AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR031454

Generalizable Permeability Prediction of Digital Porous Media via a Novel Multi-Scale 3D Convolutional Neural Network

Mohamed Elmorsy[1], Wael El-Dakhakhni[1], and Benzhong Zhao[1]

[1]Department of Civil Engineering, McMaster University, Hamilton, ON, Canada

Key Points:

- We introduce a novel three-dimensional convolutional neural network model for end-to-end permeability predictions with excellent accuracy
- The accuracy results from increasing dataset diversity and size, using multi-scale feature aggregation, and optimizing network architecture
- The model is generalizable, and it is capable of predicting the permeability of previously unseen samples

Correspondence to:

B. Zhao,
robinzhao@mcmaster.ca

Citation:

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. https://doi.org/10.1029/2021WR031454

Received 22 OCT 2021
Accepted 7 MAR 2022

Abstract Subsurface characterization is critical in understanding and controlling many natural and industrial processes including groundwater movement, oil extraction, and geological carbon dioxide sequestration. While recent advances in three-dimensional (3D) imaging of core samples have enabled digital subsurface characterization, the exorbitant computational cost associated with direct numerical simulation in 3D remains a persistent challenge. In contrast, machine learning models are much more efficient, though their use in subsurface characterization is still in its infancy. Here, we introduce a novel 3D convolutional neural network (CNN) for end-to-end prediction of permeability, which is a fundamental characteristic of subsurface porous media. We show that increasing the dataset size and diversity, utilizing multi-scale feature aggregation, and optimizing the network architecture elevate the model accuracy beyond that of existing state-of-the-art 3D CNN models for permeability prediction. We demonstrate that the model is generalizable, and it is capable of predicting the permeability of previously unseen samples with an excellent accuracy.

1. Introduction

Subsurface fluid flow is critical in many natural and industrial processes such as groundwater movement (Bear, 2013; Freeze & Cherry, 1979), energy extraction (Lake, 1989; Orr & Taber, 1984), and geological carbon sequestration (Juanes et al., 2010; Szulczewski et al., 2012). Over geological times, the Earth's subsurface evolved to encompass a wide variety of different rock and soil types, which are characterized by vastly different physical and chemical properties (Tarbuck et al., 2005). In particular, permeability—A property that measures the ease with which fluids pass through a porous medium—Exerts fundamental control over fluid flow in the subsurface. In contrast to the wealth of information on the Earth's surface, our knowledge of the subsurface is coarse, limited and often inaccurate. This is in large part due to the time-consuming and expensive nature of traditional methods for characterizing porous media.

For decades, characterization of subsurface properties have relied on the retrieval and analysis of core samples via underground drilling. Despite the advent of seismic surveys in the past decades, drilling remains the most reliable and accurate way of characterizing the subsurface (Chan & Schmitt, 2015). However, drilling is a highly expensive venture—A single well can cost hundreds of thousands to millions of dollars depending on its depth and the surrounding geology (Hossain, 2015). Even after drilling is completed, the characterization of the retrieved core samples requires complex and time-consuming laboratory tests using specialized equipment. For example, permeability is traditionally measured by flowing a pressurized gas or liquid through a core sample at varying flow rates and measuring the resulting pressure drop across the sample (Bear & Bachmat, 1991). This procedure limits the re-usability of the core sample for other tests. Additionally, since many geologic media are anisotropic in nature, such tests need to be repeated for each principal direction of anisotropy. As a result, there are millions of uncharacterized cores and drill-cutting samples that remain dormant in repositories housed by national geologic surveys and individual energy and mining corporations across the globe.

Recent advances in imaging technologies such as X-ray micro computed tomography (micro-CT) have allowed visualizing the internal structure of opaque porous media in three-dimensions (3D) and characterize their physical properties *digitally* via numerical simulations (Andrä et al., 2013a, 2013b; Berg et al., 2017; Ketcham & Carlson, 2001; Spanne et al., 1994). This technology, commonly referred to as *digital rock analysis*, is rapidly transforming how rock permeability is determined in both academic and industrial settings (Blunt et al., 2013; Mostaghimi et al., 2013). The current benchmark for digital characterization of rock permeability is achieved through high-resolution computational fluid dynamics (CFD) simulation of single-phase flow through digital rock samples. The permeability is calculated from the imposed pressure drop across the computed flow field

within the sample. The most popular computational methods for permeability characterization are lattice/particle-based models (e.g., lattice Boltzmann method (LBM); Chen & Doolen, 1998; Zhao et al., 2018) and continuum models that directly solve for the Stokes equation on the discretized pore geometry (Bijeljic et al., 2013; Guibert et al., 2016; Muljadi et al., 2016). Both methods have been successfully applied to predict the permeability of a variety of rock types including sandstones, limestones, and carbonates (Andrä et al., 2013a; Blunt et al., 2013; Boek & Venturoli, 2010; Chen & Doolen, 1998; Dong & Blunt, 2009; Manwart et al., 2002; Mostaghimi et al., 2013). However, direct numerical simulation of flow in 3D porous media remains computationally expensive, even with the implementation of parallel computing techniques (Blunt et al., 2013; Wang et al., 2019). An alternative to direct numerical simulation is pore network modelling (PNM), which models fluid flow by idealizing the complex 3D pore geometry as a lattice of pores and throats (Blunt, 2001, 2017). However, pore network models are less accurate than direct numerical simulations due to ambiguities involved in network extraction algorithms (Dong & Blunt, 2009).

To circumvent the exorbitant computational cost associated with direct numerical simulation while preserving accuracy, researchers have started to explore the use of machine learning models in digital permeability characterization (Alpak et al., 2018; Graczyk & Matyka, 2020).

Tian et al. (2020) applied a hybrid machine learning framework that combined genetic algorithm and artificial neural network (ANN) to predict the permeability of stochastically generated 3D porous media. Their dataset consisted of 1,000 synthetic samples of size $100 \times 100 \times 100$ cubic voxels with absolute permeability that ranged from 300 to 1,200 mD. The permeabilities of the synthetic samples were calculated using the LBM. The permeabilities calculated from LBM simulations were used as the training data for the hybrid machine learning model, which used 14 morphological features (e.g., porosity, tortuosity, pore throat size, etc.) as the input. Their model achieved a coefficient of determination $R^2 = 0.999$ between the predicted permeability and the LBM-calculated permeability. However, it should be noted that the study was based on a very small dataset (700 training samples and 300 testing samples) with a narrow range of permeability values (Wang et al., 2021). In another study, Tembely et al. (2020) used a variety of machine learning methods including linear regression, gradient boosting, and ANN to predict porous media permeability based on 3D micro-CT images. Their training dataset consisted of $400152 \times 152 \times 175$ cubic voxel subsamples with 2 μm/voxel resolution, and $759100 \times 100 \times 160$ cubic voxel sub-samples with 0.48 μm/voxel resolution. The permeabilities of the training dataset were obtained from LBM simulations, and they range from 100 to 400 mD. The authors concluded that ANN achieved the highest prediction accuracy ($R^2 = 0.91$).

Among the multitude of machine learning techniques used for digital permeability characterization, CNNs have shown particular promise, due in part to the image-based nature of digital rock analysis (Wang et al., 2021). CNN is a type of ANN, and it has been developed to solve difficult image-driven pattern recognition tasks (LeCun et al., 1989). CNNs consist of multiple convolution layers capable of automatically learning high-level features and representations from raw images without the need for hand-crafted features extracted by domain experts (Ghosal et al., 2018). CNNs were first developed to process 2D images (2D CNN), but they have recently been extended to analyze 3D volumetric imaging data (3D CNN; Huang et al., 2019; Korez et al., 2016; Milletari et al., 2016).

Sudakov et al. (2019) applied both 2D CNN and 3D CNN, as well as feature-based machine learning methods such as gradient regression trees and deep neural networks to predict the permeability of a single Berea sandstone sample. Their training dataset consisted of 9,261 subsamples of size $100 \times 100 \times 100$ cubic voxels, and the permeabilities of the training dataset were obtained from PNM simulations. They found that 3D CNN was the most accurate and efficient method. Santos et al. (2020) applied 3D CNN to predict the velocity field of single-phase flow through 3D digital rock samples. Their training dataset was based on CT scans of a bead pack with porosity that ranged from 0.11 to 0.298, and it consisted of 1,080 subsamples of size $80 \times 80 \times 80$ cubic voxels. The velocity field of the training dataset were obtained from LBM simulations. In addition to the 3D digital rock images, their model used four extracted features as input. The resulting model was used to predict the velocity field of single-phase flow through digital rock of varying complexities, which included different types of sandstone and limestone samples. The permeability calculated from the predicted velocity field yielded relative errors that ranged from 1.06% (Castlegate sandstone) to 27.3% (Bentheimer sandstone). Kamrava et al. (2020) trained a 3D CNN model using 500 synthetically generated porous media samples of size $200 \times 200 \times 200$ cubic voxels. The synthetic samples were stochastic reconstructions of a single sandstone core.

The permeabilities of the training dataset were obtained by numerically solving the Stokes equation, and they range from 100 to 500 mD. The authors applied the trained CNN to predict the permeability of a Fontainebleau sandstone, and the model achieved a prediction accuracy of $R^2 = 0.9$. Hong and Liu (2020) trained a 3D CNN model based on X-ray micro-CT scan of a Coconino sandstone core. The training dataset consisted of 3,158 subsamples of size $100 \times 100 \times 100$ cubic voxels. The permeabilities of the training dataset were obtained from LBM simulations, and they range from 10 to 6,000 mD. While the CNN model achieved a permeability prediction accuracy of $R^2 = 0.91$ on the same Coconino sandstone sample it was trained on, the prediction accuracy dropped to $R^2 = 0.69$ when the model was applied to a Bentheimer sandstone sample, which indicates a lack of model generalizability. Most recently, Alqahtani et al. (2021) trained a 3D CNN model on a much larger dataset, which included images of six different sandstone samples and two different carbonate samples. The training dataset consisted of ~25,000 subsamples of size $64 \times 64 \times 64$ cubic voxels, whose permeabilities range from 10 to 1,800 mD. The permeabilities of the training dataset were obtained using a fast numerical solver that provides an approximate solution to the Stokes equation (Chung et al., 2019a). The resulting permeability estimates are within 25% difference compared to those obtained from direct numerical simulation. Notably, the trained CNN model achieved a prediction accuracy of $R^2 = 0.86$ when it was applied to images of the Doddington sandstone, a rock sample that the model was not exposed to during the training process.

The existing literature clearly demonstrates the promising potential of 3D CNNs in digital rock permeability characterization. In particular, many authors have underscored CNN's short prediction time (less than a second) as its major advantage compared to traditional CFD simulations (hours to days; Alqahtani et al., 2021; Santos et al., 2020; Sudakov et al., 2019). At the same time, the lack of diverse labeled training data has been highlighted as a major obstacle in advancing the predictive capability of CNNs (Hong & Liu, 2020; Kamrava et al., 2020). Modern CNN architectures consist of a massive number of parameters that can only be properly tuned with a large amount of training data (Alzubaidi et al., 2021; Khan et al., 2020). Due to limitations in computational resources, current CNN models have either been trained on relatively small datasets consisting of a single rock type (Hong & Liu, 2020; Kamrava et al., 2020; Santos et al., 2020), with subsample size of less than or equal to $100 \times 100 \times 100$ cubic voxels (Alqahtani et al., 2021; Hong & Liu, 2020; Santos et al., 2020; Sudakov et al., 2019), or labeled with permeability values obtained from approximate solutions to the Stokes equation (e.g., PNM; Alqahtani et al., 2021; Sudakov et al., 2019).

Here, we introduce a novel 3D CNN model with multi-scale feature aggregation that has been trained with a massive dataset consisting of diverse rock types with a broad permeability range. We show that the large dataset size, in combination with the model architecture, elevate the model prediction accuracy beyond other existing 3D CNN models for permeability prediction. Furthermore, we demonstrate that the model is generalizable, and it is capable of predicting the permeability of previously unseen samples with excellent accuracy.

2. Methodology

2.1. Dataset

We use a comprehensive set of 3D images of different rock samples to develop the CNN permeability prediction model (Table 1). The images were previously aquired by researchers at Imperial College London with either a synchrotron X-ray beamline or a micro-CT scanner. These 3D images have been used to study various pore-scale processes of flow and transport in natural porous media (Blunt et al., 2013; Muljadi et al., 2016), and they are publicly-available via an online portal (Bijeljic & Raeini, 2015). The 3D images capture the pore spaces of rock cubes with sizes ranging from 1.1 to 3 mm and spatial resolutions ranging from 2.7 to 5.3 μm/voxel. We rescale the images such that each voxel corresponds to a physical dimension of 3 μm for all samples in our study. We validate the rescaling procedure by calculating the porosities of the rescaled samples, which are all within 2% of the original sample porosities. After rescaling, we extract subvolumes from the original 3D images by employing a sliding cube with size of $150 \times 150 \times 150$ cubic voxels and an overlapping stride of either 25 or 50 voxels (Table 1). The size of the subvolumes is limited by the memory capacity of currently available graphics processing units (GPUs). We foresee that the subvolume size of future models to grow as computer hardware specifically designed for machine learning continues to improve.

We obtain the permeability of the subvolumes numerically using OpenFoam®, which is an open source platform for computational fluid dynamics (CFD) simulations (Horgue et al., 2015). OpenFoam® models incompressible

Table 1
Digital Rock Samples Used to Develop the CNN Permeability Prediction Model

Rock Type	Size [mm]	Resolution [μm/voxel]	Porosity [−]	Stride [voxel]	Number of subvolumes [−]	Number of labeled permeabilies [−]
Bentheimer sandstone	3	3	0.22	50	5,096	15,288
Ketton limestone	3	3	0.13	50	4,042	12,126
Berea sandstone	2.1	5.3	0.19	25	5,387	16,161
Estaillades limestone	3.3	3.3	0.13	50	2,500	3,564
Doddington sandstone	2.7	2.7	0.19	50	2,562	7,686
Carbonate A	1.1	2.8	0.23	50	770	2,310
Carbonate B	2.1	5.3	0.16	50	402	1,206

steady-state viscous flow by solving the Navier-Stokes equation. The pressure and velocity of the fluid flow are solved iteratively using the semi-implicit method for pressure-linked equations (SIMPLE) algorithm. For each subvolume, we simulate water flow along each of the principal axis (i.e., x, y, z). Specifically, we prescribe a total pressure drop $\Delta p = 1$ Pa across the bounding surfaces that are normal to the flow direction, while the other bounding surfaces are prescribed as no-flow boundaries. Solid surfaces in the computational domain are assigned the no-slip boundary condition. The simulation yields the steady-state velocity field throughout the entire subvolume. Finally, we compute the volumetric flux Q across the subvolume by integrating the velocities normal to the outflow bounding surface (Equation 1a) and use Darcy's law to calculate the absolute permeability k of the subvolume (Equation 1b),

$$Q_i = \iint u_i dA \tag{1a}$$

$$k_i = \frac{Q_i \mu L}{\Delta p A}, \tag{1b}$$

where $i = x, y, z$ are the principal axes, μ is the dynamic viscosity of water, L is the length of the subvolume, and A is the area of the bounding surface.

We have compiled a dataset of over 57,500 subvolumes with labeled permeabilities obtained from direct numerical simulation. We perform data augmentation on the subvolumes to further increase the size of our dataset. Data augmentation using basic image manipulations such as flipping and rotation is a common practice in deep learning that has proven to be effective on datasets such as CIFAR-10 and ImageNet (Shorten & Khoshgoftaar, 2019). Here, we flip the constituent 2D image slices of each 3D subvolume horizontally and vertically to achieve 4 distinct subvolumes with the same permeability value. Then, we reverse the ordering of the consituent 2D image slices to achieve 4 more distinct subvolumes with the same permeability value (Figures 1 and 2). In total, data augmentation increases the size of the original dataset to include 460,000 distinct subvolumes with 57,500 unique permeability values.

The permeability values of our dataset range from 0 mD to over 20,000 mD, with a positively skewed distribution characterized by a long tail in the region corresponding to high permeabillity subvolumes (Figure 3a). While distributions of natural processes are typically skewed (e.g., precipitation, earthquake, etc.), such unbalanced datasets pose particular challenges for machine learning models. Specifically, machine learning models trained with highly unbalanced datasets tend to treat the tail region of the distribution as outliers and produce biased predictions that correspond to the most common values in the training data (Bauder & Khoshgoftaar, 2018; Johnson & Khoshgoftaar, 2019; Olson, 2004). As a result, it is very difficult for machine learning models to learn correctly from unbalanced datasets (Bauder et al., 2018; Bauder & Khoshgoftaar, 2018). Here, we employ the under-sampling technique to achieve a balanced dataset (Torgo et al., 2015; Fernández et al., 2018)—We first divide the augmented dataset into bins that each encompasses a permeability interval of 100 mD, then randomly sample ∼700 subvolumes from each bin. Finally, we arrive at an evenly-distributed dataset with permeability values that range from 200 to 9,600 mD (Figure 3b), where the upper limit corresponds to the point beyond which the number of subvolumes is not sufficient. We further characterize the balanced dataset by the porosity

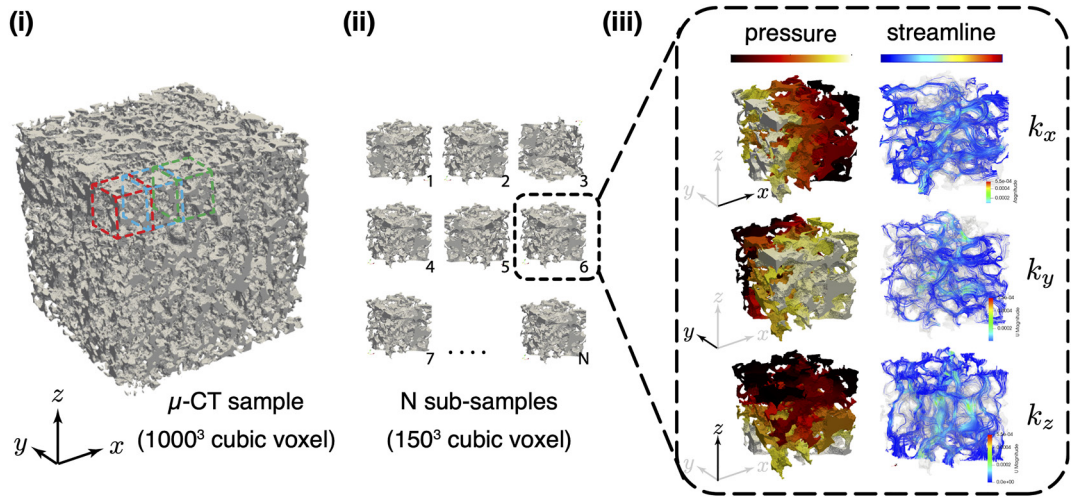

Figure 1. Dataset generation workflow. (a) We obtain three-dimensional micro computed tomography images of rock samples from the open source repository hosted by Imperial College London (Bijeljic & Raeini, 2015). The samples are rescaled such that each voxel corresponds to 3 μm in physical dimension. (b) We employ a sliding cube across the original sample to obtain subvolumes of size $150 \times 150 \times 150$ cubic voxels. (c) We numerically simulate water flow through each subvolume by prescribing pressure boundary conditions at the two bounding surfaces normal to one of the principal axes (i.e., x, y, z), while imposing no-flow boundary conditions at the other bounding surfaces. We integrate the velocities normal to the outflow bounding surface to obtain the total volumetric flux across the subvolume, and calculate its permeability using Darcy's law (Equation 1).

and permeability anisotropy of the subvolumes (Figures 3c and 3d). We measure permeability anisotropy by the anisotropy index I_a, which is a three-dimensional estimate of the deviation from isotropy (Clavaud et al., 2008).

$$k_{\text{iso}} = (k_{\min} k_{\text{int}} k_{\max})^{\frac{1}{3}}, \tag{2a}$$

$$I_a = \left[\frac{(k_{\min} - k_{\text{iso}})^2 + (k_{\text{int}} - k_{\text{iso}})^2 + (k_{\max} - k_{\text{iso}})^2}{k_{\min}^2 + k_{\text{int}}^2 + k_{\max}^2} \right]^{\frac{1}{2}}, \tag{2b}$$

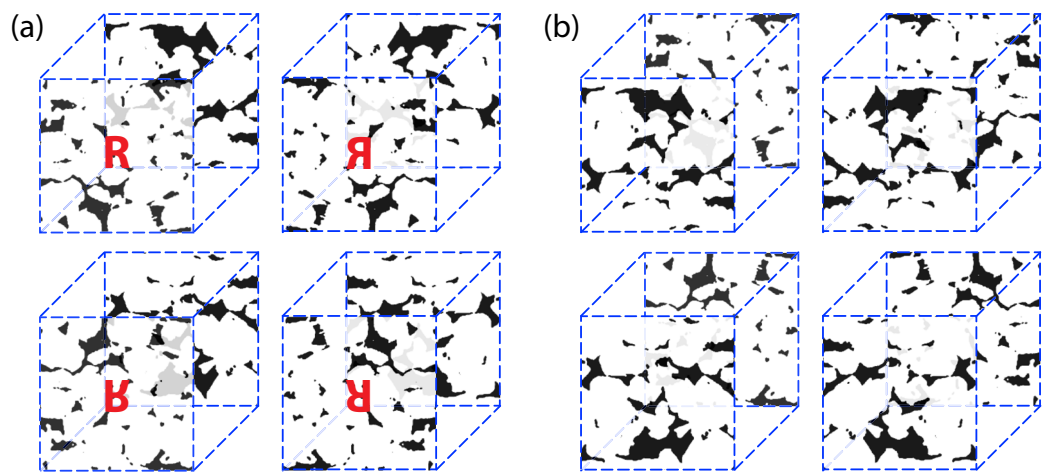

Figure 2. Data augmentation. (a) For each three-dimensional (3D)subvolume, we flip the constituent 2D image slices horizontally and vertically to achieve four distinct subvolumes with the same permeability value. The capital letter R on the face of the subvolume aids visualizing the flipping transformations. (b) For each 3D subvolume in (a), we reverse the order of the consituent 2D image slices to achieve four more distinct subvolumes with the same permeability value.

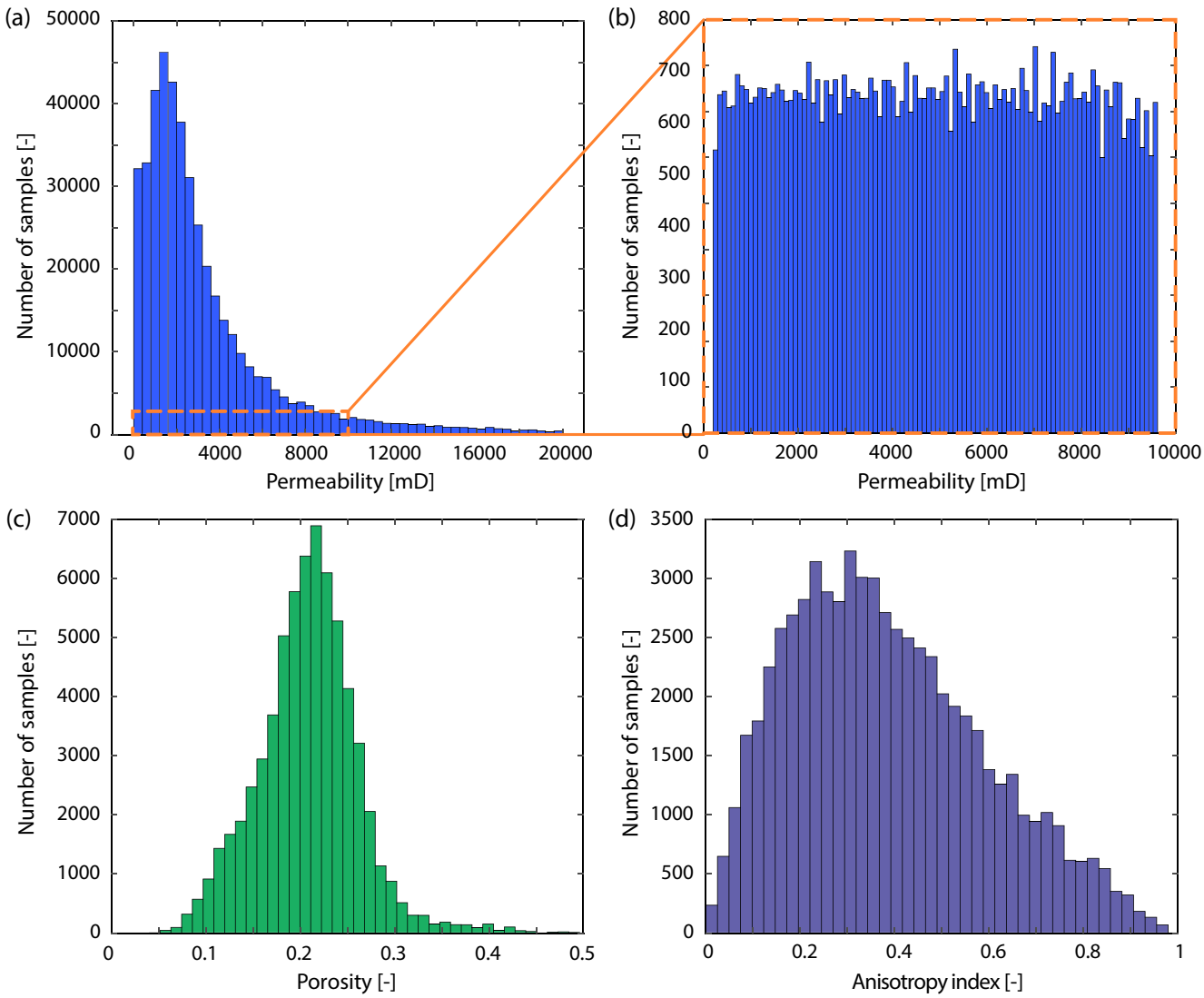

Figure 3. Dataset distributions. (a) Permeability distribution of the augmented dataset resembles a log-normal distribution, with a long tail in the region corresponding to high permeabillity values. (b) We randomly sample ~700 subvolumes for each 100 mD interval from the augmented dataset in the range between 200 and 9,600 mD to construct a balanced dataset. (c) The porosity of the balanced dataset ranges from 0.046 to 0.492. (d) The anisotropy index of the balanced dataset ranges from ~0 (i.e., isotropic) to ~1 (i.e., extremely anisotropic).

where k_{min}, k_{int} and k_{max} are the minimum, intermediate and maximum permeabilities, respectively, for a given subvolume. In the context of geologic porous media, rock samples with $I_a < 0.15$ are considered relatively isotropic, while rock samples with $I_a > 0.4$ are considered highly anisotropic (Clavaud et al., 2008). Our dataset covers a wide spectrum of I_a values, and it includes a significant portion of highly anisotropic subvolumes (Figure 3d). This is significant to ensure the generalizability of the model and that it is not limited to predict permeability of isotropic and homogeneous rock samples.

We randomly divide the balanced dataset of 65,250 subvolumes with labeled permeabilities into testing and training datasets, with a 20–80 split. The training dataset is further divided into validation and training subsets with the same split ratio during the network training process.

2.2. Deep Learning Model Architecture

Convolutional neural networks (CNNs) are a class of machine learning methods specifically designed to analyze visual imagery (Valueva et al., 2020). CNNs are widely used in applications including handwriting classification

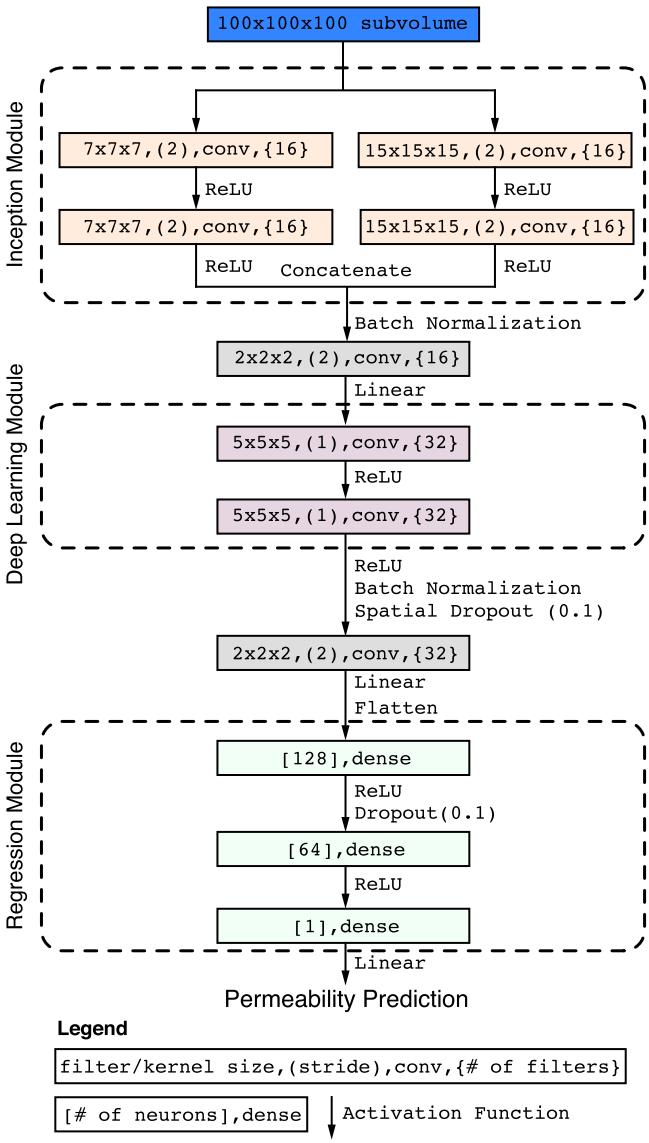

Figure 4. The architecture of our three-dimensional CNN model consists of an inception module with two parallel paths of convolutional layers of different kernel sizes, a deep learning module with two sequential convolutional layers, and a regression module with two dense fully-connected layers.

(Cireşan et al., 2010, 2011), facial recognition (Taigman et al., 2014), self-driving (Bojarski et al., 2016), material defect detection (Cha et al., 2017), and disease diagnosis (Esteva et al., 2017). The vast majority of existing CNNs are developed to perform classification tasks, whose output is a discrete set of categories. In contrast, the use of CNNs to perform regression tasks that predict continuous output variables is more challenging and still in its infancy. For instance, it was not until late 2020 when the first comprehensive analysis of deep regression techniques was published (Lathuiliere et al., 2020). In addition, while the application of CNNs on 2D images have become routine, their application on 3D volumetric images remain a considerable computational challenge (Maturana & Scherer, 2015). In our present work, training a 3D CNN with >50, 000 subvolumes of $150 \times 150 \times 150$ cubic voxels in size requires over 500 GB of memory, which exceeds the memory capacity of most existing GPU servers. To maintain the dataset size while ensuring computational feasibility, we downscale each subvolume to $100 \times 100 \times 100$ cubic voxels before feeding it to the machine learning model.

We develop a CNN model that takes 3D volumetric images as the input and predicts the permeability as the output. This approach is typically referred to as end-to-end learning, since it does not require the explicit calculation of hand-crafted geometric features (e.g., maximum inscribed sphere of the pore spaces) as an intermediate step (Santos et al., 2020). Traditional CNN models (e.g., LeNet, AlexNet) consist of sequentially stacked convolutional layers of fixed kernel size at each level (Lecun et al., 1998; Krizhevsky et al., 2012), and that a deeper network (i.e., more layers/levels) typically leads to more accurate predictions (Simonyan & Zisserman, 2015). However, very deep CNNs are computationally expensive and prone to overfitting (Li et al., 2019). Overfitting is characterized by much more accurate model predictions on the training set compared to the testing set, and it is a common problem in machine learning applications (Srivastava et al., 2014). Models with a relatively small training dataset and a large number of parameters can effectively "memorize" the training dataset (Ying, 2019). To minimize overfitting, we adopt the following regularization techniques in our CNN model architecture (Figure 4): *(a)* increase the dataset size via data augmentation (Shorten & Khoshgoftaar, 2019); *(b)* reduce the number of parameters by downsampling feature maps (Wu & Gu, 2015); *(c)* drop units and neuron connections randomly during training at a given probability (i.e., dropout rate; Srivastava et al., 2014); *(d)* stop the training process early, when the validation error stops decreasing (Ying, 2019). In addition to regularizing a neural network, adding non-linearity to it is equally important. Since convolution and pooling layers are linear operations, the output will always be a linear transformation of the input regardless of how deep the network is (Li et al., 2019). We apply a rectified linear unit (ReLU) as the activation function (Figure 4), which introduces non-linearity while minimizing added computational cost (Alzubaidi et al., 2021).

For the convolution layers, where a fixed kernel size is commonly used, it is often difficult to determine the optimal kernel size for images whose salient features have large variations in size (Sun et al., 2019). This challenge is particularly relevant to natural porous media, whose pore sizes vary greatly even in a millimeter scale sample. While large pores contain most of the pore volume of the sample, small pores exert fundamental control over the permeability. In addition, characteristic pore sizes vary greatly between different rock types. To effectively capture the disparate length scales present in natural porous media, our CNN model incorporates an inception module (Figure 4), which processes convolutional layers of different kernel sizes in parallel at the same level (Szegedy et al., 2015). Specifically, our inception module consists of two parallel paths of convolutional layers with kernel sizes of $7 \times 7 \times 7$ cubic voxels (small pore features) and $15 \times 15 \times 15$ cubic voxels (large pore

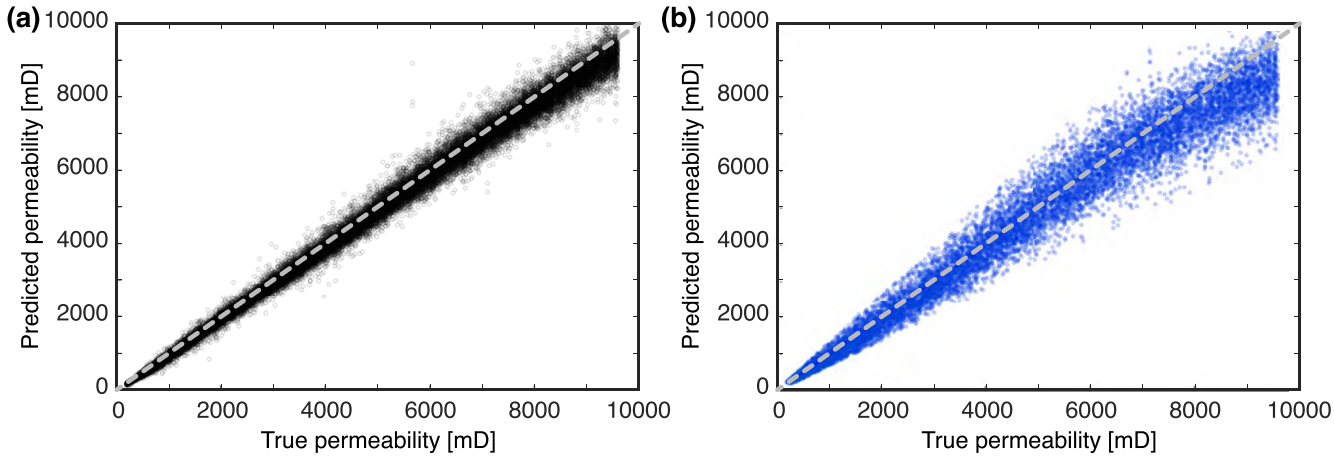

Figure 5. The permeabilities predicted by the convolutional neural network model closely match the true permeabilities obtained from direct numerical simulation for both (a) the training dataset and (b) the testing dataset. The dashed line represents perfect agreement between the predicted permeability and the true permeability.

features). The resulting feature maps of the two convolution layers from the inception module are concatenated, and passed through a convolution layer with a stride of 2 voxels. The purpose of this additional convolution layer is to reduce the number of network parameters, accelerate the training process and prevent over-fitting, and it is similar to the use of pooling layers (e.g., maximum pooling, average pooling) in most modern CNN models (Alzubaidi et al., 2021; Springenberg et al., 2014).

After the inception module, we apply another module of two sequential convolutional layers with kernel size of $5 \times 5 \times 5$ cubic voxels to extract even deeper features, and to double the number of output filters to 32 (Figure 4). We then add a spatial dropout layer with a dropout rate of 0.1 to prevent over-fitting during the training process. The final extracted feature map is compressed by a factor of 2 via a convolutional layer and passed to the regression module (Figure 4). The regression module consists of 2 dense layers of fully connected layers with 128 and 64 neurons each that process the extracted deep features with a dropout rate of 0.1. The final dense layer consists of only one neuron with a linear activation function that outputs the predicted permeability.

The entire CNN model contains a total of ~2.6 million tunable parameters. We train the model using the open-source software interface (Keras 2.4.0) and machine learning library (TensorFlow 2.3.1) on a computational cluster consisting of 7 graphics cards (NVIDIA GeForce RTX 2080 Ti GPU). The use of GPUs enables parallel computing, which significantly reduces the training time and allows the model to scale with additional resources (Owens et al., 2008).

The training process is conducted via the Adam optimizer (Kingma & Ba, 2015)—A computationally efficient extension to adaptive stochastic gradient descent. The mean absolute percentage error is used as the loss function, and all convolution layers are processed with padding and without added bias to the output values. We run the training process for 100 epochs and save the parameters yielded by the best-performing epoch.

3. Model Performance

We assess the predictive capability of the CNN model with the testing dataset, which consists of over 13,000 subvolumes of a variety of rock types (Table 1). The predictions of the CNN model closely match the true permeabilities obtained from direct numerical simulation over a wide range of permeability values (Figure 5). We quantify the accuracy of the model predictions by calculating the relative error $RE = (k_{pred} - k_{true})/k_{true}$ and the absolute relative error $ARE = |RE|$. To eliminate potential oversized impact of outliers, we report the accuracy of the model only for predictions whose ARE are within the 95th percentile. The CNN model produces balanced predictions over a wide range of permeability values ($RE \in [-0.32, 0.32]$), although it does tend to underpredict for permeabilities greater than 7,000 mD (Figures 5b and 6a). We believe the CNN model underestimates the permeability of highly-permeable samples as a result of the training dataset's finite permeability range (200–9,600 mD). In effect, the CNN model has never encountered samples with permeability greater than 9,600 mD. Therefore, the model "prefers" to err on the side of underprediction when it encounters a highly-permeable

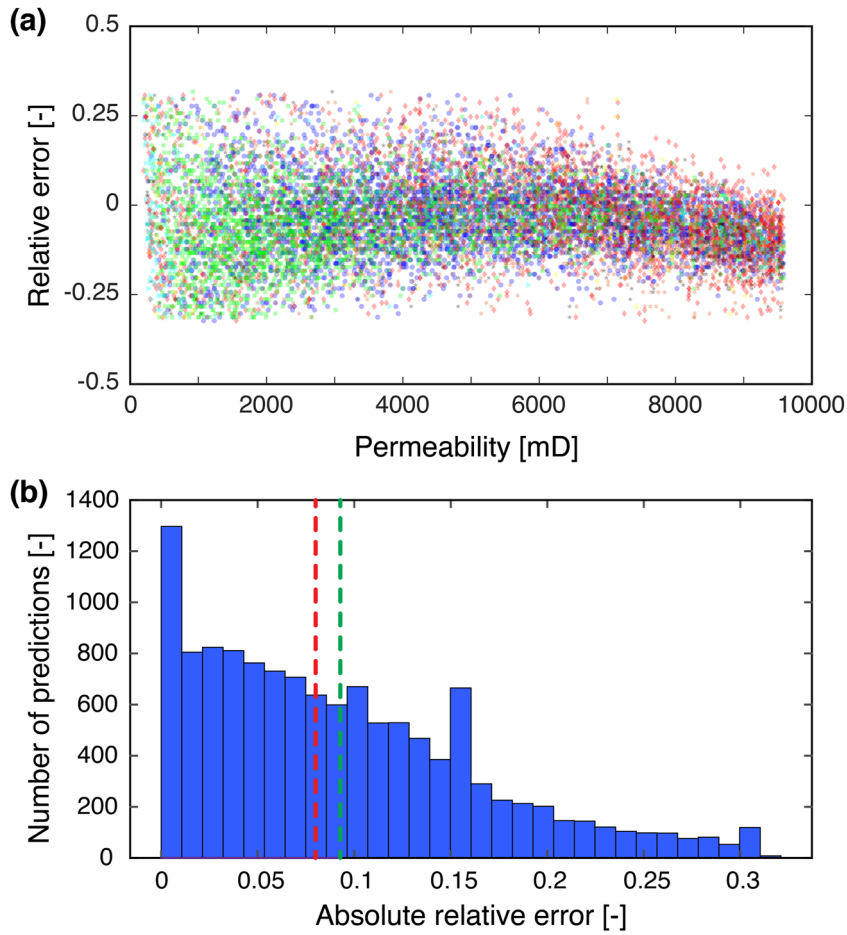

Figure 6. Prediction capability of the CNN model. (a) The relative error *RE* of the model predictions is bounded between −0.32 and 0.32 over a wide range of permeability values for all rock types: Bentheimer sandstone (blue circle), Ketton limestone (red diamond), Berea sandstone (green square), Doddington sandstone (organe hexagram), Estaillades limestone (black pentagram), Carbonate A (cyan right-pointing triangle), Carbonate B (yellow left-pointing triangle). (b) The distribution of the absolute relative error (ARE) shows that the vast majority of the model predictions are clustered in the low error range. The red dashed line and the green dashed line illustrate the median absolute relative error and the mean absolute relative error (MARE), respectively.

sample. We believe underprediction errors within this range of permeability values can be reduced by extending the permeability range of the training dataset.

Furthermore, the model predictions are equally accurate for all 7 rock types in the testing dataset (Figure 6a). The ARE distribution shows that the vast majority of the model predictions are clustered in the low error range, and the mean absolute relative error MARE = 0.093 (Figure 6b). We additionally calculate the coefficient of determination R^2, which measures how well the model predictions approximate the dataset—$R^2 = 1$ indicates perfect agreement between the model and the true permeability. It is a common statistical measure employed by other machine learning models for permeability prediction (Chung et al., 2019b; Hong & Liu, 2020; Kamrava et al., 2020; Tembely et al., 2021; Table 2). Our CNN model achieves $R^2 = 0.95$ on the testing dataset.

We attribute the accuracy of our model to *(a)* the diverse and large training dataset (52,200 subvolumes), *(b)* the novel architecture of the machine learning model, and *(c)* the use of convolutional layers in place of pooling layers in down-sampling operations. To investigate the impact of dataset size on the accuracy of model predictions, we train the same CNN model with a series of reduced datasets. The prediction error decreases monotonically as training dataset size increases, such that the model trained with 52,200 subvolumes (i.e., full training dataset) is 80% more accurate than the model trained with the subset containing 10,000 subvolumes (Figure 7a). In general, the amount of data required to adequately train a machine learning model increases with the number of trainable

Table 2
A Summary of Existing Three-Dimensional (3D) Convolutional Neural Network (CNN) Models for Permeability Prediction

	Method	Permeability range [mD]	Training set size	End-to-End	R^2	MARE
Current model	3D CNN	200–9,600	>50k	Yes	0.95	0.093
Hong and Liu (2020)	3D CNN	10–6,000	<4k	Yes	0.92	N/A
Kamrava et al. (2020)	3D CNN	100–500	>1k	Yes	0.91	N/A
Tembely et al. (2021)	3D CNN	70–400	<1k	No	0.87	N/A
Alqahtani et al. (2021)	3D CNN	10–1,800	<25k	Yes	0.86	0.189

Note. Since the models are tested on different datasets, their accuracy as quantified by the coefficient of determination (R^2) and the mean absolute relative error (MARE) are listed for reference only and not for direct comparison.

parameters and the complexity of the required prediction. In addition to the dataset size, the accuracy of our model is enhanced by the inclusion of an inception module consisting of two parallel paths of convolutional layers with different kernel sizes (Figure 4), which is more effective at extracting salient features of disparate length scales compared to sequential convolutional layers with a single kernel size. Specifically, we find that models trained with sequential convolutional layers with a single kernel size produce less accurate predictions compared to the model trained with the inception module, regardless of the value of the kernel size (Figure 7b). Finally, the accuracy of our model predictions is enhanced by the use of convolutional layers in down-sampling, instead of the more commonly-used approach of applying pooling layers. This so-called *all-convolutional-net* approach better preserves useful feature information in the neural network's receptive fields during down-sampling (Springenberg et al., 2014). Quantitatively, we find that the use of convolutional layers instead of maximum pooling layers in down-sampling contributes to ∼15% and ∼20% reduction in MARE and median ARE, respectively, while all other model configurations remain the same.

In addition to having low prediction errors within the testing dataset, generalizability is another important metric in assessing machine learning models. In this context, generalizability is defined as a model's ability to produce sensible predictions when applied on datasets that have not been used in the training process (LeCun et al., 2015). To characterize the generalizability of our model, we apply it to predict the permeability of a new sandstone core sample (Bijeljic & Raeini, 2015) previously unseen by the CNN model. The new sample has an average porosity of 0.21 with standard deviation of 0.03, an average anisotropy index of 0.38 with standard deviation of 0.18, and an average permeability of 4,158 mD with standard deviation of 2285 mD. Our model achieves excellent accuracy (MARE = 0.118, R^2 = 0.93) when it is used to predict the permeability of the 2,055 subvolumes of

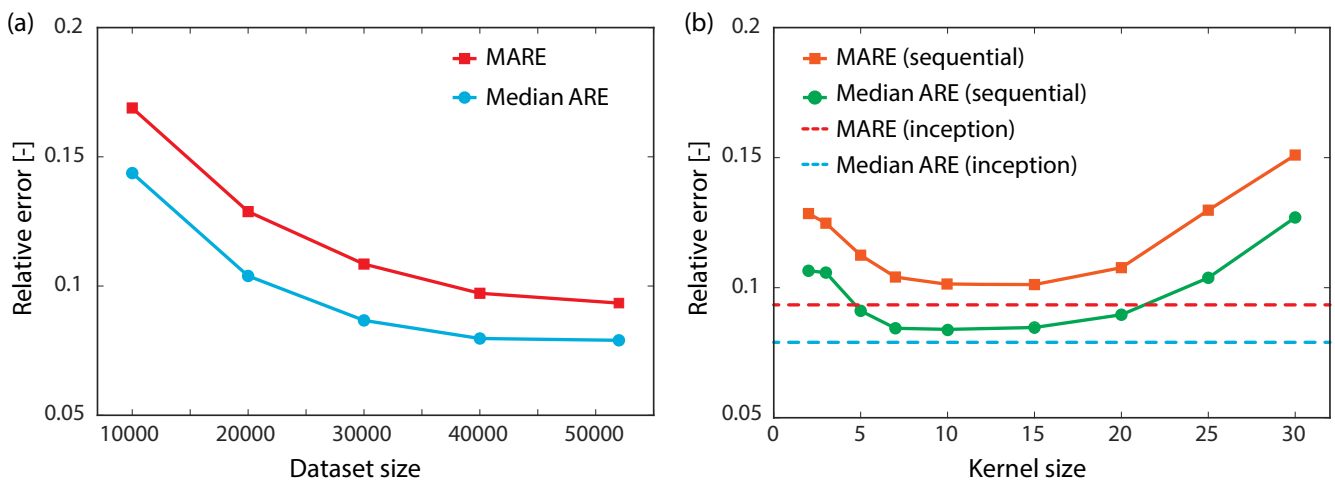

Figure 7. (a) The model prediction error decreases monotonically as the training dataset size increases. (b) The relative error of convolutional neural network models trained with sequential convolutional layers is a convex function of the kernel size. The lowest prediction error is achieved in the model with a kernel size of 10 voxels, though it is still higher than the prediction error produced by the model equipped with the inception module (dashed lines).

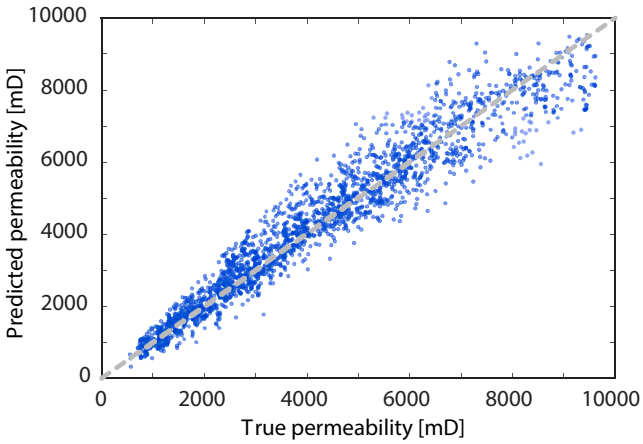

Figure 8. The three-dimensional convolutional neural network model achieves excellent accuracy when it is used to predict the permeabilities of previously unseen sandstone specimens. The dashed line represents perfect agreement between the predicted permeability and the true permeability obtained from direct numerical simulation.

the external dataset (Figure 8), which is similar to the prediction accuracy obtained from the original testing dataset (MARE = 0.093, R^2 = 0.95). We attribute the generalizability of the model to the large and diverse dataset and the data balancing procedure, which removes bias toward certain ranges of permeability values that might otherwise have a high level of occurrence.

The 3D CNN model performs permeability prediction in less than 5 milliseconds (i.e., ~200 samples/second) on a computational cluster with 7 GPUs. This timeframe is much faster than direct numerical simulation—Solving for the permeability along the three principal directions (i.e., x, y, z) of our training dataset numerically using OpenFoam® required over 6,000 simulation hours on a computational cluster of 32 × 3.20 GHz CPUs (i.e., ~10 samples/hour). We note that the CNN model's fast prediction time is built upon labeled datasets obtained from direct numerical simulations, and the model requires additional time for model architecture development and hyper-parameter optimization. Moreover, direct numerical simulations provide additional information about fluid flow (e.g., velocity field) that state-of-the-art machine learning models are still struggling to predict accurately at the pore-scale (Da Wang et al., 2020). However, the capability of providing accurate permeability predictions in real time justifies the significant computational resources used to train the 3D CNN model.

4. Conclusions

We have advanced the state-of-the-art of machine learning-driven digital rock permeability characterization. Specifically, we have assembled a diverse and massive dataset of over 57,500 unique subvolumes taken from CT scans of a variety of rock types. The true permeabilities of the subvolumes are obtained from direct numerical simulation via OpenFoam® and they range from 0 mD to over 20,000 mD (Figure 1). We perform data augmentation to further increase the dataset size, and data balancing to reduce the dataset bias (Figures 2 and 3).

Our machine learning model is a 3D CNN that enables end-to-end permeability prediction. The model adopts a novel network architecture that encompasses an inception module capable of feature extraction at multiple scales (Figure 4). The trained model achieves excellent accuracy (Figures 5 and 6) compared to other existing CNN models for permeability prediction, as signified by the model's low mean absolute relative error (MARE = 0.093) and high coefficient of determination (R^2 = 0.95). We demonstrate that the model's superior accuracy stems from the large training dataset size (Figure 7a), the implementation of an inception module in place of sequential convolutional layers for feature extraction (Figure 7b), and the use of convolutional layers instead of pooling layers in down-sampling operations. Furthermore, the model performs well when it is used to predict the permeability of previously unseen samples (Figure 8), which further demonstrates its generalizability. In addition to its accuracy, a key advantage of the model is its computational efficiency—Predicting the permeability of a single 100 × 100 × 100 cubic voxel sample takes less than 5 milliseconds on a computational cluster with 7 GPUs.

Subsurface characterization based on images of rock samples (i.e., digital rock analysis) has become an important tool in both academic and industrial applications in the past decade. Our work demonstrates the growing potential of applying machine learning models such as CNNs to achieve fast and accurate predictions in digital rock analysis. Increasing availability and variety of labeled datasets, more sophisticated network architecture (e.g., additional inception modules), and higher computing power can push the technology to new heights in the coming years.

Data Availability Statement

The data is publicly available on Zenodo, a general-purpose open-access repository (Elmorsy, 2022).

Acknowledgments
This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants and through the Canadian Nuclear Energy Infrastructure Resilience under Systemic Risk (CaNRisk) - Collaborative Research and Training Experience (CREATE). Additional support by the INTERFACE Institute and the INViSionLab, both of McMaster University, is gratefully acknowledged.

References

Alpak, F. O., Berg, S., & Zacharoudiou, I. (2018). Prediction of fluid topology and relative permeability in imbibition in sandstone rock by direct numerical simulation. *Advances in Water Resources*, *122*, 49–59. https://doi.org/10.1016/j.advwatres.2018.09.001

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*(04), 1800–1811. https://doi.org/10.2118/205376-pa

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., et al. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data*, *8*(1), 1–74. https://doi.org/10.1186/s40537-021-00444-8

Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., et al. (2013a). Digital rock physics benchmarks—Part I: Imaging and segmentation. *Computational Geosciences*, *50*, 25–32. https://doi.org/10.1016/j.cageo.2012.09.005

Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., et al. (2013b). Digital rock physics benchmarks—Part II: Computing effective properties. *Computational Geosciences*, *50*, 33–43. https://doi.org/10.1016/j.cageo.2012.09.008

Bauder, R. A., & Khoshgoftaar, T. M. (2018). The effects of varying class distribution on learner behavior for medicare fraud detection with imbalanced big data. *Health Information Science and Systems*, *6*(1). https://doi.org/10.1007/s13755-018-0051-3

Bauder, R. A., Khoshgoftaar, T. M., & Hasanin, T. (2018). An empirical study on class rarity in big data. In *2018 17th IEEE International Conference on Machine Learning and Applications* (pp. 785–790). https://doi.org/10.1109/icmla.2018.00125

Bear, J. (2013). *Dynamics of fluids in porous media*. Courier Corporation.

Bear, J., & Bachmat, Y. (1991). Introduction to modeling phenomena of transport in porous media. *Theory and Applications of Transport in Porous Media*.

Berg, C. F., Lopez, O., & Berland, H. (2017). Industrial applications of digital rock technology. *Journal of Petroleum Science and Engineering*, *157*, 131–147.

Bijeljic, B., Raeini, A., Mostaghimi, P., & Blunt, M. (2013). Predictions of non-Fickian solute transport in different classes of porous media using direct simulation on pore-scale images. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *87*(1). https://doi.org/10.1103/physreve.87.013011

Bijeljic, B., & Raeini, A. Q. (2015). *Micro-CT images and networks*. Retrieved from https://www.imperial.ac.uk/earth-science/research/research-groups/pore-scale-modelling/micro-ct-images-and-networks/

Blunt, M. J. (2001). Flow in porous media—Pore-network models and multiphase flow. *Current Opinion in Colloid & Interface Science*, *6*, 197–207. https://doi.org/10.1016/s1359-0294(01)00084-x

Blunt, M. J. (2017). *Multiphase flow in permeable media: A pore-scale perspective*. Cambridge University Press.

Blunt, M. J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., et al. (2013). Pore-scale imaging and modelling. *Advances in Water Resources*, *51*, 197–216. https://doi.org/10.1016/j.advwatres.2012.03.003

Boek, E. S., & Venturoli, M. (2010). Lattice-Boltzmann studies of fluid flow in porous media with realistic rock geometries. Computers & Mathematics With Applications, 59, 2305–2314. https://doi.org/10.1016/j.camwa.2009.08.063

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., & Goyal, P. (2016). End to end learning for self-driving cars. arXiv, arXiv:1604.07316.

Cha, Y.-J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, *32*(5), 361–378. https://doi.org/10.1111/mice.12263

Chan, J., & Schmitt, D. R. (2015). Initial seismic observations from a deep borehole drilled into the Canadian Shield in northeast Alberta. *International Journal of Earth Sciences*, *104*(6), 1549–1562. https://doi.org/10.1007/s00531-014-1110-x

Chen, S., & Doolen, G. D. (1998). Lattice Boltzmann method for fluid flows. *Annual Review of Fluid Mechanics*, *30*, 329–364. https://doi.org/10.1146/annurev.fluid.30.1.329

Chung, T., Wang, Y. D., Armstrong, R. T., & Mostaghimi, P. (2019a). Approximating permeability of microcomputed-tomography images using elliptic flow equations. *SPE Journal*, 1154–1163. https://doi.org/10.2118/191379-pa

Chung, T., Wang, Y. D., Armstrong, R. T., & Mostaghimi, P. (2019b). Approximating permeability of microcomputed-tomography images using elliptic flow equations. *SPE Journal*, *24*(03), 1154–1163. https://doi.org/10.2118/191379-PA

Cireşan, D. C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2010). Deep, big, simple neural nets for handwritten digit recognition. *Neural Computation*, *22*(12), 3207–3220.

Cireşan, D. C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2011). Convolutional neural network committees for handwritten character classification. In *2011 International Conference on Document Analysis and Recognition* (pp. 1135–1139). https://doi.org/10.1109/icdar.2011.229

Clavaud, J.-B., Maineult, A., Zamora, M., Rasolofosaon, P., & Schlitter, C. (2008). Permeability anisotropy and its relations with porous medium structure. *Journal of Geophysical Research*, *113*(B1). https://doi.org/10.1029/2007jb005004

Da Wang, Y., Chung, T., Armstrong, R. T., & Mostaghimi, P. (2020). ML-LBM: Machine learning aided flow simulation in porous media. arXiv.

Dong, H., & Blunt, M. J. (2009). Pore-network extraction from micro-computerized-tomography images. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *80*, 036307. https://doi.org/10.1103/physreve.80.036307

Elmorsy, M. (2022). *elmorsym1/permeability-prediction-via-3d-cnn*. https://doi.org/10.5281/zenodo.6276081

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Fernández, A., Garcia, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence*, *61*, 863–905. https://doi.org/10.1613/jair.1.11192

Freeze, R. A., & Cherry, J. A. (1979). *Groundwater*. Prentice-Hall.

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(18), 4613–4618. https://doi.org/10.1073/pnas.1716999115

Graczyk, K. M., & Matyka, M. (2020). Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-78415-x

Guibert, R., Horgue, P., Debenest, G., & Quintard, M. (2016). A comparison of various methods for the numerical evaluation of porous media permeability tensors from pore-scale geometry. *Mathematical Geosciences*, *48*(3), 329–347. https://doi.org/10.1007/s11004-015-9587-9

Hong, J., & Liu, J. (2020). Rapid estimation of permeability from digital rock using 3D convolutional neural network. *Computational Geosciences*, *24*(4), 1523–1539. https://doi.org/10.1007/s10596-020-09941-w

Horgue, P., Soulaine, C., Fanc, J., Guibert, R., & Debenest, G. (2015). An open-source toolbox for multiphase flow in porous media. *Computer Physics Communications*, *187*, 217–226. https://doi.org/10.1016/j.cpc.2014.10.005

Hossain, M. E. (2015). Drilling costs estimation for hydrocarbon wells. *Journal of Sustainable Energy Engineering*, *3*(1), 3–32. https://doi.org/10.7569/jsee.2014.629520

Huang, J., Zhou, W., Li, H., & Li, W. (2019). Attention-based 3D-CNNs for large-vocabulary sign language recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, *29*(9), 2822–2832. https://doi.org/10.1109/tcsvt.2018.2870740

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0192-5

Juanes, R., MacMinn, C. W., & Szulczewski, M. L. (2010). The footprint of the CO_2 plume during carbon dioxide storage in saline aquifers: Storage efficiency for capillary trapping at the basin scale. *Transport in Porous Media*, *82*, 19–30. https://doi.org/10.1007/s11242-009-9420-3

Kamrava, S., Tahmasebi, P., & Sahimi, M. (2020). Linking morphology of porous media to their macroscopic permeability by deep learning. *Transport in Porous Media*, *131*(2), 427–448. https://doi.org/10.1007/s11242-019-01352-5

Ketcham, R. A., & Carlson, W. D. (2001). Acquisition, optimization and interpretation of X-ray computed tomographic imagery: Applications to the geosciences. *Computational Geosciences*, *27*(4), 381–400. https://doi.org/10.1016/s0098-3004(00)00116-3

Khan, A., Sohail, A., Zahoora, U., & Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, *53*(8), 5455–5516. https://doi.org/10.1007/s10462-020-09825-6

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. arXiv preprint.

Korez, R., Likar, B., Pernuš, F., & Vrtovec, T. (2016). Model-based segmentation of vertebral bodies from MR images with 3D CNNs. In *Transactions on Computational Science XXXVIII* (pp. 433–441). Springer. https://doi.org/10.1007/978-3-319-46723-8_50

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*, 1097–1105.

Lake, L. W. (1989). *Enhanced oil recovery*. Prentice-Hall.

Lathuiliere, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2020). Analysis of deep regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(9), 2065–2081. https://doi.org/10.1109/tpami.2019.2910523

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *121*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Li, H., Li, J., Guan, X., Liang, B., Lai, Y., & Luo, X. (2019). Research on overfitting of deep learning. In *2019 15th International Conference on Computational Intelligence and Security* (pp. 78–81). CIS. https://doi.org/10.1109/cis.2019.00025

Manwart, C., Aaltosalmi, U., Koponen, A., Hilfer, R., & Timonen, J. (2002). Lattice-Boltzmann and finite-difference simulations for the permeability for three-dimensional porous media. *Physical Review E–Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *66*(1), 1–32. https://doi.org/10.1103/physreve.66.016702

Maturana, D., & Scherer, S. (2015). VoxNet: A 3D convolutional neural network for real-time object recognition. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 922–928). https://doi.org/10.1109/iros.2015.7353481

Milletari, F., Navab, N., & Ahmadi, S. A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D vision 3DV* (pp. 565–571). https://doi.org/10.1109/3dv.2016.79

Mostaghimi, P., Blunt, M. J., & Bijeljic, B. (2013). Computations of absolute permeability on micro-CT images. *Mathematical Geosciences*, *45*(1), 103–125. https://doi.org/10.1007/s11004-012-9431-4

Muljadi, B. P., Blunt, M. J., Raeini, A. Q., & Bijeljic, B. (2016). The impact of porous media heterogeneity on non-Darcy flow behaviour from pore-scale simulation. *Advances in Water Resources*, *95*, 329–340. https://doi.org/10.1016/j.advwatres.2015.05.019

Olson, D. L. (2004). Data set balancing. In *Chinese Academy of Sciences Symposium on Data Mining and Knowledge Management* (pp. 71–80).

Orr, F. M., Jr, & Taber, J. J. (1984). Use of carbon dioxide in enhanced oil recovery. *Science*, *224*(4649), 563–569. https://doi.org/10.1126/science.224.4649.563

Owens, J. D., Houston, M., Luebke, D., Green, S., Stone, J. E., & Phillips, J. C. (2008). GPU computing. *Proceedings of the IEEE*, *96*(5), 879–899. https://doi.org/10.1109/JPROC.2008.917757

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. https://doi.org/10.1016/j.advwatres.2020.103539

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 1–48. https://doi.org/10.1186/s40537-019-0197-0

Simonyan, K., & Zisserman, A. (2015). *Very deep convolutional networks for large-scale image recognition*. arXiv.

Spanne, P., Thover, J. F., Jacquin, C. J., Lindquist, W. B., Jones, K. W., & Adler, P. M. (1994). Synchrotron computed microtomography of porous media: Topology and transports. *Physical Review Letters*, *73*(14), 2001–2004. https://doi.org/10.1103/physrevlett.73.2001

Springenberg, J. T., Dosovitskiy, A., Brox, T., & Riedmiller, M. (2014). *Striving for simplicity: The all convolutional net*. arXiv.

Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

Sudakov, O., Burnaev, E., & Koroteev, D. (2019). Driving digital rock towards machine learning: Predicting permeability with gradient boosting and deep neural networks. *Computational Geosciences*, *127*, 91–98. https://doi.org/10.1016/j.cageo.2019.02.002

Sun, G., Huang, H., Zhang, A., Li, F., Zhao, H., & Fu, H. (2019). Fusion of multiscale convolutional neural networks for building extraction in very high-resolution images. *Remote Sensing*, *11*(3), 227. https://doi.org/10.3390/rs11030227

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. In *Proceedings IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit* (pp. 1–9). https://doi.org/10.1109/cvpr.2015.7298594

Szulczewski, M. L., MacMinn, C. W., Herzog, H. J., & Juanes, R. (2012). Lifetime of carbon capture and storage as a climate-change mitigation technology. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(14), 5185–5189. https://doi.org/10.1073/pnas.1115347109

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Computer Society Conference Computer Vision Pattern Recognition* (pp. 1701–1708). https://doi.org/10.1109/cvpr.2014.220

Tarbuck, E. J., Lutgens, F. K., Tasa, D., & Tasa, D. (2005). *Earth: An introduction to physical geology*. Prentice-Hall.

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

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 Report*, *7*, 1460–1472. https://doi.org/10.1016/j.egyr.2021.02.065

Tian, J., Qi, C., Sun, Y., Yaseen, Z. M., & Pham, B. T. (2020). Permeability prediction of porous media using a combination of computational fluid dynamics and hybrid machine learning methods. *Engineering Computations*. https://doi.org/10.1007/s00366-020-01012-z

Torgo, L., Branco, P., Ribeiro, R. P., & Pfahringer, B. (2015). Resampling strategies for regression. *Expert Systems*, *32*(3), 465–476. https://doi.org/10.1111/exsy.12081

Valueva, M., Nagornov, N., Lyakhov, P., Valuev, G., & Chervyakov, N. (2020). Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and Computers in Simulation*, *177*, 232–243. https://doi.org/10.1016/j.matcom.2020.04.031

Wang, Y. D., Blunt, M. J., Armstrong, R. T., & Mostaghimi, P. (2021). Deep learning in pore scale imaging and modeling. *Earth-Science Reviews*, *215*, 103555. https://doi.org/10.1016/j.earscirev.2021.103555

Wang, Y. D., Chung, T., Armstrong, R. T., McClure, J. E., & Mostaghimi, P. (2019). Computations of permeability of large rock images by dual grid domain decomposition. *Advances in Water Resources*, *126*, 1–14. https://doi.org/10.1016/j.advwatres.2019.02.002

Wu, H., & Gu, X. (2015). Max-pooling dropout for regularization of convolutional neural networks. In *International Conference on Neural Information Processing* (pp. 46–54). https://doi.org/10.1007/978-3-319-26532-2_6

Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, *1168*, 022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Zhao, T., Zhao, H., Ning, Z., Li, X., & Wang, Q. (2018). Permeability prediction of numerical reconstructed multiscale tight porous media using the representative elementary volume scale lattice Boltzmann method. *International Journal of Heat and Mass Transfer*, *118*, 368–377. https://doi.org/10.1016/j.ijheatmasstransfer.2017.11.004