A computationally efficient modeling of flow in complex porous media by coupling multiscale digital rock physics and deep learning: Improving the tradeoff between resolution and field-of-view

Iman Nabipour [a], Amir Raoof [b], Veerle Cnudde [b,c], Hamed Aghaei [d], Jafar Qajar [a,b,*]

[a] Department of Petroleum Engineering, School of Chemical and Petroleum Engineering, Shiraz University, Shiraz, Iran
[b] Department of Earth Sciences, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands
[c] Department of Geology, PProGRess, Ghent University, Ghent, Belgium
[d] Department of Earth and Environmental Sciences, Università di Pavia, Pavia, Italy

ARTICLE INFO

ABSTRACT

Digital rock physics is at the forefront of characterizing porous media, leveraging advanced tomographic imaging and numerical simulations to extract key rock properties like permeability. However, fully capturing the heterogeneity of natural rocks necessitates imaging increasingly larger sample volumes, presenting a significant challenge. Direct numerical simulations at these scales become either prohibitively expensive or computationally unfeasible due to limitations in resolution and field of view (FOV). This issue is particularly pronounced in carbonate rocks, known for their complex, multiscale pore structures, which exacerbate the resolution-FOV tradeoff. To address this, we introduce a machine learning strategy that merges multiscale imaging data from various resolutions with a 3D convolutional neural network (CNN) model. This approach is innovative in its ability to identify cross-scale correlations, thereby enabling the estimation of transport properties in larger volumes—properties that are difficult to simulate directly—using trainable proxies. The integration of multiscale imaging with deep learning allows for accurate permeability predictions at scales beyond those feasible with traditional direct simulation methods. By employing transfer learning across different scales during the training phase, our multiscale machine learning model achieves robust performance, with an R^2 exceeding 0.96 when evaluated on diverse lower-resolution domains with larger FOVs. Notably, this method significantly enhances computational efficiency, reducing the computational time by orders of magnitude. Originally developed for the intricate pore structures of carbonate rocks, our approach shows promise for application to a wide range of multiscale porous media, offering a viable solution to the longstanding tradeoff between imaging resolution and FOV in digital rock physics.

1. Introduction

Characterizing transport properties in porous materials represents a significant challenge across various scientific and engineering domains, including capillary water absorption in the building stones of cultural heritages (Çelik and Sert, 2021), hydrocarbon recovery from underground resources (Qajar and Arns, 2022), geological storage of CO_2 and hydrogen (Aftab et al., 2022; Schultz et al., 2023), reactive transport in the subsurface (Poonoosamy et al., 2022), and blood filtration in kidneys (Kahshan et al., 2020). The complexity of pore space geometry, often not fully understood and varying across multiple scales due to heterogeneities, plays a crucial role in this challenge. Understanding the impact of

these small-scale heterogeneities on reservoir performance is a complex task that requires the integration of multidisciplinary studies (Khodja et al., 2020).

In recent years, the use of pore-scale imaging and modeling, known as digital rock physics (DRP), has emerged as a powerful technique for accurately quantifying flow through complex porous media. X-ray micro-computed tomography (μ-CT) has become the preferred method for producing 3D pore structures with a large field of view (FOV), facilitating representative property calculations (Bultreys et al., 2016; Cnudde and Boone, 2013; Withers et al., 2021). However, the resolution limitation of digital images and the difficulty in characterizing pore space features in a computationally manageable model pose significant

---

challenges in DRP (Blunt et al., 2013). Pore-scale models bridge the gap between sub-pore and core-scale estimates of transport properties, with efforts aimed at making these techniques cost-effective. Accurate modeling of extensive porous systems requires information across various length scales, from nanometres to meters (Chaaban et al., 2022).

For media with a wide range of pore sizes, a single μ-CT image may not capture all pores. A top-to-bottom strategy involving acquiring coarse-scale 3D CT images with large FOVs followed by high-resolution μ-CT images of selected locations can overcome this limitation. Additionally, combining various imaging modalities, such as 2D SEM and/or 3D FIB-SEM, allows for collecting higher-resolution data (Anderson et al., 2020; De Boever et al., 2015). Transport features calculated from high-resolution pore structures are integrated back into the large-scale image, using information from both resolved and unresolved rock fabrics. This multiscale DRP analysis on the large-scale image determines the desired attributes (Mehmani et al., 2020; Sun et al., 2017; Sun et al., 2017; Sungkorn et al., 2015). Multiscale DRP offers a promising approach for characterizing complex porous materials like sedimentary rocks. However, the accuracy of the model, both numerically and in predicting actual properties at interest scales, such as effective permeability, is closely linked to the resolution of the image (Chen and Zhou, 2017). Classic absorption-based X-ray μ-CT imaging faces inherent FOV and resolution trade-offs, with a resolution three orders of magnitude below the FOV achievable. Thus, novel methodologies are essential for fully understanding, characterizing, and modeling multiscale porous systems (Jackson et al., 2021).

In the context of the aforementioned challenges, machine learning, particularly deep learning, has emerged as a powerful tool in data-driven, image-based calculations, propelled by advancements in GPU processing, optimization techniques, and neural network designs (Da Wang et al., 2020; Gärttner et al., 2021; Kamrava et al., 2020; Tembely et al., 2021; Tembely et al., 2020; Kalule et al., 2023; Gupta and Gupta, 2021). These methods have been instrumental in exploring the influence of pore-scale geometrical characteristics on macroscale fluid percolation processes within porous media. Deep learning algorithms have notably enhanced image resolution and expedited image analysis and image-based calculations (Da Wang et al., 2021; Sit et al., 2020; Ma et al., 2023). Deep learning's role in pore-scale modeling has evolved into two distinct approaches. The first approach involves using pore-scale modeling to estimate transport properties, which can then be upscaled. This method primarily focuses on extracting geometrical information from pore-scale images, either directly or indirectly, followed by employing regression techniques to predict physical properties, with permeability being a key focus (Kamrava et al., 2020; Tembely et al., 2020; Erofeev et al., 2019; Sudakov et al., 2019; Alqahtani et al., 2020; Rabbani et al., 2020; Alqahtani et al., 2021; Tian et al., 2021; Hong and Liu, 2020; Fu et al., 2023; Najafi et al., 2021; Jiang et al., 2022; Tang et al., 2022; Zhang et al., 2022; Tian et al., 2021).

The second approach utilizes convolutional neural networks (CNNs) either as a substitute for or in conjunction with flow simulation solvers to predict or accelerate pore-scale modeling. This results in direct predictions of velocity fields, which can be further processed to estimate permeability or model other transport phenomena (Da Wang et al., 2020; Santos et al., 2020; Santos et al., 2021; Ribeiro et al., 2020). Unlike traditional methods that rely on the geometric characterization of pore-scale images as input to artificial neural networks (ANNs), deep learning techniques for permeability prediction use the image itself as input to a CNN. This process characterizes the geometry, which is then directly fed into an ANN to leverage "hidden features" for permeability estimation. This approach offers a more comprehensive and flexible "end-to-end" implementation, independent of any extracted geometric features. The design incorporates an ANN in the form of dense layers at the end, enabling the structure to automatically identify and transfer geometric elements of interest directly to an ANN via the CNN. By entrusting data interpretation entirely to the network, the success of CNN-based regression models becomes more reliant on the architecture

and dataset rather than on the selection of input parameters. CNNs in this context refer to the combined CNN+ANN regression networks, highlighting their integral role in advancing pore-scale modeling through deep learning.

Permeability, which quantifies the resistance to fluid flow in porous media, is a variable with uncertainty, largely due to the complex, non-linear relationship between pore morphology and permeability. Capturing this relationship through machine learning, especially within the intricate geometries of pore-scale images, poses a significant challenge. Alqahtani et al. (2021) utilized ResNet and ResNext CNN networks to explore the potential of 3D regression in predicting permeability across various lithologies. Their study, focusing on sandstone images, achieved an R^2 value of 0.87 on the testing datasets, highlighting the effectiveness of these networks in permeability prediction. Kamrava et al. (2020) reported an R^2 value of 0.91, testing the accuracy of permeability estimation on 500 stochastically generated rock image samples with permeabilities ranging from 100 mD to 500 mD. Zhang et al. (2022) developed a semi-supervised machine learning approach combining an auto-encoder (AE) module with CNN to predict permeability from low-resolution images. Their AE-CNN model outperformed traditional CNN approaches, achieving an average R^2 value of 0.896 on the test dataset, thereby demonstrating high accuracy in permeability prediction.

In a comparative study, Tian et al. (2021) introduced a hybrid machine learning method that combines an artificial neural network (ANN) with a genetic algorithm (GA) to predict permeability based on pore structure parameters. While their ANN model achieved high R values on both training and testing sets, it's important to note that their dataset was prepared using artificial porous media. The concept of transfer learning has been identified as a promising technique to enhance the performance of CNNs in permeability estimation. Tang et al. (2022) introduced a CNN model that, through the integration of transfer learning, showed significant improvement in estimating the permeability of rocks from their three-dimensional images. Model fine-tuning, a deep learning technique proposed by Yosinski et al. (2014), was employed to adjust the weights of all layers in the pre-trained CNN using a new dataset, with further training conducted on 16 new samples matching the distribution of the testing sets. Santos et al. (2021) developed a computationally efficient multiscale neural network aimed at predicting velocity fields through hierarchical regression. However, their training utilized simple synthetic geometries to represent porous and vuggy media, artificially reducing the resolution of images.

Recent investigations have sought to address the challenge of balancing image resolution and field of view (FOV) in micro-CT imaging by employing deep learning super-resolution techniques to enhance the resolution of micro-CT images. Ma et al. (2023) introduced an innovative deep-learning method that significantly improves the resolution of micro-CT images of carbonate rocks. Their super-resolution algorithm, based on the diffusion model, effectively enhances image quality without introducing noise or blurring, showing great potential in generating clear and detailed pore networks essential for calculating the petrophysical properties of carbonate rocks. In a comparative study, Soltanmohammadi and Faroughi (2023) evaluated four advanced CNN-based super-resolution algorithms for their effectiveness in increasing the resolution of micro-CT images of heterogeneous carbonate rocks. They assessed the algorithms' accuracy in reconstructing high-resolution images and preserving petrophysical properties, such as porosity, using various evaluation metrics. While super-resolution techniques have demonstrated promising outcomes in improving digital rock images, they face limitations, including significant computational demands and challenges in reconstructing high-frequency details. This can lead to oversimplified textures in the super-resolved images, potentially resulting in the loss of crucial geological information (Bai and Berezovsky, 2020).

Most prior research has focused on 2D or simplified small 3D pore geometries, which may not accurately represent the complexity of real

porous materials like microporous carbonate rocks. Furthermore, many existing models assume a single-scale pore structure, which may not adequately capture the multiscale texture of carbonates. Given the proven effectiveness of CNN approaches in estimating transport properties in complex structures and their non-linear correlation with the morphological characteristics of porous media, training realistic 3D multiscale images of complex tight carbonates at different spatial resolutions, digitally aligned through an efficient custom CNN model architecture, remains a compelling proposition.

Our research is among the first to apply a multiscale-resolution aware CNN framework for permeability estimation in realistic, complex porous media, specifically focusing on heterogeneous tight carbonate rocks. This approach employs a novel transfer learning strategy across different image resolutions, encompassing a wide range of pore sizes and throat varieties. This significantly enhances the model's adaptability to various resolutions and improves the accuracy and generalizability of permeability predictions. By addressing the challenge of analyzing low-resolution micro-CT images without compromising permeability prediction accuracy, our study tackles the critical issue of the resolution-FOV trade-off, offering a novel solution to this long-standing problem and enabling more effective analysis of heterogeneous porous materials from coarser domains. The success of this study is attributed to the development of an advanced multiscale CNN model that leverages standard GPU (graphics processing unit) and hardware specifications to provide transport property estimates much faster than conventional direct numerical simulations (DNSs) without sacrificing accuracy. Capable of operating on lower-resolution domains with larger

FOVs, which DNSs find impractical, our approach has the potential to be adapted for a wide range of applications beyond tight carbonate rocks, opening new avenues in the broader field of digital rock physics and related fields.

2. General workflow

This research focuses on integrating multiscale μ-CT imaging with artificial intelligence via a deep learning approach paradigm. Our proposed method estimates the permeability across many near-representative heterogeneous volumes inside low-resolution images, providing information unattainable from conventional laboratory methods. In this context, the datasets under consideration consist of 3D stacked images produced by X-ray μ-CT scanning of real core plug samples. Our methodology leans on multiple image resolutions to detect larger-scale heterogeneity.

While lower-resolution binarized images might not entirely capture the pore structures, our assessment indicates that specific pore phase features inherently hold about the vital fluid flow pathways. As image resolution decreases, there is an increase in the fraction of intermediate (or microporous) grayscale voxels. Conversely, the fraction representing the clearly-defined macroporous phase decreases. At specific resolution thresholds, internal connections (such as pore throats), get included in the intermediate grayscale values, excluding direct modeling calculations. This study focuses on extracting information from images in this region where flow paths are not fully resolved. Furthermore, local averaging with each voxel at lower resolutions covers such a large
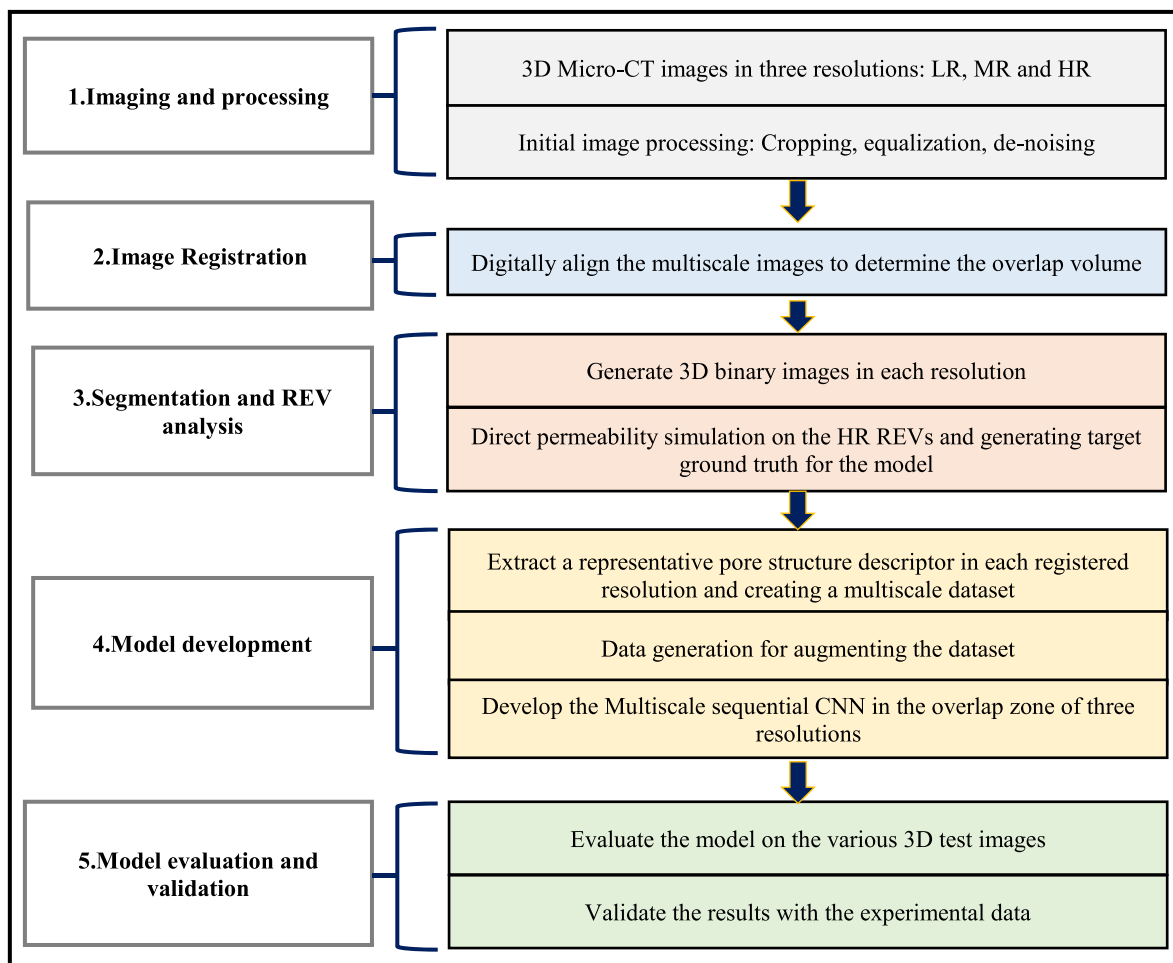

Fig. 1. Workflow of the multiscale CNN development process. The flowchart illustrates the sequential steps undertaken in the study. Each step encompasses detailed tasks, progressing from acquiring of 3D Micro-CT images in various resolutions to the final validation of the CNN model.

volume that all geometric information is lost. It appears unlikely that valuable information can be extracted from the lower-resolution images to predict transport properties.

Our proposed method is to develop a multiscale sequential convolutional network with a transfer learning scheme from high to low-resolution images and high-resolution computed permeability. This resultant permeability mapping becomes a pivotal input for extending permeability calculations to the core's entire scale and even further. The procedure adopted in this study consists of five important stages summarized in Fig. 1, and subsequent sections will delve deeper into the details of each section.

3. Rock sample, multiscale μ-CT imaging, and image analysis

The multiscale characterization was conducted on a borehole carbonate sample of 3.8 cm in diameter and 5 cm in length taken from the Asmari reservoir located in southwestern Iran. Fig. 2(a) displays a photo of the as-received whole core plug. The XRD (X-Ray Diffraction) tests indicated the rock was almost made of dolomite. The whole plug was imaged using a medical CBCT (Cone Beam Computed Tomography) scanner for the initial assessment, and no significant heterogeneity was observed (Fig. 3(a)). A subplug of 3 mm in diameter was extracted from the whole core, as shown in Fig. 2(b).

This subplug underwent μ-CT imaging at three different resolutions: 3 μm (low-resolution, termed LR), 2 μm (middle-resolution, termed MR), and 1 μm (high-resolution, termed HR) using μ-CT imaging facility at Imperial College London. To assess the pore size distribution (PSD) of the specimen, mercury intrusion porosimetry (MIP) measurements were executed on the subplug post its multi-resolution imaging, utilizing the Micromeritics' AutoPore IV 9510 mercury porosimeter, capable of reaching a peak pressure of 60,000 psia. The attributes of the subplug deduced using the MIP technique are tabulated in Table 1. Fig. 3(b) portrays both cumulative and incremental MIP graphs of the subplug, showcasing a monomodal PSD.

The sample's porosity spans pore sizes ranging from 200 nm to 4 μm. The cumulative MIP graph shows that approximately 95%, 80%, and 50% of the total sample porosity, corresponding to resolutions of 3, 2, and 1 μm, respectively, are reached solely via the sub-resolution pores (microporous phase). Consequently, there is a pronounced variance among the multi-resolution images explored in the context of the discernible macroporous phase and the associated image-based properties. It's pivotal to note that the designation of high- and low-resolution images is relative and hinges on the complexity of the specimen's pore space. The prudent choice between high and low-resolution voxel values is contingent on the pore space structure, with an emphasis on the pore size distribution.

3.1. Image processing

We undertook several image pre-processing steps prior to the principal tasks of registration and segmentation. Firstly, images were cropped to eliminate empty areas. Next, a histogram equalization was applied to enhance the contrast and balance of the histograms of the images. While this might compromise some structural data in the image (Almotairi, 2020), it remains indispensable for the ensuing stages. The third measure involved image enhancement through de-noising. We used a non-local means filter set with a one-pixel local neighborhood. This method relies on a weighted average of all the image pixels. Here, the weights are determined not by the proximity of pixels but by whole blocks, facilitating efficient noise reduction in tomographic images without diminishing contrast boundaries (Sun et al., 2017).

Within the workflow we devised, registration plays a crucial role. Distinct datasets with varying voxel sizes bridge high-resolution smaller volumes to low-resolution larger volumes. Evidently, when the scanning size diminishes, the resolution of the 3D digital core augments. Large field-of-view (FOV) low-resolution images assist in pinpointing the location and size of high-resolution scans with smaller FOV and physical sub-samples. Given the meticulous voxel-to-voxel registration of the images, this imaging tool permits the probing of cross-scale relationships, along with the high-fidelity imaging of notable sample variations across multiple scales. We utilized a 3D/3D image registration (Cui et al., 2020) to overlay the multiscale (multi-resolution) images of the core sample. Fig. 4 presents instances of 2D slices and 3D renderings of registered images at varying resolutions. Apart from x-ray micro-CT imaging and physical sub-sampling, the sample undergoes scaled-down x-ray micro-CT imaging.

In the segmentation phase, the thresholding step is often undertaken using a global histogram method, indicating open pores and microporous areas by identifying significant histogram patterns. As suggested by Verri et al. (2017), if the image resolution is sufficiently high and microporosity does not overwhelmingly surpass open pores, this process can be executed through fully automatic algorithms. Conversely, when the image resolution is lower and the rock structures are ambiguous, the thresholding procedure should be supplemented by visual image inspection, leveraging the user's expertise to confirm the estimation's accuracy. In our study, initial image segmentation was autonomously achieved using the Otsu method, a standout global thresholding technique known for its effectiveness in threshold-based segmentation, particularly in distinguishing between solid and pore spaces based on voxel intensity. It offers a straightforward, computationally efficient approach, which is advantageous when dealing with large datasets. Furthermore, in several DRP studies, this method was applied to segment the pore space (Andrä et al., 2013; Saxena et al., 2017; Abrosimov et al., 2021; Tawfeeq and Al-Sudani, 2020). Subsequently, it was refined through an attentive visual examination of the pore space
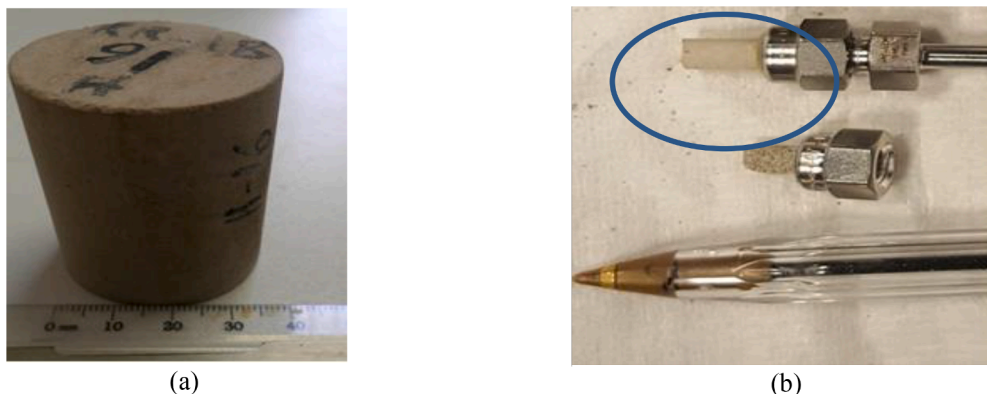

(a)   (b)

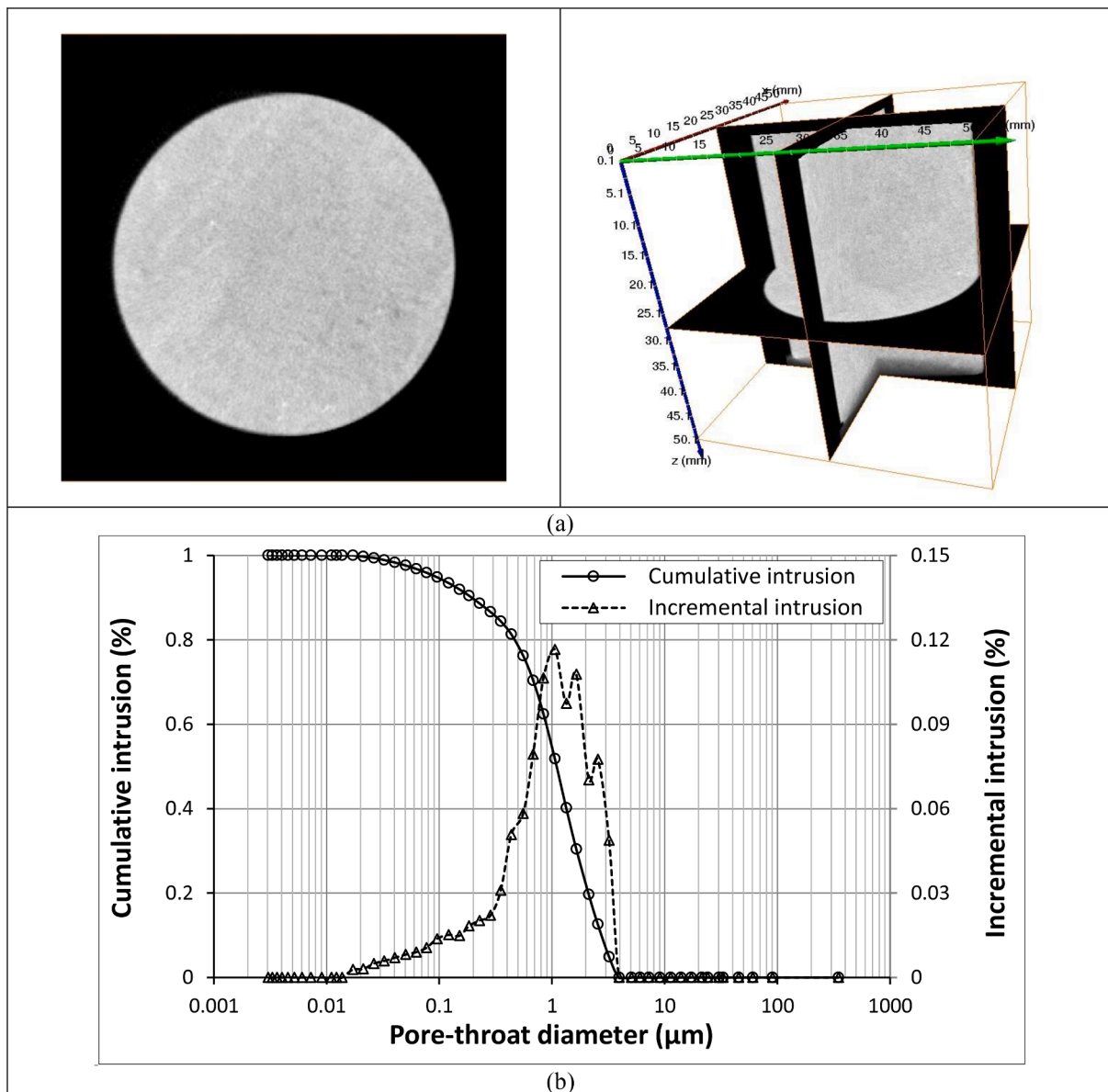**Fig. 2.** Photographs of (a) the whole core sample and (b) the mini-plug used for μ-CT imaging.

Fig. 3. (a) The medical CT scan of the whole plug, taken at a resolution of 100 μm/voxel; (b) The incremental and cumulative MIP curves of the sub-plug sample.

Table 1
The measured/estimated properties from mercury intrusion through the subplug sample.

Hg porosity (%)	Swanson permeability (md)	Threshold pressure (psi)	Total pore area (m²/g)	Tortuosity	Average pore diameter (nm)
11.075	1.289	45.967	0.506	2.637	399.257

(Miarelli and Della Torre, 2021), and ultimately, the computed porosity was validated by the MIP data. Fig. 5 displays the gray intensity histograms for the HR, MR, and LR tight carbonate images.

While the Otsu method may have limitations in handling transitional voxels related to the microporosity fraction, we addressed these through careful preprocessing and post-segmentation refinement. The choice of the Otsu method over several other advanced segmentation techniques, such as watershed segmentation, was driven by its suitability for our specific dataset and research objectives, balancing accuracy with computational efficiency and validation against lab data. The porosity

distribution of the multiscale segmented images post-registration is depicted in Fig. 6, representing local variations in porosity across the sample.

3.2. Direct simulation

Lattice Boltzmann methods (LBM) and finite volume methods, grounded on the discretization of the Navier-Stokes equations, are the predominant techniques for direct pore-scale modeling (Mehmani et al., 2020). In this research, we opted for LBM as the direct simulation method for creating ground-truth flow data, as it has been proven to reconstruct the Navier-Stokes equation accurately. However, other direct simulation techniques could be substituted. LBM directly calculates single and multi-phase flow on pore space images (Rao and Schaefer, 2020). Moreover, the bounce-back condition streamlines the handling of boundary conditions, permitting straightforward representation of intricate geometries on an orthogonal grid without the need for complicated meshing procedures. As a result, LBM has gained traction in the porous media sector, which grapples with intricate 3D pore spaces showcasing features across multiple scales (Mehmani et al., 2020). In the
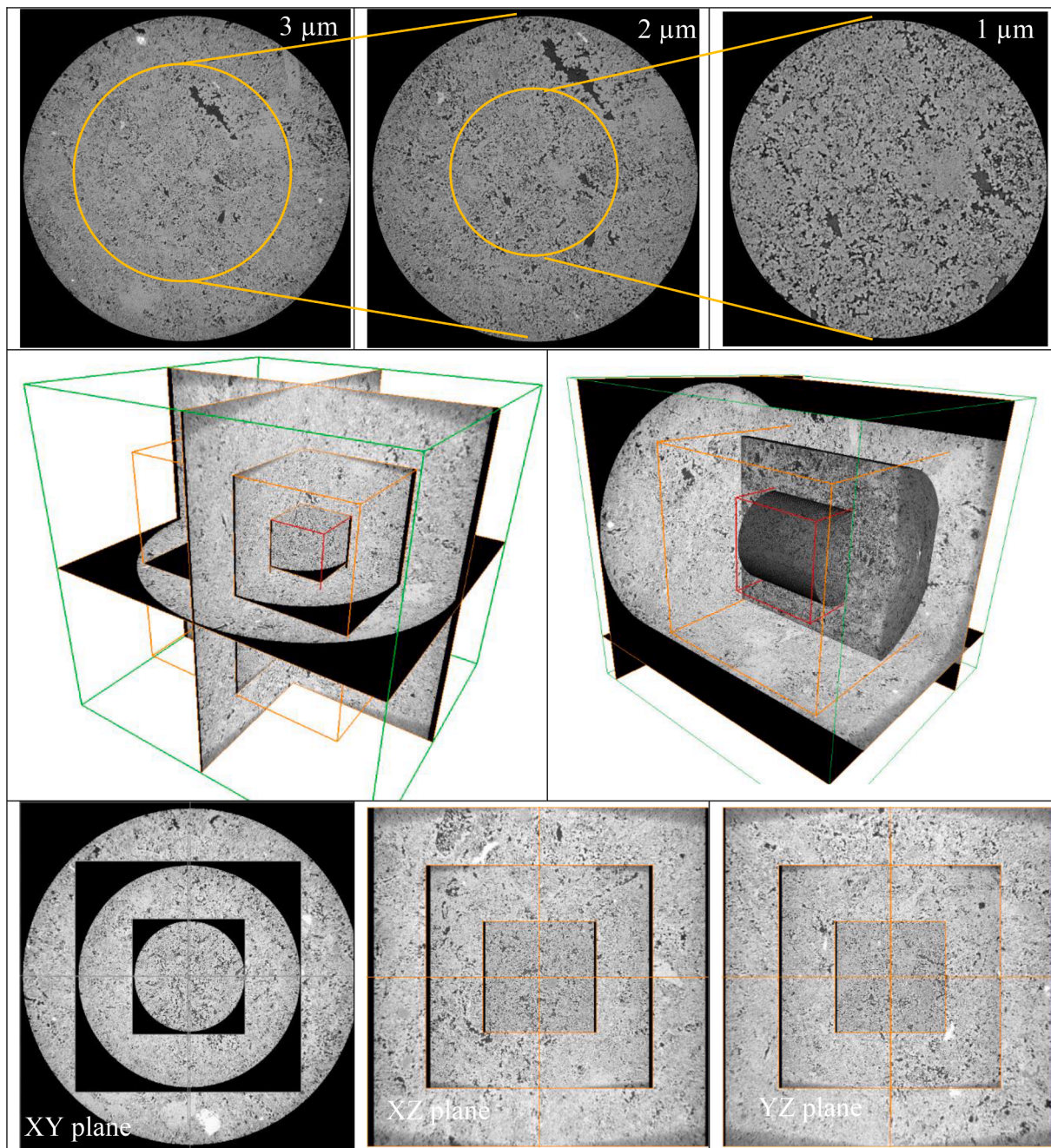
Fig. 4. Examples of 2D slices and 3D visualizations of registered μ-CT images at three resolutions of 3, 2, and 1 μm.

context of simulating fluid flow at the pore scale, the lattice boltzmann method (LBM) is more mathematically rigorous compared to pore network modeling (PNM). Additionally, LBM can provide more reliable evaluations of permeability for porous media with complex geometries, such as carbonates. This is because the accuracy of using pore networks for carbonates has not been established yet (Mehmani et al., 2020). Furthermore, unlike PNM, LBM simulation is executed directly on the voxel domain of digital microstructures without any simplification. In addition, the calculated permeability values within the lattice unit can be directly associated with the pore morphological characteristics in the voxel unit, thereby eliminating the need for further data conversion or processing (Fu et al., 2023).

LBM represents the fluid flow by simulating the propagation of fluid particles on a regular lattice using a time-dependent distribution. Lattice nodes in DRP research are situated at the center of pore voxels extracted from digital rock images, which function as the regular lattice for LBM to simulate fluid flow at the pore scale. In this investigation, the Palabos open-source code is utilized to implement the conventional LBM scheme (Chen et al., 1992), which features the D3Q19 lattice arrangement and the bhatnagar–gross–krook (BGK) collision operator. This model was chosen for its balance between computational efficiency and the ability to accurately capture the essential physics of fluid flow in porous media. The particle distribution function $f_i(x,t)$ denotes the probability of detecting a fluid particle at time t and location x with lattice velocity c_i. Initially, in a state denoted as $f_i(x,t)$, the system evolves locally while progressing from one lattice node to its adjacent nodes during each time step while remaining subject to both mass and momentum conservation. The expression for the evolution of $f_i(x,t)$ in the direction of c_i from time t to $t + \Delta t$ is as follows:
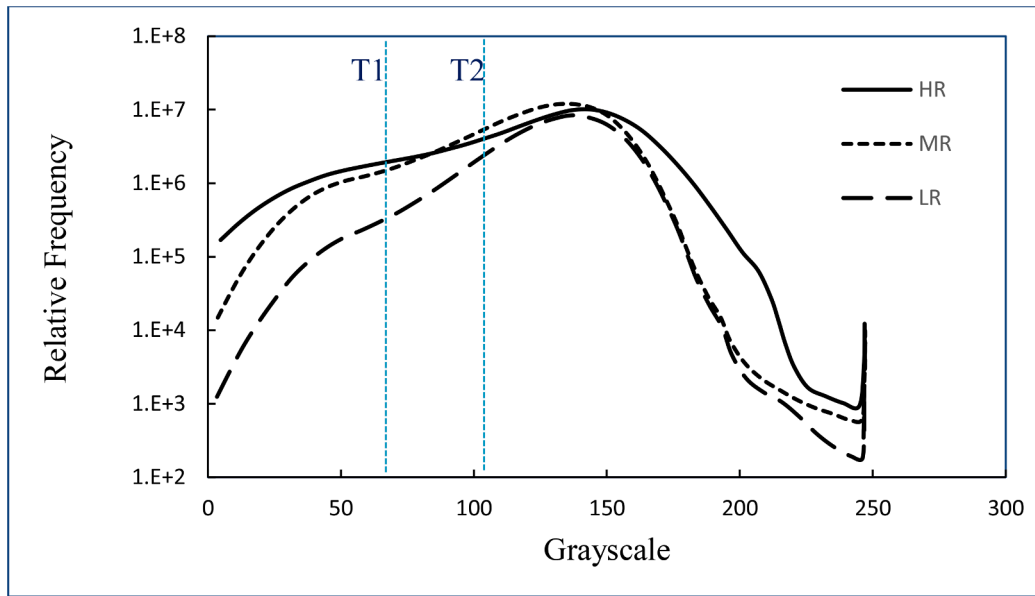
Fig. 5. The gray intensity histograms of the HR, MR and LR images. The area between lower (T1) and upper (T2) thresholds denotes the unresolved micro-porosity fraction.
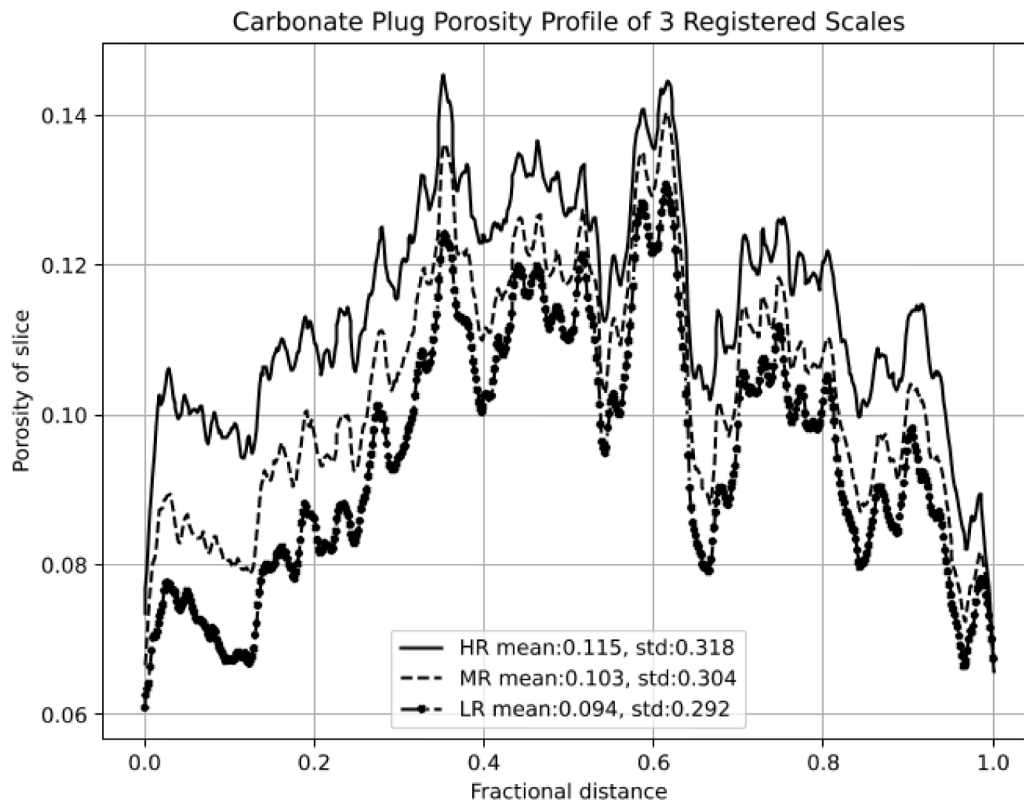

Fig. 6. Porosity profile of the tight carbonate plug for three registered segmented scales in the x-y plane.

$$f_i(x + c_i\Delta t, t + \Delta t) - f_i(x, t) = -\frac{1}{\tau}[f_i(x, t) - f_i^{eq}(x, t)] \quad (1)$$

Where $f_i^{eq}(x, t)$ denotes the equilibrium distribution function, τ is the single-relaxation time, and the subscript i represents the direction of lattice velocity in the vicinity of the lattice node.

The relaxation time τ is a function of the kinematic lattice viscosity ν of the fluid being simulated, i.e., $\nu = c_s^2\Delta t(\tau - 0.5)$. The lattice speed of sound, denoted as c_s is represented by the dimensionless value of $\sqrt{1/3}$.

In order to recover the macroscopic Navier–Stokes equation, the equilibrium distribution function $f_i^{eq}(x, t)$ depicts an ideal state in which the particle distributions approach a particular macroscopic state. $f_i^{eq}(x, t)$ is defined as follows for the D3Q19 lattice arrangement with the BGK collision operator (Chen et al., 1992)

$$f_i^{eq}(x, t) = w_i\rho\left[1 + 3(c_i.u) + \frac{9(c_i.u)^2}{2} - \frac{3(u.u)}{2}\right] \quad (2)$$

where w_i represents the weight factor of the D3Q19 lattice structure, u denotes the macroscopic fluid velocity, and ρ signifies the fluid density. The weight factors w_i for the velocity directions of the central lattice node, face-connected neighbors, and edge-connected neighbors are $\frac{12}{36}$, $\frac{2}{36}$, and $\frac{1}{36}$ respectively, for the D3Q19 lattice model. The macroscopic properties of fluid flow, such as velocity u and density ρ, can be approximated from $f_i(x,t)$ at the end of each time step using the next pair of equations. These macroscopic properties will be utilized in the LBM computation at the subsequent time step.

$$\rho = \sum_{i=1}^{n} f_i(x,t) \tag{3}$$

$$u = \frac{\sum_{i=1}^{n} f_i(x,t)c_i}{\rho} \tag{4}$$

Where n represents the number of lattice directions ($n = 19$ in the present study's D3Q19 lattice structure) (Fu et al., 2023).

By utilizing the representative elementary volume (REV) analysis and the porous phase derived from the segmentation activity, we conducted numerical permeability simulations on the finest scale sub-volumes obtained from the HR image, which were driven by a constant pressure difference between the inlet and outlet faces. All the simulations are set in the laminar steady-state flow regime, where the Reynolds number is considerably below one, aligning with the usual flow dynamics in subsurface structures without fractures or boreholes (Santos et al., 2020).

While LBM boasts high parallelization capabilities, its computational time demands surge with the intricacy of the geometry, notably in carbonates with their tight crystalline makeup. Given recent leaps in computer capabilities, modeling more expansive domains is now achievable. Yet, computational timeframes—even on formidable supercomputer clusters—can be protracted, with considerable computational resources being indispensable. The computational needs of these methods amplify, at minimum, in accordance with the cube of the domain's side length for uniform cubic samples. Therefore, running direct simulations on a representative elementary volume using standard desktop systems often isn't viable. Moreover, complex porous materials like carbonates possess pore size distributions that extend and vary across an extensive scale spectrum. This augments the size of a representative elementary volume and, consequently, the computational duration needed for simulations. Conversely, conducting direct numerical simulations such as LBM on LR images might be impractical, given that segmenting these lower-resolution images typically omits a significant portion of grayscale details relevant to pore structure.

3.3. REV analysis

Connecting pore-scale data to continuum-scale characteristics of porous media depends on defining an REV. Determining the most concise field of view used for a physics-based simulation that accurately represents the entirety of the rock sample for a specific property is pivotal in the DRP framework (Saxena et al., 2017). The REV signifies the smallest volume wherein microscopic fluctuations in the observed attributes of a porous media become negligible or where the assessed property remains scale-agnostic. Setting an REV is a fundamental step for upscaling petrophysical properties across different length scales (Singh et al., 2020). In this study, we scrutinized the shifts in derived permeability relative to sample size, gauged using the LBM for the HR carbonate imagery. Subsequently, we selected the dimensions wherein the deviations in calculated permeability were under 20% (Saxena et al., 2017). Based on this criterion, images spanning 300^3 voxels were designated as the REV for the core images. A figure of fluid flow through a sample REV captured by LBM and the histogram of ground truth permeability variations is depicted in Fig. 7.

It is essential to note that the choice of these voxel numbers for the REV is inherently linked to the voxel sizes (resolutions) and the specific characteristics of the analyzed rock sample. The decision to use the mentioned REV size is based on ensuring that this volume is sufficiently large to capture the heterogeneity and characteristic features of the rock's pore structure. While the voxel number is a primary factor, the voxel size (resolution) also plays a crucial role. Different resolutions provide varying levels of detail, and the appropriateness of the 300^3 voxel volume as the REV is evaluated in the context of the resolution used. Higher resolutions may reveal more detailed pore structures, which could affect determining what constitutes a representative volume. Based on our research question in this study, calculating permeability as our target transport property on the low-resolution images with larger and more representative field-of-view (FOV) through direct LBM is impractical because lots of resolved pores in high-resolution domains that play a significant role in reliable fluid flow simulation would be lost in the lower-resolution images.

Therefore, we ensured that the chosen voxel volume and resolution combination adequately represented the core's heterogeneity while still being manageable for computational analysis. We have chosen the highest image resolution for direct simulation, REV analysis, and, consequently, the generation of reliable reference ground truth for developing the multiscale CNN workflow. As a result, the choice of REV
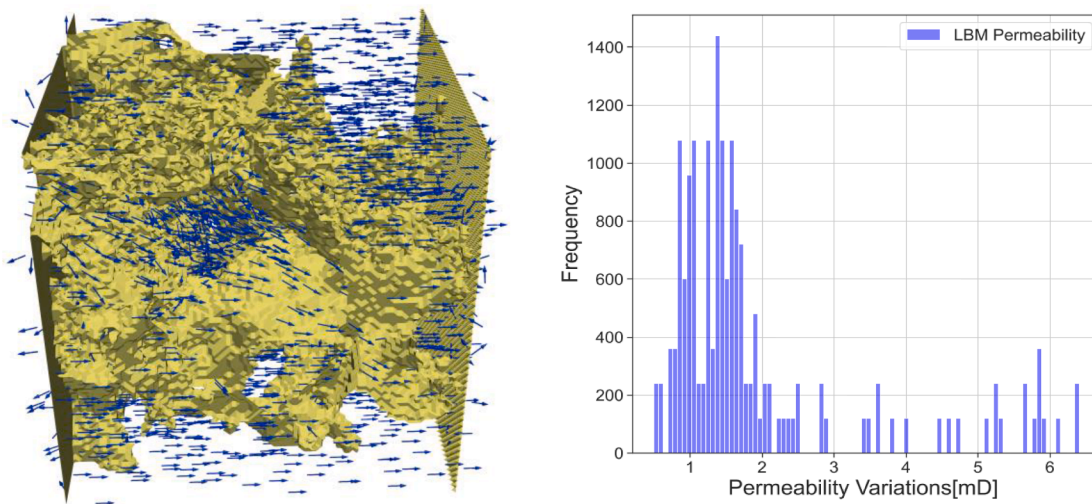

Fig. 7. LBM Velocity Field on an HR REV Subvolume (left) and the histogram of reference permeability variations (right).

size in our study was made in the context of the highest resolution used and, therefore, the calculation accuracy of specific characteristics of the rock samples, ensuring that the selected volume is genuinely representative of the core's properties.

4. The convolutional neural network model

Extensive studies have delved into predicting permeability through convolutional neural network-based deep learning techniques. However, less focus has been directed towards estimating transport attributes in multiscale heterogeneous and genuine porous media. This study emphasizes developing a workflow that harnesses a potent multiscale convolutional neural network (CNN) model in tandem with multiscale, multiresolution 3D visuals of a tight carbonate specimen. Details of the model will be discussed in the following sections. Specifically, we deployed a CNN-type spatially-aware neural network, which acknowledges the spatial configuration inherent in porous media. This network adeptly grasps 3D spatial correlations via a sequence of convolutional procedures (filters). For these filters, an optimal parameter set can mirror the LBM outcome. Additionally, this approach spawns universal and reusable models.

In this study, a multiscale CNN design is employed for sequential image regression, aiming to uncover correlations between pore morphology and a transport characteristic, predominantly permeability. This transition from LBM facilitates the estimation of permeability in the LR images where direct simulation remains infeasible. Our methodology commences by scrutinizing HR volumetric depictions of the carbonate specimen (which encompasses the intricate heterogeneities influencing flow). This step lays the foundation for the primary regression-based CNN model connecting the inputs and outputs. Subsequent training of the model progressively shifts to the synchronized lower-resolution visuals, encapsulating grander scale heterogeneities in the process. To elaborate, the model initially undergoes training with HR input in the registered overlapping region. It then transitions to the encompassing MR images and ultimately undergoes fine-tuning with the LR images, which possess the broadest FOV. This transfer learning paradigm intends to capture the target response fluctuations. Our findings indicate that this strategy aptly models the intricate physics inherent in complex porous substrates like tight carbonates. Given the intricate pore configuration of the carbonate sample in question, we perceived the entire texture within the visuals as a singular unit. The current study does not explore clustering or grouping of heterogeneities. The outlined workflow could potentially encompass more than three imaging levels to span an extended range of scales. Within the sphere of reservoir-scale modeling and simulation, the permeability values extrapolated through this methodology for distinct rock units can be harnessed to gage effective permeability at the core level via Darcy simulation techniques (Botha and Sheppard, 2016).

4.1. Creating dataset

Stable and representative samples bolster spatially aware neural networks, enhancing their capability to identify congruent patterns within data. To foster uniform behavior, the dimension of the 3D sample subset should match or surpass the REV, positively impacting the training efficacy of the convolutional network. Utilizing subset sizes smaller than the REV might induce data non-stationarities, impeding a stable assessment (Santos et al., 2020). While it is feasible to train the convolutional network with 3D images smaller than the REV, primarily due to GPU memory constraints, we prioritize using subvolumes that meet the relatively expansive near-REV size in this work. This is to effectively capture the 3D connectivity intrinsic to a heterogeneous natural porous stone, like a tight carbonate, and its correlational potential with transport properties such as permeability. Contrarily, most preceding studies predominantly trained deep learning models using 2D imagery or 3D synthetic and/or rudimentary real images of smaller

dimensions (e.g., 80, 100, 128, 200, and 256 voxels per direction) (Gärttner et al., 2021; Kamrava et al., 2020; Tembely et al., 2021; Sudakov et al., 2019; Rabbani et al., 2020; Hong and Liu, 2020; Tang et al., 2022; Santos et al., 2020; Santos et al., 2021; Araya-Polo et al., 2020; Takbiri-Borujeni et al., 2020; Srisutthiyakorn*, 2016; Wu et al., 2018; Graczyk and Matyka, 2020; Alqahtani et al., 2018; Marcato et al., 2022; Tembely and AlSumaiti, 2019), often within a singular spatial resolution. Therefore, we have opted for a 300^3 voxel size to extract subvolumes from the primary HR image sized at 600^3 voxels and the synchronized MR and LR imagery with augmented voxel dimensions and FOV, respectively. Henceforth, we will term 3D images of this precise size as "subvolumes." The extraction of these subvolumes employs the sliding window approach, a method commonly used to optimize a given data set beyond its division into non-intersecting subsets. Explicitly, a $300 \times 300 \times 300$ voxel window sweeps each directional axis of the original 600-voxel cube, yielding 59 non-overlapping and overlapping subvolumes in each resolution within the registered domain of the three resolutions (as illustrated in Fig. 8). Moreover, to train the convolutional network using a broader array of sample morphological features, we incorporated downsampled images (Alqahtani et al., 2021) with dimensions ranging from 400^3, 500^3, and 600^3 down to 300^3 voxels for each image resolution. This inclusion does come at the expense of losing certain intrinsic pore structural nuances. Given our use of a supervised learning approach, every training input has an associated label referencing a transport feature inherent to the input data or a pertinent one. To address the image regression dilemma, effective permeability is calculated using LBM for all HR subvolumes to ascertain the requisite ground truth. These LBM computations were executed on a CPU-centric machine equipped with 64 GB of RAM and quad 2.3 GHz Intel Xenon processors. It's crucial to note our exclusive consideration of connected pore spaces. The permeability within an isolated cavity is nullified. Consequently, we integrate an image processing phase to distill all interconnected pores with 26 adjacent spaces (where voxels with a solitary shared vertex are deemed connected) from the binary visuals.

In tandem with binary images, this study employed a 3D normalized Euclidean distance transform (EDT) (Rabbani et al., 2020; Santos et al., 2020; Santos et al., 2021) sourced from all binary subvolumes in each resolution distinctly, within the registered overlapping zone. This transformation marks every pore voxel based on its proximity to the closest solid boundary. Our motivation for including such supplementary inputs springs from the objective of forging a model endowed with capabilities to function across varied scales and the multifaceted geometries typical of the carbonate sample.

By supplying these supplementary descriptor factors, we boost the model's training efficacy. Consequently, the resultant model exhibits adequate generalization, empowering it to compute permeability within a permissible error bandwidth even when applied to a test dataset embodying myriad unfamiliar geometries. Essentially, the model discerns and adapts to unique patterns, thereby orchestrating a sturdy function linking the visual imagery to permeability. The preference for the normalized Euclidean distance transform (EDT) stems from its simplicity, cost-effectiveness, and potential to offer more insight than a mere binary image at the outset. This specific transformation offers a concise visual representation of the space viable for fluid flow and the relative distance to the adjacent solid (impermeable) boundary (Santos et al., 2020), and notably, it exhibits a potent correlation with transport attributes.

Furthermore, underpinning our study's hypothesis, as highlighted earlier, is the notion that LR images might not encapsulate the entirety of the pore architecture. Hence, we postulate that these discrete features, being digitally synchronized with the MR and HR images, intrinsically embody data indicative of pivotal fluid conduits. In essence, this EDT can potentially pinpoint and measure these conduits, buttressing our primary goal of approximating permeability within LR images that possess an expansive field of view (FOV). In line with this methodology, we curated 59 standardized EDT 3D datasets for each
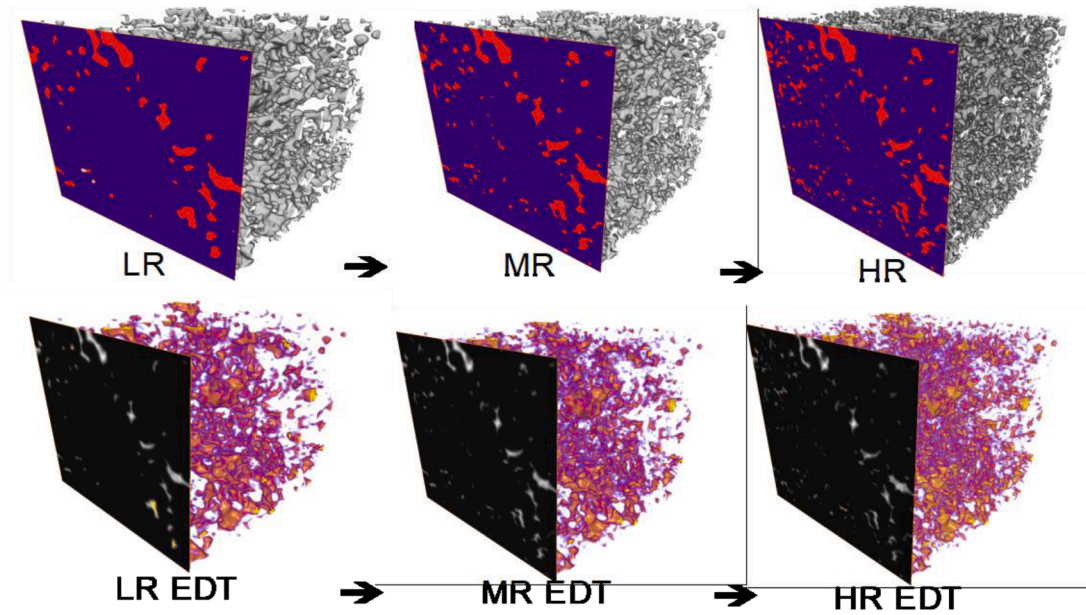
Fig. 8. 2D and 3D representations of the binary (upper row) and normalized EDT (lower row) subvolumes derived from registered images of HR, MR, and LR from left to right, which serve as input data for the multiscale CNN model. Regarding the HR image, the corresponding MR and LR volumes exhibit a successive increase in voxel size and loss of resolution of smaller pores.

resolution (refer to Fig. 8). Cumulatively, this results in a compilation of 354 images (comprising 177 binary and 177 normalized EDT visuals) spread across all resolutions (with each resolution containing 118 images), serving the exclusive purpose of training our bespoke CNN model.

4.2. Network

In the upcoming sections, we provide a brief overview of how convolutional networks operate, leading up to the introduction of our innovative model, the multiscale CNN. This model amalgamates single-scale CNNs that function sequentially to yield predictions for a specified sample. Subsequent discussions shed light on our data augmentation techniques aimed at enhancing the accuracy and universality of results. This is followed by an exposition on the transfer learning framework seamlessly integrating these CNNs.

4.2.1. Convolutional neural network (CNN)

Past studies have underscored the proficiency of convolutional neural networks (CNNs) in dissecting digital rock imagery for tasks such as classification, segmentation, resolution amplification, reconstruction, and regression. These models have consistently exhibited prowess in deciphering intricate relationships within voluminous data sets. Unlike traditional matrix multiplication found in fully connected feed-forward networks, CNNs harness discrete convolution to distill local spatial associations (sparse interactions) across the domain. Using filters that are significantly smaller than the input image, CNNs adeptly glean pertinent domain-specific details. This cascading structure of convolutional layers enables the network to capture attributes with ever-expanding receptive fields at subsequent abstraction tiers. By leveraging this layout in conjunction with the backpropagation technique, a network can be calibrated to discern intricate, nonlinear interplay between inputs and associated outputs (Santos et al., 2020). The CNN model showcased in this study is a tailor-made regression model that boasts a straightforward
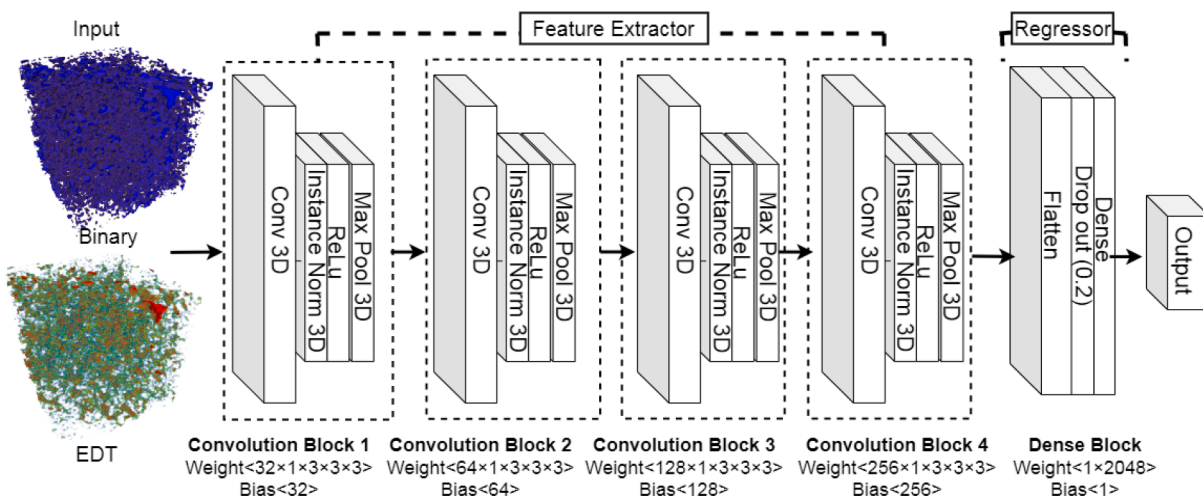
Fig. 9. A diagram of a final CNN design with a single dense or fully-connected layer and four convolution blocks for the multiscale regression problem. The inputs consist of 3D carbonate images (EDT and binary) with varying resolutions-scales, and the output is the computed permeability. The dimensions of weights and bias terms in each convolution layer depicted below each block.

design consistent across all scales. This design was honed through iterative experimentation with diverse configurations, with model efficacy gauged by monitoring accuracy metrics and the behavior of the loss function.

The network structure is shown in Fig. 9. The final architecture for feature extraction comprises four convolution blocks, each containing a 3D convolution layer, an Instance normalization layer, a nonlinear activation function ReLU, and a pooling layer. The end block of the network is composed of the layers flatten, drop out, and dense, in that order. The brief introduction of the mentioned layers is summarized below. For further reading, a detailed explanation of all the main components of a CNN can be found in (Kamrava et al., 2020; Kamrava et al., 2021):

- 3D convolution layer (Da Wang et al., 2021; Santos et al., 2021): This layer comprises 3^3 kernels (or filters) for the aforementioned four convolutional layers, and the stride is $1 \times 1 \times 1$ for all kernels that are slid along the input to generate feature maps via the convolution process (Eq. (5)):

$$Y = \sum_{i=1}^{n_f} X * K_i + b_i \tag{5}$$

Y is the output, n_f is the number of filters, X is an input array of shape (n_b, n_x, n_y, n_z) where n_b is the number of batches, and (n_x, n_y, n_z) is the number of voxels in each direction of the coordinate system. K represents a collection of convolutional filters of shape (n_f, k_x, k_y, k_z) where k refers to the size of each filter (kernel size), $*$ is the convolution operation, and b is the bias term.

- Instance normalization 3D (Ulyanov et al., 2016): This layer normalizes its inputs to have a zero mean and one standard deviation (Eq. (6)). This makes it easier to train a model using samples with significant variations. This is performed on each sample using the sample's unique statistics and without trainable parameters (Santos et al., 2021).

$$x_{out} = \frac{x_{in} - \overline{x}}{\sqrt{\sigma^2 + \varepsilon}} \tag{6}$$

$\overline{x}$, σ are the sample mean and its standard deviation, respectively, and ε is a small constant to prevent divisions per zero.

- Nonlinear activation function: This layer contributes to forming nonlinear relationships (such as the one between pore geometry and permeability). In this work, we employ a Rectified Linear Unit (ReLU) activation function. This layer transforms any data that passes through using the following equation (Eq. (7)):

$$Y_{rectified} = \max(0, Y) - \alpha \max(0, -Y) \tag{7}$$

where α equals 0 for (ReLU).

- Maximum pooling 3D (Sit et al., 2020; Rabbani et al., 2020): This technique transfers the relevant learned characteristics to the subsequent layers while transforming data into a smaller collection without losing much information and sacrificing positional knowledge. Maximum pooling computes the highest possible value for each feature map patch. In our design network (Fig. 9), the max pool 3D layer in the first and second convolution blocks has a kernel size

of 2 and stride of 2, whereas, in the third and fourth convolution blocks, The kernel size and stride are both 3.

- Dense block (Kamrava et al., 2020; Da Wang et al., 2021): Finally, the subsequent dense neural layer interprets the extracted features. These techniques are used to provide input to the fully connected layer once data from a CNN block has been "flattened" from its multidimensional shape into 1D vectors. The dense or fully-connected layer is added to the end of a CNN to produce regression results and retains the most significant outputs from all the convolutional layers. To prevent overfitting issues, this block employs a drop-out filter that randomly eliminates a certain percentage of neurons (20% in this network).

4.3. Data augmentation

Based on the preceding section, the primary dataset of the tight carbonate image consists of 59 3D micro-tomography image subvolumes, both in binary and normalized EDT formats, for each scale. This amounts to a total of 354 images across three resolutions within the registered region. Given the scarcity of varied and authentic tomography data across diverse resolutions and the pressing need for 3D CNNs to be trained on substantial image datasets, we turned to morphological filters. These filters enable the manipulation of existing data to craft novel representations by tapping into an extensive suite of shape-centric image processing capabilities.

Morphological operations determine the value of each pixel in the output image based on a comparative analysis of its counterpart in the input image and its surrounding pixels. Among these operations, dilation, and erosion take precedence. Dilation can be visualized as placing an imaginary sphere within the non-reference phase and gaging its extent, assuming the reference phase is impermeable to this sphere. In essence, spaces inaccessible to the sphere translate to an expanded reference phase. On the other hand, erosion is the reverse operation, wherein the reference phase is permeable to this virtual sphere. In the digital realm, dilation contributes pixels to the boundaries of objects, while erosion detracts pixels from these peripheries.

Frequently, dilation and erosion operate in tandem in a process termed "opening." This dual operation notably elevates the quality of reconstructed imagery, proving invaluable for intricate microstructures where capturing essential structural nuances via limited statistical correlations is challenging (Mandzhieva, 2017). In our research, we harnessed the morphological filters of dilation, erosion, and opening on both binary and normalized EDT datasets, ensuring consistency in image dimensions across scales. Fig. 10 provides a pictorial representation of the filters' implementation on a sample. These filters augment training by spanning a wide spectrum of probable pore distributions, enriching the plausible variations in pore structures and morphologies. Consequently, the CNN model is exposed to diverse pore space manifestations with differentiated morphologies (Kamrava et al., 2020).

Moreover, we employed the said morphological filters to devise pore structures that, while similar in form and heterogeneity, presented nuanced differences in porosity (Santos et al., 2022). The realizations emanating from these filters mirror the porosity distribution inherent in the original 3D imagery. As a result, the dataset culminated in 472 training images for each resolution, partitioned into 118 original images and 354 crafted realizations. In aggregate, this encompasses 1416 training 3D images – evenly split between 708 binaries and 708 normalized EDT subvolumes – spread across the three resolutions, all of which are meticulously aligned and bear dimensions of 300^3 voxels.

4.4. Transfer learning

While classic deep learning algorithms have their merits, they also come with challenges. Neural networks tend to excel with larger datasets. With smaller datasets, the risk of overfitting increases, potentially leading to inaccurate test set predictions (Tang et al., 2022). Given the
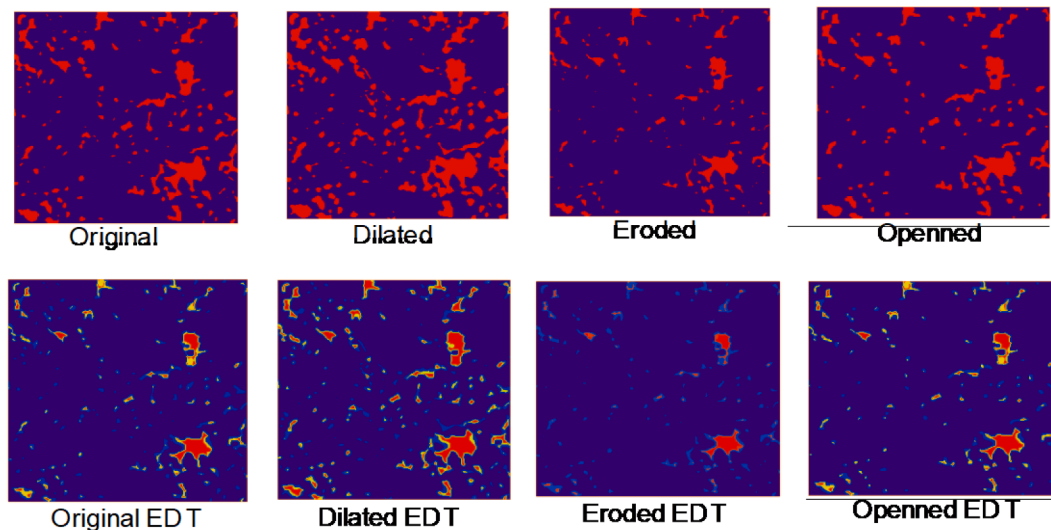
Fig. 10. Visualizations of the 2D cross-sections of the 3D binary and normalized EDT subvolumes generated using morphological filters. All of the resulting images are digitally aligned.

cost of acquiring data via numerical simulations and experiments, the effectiveness of our proposed method for a small dataset in a multiscale approach is of interest. Transfer learning offers a way to repurpose existing deep learning models for new tasks with scarce data. Unlike traditional algorithms focusing on specific tasks, transfer learning leverages knowledge from one task to benefit another.

CNNs have proven their efficacy in learning image characteristics and generalizing image features from vast training sets. Transfer learning has the potential to produce outcomes that are comparable to, or potentially superior to, initiating a CNN model with random parameters. This approach involves adapting an already trained neural network to a new dataset, capitalizing on previously learned features (Anzum, 2021). Especially when computational resources are constrained, transfer learning offers an advantage, as training extensive models on powerful GPUs is time-consuming. It allows leveraging pre-trained weights as an initial point (Weiss et al., 2016). Given the limited image dataset for the multiscale regression model, our first step is to use transfer learning. We then fine-tune the model using various transformation techniques to enhance its performance. This fine-tuning not only retrains the model but also extends backpropagation to adjust the pre-trained network's weights. While it's possible to fine-tune the entire network, we might retain some early layers and only modify higher-level components to mitigate overfitting. For this study, each scale uses a distinct dataset to adjust all layers of the previously-trained CNN.

Different transfer learning techniques exist, and our choice depends on the specifics of the new dataset relative to the dataset of the pre-trained models. We aim to combine scales to predict transport properties, integrating transfer learning within the 3D CNN framework, as shown in Fig. 11. To build the multiscale model, we curate datasets for each scale, subdividing them into binary and normalized EDT 3D subvolumes. The first step involves pre-training the model using three-scale datasets of normalized EDT maps. The model is designed to recognize the physical correlation between the normalized EDT and the velocity field and, subsequently, the permeability of the porous media, which is inherently nonlinear.

At each scale, we commence with the weights from the previously pre-trained network and then train the system to update these weights. As a result, the network starts to grasp the intricate relationship between the diverse pore structures and the target transport property, i.e., permeability, within the designated region. This foundational understanding is then built upon through sequential transfer learning, employing the binary datasets for each scale and refining them with the LR images. For permeability estimation, only the binary LR data from the entire core sample will be necessary, streamlining the evaluation process with the refined model. It is crucial to note that high-resolution (HR) data plays a pivotal role in deciphering the nuanced 3D feature distribution's non-linear correlation with permeability at its respective scale. This vital information is relayed to the networks functioning at lower resolutions within the registered zone. Here, the less detailed images offer insights into the broader scale variances that influence fluid movement, as depicted in Fig. 11.

We achieved a robust and generalizable model for the corresponsive heterogeneity of this carbonate sample through the suggested data
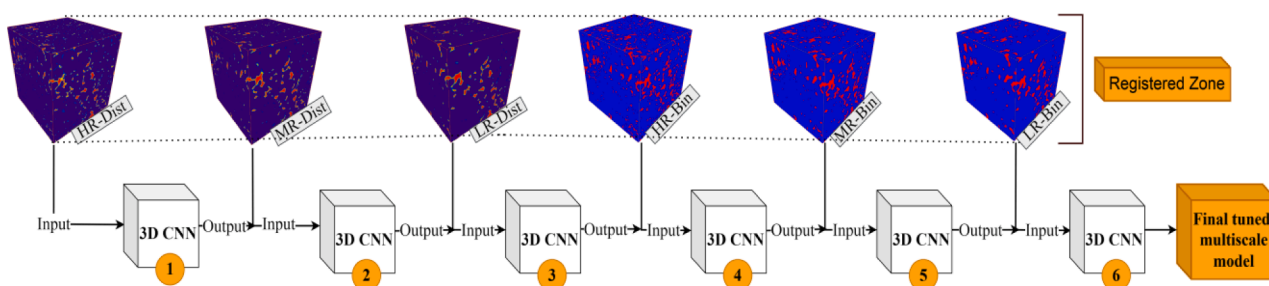
Fig. 11. Workflow for training a multiscale CNN. Initially, the model is pre-trained with the 3D normalized EDT dataset obtained progressively from the HR, MR, and LR carbonate images, respectively, in the registered overlap region that can provide information about the original binary map for initial optimization of the model parameters. Each scale learns the association between pore space morphology and permeability response at the specific image resolution. The primary image regression is accomplished by training binary images in the same order. The final calibrated model calculates the permeability of LR image volumes.

augmentation and transfer learning processes, despite the fact that the model is calibrated in the registered region of the HR images and limited data for training the CNN were available for each scale. Furthermore, the relatively large image volumes for inputting the CNNs (300^3 voxels) and memory restriction of the computational GPU necessitate that we accomplish the desired result within an optimal and simplistic network architecture. In other words, this method is rigorously optimized inside a constrained frame of limited data at each scale underneath a GPU with limited memory.

4.5. Model training

The multiscale CNN model was developed and trained using Python in the PyTorch framework. We also utilized open-source Python packages like Pandas, tifffile, SciPy, Numpy, Seaborn (Waskom, 2021), Plotly, and Matplotlib (Hunter, 2007) for pre- and post-processing tasks. The backpropagation Adam algorithm (Kingma and Ba, 2014) with a learning rate of 1e-4 was used for backpropagation during CNN training, aiming to minimize the root mean squared error (RMSE) prediction loss. In the realm of supervised learning, RMSE (as given in Eq. (8)) is a widely used metric to gage model performance and ascertain prediction accuracy. Leveraging the Euclidean distance of the residuals quantifies the deviation of predictions from actual values. Furthermore, our dataset was divided as 90% for training and 10% for validation. Larger, unregistered MR and LR images served as an independent test set. We trained our model over 60 epochs per scale, using a batch size of one. The input data is provided in tiff format, and Table 2 lists the model's hyper-parameters. The training was executed on a single Nvidia GeForce RTX 3060 GPU boasting 12 GB of memory, complemented by 48 GB of RAM.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \| y(i) - \widehat{y}(i)\|^2}{N}} \qquad (8)$$

Where N represents the number of samples, $y(i)$ is the ith measurement, and $\widehat{y}(i)$ is the corresponding prediction.

5. Results and discussion

The primary goal of this study is to establish a swift and precise regression connection between the intricate geometry of a tight multiscale carbonate porous medium and its permeability variations using the introduced multiscale CNN model. The workflow for predicting permeability encompasses a sequential CNN calibration anchored to the registered region across three resolutions.

The coefficient of determination, R^2, serves as an index of the ability of the microstructural predictor features to explain variability in the response variable. R^2 (given by Eq. (9)) is the disparity between 1 and the quotient of the total sum of squares (SStot) and the residual sum of squares (SSres):

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \qquad (9)$$

$$SS_{res} = \sum_{i=1}^{N} \left(k_i^{Model} - k_i^{LBM} \right) \qquad (10)$$

$$SS_{tot} = \sum_{i=1}^{N} \left(k_i^{LBM} - \overline{k_i^{LBM}} \right) \qquad (11)$$

where N is the number of samples and k_i the permeability of each sample. We employ R^2 to evaluate the predictive accuracy of 3D sub-volumes of pore structure images and their associated contribution to estimating permeability variations in multiscale CNN models. Our methodology is adaptable across all three scales and particularly tailored to map permeability in LR volumes. To refine the CNN architecture, we study the influence of various hyper-parameters on prediction performance. These include the number of convolution blocks, learning rate, momentum, and weight decay, whose significance in impacting CNN's predictive capability has been underscored in prior research (Tang et al., 2022). The optimal model was selected based on criteria such as the lowest RMSE loss and the highest R^2 value. The definitive model structure is showcased in Fig. 9, and the finalized hyper-parameters are detailed in Table 3.

Apart from the data generation method delineated in Section 4.3, we also explored data augmentation techniques, like introducing random horizontal and vertical flips to the images. This augmentation led to discernible improvements in both loss values and R-squared scores. The outcomes from the multiscale modeling and characterization process are expounded upon in this section and are divided into two parts. The first part delves into the results of the calibration model within the registered areas across the three image resolutions. The second part focuses on assessing the model's performance on the test data, then scaling up the findings from the test phase and validating the overarching model.

5.1. Calibration model

The model's training and subsequent evaluation begin with analyzing the learning curves from its final iteration. For each training epoch, the model's validation and training losses are measured in terms of the root mean square error to ensure no overfitting and verify that the training has attained its optimal stride. The sequential reduction in both training and validation losses is detailed in Table 3. The decreasing trend of the loss function's mean error and standard deviation from the HR to LR data replicates that while the network continued to learn and adapt its parameters, the pre-learned features from the HR dataset provided a foundation that required fewer drastic adjustments. This could be interpreted as the network sequentially assimilating knowledge from datasets of varying resolutions by adopting efficient transfer learning. Extending training beyond 60 epochs scarcely reduced the validation loss below 0.25. Hence, training was stopped at 60 epochs.

This section presents the calibration model results to showcase the multiscale CNN model's capacity to interpret the geometric heterogeneities at different resolutions. Optimal hyper-parameters were selected through training, and the model's accuracy was subsequently validated using test data. Our analysis yields statistical insights by examining the model's predictions and loss values. The R^2 metric serves as the chosen indicator of model precision. The first evaluation phase yielded permeability estimates with actual sample data inputs for each resolution within the overlapping zone of the three aligned images.

Fig. 12 displays the permeability values derived from flow simulations in HR images against those predicted from HR, MR, and LR images within the registered area using our multiscale regression approach. Each scatterplot includes a regression line, a robust method for showcasing the linear relationship between two variables. The plots are annotated with the R^2 value for the test data permeability comparison, registering an R^2 of approximately 0.96 across all three resolutions. Fig. 12(d) also directly compares voxel permeability predictions for LR

Table 2
The hyper-parameters of the regulated multiscale CNN model.

Hyper-parameters	Value
Batch size	1
Number of epochs for training on each resolution	60
Momentum	0.8
Weight decay	5e-3
Number of initial feature maps	32
Learning rate	1e-4

Table 3
The statistics of RMSE for training and validation data.

| Row	Dataset	RMSE for training data				RMSE for validation data			
		Min	Max	Mean	STD	Min	Max	Mean	STD
1	HR-EDT	0.6046	1.3002	0.6957	0.1056	0.4714	1.1356	0.6557	0.0905
2	MR-EDT	0.5281	1.0173	0.6281	0.0766	0.3302	1.0872	0.4524	0.1152
3	LR-EDT	0.4785	0.8525	0.5603	0.0606	0.3211	0.8199	0.4512	0.0843
4	HR-Binary	0.4521	0.7618	0.5351	0.0532	0.2662	0.6623	0.3692	0.0675
5	MR-Binary	0.4481	0.6838	0.5374	0.0451	0.2749	0.5398	0.3741	0.0501
6	LR-Binary	0.4263	0.6891	0.4887	0.0471	0.2541	0.6681	0.3689	0.0704
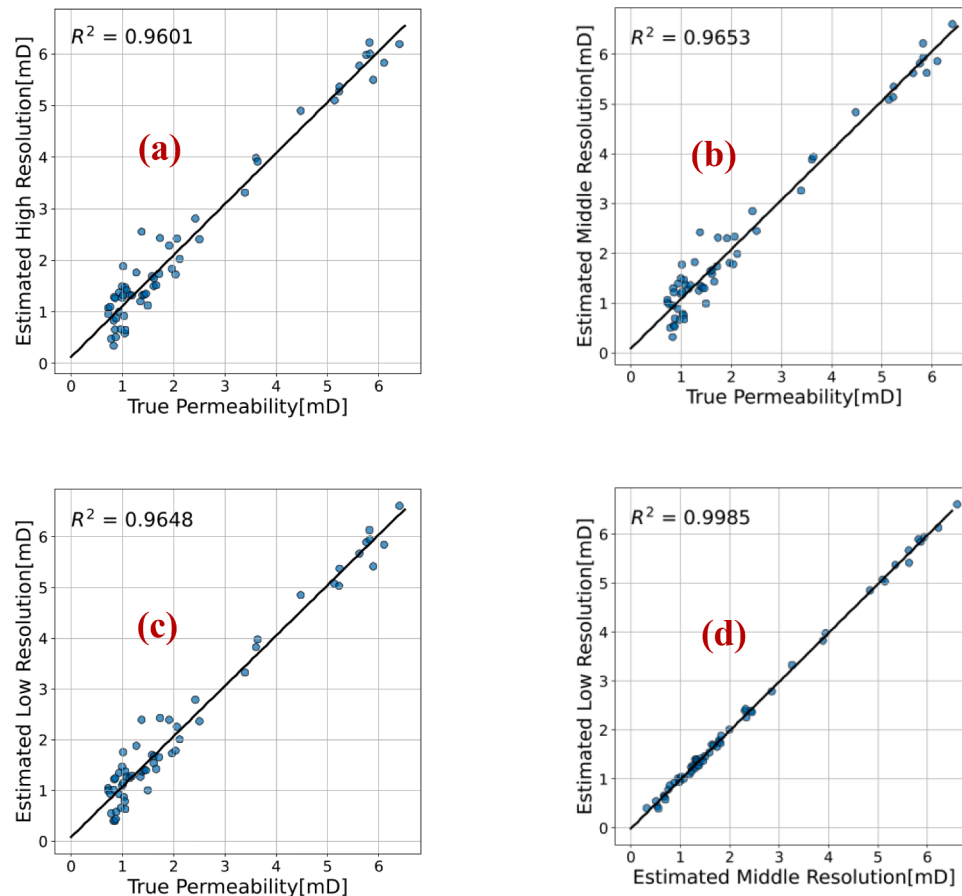
Fig. 12. Permeability estimation performance for the corresponding real samples in HR, MR, and LR subvolumes of the registered area and their calculated R^2. A solid black 1:1 line shows the perfect correspondence of the LR/MR predictions with the HR data.

and MR images with an R^2 of 0.99, affirming the close similarity between the computed values. These plots not only illustrate the model's accuracy but also its generalizability across various resolutions of input data, which is a cornerstone of robust machine learning models. It is important to point out that these results have been underpinned by the carefully registered images in three resolutions. Given the variability in pore structures and permeability estimates within the registered region, the precision achieved is noteworthy. It reflects well on the model's ability to accurately correlate the predicted with the actual permeabilities for this carbonate study. The consistency of the approach is further evidenced by the minimal number of outliers in these plots. Both the plots and Fig. 13—which includes a histogram and box plot comparison of true vs. predicted permeabilities—indicate that errors are marginally higher for lower permeability values, which occur more frequently. This could be attributed to the complex pore structure in denser samples, which presents a more significant challenge for the model in capturing and generalizing the intricate local spatial interactions related to permeability (Dehghan Khalili, 2013).

Furthermore, direct comparisons in Fig. 12(a,b, and c) show a high correlation with the HR target values at higher permeabilities, while at permeabilities under one millidarcy (mD), the correlation diminishes. This suggests differences in connectivity within the micro-structure of tighter areas that are not as precisely captured by lower-resolution images. A consistent offset at higher permeabilities suggests that, although porosity is well-aligned across scales, the model does not capture certain fine details below a specific resolution threshold. This consistent pattern of missing features, as shown in each image scale representing a permeability REV, results in an almost constant disparity between predictions and actual targets, mirroring observations from earlier studies such as those by Jackson et al. (2021).

Another interesting observation regarding comparing the R^2 values in Fig. 12 is the slightly lower R^2 for the estimated high-resolution case compared to the middle and low-resolutions. This result is related to the effective transfer learning scheme implemented during the model development phase. Our model was trained using multiscale data, aiming to generalize across different resolutions. To achieve the goal in
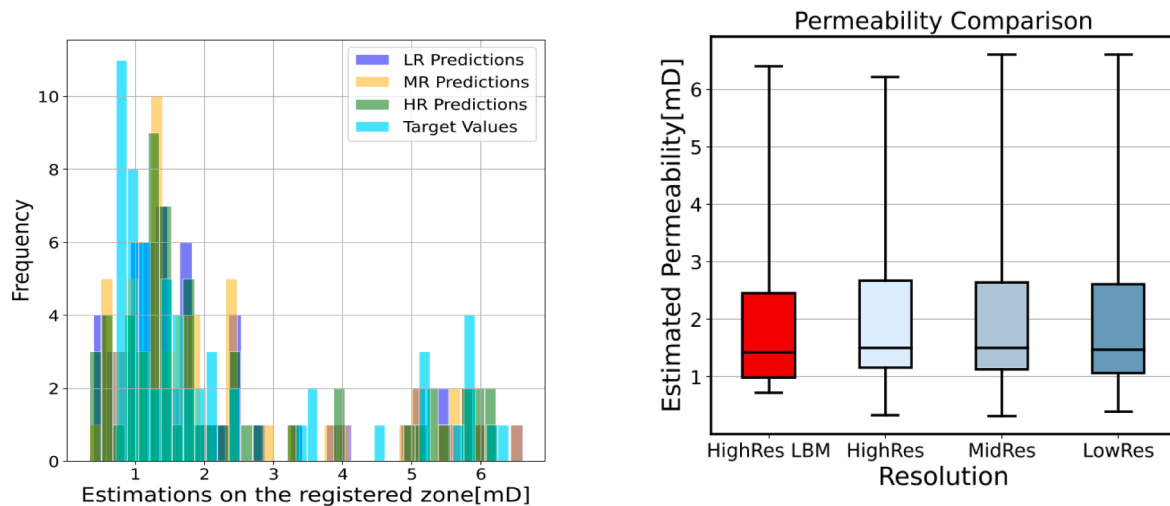
Fig. 13. A histogram (left) and box plot (right) of permeability predictions in the registered calibration area of the three resolution images and the target values obtained from the LBM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

this work in order to computationally efficient estimation of permeability on the low-resolution domains with larger field-of-view (which is infeasible by direct numerical simulation) and with high accuracy and generalizability but less computational time compared to LBM, the final model sequentially fine-tuned on the middle-resolution and low-resolution domains respectively after pre-training on the high-resolution images. Secondly, the high-resolution datasets inherently contain more detailed and complex pore structures than the registered middle and low resolutions. This complexity includes finer resolved features and subtle variations in pore geometry, which are not present in the lower-resolution images and are more challenging to capture and accurately predict. As a result, the model's performance, as measured by R^2, may appear slightly lower for high-resolution estimations due to the increased difficulty in capturing these intricate details. Thirdly, this slight increase in R^2 from HR to MR and LR data, also implies that the model, when trained sequentially, can potentially avoid overfitting to the HR details that are not present in lower resolutions. Finally, despite the marginally lower R^2 in high-resolution predictions, it is essential to emphasize that the values obtained are still indicative of robust model performance, particularly considering the complexity of the multiscale data. The results still validate the effectiveness of our approach in multiscale modeling, since the final trained model is a multiscale-resolution-aware network and can predict the permeability of multi-resolution images with R² more than 0.96.

Furthermore, The histogram and box plot of Fig. 13 overlay the frequency distributions of predicted permeability values for each resolution (LR, MR, HR) and the ground truth. These distributions indicate the most and least common permeability values predicted by the model and demonstrate that the model's estimations at different resolutions capture the overall trend of the ground truth data. As a result, the multiscale CNN suggests a well-tuned model that encompasses the essential features necessary for accurate permeability prediction.

Table 4 details the average predicted and actual permeabilities and their average relative error. The results show a low average relative error of just 4.5% for the HR data predictions, with a modest stepwise decrease in error observed as the network progresses through scales—down to 3% for LR data predictions. This improvement in the network's predicting ability is attributed to the transfer learning process by further fine-tuning or incorporating additional features that compensate for the loss of resolution.

The data also reveal a marginal increase in the standard deviation for permeability predictions at the MR and LR scales, reflecting the wider permeability range encountered at these resolutions. Moreover, Table 4 includes the Dykstra-Parson coefficient (V_k), which is the quotient of

Table 4
The performance of the model in terms of averaged loss values, relative errors, standard deviations, and the Dykstra-Parson coefficient of the actual and estimated values.

Transport property	Mean [mD]	Relative error (%)* of Mean	STD	Dykstra-Parson coefficient, V_K [-]
K-LBM	2.2067	0	1.7551	0.7953
K-pred-HR	2.3070	4.5452	1.7652	0.7651
K-pred-MR	2.2890	3.7295	1.7740	0.7750
K-pred-LR	2.2740	3.0498	1.7719	0.7792

* Relative error $= |\frac{K_{LBM} - K_{Model}}{K_{LBM}}| \times 100$.

the standard deviation to the mean effective permeability (Jackson et al., 2021). Notably, LR predictions exhibit a marginally higher V_k value, suggesting a more comprehensive capture of the heterogeneity inherent in the plug images, resulting in a broader range of permeability variations at a larger scale. Conversely, HR predictions within the overlap region typically yield a lower V_k due to the more discernible pore-space geometries and their higher connectivity, which equates to less permeability variation among the REVs.

Overall, the findings of this subsection, in conjunction with the RMSE loss presented in Table 4, attest to the efficacy of a transfer learning workflow augmented by image registration and data generation techniques using normalized EDT and binary maps. Such a strategy significantly bolsters the CNN model's predictive power and generalization capabilities across various image resolutions within the registered area.

Fig. 14 plots the modeled response values against the residuals, which are the discrepancies between predicted and actual permeabilities at each resolution, to showcase the model's robustness and effectiveness. The figure suggests that residuals' variance tends to decrease as the magnitude of the response variable increases, indicating that the model exhibits homoscedasticity, which is a characteristic of a high-quality model. Nonetheless, while the coefficient of determination is a useful indicator of precision, it cannot be the sole judge of model quality. Ideally, model residuals should center around zero without obvious clustering or patterns at any scale, as such patterns could suggest systematic bias within the model (Botha and Sheppard, 2016). A visual examination of the residuals presented in Fig. 14 reveals a relatively scattered distribution, reinforcing the conclusion that our model operates without bias and is suitable for the data it was designed to interpret.
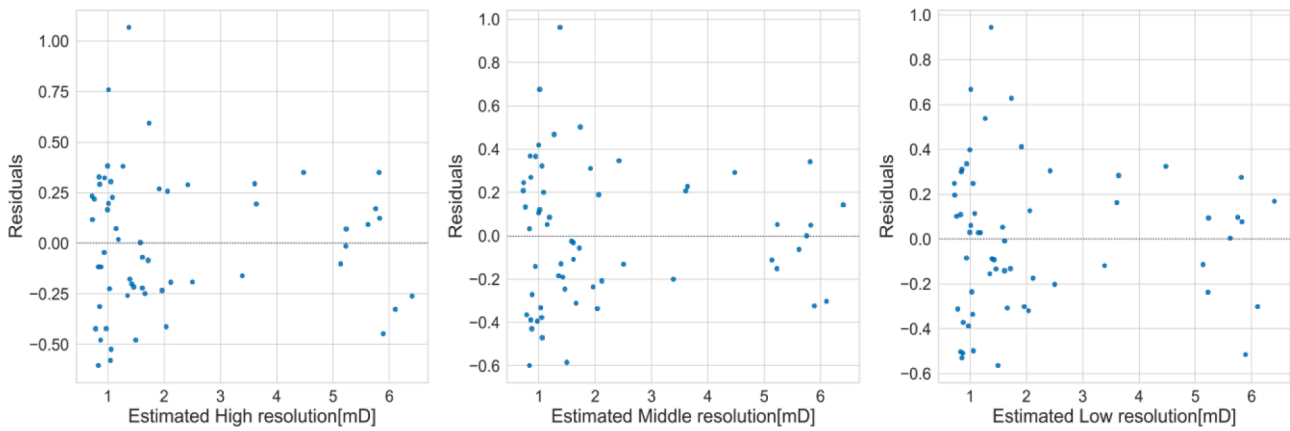
Fig. 14. Multiscale CNN model calibration over the field of view of the image with the highest resolution. Plotting the residuals of a multiscale CNN model against the predicted permeability values at three resolutions reveals no apparent systematic bias.

5.2. Comparison with traditional (Vanilla) CNN

To validate the effectiveness of our multiscale CNN approach, a comparative analysis was conducted against a traditional (or vanilla) CNN architecture. The vanilla CNN was trained and tested on the same LR dataset and processed just the LR images independently without the benefit of multiscale integration or sequential learning. The resulting regression plot of this model is depicted in Fig. 15.

As can be seen, the multiscale CNN demonstrated superior performance in terms of prediction accuracy, with higher R-squared values (~0.96) (Fig. 12) compared to a traditional CNN ($R^2 \approx 0.83$). It means the vanilla CNN struggled to capture the multiscale heterogeneity inherent in tight carbonate rocks, resulting in less accurate permeability predictions, especially at lower resolutions, which is our target. In contrast, with its integrated approach, the multiscale CNN showed a more robust understanding of these complexities. In traditional CNNs training, the CNN is trained by each resolution dataset without integration, which is insufficient to reach the desired accuracy and generalization.

Therefore, here we clearly highlight the novelty and distinctions of the proposed multiscale CNN architecture:

(1). Integration of multiscale approach: Our multiscale 3D CNN uniquely combines 3D information across high, middle, and low
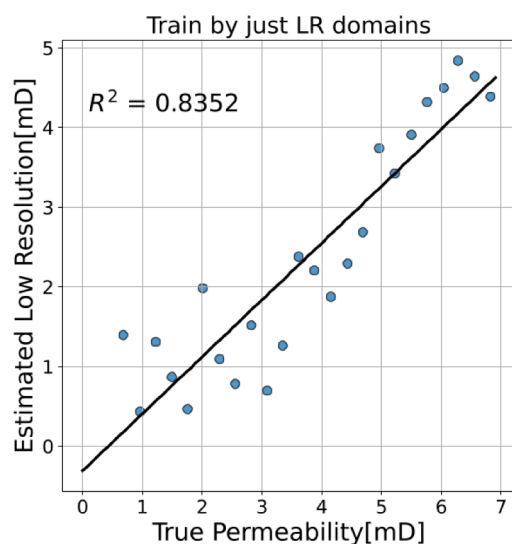

Fig. 15. Permeability estimation performance for the corresponding real samples in just LR subvolumes of the registered area and their calculated R².

resolutions. This integration addresses the generalization issue by training the model on a variety of scales, which is crucial for capturing the comprehensive heterogeneity of tight carbonate rocks, where pore structures significantly vary across scales. This is an advancement over traditional CNNs implemented in previous studies (Kamrava et al., 2020; Rabbani et al., 2020; Jiang et al., 2022; Santos et al., 2021; Takbiri-Borujeni et al., 2020; Marcato et al., 2022), which typically process single-scale images and may overlook critical multiscale features relevant to permeability estimation.

(2). Sequential learning across scales: Unlike a vanilla CNN, which learns features from a single resolution (such as these recent studies (Elmorsy et al., 2022; Liu et al., 2023)), we employ a sequential approach in our CNN, where learning is distributed across multiple stages, each focused on a different resolution scale. This method is computationally more efficient than attempting to process all scales in a single, monolithic model. It starts by learning fine-scale details in high-resolution images and then progressively adapts to coarser details in lower resolutions. This sequential learning strategy, employed by transfer learning, enables the model to leverage learned features from high-resolution images and efficiently apply them to lower-resolution images because of establishing the prior registration of three-resolution images. In the training phase, the model was initially pre-trained with the normalized EDT maps sequentially in three resolutions, and then the model was fine-tuned with multiscale binary maps. In the evaluation phase for the model usage, we simply need low-resolution binary images as input. Therefore, the transfer learning workflow in a company with data augmentation and generation technique in each resolution (generation of abundant various pore structure realizations through morphological filters) resulted in a multi-scale aware network with simplistic and straightforward usage that can accurately estimate the permeability of the low-resolution images with larger FOV that is not possible through the LBM. This approach is not commonly seen in standard CNNs implemented in previous studies, where each resolution is typically treated independently.

(3). Overfitting: Traditional ANNs and CNNs, mainly when they are deep and have many parameters, are prone to overfitting, especially with limited training data. Overfitting results in models that perform well on training data but poorly on unseen data. By employing the transfer learning strategy, the model leverages learned features from HR data to enhance performance on LR images, thereby mitigating the risk of overfitting, even with limited data per scale.

(4). Optimization for limited computational resources: We have designed our network to achieve optimal results within the constraints of limited data and GPU memory capacity. Our model uses larger image volumes (300^3 voxels) for input, a considerable step up from traditional models that often use smaller dimensions. This is achieved by carefully crafting a network architecture that is both robust and computationally efficient, ensuring that the model is feasible for practical applications despite hardware limitations.

(5). Handling of spatial correlations: Traditional ANNs do not inherently account for spatial hierarchies and correlations in image data and often require extensive feature engineering to achieve optimal performance. CNNs, on the other hand, require less manual feature engineering due to their ability to automatically learn spatial hierarchies by using convolutional layers, but they can still struggle with capturing long-range dependencies within 3D images. The multiscale nature of our CNN allows for better handling of spatial correlations and textures at different scales, which is crucial for accurately capturing the heterogeneity in porous materials like carbonate rocks.

5.3. Test data results

The training data performance, while promising, does not guarantee model effectiveness on the diverse range of heterogeneous carbonate microstructures encountered in practical scenarios. Consequently, the next logical step is to assess the trained multiscale CNN on additional test microstructures derived from the various complexities across entire volumes of the MR and LR carbonate images. This evaluation will provide a more comprehensive understanding of the model's practical application range, as we explore in this sections. The permeability estimates generated from our model for individual rock units may be used to infer effective permeability at the core scale using Darcy's simulation techniques (Botha and Sheppard, 2016). The modeled porosity was in agreement with laboratory observations.

We cross-verified the model's predictions with MIP data obtained from the same sub-plug for the test data, which includes binary images of the whole sub-plug at relatively lower resolutions without corresponding target labels as in the registered calibration zone. Consequently, the test subvolumes are scaled to match the REV dimensions used during training, allowing for comparing their permeability estimates to the experimental data after scaling up. The chosen image dimension for training was based on a standard REV for the carbonate images under study, which typically exhibit 8 to 12% porosity and encompass a various range of pore geometries; notably, the LR and MR images display a wider pore size distribution than those in the training set. We applied two methods to evaluate the calibrated multiscale CNN model:

1) Applying the moving window technique outlined in the methodology, we manoeuvred a computational cube of 300^3 voxels—corresponding to a REV—across the entire binary carbonate images at MR and LR resolutions to extract 102 subvolumes from each resolution for testing. These images were then used to assess the model's performance. In the subsequent phase, permeability estimated from the lower-resolution cube images of the carbonate core

plug is scaled up. We adopt a voting-based scale-up method for result validation using MIP laboratory data, selecting the geometric mean for this purpose (Miarelli and Della Torre, 2021). This approach is backed by research indicating that the geometric mean is apt for characterizing the most probable permeability behavior in a heterogeneous porous medium consisting of n regions with distinct uniform permeabilities from K1 to Kn. The mean aligns with the mode of a log-normal distribution, which often models the permeability distribution in such media. Our findings are presented in Table 5. The prediction relative error is calculated to be 7.4828% for LR test images and 16.9605% for MR test images, respectively. This discrepancy may be due to the capacity of the differing resolutions to resolve the grain/pore structure despite a similar distribution. The model, which was fine-tuned using LR images, yielded a lower error rate, showcasing the network's adaptability across different pore geometries found in the non-registered, full-plug lower-resolution images. Considering the training set's representative coverage of pore texture and morphology, the results align impressively with the MIP laboratory measurements.

1) In the second testing method, we resize (or downsample) the lower-resolution micro-CT images of our carbonate, similar to what Alqahtani et al. (2018) carried out with sandstone images. Consequently, in this investigation, the LR images of various cube sizes, including 400^3, 500^3, 550^3, 600^3, and 650^3 voxels, are resampled to 300^3 voxels using the bicubic interpolation method. With this procedure, the porosities of the cubes were maintained in the range of 10 to 11%, as opposed to other approaches such as bilinear (which resulted in additional variations in porosity), since we preferred to preserve the porosity in this range with the fewest of pore structure destruction. The resolution or voxel size of the cubes varied in the following order after resampling: 4, 5, 5.5, 6, and 6.5 μm. Again, the model was evaluated using these new images with lower resolutions than the training LR images but with the same dimensions (300^3). As shown in Table 6, although the mean estimated permeability is in the range of the predicted values in the registered area (Table 4), the results in this instance were not as encouraging as anticipated relative to the MIP data. These results imply that downsampling the LR larger volumes of such a complex, tight carbonate may not be an appropriate method for estimating permeability through the model, and even when preserving the porosity of the volumes in the REV range, significant features of the original pore structures are sacrificed. Like all deep learning algorithms, the performance of the network is largely dependent on the quality/quantity of the training data and its representativeness with respect to the test data (Jackson

Table 6
Model evaluation with resampled test images.

Row	Resampled LR images to 300^3 voxels	Predicted Permeability (mD)	Relative Error (%)
1	400^3	2.5455	69.3333
2	500^3	3.0924	106.16
3	550^3	2.0176	34
4	600^3	1.8276	21.3333
5	650^3	2.1885	40
6	Geometric mean (mD)	2.2939	52.9318

Table 5
Model evaluation with the original test images.

Row	Test images	Mean Permeability(mD)	STD	Geometric mean (mD)	Relative Error* (%)	Total data
1	Middle resolution	1.9159	0.7255	1.7544	16.9605	102
2	Low resolution	1.8166	0.7052	1.6122	7.4828	102

* Relative Error $= \left| \dfrac{K_{MICP} - K_{Model}}{K_{MICP}} \right| \times 100$.

et al., 2021). Consequently, we must be more cautious when applying digital 3D pore-scale modeling to heterogeneous carbonates in general.

It is typically more challenging for neural networks to handle test data with out-of-range characteristics than training samples because the features of the out-of-range samples (resampled images in this scenario) may differ from the characteristics of the training samples. Nevertheless, the results indicate that the multiscale CNN achieves accurate performance when dealing with unseen test data with different structures in the non-registered regions of the actual lower-resolution images and successfully links the complex multiscale morphology of the carbonate porous media with this degree of heterogeneity to their permeability variations. The efficiency of the computations is such that the permeability of a new sample from the LR and MR images with larger FOVs can be estimated in a few seconds after the training is complete, thereby allowing the avoidance of the intensive computations for effective permeability, which are not applicable by the LBM on low-resolution domains.

5.4. Normal probability plot of results

Fig. 16 presents the permeability prediction histograms and box plots for the entire set of real test images at both LR and MR resolutions. Notably, the model retains its precision for MR and LR samples, which possess a range of complex geometries. The permeability distribution at both resolutions shows comparable trends, with histograms that approximate a normal bell curve, indicating the model's capability to generalize across different sample complexities.

In order to scrutinize the normality of the estimated permeabilities, we turn to Q-Q plots, which are statistical tools used to compare the quantiles of two distributions. A Q-Q plot can determine if a dataset follows a certain theoretical distribution, such as normal, exponential, or uniform, or assess if two datasets come from similar distributions. This is particularly useful in regression problems, where confirming distribution uniformity between separate training and test datasets is essential. Such plots can validate whether the distributions generated by the multiscale CNN model are normal, thereby facilitating data experimentation. Considering the prevalence of normal distributions in natural phenomena, establishing normality in our model's outputs is advantageous.

Fig. 17(a) features the Q-Q plot for the LR residuals, which compares the differences between the LBM target responses and the predicted LR permeabilities within the registered area. The normal probability plot for the residuals is constructed to verify the normality assumption. A roughly linear plot suggests that the residuals are normally distributed. The diagonal reference line helps ascertain the linear correlation between theoretical and sample quantiles. However, the plot reveals some non-linearity, especially at the lower quantiles, indicating that the error terms for this data subset do not follow a normal distribution. This deviation is primarily attributed to the lower permeability estimates for tighter samples, which tend to have less precision, as previously mentioned. The more accurate predictions for each LR sample are those that closely align with the normal distribution for their residuals.

Fig. 17(b) displays the Q-Q plot for the distributions of reference permeabilities and LR estimates within the registered zone. This plot confirms that the permeability estimates from the model closely resemble the reference data, as evidenced by the points aligning nearly linearly along the reference line. This alignment suggests that the network has effectively captured the core structures and features common to all registered images despite variations in resolution.

Additionally, for the test results, we offer a plot of estimated permeabilities from MR and LR test images compared with their respective theoretical quantiles, as demonstrated in Fig. 18. This figure confirms that the permeability estimations for both MR and LR have statistical distributions that predominantly align with normal distributions, echoing the trends observed in Fig. 16. Detailed analysis in Fig. 18(a and b) provides Q-Q plots from the perspective of kurtosis—a measure of the "tailedness" of a distribution. These plots exhibit minimal or insignificant deviations at the tails, indicating a leptokurtic (thin-tailed) distribution, corroborating the model's high prediction accuracy and unity with a normal distribution.

The multiscale CNN model is acclaimed for its precision and efficiency. It takes 531 h to calculate the permeability of 59 HR 3D samples using LBM, whereas training the CNN with all three sets of registered images on the specified hardware takes only 48 h. Moreover, once the model is built, it can predict permeability in mere seconds. This rapid prediction capability underscores the potential of deep learning algorithms to drastically curtail computation time while elevating permeability estimation accuracy. The trained network can also be applied immediately to new rock images previously not trained on.

However, it should be noted that the high-resolution data from the registered area substantiates the permeability estimates across entire image volumes with larger voxel sizes from MR and LR images. If the
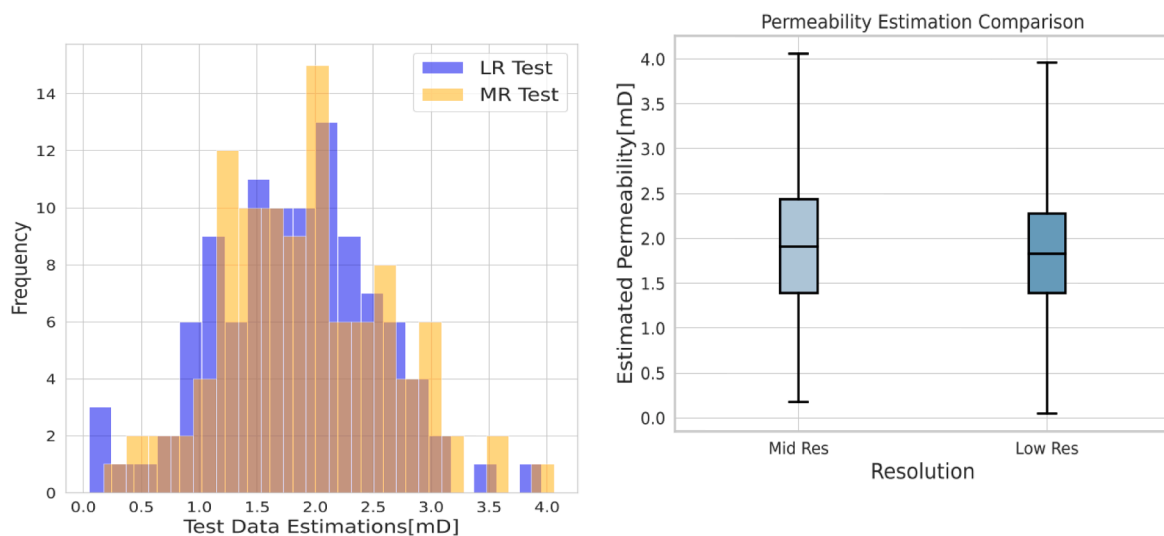

Fig. 16. Histograms (left) and box plot (right) of estimated permeabilities for the whole LR and MR test images in the non-registered area. The comparison reveals that the distributions of predictions at both resolutions are broadly similar and close to a normal bell-shaped distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
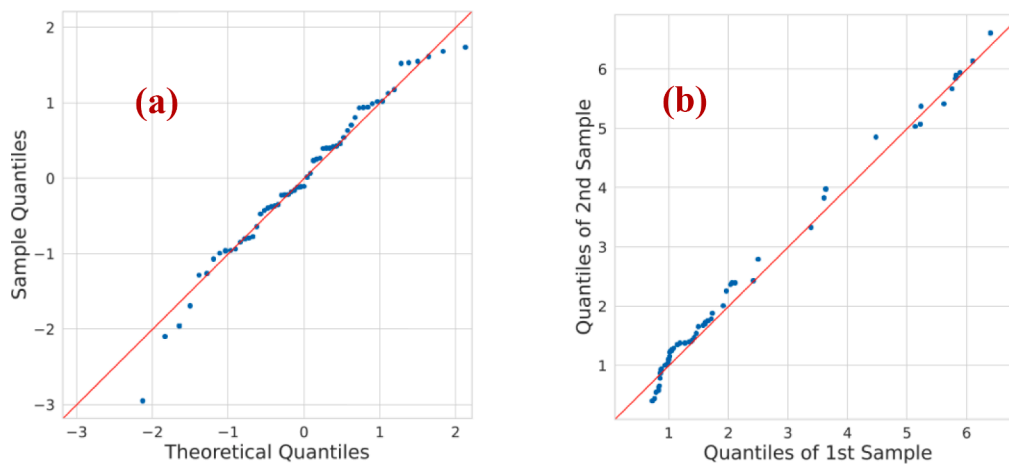
Fig. 17. The Q-Q plots for the LR residuals (a) and for the target references against LR permeability predictions in the registered area (b).
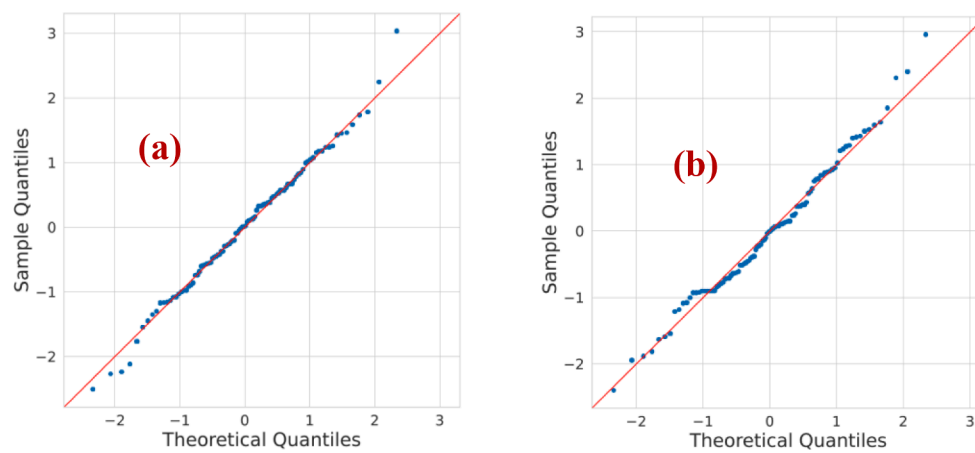

Fig. 18. The Q-Q plot for permeability estimations from the a) LR and b) MR test images.

calibration registered region fails to capture a comprehensive range of rock textures and morphologies—especially without data augmentation across resolutions—prediction uncertainties could escalate. Thus, the quality of the image and the computed permeability are crucial to the calibration model's success and the fidelity of predicted outcomes. The synergy of transfer learning with data augmentation—utilizing normalized EDT and binary maps for each resolution—has proven to enhance the model's predictive reach. The proposed method shows promise for application to other complex multiscale porous media with varying transport properties.

It is essential to acknowledge that the efficacy of the transfer learning operation employed in the described workflow could attract our interest if it were to be extended to additional rock samples. However, significant factors need to be addressed before conducting further investigations. The most critical aspects to remember are the following: The applicability of transfer learning across diverse rock samples depends on the degree of similarity evidenced by their characteristics. For instance, transferring knowledge from carbonate to sandstone samples may be effective if the pore structures or other relevant properties exhibit underlying similarities. However, the more dissimilar the samples, the less effective the transfer learning might be. In our workflow, the foundation of image registration in multiscale imaging strengthens the transfer learning strategy. As a result, during the training phase, the heterogeneous pore structures at various resolutions are initially aligned and comparable, albeit with varying degrees of detail at each resolution scale.

Implementing transfer learning across different rock samples

involves leveraging a model trained on one rock type (or dataset) to enhance or expedite learning on a different but related rock type. This approach is particularly beneficial in digital rock physics problems, where obtaining varied and extensive datasets for each rock type is challenging. By transferring knowledge from one sample to another, the model can efficiently adjust to new data, reducing the need for extensive training from scratch. In addition, this methodology can prove to be especially advantageous in situations involving restricted data about specific rock types. A model trained on a vast and varied dataset of rock images has the potential to acquire characteristics that can be applied to different kinds of rocks.

Nonetheless, several potential limitations still exist that need to be carefully treated. A notable limitation emerges due to the intrinsic heterogeneity of the rock samples. Different rock types can vastly exhibit different pore geometries, differences in mineral composition, REV and physical properties. Such variations can affect the model's ability to accurately transfer learning from one sample to another since the success of transfer learning depends on the degree of similarity between the training and target datasets. Another issue is related to the model overfitting. Overfitting may occur if the model fails to generalize effectively from the training dataset to the new dataset despite its successful performance on the training set. This necessitates careful tuning and validation when applying the model to different rock types, especially in the case of small datasets.

Furthermore, domain adaptation presents a significant challenge, requiring the model to adjust according to the new dataset's particular characteristics. This often requires fine-tuning the model with a smaller

subset of the target dataset to ensure better performance. These methods can help adjust the model to the specific characteristics of the new rock type, improving its performance. Data preprocessing needs are another critical factor. Different rock samples require unique preprocessing steps to make the datasets compatible with effective transfer learning. This adds complexity to the implementation process. Therefore, our study demonstrates the potential of this approach, and future research will delve deeper into optimizing transfer learning for broader applications in diverse rock types.

The triumph of the present study lies in synthesizing a scalable LBM for fluid flow property predictions, such as effective permeability in low-resolution micro-CT images of a complex carbonate with a larger FOV, something out of reach for direct simulation. Here, "synthetic" LBM refers to creating a sophisticated multiscale CNN model using standard GPU and hardware specifications that estimate transport properties much faster than traditional LBM and with comparable accuracy. Prediction times for each test image were on the order of a few seconds, while LBM simulations for each tight carbonate subvolume required about 9 h for convergence on our hardware. In the digital rock physics framework outlined in this paper, the CNN model's hierarchical benefits are most significant in estimating effective permeability, integrating information from a vast spectrum of pore sizes and throat varieties at different resolutions.

The current workflow also has the potential to provide a solid foundation for extending to multiphase flow properties. The first step involves integrating a comprehensive dataset specific to multiphase flow properties, such as capillary pressure and relative permeability, into the training and validation process. This dataset can be obtained from laboratory measurements or high-fidelity numerical simulations that capture multiphase flow behavior in complex rocks. One approach offers that this dataset can consist of multiscale images of rock samples under various wetting conditions and at different saturation levels, along with corresponding capillary pressure and relative permeability measurements. The next factor reflects the CNN architecture that may need adaptations to effectively process and learn from multiphase flow data. This could involve modifying the network layers to handle additional inputs or outputs (e.g., outputting multiple properties or training separate models for each property). Also, physics-based constraints related to multiphase flow (e.g., capillary pressure and relative permeability principles) can be incorporated to enhance the model's predictive capability. This could be achieved through hybrid modeling, where machine learning is combined with traditional physics-based models. For instance, the method proposed by (Tang et al., 2022) but implemented for single scale and single phase flow or considering a hybrid approach that combines CNNs with traditional simulation methods (e.g., pore-network modeling) for multiphase flow, leveraging the strengths of both data-driven and physics-based models such as the works done by (Xie et al., 2023) and (Rabbani et al., 2020).

Nevertheless, extending the model to predict multiphase flow properties introduces additional challenges. These include capturing the intricate interplay between fluid properties, rock properties, and flow dynamics through multiscale imaging and dealing with the increased variability in multiphase flow properties in different resolutions. Another critical challenge is related to the REV size of the simulated ground truth data since multiphase flow properties like relative permeability simulation need even larger REV sizes (Saxena et al., 2018), which increase the computational costs. Therefore, extending our current workflow to predict multiphase flow properties is feasible and would increase the model's applicability in realistic scenarios and contribute significantly to the field of multiscale digital rock physics while acknowledging the challenges and the need for further research in this area.

6. Summary and conclusions

In this study, we introduce a novel and computationally efficient

workflow designed to address the resolution field-of-view (FOV) trade-off in micro-CT imaging within the realm of digital rock physics (DRP). By integrating multiscale/multiresolution imaging with a 3D convolutional neural network (CNN)-based deep learning methodology, our workflow estimates transport properties, specifically effective permeability, from low-resolution micro-CT images of complex carbonate samples with large FOV. Our multiscale approach spans three different resolutions, capturing the heterogeneity of the sample and successfully extrapolating high-resolution image properties to larger FOVs with lower resolution. The key findings and contributions of our research are summarized as follows:

- The model demonstrates high accuracy in permeability estimation (R-squared $\approx$ 0.96) across various pore geometries and scales, aligning with lattice Boltzmann method (LBM)-derived reference values on representative elementary volume (REV) subvolumes, despite the complexities of the carbonate samples and the limited size of the training dataset per scale.
- Residual analysis for each scale within the registered region shows homoscedasticity, with residuals distributed randomly around zero, indicating a high-quality and unbiased model capable of managing the complexity of carbonate media.
- During the testing phase, the model effectively predicted permeability using various binary domains from the entire, lower-resolution image volumes, showcasing the model's ability to generalize across different pore structures and link these variations to permeability.
- Our research reveals that down sampling low-resolution images to estimate permeability is less reliable, as essential pore structure features crucial for prediction accuracy are often lost, even when porosity is preserved within the REV range.
- A significant discovery is the successful application of transfer learning and data augmentation strategies, which significantly improve the model's performance and generalization capabilities on the small datasets of each scale, particularly when pre-training on 3D normalized Euclidean distance transform (EDT) maps.
- Histogram and quantile-quantile (Q-Q) plot analyses suggest that most deviations from normality in residuals are linked to lower precision in estimating the permeability of tighter samples.

7. Future directions

The study underscores the importance of image quality and sample representativeness in digital rock physics (DRP) modeling, as these factors significantly affect the diversity observed in pore size distributions and morphologies. To build upon the findings of this research and further enhance the capabilities of DRP models, future research directions should include:

- Expanding the generalization capability of our workflow to apply it to a wide array of porous materials. This expansion would increase the utility of our model across different geological and environmental scenarios. Future studies could involve acquiring sub-plugs or region-of-interest scans from additional rock samples to test the model's predictive accuracy across entire lower-resolution images from multiple calibration volumes.
- Broadening the research scope to include a wider variety of heterogeneities, such as actual fractures, vugs, and two-phase transport properties. By incorporating these complex features into the modeling process, the predictive power and applicability of the model to real-world scenarios will be significantly improved.
- Exploring lower-resolution systems at scales ranging from 20 μm to high-resolution sub-micron levels. This investigation would aim to enhance the transfer of information in situations where full core images are unable to capture primary flow pathways, yet there is still a need for accurately mapping variations in transport properties.

CRediT authorship contribution statement

Iman Nabipour: Validation, Software, Conceptualization, Formal analysis, Methodology, Writing – original draft. **Amir Raoof:** Resources, Writing – review & editing. **Veerle Cnudde:** Funding acquisition, Resources, Writing – review & editing. **Hamed Aghaei:** Resources, Writing – review & editing. **Jafar Qajar:** Supervision, Formal analysis, Resources, 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.

Acknowledgement

References

Abrosimov, K., Gerke, K., Semenkov, I., Korost, D., 2021. Otsu's algorithm in the segmentation of pore space in soils based on tomographic data. Eurasian Soil Science 54, 560–571.

Aftab, A., Hassanpouryouzband, A., Xie, Q., Machuca, L.L., Sarmadivaleh, M., 2022. Toward a Fundamental Understanding of Geological Hydrogen Storage. Industrial & Engineering Chemistry Research 61, 3233–3253.

Almotairi, K.H., 2020. A Global Two-Stage Histogram Equalization Method for Gray-Level Images. J. ICT Res. Appl 14, 95–114.

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.

Alqahtani, N., Armstrong, R.T., Mostaghimi, P., 2018. Deep learning convolutional neural networks to predict porous media properties. In: SPE Asia Pacific oil and gas conference and exhibition. Society of Petroleum Engineers.

Alqahtani, N.J., Chung, T., Wang, Y.D., Armstrong, R.T., Swietojanski, P., Mostaghimi, P., 2021. Flow-based characterization of digital rock images using deep learning. SPE Journal 26, 1800–1811.

Anderson, T.I., Vega, B., Kovscek, A.R., 2020. Multimodal imaging and machine learning to enhance microscope images of shale. Computers & Geosciences 145, 104593.

Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., Keehm, Y., Krzikalla, F., Lee, M., Madonna, C., 2013. Digital rock physics benchmarks—Part I: Imaging and segmentation. Computers & Geosciences 50, 25–32.

Anzum, F., 2021. Exploring Convolutional Neural Networks and Transfer Learning for Oil Sands Drill Core Image Analysis. Science.

Araya-Polo, M., Alpak, F.O., Hunter, S., Hofmann, R., Saxena, N., 2020. Deep learning–driven permeability estimation from 2D images. Computational Geosciences 24, 571–580.

Bai, Y., Berezovsky, V., 2020. Digital rock core images resolution enhancement with improved super resolution convolutional neural networks. In: Proceedings of the 2020 12th International Conference on Machine Learning and Computing, pp. 401–405.

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 Resources 51, 197–216.

Botha, P.W., Sheppard, A.P., 2016. Mapping permeability in low-resolution micro-CT images: A multiscale statistical approach. Water Resources Research 52, 4377–4398.

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-Science Reviews 155, 93–128.

Çelik, M.Y., Sert, M., 2021. An assessment of capillary water absorption changes related to the different salt solutions and their concentrations ratios in the Döğer tuff (Afyonkarahisar-Turkey) used as building stone of cultural heritages. Journal of Building Engineering 35, 102102.

Chaaban, M., Heider, Y., Markert, B., 2022. A multiscale LBM–TPM–PFM approach for modeling of multiphase fluid flow in fractured porous media. International Journal for Numerical and Analytical Methods in Geomechanics 46, 2698–2724.

Chen, H., Chen, S., Matthaeus, W.H., 1992. Recovery of the Navier-Stokes equations using a lattice-gas Boltzmann method. Physical review A 45, R5339.

Chen, X., Zhou, Y., 2017. Applications of digital core analysis and hydraulic flow units in petrophysical characterization. Advances in Geo-Energy Research 1, 18–30.

Cnudde, V., Boone, M.N., 2013. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. Earth-Science Reviews 123, 1–17.

Cui, L.-K., Sun, J.-M., Yan, W.-C., Dong, H.-M., 2020. Multi-scale and multi-component digital core construction and elastic property simulation. Applied Geophysics 17, 26–36.

Da Wang, Y., Blunt, M.J., Armstrong, R.T., Mostaghimi, P., 2021. Deep learning in pore scale imaging and modeling. Earth-Science Reviews 215, 103555.

Da Wang, Y., Chung, T., Armstrong, R.T., Mostaghimi, P., 2020. ML-LBM: Machine Learning Aided Flow Simulation in Porous Media. arXiv preprint arXiv:2004.11675.

De Boever, W., Derluyn, H., Van Loo, D., Van Hoorebeke, L., Cnudde, V., 2015. Data-fusion of high resolution X-ray CT, SEM and EDS for 3D and pseudo-3D chemical and structural characterization of sandstone. Micron 74, 15–21.

Dehghan Khalili, A., 2013. Transport properties for heterogeneous carbonate rocks. UNSW Sydney.

Elmorsy, M., El-Dakhakhni, W., Zhao, B., 2022. Generalizable Permeability Prediction of Digital Porous Media via a Novel Multi-Scale 3D Convolutional Neural Network. Water Resources Research 58 e2021WR031454.

Erofeev, A., Orlov, D., Ryzhov, A., Koroteev, D., 2019. Prediction of porosity and permeability alteration based on machine learning algorithms. Transport in Porous Media 128, 677–700.

Fu, J., Wang, M., Chen, B., Wang, J., Xiao, D., Luo, M., Evans, B., 2023. A data-driven framework for permeability prediction of natural porous rocks via microstructural characterization and pore-scale simulation. Engineering with Computers 1–32.

Gärttner, S., Alpak, F.O., Meier, A., Ray, N., Frank, F., 2021. Estimating permeability of 3D micro-CT images by physics-informed CNNs based on DNS. arXiv preprint arXiv: 2109.01818.

Graczyk, K.M., Matyka, M., 2020. Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. Scientific reports 10, 1–11.

Gupta, S., Gupta, S.K., 2021. Development and evaluation of an innovative Enhanced River Pollution Index model for holistic monitoring and management of river water quality. Environmental Science and Pollution Research 28, 27033–27046.

Hong, J., Liu, J., 2020. Rapid estimation of permeability from digital rock using 3D convolutional neural network. Computational Geosciences 24, 1523–1539.

Hunter, J.D., 2007. Matplotlib: A 2D graphics environment. Computing in science & engineering 9, 90–95.

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 arXiv:2111.01270.

Jiang, F., Guo, Y., Tsuji, T., Kato, Y., Shimokawara, M., Esteban, L., Seyyedi, M., Pervukhina, M., Lebedev, M., Kitamura, R., 2022. Upscaling the permeability properties using multiscale X-ray-CT images with digital rock modeling and deep learning techniques. Authorea Preprints.

Kahshan, M., Lu, D., Abu-Hamdeh, N.H., Golmohammadzadeh, A., Farooq, A.A., Rahimi-Gorji, M., 2020. Darcy-Brinkman flow of a viscous fluid through a porous duct: Application in blood filtration process. Journal of the Taiwan Institute of Chemical Engineers 117, 223–230.

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, 9855.

Kamrava, S., Im, J., de Barros, F.P., Sahimi, M., 2021. Estimating Dispersion Coefficient in Flow Through Heterogeneous Porous Media by a Deep Convolutional Neural Network. Geophysical Research Letters 48 e2021GL094443.

Kamrava, S., Tahmasebi, P., Sahimi, M., 2020. Linking morphology of porous media to their macroscopic permeability by deep learning. Transport in Porous Media 131, 427–448.

Khodja, M.R., Li, J., Hussaini, S.R., Ali, A.Z., Al-Mukainah, H.S., Jangda, Z.Z., 2020. Consistent prediction of absolute permeability in carbonates without upscaling. Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles 75, 14.

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

Liu, M., Ahmad, R., Cai, W., Mukerji, T., 2023. Hierarchical Homogenization With Deep-Learning-Based Surrogate Model for Rapid Estimation of Effective Permeability From Digital Rocks. Journal of Geophysical Research: Solid Earth 128 e2022JB025378.

Ma, Z., Sun, S., Yan, B., Kwak, H., Gao, J., 2023. Enhancing the Resolution of Micro-CT Images of Rock Samples via Unsupervised Machine Learning based on a Diffusion Model. In: SPE Annual Technical Conference and Exhibition? SPE p. D021S028R005.

Mandzhieva, R., 2017. Introduction to digital core analysis: 3D reconstruction, numerical flow simulations and pore network modeling. NTNU.

Marcato, A., Boccardo, G., Marchisio, D., 2022. From Computational Fluid Dynamics to Structure Interpretation via Neural Networks: An Application to Flow and Transport in Porous Media. Industrial & Engineering Chemistry Research.

Mehmani, A., Verma, R., Prodanović, M., 2020. Pore-scale modeling of carbonates. Marine and Petroleum Geology 114, 104141.

Miarelli, M., Della Torre, A., 2021. Workflow Development to Scale up Petrophysical Properties from Digital Rock Physics Scale to Laboratory Scale. Transport in Porous Media 140, 459–492.

Najafi, A., Siavashi, J., Ebadi, M., Sharifi, M., Fahimpour, J., Koroteev, D., 2021. Upscaling permeability anisotropy in digital sandstones using convolutional neural networks. Journal of Natural Gas Science and Engineering 96, 104263.

Poonoosamy, J., Lu, R., Lönartz, M.I., Deissmann, G., Bosbach, D., Yang, Y., 2022. A lab on a chip experiment for upscaling diffusivity of evolving porous media. Energies 15, 2160.

Qajar, J., Arns, C.H., 2022. Chemically Induced Evolution of Morphological and Connectivity Characteristics of Pore Space of Complex Carbonate Rock via Digital Core Analysis. Water Resources Research 58 e2021WR031298.

Rabbani, A., Babaei, M., Shams, R., Da Wang, Y., Chung, T., 2020. DeePore: a deep learning workflow for rapid and comprehensive characterization of porous materials. Advances in Water Resources 146, 103787.

Rao, P., Schaefer, L., 2020. Permeability estimation on tomographic images using curved boundary schemes in the lattice Boltzmann method. Advances in Water Resources 143, 103685.

Ribeiro, M.D., Rehman, A., Ahmed, S., Dengel, A., 2020. DeepCFD: Efficient steady-state laminar flow approximation with deep convolutional neural networks. arXiv preprint arXiv:2004.08826.

Santos, J.E., Pyrcz, M.J., Prodanović, M., 2022. 3D Dataset of binary images: A collection of synthetically created digital rock images of complex media. Data in Brief, 107797.

Santos, J.E., Xu, D., Jo, H., Landry, C.J., Prodanović, M., Pyrcz, M.J., 2020. PoreFlow-Net: A 3D convolutional neural network to predict fluid flow through porous media. Advances in Water Resources 138, 103539.

Santos, J.E., Yin, Y., Jo, H., Pan, W., Kang, Q., Viswanathan, H.S., Prodanović, M., Pyrcz, M.J., Lubbers, N., 2021. Computationally efficient multiscale neural networks applied to fluid flow in complex 3D porous media. Transport in porous media 140, 241–272.

Saxena, N., Hofmann, R., Alpak, F.O., Berg, S., Dietderich, J., Agarwal, U., Tandon, K., Hunter, S., Freeman, J., Wilson, O.B., 2017a. References and benchmarks for pore-scale flow simulated using micro-CT images of porous media and digital rocks. Advances in Water Resources 109, 211–235.

Saxena, N., Hofmann, R., Alpak, F.O., Dietderich, J., Hunter, S., Day-Stirrat, R.J., 2017b. Effect of image segmentation & voxel size on micro-CT computed effective transport & elastic properties. Marine and Petroleum Geology 86, 972–990.

Saxena, N., Hows, A., Hofmann, R.O., Alpak, F., Freeman, J., Hunter, S., Appel, M., 2018. Imaging and computational considerations for image computed permeability: Operating envelope of Digital Rock Physics. Advances in Water Resources 116, 127–144.

Schultz, R.A., Williams-Stroud, S., Horváth, B., Wickens, J., Bernhardt, H., Cao, W., Capuano, P., Dewers, T.A., Goswick, R.A., Lei, Q., 2023. Underground energy-related product storage and sequestration: site characterization, risk analysis and monitoring, 528. Geological Society, London, Special Publications. SP528-2022-2066.

Singh, A., Regenauer-Lieb, K., Walsh, S.D., Armstrong, R.T., van Griethuysen, J.J., Mostaghimi, P., 2020. On representative elementary volumes of grayscale micro-CT images of porous media. Geophysical Research Letters 47 e2020GL088594.

Sit, M., Demiray, B.Z., Xiang, Z., Ewing, G.J., Sermet, Y., Demir, I., 2020. A comprehensive review of deep learning applications in hydrology and water resources. Water Science and Technology 82, 2635–2670.

Soltanmohammadi, R., Faroughi, S.A., 2023. A comparative analysis of super-resolution techniques for enhancing micro-CT images of carbonate rocks. Applied Computing and Geosciences 20, 100143.

Srisutthiyakorn, N., 2016. Deep-learning methods for predicting permeability from 2D/3D binary-segmented images, SEG technical program expanded abstracts 2016. Society of Exploration Geophysicists 3042–3046.

Sudakov, O., Burnaev, E., Koroteev, D., 2019. Driving digital rock towards machine learning: Predicting permeability with gradient boosting and deep neural networks. Computers & geosciences 127, 91–98.

Sun, H., Di, D., Tao, G., Vega, S., Li, K., Liu, L., Belhaj, H., 2017a. Carbonate rocks: a case Study of rock properties evaluation using multi-scale digital images. In: Abu Dhabi International Petroleum Exhibition & Conference. OnePetro.

Sun, H., Tao, G., Vega, S., Wang, B., Liu, H., Li, K., 2017b. Multi-Scale Image Analysis of Digital Carbonate Rock. In: 79th EAGE Conference and Exhibition 2017. European Association of Geoscientists & Engineers, pp. 1–5.

Sungkorn, R., Morcote, A., Carpio, G., Davalos, G., Mu, Y., Grader, A., Derzhi, N., Toelke, J., 2015. Multi-scale and upscaling of digital rock physics with a machine that can learn about rocks. Inter. Sym. Soc. Cor. Anal.

Takbiri-Borujeni, A., Kazemi, H., Nasrabadi, N., 2020. A data-driven surrogate to image-based flow simulations in porous media. Computers & Fluids 201, 104475.

Tang, P., Zhang, D., Li, H., 2022. Predicting permeability from 3D rock images based on CNN with physical information. Journal of Hydrology 606, 127473.

Tawfeeq, Y.J., Al-Sudani, J.A., 2020. Digital Rock Samples Porosity Analysis by OTSU Thresholding Technique Using MATLAB. Iraqi Journal of Chemical and Petroleum Engineering 21, 57–66.

Tembely, M., AlSumaiti, A., 2019. Deep learning for a fast and accurate prediction of complex carbonate rock permeability from 3D micro-CT images. In: Abu Dhabi International Petroleum Exhibition & Conference. OnePetro.

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, 1541–1556.

Tembely, M., AlSumaiti, A.M., Alameri, W.S., 2021. Machine and deep learning for estimating the permeability of complex carbonate rock from X-ray micro-computed tomography. Energy Reports 7, 1460–1472.

Tian, J., Qi, C., Sun, Y., Yaseen, Z.M., Pham, B.T., 2021. Permeability prediction of porous media using a combination of computational fluid dynamics and hybrid machine learning methods. Engineering with Computers 37, 3455–3471.

Ulyanov, D., Vedaldi, A., Lempitsky, V., 2016. Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022.

Verri, I., Della Torre, A., Montenegro, G., Onorati, A., Duca, S., Mora, C., Radaelli, F., Trombin, G., 2017. Development of a digital rock physics workflow for the analysis of sandstones and tight rocks. Journal of Petroleum Science and Engineering 156, 790–800.

Waskom, M.L., 2021. Seaborn: statistical data visualization. Journal of Open Source Software 6, 3021.

Weiss, K., Khoshgoftaar, T.M., Wang, D., 2016. A survey of transfer learning. Journal of Big data 3, 1–40.

Withers, P.J., Bouman, C., Carmignato, S., Cnudde, V., Grimaldi, D., Hagen, C.K., Maire, E., Manley, M., Du Plessis, A., Stock, S.R., 2021. X-ray computed tomography. Nature Reviews Methods Primers 1, 18.

Wu, J., Yin, X., Xiao, H., 2018. Seeing permeability from images: fast prediction with convolutional neural networks. Science bulletin 63, 1215–1222.

Xie, C., Zhu, J., Yang, H., Wang, J., Liu, L., Song, H., 2023. Relative permeability curve prediction from digital rocks with variable sizes using deep learning. Physics of Fluids 35.

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks? Advances in neural information processing systems 27.

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.