PoreSkel: Skeletonization of grayscale micro-CT images of porous media using deep learning techniques

Mehdi Mahdaviara [a], Mohammad Sharifi [a,*], Amir Raoof [b,*]

[a] *Department of Petroleum Engineering, Amirkabir University of Technology, Tehran Polytechnic, Tehran 15875-4413, Iran*
[b] *Department of Earth Sciences, Utrecht University, P.O. Box 80.021, TA, Utrecht 3508, the Netherlands*

ARTICLE INFO

ABSTRACT

Skeletonization, a crucial step in pore network modeling, traditionally involves the extraction of skeleton pixels from binarized, segmented X-ray images of porous materials. However, this conventional approach often suffers from user bias during segmentation, potentially leading to the loss of essential image details. This study addresses this limitation by developing deep learning model, called PoreSkel, designed to directly perform skeletonization and distance map extraction from unprocessed grayscale images, thus eliminating the need for additional image processing steps. The model was trained, validated, and tested using an expansive databank of micro-CT images from 20 distinct sandstones, carbonates, and sand pack samples, a total of 10,240 images, with each sample represented by a cube of size 512^3. A fifth of these images, specifically 15.6 % from sixteen sandstone and sand pack samples, were used for training, while the remainder served for model validation (4.4 %) and extensive testing (80 %). PoreSkel showed an excellent performance, achieving a mean f1-score of 0.964 for skeletonization and an RMSE of 0.057 for distance map extraction during the testing phase. Our assessments revealed that the model is robust to bias toward the majority class, namely the background pixels. Furthermore, the model showed high generality, maintaining its performance when tested using unseen images from three carbonates and an additional sandstone. Notably, PoreSkel effectively handles disruptions caused often by the presence of minerals in pore spaces and perturbations on pore boundaries - a common challenge for the medial axis technique - resulting in fewer nodes (i.e., pore junctions) and pore coordination numbers, but a higher number of connected skeletons. Therefore, PoreSkel provided a more precise and representative pore structures of porous material that is needed for accurate pore network generation and modeling.

1. Introduction

Characterizing porous materials, such as sedimentary deposits, has been subject of many studies (Blunt, 2017; Mahdaviara et al., 2022; Raoof and Hassanizadeh, 2010). Advances in high-resolution X-ray tomography have enabled the investigation of internal structures within porous materials (Blunt et al., 2013; Bultreys et al., 2016). The acquired images allow us to accurately identify textures and even constituent minerals in rocks (Xiong et al., 2016), determine static parameters like porosity and tortuosity (Alqahtani et al., 2020; Graczyk and Matyka, 2020), and simulate flow of single and multiple phases (Najafi et al., 2023; Raoof et al., 2013). Representation and characterization of dynamic processes can be achieved using direct numerical simulation (DNS, e.g., by Armstrong et al. 2016, Wang et al., 2020, Fathi et al. 2017a) and pore network modeling approaches (PNM, e.g., by Raoof

et al. 2013). In DNS, the governing equations of flow are numerically solved without simplifying the pore morphology, while in PNM, calculations are performed by mapping the complex pore space structure into a reduced representation often consisting of regular geometrical shapes. Although DNS yields more realistic results, its notably high computational costs limit its applicability to small portions of porous media (Blunt, 2001; Blunt et al., 2013; Fathi et al., 2017b)). Consequently, the PNM method is indispensable for large samples sizes.

PNM reconstruction begins with the topological and geometrical characterization of porous media (Mahmoodlu et al., 2020; Vogel and Roth, 2001). Topological characterization involves identifying the locations of pores and throats, as well as quantifying their interrelationships through connectivity and mean coordination number (Lindquist et al., 1996; Raoof and Hassanizadeh, 2010). Connectivity is often determined by the Euler number (or Euler–Poincaré characteristic),
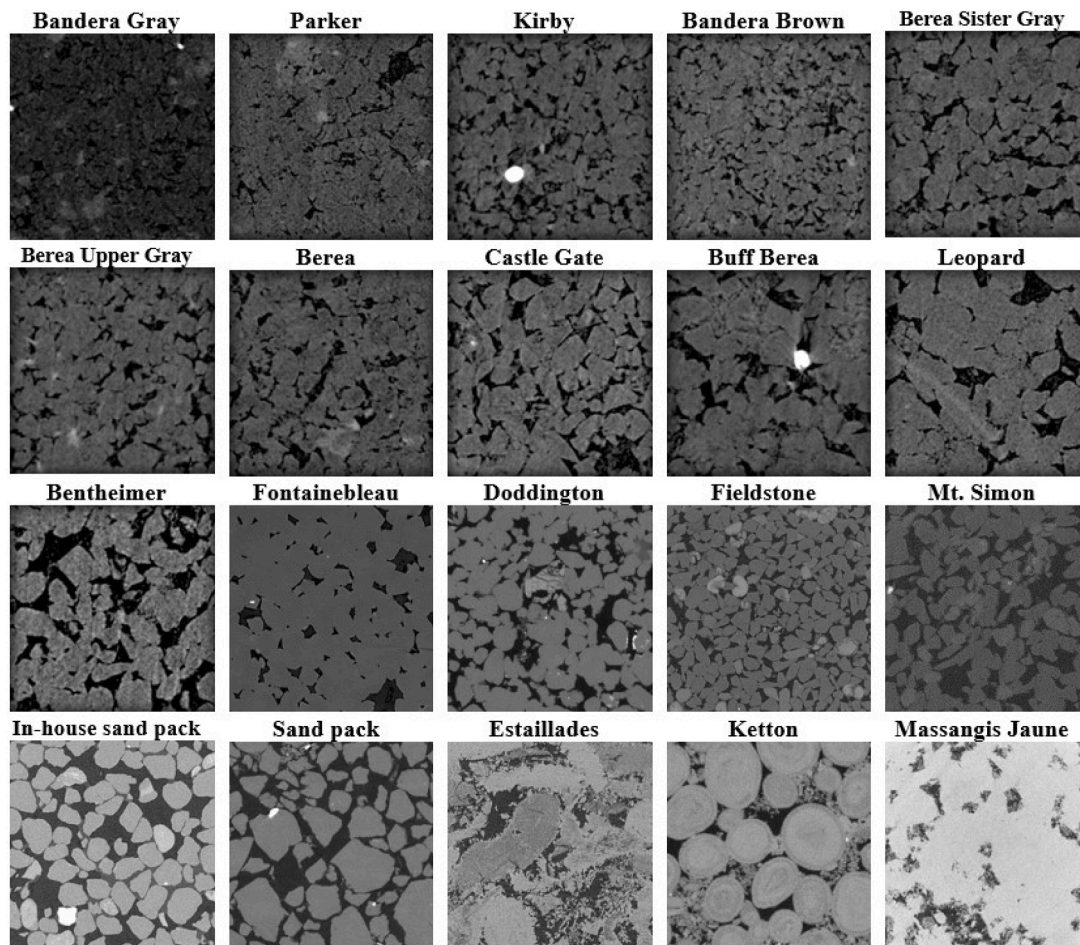
Fig. 1. The x-ray images of diverse samples used for both training and validation of the constructed deep learning models. Within the provided samples, there are 15 sandstones (located in rows 1 through 3 from the top), along with two sand packs and three carbonates (the last row). Samples show a wide range of petrophysical properties, pore structures, and imaging resolutions.

Table 1

Detailed description of the dataset utilized in this study. The samples have various porosity, permeability, and resolutions, changing in the wide ranges of 11.6–38 %, 9–386 mD, and 2.25–8.30 μm, respectively.

	Sample	Porosity (%)	Permeability (mD)	Resolution (μm)	Dimension	Refs.
Sandstone	**Bandera Gray**	18.10	9	2.25	512^3	Neumann et al. (2020)
	Parker	14.77	10	2.25	512^3	Neumann et al. (2020)
	Kirby	19.95	62	2.25	512^3	Neumann et al. (2020)
	Bandera Brown	24.11	63	2.25	512^3	Neumann et al. (2020)
	Berea Sister Gray	19.07	80	2.25	512^3	Neumann et al. (2020)
	Berea Upper Gray	18.56	86	2.25	512^3	Neumann et al. (2020)
	Berea	18.96	121	2.25	512^3	Neumann et al. (2020)
	Castle Gate	26.54	269	2.25	512^3	Neumann et al. (2020)
	Buff Berea	24.02	275	2.25	512^3	Neumann et al. (2020)
	Leopard	20.22	327	2.25	512^3	Neumann et al. (2020)
	Bentheimer	22.64	386	2.25	512^3	Neumann et al. (2020)
	Mt. Simon	27.10	NA	2.80	512^3	Kohanpur et al. (2019)
	Doddington	NA	NA	5.38	512^3	Andrew et al. (2020)
	Fieldstone	18.16	NA	5.20	512^3	Bultreys and De Boever (2020)
	Fontainebleau	NA	NA	3.28	512^3	Garing et al. (2017)
Sand pack	**Sand pack**	35.40	48.72	4.98	512^3	Singh et al. (2020)
	In-house sand pack	38.0	NA	8.30	512^3	Mahmoodlu et al. (2020)
Carbonate	**Estaillades**	NA	NA	3.10	512^3	Bultreys (2016a)
	Ketton	NA	NA	5.00	512^3	Scanziani et al. (2018)
	Massangis Jaune	11.60	0.03–70	4.54	512^3	Bultreys (2016b)

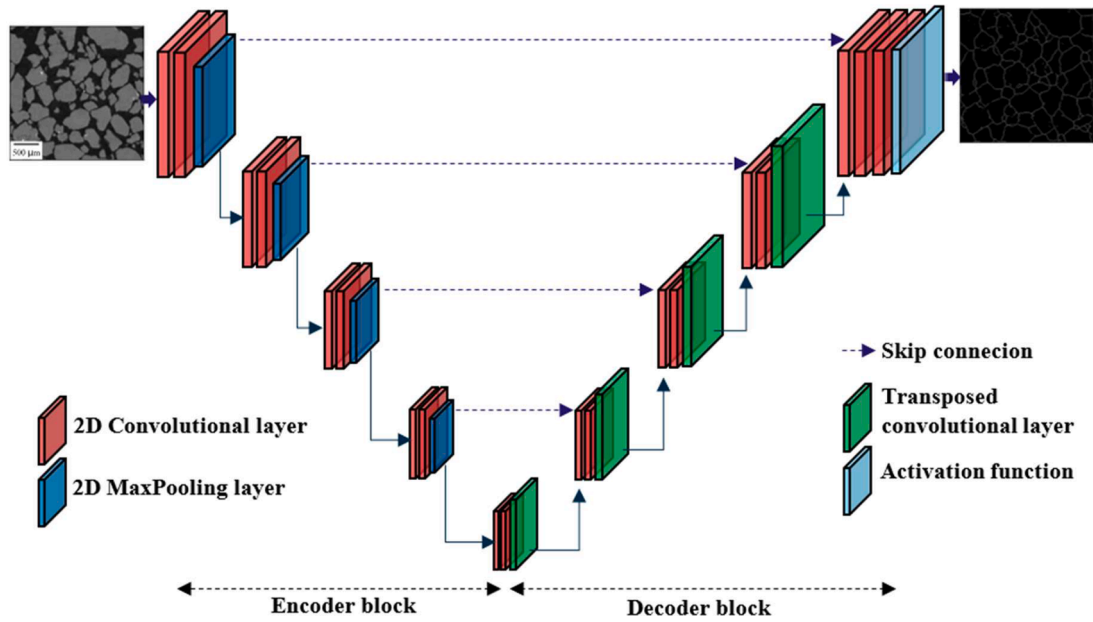NA stands for "Not Available".

Fig. 2. Schematic representation of the deep learning structure utilized in this study. The process begins with unprocessed grayscale images being used as input for skeleton extraction, which is done directly from the grayscale images. The model utilizes several layers such as convolutional, MaxPooling, and Transposed convolutional layers, along with skip connections and an activation function. These layers are trained through an encoding and decoding process. A similar procedure is also used to extract the distance map.
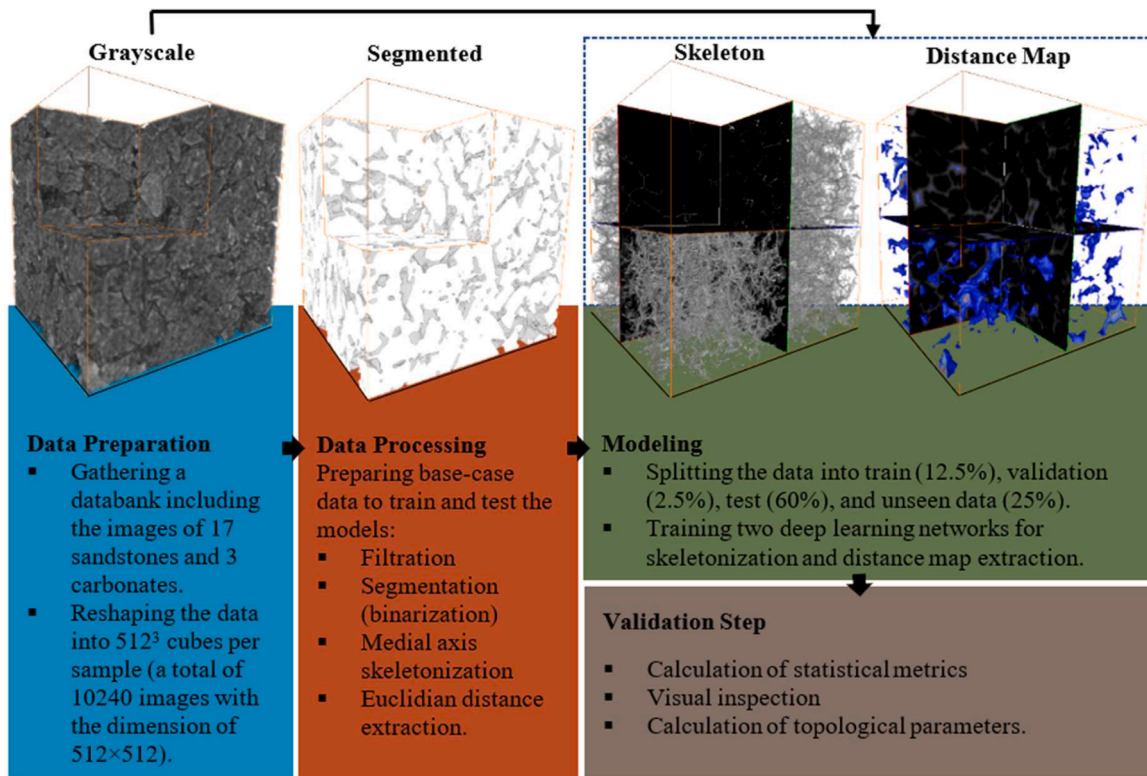

Fig. 3. Workflow describing the development of PoreSkel for direct skeletonization and distance map extraction. Preskel is trained to deduce the skeleton and distance map of porous media from grayscale images, eliminating the need for image processing, such as segmentation.

which quantifies the structure of a topological space, independent of any deformation or distortion it may undergo (Vogel and Roth, 2001). The mean coordination number refers to the average number of throats connected to each pore in porous media (Bakke and Øren, 1997; Thovert et al., 1993). Geometrical representation describes the shape and size of pores and throats using shape factors and aspect ratios. The shape factor is the dimensionless ratio of a pore space perimeter to its area, while the aspect ratio pertains to the diameter of a pore body relative to a pore throat (Joekar Niasar et al., 2009). Pore bodies are typically considered spherical or cubic (Joekar-Niasar et al., 2010; Koplik and Lasseter, 1985), while throat cross-sections can assume a wide variety of shapes to effectively represent the two-phase flow of non-wetting and
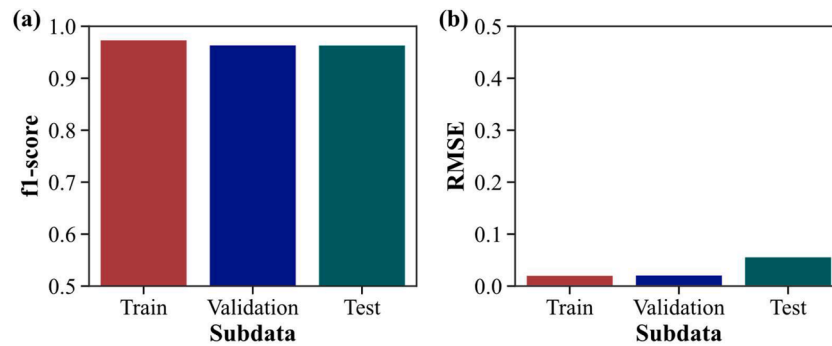
Fig. 4. The average metrics of (a) f1-score and (b) root mean squared error (RMSE) to assess the performance of PoreSkel in extracting skeletons and distance maps, respectively. The number of images (with the dimensions of 512×512) used for training, validation, and testing are 1200, 300, and 8240, respectively.
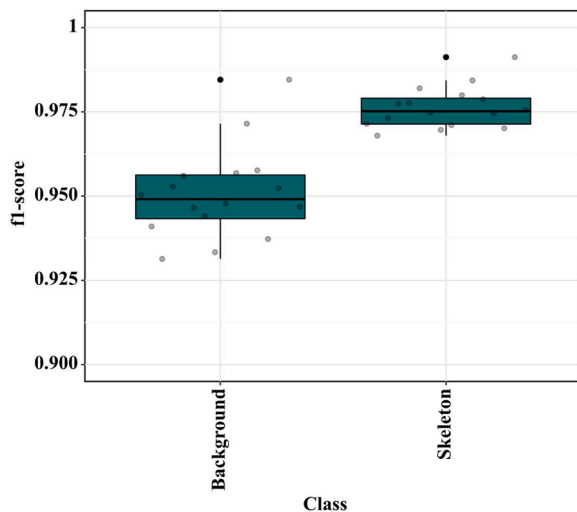

Fig. 5. The f1-score of the individual classes of skeleton and background. Each point corresponds to the f1-score of a specific rock sample. The median f1-score, as depicted by the line within the interquartile range (IQR), is near to 1 for both the skeleton and background classes. This suggests that the model has successfully addressed the issue of class imbalance. Moreover, the size of the boxes is small, indicating the consistency of results for all samples.

wetting phases in the pore space (Joekar-Niasar and Hassanizadeh, 2012; Mason and Morrow, 1991; Pereira et al., 1996; van Dijke and Sorbie, 2006). The aforementioned parameters can be measured using statistical, geological (grain-based), and direct mapping techniques (Bryant and Blunt, 1992; Ioannidis and Chatzis, 2000). Statistical and geometrical methods employ information obtained from 2D thin sections of rocks to reconstruct the pore structure. In contrast, the direct mapping approach extracts an irregular pore network directly from 3D computed tomography (CT) images, providing more reliable results.

The extraction of pore body and throat coordinates from CT images is a vital aspect of pore network extraction (Jiang et al., 2007; Raoof and Hassanizadeh, 2010). This process is typically executed using a morphological technique known as skeletonization. This method captures the essential depiction of porous media by extracting single-pixel or voxel skeleton lines that align with the medial axis of the pore space. The skeleton's junction and non-junction points are viewed as pore and throat bodies, respectively. Moreover, it provides valuable data concerning throat length, redundant connections, connection coefficients, isolated pores, enclosed cavities, and coordination numbers. The precision of this data is heavily dependent on the quality of the extracted skeleton. An optimal skeleton is centrally positioned within the pore space (medial-ness) and maintains the topological and geometrical characteristics of the original image (Jiang et al., 2007).

Traditional methods for producing image skeletons encompass medial axis transformations (MAT) (Lee, 1982), thinning approaches (Pudney, 1998), and a blend of both to satisfy the above criteria. The MAT procedure involves obtaining a distance transform, or distance map, which represents the distance of each pixel or voxel of pore space from the nearest boundary. Local maxima, located in the center of the pore space, are extracted as skeleton pixels or voxels. Thinning methods utilize morphological operators for skeletonization, such as eroding pore space pixels or voxels from the boundaries repeatedly. Nevertheless, MAT and thinning techniques may struggle to preserve topology and medial-ness, respectively (Jiang et al., 2007). Several modifications and hybrid methods have been suggested to mitigate these limitations (Delerue et al., 1999; Lohou and Bertrand, 2005; Morgenthaler, 1981; Saito, 1995). Despite these improvements, the impact of image processing quality on skeletonization outcomes remains under-explored. To our knowledge, most, if not all, algorithms necessitate segmented image representations (binarized into voids and solids) for skeleton extraction. Image segmentation is considerably influenced by variables such as image resolution, overlapping noise/artifacts, segmentation methods, and user bias (Mahdaviara et al., 2023). Minor segmentation errors can cause substantial deviations in the extracted skeleton and the resultant topological and geometrical information. As a result, there exists disparity between PNM calculations and actual measurements.

In this study, we aim to close this gap by circumventing segmentation and directly extracting the skeleton and distance map from the original (grayscale) images. We trained robust deep-learning algorithms on a substantial dataset comprising micro-CT images of 20 sandstones, carbonates, and sand packs for this purpose. These trained models, called PoreSkel, were evaluated using statistical metrics and visual inspections. Additionally, their universality was assessed by testing them on unseen sandstone and carbonate images. This study is structured as follows: Section 2 provides the dataset used for training and evaluation, along with the modeling procedure. Section 3 discusses the results, and the final section presents the conclusions.

2. Materials and methods

2.1. Datasets

To enhance the generality of PoreSkel, we assembled a diverse collection of X-ray CT data from 20 rock samples, each exhibiting distinct textures and petrophysical properties (Andrew et al., 2020; Bultreys, 2016a, 2016b, 2016c; Garing et al., 2017; Kohanpur et al., 2019; Neumann et al., 2020; Scanziani et al., 2018; Singh et al., 2020). The databank encompasses Bandera Gray, Parker, Kirby, Bandera Brown, Berea, Berea Sister Gray, Berea Upper Gray, Castle Gate, Buff Berea, Leopard, Bentheimer, Doddington, Fieldstone, and Fontainebleau sandstones, two Sand packs, along with Estaillades, Ketton, and Massangis Jaune carbonates (refer to Fig. 1). The description of these samples can be found in Table 1.
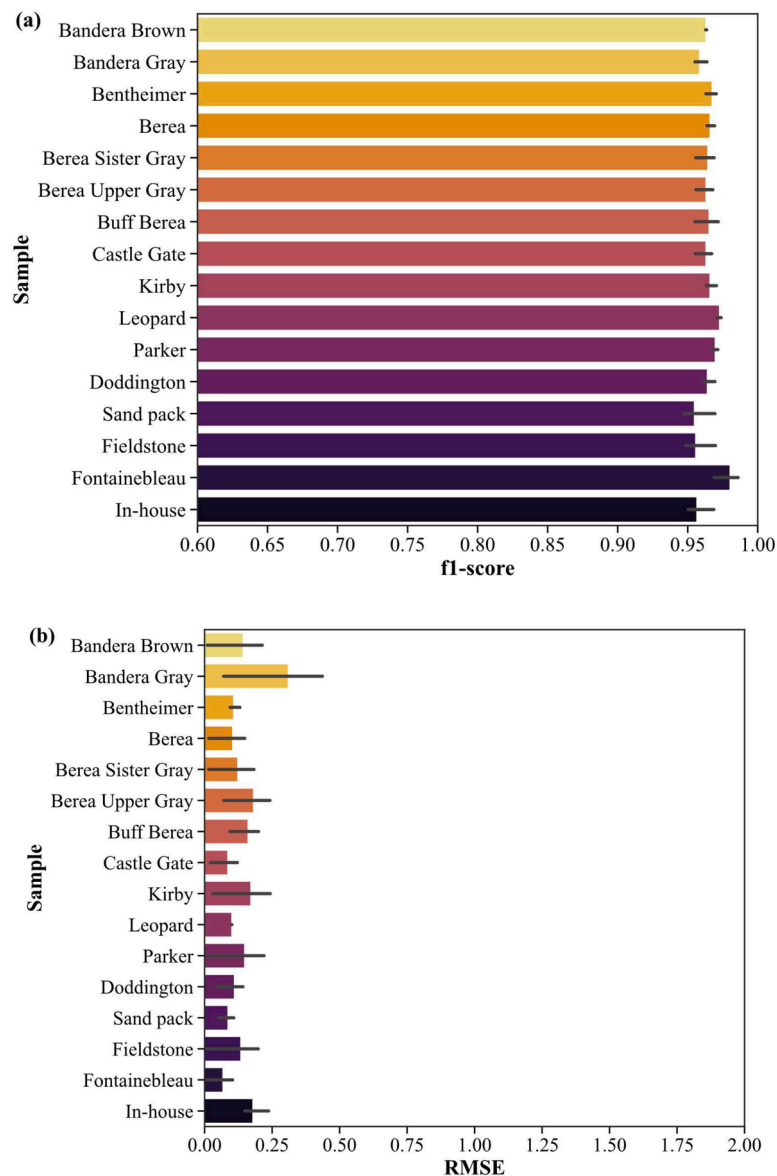
Fig. 6. Performance evaluation metrics for PoreSkel in extracting skeletons and distance maps across varied rock samples: (a) F1-score and (b) Root Mean Squared Error (RMSE). Horizontal bars denote the average scores for training, validation, and testing phases. The images of 16 sandstones and sand packes were used in thid step, while the images of carbonates and the Mt. Simon sandstone were excluded for further evaluation.

2.2. Image processing

We prepared a total of 10,240 2D grayscale images by picking 512^3 voxel volumes from each rock sample. The image histogram is illustrated in Fig. A.1, which reveals significant intensity differences among the various rocks, thus highlighting the challenge of training single models for skeletonization and distance map extraction across all samples. To establish the base-case data, the following procedure was performed; (i) a non-local mean filter was applied to the images to eliminate possible noise and artifacts, (ii) images were binarized (segmenting them into voids and solids) using a thresholding approach, and (iii) skeletons and distance maps were extracted from the binarized images by employing the medial axis transform (MAT) and Euclidean distance methods, respectively. Consequently, we compiled a databank of these grayscale images (each measuring 512 × 512) and their corresponding base-case data, which served as the foundation for training and evaluating the model in a supervised manner.

2.3. Deep autoencoder modeling

In this study, we utilized U-Net method for skeletonization and distance map extraction from grayscale images. Autoencoders, which are feed-forward neural networks, learn a compressed representation of the input data, commonly referred to as a latent representation or encoding (Géron, 2022). The encoded data possess a lower dimensionality, which makes autoencoders excellent candidates for feature extraction and dimensionality reduction (Goodfellow et al., 2016). Contemporary autoencoders incorporate a decoder block that decodes the representations back into their original dimensions.

An autoencoder typically consists of two components: an encoder and a decoder. The encoder, a network that maps the data into a reduced dimensionality space, processes the input data. This dense representation is then passed through the decoder, another neural network that translates the codes back to their primary dimensions. The output from the final layer is a reconstructed version of the input. The autoencoder's objective is to minimize this reconstruction error, achieved by using an optimization algorithm to adjust the network weights.
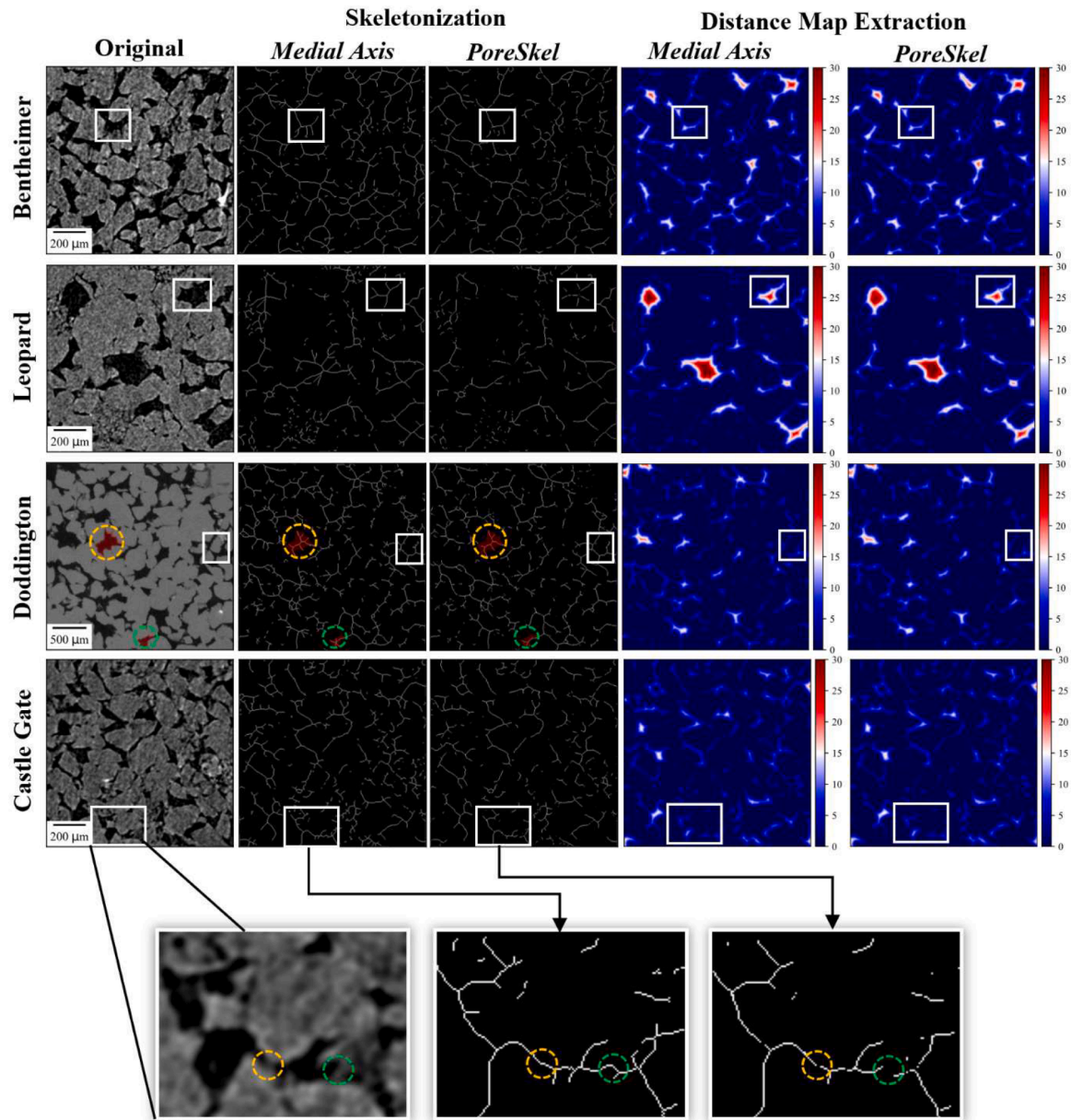
Fig. 7. A comparative visual depiction of skeletons and distance maps generated by the Medial Axis and PoreSkel methods for a range of samples. Notable areas where PoreSkel demonstrates superior performance to the base case technique are highlighted within squares. PoreSkel effectively accounts for discontinuities induced by minerals within the pore space, as illustrated by dashed circles in the Castle Gate sample (bottom plots). It also shows a distinct ability to maintain topological consistency despite boundary perturbations of pores, a feature that the base case method struggles to handle, as demonstrated in the circled regions of the Doddington sample.
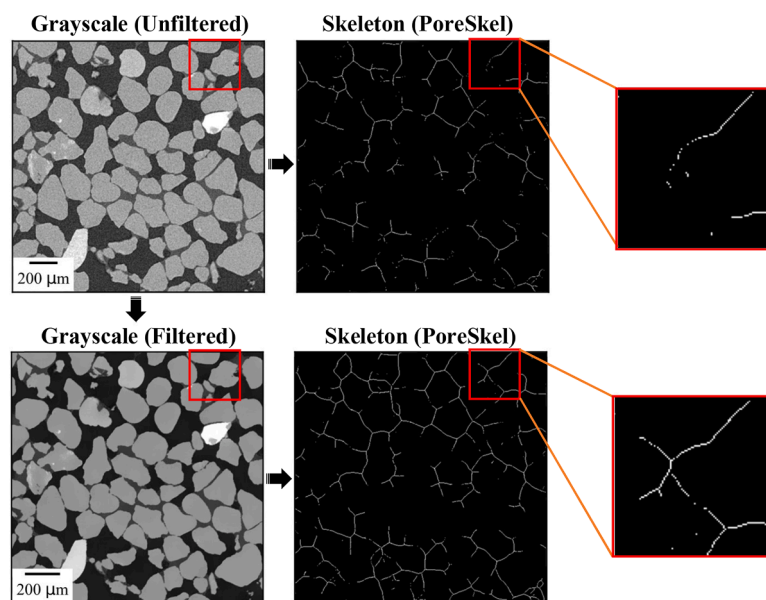
Fig. 8. Evaluation of PoreSkel's performance with both filtered and unfiltered input data, derived from in-house sand pack sample. The figure clearly illustrates the enhancement of PoreSkel's performance when image pre-filtering is applied prior to the skeletonization process.

Various autoencoder extensions exist to address specific challenges, including the denoising autoencoder (DAE) (Lu et al., 2013), variational autoencoder (VAE) (Kingma and Welling, 2013), recurrent autoencoder (RAE) (Gonzalez and Balajewicz, 2018), sparse autoencoder (Hinton and Salakhutdinov, 2006), and U-Net (Ronneberger et al., 2015). U-Net is a deep autoencoder specifically designed for image segmentation tasks, wherein image objects are identified and outlined. The encoder and decoder blocks of U-Net traditionally employ a multi-neuron convolutional neural network (CNN), a deep-learning architecture widely used for image recognition, utilizing convolutional layers for feature extraction and maxpooling layers for reducing the output's spatial dimensions. For a comprehensive overview of CNNs, readers are referred to LeCun et al. (2015), Goodfellow et al. (2016).

U-Net incorporates CNNs in a symmetric U-shape architecture, gradually decreasing and increasing the feature map scales in the encoder and decoder, respectively. This strategy reduces the network's parameter count and helps prevent overfitting. The unique innovation of U-Net is the use of skip connections from the encoder to the decoder, facilitating the transfer of more information and leading to higher segmentation accuracy. Owing to its effectiveness in various medical image segmentation tasks, several variations and extensions of the original U-Net architecture, such as Attention U-Net (Oktay et al., 2018) and UNet++ (Zhou et al., 2018), have been developed to further enhance the accuracy and efficiency of image segmentation.

The configuration employed in this study is depicted in Fig. 2. We divided the data bank into four subsets: training set (12.5 %), validation set (2.5 %), test set (60 %), and unseen data (25 %). The first three sets consist of images of sixteen sandstones and sand packs. The carbonates and the Mt. Simon sandstone were exclusively assigned to the unseen data, enabling a more comprehensive evaluation of the models on data excluded from the training and testing stages. The training sets were used for feature extraction in the encoder block. This block consists of five convolutional layers with 3×3 kernels, ranging from 64 to 1024. To ensure that the kernel convolves over the image one step at a time, the stride length was set to 1. The outputs of each convolutional layer were passed through a maxpooling layer to reduce the size of extracted features while preserving the most prominent features.

The decoder block mirrors the architecture of the encoder but uses transposed convolutional layers for upsampling. The network's output layer was tailored for classification and regression tasks of skeletonization and distance map extraction, respectively. We employed the Softmax (normalized exponential function) and Softplus activation functions for skeletonization and distance map extraction, respectively (Bishop and Nasrabadi, 2006; Goodfellow et al., 2016).

The skeletonization model was trained based on the categorical cross-entropy loss function, while the mean square error (MSE) was used for distance map extraction. In both models, the weights were tuned by the Adam optimizer (Kingma and Ba, 2014) with a learning rate of 1×10^{-4}. After conducting multiple trials, we determined 10 epochs to be the optimal number to mitigate potential issues of overfitting and bias. For the sake of clarity, we named the trained models for skeleton and distance map extraction as PoreSkel.

3. Results and discussions

As previously highlighted, deep learning has been utilized to deduce the skeleton and distance map of porous media from grayscale images, eliminating the need for image processing, such as filtering and segmentation. The dataset employed spans micro-CT scans from 20 distinct sandstone and carbonate samples (a 512^3 cube per sample, totaling 10,240 images), all exhibiting various topological and geological attributes (refer to Table 1 and Fig. 1). A fifth (20 %) of sixteen sandstone images were earmarked for training (15.6 %) and evaluation (4.4 %), while the remaining bulk (80 %) was set aside for testing the model. The trained model, PoreSkel, was tested using statistical metrics, visual assessment, and the computation of topological and geometrical properties. The steps undertaken during the development of PoreSkel are shown in Fig. 3.

3.1. Evaluation by calculation of statistical metrics

The overall efficiency of PoreSkel for skeletonization was assessed through accuracy, intersection over union (IoU) (Jaccard, 1912), and f1-score (Dice, 1945), also known as the dice similarity coefficient. Conversely, the deduced distance maps were evaluated via mean squared error (MSE) and root mean squared error (RMSE) metrics. Accuracy is defined as the ratio of the number of predictions congruent with the base-case data (medial axis) to the total number of pixels:

$$\text{Accuracy} = \frac{\text{The number of accurate predictions}}{\text{The total number of pixels}} \tag{1}$$
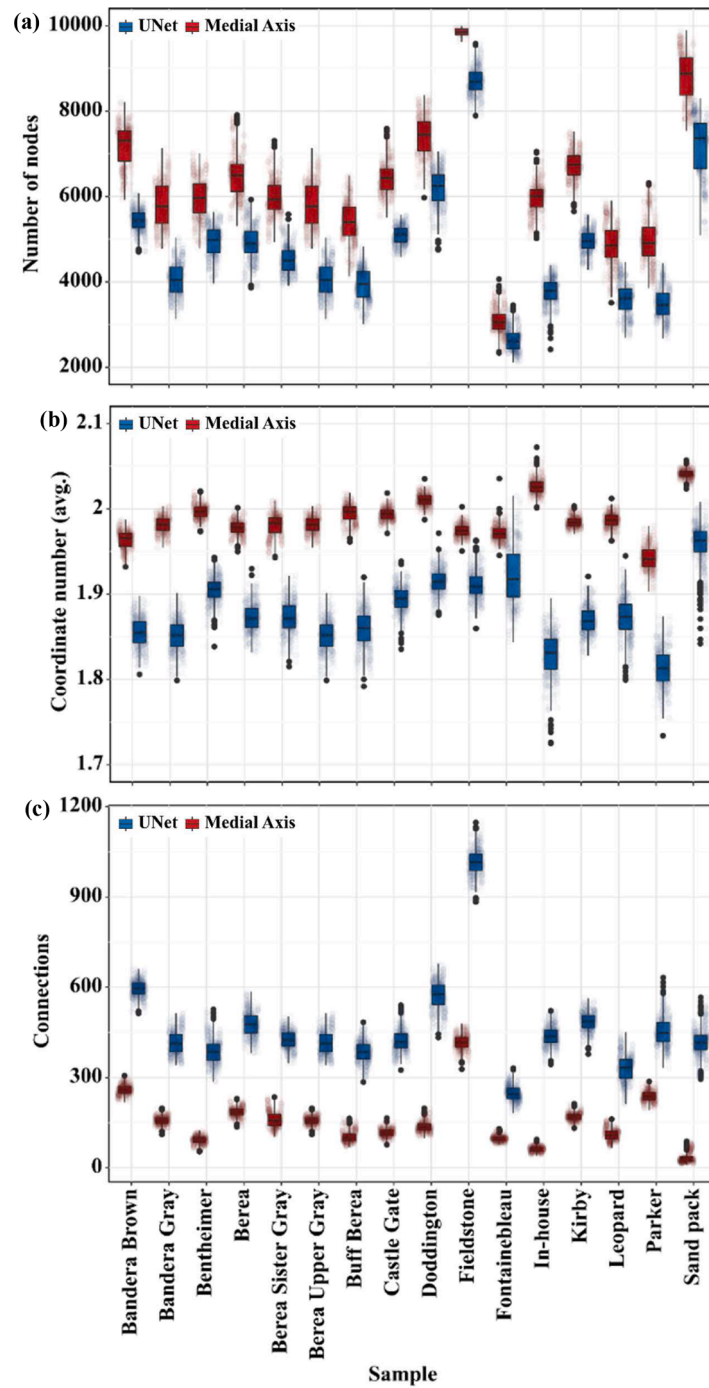
Fig. 9. Analysis of the extracted skeletons using (a) the number of nodes, (b) average coordinate number, and (c) connected skeletons. In compared to media axis method, PoreSkel provides a lower number of nodes and coordination numbers, while showing a higher number of connected skeletons.
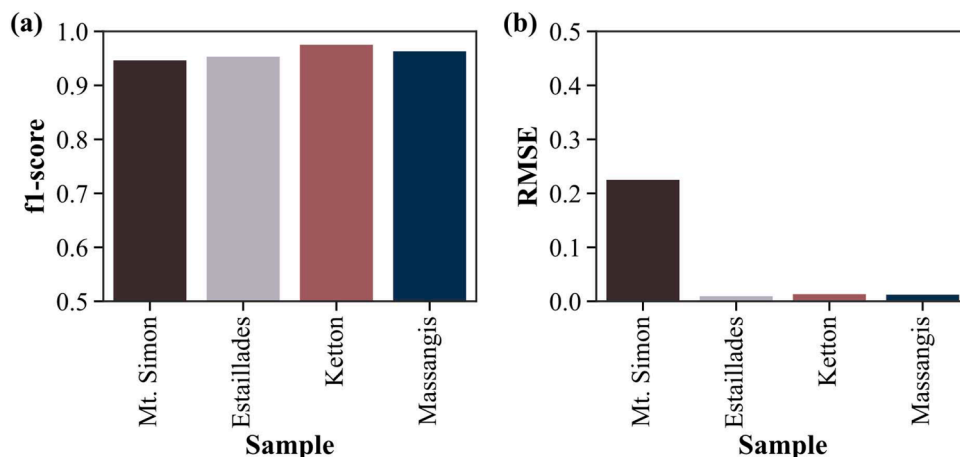
Fig. 10. The accuracy metrics of (a) f1-score and (b) root mean squared error (rmse) used to assess PoreSkel's performance in extracting skeletons and distance maps from unseen data. The unseen data, which include the images of Estaillades, Ketton, and Massangis carbontes, as well as the Mt. Simon sandstone, were excluded from the training and the primary test and validation.
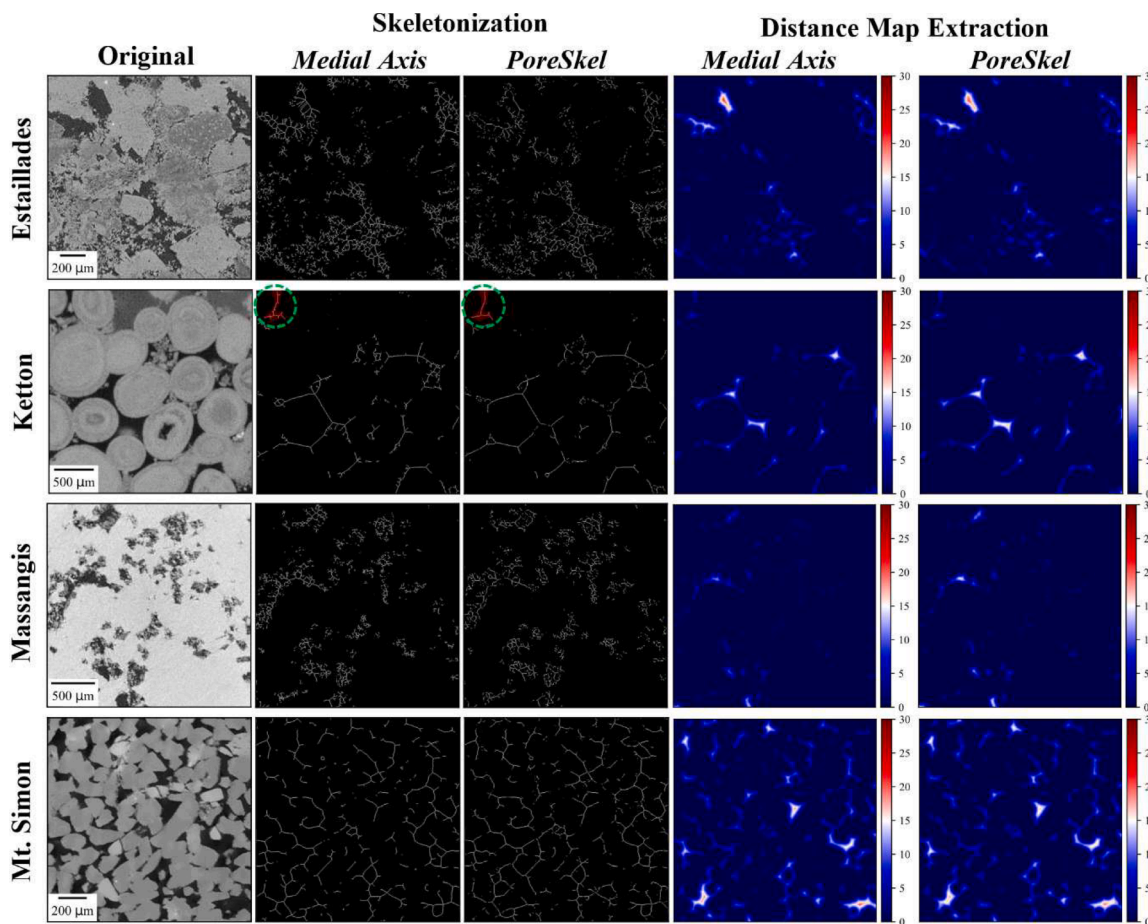
Fig. 11. Visual comparison of skeleton and distance maps extracted using PoreSkel and base-case data for unseen carbonate (Estaillades, Ketton, and Massangis samples) and sandstone (Mt. Simon) samples. The distances are color-coded, ranging from dark blue to dark red. As marked by circles on the Ketton sample (as an example), the PoreSkel could effectively detect the boundary perturbations.

Both the f1-score and IoU represent the ratios of the overlapping area between two images to the total pixel count and to the union of the areas between images, respectively. Let's denote the predicted and base-case (medial axis) images as P and B. The corresponding f1-score and IoU can be calculated by the subsequent equations:

$$f1 - score = \frac{2|P \cap B|}{|P| + |B|} \tag{2}$$

$$IoU = \frac{|P \cap B|}{|P \cup B|} \tag{3}$$

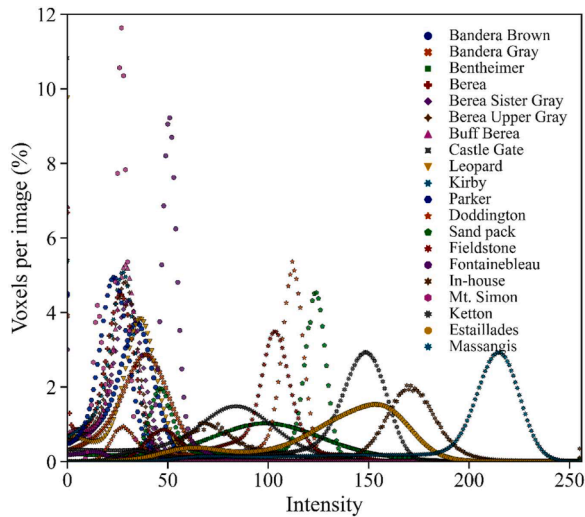**Fig. A.1.** The histogram of the databank (i.e., the collection of x-ray images of 20 rock samples) used in this study. The figure shows the percentage of voxels having certain intensities in each 3D image. There are high discrepancies between the intensities of various rocks, highlighting the challenge of training single models for skeletonization and distance map extraction across all samples.

The Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) can be utilized to estimate the deviation of the predicted distance maps from the baseline or Euclidean distance as:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (P_i - B_i)^2 \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_i - B_i)^2} \tag{5}$$

where, 'i' and 'N' denote the pixel number and total number of pixels, respectively. The higher the accuracy, f1-score, and Intersection over Union (IoU) values, the lower the discrepancies between the predicted and baseline images. Conversely, higher MSE and RMSE metrics imply a greater deviation from the base-case.

The aforementioned metrics were computed for each pair of predicted and baseline images, with the averages taken across all images subsequently. The mean f1-scores and RMSE values for the training (comprising 1200 images), validation (300 images), and testing (8240 images) stages of the model are presented in Fig. 4, with additional metrics detailed in Table A.1. The data reveals mean f1-scores for skeletonization of 0.974, 0.964, and 0.964; and IoUs of 0.974, 0.974, and 0.974 for the training, validation, and testing stages, respectively. Additionally, the distance maps yielded RMSE values of 0.021, 0.022, and 0.057, respectively. The negligible differences between the training and testing phases suggest that PoreSkel is resilient to overfitting. Overall, these results underscore the impressive performance of deep learning in extracting the skeleton and distance map of porous media images, even when unprocessed grayscale images are employed as inputs.

It is noteworthy that the allocation of the skeleton and background classes in the images is significantly skewed, with the skeleton representing the minority class. This disparity presents a challenge for PoreSkel as there is potential for bias toward the more dominant class. To understand how this class imbalance influenced the derived metrics, we calculated the f1-score for each individual class. Fig. 5 presents these results in a box plot format. The median f1-score, as depicted by the line within the interquartile range (IQR), stands at 0.975 and 0.949 for the skeleton and background classes, respectively. This outcome implies

that PoreSkel effectively manages the class imbalance.

Additionally, the f1-score and RMSE were computed for individual rock samples to provide a more granular assessment of PoreSkel. The results are illustrated in Fig. 6, with Table A.2 providing further details. The figure suggests the model is highly capable of extracting the skeleton and distance map, as all samples exhibit f1-scores exceeding 0.954 and RMSEs below 0.310.

3.2. Visual evaluation

The aforementioned statistical metrics were derived through a comparison between the predicted and base-case data. As previously highlighted, the base-case data can be prone to errors based on the generation technique and the image processing procedure (filtration and segmentation), which could impact the computed metrics. Thus, a thorough visual inspection was considered for a comprehensive evaluation. Fig. 7 showcases examples of PoreSkel outcomes alongside the base-case data (refer to Fig. A.2 for additional examples). Evidently, deep learning has detected the stochastic context of the pore space, extracting the skeleton and distance map within the midst of the pore space. As highlighted by squares in Fig. 7, PoreSkel demonstrates superior performance in some areas compared to the base case (highlighted by the squares). A detailed examination of the Castle Gate skeletons reveals that the model accounts for discontinuities introduced by the minerals present in the pore space (see dashed circles in the Castle Gate sample). Furthermore, the circles highlighted in the Doddington sample demonstrate the model's ability to preserve topology by addressing the perturbations on the pore boundaries – a significant challenge for the medial axis technique. Therefore, PoreSkel enhances the depiction of porous media and can potentially yield more representative pore networks.

It is important to note that the efficacy of PoreSkel can be influenced by the quality of the input images. For example, artificial contrasts caused by salt and pepper (i.e., impulse) noise could lead the model astray in identifying pore space, subsequently affecting skeleton extraction. This issue could be mitigated by applying suitable filters, such as non-local mean, to the input images (Refer to Fig. 8).

3.3. Skeleton analysis

We assessed topological parameters such as connectivity, coordinate number, and node count from medial axis and PoreSkel-extracted skeletons of various rock samples (i.e., a volume of 512^3 cubes per sample). To accomplish this, we generated graph representations of the 2D images. These graphs were then presented as adjacency lists, outlining the indices of each node (junction) in the skeleton and their corresponding neighbors. We employed the breadth-first search (BFS) algorithm to calculate these parameters. BFS, a graph traversal algorithm, explores the skeleton starting from the root node and examines adjacent nodes queued in a list. This approach is commonly used in graph analysis to identify nodes, branches, and endpoints. The determined parameters are displayed in the box plots of Fig. 9. As illustrated in this figure, PoreSkel-extracted skeletons possess fewer nodes and coordinate numbers than the medial axis. This is attributed to the medial axis's high sensitivity to perturbations. Minor perturbations can generate false nodes (pores) and branches (throats), leading to inaccurate representation of the pore space. In contrast, PoreSkel skeletons display a higher number of connected skeletons as PoreSkel accounts for minerals found in the pores. It's noteworthy that these distinctions may have significant implications for the measurement of petrophysical properties, particularly absolute permeability. A network characterized by artificially elevated coordination numbers, as observed with the medial axis, can potentially lead to an overestimation of permeability due to the greater number of interconnections between pores. Furthermore, the divergence between the network's characteristics and the actual morphology of porous media may introduce deviations in predicting fluid flow pathways and trapping
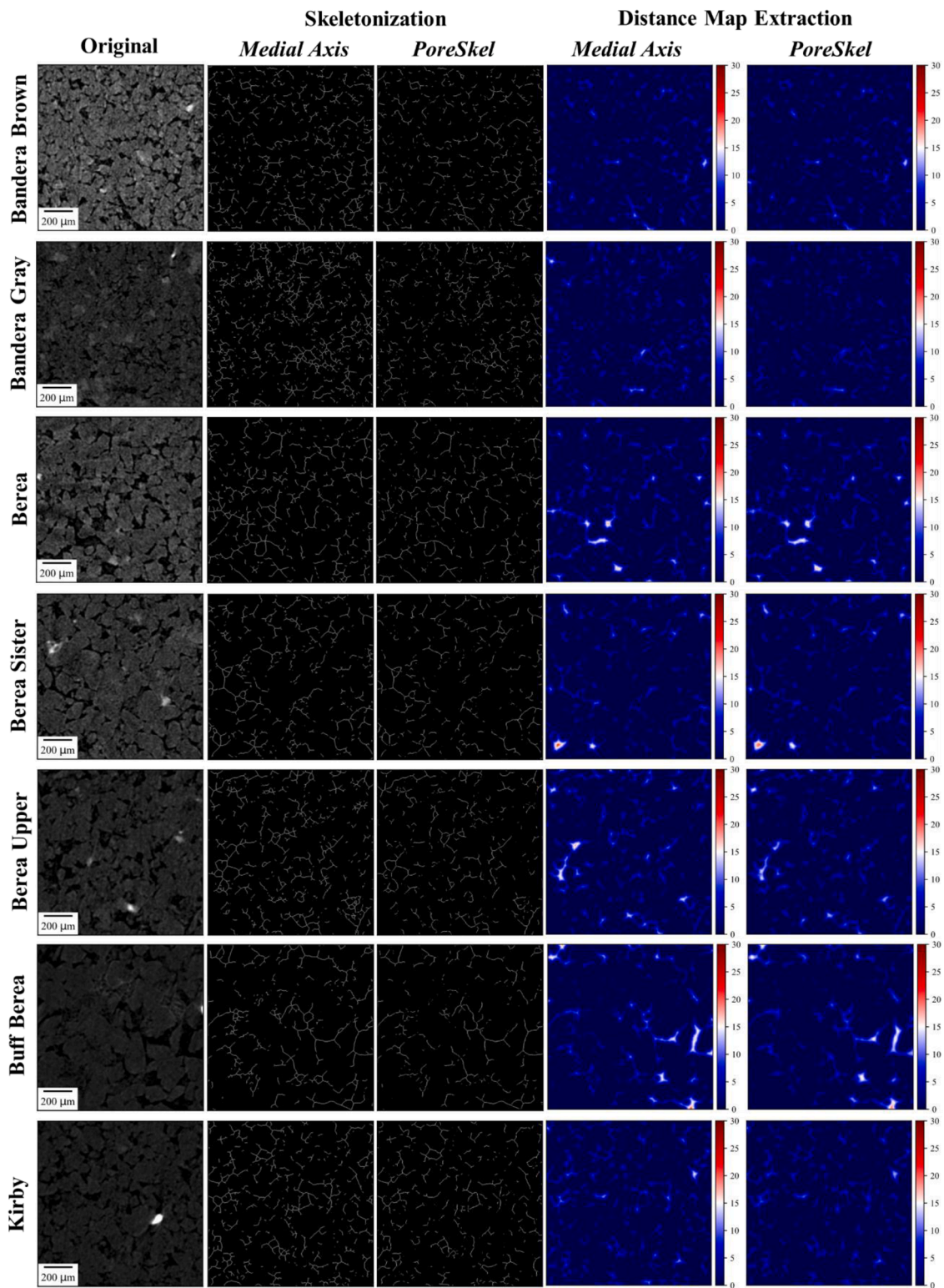
Fig. A.2. A visual description of the skeleton and distance maps extracted by the medial axis method and PoreSkel for several samples. The distances are demonstrated in a spectrum of colors ranging from dark blue to dark red.
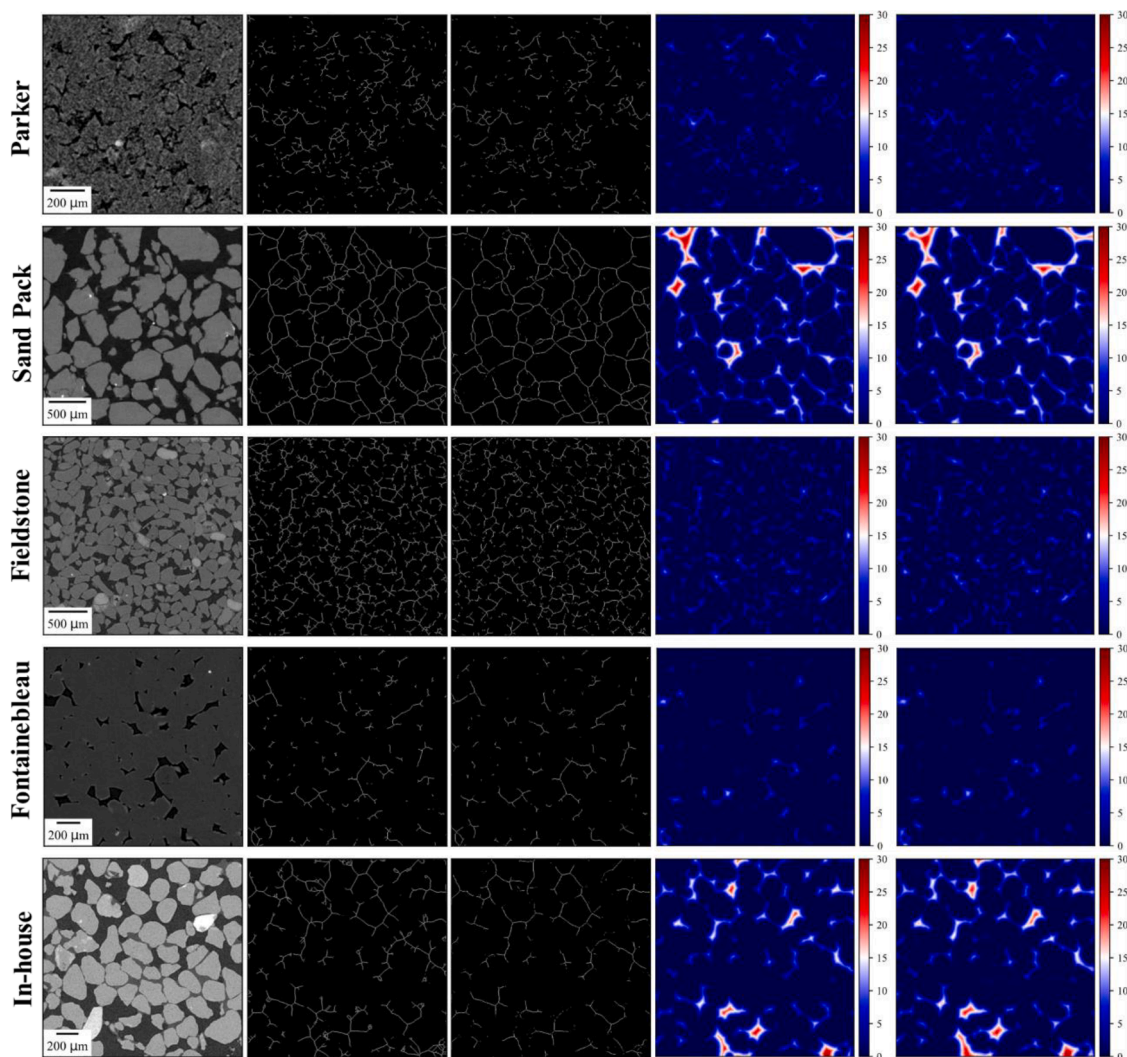
Fig. A.2. (*continued*).

Table A.1

The performance of the PoreSkel for skeletonization and distance map extraction in the train, validation, and test steps.

	Skeletonization			Distance Map Extraction	
	Accuracy	Mean IoU	f1-score	MSE	RMSE
Train	0.986	0.974	0.974	0.145	0.021
Validation	0.985	0.974	0.964	0.147	0.022
Test	0.985	0.974	0.964	0.206	0.057
Avg.	0.985	0.974	0.967	0.166	0.033

mechanisms. These differences may subsequently influence the multiphase flow parameters such as relative permeability and capillary pressure.

3.4. Generality of the model

As previously stated, PoreSkel was trained, validated, and tested using images of 16 different sandstones. To evaluate the model's ability to generalize to unseen images, we introduced 512^3 cubes from three carbonates (Estaillades, Ketton, and Massangis) and one sandstone (Mt. Simon) to the models. The results were compared to the base-case data, with findings illustrated in Figs. 10 and 11 (see Table A.3 for more information). As shown in Fig. 10, the skeletons and distance maps were

extracted with an average f1-score and RMSE of 0.961 and 0.066, respectively. Consequently, the models' performance did not significantly decline when faced with unseen data. A visual examination of the results in Fig. 11 suggests that PoreSkel accurately identified the pore space and extracted the skeleton and distance map in relation to the boundaries. Additionally, a closer look at the Ketton sample reveals that PoreSkel effectively addresses the boundary perturbations, a common challenge with the medial axis technique (see dashed circles).

Based on the aforementioned assessments, the PoreSkel model is capable of creating skeletonization and distance map extraction from grayscale images of porous media. This model provides a more comprehensive representation of pore space, ensuring that minerals within the pores are not overlooked. Additionally, the model remains unaffected by distortions in the pore boundaries, a complication that traditionally hinders the medial axis technique (e.g., shown in Figs. 7 and 11). Owing to its extensive training and testing on 20 distinct sandstone and carbonate samples, PoreSkel boasts high generality. As a result, this approach can be broadly applied for pore network extraction from diverse porous materials, including but not limited to, sedimentary rocks, soil, membranes, and gas diffusion layers in fuel cells.

4. Conclusions

This study aimed to use deep learning in the characterization of porous media, thereby enhancing pore network analysis and modeling.

Table A.2
The performance of the PoreSkel for skeletonization and distance map extraction for individual rock samples.

Sample	Subdata	Skeletonization		Distance map extraction	
		IoU	f1-score	MSE	RMSE
Bandera Brown	Train	0.973	0.963	0.089	0.205
	Validation	0.973	0.964	0.094	0.214
	Test	0.972	0.963	0.098	0.010
Bandera Gray	Train	0.973	0.964	0.281	0.436
	Validation	0.966	0.955	0.271	0.424
	Test	0.967	0.957	0.265	0.070
Bentheimer	Train	0.973	0.963	0.109	0.094
	Validation	0.978	0.970	0.111	0.098
	Test	0.976	0.969	0.362	0.131
Berea	Train	0.973	0.964	0.077	0.149
	Validation	0.977	0.969	0.077	0.148
	Test	0.976	0.965	0.123	0.015
Berea Sister Gray	Train	0.967	0.956	0.107	0.170
	Validation	0.977	0.969	0.128	0.183
	Test	0.978	0.969	0.124	0.015
Berea Upper Gray	Train	0.967	0.956	0.185	0.231
	Validation	0.975	0.966	0.209	0.242
	Test	0.977	0.968	0.266	0.071
Buff Berea	Train	0.966	0.955	0.147	0.189
	Validation	0.981	0.972	0.157	0.200
	Test	0.986	0.969	0.307	0.094
Castle Gate	Train	0.967	0.955	0.094	0.117
	Validation	0.975	0.967	0.100	0.121
	Test	0.975	0.967	0.144	0.021
Doddington	Train	0.977	0.969	0.091	0.140
	Validation	0.971	0.962	0.099	0.142
	Test	0.971	0.961	0.221	0.049
Fieldstone	Train	0.978	0.970	0.075	0.199
	Validation	0.961	0.949	0.077	0.198
	Test	0.960	0.948	0.076	0.006
Fontainebleau	Train	0.977	0.969	0.018	0.100
	Validation	0.989	0.986	0.020	0.104
	Test	0.989	0.986	0.018	0.000
In-house sand pack	Train	0.977	0.969	0.461	0.152
	Validation	0.964	0.951	0.422	0.148
	Test	0.964	0.950	0.487	0.237
Kirby	Train	0.978	0.971	0.140	0.243
	Validation	0.975	0.963	0.137	0.243
	Test	0.974	0.964	0.170	0.029
Leopard	Train	0.978	0.971	0.152	0.101
	Validation	0.980	0.974	0.157	0.101
	Test	0.980	0.973	0.317	0.101
Parker	Train	0.977	0.970	0.102	0.217
	Validation	0.978	0.968	0.112	0.220
	Test	0.980	0.972	0.094	0.009
Sand pack	Train	0.977	0.969	0.204	0.108
	Validation	0.962	0.947	0.177	0.101
	Test	0.962	0.948	0.227	0.051

In contrast to the traditional use of segmented (binarized) images, the developed model, PoreSkel, leverages deep learning to infer the skeleton and distance map of porous media from grayscale images, thereby eliminating the need for traditional image processing methods.

A diverse collection of x-ray images, comprising 512^3 cubes obtained from 20 different sandstones and carbonates, were gathered for this purpose to cover a wide range of pore structures and topological attributes. A fraction of 15.6 % of sixteen sandstones was designated for training, with the remaining samples reserved for validation and testing.

Additionally, four samples - three carbonates and one sandstone - were set aside as unseen test data, not included in the aforementioned procedure. PoreSkel was evaluated using both visual inspection and statistical metrics and topological parameters. The results are as follows:

- The PoreSkel model demonstrated high accuracy and flexibility, as demonstrated by the statistical metrics calculated for each phase of the model's deployment. The mean f1-scores for skeletonization remained consistently high at 0.974, 0.964, and 0.964 for the training, validation, and testing stages, respectively. In parallel, the distance maps achieved low RMSE values of 0.021, 0.022, and 0.057, respectively. Despite the inherent class imbalance presented by the skeleton and background classes, PoreSkel could effectively navigate this disparity, achieving a median f1-score of 0.975 for the skeleton class.
- In terms of visual evaluation, PoreSkel demonstrated a superior performance in capturing the stochastic context of the pore space, as well as addressing the perturbations on the pore boundaries, a significant challenge for the medial axis technique. This was further corroborated by the skeleton analysis, which revealed that PoreSkel-extracted skeletons possessed fewer nodes and coordinate numbers than those obtained from the medial axis, but a larger number of connected skeletons.
- PoreSkel's generalizability was also validated, which maintained high accuracy when introduced to unseen data from additional sandstone and carbonate samples, indicating the potential of PoreSkel to be applied for pore network extraction across a diverse array of porous materials.

To our knowledge, PoreSkel is the first methodology capable of extracting both skeletons and distance maps from grayscale images. Opportunities exist for refining PoreSkel to achieve superior precision and resilience against image artifacts. Future research should further explore expanding the model's applicability to 3D images.

Author statement

- Mehdi Mahdaviara: Writing – original draft, Methodology, Software, Visualization
- Mohammad Sharifi: Supervision, Conceptualization, Writing – review & editing
- Amir Raoof: Supervision, Conceptualization, Writing – review & editing

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.

Data availability

Data will be made available on request.

Table A.3
The performance of the PoreSkel for skeletonization and distance map extraction for the unseen rock samples.

Sample	Skeletonization					Distance Map Extraction	
	Accuracy	mean IoU	f1-score	f1-score (background)	f1-score (skel.)	MSE	RMSE
Mt. Simon	0.971	0.959	0.948	0.945	0.979	0.476	0.226
Estaillades	0.974	0.960	0.955	0.965	0.976	0.104	0.011
Ketton	0.989	0.981	0.977	0.970	0.987	0.122	0.015
Massangis	0.978	0.968	0.965	0.962	0.982	0.117	0.014
Average	0.978	0.967	0.961	0.961	0.981	0.205	0.066

Appendix

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. J. Pet. Sci. Eng. 184, 106514.

Andrew, M., Bijeljic, B., Blunt, M., 2020. In: Doddington sandstone, Digital Rocks Portal https://doi.org/10.17612/NHDA-WZ02.

Armstrong, R.T., et al., 2016. Beyond Darcy's law: the role of phase topology and ganglion dynamics for two-fluid flow. Phys. Rev. E 94 (4), 043113.

Bakke, S., Øren, P.E., 1997. 3-D pore-scale modelling of sandstones and flow simulations in the pore networks. SPE J. 2 (02), 136–149.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning, 4. Springer.

Blunt, M.J., et al., 2013. Pore-scale imaging and modelling. Adv. Water Resour. 51, 197–216.

Blunt, M.J., 2001. Flow in porous media — Pore-network models and multiphase flow. Curr. Opin. Colloid Interface Sci. 6 (3), 197–207.

Blunt, M.J., 2017. Multiphase Flow in Permeable Media: a Pore-Scale Perspective. Cambridge University Press, Cambridge.

Bryant, S., Blunt, M., 1992. Prediction of relative permeability in simple porous media. Phys. Rev. A 46 (4), 2004.

Bultreys, T. and De Boever, W., 2020. Belgian Fieldstone, Digital Rocks Portal. https://doi.org/10.17612/C13T-AH38.

Bultreys, T., De Boever, W., Cnudde, V., 2016. Imaging and image-based fluid transport modeling at the pore scale in geological materials: a practical introduction to the current state-of-the-art. Earth Sci. Rev. 155, 93–128.

Bultreys, T., 2016. Estaillades carbonate #2, Digital Rocks Portal. https://doi.org/10.17612/P7C09.

Bultreys, T., 2016. Massangis Jaune carbonate. Digital Rock Portal. https://doi.org/10.17612/P7RG6N.

Bultreys, T., 2016. Savonnières carbonate, Digital Rocks Portal. https://doi.org/10.17612/P7W88K.

Delerue, J.F., Perrier, E., Yu, Z.Y., Velde, B., 1999. New algorithms in 3D image analysis and their application to the measurement of a spatialized pore size distribution in soils. Phys. Chem. Earth Part A. 24 (7), 639–644.

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26 (3), 297–302.

Fathi, H., Raoof, A., Mansouri, S., 2017a. Insights into the role of wettability in cathode catalyst layer of proton exchange membrane fuel cell; pore scale immiscible flow and transport processes. J. Power Sources 349, 57–67.

Fathi, H., Raoof, A., Mansouri, S., van Genuchten, M.T., 2017b. Effects of porosity and water saturation on the effective diffusivity of a cathode catalyst layer. J. Electrochem. Soc. 164 (4), F298.

Géron, A., 2022. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, Inc.

Garing, C., Ajo-Franklin, J., Voltolini, M., Benson, S., 2017. In: Residual air in air/brine systems, Digital Rocks Portal https://doi.org/10.17612/P7ZW24.

Gonzalez, F.J. and Balajewicz, M., 2018. Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems. arXiv preprint. *arXiv:1808.01346*.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT press.

Graczyk, K.M., Matyka, M., 2020. Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. Sci. Rep. 10 (1), 21488.

Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313 (5786), 504–507.

Ioannidis, M., Chatzis, I., 2000. On the geometry and topology of 3D stochastic porous media. J. Colloid Interface Sci. 229 (2), 323–334.

Jaccard, P., 1912. The distribution of the flora in the alpine zone.1. New Phytol. 11 (2), 37–50.

Jiang, Z., et al., 2007. Efficient extraction of networks from three-dimensional porous media. Water Resour. Res. 43 (12).

Joekar Niasar, V., Hassanizadeh, S., Pyrak-Nolte, L., Berentsen, C., 2009. Simulating drainage and imbibition experiments in a high-porosity micromodel using an unstructured pore network model. Water Resour. Res. 45 (2), 2–15.

Joekar-Niasar, V., Hassanizadeh, S., 2012. Analysis of fundamentals of two-phase flow in porous media using dynamic pore-network models: a review. Crit. Rev. Environ. Sci. Technol. 42 (18), 1895–1976.

Joekar-Niasar, V., Hassanizadeh, S.M., Dahle, H., 2010. Non-equilibrium effects in capillarity and interfacial area in two-phase flow: dynamic pore-network modelling. J. Fluid Mech. 655, 38–71.

Kingma, D.P. and Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint. *arXiv:1412.6980*.

Kingma, D.P. and Welling, M., 2013. Auto-encoding variational bayes. arXiv preprint. *arXiv:1312.6114*.

Kohanpur, A.H.A.V., Albert, Crandall, D., 2019. In: Micro-CT images of a heterogeneous Mt. Simon sandstone sample, Digital Rocks Portal. https://doi.org/10.17612/1dvh-1n64.

Koplik, J., Lasseter, T., 1985. Two-phase flow in random network models of porous media. Soc. Pet. Eng. J. 25 (01), 89–100.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Lee, D.T., 1982. Medial axis transformation of a planar shape. IEEE Trans. Pattern Anal. Mach. Intell. (4), 363–369.

Lindquist, W.B., Lee, S.M., Coker, D.A., Jones, K.W., Spanne, P., 1996. Medial axis analysis of void structure in three-dimensional tomographic images of porous media. J. Geophys. Res. Solid Earth 101 (B4), 8297–8310.

Lohou, C., Bertrand, G., 2005. A 3D 6-subiteration curve thinning algorithm based on P-simple points. Discret. Appl. Math. 151 (1), 198–228.

Lu, X., Tsao, Y., Matsuda, S., Hori, C., 2013. Speech enhancement based on deep denoising autoencoder. Interspeech 2013, 436–440.

Mahdaviara, M., Sharifi, M., Bakhshian, S., Shokri, N., 2022. Prediction of spontaneous imbibition in porous media using deep and ensemble learning techniques. Fuel 329, 125349.

Mahdaviara, M., Shojaei, M.J., Siavashi, J., Sharifi, M., Blunt, M.J., 2023. Deep learning for multiphase segmentation of X-ray images of gas diffusion layers. Fuel 345, 128180.

Mahmoodlu, M.G., Raoof, A., Bultreys, T., Van Stappen, J., Cnudde, V., 2020. Large-scale pore network and continuum simulations of solute longitudinal dispersivity of a saturated sand column. Adv. Water Resour. 144, 103713.

Mason, G., Morrow, N.R., 1991. Capillary behavior of a perfectly wetting liquid in irregular triangular tubes. J. Colloid Interface Sci. 141 (1), 262–274.

Morgenthaler, D.G., 1981. Three-dimensional simple points: sereial erosion, parallel thinning and skeletonization. University of Maryland 1005.

Najafi, A., et al., 2023. Using computational fluid dynamics to compute the pore-scale CO_2-brine relative permeability. Fuel 341, 127715.

Neumann, R., Andreeta, M., Lucas-Oliveira, E., 2020. In: 11 Sandstones: raw, filtered and segmented data, Digital Rocks Portal. https://doi.org/10.17612/f4h1-w124.

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al., 2018. Attention u-net: learning where to look for the pancreas. arXiv preprint. *arXiv: 1804.03999*.

Pereira, G., Pinczewski, W., Chan, D., Paterson, L., Øren, P., 1996. Pore-scale network model for drainage-dominated three-phase flow in porous media. Transp. Porous Media 24, 167–201.

Pudney, C., 1998. Distance-ordered homotopic thinning: a skeletonization algorithm for 3D digital images. Comput. Vis. Image Underst. 72 (3), 404–413.

Raoof, A., Hassanizadeh, S.M., 2010. A New Method for Generating Pore-Network Models of Porous Media. Transp Porous Media 81 (3), 391–407.

Raoof, A., Nick, H.M., Hassanizadeh, S.M., Spiers, C.J., 2013. PoreFlow: a complex pore-network model for simulation of reactive transport in variably saturated porous media. Comput. Geosci. 61, 160–174.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer, pp. 234–241.

Saito, T., 1995. A sequential thinning algorithm for three dimensional digital pictures using the Euclidean distance transformation. In: Proceedings of the 9th Scandinavian Conference on Image Analysis, 1995, p. 6.

Scanziani, A., Singh, K., Blunt, M., 2018. Water-wet three-phase flow micro-CT tomograms. Digital Rocks Portal. https://doi.org/10.17612/P7HT11.

Singh, A., Mostaghimi, P., Regenauer-Lieb, K., Armstrong, R., Walsh, S., 2020. In: Grayscale REV analysis, Digital Rocks Portal. https://doi.org/10.17612/8F3W-EM15.

Thovert, J., Salles, J., Adler, P., 1993. Computerized characterization of the geometry of real porous media: their discretization, analysis and interpretation. J. Microsc. 170 (1), 65–79.

van Dijke, M.I.J., Sorbie, K.S., 2006. Existence of fluid layers in the corners of a capillary with non-uniform wettability. J. Colloid Interface Sci. 293 (2), 455–463.

Vogel, H.J., Roth, K., 2001. Quantitative morphology and network representation of soil pore structure. Adv. Water Resour. 24 (3), 233–242.

Da Wang, Y., Chung, T., Armstrong, R.T. and Mostaghimi, P., 2020. ML-LBM: machine learning aided flow simulation in porous media. arXiv preprint. *arXiv:2004.11675*.

Xiong, Q., Baychev, T.G., Jivkov, A.P., 2016. Review of pore network modelling of porous media: experimental characterisations, network constructions and applications to reactive transport. J. Contam. Hydrol. 192, 101–117.

Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., 2018. Unet++: a nested u-net architecture for medical image segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer, pp. 3–11.