineral boundaries enable the quantification of mineral-fluid reaction interfaces, which supports the kinetic modeling of dissolution-precipitation processes. Thus, TransUNet establishes a new paradigm for core microstructure analysis and emerges as a foundational tool for investigating seepage mechanisms and predicting productivity in reservoirs.

Table 3
Quantitative comparison of U-Net and TransUNet on the core image segmentation dataset.

Parameter	U-Net	TransUNet
Overall PA value	0.8416 (Pixel Accuracy)	0.8720
CPA value for each category	[0.9043,0.8536,0.8135, 0.7191]	[0.9157,0.8614, 0.8266,0.7634]
Average MPA by class	0.8227 (Average Pixel Accuracy)	0.8417
IoU value for each category	[0.8731,0.8324,0.7618, 0.6983]	[0.8956,0.8533, 0.7854,0.7492]
MIoU by class average	0.7914 (Average intersection ratio)	0.8208

Note: The four categories are that boundary, skeleton minerals, pores, and argillaceous microporous networks.

3.3. Low-resolution CT image segmentation

Following the successful development of the TransUNet segmentation model, we proceeded to apply it to the segmentation of low-resolution core CT images (Fig. 6(c)). For a discussion on the validation of the model, which was trained on high-resolution images and applied to low-resolution image predictions (Section S3.3 and Fig. S14). We conducted a comparative analysis of the

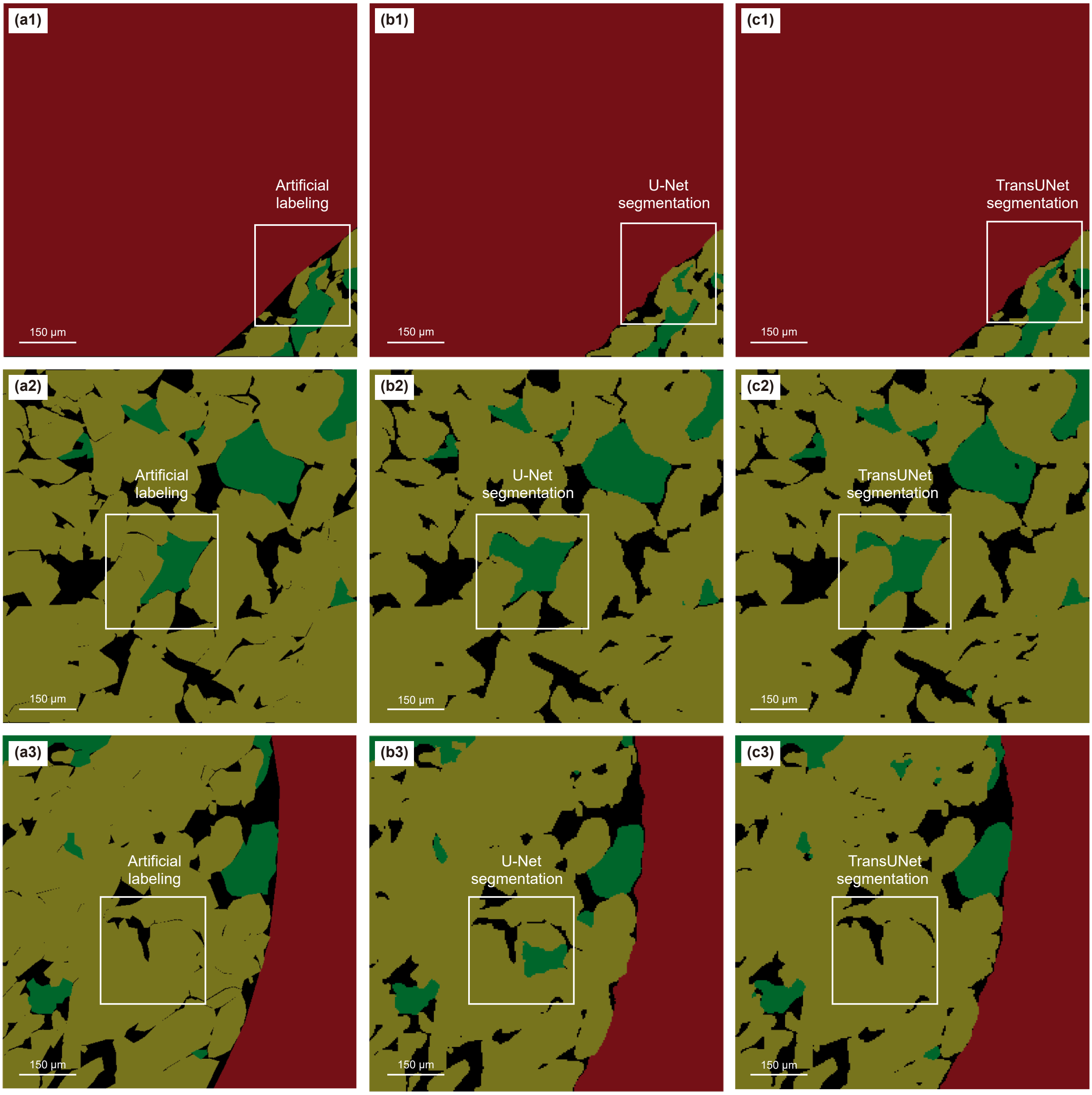

Fig. 5. Comparison of image segmentation results: (**a1**)–(**a3**) manually annotated ground truth images; (**b1**)–(**b3**) segmentation results predicted by the U-Net model; (**c1**)–(**c3**) segmentation results predicted by TransUNet model.

model's performance against traditional thresholding methods (see Fig. 6(b) and Fig. S13). The CNN encoder demonstrates effective extraction of local features, including quartz edges, pore shapes, and an argillaceous microporous network. Also, these features remain consistent across different image resolutions. The Transformer module captures long-range dependencies between image patches via a self-attention mechanism. This enables the model to comprehend global semantic context and learn information about core mineral structures, their distribution, and pore geometry from large datasets. Consequently, even when fine details are blurred, the model can leverage global semantic information to make reasonable inferences based on the retained large-scale structural information. This combined approach achieves high-precision segmentation of low-resolution CT images. Consequently, it effectively characterizes the evolution process of pore structures. In practice, all input images are cropped to a uniform size before being fed into the network for inference. This step ensures consistent input dimensions and leverages the model's inherent scale-invariant feature learning capability to adapt to resolution variations. However, a fundamental limitation must be acknowledged: low-resolution images inherently suffer from unavoidable information loss. The model's capability is based on reasoning from the information present within the image; it cannot create detail that was not originally captured. Therefore, if critical details are entirely absent from the low-resolution input,

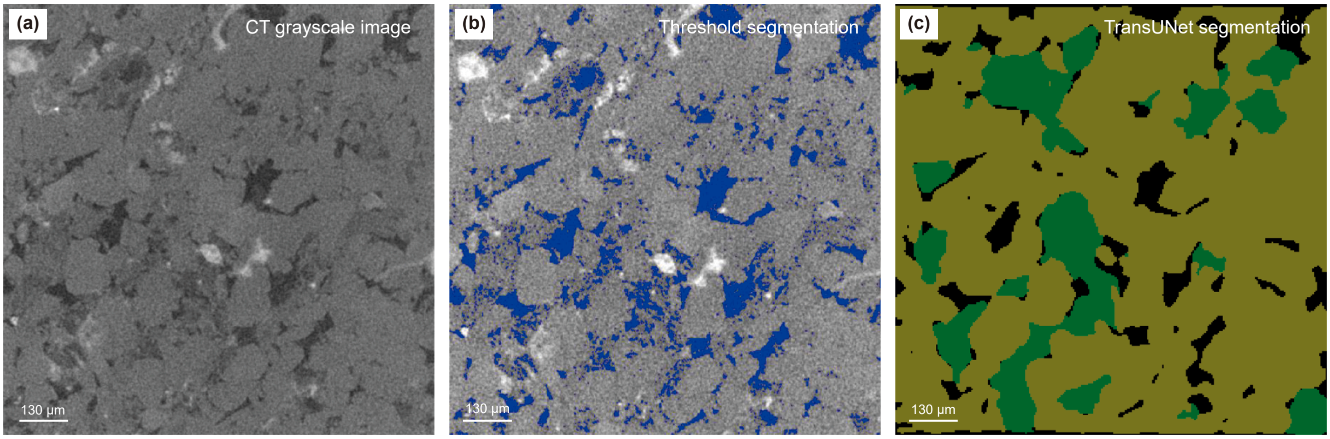

Fig. 6. Comparison of core CT images under different resolutions: (**a**) original CT image; (**b**) segmentation result using the thresholding method; (**c**) segmentation result generated by the TransUNet model.

the upper limit of segmentation performance will inevitably be constrained.

The thresholding approach, while capable of accurately identifying larger pore regions, exhibited significant over-segmentation within argillaceous and micro-porosity networks. This method frequently misclassified clay minerals as pores and failed to detect fine micropores. This limitation stems from its reliance on a single grayscale threshold, an approach which overlooks the spatial relationships within microstructures. In contrast, the TransUNet model maintained high segmentation accuracy for macro-pores, exceeding 90%. It also successfully identified micro-porosity networks within clay matrices (shown in green in Fig. 6(c)), where the continuity and topology of the segmented regions align closely with the actual pore distribution. Unlike threshold-based methods, which often misinterpret water-absorbent clay particles as pores due to their neglect of the spatial coupling between micropores and clay, TransUNet utilizes a Transformer-based attention mechanism. This mechanism effectively captures local textural patterns, such as the flocculent structure of clay and variations in pore connectivity. Consequently, TransUNet significantly improves recognition accuracy for micro-porosity networks, thereby providing a reliable area parameter for subsequent porosity correction. Owing to the resolution limitations of conventional CT and constraints in manual annotation, the model cannot directly output absolute sizes of micropores. To address this, we introduced a porosity correction mechanism. Since the total porosity of the core after image segmentation comprises both the segmented pore fraction and the micropores within AMN, the following correction formula was applied:

$$\Phi\_\text{corrected} = \Phi + \text{A\_AMN} \times 0.45 \tag{3}$$

where Φ_corrected is the corrected porosity, Φ is the originally segmented porosity, and A_AMN is the area ratio of the AMN.

The correction coefficient of 0.45 was established through a regression analysis that compared model-derived porosity measurements with true values obtained from pre-flooding core samples. The goodness-of-fit for this regression, quantified by a coefficient of determination (R^2) of 0.98, is shown in Fig. S15. This coefficient was subsequently validated using core samples from both the during-flooding and post-flooding stages. Results indicated a significant improvement in the accuracy of corrected porosity calculations, with accuracy reaching 94% and 96% for the respective stages, confirming the coefficient's effectiveness. In order to validate the porosity correction formula comprehensively, three sandstone samples representing high, medium, and low permeability were selected from various blocks of the Daqing Oilfield (Fig. S16). And, the accuracy of the corrected porosity exceeded 90% across all samples (Table S2). Consequently, the correction coefficient possesses significant reference value and demonstrates statistical relevance for the sandstones analyzed in this study. Ultimately, by integrating deep learning segmentation with physical empirical correction under low-resolution CT conditions (Table 4), a significant improvement in porosity prediction accuracy—from 53.8% to 94.8%—was achieved. To extend the practical applicability of our lithology research, we experimentally validated four rock types: sandstone, mudstone, shale, and carbonate. As shown in Fig. S17, high-resolution CT images of mudstone and shale clearly reveal fractures but not pores, which require visualization via scanning electron microscopy (SEM). The pore-type distribution observed in shale SEM images differs from that in sandstone (Fig. S18). Moreover, SEM provides only limited 2D images, insufficient for 3D reconstruction and accurate porosity calculation. Carbonate rocks contain abundant vugs, and their pore structures differ significantly from sandstone.

Therefore, extending this approach to mudstone, shale, and carbonate rocks would require a systematic investigation into pore formation mechanisms associated with heterogeneous mineral compositions. This entails building a comprehensive database linking porosity with high-resolution CT images, and employing

Table 4
Comparison of measured and algorithm-calculated core porosity during alkaline flooding.

Parameter	Pre-flooding	During-flooding	Post-flooding
Porosity (measured value), %	18.20	19.30	19.60
Porosity (grayscale threshold segmentation algorithm), %	5.90	11.00	14.10
Accuracy (grayscale threshold segmentation algorithm), %	32.42	56.99	71.94
Porosity (image segmentation method), %	16.63	18.76	20.40
Accuracy (image segmentation method), %	91.38	97.20	95.94

Note: The porosity values calculated in this study using the grayscale threshold method and the image segmentation method are not based on direct estimations from 2D images. Instead, they are derived from 3D reconstructions performed on extensive data. Consequently, these results are comparable to the core porosity obtained through actual measurement by the gas method.

machine learning to explore potential functional relationships among mineral composition, pore type, and correction coefficients. Such efforts would improve the accuracy of correction parameters and support the development of a more generalizable predictive model. Ultimately, this approach could establish a new paradigm for the digital characterization of unconventional reservoirs.

3.4. Time resolution law of pore change in the flooding core

As shown by the pressure curve in Fig. 7(a) and the permeability data in Table 5, the pressure increased rapidly from 0 to 10.6 MPa during the initial flooding stage (0–1.5 days). The sharp rise is attributed to the core's low initial permeability (28.7 mD) and porosity (18.2%). Under these conditions, capillary resistance and clay mineral swelling (hydration expansion) collectively increased flow resistance. This interpretation references the XRD data for various mineral crystal planes shown in Fig. 9. The clay minerals referenced include illite, montmorillonite, kaolinite, chlorite, and others (Xiang and Ye, 2020). Subsequently, the pressure dropped sharply to 8.9 MPa and then gradually recovered to 9.5 MPa (1.5–12.5 days), indicating that the alkaline solution dissolved clay cement and formed dominant flow channels. Alkaline solutions can damage the crystal structure of clay minerals, primarily including illite, montmorillonite, and chlorite (Fig. 9(f)) (Feng et al., 2016). The OH^- ions readily attack the bonds between silicon (Si) and aluminum (Al), specifically the –Si–O–Al– linkages. This reaction generates soluble silicate ions, such as $H_3SiO_4^-$, and aluminate ions, like $[Al(OH)_4]^-$ (Zhang et al., 2024). During the mid-term flooding phase, permeability increased sharply to 46.4 mD. This rise occurred as the Jamin effect diminished. Concurrently, the development of secondary pores enhanced the flow capacity. These two processes acted synergistically. With continued flooding over 12.5 to 30 days, the pressure steadily decreased to 6.0 MPa. By day 30, it stabilized at around 5.5 MPa. Further analysis of the pH curve (Fig. 7(b)) reveals the underlying mechanism: alkali consumption occurred in three stages. Initially (0–1.5 days), rapid adsorption of OH^- by clay minerals caused the pH to drop sharply (from 13.2 to 13.0) (Peng et al., 2016).

Subsequently, the pH rebounded to 13.16 due to the partial desorption and rapid dissolution of clay minerals, as shown in

Table 5
Variations in permeability and porosity during the core flooding experiment.

Parameter	Pre-flooding	During-flooding	Post-flooding
Permeability, mD	28.7	46.4	55.6
Porosity (measured value), %	18.2	19.3	19.6

Fig. S22 (Table S3)—a mechanism confirmed by clay-alkali immersion simulation experiments. The pH then gradually declined to 12.9 and eventually stabilized, a result of the slow dissolution of quartz and feldspar, which was verified by the shifted quartz peak in the XRD pattern (Fig. 9(e)). Throughout this process, continuous slow dissolution enlarged the pore-throat diameters. This expansion is evidenced by a pressure drop from 5.5 to 4.5 MPa between day 30 and 50 (Fig. 7(a)), leading to a late-stage permeability of 55.6 mD and a porosity increase to 19.6%. These results confirm that dissolution expanded the connected flow pathways. During the period from day 40 to day 50 (represented by the purple bar), both the flooding pressure and the pH value stabilized at 4.5 MPa and a constant level, respectively. This steady state was achieved through a dual balance mechanism: the limited core length (1 cm) prevented dissolution products such as calcium silicate from reaching precipitation saturation, thereby avoiding the risk of pore-throat blockage by secondary minerals. This outcome contrasts with the permeability decline typically observed in traditional long-core experiments.

Concurrently, the dissolution rates of quartz/feldspar and OH^- reached a dynamic balance, establishing a new steady-state flow network. This state is clearly demonstrated by the horizontal pressure baseline from day 30 to 50 in Fig. 7(a). The underlying mechanism involves the initial rapid dissolution products (silicate and aluminate ions) reacting with solution cations (e.g., Ca^{2+}, Na^+) to form an aluminosilicate gel layer on mineral surfaces. This layer hinders further fluid contact with the reactive substrate, while the consumption of OH^- ions lowers the pH (Khorshidi et al., 2018; Hou et al., 2018). The increasing concentration of dissolved ions further reduces the fluid's chemical aggressiveness, causing the reaction rate to gradually slow and eventually stabilize. The aluminosilicate gel can also migrate into larger pores where, under pressure, it dehydrates to form clay-like substances. This process contributes to the observed increase in flocculent AMN regions.

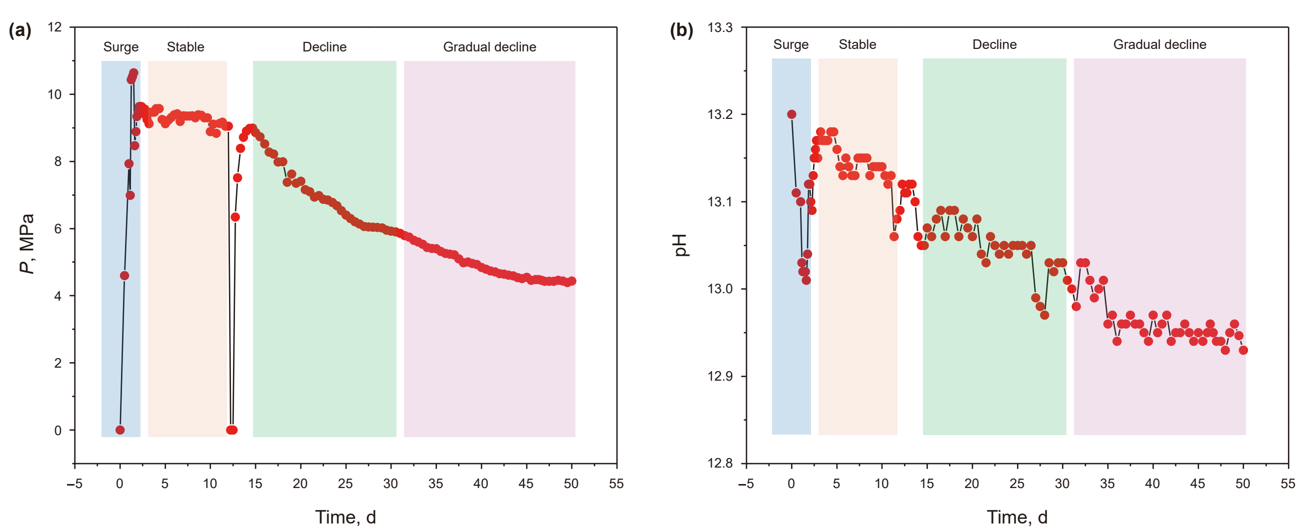

Fig. 7. Results of core flooding experiments. (**a**) The variation in pressure during the flooding process, while (**b**) the corresponding changes in pH value of the effluent.

Together, these results elucidate the co-evolution of pore networks and flow capacity from a multi-scale perspective, clarifying the dynamic response mechanisms of core structures during alkaline flooding. To accurately capture the details of this pore network evolution, further validation via high-resolution CT imaging is warranted.

Through systematic comparison of CT images before, during, and after flooding (Fig. 8), the evolution of two distinct regions was elucidated. In Region 1, the initial structure consisted of high-density minerals (e.g., dolomite, identified based on rock density literature, SEM-EDS comparisons (Yang et al., 2023), and shifting XRD peaks in Fig. 9(b)) enveloping AMN—a coexisting structure of clay and pores (Fig. 8(a1)). By the mid-term stage of strong alkali injection (Fig. 8(b1)), significant dissolution of dolomite occurred, leading to the expansion of the originally enclosed flocculated clay-pore network. By the flooding endpoint (Fig. 8(c1)), these high-attenuation minerals had completely disappeared, transforming the region into a dark clay-pore mixed structure. This sequence directly demonstrates that dissolution of high-density minerals preferentially releases pore space. Simultaneously, Region 2 exhibited a distinct mineral-pore boundary evolution. Initially sharp before flooding (Fig. 8(a2)), the interface became blurred and passivated after mid-term alkaline action (Fig. 8(b2)), eventually forming a connected network of secondary dissolution pores by the later stage, with pore diameters increasing by 60%–70% (Fig. 8(c2)). This boundary reconfiguration originated from the slow dissolution of quartz/feldspar, described by the reaction $2NaOH + SiO_2 \rightarrow Na_2SiO_3 + H_2O$, which gradually dissolved mineral edges.

Focusing on the clay-dominated regions (Regions 3 and 4), Region 3 initially exhibited open pores, which sharply contracted during the mid-term stage due to clay swelling and flocculent growth. However, sustained alkali action later dissolved some flocculent structures, regenerating them into irregular micropores. More strikingly, Region 4 underwent a complete disappearance-regeneration cycle of micropores: the pre-flooding micro-pore

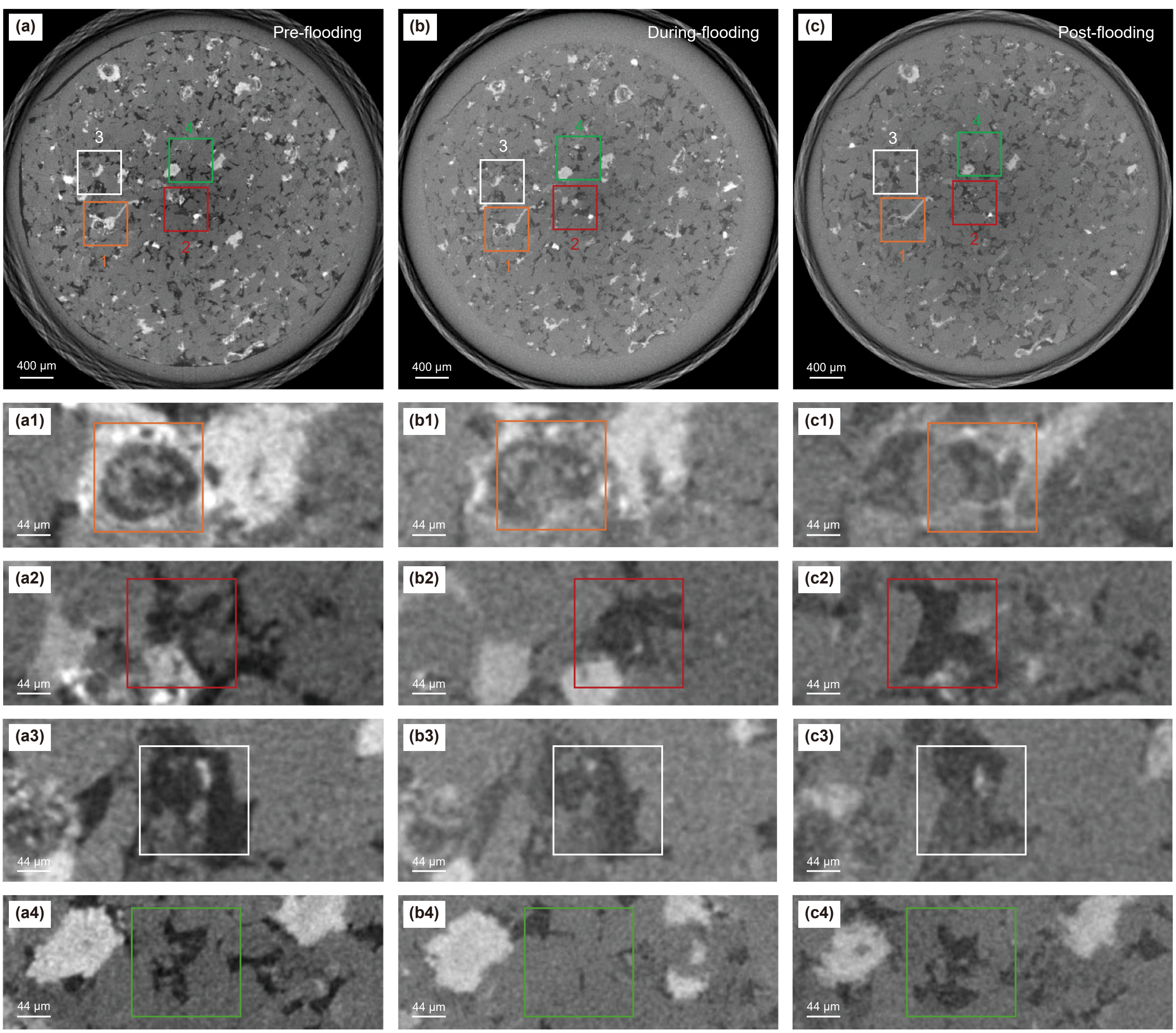

Fig. 8. CT image series tracking microstructural changes in a core sample during alkaline flooding. The varying grayscale contrast directly correlates with atomic number: high-Z elements (e.g., Fe) appear bright (likely Fe-dolomite), whereas low-Z elements (e.g., Si, Al in clay minerals) are shown in gray. (**a**)–(**c**) Full views at different stages. (1)–(4) Magnified views revealing detailed evolution. For more CT slice images see Figs. S19–S21.

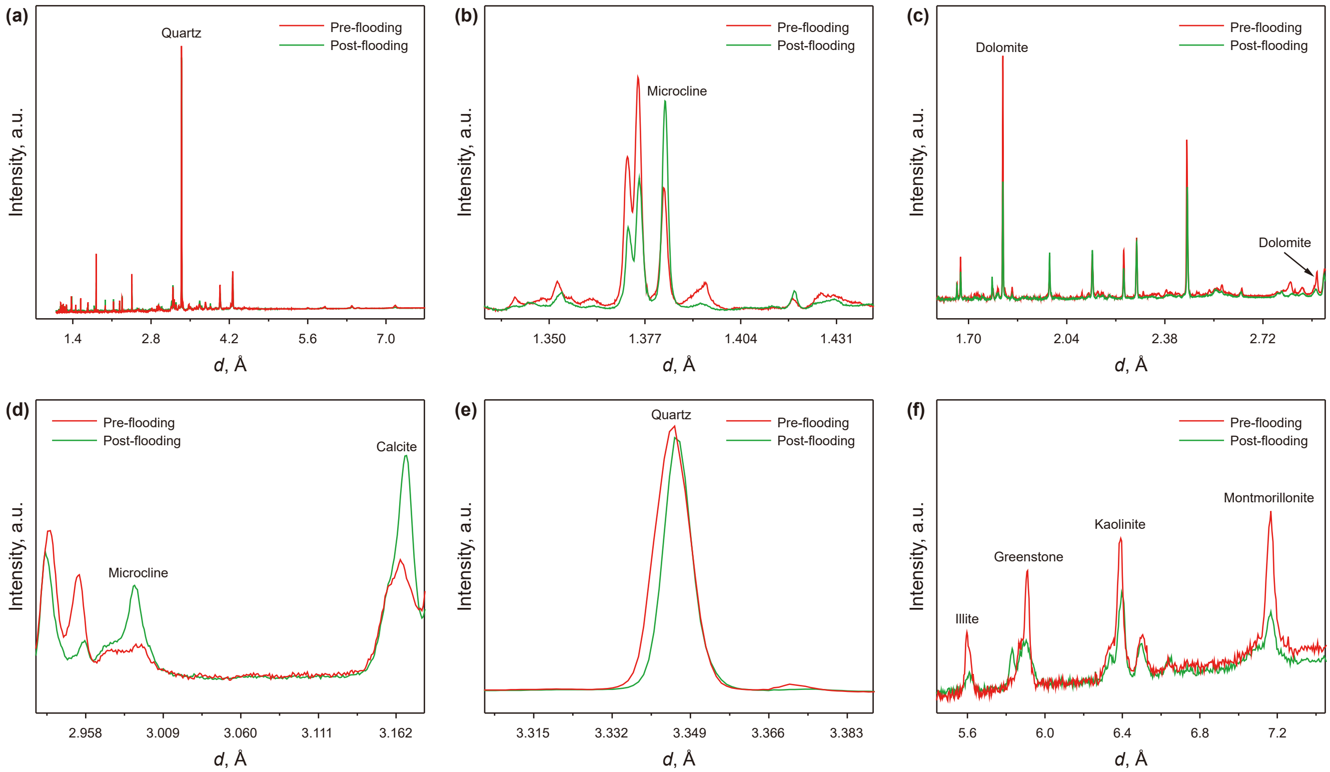

Fig. 9. Synchrotron radiation X-ray diffraction (SR-XRD) patterns of the core sample. (**a**) Full-range spectrum (d = 1–7 Å). Magnified views of characteristic regions for (**b**) plagioclase (d = 1.33–1.45 Å), (**c**) mica (d = 1.7–2.8 Å), (**d**) plagioclase and calcite (d = 2.95–3.09 Å and 3.162 Å), (**e**) quartz (d = 3.34 Å), and (**f**) clay minerals (d = 5.5–7.3 Å), showing characteristic peaks of illite, montmorillonite, kaolinite, and chlorite. The intensities of the quartz peaks are normalized to highlight peak shifts.

and mineral assemblage was first completely filled by mid-term clay expansion, reducing porosity by 80%–90%, before clay migration and reorganization ultimately reshaped new micropore structures in the later stage. Integrating these changes reveals a three-stage cooperative evolution (Ahmadullah and Chrysochoou, 2025): (1) pore shrinkage from clay swelling, (2) secondary pore expansion from silicate dissolution, and (3) micro-pore regeneration via clay reconstruction. This process is driven by differential mineral phase transformation: high-density dolomite dissolved rapidly within 0–12.5 days; low-density silicate minerals dissolved slowly between 12.5 and 25 days; and clay minerals underwent dynamic equilibrium through swelling, migration, and reorganization (Ainsworth et al., 2005). These visual findings correlate with a 93.7% enhancement in macroscopic flow capacity. They clarify that, in short-core experiments, the continuous enlargement of silicate dissolution channels avoids secondary precipitation—a key contrast to traditional long-core tests. Consequently, this mechanism drives the steady pressure reduction observed in Fig. 7(a) and provides a theoretical basis for designing alkaline flooding in reservoirs.

3.5. Evolution to log-normal pore-size distribution in cores after alkaline flooding

The CT sequence images of the core were accurately segmented using a neural network. The segmentation results were then converted into a three-dimensional model using commercial 3D reconstruction software (e.g., AVIZO), which interactively reads RAW-format CT images and reconstructs the segmented regions. This model intuitively reveals the microstructural evolution of the core sample across different flooding stages. The pre-flooding 3D reconstruction (Fig. 10(a1)) clearly shows the initial composition: yellow areas represent primary minerals such as quartz, forming the main rock skeleton; green areas indicate the clay mineral network, displaying a layered distribution consistent with natural sedimentation; and black areas signify pore space. By the mid-flooding stage (Fig. 10(a2)), significant structural changes occurred. The pore space (black) increased in abundance and exhibited markedly enhanced connectivity. Simultaneously, the clay mineral network (green) became more uniformly and widely distributed throughout the core. These alterations are closely related to chemical interactions with the flooding fluid (e.g., alkaline solution), which can dissolve certain clay minerals, induce clay particle migration, or generate new mineral phases, thereby modifying the original microstructure (Gaboreau et al., 2012).

By the late flooding stage (Fig. 10(a3)), a new and more stable pore structure had formed within the core. Pores of different sizes (macro-pores and micro-pores) developed a well-integrated, nested relationship and exhibited an overall uniform distribution. Compared to the middle stage, the relative content of the clay mineral network (green) decreased later on, likely due to strong fluid flushing or ongoing chemical reactions that may have degraded or displaced some clay structures from the system. To specifically track the evolution of the pore space, the black pore distributions were extracted and displayed separately in Fig. 10(b1)–10(b3). This sequence clearly illustrates how the pore network evolved as flooding progressed: initially relatively isolated and unevenly distributed before flooding, it transformed into a connected, evenly distributed, and complex network afterward. In conclusion, integrating neural network–based image segmentation with 3D reconstruction enables precise nanoscale characterization of the dynamic microstructural evolution in cores during chemical flooding. This approach provides powerful visual evidence that supports both the quantitative evaluation of

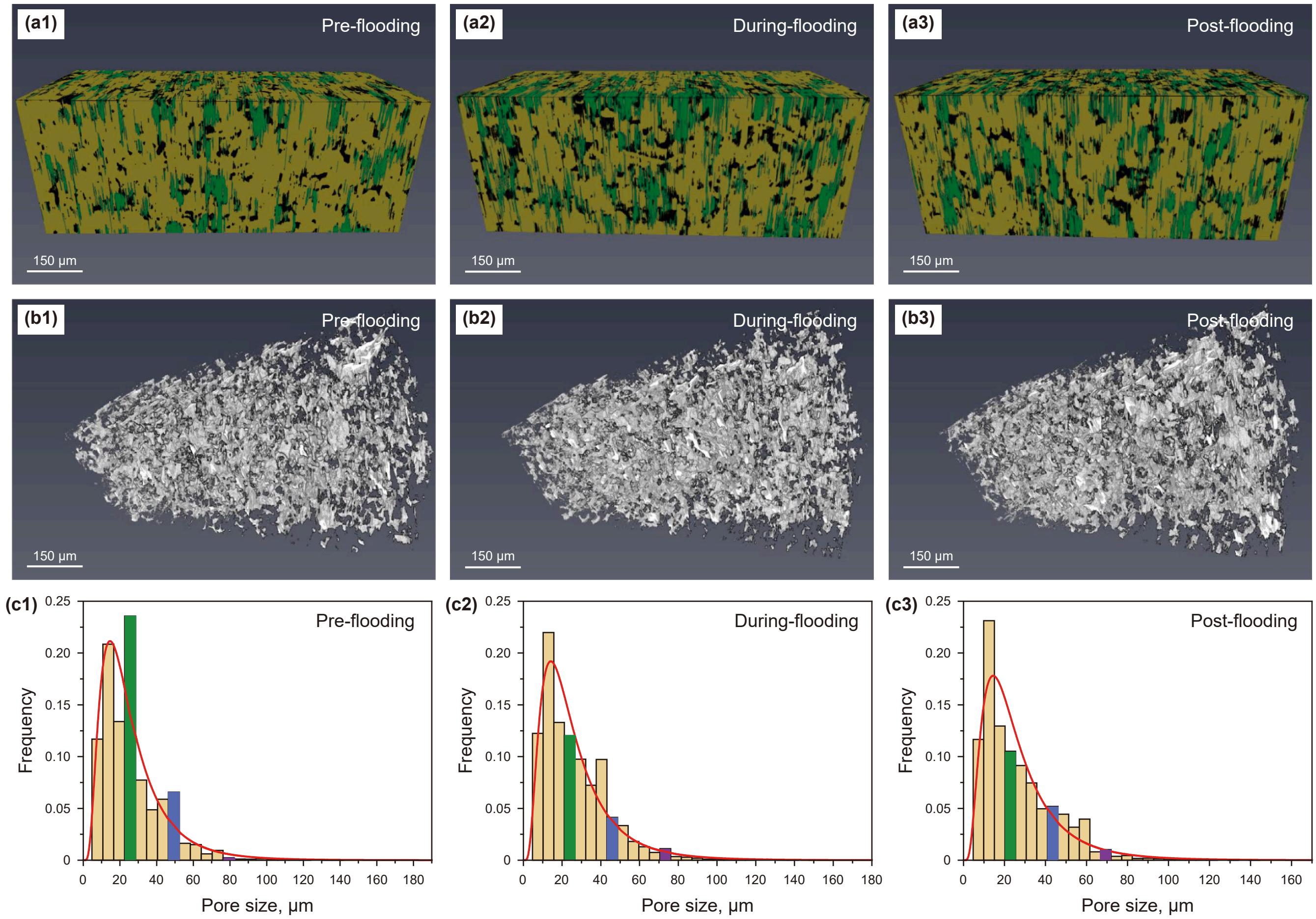

Fig. 10. Three-dimensional pore structure and pore size distribution of a core sample during alkaline flooding. (**a1**)–(**a3**) Reconstructed microstructure before, during, and after flooding (corresponding to the 3D reconstruction results in Figs. S26–S28). (**b1**)–(**b3**) Spatial distribution of discrete pores. (**c1**)–(**c3**) Pore size distribution histograms derived from the 3D reconstruction data.

flooding effectiveness and the understanding of pore evolution mechanisms.

The pore size distribution histograms (Figs. 10(c1) and S23–S25) reveal a bimodal and heterogeneous distribution in the original core, with data derived from extensive CT data processed through image segmentation and 3D reconstruction. Pores of 30 and 50 μm constitute a significant proportion, exceeding 20%, whereas those between 60 and 100 μm account for only about 7%. Pores larger than 100 μm decrease sharply, and those around 160 μm are nearly absent. This pattern reflects a dual-pore system dominated by pores associated with clay cementation (30 μm peak) and micro-fractures (50 μm peak) in the undisturbed core. As indicated in Table 6, AMN constituted 17.75% before flooding. By the middle stage of alkaline flooding (12.5 days), the proportion of 50 μm pores decreased by 50% (Fig. 10(c2)), while the proportion of 80 μm pores increased approximately 1.8-fold. Table 6 also shows that the overall pore proportion rose to 10.61%. This shift results from two competing effects. First, clay minerals such as montmorillonite swell upon contact with alkali, compressing smaller throats and reducing the 30–50 μm pores. Second, strong alkali preferentially dissolves aluminosilicates like montmorillonite and chlorite ($Al_2Si_2O_5(OH)_4 + 2NaOH \rightarrow 2NaAlSiO_4 + 3H_2O$), forming secondary pores around 80 μm in size.

Synchrotron radiation X-ray diffraction (SR-XRD) analysis further clarified the structural impact of alkaline flooding on mineral composition. The full-range XRD spectrum of the sample (Fig. 9(a)) indicates that the mineral composition consists primarily of five components, corresponding to the characteristic diffraction regions of feldspar, mica, calcite, quartz, and clay minerals, which together form the main mineral assemblage of the sandstone core. The diffraction peaks at *d*-spacings of 1.38 and 2.95 Å (Fig. 9(b)) are both attributed to plagioclase (Hanafi et al., 2025). Their presence provides key evidence for the formation of new crystalline phases, such as zeolite minerals, resulting from reactions between the alkaline fluid and silicates/aluminosilicates. The peaks at 1.81 and 2.89 Å (Fig. 9(c)) are assigned to mica. After alkaline flooding, the intensities of these characteristic peaks decreased significantly, indicating dissolution or structural alteration of mica under alkaline conditions, consistent with its

Table 6
Variation in component segmentation proportions from Core CT images based on TransUNet during alkaline flooding.

Core	Framework minerals, %	Pore, %	AMN, %	Porosity correction value, %
Pre-flooding	73.61	8.65	17.75	16.63
During-flooding	71.28	10.61	18.11	18.76
Post-flooding	65.34	8.73	25.93	20.40

reactivity in water–rock reactions under such environments (Lammers et al., 2017). The enhancement of the diffraction peak at 3.162 Å (calcite; Fig. 9(d)) reflects the reaction between carbonate and calcium ions. Although the diffraction intensity at 3.34 Å did not decrease after flooding, its d-spacing shifted toward lower values (Fig. 9(e)), likely due to lattice contraction induced by the alkaline fluid.

In the characteristic region for clay minerals, diffraction peaks at approximately 5.60, 5.91, 6.40, and 7.17 Å were observed both before and after flooding (Fig. 9(f)), corresponding to illite, chlorite, kaolinite, and montmorillonite (Ferrage et al., 2005). After flooding, the intensity of the kaolinite peak changed the least, whereas the intensities of the other peaks dropped sharply to about one-quarter of their original levels, indicating substantial dissolution or structural degradation of layered silicate clay minerals such as illite and chlorite under alkaline conditions. These mineral transformations promoted a shift in pore size distribution toward a log-normal pattern, contributing to a 61.7% increase in permeability (from 28.7 to 46.4 mD). In the later stage, although the proportion of 80 μm pores slightly decreased, the overall distribution retained its log-normal character. This is attributed primarily to the slow dissolution of quartz and feldspar, which enlarged pores from around 80 to over 90 μm, while dissolution of clay minerals released additional micropores. Together, these mechanisms enhanced pore homogeneity, further increasing permeability and resulting in a corrected porosity of 20.40%.

These findings illustrate that pore evolution is essentially a dynamic balance between mineral reactions and pore structure reorganization. In the short term, clay swelling compresses smaller pores, while soluble mineral dissolution generates mid-sized pores (transforming 50→80 μm pores) (Liu et al., 2023b). Over the long term, slow silicate dissolution enlarges connected pore throats, and regenerated clay micropores (increased by 8.18%) optimize pore-throat coordination. The eventual steady state, reflected in the log-normal distribution (Fig. 10(c3)), confirms the establishment of dominant dissolution pathways (Dou et al., 2021). A 20.4% increase in corrected porosity and a doubling of permeability demonstrate substantial improvement in pore connectivity. Alkaline flooding increased the absolute porosity. The porosity rose from 16.63% to 20.40%. It also transformed the pore size distribution. The distribution changed from a bimodal pattern to a log-normal pattern. This transformation is shown in Fig. 10(c1) to Fig. 10(c3). The process restructured the topology of the pore network. Expanded throat sizes (>70 μm pores increased 2.2-fold) and improved connectivity (argillaceous micro-porosity network increased by 46.1%) synergistically reduced flow resistance. This mechanism provides a theoretical foundation for enhancing chemical flooding efficiency in reservoirs.

4. Conclusion

This study established and validated a novel cross-scale analytical workflow integrating dynamic flooding experiments, multi-resolution CT imaging, intelligent image segmentation, and quantitative pore analysis. For the first time, this approach enables quantitative characterization of dynamic pore structure evolution during fluid flooding in reservoirs across scales from sub-micron pores to macropores. The proposed framework combines synchrotron radiation CT (0.8 μm resolution) with conventional CT (5.4 μm resolution) and employs a Transformer-based TransUNet image segmentation model to effectively address long-range spatial dependencies in CT images. The main conclusions are as follows:

(1) Segmentation performance:

The TransUNet model demonstrates superior performance in segmenting complex pore structures. For the argillaceous microporous network, an IoU of 74.92% is achieved, representing a 7% improvement over previous U-Net models. Boundary recognition accuracy for clay minerals is increased by 6%, providing a reliable basis for quantitative analysis.

(2) Porosity correction:

A porosity correction formula that integrates physical mechanisms with deep learning outputs was proposed (Φ_corrected = Φ + A_AMN × 0.45). This method significantly improves the calculation accuracy of low-resolution CT data and reduces the porosity error calculated from conventional CT images to 2.7%, achieving an accuracy of 94% and effectively addressing the porosity underestimation caused by limited resolution.

(3) Dynamic pore evolution mechanism:

The research revealed a dynamic pore equilibrium mechanism driven by mineral phase transformation during alkaline flooding. Pore evolution exhibited distinct spatiotemporal characteristics. In the short term (0–1.5 days), clay swelling compresses small pores (<50 μm decreased by 50%), while dolomite dissolution generates secondary pores (~80 μm increased by 1.8 times). In the long term (30–50 days), slow silicate dissolution continuously enlarges pores, resulting in a 93.7% increase in permeability and the clay mineral reorganization leads to a 46.1% increase in the proportion of the argillaceous microporous network. These processes collectively shifted the pore size distribution from a bimodal pattern to a more stable log-normal distribution.

(4) Methodological superiority:

Compared with grayscale threshold segmentation (37% error) and the conventional U-Net model (IoU = 0.69), the proposed multi-scale CT-TransUNet framework significantly improves quantitative accuracy, reducing recognition error for the argillaceous microporous network within the dynamic pore system to below 20%, providing a reliable tool for more accurate description of pore structure dynamics.

Overall, this study not only advances a specialized Transformer-based artificial neural network model for handling long-range dependencies in CT images but also clarifies the spatiotemporal coupling between mineral phase transformation and pore evolution. The proposed methodology provides a robust theoretical and technical foundation for the dynamic regulation and efficient development of oil and gas reservoirs.

CRediT authorship contribution statement

Shao-Hua Zhou: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Tian-Bao Liu:** Writing – review & editing, Validation, Methodology, Formal analysis. **Yu Zhao:** Writing – review & editing, Validation, Methodology, Formal analysis. **Shao-Hao Yin:** Software, Methodology, Investigation, Conceptualization. **Zhi-Yu Liu:** Software, Methodology, Investigation, Conceptualization. **Ji-Jun Liu:** Methodology, Formal analysis, Data curation, Conceptualization. **Ling-Wei Du:** Methodology, Formal analysis, Data curation, Conceptualization. **Yue-Tong Zhao:** Methodology, Formal analysis, Data curation, Conceptualization. **Wei-Guang Shi:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

No conflict of interest exists in the submission of this manuscript. I would like to declare on behalf of my co-authors that the work described is original research that has neither been published previously, nor under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

Acknowledgements

This work was supported by Program for Young Talents of Basic Research in Universities of Heilongjiang Province (YQJH2024036), Collaborative Innovation Projects of "Double First-class" Disciplines in Heilongjiang Province (LJGXCG2024-P20), and Outstanding Talent Cultivation Foundation of Northeast Petroleum University (SJQHB202004). We gratefully acknowledge the technical support provided by beamline BL13HB (2024-SSRF-PT-508984) in-situ CT characterization, as well as beamline BL14B1 (2024-SSRF-PT-508987) for in-situ XRD characterization at Shanghai Synchrotron Radiation Facility (SSRF).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petsci.2026.04.008.

References

Ahmadullah, T., Chrysochoou, M., 2025. Development of a kinetic-thermodynamic model for lime-stabilization of Na-bentonite. Geochem. Trans. 26 (1), 6. https://doi.org/10.1186/s12932-025-00103-4.

Ainsworth, C.C., Zachara, J.M., Wagnon, K., et al., 2005. Impact of highly basic solutions on sorption of Cs+ to subsurface sediments from the Hanford site, USA. Geochem. Cosmochim. Acta 69 (20), 4787–4800. https://doi.org/10.1016/j.gca.2005.06.007.

Cao, S., Sang, Q., Zhao, G., et al., 2024. CO_2–Water–rock interaction and its influence on the physical properties of Continental shale oil reservoirs. Energies 17 (2), 477. https://doi.org/10.3390/en17020477.

Chen, J., Mei, J., Li, X., et al., 2024. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. Med. Image Anal. 97, 103280. https://doi.org/10.1016/j.media.2024.103280.

Dou, W., Liu, L., Jia, L., et al., 2021. Pore structure, fractal characteristics and permeability prediction of tight sandstones: A case study from Yanchang Formation, Ordos Basin, China. Mar. Petrol. Geol. 123, 104737. https://doi.org/10.1016/j.marpetgeo.2020.104737.

Feng, G., Cheng, P., Yan, W., et al., 2016. Accelerated crystallization of zeolites via hydroxyl free radicals. Science 351 (6278), 1188–1191. https://doi.org/10.1126/science.aaf1559.

Ferrage, E., Lanson, B., Sakharov, B., et al., 2005. Investigation of smectite hydration properties by modeling experimental X-ray diffraction patterns: Part I. Montmorillonite hydration properties. Am. Mineral. 90, 1358–1374. https://doi.org/10.2138/am.2005.1776.

Gaboreau, S., Claret, F., Crouzet, C., et al., 2012. Caesium uptake by Callovian–Oxfordian clayrock under alkaline perturbation. Appl. Geochem. 27 (6), 1194–1201. https://doi.org/10.1016/j.apgeochem.2012.02.002.

Gu, X., Angelov, P., Han, J., et al., 2023. Multilayer evolving fuzzy neural networks. IEEE Trans. Fuzzy Syst. 31 (12), 4158–4169. https://doi.org/10.1016/j.apgeochem.2012.02.002.

Hanafi, M.I.J., Murshed, M.M., Robben, L., et al., 2025. Plagioclase feldspars ($Ca_{1-x}Na_x$)($Al_{2-x}Si_{2+x}$)O_8: Synthesis and characterizations of mechanical weathering relevant to Martian regolith. Z. Kristallogr. Cryst. Mater. 240 (1–2), 13–27. https://doi.org/10.1515/zkri-2024-0117.

Hou, D., Li, T., Wang, P., 2018. Molecular dynamics study on the structure and dynamics of NaCl solution transport in the nanometer channel of CASH gel. ACS Sustain. Chem. Eng. 6 (7), 9498–9509. https://doi.org/10.1021/acssuschemeng.8b02126.

Hoyer, L., Dai, L., Van Gool, L., 2024. Domain adaptive and generalizable network architectures and training strategies for semantic image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 46 (1), 220–235. https://doi.org/10.1109/tpami.2023.3320613.

Hu, C., Jia, Y., Duan, Z., 2022. Pore and permeability properties of reservoir sandstone under a uniaxial compression CT test. J. Nat. Gas Sci. Eng. 104, 104666. https://doi.org/10.1016/j.jngse.2022.104666.

Ji, J.F., Guo, H., Xue, Y.L., et al., 2023. The new X-ray imaging and biomedical application beamline BL13HB at SSRF. Nucl. Sci. Tech. 34, 197. https://doi.org/10.1007/s41365-023-01349-2.

Kalinaki, K., Malik, O.A., Lai, D.T.C., 2023. FCD-AttResU-Net: An improved forest change detection in Sentinel-2 satellite images using attention residual U-Net. Int. J. Appl. Earth Obs. Geoinf. 122, 103453. https://doi.org/10.1016/j.jag.2023.103453.

Khorshidi, Z.N., Tan, X., Liu, Q., et al., 2018. Molecular dynamics study of the dissolution mechanism of kaolinite basal surfaces in alkali media. Appl. Clay Sci. 152, 29–37. https://doi.org/10.1016/j.clay.2017.10.025.

Kudasik, M., Anioł, Ł., Gajda, A., et al., 2023. Influence of confining pressure on permeability and structural properties of selected sedimentary, igneous, and metamorphic rocks. Bull. Eng. Geol. Environ. 82 (12), 455. https://doi.org/10.1016/j.clay.2017.10.025.

Lai, N., Deng, J., Wang, Z., et al., 2023. Evaluation of the lower producing limit of the pore radius for imbibition and flooding based on shale oil reservoir characteristics. Energy Fuels. 37 (15), 11453–11464. https://doi.org/10.1021/acs.energyfuels.3c01421.

Lammers, K., Smith, M.M., Carroll, S.A., 2017. Muscovite dissolution kinetics as a function of pH at elevated temperature. Chem. Geol. 466, 149–158. https://doi.org/10.1016/j.chemgeo.2017.06.003.

Liang, Y.J., Liang, W.F., He, W., et al., 2021. Response to the variation of clay minerals during ASP flooding in the saertu oilfield in the Songliao Basin. Front. Earth Sci. 9, 764052. https://doi.org/10.3389/feart.2021.764052.

Lin, C., Jia, X., Chen, X., et al., 2025. Failure and fractal characteristics of anisotropic oil shale using real-time CT technology. Powder Technol. 466, 121522. https://doi.org/10.1016/j.powtec.2025.121522.

Liu, Q., Sun, M., Sun, X., et al., 2023a. Pore network characterization of shale reservoirs through state-of-the-art x-ray computed tomography: a review. Gas Sci. Eng. 113, 204967. https://doi.org/10.1016/j.jgsce.2023.204967.

Liu, S., Zhang, L., Su, X., et al., 2023b. Micro-CT characterization on pore structure evolution of low-permeability sandstone under acid treatment. Appl. Geochem. 152, 105633. https://doi.org/10.1016/j.apgeochem.2023.105633.

Lyu, Y., Dasani, D., Tsotsis, T., et al., 2021. Characterization of shale using Helium and Argon at high pressures. J. Petrol. Sci. Eng. 206, 108952. https://doi.org/10.1016/j.petrol.2021.108952.

Peng, C., Min, F., Liu, L., et al., 2016. A periodic DFT study of adsorption of water on sodium-montmorillonite (001) basal and (010) edge surface. Appl. Surf. Sci. 387, 308–316. https://doi.org/10.1016/j.apsusc.2016.06.079.

Song, J., Chen, X., Zhu, Q., et al., 2022. Global and local feature reconstruction for medical image segmentation. IEEE Trans. Med. Imag. 41 (9), 2273–2284. https://doi.org/10.1109/tmi.2022.3162111.

Taheri, K., Majd, S.H.H., Ghanbarian, B., et al., 2025. A synergistic approach to enhanced oil recovery by combining in-situ surfactant production and wettability alteration in carbonate reservoirs. Sci. Rep. 15 (1), 11688. https://doi.org/10.1038/s41598-025-96199-w.

Vásárhelyi, L., Kónya, Z., Kukovecz, Á., et al., 2020. Microcomputed tomography-based characterization of advanced materials: A review. Mater. Today Adv. 8, 100084. https://doi.org/10.1016/j.mtadv.2020.100084.

Wang, D., Chen, G., 2021. Seismic Stratum segmentation using an encoder–decoder convolutional neural network. Math. Geosci. 53 (6), 1355–1374. https://doi.org/10.1007/s11004-020-09916-8.

Wang, X., Li, S., Tong, B., et al., 2024b. Multiscale wettability characterization under CO_2 geological storage conditions: a review. Renew. Sustain. Energy Rev. 189, 113956. https://doi.org/10.1016/j.rser.2023.113956.

Wang, X., Mao, L., Kong, X., 2025. Microscopic effects of long-term water injection on reservoir pore structure. Geoenergy Sci. Eng. 253, 213952. https://doi.org/10.1016/j.geoen.2025.213952.

Wang, W., Yan, Z., Chen, D., et al., 2024a. The mechanism of mineral dissolution and its impact on pore evolution of CO_2 flooding in tight sandstone: a case study from the Chang 7 member of the Triassic Yanchang formation in the Ordos Basin, China. Geoenergy Sci. Eng. 235, 212715. https://doi.org/10.1016/j.geoen.2024.212715.

Xiang, G., Ye, W., 2020. Swelling of bentonite-sand mixtures after long-term dissolution in alkaline solution. Clays Clay Miner. 68 (5), 491–498. https://doi.org/10.1007/s42860-020-00090-w.

Xu, S., Yang, Z., Wu, S., et al., 2022. Fractal analysis of pore structure differences between shale and sandstone based on the nitrogen adsorption method. Nat. Resour. Res. 31 (3), 1759–1773. https://doi.org/10.1007/s11053-022-10056-5.

Xu, X., Zhang, M., 2025. Integrating multi-scale features and transformer for high-resolution remote sensing image semantic segmentation. J. Electron. Imag. 34 (3), 33024. https://doi.org/10.1117/1.jei.34.3.033024, 033024.

Yang, L., Liu, Y., Wang, W., et al., 2025. Pore-scale flow mechanism of immiscible gas-assisted gravity drainage in strongly heterogeneous glutenite reservoirs. Int. J. Coal Sci. Technol. 12 (1), 12. https://doi.org/10.1007/s40789-025-00758-5.

Yang, R., Cong, R., Gong, Y., et al., 2023. Micromechanical contrast of ordos Basin Sandstone-Mudstone interbedded layered rocks. J. Geophys. Res. Solid Earth 128 (10). https://doi.org/10.1029/2023jb027190 e2023JB027190.

Yeh, B.M., FitzGerald, P.F., Edic, P.M., et al., 2017. Opportunities for new CT contrast agents to maximize the diagnostic potential of emerging spectral CT technologies. Adv. Drug Deliv. Rev. 113, 201–222. https://doi.org/10.1016/j.addr.2016.09.001.

Zhang, X.Y., Sun, H.J., Peng, T.J., et al., 2024. Purification mechanism of microcrystalline graphite and dissolution of non-carbon impurity during alkali autoclave-acid leaching. Phys. Chem. Miner. 51 (3), 25. https://doi.org/10.1007/s00269-024-01290-9.
Zhang, Y., Wang, L., Yang, D., et al., 2024a. CT imaging to study meso-structure evolution of fractured oil shale during in-situ pyrolysis by high-temperature water vapor injection. Geoenergy Sci. Eng. 241, 213104. https://doi.org/10.1016/j.geoen.2024.213104.
Zhao, B., Saxena, N., Hofmann, R., et al., 2023. Enhancing resolution of micro-CT images of reservoir rocks using super resolution. Comput. Geosci. 170, 105265. https://doi.org/10.1016/j.cageo.2022.105265.
Zhao, Y., Sun, Y., Yuan, L., et al., 2020. Impact of nanopore structure on coal strength: A study based on synchrotron radiation nano-CT. Results Phys. 17, 103029. https://doi.org/10.1016/j.rinp.2020.103029.