www.georesursy.ru

ORIGINAL ARTICLE

DOI: https://doi.org/10.18599/grs.2025.4.16

Core Plug Porosity Prediction Using Microtomography, Supervised Labeling, and a Shifted Window Transformer

R.I. Kadyrov[*]*, E.O. Statsenko, T.H. Nguyen, M.A. Skorobogatova*
Kazan Federal University, Kazan, Russian Federation

Recent advances in machine learning have enabled the automatic analysis of microtomography (μCT) images, facilitating more efficient rock property identification. This study aims to predict the experimentally measured open porosity of reservoir rocks using μCT images of standard core plugs. A dataset of 136 core plugs was collected, including 49 sandstone and 87 carbonate samples. Open porosity was experimentally determined using gas volumetry. The core plugs (30 ± 1 mm in height and diameter) were scanned using μCT with a resolution of 34.6–38.0 μm, producing 16-bit image stacks. The dataset consisted of 100,232 images (64,119 carbonate and 36,113 sandstone). To label the images, we introduced a supervised method called Segmentation of Unresolved Pores via Experimental Reference (SUPER), which segments dark voxels to match the experimentally measured open porosity, adapting to each sample's characteristics. Three shifted window (Swin) transformer models were trained: a universal model and specialized models for sandstone and carbonate. The models used transfer learning with ImageNet weights, followed by fine-tuning. Testing confirmed that specialized models outperformed the universal model. This highlights that training an ensemble of models adapted to specific rock types leads to better performance than a single general model for porosity prediction. A key challenge arose with sandstones, especially fine-grained types, where small pores merged due to resolution limitations. Future work should improve image resolution and feed detailed images into the model. The method has potential for full-scale core scans and early porosity assessment in raw core plugs, including fragile reservoirs with oil or bitumens.

Keywords: μCT (microtomography), open porosity, reservoir rock, core plug, machine learning, transfer learning, Swin (shifted window) transformer, porosity prediction

Recommended citation: Kadyrov R.I., Statsenko E.O., Nguyen T.H., Skorobogatova M.A. (2025). Core Plug Porosity Prediction Using Microtomography, Supervised Labeling, and a Shifted Window Transformer. *Georesursy = Georesources,* 27(4), pp. 67–82. https://doi.org/10.18599/grs.2025.4.16

1. Introduction

Porosity is the proportion of voids in a rock or sediment that can be occupied by fluids (McPhee et al., 2015). It is a basic parameter influencing the storage and flow characteristics of reservoir rocks, which is widely used in petroleum geology and is essential for resource assessment, reservoir modeling, and the optimization of production strategies (Kadyrov et al., 2018). Porosity data supports enhanced oil recovery methods (EOR), such as water flooding and gas injection, by providing an accurate assessment of remaining hydrocarbons in the reservoir during the extraction processes (Yang, 2017). Moreover, the interplay between porosity, permeability, and fluid dynamics is a key factor controlling the economic viability of reservoirs and, thus, their sustainable resource management (Dullien, 1992; Tiab, Donaldson, 2016). High-quality porosity analysis using data from core studies is essential to obtain reliable and comprehensive reserve estimates.

Several types of porosity are recognized, the most important of which are absolute porosity, open porosity, and effective porosity. Absolute porosity (bulk porosity) is defined as the ratio of the total pore volume in a rock, including isolated (closed) pores, to the volume of the entire rock (McPhee et al., 2015). Open porosity reflects the interconnected pore space accessible from the sample surface (including dead-end pores) (Naftaly et al., 2020). Effective porosity is defined as the percentage of interconnected pore space relative to the bulk volume (Yang, 2017).

Microtomography (μCT or micro-CT) is the key element of research for 3D-structure and physical properties of rocks in the domains of geology and rock physics (Carmignato et al., 2018; Kadyrov et al., 2022; Liu et al., 2022; Zhang et al., 2019). This non-destructive method permits high-resolution three-dimensional imaging of pores, microfractures, and cracks allowing for fine-detailed observation and quantification of pore structure, which determine the flow behavior and mechanical properties of rocks (Andrä et al., 2013; Blunt et al., 2013; Kadyrov et al., 2020; Van Geet et al., 2000; Zhang et al., 2019). Microtomography pore imaging

[*]Corresponding author: Rail I. Kadyrov
e-mail: rail7777@gmail.com

enables advanced computational simulations of fluid flow for improved accuracy of permeability assessments (Latief et al., 2017). Also, microtomographic reconstructions provide insights into the influence of pore structures on mechanical behavior, such as stress distributions and yield properties of the materials (Song et al., 2020). The sensitivity of pore-scale measurements through petrophysical imaging is a direct reminder that the resolution is extremely important for an accurate assessment of porosity and permeability (Alyafei et al., 2015). On a scale of standard core plugs, due to deficient resolution, it is typically used in special core analysis (SCAL) as a screening instrument prior to the filtration test to study the cores and point out any possible defects, fractures, or heterogeneities that may influence flow behavior during the process (Gerke et al., 2021; McPhee et al., 2015). Overall, these attributes position microtomography as an invaluable technique for advancing knowledge in rock physics and optimizing the evaluation of hydrocarbon reservoirs.

Machine learning (ML) has increasingly been applied to examine CT and μCT images of rocks, allowing automatic and effective identification of rock properties. Such methods then harness the power of neural networks to extract information from the complex image data. Previous studies have applied convolutional neural networks for the classification of different rock types and segmentation of formations in CT and μCT images (Chawshin et al., 2021; dos Anjos et al., 2021; Kadyrov et al., 2024; Zheng et al., 2024). Deep learning algorithms are also used to enhance resolution and improve the quality of the μCT images, thus enabling a detailed reconstruction of rock structure and extracting more information on the properties of a reservoir (Jackson et al., 2021; Karimpouli et al., 2024; Karimpouli, Kadyrov, 2022; Shan et al., 2022; Wang et al., 2019). Permeability modeling is another major use of machine learning methods, which apply geometric and mineralogical information from μCT scans to fluid flow simulation (Gärttner et al., 2023; Shaik et al., 2019; Zhang et al., 2022). Recent studies have also shown that segmentation accuracy directly affects the reliability of simulated flow properties, with minor segmentation errors leading to substantial deviations in permeability estimates (Lavrukhin et al., 2021). All of these improvements greatly enhance the efficiency of pore network analysis, the process of permeability estimation, and even stress distribution modeling in geological samples (Alzahrani et al., 2023; Qu et al., 2023; Tembely et al., 2020). Coupling such methods provides not just faster data processing but also more reliable results from μCT-based analyses, thus enabling more predictive and comprehensive geological modeling.

In recent times, machine learning advances have paved the way for dramatic success in predicting porosity from CT and μCT images of rock samples. Many studies have shown great promise in using convolutional neural networks (CNNs) and artificial neural networks (ANNs) for accurately predicting porosity by extracting complex patterns from images. CNN-based regression models have been successfully used on full-size CT core scans, surpassing traditional linear models in porosity estimation. The beneficial outcomes from this approach not only pertain to enhancement in porosity estimation but also in well log calibration and core sampling in heterogeneous reservoir intervals (Chawshin et al., 2022).

In addition to full-size core analysis, ML models have been trained on extensive datasets of CT and μCT images for quick property prediction. CNNs demonstrated accuracy and efficiency by accurately predicting porosity, specific surface area, and average pore size using 2D images of sandstone samples, with errors below 6.3% (Alqahtani et al., 2020). In a similar vein, dual-energy CT data combined with ANNs have produced submillimeter-resolution porosity models that provide insight into the precise distribution of petrophysical properties within rock cores, demonstrating the ability to extract meaningful information even from multiple images with varying modalities and resolutions (Ortiz et al., 2020). Even for clay-rich formations, ResNet architectures have shown good validity and minimal errors when predicting porosity on whole-core CT images (Wu et al., 2025).

Additional applications have focused on predicting porosity and other morphological properties. CNN and XGBoost models were trained using μCT images of various sandstone types to predict porosity and permeability. It was demonstrated that including different rock types improves model generalization and predictive performance (Khan, Khanal, 2024). Ensemble models trained on segmented 2D slices showed high porosity prediction accuracy, which was validated using experimental helium pycnometry data (Kalule et al., 2023). Furthermore, convolutional neural networks trained on synthetic porous media datasets produced low prediction errors for porosity, permeability, and tortuosity, significantly speeding up the estimation process compared to traditional numerical methods (Graczyk, Matyka, 2020). Other advanced techniques, such as self-supervised learning with CNN-Transformer hybrids, obtained faster and more reliable predictions despite the limited training data (Iklassov et al., 2022). In some fields, such as additive manufacturing, adaptations of ML-based workflows can be seen, such as porosity estimation by ANNs from synthetic and experimental CT data in 3D-printed components (Mohammed et al., 2023).

It should be noted that porosity measurements from laboratory methods, such as gas porosimetry or pycnometry, are indeed relatively simple and cost-effective, being deeply established in traditional petroleum studies. However, these experimental methods require physical sample preparation, including hydrocarbon extraction, which can be time-consuming (up to several weeks), and may lead to the destruction of fragile rock samples (e.g., bituminous sandstones). Digital rock physics traditionally calculates porosity based on pore-space segmentation of high-resolution μCT images, which is feasible only for small-sized samples. Standard core plugs, widely used in petroleum research, are typically imaged at lower resolutions insufficient for detailed pore segmentation, thus limiting direct porosity calculation. By applying advanced machine learning models to predict porosity directly from standard low-resolution μCT images, we can utilize vast datasets of existing samples, overcome limitations related to fragile or complex rock types, and potentially extend these predictive capabilities to full-size cores due to relatively small differences in imaging resolution. Developing methods to reliably predict porosity from μCT images in meso-scale addresses critical gaps in reservoir characterization and enhances the efficiency and applicability of digital core workflows.

The objective of this study is to predict the experimentally measured open porosity of reservoir rocks using μCT images of standard core plugs. Specifically, the research aims

to develop and apply deep learning models based on the shifted window (Swin) transformer architecture for accurate porosity prediction using μCT images, while addressing challenges posed by unresolved porosity and image artifacts. Additionally, the study intends to investigate the efficacy of specialized models designed for various rock types compared to a universal model for predicting porosity across a variety of lithologies. This research seeks to provide a more efficient, reliable, and scalable solution for porosity prediction, advancing digital core analysis in the petroleum industry.

2. Methods

2.1. Dataset preparation

A total of 136 standard core plug samples were used for this study: 87 carbonate samples and 49 sandstone samples, representing approximately 64% and 36%, respectively. Every standard sample was drilled perpendicular to the wellbore direction and had a height and diameter of 30 ± 1 mm. The core plugs were scanned using the General Electric Phoenix v|tome|x S 240 micro- and nano-focus X-ray computed tomography system (Germany), offering a resolution of 34.6–38.0 μm. During the scanning process, the following key analytical parameters were maintained: an X-ray current of 150 μA, a voltage of 150 kV, 1200 projections with an averaging factor of 3, and an exposure time of 200 ms per projection. The resulting projections were reconstructed into high-resolution 3D models using phoenix datos|x reconstruction software.

The open porosity of each sample was measured using a gas volumetric method with the PIK-PP device (Russia) to provide benchmark values. The boxplot (Fig. 1A) compares the porosity distribution of sandstone and carbonate plugs in the sample set. In the sandstone samples, the median porosity is 16.3%, with an interquartile range (IQR) from 14.3% to 17.7%, indicating relatively low variability. Sandstone has a porosity range from a minimum of 10.1% to a maximum of 24.3%, with a standard deviation of 2.9%, indicating that most sandstone samples are closely grouped around the median porosity value, and sandstone tends to have a more uniform pore network. In contrast, carbonate samples were found to be highly variable, with a median porosity of around 10.1%, an IQR from 6.9% to 13.5%. The values vary from 0.9% to a maximum of 26.1%, with a rather high standard deviation of 6.0%. It indicates that this greater variability is due to the intrinsic heterogeneity of carbonate rocks, which possess complex pore systems with structural variability due to varying pore types, such as interparticle and vuggy pores.

The histogram of open porosity for the overall set (Fig. 1B) shows a bimodal distribution with distinct peaks at about 10% and 16%. This pattern parallels the two main rock types, sandstone and carbonate. The first peak near 10% primarily corresponds to carbonate samples, which exhibit lower porosity values, while the second peak near 16% corresponds to sandstone samples, which generally have higher porosity. Porosity ranges from 0.9% to 26.1% in the sample set, indicating significant variation. The wider spread for carbonate samples gives the histogram a long tail on the lower end, while sandstone samples correspond with the peak and upper range. This distribution highlights the contrasting geological properties of the two rock types, with sandstone maintaining a relatively stable porosity value, while carbonate appears to have a wider range because of its more heterogeneous nature.

2.2. Pre-processing of μCT Images

The μCT scanning and reconstruction produced high-resolution 3D images (~800×800×950 voxels, 16-bit depth) that captured both the rock samples and surrounding air. The resulting volumes were converted into TIFF stacks. Then we applied Tiff Stack Access module from the Avizo software package to normalize the slices through the stack. This ensured that all slices had consistent mean intensity and standard deviation and enhanced uniformity across the stack preparing it for machine learning analysis.

Despite these correctives, edge artifacts inevitably formed near the edges of the core plugs. These artifacts appeared as darkened voxels, particularly noticeable at the top and bottom ends of the cylindrical samples. These darkened areas will always result in over-segmentation when porosity is segmented in these areas, misclassifying non-porous areas as porous. Additionally, there may be some variation in the core plugs' sizes, with heights and diameters varying within a comparable range of 29 and 31 mm. Also, the scanning

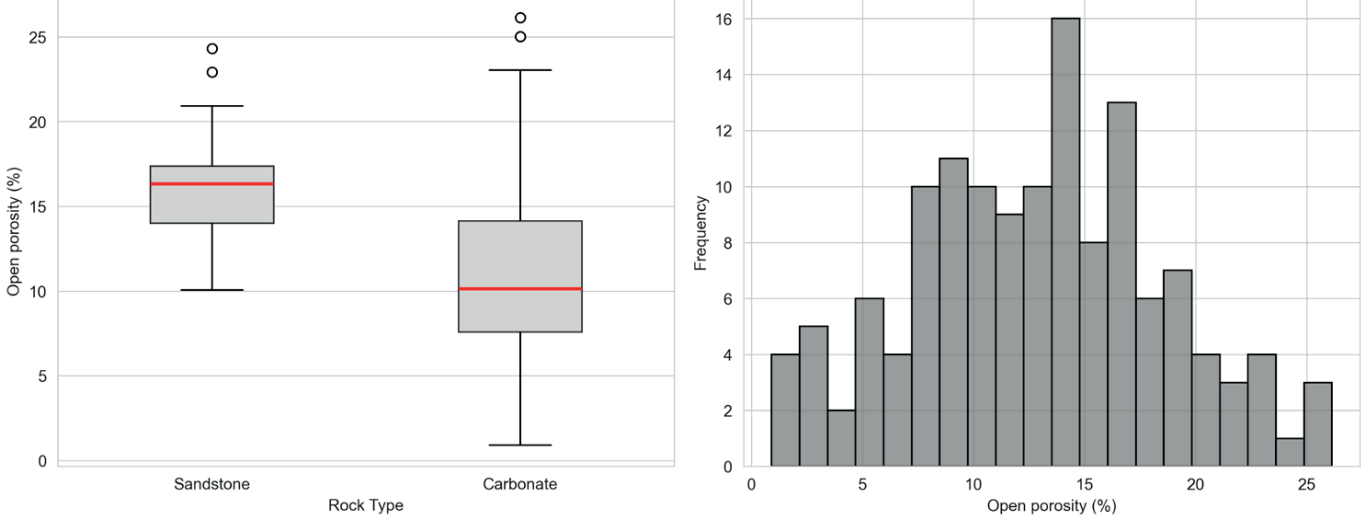

Fig. 1. Open porosity distribution of sandstone and carbonate samples: the boxplot A illustrates the distribution of porosity for two rock types: sandstone and carbonate; the histogram B represents the overall porosity distribution for the entire samples collection

resolution for each sample can differ slightly between 34.6 μm and 38.0 μm. These variations in both physical size and resolution present significant challenges when defining a standard cylindrical region of interest (ROI) for machine learning-based porosity prediction.

If a fixed-size cylindrical volume, such as 30 mm in diameter and height, is used for segmentation, there are two major risks. First, the selected volume may extend beyond the actual sample boundary, especially in cases where the sample diameter is smaller than the defined ROI. Second, darkened edge regions contain artifact-induced grayscale variations that do not accurately represent the rock's porosity, and their inclusion can result in significant errors during machine learning predictions. Surface flaws like chipping or irregularities can also be a problem because they can extend the darkened artifact zones into the sample's interior.

Analysis of the dataset revealed that the darkened areas typically extend 0.8–1.4 mm from the top and bottom surfaces of the core in the vertical direction (depending on the precise scanning resolution). The horizontally affected areas are typically narrower, but still significant. Therefore, to ensure that these problematic areas are excluded while maintaining a consistent dataset, an optimal and universal approach was developed. A cylindrical ROI of 737×737 voxels in diameter and height was empirically defined, ensuring that the ROI would be contained within the boundaries of even the smallest core and would exclude edge areas affected by artifacts, regardless of the specific sample size or resolution (Fig. 2).

Application of a cylindrical ROI guarantees that the grayscale values within the chosen volume stay constant and are unaffected by scanning artifacts, while also preserving the original 3D image quality and resolution. A cylindrical shape matches the geometry of the core plug and maximizes the amount of volumetric information available for machine learning, which means that the extracted features quite accurately reflect the physical sample used for open porosity measurements. Although the use of cubic ROIs simplifies certain machine learning tasks, this leads to the exclusion of significant parts of the sample. Additionally, the corners of the cube can contain areas susceptible to artifacts, which can lead to inaccuracies in the predicted porosity.

2.3. Labeling of 2D stack images

Machine learning applied to 2D images is more efficient in terms of computational power and processing time compared to 3D CNN models. Therefore, a key step in this study is assigning open porosity values to individual 2D images in the μCT stack. Since 3D porosity (bulk porosity) is defined as the ratio of the total pore volume to the total object volume, it can be expressed through the average 2D porosity (areal porosity) of parallel slices using the stereological principle, also known as the Cavalieri principle (Ambartzumian, 1990):

$$\varphi_{3D} = \frac{1}{N}\sum_{i=1}^{N} \varphi_{2D}\, z(i), \quad (1)$$

where φ_{3D} is bulk porosity and $\varphi_{2D}z(i)$ represents the areal porosity at the i-th slice along the height z. This approach assumes that all slices are parallel, have uniform thickness, and share a consistent region of interest (ROI), ensuring that slice area remains constant throughout the stack. Importantly, this equivalence between bulk and average 2D porosity holds regardless of whether the material is isotropic or anisotropic, as the stereological principle does not require isotropy.

A major issue arises due to the μCT resolution used for standard core plugs (ranging from 34.6 μm to 38 μm), which is insufficient to accurately resolve smaller pores that are prevalent in typical reservoir rocks. This results in what is commonly referred to as unresolved porosity (Alqahtani et al., 2022; Karimpouli, Kadyrov, 2022; Shah et al., 2016) or mixels (Brenne et al., 2021; Kadyrov et al., 2020), where sub-resolution pores contribute to the overall porosity but are not clearly distinguishable in the μCT images. If a fixed segmentation threshold is applied uniformly across all samples, the resulting porosity distribution function will be inaccurate and vary significantly depending on whether the sample has larger, well-resolved pores or is denser with finer, unresolved pores. Moreover, adjusting the segmentation threshold for different samples further complicates the process, as it directly influences the 2D porosity distribution function along the stack. The threshold sensitivity depends on both pore size and the X-ray absorption properties of the minerals comprising the reservoir rock.

To address this issue, an original approach was developed, termed Segmentation of Unresolved Pores via Experimental Reference (SUPER). The essence of this method lies in segmenting dark voxels that contain information about pores in such a way that ensures the total porosity of the segmented 3D volume accurately matches the experimentally measured open porosity. This approach ensures that for each sample, an individualized porosity distribution function along the stack is obtained. As a result, each 2D image is assigned a corresponding 2D porosity value that accounts for a precisely defined grayscale range representing unresolved porosity, tailored to the specific properties of the sample. The SUPER method overcomes the limitations of applying a fixed-value thresholding across all samples by adapting the segmentation process to the unique characteristics of each core plug. This ensures that the obtained 2D porosity values reflect the true internal pore structure of the rock, independent of resolution constraints and variations in pore size. Using the experimentally measured open porosity as a reference, the approach offers a reliable way to produce precise and reliable image labeling for the following training of ML models.

The labeling method applied to the μCT stacks results in assigning a specific porosity value to each 2D image within the stack, which ensures that the average of all 2D porosity values equals the experimentally measured 3D porosity of the

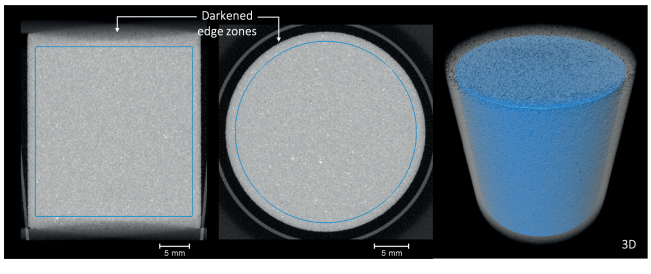

Fig. 2. Cylindrical region of interest (ROI) selection for μCT images of core plugs. The blue cylindrical ROI represents the standardized volume used for porosity prediction, which ensures that only unaffected interior regions of the sample are included in the analysis.

core sample. This creates a basis for an accurate representation of porosity across the entire dataset, accommodating the unresolved pore regions and the specific characteristics of each core sample. Figures 3 and 4 illustrate the porosity distribution along the stacks of a sandstone and a carbonate sample, respectively. In the case of sandstone, the porosity distribution remains relatively uniform across the entire stack, fluctuating only slightly. The 3D porosity of the sandstone specimen is 16.5% and is almost consistent with fluctuating 2D porosity values on the slices. This uniformity indicates a homogenous pore structure typical in sandstone with a stable pore network and with minimal fluctuations along the vertical axis. On the other hand, the heterogeneity in the porosity distribution of the carbonate core throughout the stack is considerable. The porosity decreases greatly from the top to the bottom of the core, indicating highly variable sizes and distribution of pores.

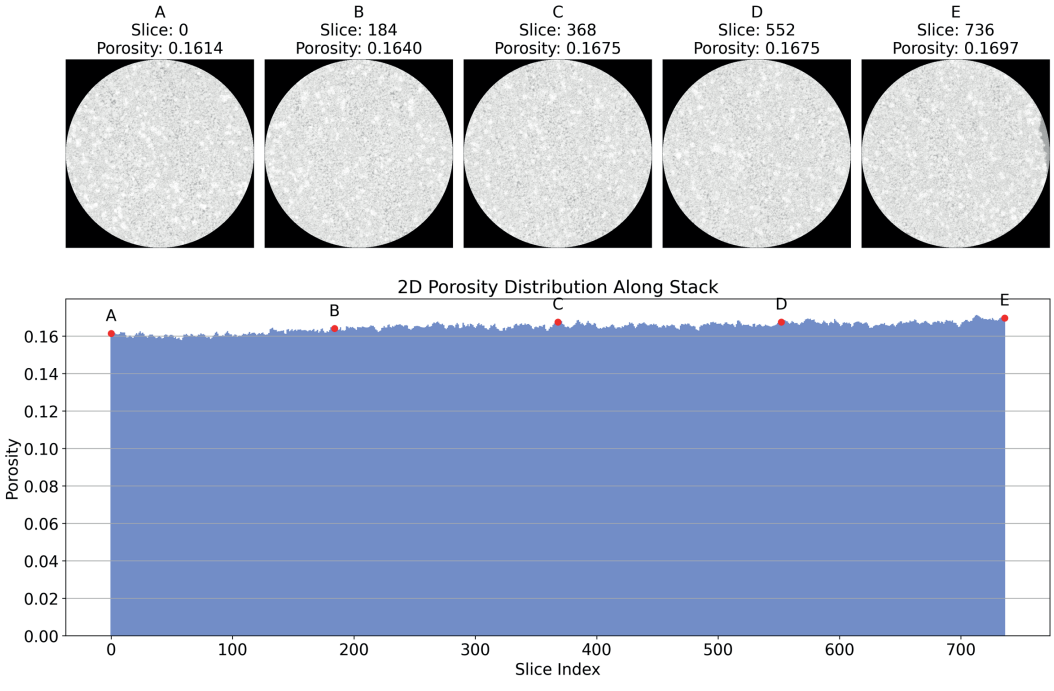

Fig. 3. Porosity distribution along the stack of a sandstone sample. The top row shows selected slices (A to E) at different positions in the stack, with their corresponding 2D porosity values. The bottom plot illustrates the uniform porosity distribution along the stack, with minor fluctuations. The average 2D porosity along the stack corresponds to the experimentally measured open porosity of 16.5%.

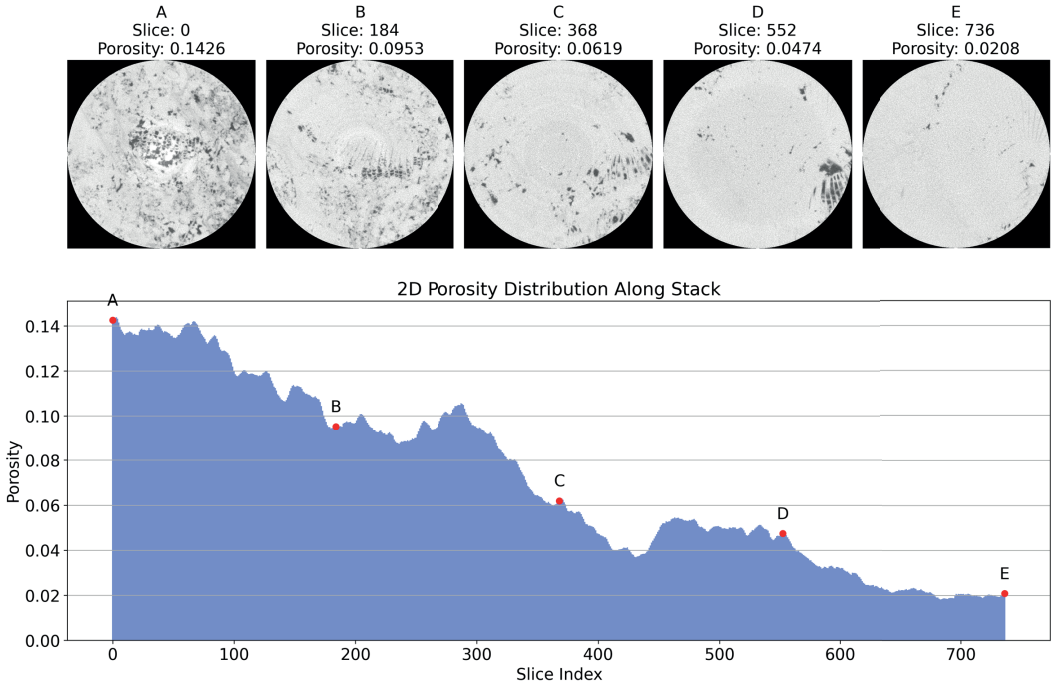

Fig. 4. Porosity distribution along the stack of a carbonate sample. The top row displays selected slices (A to E) at different positions in the stack, highlighting significant heterogeneity in pore structure. The bottom plot demonstrates the decreasing trend of 2D porosity along the stack, reflecting the heterogeneous nature of carbonate reservoir. The average 2D porosity along the stack corresponds to the experimentally measured open porosity of 7.21%.

The measured 3D porosity of the carbonate is 7.21%. The large difference in 2D porosity in the various slices reveals the complex pore structure typically found in carbonate rocks. This heterogeneity is a result of the different depositional environments and diagenetic processes that control the modes of formation, producing a variety of pore types, distributed throughout carbonate rocks.

Thus, a comprehensive dataset was created, consisting of a total of 100,232 images (136 core plug samples, each containing 737 slices per stack) with 16-bit depth, where each image is assigned a specific 2D porosity value. This dataset includes 64,119 images of carbonates and 36,113 images of sandstones, and provides a basis for training machine learning models for porosity prediction on μCT images.

2.4. Model architecture

The Swin Transformer architecture has been employed in this study for supervised learning (Liu et al., 2021). The Swin Transformer is a shift from the normal convolutional neural networks (CNNs) by introducing a self-attention mechanism for modeling the long-range dependencies in image data. In contrast to standard vision transformers that work on the whole image at once, the Swin Transformer introduces a hierarchical structure by splitting the input image into smaller overlapping patches, or windows. Local self-attention is computed within each window followed by a shifting mechanism realizing communication between the neighbor windows at the next layer. Therefore, this is the most efficient technique with which to harness local as well as global features.

One of the Swin Transformer's strong points is its ability to work with high-resolution data while still being computationally efficient. The architecture is hierarchical, and just like CNNs, as the resolution goes down with depth, self-attention picks up and models complex relationships between distant regions in the image. This feature is particularly suited for applications with large image datasets such as μCT stacks of core plugs, where the modeling of fine pore structures as well as long-range patterns is quite crucial.

In addition, the Swin Transformer has somewhat mitigated the restrictions placed on the classical attention method with non-overlapping windows by imposing continuity of the features across window boundaries. This design increases the contextual modeling capacity and thus aids in its general performance on the unseen data. With the stated attributes, the Swin Transformer has exhibited state-of-the-art results in many computer vision tasks, such as image classification (Wang et al., 2023), object detection (Gong et al., 2022), and semantic segmentation (Xu et al., 2021). Thus, it is a reasonable candidate for porosity prediction in challenging geological samples.

The Swin Transformer models used in this project were implemented in the PyTorch framework, which is flexible enough to deal with complex neural networks and large-scale datasets (Paszke et al., 2019). One of the major problems faced during this study was the relatively small size of the dataset – 100,232 images with 16-bit depth. Though this might look quite huge, when training deep models like the Swin Transformer from scratch, one generally requires much larger datasets, usually millions of samples, in order to ensure reasonable robustness in generalization. Because of this shortfall, it was decided that transfer learning would be an appealing option. It would take advantage of pre-trained weights of large diverse datasets and would thus give the model a strong starting point instead of having to learn everything from scratch (Kadyrov et al., 2024).

For the purpose of this project, the pretrained weights from ImageNet 1K were incorporated since the dataset consists of more than a million images with 8-bit depth, distributed into 1,000 classes, with a broad set of rich features (Jia Deng et al., 2009). Although the depth of images varies, high-level features obtained by learning from ImageNet 1K are transferrable when processed well to assure feature extraction for prediction of porosity in 16-bit images from μCT. In this way, it reduced the impact of data scarcity while achieving better performance.

Another notable challenge was the shape of the core plug images. Normally, a computer vision task would require performance on rectangles or squares. Core plug μCT images, however, are circular with an unrelated black background surrounding the core region. Circular masking was therefore applied to ensure that only the core region contributed to the training process. The mask was uniformly applied across all the images and preserved the cylindrical shape of the core plugs, enhancing consistency in input data.

Additionally, normalization was applied to bring the pixel intensity values to a common scale. Given that μCT images have a 16-bit depth with intensity values from 0 to 65,535, normalization becomes imperative to compress all these values into the range 0–1. This step also fundamentally improved the robustness of the training process and the network under variable intensity – as affected by different scanning conditions or material heterogeneity.

Finally, to introduce diversity into the training set and, hence, improve the generalization capacity of the model, data augmentation was used. Random rotations of 0°, 90°, 180°, and 270° were applied to simulate different orientations of the core plugs. Since they don't have a preferred orientation in practice, the augmentation helped the model adapt to orientation-invariant features.

After masking and normalization, we resized the circular region (originally 737 pixels in diameter) down to a 224×224 pixels. This step assures that each input is the same size for the Swin Transformer. We used cv2.resize function with cv2.INTER_AREA interpolation mode from OpenCV library for resizing. Conceptually, every pixel of the resized image represents an aggregated region of the original image. Area interpolation essentially calculates the pixel value of the resized image by taking the average of all the input pixels that come into the region of the output pixel. This method keeps more of the information than the simpler methods, reducing aliasing effects and keeps edges fairly decent. Unfortunately, downscaling filters are bound to bring spatial information together: minor structures seen in an original high-definition image would probably be shown less distinctively or might be lost altogether (Fig. 5). However, that is the typical compromise taken in most deep learning workflows to balance memory limitations with the receptive field requirements of convolution-based or transformer-based architectures.

The Swin Transformer architecture used in this study begins by dividing the input image into non-overlapping patches, each of size 4x4 pixels (Fig. 6). The Swin Transformer architecture used in this study first performs a linear embedding of these patches into feature vectors to generate the first input

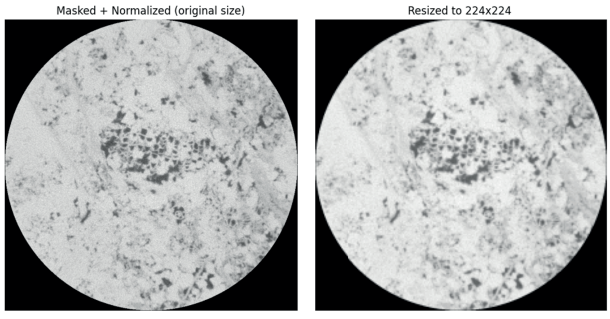

Fig. 5. Comparison of μCT images before and after resizing. The left image shows the original masked and normalized μCT image at its full resolution (737×737 pixels), while the right image depicts the resized version at 224×224 pixels. The resizing step averages pixel intensities over the input region corresponding to each output pixel. In this case, each pixel in the resized image aggregates information from approximately a 3.3×3.3 pixel area of the original image. As a result, fine features smaller than this scale are likely to be smoothed or lost. This trade-off balances computational efficiency with structural detail retention for training deep learning models.

tokens to the model. The first stage of the Swin Transformer processes the tokens through self-attention mechanisms within local windows of size 7×7 for processing local features while keeping the computation cost low. The introduction of shifted windows is one of the innovations of this architecture. These windows introduce overlapping between adjacent windows in the upper layers and subsequent layers, allowing them to interact with each other through a continuum of features across window boundaries.

The system is hierarchical and consists of four main stages. It progressively reduces the resolution of feature maps similar to the CNNs. Each stage contains multiple Swin Transformer blocks, and within each block, the self-attention mechanism is augmented with multi-head attention, layer normalization, and a multi-layer perceptron (MLP) for feature transformation. The first stage uses 96 feature channels, which double at each subsequent stage to 192, 384, and 768, respectively.

The final output of the Swin Transformer is a feature map from the last stage, which is subjected to global average pooling to aggregate spatial information. The pooled output is then fed into a fully connected (fc) regression layer specifically designed for this study to predict scalar values denoting porosity. This regression layer replaces typical ImageNet classification layers to enable the model to render continuous outputs in place of discrete classes.

To improve performance, ImageNet-1k pre-training weights were leveraged to provide a strong initialization to the network. This means that the pre-training gives the model a strong start in learning low-level and mid-level feature representations and then fine-tuning them for porosity prediction. The fully connected layers use a ReLU activation function to help the network capture nonlinear patterns, while dropout layers are added to avoid overfitting.

Overall, the Swin Transformer combines the advantages of hierarchical feature extraction, self-attention mechanisms, and dislocated windows to process high-resolution μCT images efficiently while capturing both fine pore structures and long-range dependencies. Such an architecture seems appropriate for a complex task like porosity prediction from geological core plug samples.

2.5. Workflow

This research workflow is subdivided into two major branches. The first branch involves the training of a single universal Swin Transformer model on the entire dataset of 100,232 images to predict porosity across rock types. The second branch pertains to the training of two specialized Swin Transformer models, each trained to predict porosity for either carbonates or sandstones. In the first case, the dataset consisted of 64,119 carbonate images, while the second dataset included 36,113 sandstone images. The experimental setup for model training included an NVIDIA Quadro P5000 GPU, Windows 10, PyTorch 1.13.1, CUDA 11.7, and Python 3.8. All experiments were conducted for 10 epochs.

Figure 7 illustrates the workflow process applied to each dataset. As described in Section 2.4, the original 16-bit μCT images undergo circular masking, normalization to the 0–1 range, and resizing to 224×224 pixels prior to training. After that, each dataset is split into training (80%) and validation (20%) subsets, with data augmentation by random rotations applied to the training set only. The Swin Transformer is then fine-tuned using ImageNet-1k pre-trained weights, minimizing mean squared error (MSE) through the AdamW optimizer (learning rate of 10^{-4}, batch size of 16, and 10 training epochs). Model performance on the validation and testing sets is monitored via MSE, mean absolute error (MAE), and the coefficient of determination (R^2) (Zollanvari, 2023).

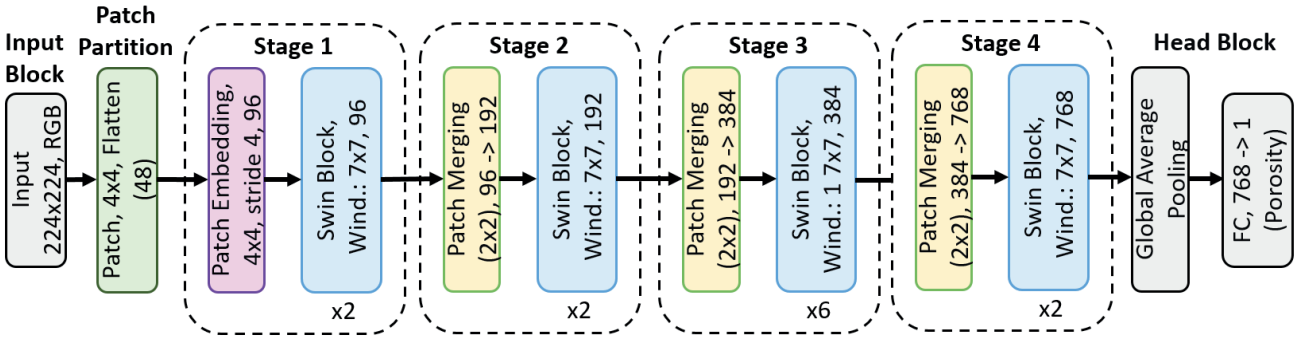

Fig. 6. Block diagram of the Swin Transformer (Swin-Tiny) architecture used in this study for porosity prediction. The network takes a 224×224 RGB image as input. The Patch Partition step divides the image into non-overlapping 4×4 patches, which are projected to 96-dimensional tokens via patch embedding. The backbone consists of four hierarchical stages with window-based self-attention using a 7×7 window size and shifted windows in alternating blocks. Stage 1, Stage 2, Stage 3, and Stage 4 contain 2, 2, 6, and 2 Swin Transformer blocks, respectively, with channel dimensions of 96, 192, 384, and 768. Between stages, Patch Merging (2×2) reduces spatial resolution while increasing the number of channels. Finally, global average pooling and a fully connected regression layer output a single porosity value.

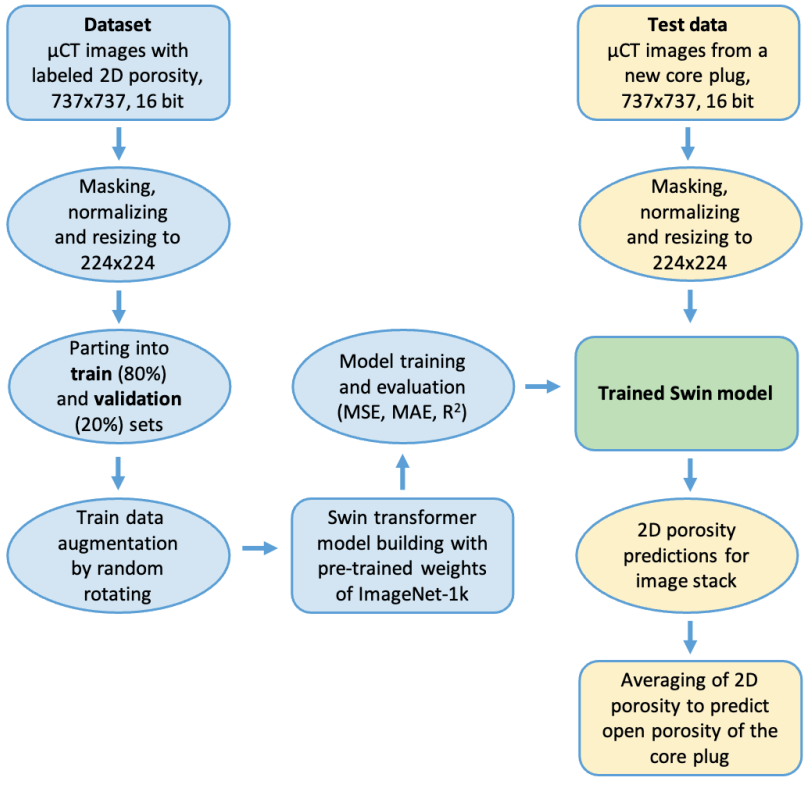

Fig. 7. Workflow schematic for porosity prediction using Swin Transformer models. Blue elements represent the training and evaluation track, while yellow elements indicate the test-phase pipeline for new core plugs.

After training and validation were completed, the model was tested on 10 new core plug samples, which consist of five sandstones and five carbonates. In this application phase, µCT stacks of the new samples underwent the same pre-processing steps as the training data, including masking, normalization, and resizing. The model then predicts porosity slice by slice, which is averaged to give an estimate of the overall 3D porosity. This strategy allows for obtaining a detailed view of porosity distribution across the core plug and enables verification of whether average 2D predictions agree with experimental measurements. By comparing the performance of the universal model and specific models, the workflow helps to determine how different training strategies affect porosity prediction for heterogeneous geological materials.

3. Results

3.1. Training and validation results

An analysis of the training dynamics across the three models reveals each effectively adapts to the features of each dataset (Fig. 8). The universal model trained on the entire dataset of 100,232 images quickly captured the overall feature distribution, completing the training process in 22 hours and 38 minutes; the loss values rapidly decreased during the first few epochs, which is a strong indication that this model learned a good representation of the data. The sandstone model showed higher loss values at the onset of training, likely due to the complexities of sandstone features. The total training time for this model was 6 hours and 30 minutes. Nevertheless, it demonstrated a significant drop in loss from epoch one to two, indicating that the model quickly adjusted itself to the features specific to sandstone. The improvement that followed, however swift, brought this model back to a loss level comparable to that of the universal model, even from a state where it was heavily challenged. The carbonate model demonstrated an intermediate behavior: it started at a loss lower than that of the sandstone model but higher than that of the universal model, completing the training process in approximately 11 hours and 38 minutes. After some small fluctuations in the mid-epochs, it maintained a steady loss decrease toward low values, suggesting some gradual resilience to the inherent variability among carbonate samples. Regardless of differences in early convergence rates and initial loss values, all three models were able to reach similarly low loss values by the final epoch.

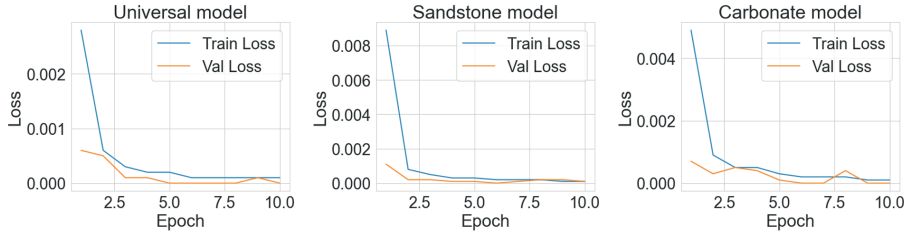

Fig. 8. Training and validation loss evolution across epochs for the universal model (left), the sandstone model (center), and the carbonate model (right)

The final validation output suggests very good performance by these three models in all cases of low errors and high values of determination coefficients (Table 1). The universal model trained on the combined dataset had an MSE of 0.0000, MAE of 0.0039, and an R^2 of 0.9920, showing its strong ability to generalize to both sandstone and carbonate samples. In comparison, the specialized sandstone model had a higher MSE at 0.0001 and higher MAE at 0.0049, with a lower value of R^2 (0.9459) than the universal model; this is indicative of the increased complexity that characterizes sandstone features. On the other hand, the carbonate model performed excellently with an MSE of 0.0000, an MAE of 0.0039, and the highest R^2 of 0.9935, implying that it has effectively dealt with the features of carbonate rocks. The metrics highlight the strength of transfer learning and the Swin Transformer architecture in predicting porosity using μCT images at this core plug scale, while also revealing slight differences in model performance that can be attributed to distinctions in rock features.

Model	MSE	MAE	R^2
Swin_universal	0.0000	0.0039	0.9920
Swin_sandstone	0.0001	0.0049	0.9459
Swin_carbonate	0.0000	0.0039	0.9935

Table 1. Final validation metrics for all models

3.2. Testing results

After completing the training and validation phases, the performance of the developed models was evaluated using a distinct test dataset composed of 10 standard core plug samples from reservoir rocks. It contained five sandstone samples and five carbonate samples, which had not been seen by the models during the training or validation process. Scanning and pre-processing of core plug samples followed the pipeline adopted in Section 2.5 to remain consistent with the training dataset. Results of the testing exercise are given in Table 2.

The Swin_universal model exhibited a balanced performance across all samples, producing porosity predictions reasonably close to measured values, sometimes with some discrepancies. For sandstone samples, Swin_universal showed sufficient agreement with measured values in many cases: e.g., Sample 2 had a prediction of 25.19% quite close to the measured porosity of 22.76%. Its predictions for Samples 1 and 5 were off significantly from the measured results, while for carbonate samples, the Swin_universal presented an appropriate and consistent prediction, especially near porosity values for Samples 6, 7, and 9.

For the sandstone subset, the Swin_sandstone model has performed acceptably with predictions having relatively low error values, especially for Samples 2 and 4. However, on the other hand, the predictions made by this model for carbonate samples were not as accurate, further reflecting its specialization to sandstone textures, and its limited capacity to generalize to carbonate rocks. In contrast, the Swin_carbonate model performed excellently and very accurately in predicting carbonate porosity, particularly for Samples 8, 9, and 10, while underestimating sandstone samples, which was expected due to its specialization in carbonate textures.

The performance parameters of the three Swin Transformer models on the test dataset are shown in Table 3. This includes overall results and specific evaluations for sandstone and carbonate samples. In addition to traditional error metrics such as MSE, MAE, and R^2, we also included the Mean Absolute Percentage Error (MAPE) for evaluating the relative error between measured and predicted porosity values:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%, \quad (2)$$

where y_i represents the true porosity and $\hat{y}_i$ the corresponding prediction. This metric provides a normalized measure of prediction accuracy to compare performance across samples with different porosity scales (Kim, Kim, 2016). Swin_universal demonstrated the best results across the entire test dataset, having an MSE of 13.11, MAE of 2.56, R^2 of 0.358, and MAPE of 13.13%, showing that it was moderately accurate across all rock types. In contrast, both Swin_sandstone and Swin_carbonate yielded higher errors with negative R^2 values, which suggests limited generalization. When evaluating the sandstone samples separately, the Swin_sandstone model outperformed the other models with an MSE of 11.06, MAE of 2.38, and the highest R^2 of 0.014, which means its specialization in predicting porosity in sandstones. However, its performance deteriorated significantly on carbonate samples, noting the challenges of cross-rock-type prediction. The results of Swin_carbonate were effective in evaluating carbonate samples, achieving the best performance

Sample No	Rock type	Measured porosity (%)	Swin_universal porosity (%)	Swin_sandstone porosity (%)	Swin_carbonate porosity (%)
1	sandstone	22.12	12.81	15.16	10.59
2	sandstone	22.76	25.19	24.22	10.22
3	sandstone	14.72	17.17	15.31	20.4
4	sandstone	16.2	16.13	15.03	17.24
5	sandstone	16.21	13.51	17.94	15.82
6	carbonate	13.74	15.63	14.95	14.67
7	carbonate	13.67	14.89	14.99	14.73
8	carbonate	25.57	21.2	16.94	23.22
9	carbonate	23.76	23.94	17.95	24.83
10	carbonate	24.02	23.69	17.13	24.59

Table 2. Measured and predicted porosity values for the test dataset of sandstone and carbonate core plug samples using Swin transformer models

Model	MSE	MAE	R^2	MAPE (%)
Swin_universal (all test dataset)	13.11	2.56	0.358	13.13
Swin_sandstone (all test dataset)	21.42	3.58	-0.049	16.51
Swin_carbonate (all test dataset)	33.27	3.72	-0.629	18.52
Swin_universal (sandstone)	21.17	3.39	-0.889	17.30
Swin_sandstone (sandstone)	11.06	2.38	0.014	11.95
Swin_carbonate (sandstone)	64.74	6.24	-4.776	30.93
Swin_universal (carbonate)	5.05	1.74	0.820	8.95
Swin_sandstone (carbonate)	31.78	4.77	-0.131	21.07
Swin_carbonate (carbonate)	1.80	1.20	0.936	6.12

Table 3. Performance metrics of Swin transformer models on the test dataset for porosity prediction

with an MSE of 1.80, MAE of 1.20, R^2 of 0.936, and MAPE of 6.12%, which is indeed an indication of its strong capability for accurate prediction of carbonate porosity.

To understand these performance differences in greater detail, porosity distributions along the stack for sandstone and carbonate sample types in the test dataset are represented in Figures 9 and 10. These figures demonstrate the models' ability or inability to capture porosity variations along the core plug and highlight their respective strengths and limitations.

Figure 9 shows measured and predicted porosity distributions for a sandstone core plug (Sample 3 from Table 2). There are high fluctuations in the porosity predictions across the 2D images from both Swin_sandstone and Swin_universal models, which deviate from measured values on individual slices. Such high variability indicates a problem in the models' ability to correctly identify fine-scale features in the sandstone images, possibly due to the resolution limitations of the μCT imaging. The high-frequency oscillations in predicted porosity indicate that local variations in grayscale intensity may be misinterpreted as structural changes, leading to inconsistent estimates. Nonetheless, despite these fluctuations, the Swin_sandstone model predicts more reasonably both the general porosity trend and the mean level of porosity across the core plug, while the Swin_universal tends to deviate further from the measured values. This difference could imply that the specialized model, which is only trained on sandstone samples, is better suited for capturing the general porosity characteristics of this lithology even though it has difficulties with fine-scale variations.

Figure 10 presents the results for a carbonate core plug (Sample 6 from Table 2). Unlike the sandstone case, the Swin_carbonate model shows strong agreement with the measured porosity distribution and maintains stable predictions along the core length. Also, the large oscillations observed in the sandstone predictions are absent, meaning that the model effectively captures the overall porosity pattern without excessive local fluctuations. This stability suggests that the Swin_carbonate model successfully identifies and generalizes the characteristic pore patterns of carbonate rocks, which are typically determined by well-defined but complex pore networks. The Swin_universal model also provides reasonable estimates, though it tends to slightly overestimate porosity in certain regions. However, its predictions remain relatively stable, which reinforces the notion that carbonate samples are generally easier to model than sandstones, likely due to the more distinct texture and high contrast in grayscale intensities associated with pore spaces in carbonate lithologies.

The test results give an overall picture of the strong and weak points of each model. Among the models, the Swin_universal model performs the most balanced way in terms of both rock types, whereas the specialty models, Swin_sandstone and Swin_carbonate, yielded better accuracy for their own classes but have limited generalization to other lithologies. The benefits of using specialized models for different rock types become evident, while the flip side is a loss of generalization in a single universal model.

4. Discussions

Predicting experimentally measured open porosity of reservoir rocks based on μCT images of standard core plugs is a challenging task due to the complex pore structures and limitations of imaging resolution. In this study, Swin transformer models demonstrated a promising predictive capability, offering a data-driven approach to estimate open porosity directly from tomographic images. The Swin transformer's self-attention mechanism and hierarchical architecture enabled more effective modeling of both local and global features. This contributes to accurate predictions even with a relatively limited dataset. Unlike classical methods that rely on segmentation and geometric analysis, the proposed approach bypasses the need for pore segmentation and reduces potential errors associated with image artifacts and unresolved porosity. It should be noted, however, that this approach is focused on porosity estimation and does not account for flow-related properties such as permeability, which require accurate pore connectivity modeling through segmentation and flow simulations (Lavrukhin et al., 2021; Varfolomeev et al., 2019).

The test results indicated that Swin Transformer model performance was rock-type dependent. The Swin_universal model performed rather well in both lithologies, whereas Swin_sandstone and Swin_carbonate models achieved better metrics when applied to their respective rock but performed poorly when tested on the other lithology. This clearly indicates that an ensemble of specialized models, trained individually, will perform better than a single universal model for porosity prediction. Moreover, each of these models could be fine-tuned, not only based on broad rock types but also based on lithological and structural features specific to a single rock type. Implementing such an approach could be based on automated characterization and classification of reservoir rocks from μCT imaging of core plugs, as described in (Kadyrov et al., 2024).

The study identified several key challenges and limitations related to the prediction of porosity using Swin transformer

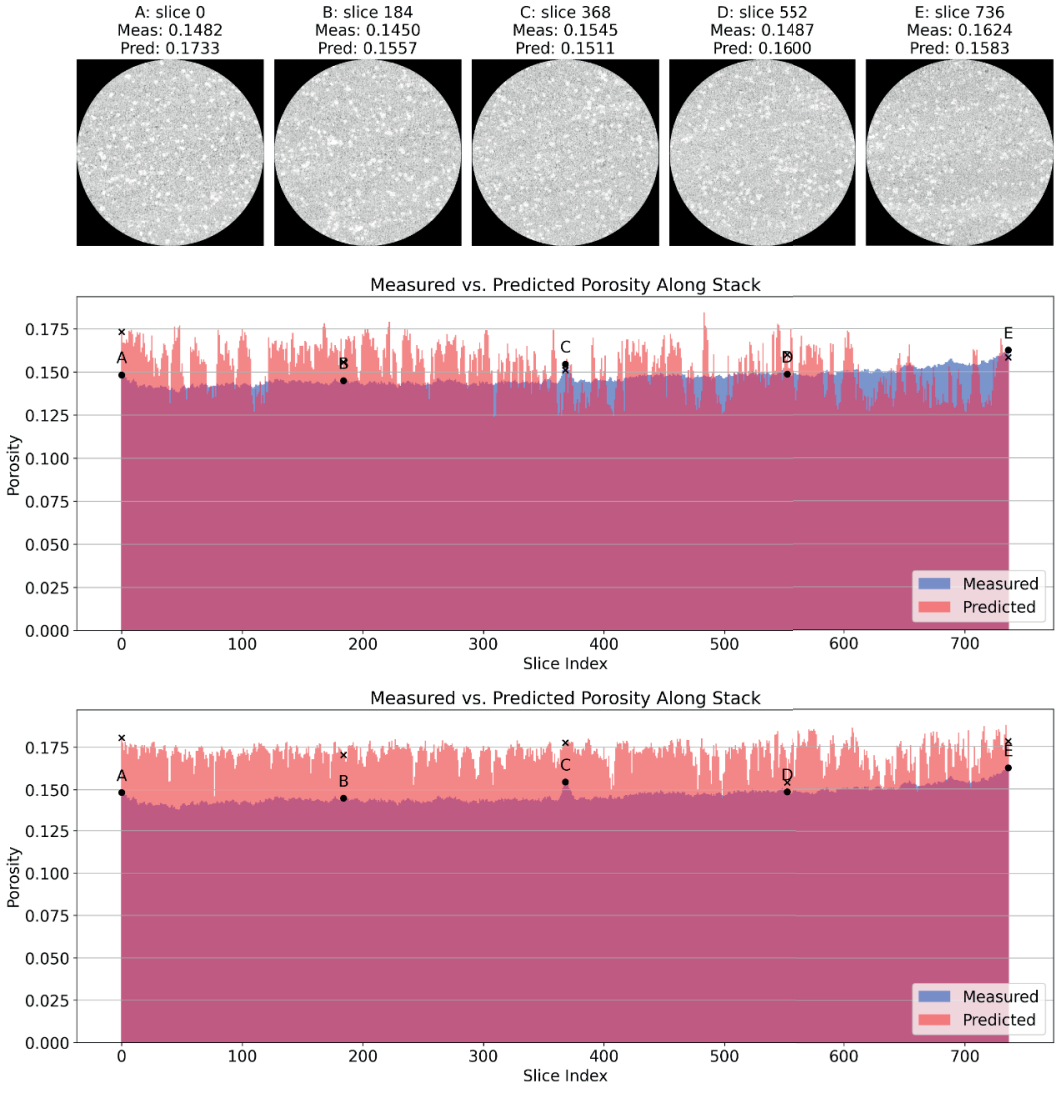

Fig. 9. Measured versus predicted 2D porosity along the stack of a sandstone core plug (Sample 3, Table 2). Both the Swin_sandstone and Swin_universal models show fluctuations due to resolution limits, with the Swin_sandstone model more closely following the overall porosity trend, while the universal model tends to slightly overestimate porosity in some regions.

models, particularly concerning the differences between carbonate and sandstone rock types. The models showed higher predictive ability on carbonate rocks compared to sandstone. This can be attributed to the differences in the porosity characteristics of these lithologies: in carbonates, the porosity is often primarily determined by larger pores, which are more easily detected by the models given the resolution of the μCT images used. Additionally, denser areas with nearly invisible porosity, formed by micrite or sparite, generally contribute less to the overall porosity and have stable porosity values. This characteristic makes carbonate rocks more suitable for accurate porosity prediction, because textural features manifest clearly as variations in grayscale intensity, allowing the models to capture these features with higher accuracy.

In contrast, the intergranular porosity structure of sandstones, especially fine-grained varieties, which contributes significantly to the measured porosity, tends to blend together at the resolutions used in this study. The fine pores often lack distinct textural features related to porosity, except for a few relatively large pores, which leads to fluctuations in the predictions for individual slices and worsens the overall prediction for the entire sample. The specialized Swin_sandstone model, which showed the best performance on sandstones, appears to rely more heavily on generalized texture features and compares them with those encountered during training.

This can be clearly observed in Figure 11, in which the performances obtained on Bentheimer sandstone (the first sample in the testing dataset) and Berea sandstone (used in the training dataset) are illustrated. Both types of samples have similar dimensions of large pores; however, the Swin_sandstone model fails to predict porosity values for Bentheimer. This suggests that in the absence of fine details of pore structures, the model leans on generalized textures, which may differ significantly from those seen during training. For example, the Berea sandstone is characterized by well-defined layering and abundant occurrences of sulfide minerals, features that are absent in Bentheimer. These variations in texture between different types of sandstone create challenges for the model, especially when it is faced with a new sample that exhibits a texture not well-represented in the training set.

The situation with low resolution is compounded by the fact that the original 737×737 images are further downscaled

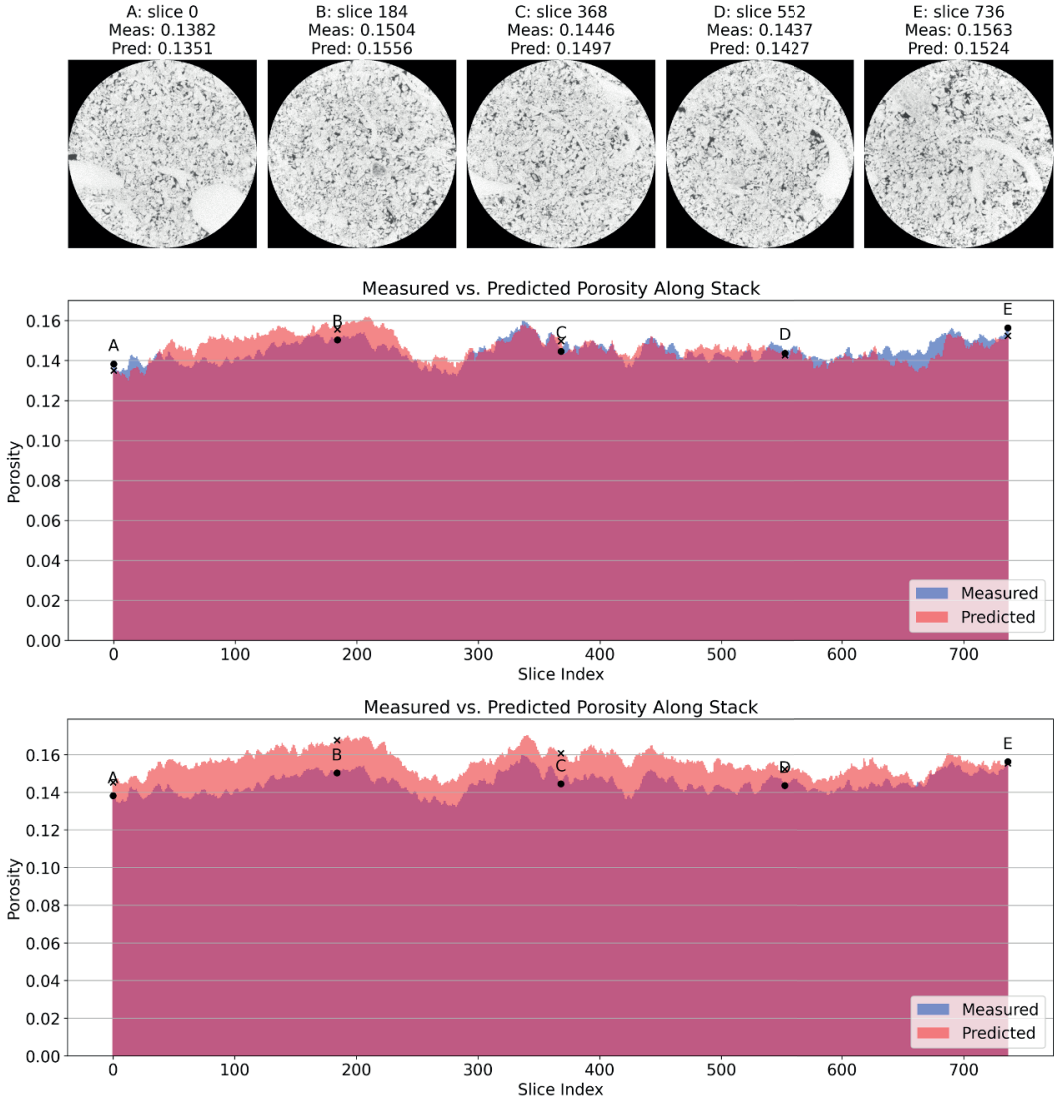

Fig. 10. Measured versus predicted 2D porosity along the stack of a carbonate core plug (Sample 6, Table 2). The Swin_carbonate model closely follows the measured porosity with minimal fluctuations, while the Swin_universal model tends to slightly underestimate porosity in certain regions.

to 224×224 for model input, resulting in a resolution of over 100 μm per pixel. Even when using Swin models with larger input sizes, such as 384×384, the resolution issue remains largely unresolved, while the computational cost of training such models increases significantly. This makes it difficult to train these models on local, modern computers (equipped with consumer-grade GPUs) and requires significantly extended training time. Therefore, future research may focus on optimizing the input resolution for the models, potentially increasing it to 2000×2000 pixels (in the case of scanning core plugs of sandstones with higher resolutions of 16–20 μm) or feeding fragmented high-resolution images to the model. Additionally, synthetic μCT datasets with known ground truth may help evaluate model sensitivity to resolution and noise in a controlled environment (Lavrukhin et al., 2025).

On the other hand, the models presented, even with such low input resolution, have demonstrated good performance on carbonates. When training specialized models for specific sandstone lithotypes, they may prove useful for predicting their porosity as well. This opens up possibilities for applying these models not only to standard images but also to tomographical images of full-sized cores, with resolutions approaching 100 μm. Therefore, advancing this concept could be beneficial for improving porosity predictions for full-scale cores and linking them with geophysical well logging data.

Another promising direction is the application of these models for porosity prediction on raw core plugs without hydrocarbon extraction. Since hydrocarbon extraction can take several weeks to several months, and standard research is only conducted afterward (during which many samples are discarded as unsuitable for further studies), a predictive system like this could save considerable time and effort. It is expected that light hydrocarbons will have minimal impact on pore contrast and the predictive capabilities of the models.

As for reservoirs containing heavy hydrocarbons and bitumens, in the case of terrigenous types, they often break down during extraction since hydrocarbons act as a cementing agent between the grains. Therefore, adapting the described approach to such raw core plugs (for example, by scanning samples before hydrocarbon extraction and experimentally evaluating them after extraction, in case the structure of the sample is preserved) could be a viable solution for assessing the porosity of such samples, thereby enabling more accurate reservoir assessments.

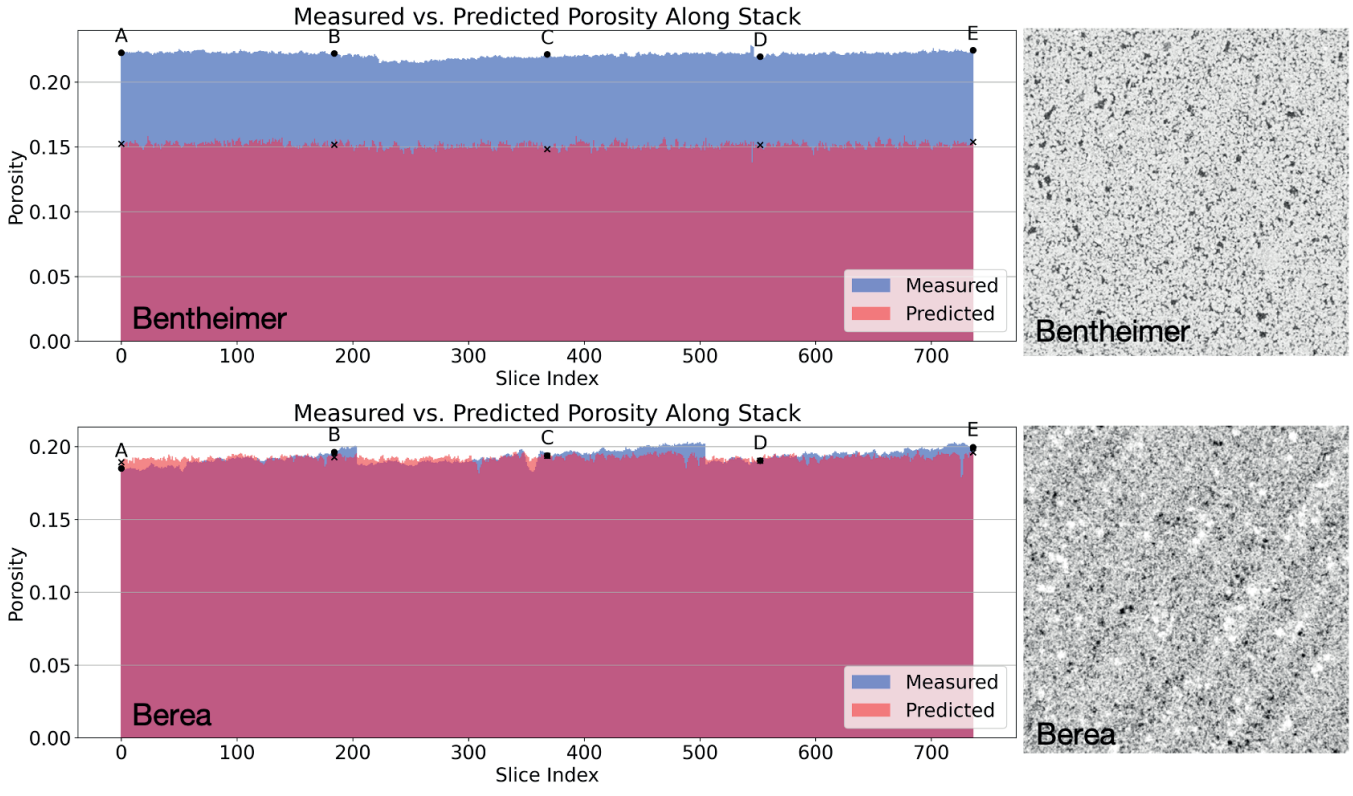

Fig. 11. Comparison of Measured vs. Predicted Porosity for Bentheimer and Berea Sandstones by the Swin_sandstone Model. The Berea sandstone was part of the training dataset, whereas the Bentheimer sandstone was included in the test dataset. Despite the similarity in the size of large pores in both sandstones, the Swin_sandstone model struggled to accurately predict the porosity of the Bentheimer sample, highlighting the model's challenges in generalizing to rock textures not encountered during training.

5. Conclusions

In this study, the Swin Transformer architecture was demonstrated to provide promising results for predicting experimental open porosity in core plugs based on μCT images. The proposed approach offers notable advantages by eliminating the need for explicit pore segmentation and effectively handling unresolved porosity. By leveraging the Swin Transformer's hierarchical self-attention mechanism, it successfully captures both localized and global features in 3D image stacks, leading to more reliable predictions compared to traditional methods that depend heavily on segmentation thresholding. Multiple Swin Transformer-based deep learning models were trained separately for carbonate and sandstone samples, as well as a universal model combining both lithologies. The results showed that specialized models achieved higher accuracy within their respective rock types, especially in resolving heterogeneous pore structures, while the universal model displayed robust performance across different lithologies. This indicates that an ensemble of specialized models can surpass a single model in capturing the distinct characteristics of various reservoir rocks.

A key limitation of the study is the insufficient resolution of μCT data, particularly for fine-grained sandstone samples, which prevents accurate differentiation of smaller pores and leads to fluctuations in slice-by-slice predictions. Future research should focus on improving image resolution, not only by enhancing μCT scanning techniques and integrating higher-resolution imaging methods to capture finer pore structures but also by optimizing input sizes for the model. Increasing the model's capacity to process high-resolution images, either through larger input dimensions or by implementing an image fragmentation approach, could help retain critical pore-scale details while maintaining computational efficiency. Applying this approach to raw core plugs before hydrocarbon extraction may accelerate the analysis workflow, offering rapid porosity estimates without the need for extended laboratory procedures. Additionally, extending the proposed approach to tomographic images of full-sized cores with resolutions approaching 100 μm could enhance porosity prediction at larger scales. Such improvements in accuracy and efficiency have direct implications for enhanced reservoir modeling, more reliable reserve estimation, and better-informed production strategies in the petroleum industry.

In conclusion, the presented Swin Transformer-based framework demonstrates the feasibility of using deep learning for automated, high-precision porosity prediction from μCT images of core plugs, offering a scalable solution that can be adapted to diverse reservoir conditions and contributing to the ongoing digital transformation of core analysis.

Acknowledgments

This paper is performed as part of the grant of the Tatarstan Academy of Sciences, provided to young candidates of science (postdoctoral fellows) for the purpose of defending their doctoral dissertation, conducting research, as well as performing their work duties in scientific and educational organizations of the Republic of Tatarstan within the framework of the State Program of the Republic of Tatarstan «Scientific and Technological Development of the Republic of Tatarstan» (Agreement No.20/2024-PD).

References

Alqahtani N., Alzubaidi F., Armstrong R.T., Swietojanski P. & Mostaghimi P. (2020). Machine learning for predicting properties of porous media from 2d X-ray images. *Journal of Petroleum Science and Engineering*, 184, 106514. https://doi.org/10.1016/j.petrol.2019.106514

Alqahtani N.J., Niu Y., Wang Y.Da, Chung T., Lanetc Z., Zhuravljov A., Armstrong R.T. & Mostaghimi P. (2022). Super-Resolved Segmentation of X-ray Images of Carbonate Rocks Using Deep Learning. *Transport in Porous Media*, 143(2), pp. 497–525. https://doi.org/10.1007/s11242-022-01781-9

Alyafei N., Raeini A.Q., Paluszny A., & Blunt M.J. (2015). A Sensitivity Study of the Effect of Image Resolution on Predicted Petrophysical Properties. *Transport in Porous Media*, 110(1), pp. 157–169. https://doi.org/10.1007/s11242-015-0563-0

Alzahrani M.K., Shapoval A., Chen Z. & Rahman S.S. (2023). Pore-GNN: A graph neural network-based framework for predicting flow properties of porous media from micro-CT images. *Advances in Geo-Energy Research*, 10(1), pp. 39–55. https://doi.org/10.46690/ager.2023.10.05

Ambartzumian R.V. (1990). Cavalieri principle and other prerequisities. In Factorization Calculus and Geometric Probability. Cambridge University Press, 286 p. https://doi.org/10.1017/cbo9781139086561.002

Andrä H., Combaret N., Dvorkin J., Glatt E., Han J., Kabel M., Keehm Y., Krzikalla F., Lee M., Madonna C., Marsh M., Mukerji T., Saenger E.H., Sain R., Saxena N., Ricker S., Wiegmann A., & Zhan X. (2013). Digital rock physics benchmarks—Part I: Imaging and segmentation. *Computers & Geosciences*, 50, pp. 25–32. https://doi.org/10.1016/j.cageo.2012.09.005

Blunt M.J., Bijeljic B., Dong H., Gharbi O., Iglauer S., Mostaghimi P., Paluszny A. & Pentland C. (2013). Pore-scale imaging and modelling. *Advances in Water Resource*s, 51, pp. 197–216. https://doi.org/10.1016/j.advwatres.2012.03.003

Brenne E.O., Dahl V.A., & Jørgensen P.S. (2021). A physical model for microstructural characterization and segmentation of 3D tomography data. *Materials Characterization*, 171, 110796. https://doi.org/10.1016/j.matchar.2020.110796

Carmignato S., Dewulf W. & Leach R. (2018). Industrial X-Ray Computed Tomography. *Springer International Publishing*, 369 p. https://doi.org/10.1007/978-3-319-59573-3

Chawshin K., Berg C.F., Varagnolo D., & Lopez O. (2021). Lithology classification of whole core CT scans using convolutional neural networks. *SN Applied Sciences*, 3(6), pp. 1–21. https://doi.org/10.1007/s42452-021-04656-8

Chawshin K., Berg C.F., Varagnolo D. & Lopez O. (2022). Automated porosity estimation using CT-scans of extracted core data. *Computational Geosciences*, 26(3), pp. 595–612. https://doi.org/10.1007/s10596-022-10143-9

dos Anjos C.E.M., Avila M.R.V, Vasconcelos A.G.P., Pereira Neta A.M., Medeiros L.C., Evsukoff A. G., Surmas R., & Landau L. (2021). Deep learning for lithological classification of carbonate rock micro-CT images. *Computational Geosciences*, 25(3), pp. 971–983. https://doi.org/10.1007/s10596-021-10033-6

Dullien F.A.L. (1992). Porous Media. Elsevier, 566 p. https://doi.org/10.1016/c2009-0-26184-8

Gärttner S., Alpak F. O., Meier A., Ray N. & Frank F. (2023). Estimating permeability of 3D micro-CT images by physics-informed CNNs based on DNS. *Computational Geosciences*, 27(2), pp. 245–262. https://doi.org/10.1007/s10596-022-10184-0

Gerke K.M., Korost D.V., Karsanina M.V., Korost S.R., Vasiliev R.V., Lavrukhin E.V. & Gafurova D.R. (2021). Modern approaches to pore space scale digital modeling of core structure and multiphase flow. *Georesursy = Georesources*, 23(2), pp. 197–213. https://doi.org/10.18599/GRS.2021.2.20

Gong H., Mu T., Li Q., Dai H., Li C., He Z., Wang W., Han F., Tuniyazi A., Li H., Lang X., Li Z., & Wang B. (2022). Swin-Transformer-Enabled YOLOv5 with Attention Mechanism for Small Object Detection on Satellite Images. *Remote Sensing*, 14(12), 2861. https://doi.org/10.3390/rs14122861

Graczyk K.M., Matyka M. (2020). Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. *Scientific Reports*, 10(1), 21488. https://doi.org/10.1038/s41598-020-78415-x

Iklassov Z., Medvedev D., Nazarov O. & Razzokov S. (2022). AI for Porosity and Permeability Prediction from Geologic Core X-Ray Micro-Tomography. *ArXiv preprint*, pp. 1–7. https://doi.org/10.48550/arXiv.2205.13189

Jackson S.J., Niu Y., Manoorkar S., Mostaghimi P., & Armstrong R.T. (2021). Deep learning of multi-resolution X-Ray micro-CT images for multi-scale modelling. *ArXiv preprint*, pp. 1–21. http://arxiv.org/abs/2111.01270

Jia Deng, Wei Dong, Socher R., Li-Jia Li, Kai Li, & Li Fei-Fei. (2009). ImageNet: A large-scale hierarchical image database. CVPR09, pp. 248–255. https://doi.org/10.1109/cvprw.2009.5206848

Kadyrov R., Glukhov M., Statsenko E. & Galliulin B. (2020). Enigma of ferruginous inclusions in Permian evaporites. *Arabian Journal of Geosciences*, 13(20), 1058. https://doi.org/10.1007/s12517-020-05995-3

Kadyrov R., Nurgaliev D., Saenger E.H., Balcewicz M., Minebaev R., Statsenko E., Glukhov M., Nizamova A., & Galiullin B. (2022). Digital rock physics: Defining the reservoir properties on drill cuttings. *Journal of Petroleum Science and Engineering*, 210, 110063. https://doi.org/10.1016/j.petrol.2021.110063

Kadyrov R., Starovoytov A. & Utemov E. (2020). Application of wavelet analysis of X-Ray computed tomography histogram for phase segmentation. *IOP Conference Series: Earth and Environmental Science*, 516(1), 012005. https://doi.org/10.1088/1755-1315/516/1/012005

Kadyrov R., Statsenko E. & Galiullin B. (2018). The porous space structure of domanik shales in the east of Russian plate. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM*, 18(1.4), pp. 907–914. https://doi.org/10.5593/sgem2018/1.4/S06.118

Kadyrov R., Statsenko E. & Nguyen T.H. (2024). Integrating μCT imaging of core plugs and transfer learning for automated reservoir rock characterization and tomofacies identification. *Marine and Petroleum Geology*, 168, 107014. https://doi.org/10.1016/j.marpetgeo.2024.107014

Kalule R., Abderrahmane H.A., Alameri W. & Sassi M. (2023). Stacked ensemble machine learning for porosity and absolute permeability prediction of carbonate rock plugs. *Scientific Reports*, 13(1), 9855. https://doi.org/10.1038/s41598-023-36096-2

Karimpouli S., Kadyrov R. (2022). Multistep Super Resolution Double-U-net (SRDUN) for enhancing the resolution of Berea sandstone images. *Journal of Petroleum Science and Engineering*, 216, 110833. https://doi.org/10.1016/j.petrol.2022.110833

Karimpouli, S., Kadyrov, R., Siegert, M., & Saenger, E. H. (2024). Applicability of 2D algorithms for 3D characterization in digital rocks physics: an example of a machine learning-based super resolution image generation. *Acta Geophysica*, 72(2), pp. 861–874. https://doi.org/10.1007/s11600-023-01149-7

Khan M.I., Khanal A. (2024). Machine Learning Assisted Prediction of Porosity and Related Properties Using Digital Rock Images. *ACS Omega*, 9(28), pp. 30205–30223. https://doi.org/10.1021/acsomega.3c10131

Kim S., Kim H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), pp. 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003

Latief, F. D. E., Fauzi, U., Irayani Z. & Dougherty G. (2017). The effect of X-ray micro computed tomography image resolution on flow properties of porous rocks. *Journal of Microscopy*, 266(1), pp. 69–88. https://doi.org/10.1111/jmi.12521

Lavrukhin E.V., Gerke K.M., Romanenko K.A., Abrosimov K.N., & Karsanina M.V. (2021). Assessing the fidelity of neural network-based segmentation of soil XCT images based on pore-scale modelling of saturated flow properties. *Soil and Tillage Research*, 209, 104942. https://doi.org/10.1016/j.still.2021.104942

Lavrukhin E.V., Murygin D.A., Toropov K.V., Khlyupin A.N. & Gerke K.M. (2025). Development of Synthetic Tomography Methods for Porous Media. *Mathematical Models and Computer Simulations*, 17(1), pp. 34–45. https://doi.org/10.1134/S2070048224700704

Liu H., Ma C. & Zhu C. (2022). X-ray Micro CT Based Characterization of Pore-Throat Network for Marine Carbonates from South China Sea. *Applied Sciences* (Switzerland), 12(5), 2611. https://doi.org/10.3390/app12052611

Liu Z., Lin Y., Cao Y., Hu H., Wei Y., Zhang Z., Lin S. & Guo B. (2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *Proceedings of the IEEE International Conference on Computer Vision*, pp. 9992–10002. https://doi.org/10.1109/ICCV48922.2021.00986

McPhee C., Reed J. & Zubizarreta I. (2015). Developments in Petroleum Science. Volume 64. Core Analysis: A BestPractice Guide (1st Edition). Elsevier Science Ltd, 852 p.

Mohammed A.S., Almutahhar M., Sattar K., Alhajeri A., Nazir A. & Ali U. (2023). Deep learning based porosity prediction for additively manufactured laser powder-bed fusion parts. *Journal of Materials Research and Technology*, 27, pp. 7330–7335. https://doi.org/10.1016/j.jmrt.2023.11.130

Naftaly M., Tikhomirov I., Hou P. & Markl D. (2020). Measuring Open Porosity of Porous Materials Using THz-TDS and an Index-Matching Medium. *Sensors*, 20(11), 3120. https://doi.org/10.3390/s20113120

Ortiz A.F., Herrera E.H. & Santos N. (2020). Porosity prediction from X-ray computed tomography logs (RHOB and PEF) using Artificial Neural Networks (ANN). *Boletín de Geología*, 42(3), pp. 141–149. https://doi.org/10.18273/revbol.v42n3-2020006

Paszke A., Gross S., Massa F., Lerer A., Bradbury J., Chanan G., Killeen T., Lin Z., Gimelshein N., Antiga L., Desmaison A., Köpf A., Yang E., DeVito Z., Raison M., Tejani A., Chilamkurthy S., Steiner B., Fang L., Bai J., Chintala S. (2019). PyTorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. Alché-Buc, E. Fox, & R. Garnett (Eds.). *Advances in Neural Information Processing Systems*, 32, pp. 1–12. https://doi.org/10.48550/arXiv.1912.01703

Qu H., Peng Y., Huang J., Pan Z. & Zhou F. (2023). Modelling of the impact of stress concentration on permeability in porous medium based on machine learning method. *Geoenergy Science and Engineering*, 224, 211655. https://doi.org/10.1016/j.geoen.2023.211655

Shah S.M., Gray F., Crawshaw J.P. & Boek E.S. (2016). Micro-computed tomography pore-scale study of flow in porous media: Effect of voxel resolution. *Advances in Water Resources*, 95, pp. 276–287. https://doi.org/10.1016/j.advwatres.2015.07.012

Shaik A.R., Al-Ratrout A.A., AlSumaiti A.M. & Jilani A.K. (2019). Rock Classification Based on Micro-CT Images using Machine Learning Techniques. D021S060R003. https://doi.org/10.2118/197651-MS

Shan L., Liu C., Liu Y., Kong W. & Hei X. (2022). Rock CT Image Super-Resolution Using Residual Dual-Channel Attention Generative Adversarial Network. *Energies*, 15(14), 5115. https://doi.org/10.3390/en15145115

Song R., Zheng L., Wang Y. & Liu J. (2020). Effects of Pore Structure on Sandstone Mechanical Properties Based on Micro-CT Reconstruction Model. *Advances in Civil Engineering*, (1), pp. 1–21. https://doi.org/10.1155/2020/9085045

Tembely M., AlSumaiti A.M. & Alameri W. (2020). A deep learning perspective on predicting permeability in porous media from network modeling to direct simulation. *Computational Geosciences*, 24(4), pp. 1541–1556. https://doi.org/10.1007/s10596-020-09963-4

Tiab D., Donaldson E.C. (2016). Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties: Fourth Edition (Fourth Ed.). Elsevier, 894 p. https://doi.org/10.1016/C2014-0-03707-0

Van Geet M., Swennen R. & Wevers M. (2000). Quantitative analysis of reservoir rocks by microfocus X-ray computerised tomography. *Sedimentary Geology*, 132(1–2), pp. 25–36. https://doi.org/10.1016/S0037-0738(99)00127-X

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

Wang D., Yang R., Zhang Z., Liu H., Tan J., Li S., Yang X., Wang X., Tang K., Qiao Y. & Su P. (2023). P-Swin: Parallel Swin transformer multi-scale semantic segmentation network for land cover classification. *Computers and Geosciences*, 175, 105340. https://doi.org/10.1016/j.cageo.2023.105340

Wang Y.Da, Armstrong R.T. & Mostaghimi P. (2019). Enhancing Resolution of Digital Rock Images with Super Resolution Convolutional Neural Networks. *Journal of Petroleum Science and Engineering*, 182, 106261. https://doi.org/10.1016/j.petrol.2019.106261

Wu H.Y., Kuo C.L., Lin C.Y. & Chu W.T. (2025). Deep learning in the advanced core sample porosity determination with XCT image. *Earth Science Informatics*, 18(1), 17. https://doi.org/10.1007/s12145-024-01602-1

Xu X., Feng Z., Cao C., Li M., Wu J., Wu Z., Shang Y. & Ye S. (2021). An improved swin transformer-based model for remote sensing object detection and instance segmentation. *Remote Sensing*, 13(23), 4779. https://doi.org/10.3390/rs13234779

Yang S. (2017). Fundamentals of Petrophysics. Springer Berlin Heidelberg, 509 p. https://doi.org/10.1007/978-3-662-55029-8

Zhang H., Yu H., Yuan X., Xu H., Micheal M., Zhang J., Shu H., Wang G. & Wu H. (2022). Permeability prediction of low-resolution porous media images using autoencoder-based convolutional neural network. *Journal of Petroleum Science and Engineering*, 208, 109589. https://doi.org/10.1016/j.petrol.2021.109589

Zhang L., Jing W., Yang Y., Yang H., Guo Y., Sun H., Zhao J. & Yao J. (2019). The Investigation of Permeability Calculation Using Digital Core Simulation Technology. *Energies*, 12(17), 3273. https://doi.org/10.3390/en12173273

Zhang P., Lee Y.Il, & Zhang J. (2019). A review of high-resolution X-ray computed tomography applied to petroleum geology and a case study. *Micron*, 124, 102702. https://doi.org/10.1016/j.micron.2019.102702

Zheng D., Zhong H., Camps-Valls G., Cao Z., Ma X., Mills B., Hu X., Hou M., & Ma C. (2024). Explainable deep learning for automatic rock classification. *Computers & Geosciences*, 184, 105511. https://doi.org/10.1016/j.cageo.2023.105511

Zollanvari A. (2023). Machine learning with Python: Theory and implementation. Machine Learning with Python: Theory and Implementation. Springer International Publishing, 452 p. https://doi.org/10.1007/978-3-031-33342-2

About the Authors

Rail I. Kadyrov – Cand. Sci. (Geology and Mineralogy), Senior Researcher, Institute of Geology and Petroleum Technologies, Kazan Federal University
4/5 Kremlevskaya st., Kazan, 420008, Russian Federation
e-mail: rail7777@gmail.com

Evgeny O. Statsenko – Junior Research Fellow, Institute of Geology and Petroleum Technologies, Kazan Federal University
4/5 Kremlevskaya st., Kazan, 420008, Russian Federation

Thanh Hung Nguyen – Junior Research Fellow, Institute of Geology and Petroleum Technologies, Kazan Federal University
4/5 Kremlevskaya st., Kazan, 420008, Russian Federation

Marta A. Skorobogatova – Engineer, Institute of Geology and Petroleum Technologies, Kazan Federal University
4/5 Kremlevskaya st., Kazan, 420008, Russian Federation

Manuscript received 11 March 2025;
Accepted 30 June 2025; Published 20 December 2025

IN RUSSIAN

ОРИГИНАЛЬНАЯ СТАТЬЯ

Прогнозирование пористости стандартных образцов керна на основе микротомографии, разметки с учителем и трансформера со сдвигаемыми окнами

Р.И. Кадыров, Е.О. Стаценко, Т.Х. Нгуен, М.А. Скоробогатова*
Институт геологии и нефтегазовых технологий, Казанский федеральный университет, Казань, Россия
**Ответственный автор: Раиль Илгизарович Кадыров, e-mail: rail7777@gmail.com*

Недавние достижения в области машинного обучения позволили автоматически анализировать изображения микротомографии (микро-КТ), способствуя более эффективной идентификации свойств горных пород. Целью данного исследования является прогнозирование экспериментально измеренной открытой пористости пород-коллекторов с использованием изображений микро-КТ стандартных образцов керна. Был собран набор данных из 136 образцов керна, включая 49 образцов песчаника и 87 образцов карбоната. Открытая пористость была экспериментально определена с использованием газового волюметра. Образцы керна (30 ± 1 мм в высоту и диаметр) были отсканированы с помощью микро-КТ с разрешением 34,6–38,0 мкм, что дало 16-битные стеки изображений. Набор данных состоял из 100 232 изображений (64 119 карбоната и 36 113 песчаника). Для маркировки изображений мы ввели контролируемый метод под названием «Сегментация неразрешенных пор с помощью экспериментального эталона» (SUPER), который сегментирует темные воксели для соответствия экспериментально измеренной открытой пористости, адаптируясь к характеристикам каждого образца. Были обучены три модели трансформера со сдвигаемыми окнами (Swin): универсальная модель и специализированные модели для песчаника и карбоната. Модели использовали трансферное обучение с весами ImageNet, за которым последовала тонкая настройка. Тестирование подтвердило, что специализированные модели превзошли универсальную модель. Это подчеркивает, что обучение ансамбля моделей, адаптированных к определенным типам пород, приводит к лучшей производительности, чем одна общая модель для прогнозирования пористости. Основная проблема возникла с песчаниками, особенно мелкозернистыми типами, где мелкие поры сливались из-за ограничений разрешения. Последующая работа должна быть направлена на улучшение разрешения изображений и непосредственное введение детализированных изображений в модель. Метод имеет потенциал применения для полноразмерного керна и ранней оценки пористости в неэкстрагированных стандартных образцах, включая хрупкие коллекторы с нефтью или битумами.

Ключевые слова: микро-КТ (микротомография), открытая пористость, коллектор, стандартный образец керна, машинное обучение, трансферное обучение, Swin (трансформер со сдвигаемыми окнами), прогнозирование пористости

Для цитирования: Kadyrov R.I., Statsenko E.O., Nguyen T.H., Skorobogatova M.A. (2025). Core Plug Porosity Prediction Using Microtomography, Supervised Labeling, and a Shifted Window Transformer. *Georesursy = Georesources,* 27(4), pp. 67–82. https://doi.org/10.18599/grs.2025.4.16

Благодарности

Работа выполнена за счет гранта Академии наук Республики Татарстан, предоставленного молодым кандидатам наук (постдокторантам) с целью защиты докторской диссертации, выполнения научно-исследовательских работ, а также выполнения трудовых функций в научных и образовательных организациях Республики Татарстан в рамках Государственной программы Республики Татарстан «Научно-технологическое развитие Республики Татарстан» (соглашение № 20/2024-ПД).

Сведения об авторах

Раиль Илгизарович Кадыров – кандидат геол.-минерал. наук, старший научный сотрудник, Институт геологии и нефтегазовых технологий, Казанский федеральный университет
Россия, 420008, Казань, ул. Кремлевская, д. 4/5
e-mail: rail7777@gmail.com

Евгений Олегович Стаценко – младший научный сотрудник, Институт геологии и нефтегазовых технологий, Казанский федеральный университет
Россия, 420008, Казань, ул. Кремлевская, д. 4/5

Тхань Хынг Нгуен – младший научный сотрудник, Институт геологии и нефтегазовых технологий, Казанский федеральный университет
Россия, 420008, Казань, ул. Кремлевская, д. 4/5

Марта Андреевна Скоробогатова – инженер, Институт геологии и нефтегазовых технологий, Казанский федеральный университет
Россия, 420008, Казань, ул. Кремлевская, д. 4/5

Статья поступила в редакцию 11.03.2025;
Принята к публикации 30.06.2025; Опубликована 20.12.2025