


# Porous Structure Reconstruction Using Convolutional Neural Networks

Yuzhu Wang<sup>1</sup>  · Christoph H. Arns<sup>1</sup> ·  
Sheik S. Rahman<sup>1</sup> · Ji-Youn Arns<sup>2</sup>

Received: 18 November 2017 / Accepted: 1 May 2018  
© International Association for Mathematical Geosciences 2018

**Abstract** The three-dimensional high-resolution imaging of rock samples is the basis for pore-scale characterization of reservoirs. Micro X-ray computed tomography ( $\mu$ -CT) is considered the most direct means of obtaining the three-dimensional inner structure of porous media without deconstruction. The micrometer resolution of  $\mu$ -CT, however, limits its application in the detection of small structures such as nanochannels, which are critical for fluid transportation. An effective strategy for solving this problem is applying numerical reconstruction methods to improve the resolution of the  $\mu$ -CT images. In this paper, a convolutional neural network reconstruction method is introduced to reconstruct high-resolution porous structures based on low-resolution  $\mu$ -CT images and high-resolution scanning electron microscope (SEM) images. The proposed method involves four steps. First, a three-dimensional low-resolution tomographic image of a rock sample is obtained by  $\mu$ -CT scanning. Next, one or more sections in the rock sample are selected for scanning by SEM to obtain high-resolution two-dimensional images. The high-resolution segmented SEM images and their corresponding low-resolution  $\mu$ -CT slices are then applied to train a convolutional neural network (CNN) model. Finally, the trained CNN model is used to reconstruct the entire low-resolution three-dimensional  $\mu$ -CT image. Because the SEM images are segmented and have a higher resolution than the  $\mu$ -CT image, this algorithm integrates the super-resolution and segmentation processes. The input data are low-resolution  $\mu$ -CT images, and the output data are high-resolution segmented porous structures. The

---

✉ Yuzhu Wang  
yuzhu.wang@unsw.edu.au

<sup>1</sup> School of Petroleum Engineering, University of New South Wales, Sydney 2052, Australia

<sup>2</sup> Research School of Physics and Engineering, Australian National University, Canberra 0200, Australia

experimental results show that the proposed method can achieve state-of-the-art performance.

**Keywords** Convolutional neural network · Porous structure reconstruction ·  $\mu$ -CT

## 1 Introduction

Three-dimensional pore-scale numerical modeling of reservoirs dates back to 1984, when Quiblier generated a three-dimensional porous structure based on the measurements of the characteristics using two-dimensional thin sections of porous media (Quiblier 1984). This method involves successively passing a normalized uncorrelated Gaussian random field through a linear and then a nonlinear filter to generate the discrete values that represent the phases of the structure (Yeong and Torquato 1998a). This filtering method is quite general (Hiroshi 2004) and robust (Adler et al. 1990), and is universally used to model various porous media from food to rock samples (Adler et al. 1992; Frederic and Dominique 2004; Liang et al. 1998; Roberts 1997; Roberts and Torquato 1999; Spanne et al. 1994). However, the use of only one- and two-point probability functions has been found to be insufficient for reproducing more complex porous structures (Yeong and Torquato 1998a), because the Gaussian filtering reconstruction methods retain a similar degree of isotropy and homogeneity as the original training image but do not reproduce the connectivity properties, which are critical for transport (Biswal and Hilfer 1999). Additional morphological descriptors containing some connected information were then introduced to import more constraints into the reconstruction process, such as a two-point cluster function, lineal-path length and chord-length function. Because it is difficult to incorporate these new descriptors into the aforementioned filtering method, the simulated annealing (SA) method was proposed to reconstruct the porous media subject to various statistical constraints (Pant et al. 2015; Rintoul and Torquato 1997; Yeong and Torquato 1998a, b).

The SA algorithm allows arbitrary structural descriptors to be imported into the reconstruction procedure as constraints, which improve the accuracy of the modeling. Although this method has been used successfully for the reconstruction of several materials, including sandstone (Manwart et al. 2000) and chalk (Talukdar et al. 2002), the resulting images do not always capture the long-range connectivity of the pore space, specifically for media exhibiting low porosity or small-scale connectivity (i.e., because of high pore-throat aspect ratios or the presence of fractures) (Čapek et al. 2009). This is because it is more difficult for these low-porosity media to obtain a global minimum, since it requires the stepwise construction of small channels. The tedious computation of the SA modeling is an additional obstacle to its practical application.

Another novel method for modeling porous structures is the process-based reconstruction method, which numerically mimics the geological process of the formation of sedimentary rocks (Bryant and Blunt 1992; Coelho et al. 1997; Latham et al. 2001; Øren and Bakke 2002). However, a precondition for using this method is an adequate understanding of the physical and chemical processes that are essential to the formation of the porous structure, which is always difficult (Lucia et al. 2003), especially in carbonates and shale.

In 2004, a pattern recognition approach that uses a small template was developed by Wu et al. (2004). This approach is an effective method for generating a wide variety of pore structures in two dimensions by replicating patterns in a five-point stencil using a Markov chain Monte Carlo (MCMC) method. The authors subsequently extended this algorithm to the modeling of three-dimensional porous structures by using three perpendicular two-dimensional sections as training images. These three sections are treated as priors to determine the transition probabilities that control the Markov chain process (Wu et al. 2006). The MCMC algorithm was soon generalized to the multiple-point statistics (MPS) method (Okabe and Blunt 2004; Tahmasebi and Sahimi 2016a, b). In a later study, Okabe and Blunt (2007) expanded upon their work to enable the recovery of small features that are lost in micro X-ray computed tomography ( $\mu$ -CT) images, by combining two-dimensional scanning electron microscopy (SEM) images with three-dimensional low-resolution  $\mu$ -CT images. In this approach, macro-pores are identified by  $\mu$ -CT, while micro-pores are generated by the MPS reconstruction method based on an SEM image. The two images are then superimposed to obtain the final image, which has improved image quality. This hybrid strategy has been extended to multiresolution and multiscale modeling to address the problems associated with the computational cost and multiscale feature reproduction (Tahmasebi et al. 2015, 2016, 2017). In this method, a low-resolution image is applied to reconstruct the large-scale features in a shale sample, and high-resolution two-dimensional images are used to reconstruct the nanoscale structures. The multiscale pore network is then obtained by overlaying the micro and macro three-dimensional models. In some cases, the macro-scale reconstruction can be skipped if a three-dimensional macro image can be obtained directly. The MPS method can be further extended to a neighbor embedding reconstruction method to improve the resolution of the  $\mu$ -CT image, where low-frequency information is provided by the  $\mu$ -CT image itself, while high-frequency information is supplemented by high-resolution SEM image(s) (Wang et al. 2018). In theory, it is possible to perfectly reproduce the porous structure if the template size is sufficiently large and the training image can provide a sufficient number of replicates for the actual data events. However, due to limitations in computational resources including computing time, random-access memory (RAM), and thus training image size, the challenge of reproducing the long-range connectivity still exists in practice.

The deep learning algorithms in wide use today offer us a promising choice for improving the resolution of the porous structure extracted from  $\mu$ -CT images (Mosser et al. 2017). The convolutional neural network (CNN) is one of the most notable deep learning approaches, in which multiple layers are trained in a robust manner (LeCun et al. 1998). It has been found to be highly effective and is also the most commonly used approach in diverse computer vision applications (Guo et al. 2016). In addition, a large number of studies have investigated the CNN model for image processing, including image denoising (Jain and Seung 2008; Xuejiao et al. 2015; Zhang et al. 2017), image restoration (Cheong and Park 2017; Deepak and Ghanekar 2017; Dong et al. 2016; Jia et al. 2017; Liu et al. 2016; Samuel et al. 2015; Yamanaka et al. 2017) and image segmentation (Liu et al. 2015; Long et al. 2015). In this paper, a convolutional neural network reconstruction (CNNR) method that combines super-resolution and image segmentation is used to generate high-resolution segmented images based on low-

resolution tomographic  $\mu$ -CT images and high-resolution two-dimensional sections such as SEM image(s).

The remainder of this work is organized as follows. In Sect. 2, the architecture of a typical CNN model and the training process are briefly expounded. In Sect. 3, a  $\mu$ -CT image of Indiana limestone is used as an example to illustrate the CNNR method, after which the reconstructed image is evaluated in two respects, quantitative assessment and morphological measurements, in Sect. 4. In terms of the morphological measurements, local porosity theory and Minkowski functionals are applied to compare the performance of the CNNR method with that of the popular MPS method. Finally, the conclusions are given in Sect. 5.

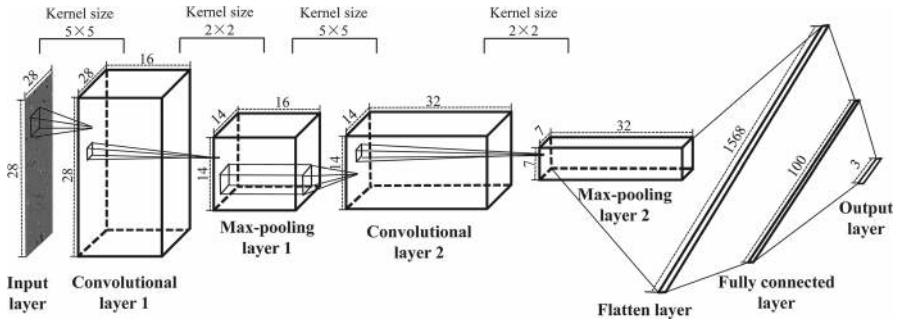
## 2 Related Work

### 2.1 Architecture of the CNN Model

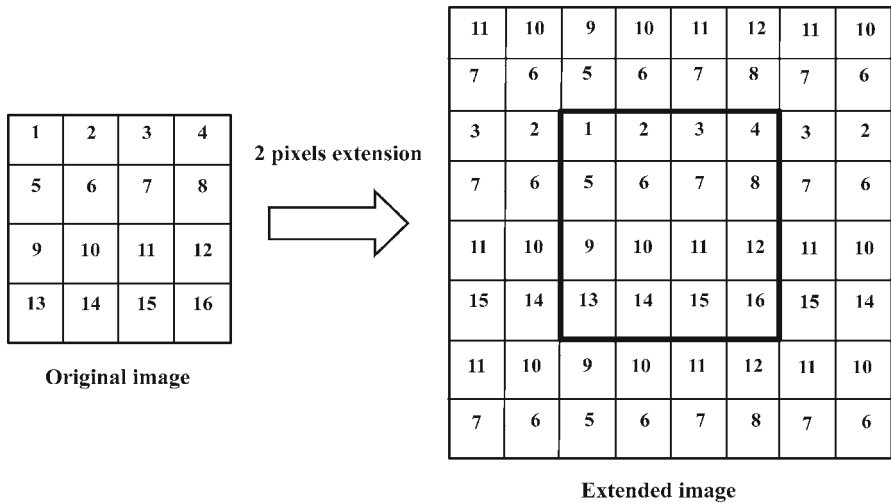
Convolutional neural networks can be dated back to 1989 (LeCun et al. 1989), but were largely neglected for many years, their recent rise in popularity chiefly attributable to two factors: (1) the efficient training implementation on modern powerful GPUs (Krizhevsky et al. 2017) and (2) the easy access to an abundance of data for training large models (Deng et al. 2009). Similar to the conventional neural network model, a CNN model also has one input layer and one output layer, as well as a number of hidden layers. In a typical CNN model, there are three main types of layers: convolutional layer, pooling layer and fully connected layer (Guo et al. 2016). In some cases, a flatten layer is needed between the fully connected layer and the convolutional or pooling layer to reshape the two-dimensional or three-dimensional neural nodes to be one-dimensional. A scheme was recently proposed in which pooling is replaced by convolution with a stride of 2 to improve the training stability (Chaoben and Shesheng 2018; Springenberg et al. 2014). All of the layers in a CNN model will be stacked to establish the convolutional neural network architecture. In general, most deep learning applications use a feedforward neural network architecture, which learns to map a fixed-size input (e.g., an image) to a fixed-size output (e.g., a probability for each of several categories). To move from one layer to the next, a set of units computes a weighted sum of their inputs from the previous layer and passes the result through a nonlinear function (LeCun et al. 2015). Figure 1 illustrates the CNN model used in this paper for high-resolution porous structure reconstruction (see Sect. 3).

The CNN model presented in Fig. 1 consists of an input layer, two convolutional layers, two pooling layers, one fully connected layer and an output layer, which are described in greater details as follows:

1. A  $28 \times 28 \times 1$  voxel grayscale image is inputted into the model. In this study, the width and height of the input image are 28, and the color channel is 1.
2. Next, 16 filters  $5 \times 5$  voxels in size convolve with the input layer to generate 16 feature maps ( $28 \times 28$  pixels) of the input layer. These feature maps extract features of the input image under different frequency domains. The elements of these 16



**Fig. 1** A CNN model with two convolutional layers, two max-pooling layers and one fully connected layer



**Fig. 2** Schematic graph of symmetric extension

filters are called “weights” in the CNN model. For example, the value of the  $(i, j)$  voxel of the  $m$ th feature map is denoted by  $f_m(i, j)$ , which is given by

$$f_m(i, j) = R((I * W_m)(i, j) + b_m), \tag{1}$$

where  $R(*)$  is an activation function such as ReLU or Sigmoid (LeCun et al. 2015),  $I$  is the input image,  $*$  is the convolution operation,  $W_m$  is the  $m$ th filter and  $b_m$  is bias. Note that a symmetric padding method (see Fig. 2) is used to guarantee that the output of the convolution has the same size as the input image.

3. The first convolutional layer is followed by a pooling layer that will perform a downsampling operation along the spatial dimensions (width and height) and result in a volume of  $14 \times 14 \times 16$  voxels. The most common conventional pooling methods are max-pooling and average-pooling (Guo et al. 2016); in addition, stochastic pooling (Zeiler and Fergus 2013), spatial pyramid pooling (Kaiming

et al. 2015) and def-pooling (Ouyang et al. 2015) have been proposed for solving various problems. In this paper, the max-pooling strategy is applied.

4. The second convolutional process is then conducted based on max-pooling layer 1. In this step, 32 filters  $5 \times 5$  voxels in size are used to convolve with max-pooling layer 1. The value of the  $(i, j)$  voxel of the  $m$ th feature map is denoted as  $f_m(i, j)$  and can be calculated by

$$f_m(i, j) = R \left( \sum_{t=1}^T (P_t * W_m)(i, j) + b_m \right), \quad (2)$$

where  $R(*)$  is the activation function,  $P_t$  is the  $t$ th image of max-pooling layer 1,  $*$  is the convolution operation,  $W_m$  is the  $m$ th filter and  $b_m$  is the bias.

5. Next, a pooling operation is undertaken based on convolutional layer 2. The kernel size is set to  $2 \times 2$ , and the max-pooling method is selected.
6. Because max-pooling layer 2 is three-dimensional, a flatten layer is required to convert the three-dimensional neural nodes into one dimension. Through the flattening step, fully connected layers can be used after some convolutional layers or pooling layers.
7. A fully connected layer with 100 elements then converts the two-dimensional feature maps into a one-dimensional feature vector. Fully connected layers perform like a traditional neural network and contain the main part of the parameters in the CNN model (Guo et al. 2016). Fully connected layers can be either fed forward into certain number categories for classification or used as a feature vector for follow-up processing.
8. In Fig. 1, the fully connected layer is immediately followed by an output layer with three categories.

## 2.2 Training

The entire network is composed of neurons that have learnable weights and biases. Learning the end-to-end mapping function requires an estimation of the network parameters  $\theta$ , which can be described as

$$\theta = \{W_1, W_2, \dots, W_n, B_1, B_2, \dots, B_n\}, \quad (3)$$

where  $W_i (i=1, 2, \dots, n)$  denotes the convolutional kernel of the  $i$ th layer (e.g., in Fig. 1, the size of the kernel of the first convolutional layer  $W_1$  is  $5 \times 5 \times 16$ ).  $B_i (i=1, 2, \dots, n)$  is the bias of the  $i$ th convolutional layer (e.g., in Fig. 1, the size of  $B_1$  is 16). The parameter  $n$  defines the depth of the CNN model.

The training process is conducted via minimizing the loss between the estimated result and the corresponding labels. In this paper, cross entropy is used as the loss function

$$L(\theta) = (-1/S) \sum_{i=1}^S [E_i \ln O_i + (1 - E_i) \ln(1 - O_i)], \quad (4)$$

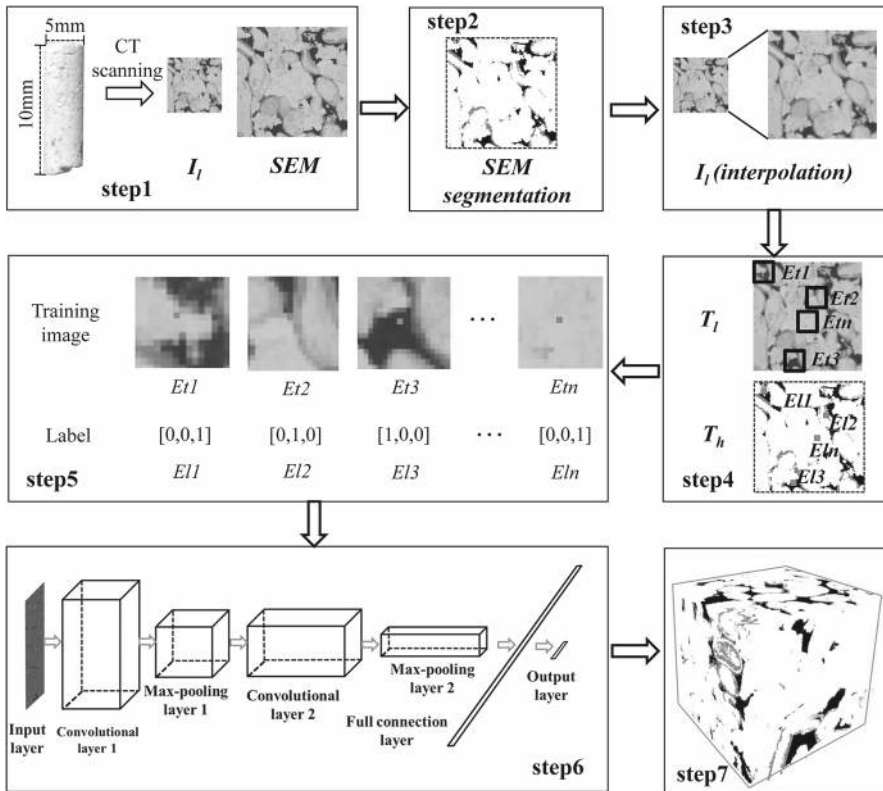
where  $S$  is the number of training images,  $E_i$  is the label of the  $i$ th training image, and  $O_i$  is the predicted value. The loss is then minimized using stochastic gradient descent with standard backpropagation (LeCun et al. 1998). Specifically, the weight matrices are updated as follows:

$$\Delta_{i+1} = 0.9 \cdot \Delta_i + \alpha \cdot \left( \partial L / \partial W_i^l \right), W_{i+1}^l = W_i^l + \Delta_{i+1}, \quad (5)$$

where  $l \in \{1, 2, 3, \dots, n\}$  is the index of the layers, and  $i$  is iteration index. Parameter  $\alpha$  is the learning rate, and  $\partial L / \partial W_i^l$  is the derivative. The filter weights of each layer are initialized by drawing randomly from a Gaussian distribution with a zero mean and a standard deviation of 0.001. Additionally, all of the biases are initialized as 0. The learning rate is set to 0.001.

### 3 Convolutional Neural Network Reconstruction

In this section, an Indiana limestone sample is applied as an example to explain the proposed CNNR method (see Fig. 3). First, a 5-mm-diameter and 10-mm-long Indiana limestone sample is scanned by  $\mu$ -CT under a resolution of  $4.05 \times 4.05 \times 4.05 \mu\text{m}^3/\text{voxel}$  and  $16.2 \times 16.2 \times 16.2 \mu\text{m}^3/\text{voxel}$ , respectively. Note that the high-resolution  $\mu$ -CT image is scanned to evaluate the performance of the reconstruction. In practice, one or more high-resolution SEM images are scanned for training the model instead of scanning the entire  $\mu$ -CT sample in high resolution. The  $\mu$ -CT imaging and the subsequent registration work are completed by the Tyree X-ray CT Lab of the University of New South Wales. Two image subsets, a  $1.62 \times 1.62 \times 1.62 \text{ mm}^3$  low-resolution subset and its corresponding high-resolution subset, are then extracted from the registered low- and high-resolution images and are denoted by  $I_l$  ( $100 \times 100 \times 100$  voxels) and  $I_h$  ( $400 \times 400 \times 400$  voxels). Second, the high-resolution subset  $I_h$  is segmented into three phases: micro-pore, porous matrix and solid. The segmented  $I_h$  is denoted as  $IS_h$ . Third, the low-resolution subset  $I_l$  is interpolated to the  $400 \times 400 \times 400$  voxel size using a cubic spline interpolation algorithm. Although the voxel size of this interpolated low-resolution image is identical to that of the high-resolution image, it is still referred to as “low-resolution” image  $I_l$  for ease of presentation. Fourth, the first slice of  $I_l$  and  $IS_h$  is selected and treated as the training image for the following training process of the CNN model. These low- and high-resolution training image pairs are denoted by  $T_l$  and  $T_h$ , respectively. In the fifth step, the training dataset is extracted from  $T_l$  and  $T_h$ . Every element of the training dataset contains two parts: the training image and the label. For example, if a  $28 \times 28$  voxel template is used to extract the training images from  $T_l$ , the image patch  $(i - 13:i + 14, j - 13:j + 14)$  in  $T_l$  and the value of pixel  $(i, j)$  in  $T_h$  consists of the training image and its label, respectively. In step 5 of Fig. 3, vectors  $[1, 0, 0]$ ,  $[0, 1, 0]$  and  $[0, 0, 1]$  represent the labels of the pore space, porous matrix and solid phases, respectively. To eliminate the influence of the boundary,  $i$  and  $j$  must satisfy the condition that  $14 \leq i \leq 386$  and  $14 \leq j \leq 386$ . In this study, 100,000 elements are prepared for the CNN training. In the sixth step, the CNN parameters are trained based on the prepared training dataset. Lastly, an empty grid system  $G$  with a size of  $400 \times 400 \times 400$  voxels



**Fig. 3** Schematic graph of the convolutional neural network reconstruction of the porous structure of rock samples

is established. To any uncertain voxel  $(i, j, k)$  in  $G$ , its corresponding training image is the image patch  $V(i:i+27, j:j+27, k)$  in  $I_1$ . The estimation of  $G(i, j, k)$  is calculated by the trained CNN model in the last step. This process is repeated until the whole  $G$  is estimated. Considering the impact of the boundary, low-resolution image  $I_1$  is symmetrically extended 13 and 14 voxels in the head and tail along the  $i$  direction, and then symmetrically extended 13 and 14 voxels in the head and tail along the  $j$  direction.

Note that the reconstruction process is conducted layer by layer within a two-dimensional plane (e.g.,  $i-j$  plane), and every layer is reconstructed separately (see Fig. 4). These reconstructed two-dimensional layers are then stacked to build a three-dimensional high-resolution segmented image. Although the reconstruction is undertaken within a two-dimensional plane ( $i-j$  plane), there is no discontinuity noise in the  $i-k$  and  $j-k$  planes in the reconstructed image (see Sect. 4). This finding occurs mainly because the interpolation process smooths the initial low-resolution  $\mu$ -CT image in the  $i, j$  and  $k$  direction.

Figure 5 illustrates the procedure for reconstructing a single layer of the grid system  $G$ . Before reconstruction, the target low-resolution  $\mu$ -CT image is interpolated from



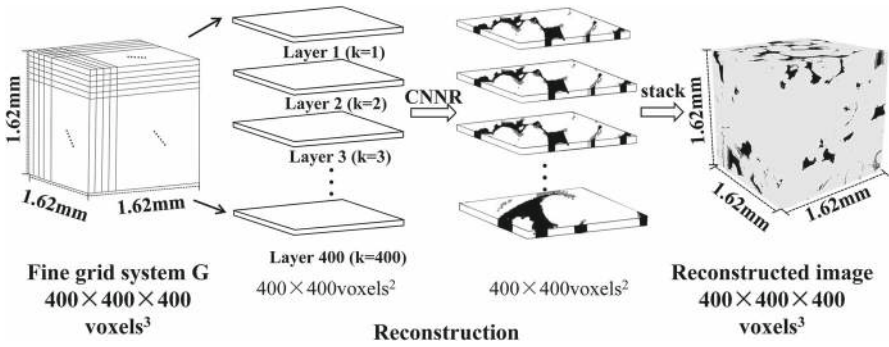


Fig. 4 Schematic graph of two-dimensional to three-dimensional reconstruction process

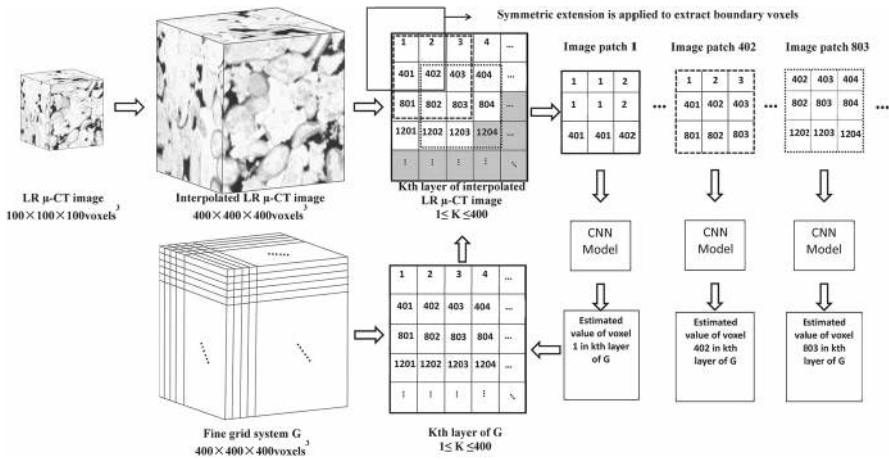


Fig. 5 Schematic graph of the reconstruction of a single layer

100 × 100 × 100 voxels to 400 × 400 × 400 voxels. This interpolated low-resolution tomographic image is used as a constraint during the subsequent reconstruction work. Note that the interpolated low-resolution μ-CT image has the same voxel size as that of the fine grid system, which means that every empty layer in  $G$  has a corresponding layer in the interpolated tomographic image. In Fig. 5, the reconstruction of the  $k$ th layer of  $G$  is used as an example to demonstrate the reconstruction. In both the  $k$ th layer of  $G$  and the interpolated μ-CT image, the voxels are labeled from 1 to 160,000 according to their location. For the first voxel in  $G$ , an image patch 3 × 3 voxels in size is extracted from the  $k$ th layer of the interpolated μ-CT image, centered at the first voxel. This image patch is used as the input data for the CNN model, and the output value is the estimated phase of the first voxel in the  $k$ th layer of  $G$ . Note that a symmetric extension (see Fig. 2) is applied for the boundary voxels to extract an image patch. This process is undertaken voxel by voxel using a maximally overlapping moving window scheme, until the whole  $k$ th layer of  $G$  is estimated.

## 4 Results

Figure 6 depicts three slices that were extracted from a low-resolution tomographic image (LR\_TOMO), low-resolution segmented image (LR\_SEG), high-resolution segmented image (HR\_SEG) and high-resolution CNNR reconstructed image (HR\_CNNR) of Indiana limestone with three phases. These three phases are pore, solid and an intermediate phase, which represents the sub-resolution porosity that is typical for oolitic limestone. The three slices are extracted from the  $i$ - $j$  plane (top row),  $i$ - $k$  plan (middle row) and  $j$ - $k$  plane (bottom row), respectively. Because of the insufficient resolution, the LR\_TOMO image lost part of its high-frequency information, and its segmented image (LR\_SEG) lost some pore connections, which greatly simplified the pore geometry features. In terms of the visual sensitivity, two conclusions can be obtained from Fig. 6: (1) the CNNR reconstructed structure improved the connectivity of the pore space, and (2) the reconstructed image successfully reproduced the contact relationships among three phases. Here, “contact relationship” refers to the spatial configuration of the different phases. The porous matrix occurs as an outer shell for some special grains, but not all grains, which means that it is necessary to identify which grain has a porous matrix shell when segmenting the tomographic images. In this study, this problem is solved by setting a gradient threshold to determine whether a grain has a porous matrix shell. The existing porous matrix increases the complexity of the segmentation work. However, the proposed CNNR method reproduces this feature very well.

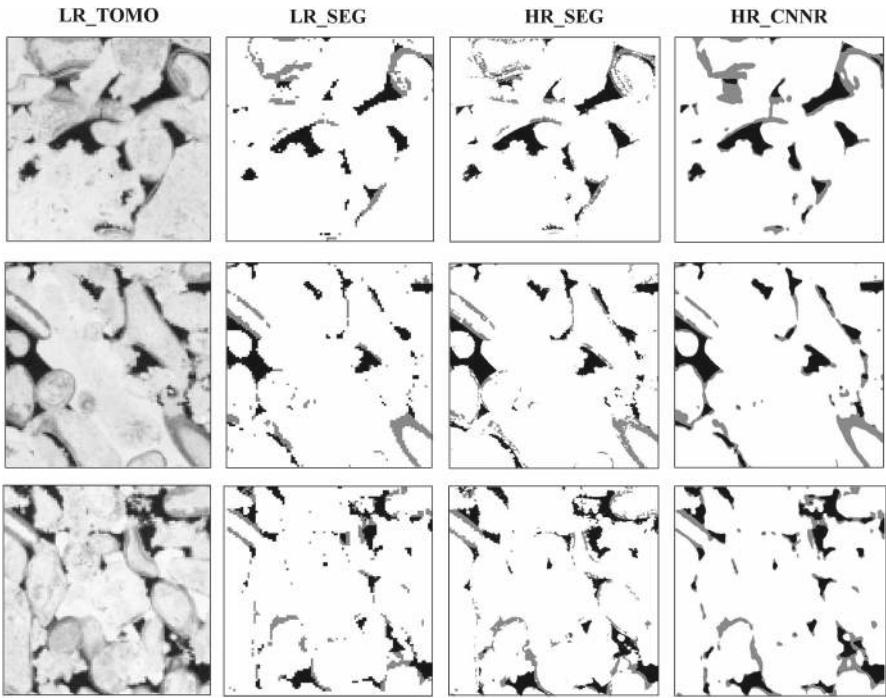
From Fig. 6, it is obvious that the reconstructed image is “smoother” than the segmented images. This arrangement occurs because the CNN model provides the probability of every phase for every voxel, and the phase with the largest probability is selected. A promising approach for solving this problem is to apply Monte Carlo sampling according to the probability of every phase instead of selecting the phase with the largest probability directly. However, this process will result in some unexpected noise. Thus, further study is needed to explore ways to reduce the smoothness without increasing the noise.

Except for the visual sensitivity analysis discussed earlier, this study also estimates the performance of the reconstruction work based on two aspects: (1) quantitative assessment via computing the Hamming distance between the CNNR-generated image and the high-resolution reference image, and (2) comparing the morphological measurements of the CNNR model with that of the low-resolution image and high-resolution reference image, as well as models derived by MPS.

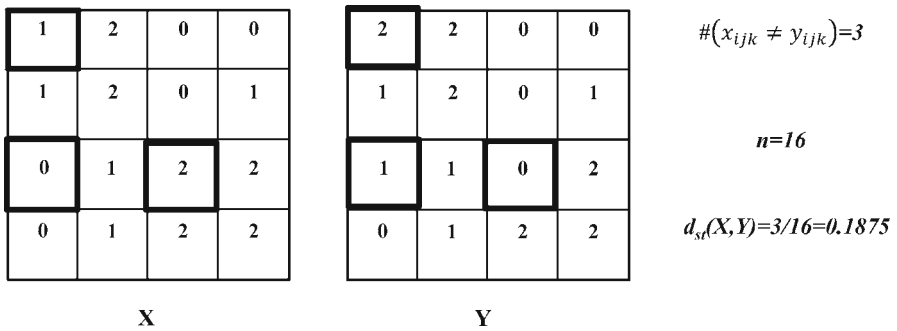
### 4.1 Quantitative Assessment

As discussed earlier, the first slice of the high-resolution segmented image is used as a training image, and another part is preserved as reference images for validating the performance of the CNNR method. In this paper, the Hamming distance  $d_{st}$  is chosen as the similarity measurement and is given by

$$d_{st} = (\#(x_{ijk} \neq y_{ijk})) / n, \quad (6)$$

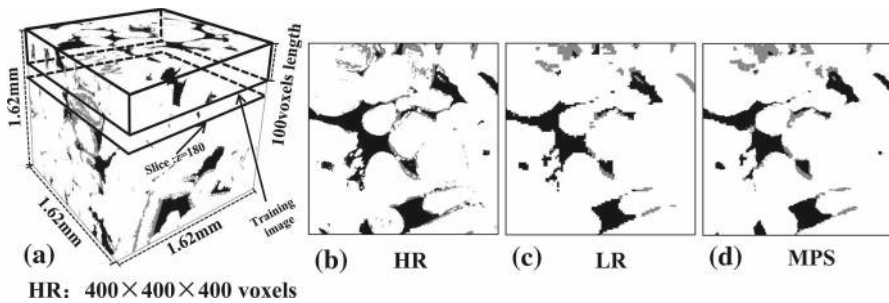


**Fig. 6** Three slices of the HR\_CNNR image, LR\_TOMO image, LR\_SEG image and HR\_SEG image



**Fig. 7** Calculation of the Hamming distance

where  $x_{ijk}$  and  $y_{ijk}$  are the elements that are located at voxel  $ijk$  of datasets  $X$  and  $Y$ , respectively, and  $i, j, k$  denote the grid indices.  $\#(x_{ijk} \neq y_{ijk})$  is the number of non-identical elements in  $X$  and  $Y$  (see Fig. 7). The Hamming distance thus measures the fraction of the non-identical elements between two datasets; a Hamming distance of 0 shows that the two datasets are identical, and a Hamming distance of 0.1 shows that ten percent of the elements of the two datasets are different. The Hamming distance between the CNNR reconstructed image and the high-resolution reference segmented image is 0.0523.



**Fig. 8** a) The HR image, with the top quarter of the HR selected as the training image; b–d three slices ( $z = 180$ ) extracted from HR, LR and MPS, respectively

## 4.2 Morphological Measurements

The main goal with the improved  $\mu$ -CT image resolution is to characterize the morphological features of the porous structure. Compared to quantitative measurement of the accuracy of the reconstructed image, its morphological measurements are more concerned with being practical. In this section, the authors compare the morphological measurements of the CNNR reconstructed image with those of the high- and low-resolution segmented images, and also compare the performance of the proposed method with the MPS method, which is currently the most popular applied reconstruction algorithm.

### 4.2.1 MPS Reconstruction

Because the MPS method is applied as a reference method to evaluate the proposed CNNR method, and because the performance of the MPS method is highly dependent on the implementation process, it is necessary to briefly introduce the MPS reconstruction procedure. As Fig. 8 shows, the top quarter of the high-resolution image is used as a training image to extract multiple-point statistics, and these statistics are then used to reconstruct the whole low-resolution image. The detailed reconstruction process is based on Comunian et al. (2012), with a  $4 \times 4 \times 4$  template size and three multigrid levels. Figure 8 also illustrates a slice with  $z = 180$  of the high-resolution (HR), low-resolution (LR) and MPS reconstructed image (MPS), respectively (see Fig. 8b–d). The MPS reconstructed image appears to slightly improve the connectivity of the porous structures compared with the low-resolution image, for which there is quantitative proof via morphological measurements (local porosity theory and Minkowski functionals), which is shown in the subsequent section.

### 4.2.2 Local Porosity Theory

Local porosity theory (LPT) is universally used to characterize porosity and connectivity fluctuations at different length scales in three-dimensional digitized models (Hilfer 1992; Latief et al. 2010). Here, the basic definitions of the quantities of LPT are briefly introduced.

*Local Porosity Distribution* A measurement cell  $\mathbf{K}(\mathbf{r}, L)$  denotes a cube that is centered at the lattice location  $\mathbf{r}$  with length  $L$ . Within the cubic measurement cell  $\mathbf{K}(\mathbf{r}, L)$ , the local porosity  $\varnothing(\mathbf{r}, L)$  is defined as

$$\varnothing(\mathbf{r}, L) = \frac{V[\mathbf{P} \cap \mathbf{K}(\mathbf{r}, L)]}{V[\mathbf{K}(\mathbf{r}, L)]}, \tag{7}$$

where  $V(\mathbf{K}(\mathbf{r}, L))$  denotes the volume of cell  $\mathbf{K} \in \mathbf{R}^3$ , and  $V(\mathbf{P} \cap \mathbf{K}(\mathbf{r}, L))$  is the pore space of  $\mathbf{K}(\mathbf{r}, L)$ . The local porosity distribution  $\mu(\varnothing, L)$  is given by

$$\mu(\varnothing, L) = \frac{1}{N} \sum_{\mathbf{r}} \delta[\varnothing - \varnothing(\mathbf{r}, L)], \tag{8}$$

where  $N$  is the number of cells, and  $\delta(\varnothing - \varnothing(\mathbf{r}, L))$  is the Dirac delta function.

*Local Percolation Probabilities* Local percolation probabilities characterize the connectivity of the pore space, which controls the transport and propagation in porous media. The connectivity function is defined as

$$\Lambda_c(\mathbf{r}, L) = \begin{cases} 1 & \text{if there is path through } \mathbf{K}(\mathbf{r}, L) \text{ in } c \text{ direction} \\ 0 & \text{otherwise} \end{cases}, \tag{9}$$

where  $c$  denotes the direction of the measurement, including the  $x, y$  and  $z$  directions. Additionally,  $\Lambda_3$  indicates that the cell is percolating in all directions, while  $\Lambda_\alpha$  indicates percolation in the  $x, y$  or  $z$  direction. The local percolation probabilities  $\lambda_c(\varnothing, L)$  are then given by

$$\lambda_c(\varnothing, L) = \frac{\sum_{\mathbf{r}} \Lambda_c(\mathbf{r}, L) \delta_{\varnothing\varnothing(\mathbf{r}, L)}}{\sum_{\mathbf{r}} \delta_{\varnothing\varnothing(\mathbf{r}, L)}}, \tag{10}$$

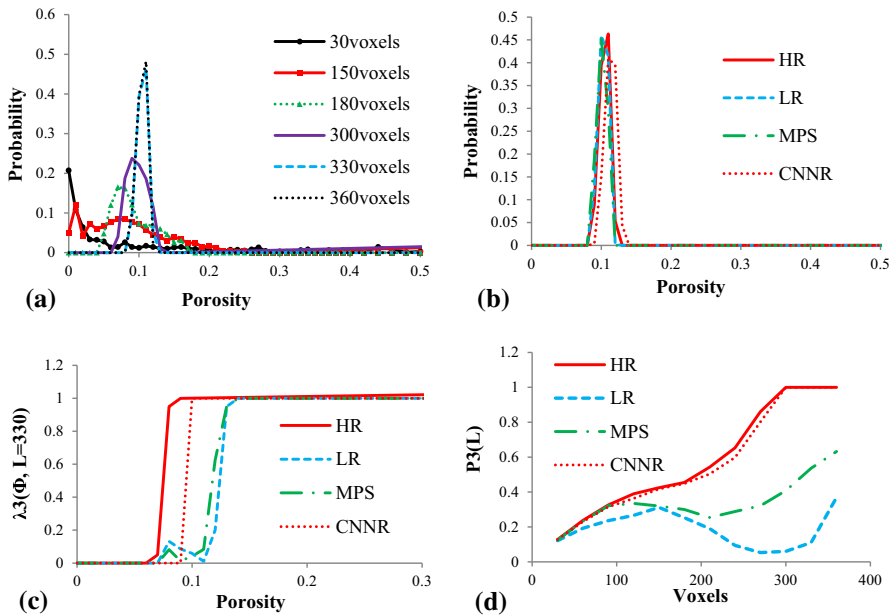
where parameter  $\lambda_c(\varnothing, L)$  describes the fraction of the analyzed cells with side length  $L$  and porosity  $\varnothing$  that are percolating in the  $c$  direction.  $\delta_{\varnothing\varnothing(\mathbf{r}, L)}$  is the Kronecker delta and can be calculated by

$$\delta_{\varnothing\varnothing(\mathbf{r}, L)} = \begin{cases} 1 & \text{if } \varnothing = \varnothing(\mathbf{r}, L) \\ 0 & \text{otherwise} \end{cases}. \tag{11}$$

*Total Fraction of Percolating Cells* Total fractional of percolating cells  $P_c(L)$  describes the percolating probability of measurement cells with side-length  $L$  in the  $c$  direction. It is given by

$$P_c(L) = \int_0^1 \mu(\varnothing, L) \lambda_c(\varnothing, L) d\varnothing. \tag{12}$$

Figure 9 illustrates the local porosity distribution curve, the local percolation probabilities curve, and the total fraction of the percolation cells curve of the porous structure



**Fig. 9** Local porosity theory of the CNNR image and its reference images

extracted from the high-resolution image (HR), low-resolution image (LR), CNNR reconstructed image and MPS reconstructed image. The original low-resolution segmented image ( $100 \times 100 \times 100$  voxels) is interpolated to a  $400 \times 400 \times 400$  voxel size via nearest-neighbor interpolation. Thus, all of the images have a voxel size of  $400 \times 400 \times 400$  voxels. Figure 9a presents the local porosity distribution of HR with different cube sizes. Along with an increase in the cube size, the distribution curve of the local porosity approaches stability, and the minimum stable size is approximately  $330 \times 330 \times 330$  voxels. Thus, the cube size of  $330 \times 330 \times 330$  voxels is applied to compare the various LPT parameters of different images. The local porosity distribution curves of four images are very close to each other (see Fig. 9b). However, in terms of the local percolation probabilities distribution (see Fig. 9c) and the total fraction of the percolation cells curve (see Fig. 9d), the performance of the CNNR method is significantly better than that of MPS, and both the CNNR and MPS methods improve the connectivity of the porous structure significantly.

#### 4.2.3 Minkowski Functionals

Minkowski functionals are used as basic integral geometric measures to quantify the porous structure, and are known to be sensitive descriptors of the morphology (Arns et al. 2004, 2010). After the segmentation, the images to be analyzed are segmented into multiple phases. Integral geometry provides a complete set of  $d+1$  additive Minkowski functionals for each phase, where  $d$  is the dimension of the embedding space. For ambiguous configurations on the cubic lattice, this study used 8 and 26 neighborhoods

for the mean and total curvature, respectively, using the algorithm given in Arns et al. (2001), and reports the Minkowski functionals normalized to the volume of the considered domain  $V_t$ . The first functional  $M_0$  is simply the total fraction of the target phase, which is given by

$$m_{0X} = M_0(X) / V_t = V(X) / V_t, \tag{13}$$

where  $X \subset \Omega$  ( $\Omega$  is the embedding space) is the space occupied by the target phase (e.g., the pore space in the binary image). The other Minkowski functionals are defined through integrals over the surfaces of the pores, denoted as  $\delta X$ , which unambiguously defines the shape or morphology of the pore structure at the given resolution. It can be described by

$$m_{1X} = M_1(X) / V_t = \frac{1}{6} \int_{\delta X} ds / V_t, \tag{14}$$

where  $ds$  is a surface element. The second integral measures the mean curvature of the interface

$$m_{2X} = M_2(X) / V_t = \frac{1}{3\pi} \int_{\delta X} \left( \frac{1}{r_1} + \frac{1}{r_2} \right) ds / V_t, \tag{15}$$

where  $r_1$  and  $r_2$  are the minimum and maximum radius of curvature for the surface element  $ds$ . This radius is positive for convex curvatures and negative for concave curvatures. The third integral measures the total curvature

$$m_{3X} = M_3(X) / V_t = \int_{\delta X} \left( \frac{1}{r_1 r_2} \right) ds / V_t, \tag{16}$$

which is related to the connectivity of the considered phase. For well-connected phases and few isolated components, this measure is typically negative, and it crosses zero to become positive close to the percolation threshold of the material.

Table 1 illustrates the Minkowski functionals of the high-resolution, low-resolution, CNNR reconstructed and MPS reconstructed images of Indiana limestone. In general, both reconstruction methods (CNNR and MPS) improve the Minkowski functionals, except for the volume fraction of the porous matrix ( $m_{0i}$ ) and the mean curvature of the pore space ( $m_{2p}$ ). Compared to the MPS method, CNNR performs better in terms of the volume fraction of the pore space ( $m_{0p}$ ) and solid space ( $m_{0s}$ ), and the surface density of the pore space ( $m_{1p}$ ), porous matrix ( $m_{1i}$ ) and solid ( $m_{1s}$ ). However, in curvature-related measurements ( $m_{2x}$  and  $m_{3x}$ ), the MPS method exhibits better performance than the CNNR method. The reason is that the image reconstructed using the CNNR method is “too smooth,” and the curvature-related measurements are extremely sensitive to high-frequency information (noise and edges (Arns et al. 2010)). As mentioned previously, further study is still needed to determine a reasonable means of reducing the smoothness of the CNNR reconstructed image without increasing the noise.

**Table 1** Minkowski functionals of Indiana limestone

Method	HR	CNNR	MPS	LR
$m_{op}$	0.0707	0.0686	0.0542	0.0479
$m_{oi}$	0.0389	0.0614	0.0356	0.0365
$m_{os}$	0.8895	0.870	0.9103	0.9156
$m_{1p}$ ( $\mu\text{m}^{-1}$ )	0.0075	0.005	0.0047	0.0034
$m_{1i}$ ( $\mu\text{m}^{-1}$ )	0.0119	0.0119	0.0076	0.0041
$m_{1s}$ ( $\mu\text{m}^{-1}$ )	0.0134	0.0089	0.0079	0.0063
$m_{2p}$ ( $\mu\text{m}^{-2}$ )	2.84E-05	5.93E-05	5.45E-05	3.04E-05
$m_{2i}$ ( $\mu\text{m}^{-2}$ )	4.50E-04	1.74E-04	2.86E-04	1.12E-04
$m_{2s}$ ( $\mu\text{m}^{-2}$ )	- 2.00E-04	- 6.74E-05	- 1.05E-04	- 7.63E-05
$m_{3p}$ ( $\mu\text{m}^{-3}$ )	1.36E-06	1.17E-08	5.61E-08	- 1.68E-08
$m_{3i}$ ( $\mu\text{m}^{-3}$ )	- 1.66E-06	- 8.88E-07	- 1.38E-06	- 5.91E-07
$m_{3s}$ ( $\mu\text{m}^{-3}$ )	1.16E-07	- 3.11E-08	- 2.16E-08	- 4.00E-08

The reference phases are denoted as “*p*” (pore space), “*i*” (porous matrix) and “*s*” (solid)

## 5 Conclusions

In this paper, a convolutional neural network reconstruction method is introduced to improve the resolution of the porous structure extracted from  $\mu$ -CT images of rock samples. A slice of a low-resolution tomographic  $\mu$ -CT image and its corresponding high-resolution segmented two-dimensional section (e.g., SEM image) are used to train the neural network model. The trained model is then applied to estimate the whole low-resolution  $\mu$ -CT image to reconstruct a high-resolution porous structure. The CNNR method integrates the super-resolution and segmentation processes. Its application to Indiana limestone samples shows that the proposed method greatly improves the connectivity of the pores and exhibits better performance than conventional methods, including MPS. In addition, the reconstructed high-resolution porous structure successfully reproduces the features of the contact relationships among the three phases.

## References

- Adler PM, Jacquin CG, Quiblier JA (1990) Flow in simulated porous media. *Int J Multiph Flow* 16:691–712. [https://doi.org/10.1016/0301-9322\(90\)90025-E](https://doi.org/10.1016/0301-9322(90)90025-E)
- Adler PM, Jacquin CG, Thovert JF (1992) The formation factor of reconstructed porous media. *Water Resour Res* 28:1571–1576. <https://doi.org/10.1029/92WR00059>
- Arns CH, Knackstedt MA, Pinczewski WV, Mecke KR (2001) Euler-Poincaré characteristics of classes of disordered media. *Phys Rev E* 63:031112
- Arns CH, Knackstedt MA, Mecke KR (2004) Characterisation of irregular spatial structures by parallel sets and integral geometric measures. *Colloids Surf A Physicochem Eng Aspects* 241:351–372. <https://doi.org/10.1016/j.colsurfa.2004.04.034>
- Arns CH, Knackstedt MA, Mecke K (2010) 3D structural analysis: sensitivity of Minkowski functionals. *J Microsc* 240:181–196. <https://doi.org/10.1111/j.1365-2818.2010.03395.x>



- Biswal B, Hilfer R (1999) Microstructure analysis of reconstructed porous media. *Phys A Stat Mech Appl* 266:307–311
- Bryant S, Blunt M (1992) Prediction of relative permeability in simple porous media. *Phys Rev A* 46:2004–2011
- Čapek P, Hejtmánek V, Brabec L, Zikánová A, Kočířík M (2009) Stochastic reconstruction of particulate media using simulated annealing: improving pore connectivity. *Transp Porous Media* 76:179–198. <https://doi.org/10.1007/s11242-008-9242-8>
- Chaoben D, Shesheng G (2018) Multi-focus image fusion with the all convolutional neural network. *Opto-electron Lett* 14:71–75
- Cheong JY, Park IK (2017) Deep CNN-based super-resolution using external and internal examples. *IEEE Signal Process Lett* 24:1252–1256
- Coelho D, Thovert JF, Adler PM (1997) Geometrical and transport properties of random packings of spheres and aspherical particles. *Phys Rev E* 55:1959–1978
- Comunian A, Renard P, Straubhaar J (2012) 3D multiple-point statistics simulation using 2D training images. *Comput Geosci* 40:49–65. <https://doi.org/10.1016/j.cageo.2011.07.009>
- Deepak AVS, Ghanekar U (2017) RDCN-SR: Integrating regression model with deep convolutional networks for image super-resolution. In: 2017 international conference on intelligent computing and control systems (ICICCS), 15–16 June 2017, pp 623–628. <https://doi.org/10.1109/iccons.2017.8250537>
- Deng J, Dong W, Socher R, Li LJ, Kai L, Li F-F (2009) ImageNet: a large-scale hierarchical image database. In: Paper presented at the 2009 IEEE conference on computer vision and pattern recognition, 20–25 June 2009
- Dong C, Loy CC, He K, Tang X (2016) Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell* 38:295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>
- Frederic B, Dominique J (2004) Modelling a food microstructure by random sets. *Image Anal Stereol* 23:33–44
- Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS (2016) Deep learning for visual understanding: a review. *Neurocomputing* 187:27–48. <https://doi.org/10.1016/j.neucom.2015.09.116>
- Hilfer R (1992) Local-porosity theory for flow in porous media. *Phys Rev B* 45:7115–7121
- Hiroshi O (2004) Pore-scale modelling of carbonates. Imperial College London, London
- Jain V, Seung S (2008) Natural image denoising with convolutional networks. In: Bottou DKaDSaYBaL (ed) *Advances in neural information processing systems*, 21. Curran Associates, Inc., pp 769–776
- Jia X, Xu X, Cai B, Guo K (2017) Single image super-resolution using multi-scale convolutional. *Neural Netw. ArXiv e-prints* 1705
- Kaiming H, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell* 37:1904–1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
- Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. *Commun ACM* 60:84–90. <https://doi.org/10.1145/3065386>
- Latham J-P, Lu Y, Munjiza A (2001) A random method for simulating loose packs of angular particles using tetrahedra. *Géotechnique* 51:871–879. <https://doi.org/10.1680/geot.2001.51.10.871>
- Latif FDE, Biswal B, Fauzi U, Hilfer R (2010) Continuum reconstruction of the pore scale microstructure for Fontainebleau sandstone. *Phys A Stat Mech Appl* 389:1607–1618. <https://doi.org/10.1016/j.physa.2009.12.006>
- LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1:541–551
- LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86:2278–2324
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444
- Liang ZR, Fernandes CP, Magnani FS, Philippi PC (1998) A reconstruction technique for three-dimensional porous media using image analysis and Fourier transforms. *J Petrol Sci Eng* 21:273–283. [https://doi.org/10.1016/S0920-4105\(98\)00077-1](https://doi.org/10.1016/S0920-4105(98)00077-1)
- Liu F, Lin G, Shen C (2015) CRF learning with CNN features for image segmentation. *Pattern Recogn* 48:2983–2992. <https://doi.org/10.1016/j.patcog.2015.04.019>
- Liu D, Wang Z, Wen B, Yang J, Han W, Huang TS (2016) Robust Single Image Super-Resolution via Deep Networks With Sparse Prior. *IEEE Trans Image Process* 25:3194–3207

- Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Paper presented at the CVPR2015
- Lucia FJ, Kerans C, Jennings JW Jr (2003) Carbonate reservoir characterization. *J Petrol Technol*. <https://doi.org/10.2118/82071-JPT>
- Manwart C, Torquato S, Hilfer R (2000) Stochastic reconstruction of sandstones. *Phys Rev E* 62:893–899
- Mosser L, Dubrulle O, Blunt MJ (2017) Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Phys Rev E* 96:043309
- Okabe H, Blunt MJ (2004) Prediction of permeability for porous media reconstructed using multiple-point statistics. *Phys Rev E* 70:066135
- Okabe H, Blunt M (2007) Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. *Water Resour Res* 43:179–183
- ØREN P-E, Bakke S (2002) Process based reconstruction of sandstones and prediction of transport properties. *Transp Porous Media* 46:311–343. <https://doi.org/10.1023/a:1015031122338>
- Ouyang W et al. (2015) DeepID-Net: deformable deep convolutional neural networks for object detection. In: Proceedings of the CVPR
- Pant LM, Mitra SK, Secanell M (2015) Multigrid hierarchical simulated annealing method for reconstructing heterogeneous media. *Phys Rev E* 92:063303
- Quiblier JA (1984) A new three-dimensional modeling technique for studying porous media. *J Colloid Interface Sci* 98:84–102. [https://doi.org/10.1016/0021-9797\(84\)90481-8](https://doi.org/10.1016/0021-9797(84)90481-8)
- Rintoul MD, Torquato S (1997) Reconstruction of the structure of dispersions. *J Colloid Interface Sci* 186:467–476. <https://doi.org/10.1006/jcis.1996.4675>
- Roberts AP (1997) Statistical reconstruction of three-dimensional porous media from two-dimensional images. *Phys Rev E* 56:3203–3212
- Roberts AP, Torquato S (1999) Chord-distribution functions of three-dimensional random media: approximate first-passage times of Gaussian processes. *Phys Rev E* 59:4953–4963
- Samuel S, Christian L, Horst B (2015) Fast and accurate image upscaling with super-resolution forests. In: Paper presented at the 2015 IEEE conference on computer vision and pattern recognition (CVPR), 7–12 June 2015
- Spanne P, Thovert JF, Jacquin CJ, Lindquist WB, Jones KW, Adler PM (1994) Synchrotron computed microtomography of porous media: topology and transports. *Phys Rev Lett* 73:2001–2004
- Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M (2014) Striving for simplicity: the all convolutional Net. ArXiv e-prints 1412
- Tahmasebi P, Sahimi M (2016a) Enhancing multiple-point geostatistical modeling: 1. Graph theory and pattern adjustment. *Water Resour Res* 52:2074–2098. <https://doi.org/10.1002/2015WR017806>
- Tahmasebi P, Sahimi M (2016b) Enhancing multiple-point geostatistical modeling: 2. Iterative simulation and multiple distance function. *Water Resour Res* 52:2099–2122. <https://doi.org/10.1002/2015WR017807>
- Tahmasebi P, Javadpour F, Sahimi M (2015) Multiscale and multiresolution modeling of shales and their flow and morphological properties. *Sci Rep* 5:16373. <https://doi.org/10.1038/srep16373>
- Tahmasebi P, Javadpour F, Sahimi M (2016) Stochastic shale permeability matching: three-dimensional characterization and modeling. *Int J Coal Geol* 165:231–242. <https://doi.org/10.1016/j.coal.2016.08.024>
- Tahmasebi P, Sahimi M, Andrade JE (2017) Image-based modeling of granular porous media. *Geophys Res Lett* 44:4738–4746. <https://doi.org/10.1002/2017GL073938>
- Talukdar MS, Torsæter O, Ioannidis MA, Howard JJ (2002) Stochastic reconstruction of chalk from 2D images. *Transp Porous Media* 48:101–123. <https://doi.org/10.1023/a:1015650705205>
- Wang Y, Rahman SS, Arns CH (2018) Super resolution reconstruction of  $\mu$ -CT image of rock sample using neighbour embedding algorithm. *Phys A Stat Mech Appl* 493:177–188
- Wu K, Nunan N, Crawford JW, Young IM, Ritz K (2004) An efficient Markov chain model for the simulation of heterogeneous soil structure. *Soil Sci Soc Am J* 68:6
- Wu K et al (2006) 3D stochastic modelling of heterogeneous porous media—applications to reservoir rocks. *Transp Porous Media* 65:443–467. <https://doi.org/10.1007/s11242-006-0006-z>
- Xuejiao W, Qiuyan T, Lianghao W, Dongxiao L, Ming Z (2015) Deep convolutional architecture for natural image denoising. In: Paper presented at the 2015 international conference on wireless communications & signal processing (WCSP), Nanjing, China
- Yamanaka J, Kuwashima S, Kurita T (2017) Fast and accurate image super resolution by deep CNN with skip connection and network in network. ArXiv e-prints 1707

- Yeong CLY, Torquato S (1998a) Reconstructing random media. *Phys Rev E* 57:495–506
- Yeong CLY, Torquato S (1998b) Reconstructing random media. II. Three-dimensional media from two-dimensional cuts. *Phys Rev E* 58:224–233
- Zeiler MD, Fergus R (2013) Stochastic pooling for regularization of deep convolutional neural networks. ArXiv e-prints 1301
- Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian Denoiser: residual learning of deep CNN for image denoising. *IEEE Trans Image Process* 26:3142–3155. <https://doi.org/10.1109/TIP.2017.2662206>