Statistical characterization and reconstruction of heterogeneous microstructures using deep neural network

Jinlong Fu[a], Shaoqing Cui[a], Song Cen[c], and Chenfeng Li [*a,b]

[a]Zienkiewicz Centre for Computational Engineering, College of Engineering, Swansea University Bay Campus, Swansea SA1 8EN, United Kingdom
[b]Energy Safety Research Institute, College of Engineering, Swansea University Bay Campus, Swansea SA1 8EN, United Kingdom
[c]School of Aerospace, Tsinghua University, Beijing 100084, China

Abstract

Heterogeneous materials, whether natural or artificial, are usually composed of distinct constituents present in complex microstructures with discontinuous, irregular and hierarchical characteristics. For many heterogeneous materials, such as porous media and composites, the microstructural features are of fundamental importance for their macroscopic properties. This paper presents a novel method to statistically characterize and reconstruct random microstructures through a deep neural network (DNN) model, which can be used to study the microstructure-property relationships. In our method, the digital microstructure image is assumed to be a stationary Markov random field (MRF), and local patterns covering the basic morphological features are collected to train a DNN model, after which statistically equivalent samples can be generated through a DNN-guided reconstruction procedure. Furthermore, to overcome the short-distance limitation associated with the MRF assumption, a multi-level approach is developed to preserve the long-distance morphological features of heterogeneous microstructures. A large number of tests have been conducted to compare the reconstructed and target microstructures in both morphological characteristics and physical properties, and good agreements are observed in all test cases. The proposed method is efficient, accurate, versatile, and especially beneficial to the statistical reconstruction of 2D/3D microstructures with long-distance correlations.

Keywords: Heterogeneous material; Random microstructure; Characterization and reconstruction; Statistical equivalence; Physical property.

1 Introduction

Random heterogeneous materials comprised of at least two distinct phases (including void) are ubiquitous in engineering and natural environment, and typical examples include composites, alloys, concretes, soils and rocks. Their physical properties (i.e. transport, elastic, and conductive properties) usually exhibit strong uncertainty, due to the random distributions of different phases and the phase discontinuity on the interfaces (Torquato, 2013). Microstructural characteristics of heterogeneous media are considered to be the key to understanding their macroscopic behaviors (Berg, 2014; Gupta et al., 2015). However, it is not trivial to quantitatively characterize the random microstructures because of various reasons: (1) heterogeneous media often contain internal structures at the microscale or even nanoscale, and thus advanced microscopy imaging techniques are the only option for non-invasive and non-destructive measurements (Brandon and Kaplan, 2013); (2) hierarchical microstructures are often present such that the characteristic component itself comprises a structure at a smaller scale (Yang et al., 2017); (3) the material morphology is of a random nature in terms of shape, size and spatial distribution of different phases (Vogel et al., 2010).

---

[*]Corresponding author: c.f.li@swansea.ac.uk

Modern microscopy imaging techniques have been able to digitize the geometry of a heterogeneous medium, usually called a microstructure, into a 2D or 3D image at different resolution levels and various length scales. The digital microstructure can be used for a variety of simulation or diagnostic purposes. Scanning electron microscopy (SEM) (Stutzman, 2004), backscattered electron (BSE) imaging, atomic force microscopy (AFM), optical microscopy and transmission electron microscopy (TEM) are commonly-used techniques to provide 2D visualizations of microstructures. Besides, 3D digital microstructures can be acquired through X-ray micro-computed tomography (μ-CT) (Schlüter et al., 2014), magnetic resonance imaging (MRI), and focused ion beam scanning electron microscopy (FIB-SEM). After image acquisition, image filtering and segmentation are performed on the raw scanning images, where the raw grayscale images are converted to a more discrete form permitting quantitative characterization and numerical modeling (Schlüter et al., 2014). For illustration, an SEM image of a composite sample and its binary segmentation are shown in Figures 1a and 1b, respectively; and a μ-CT image of a Gildehauser sandstone sample and its binary segmentation are plotted in Figures 1c and 1d, respectively.
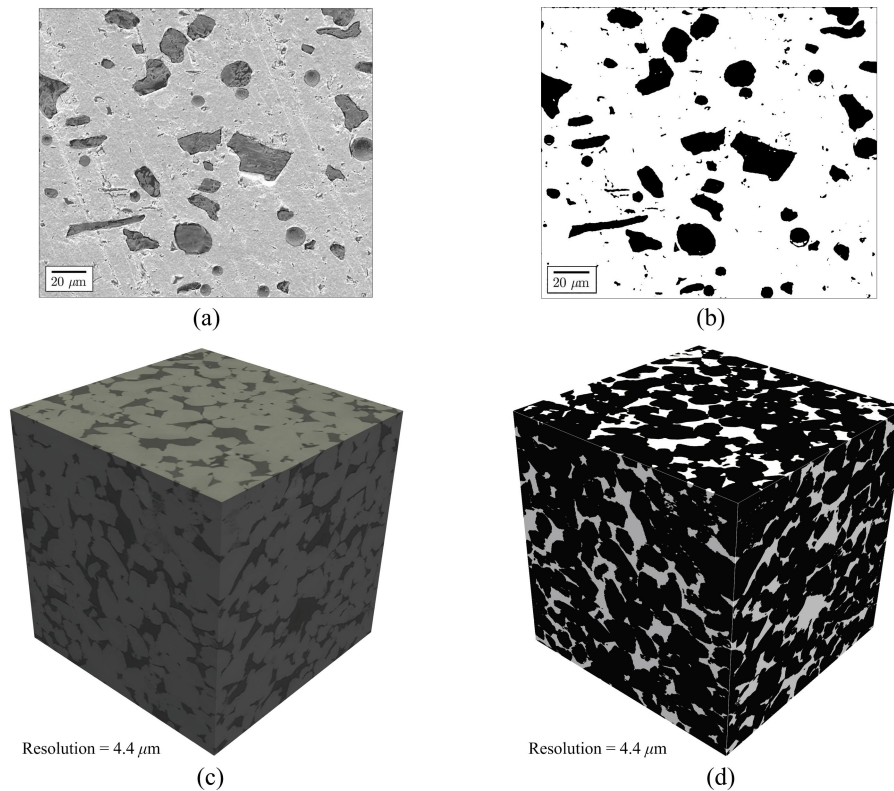

Figure 1: (a) SEM image of a composite sample (Yang et al., 2018); (b) The binary segmentation of the SEM image (white areas denote epoxy matrix, and black regions denote silica phase); (c) μ-CT image of a Gildehauser sandstone sample; (d) The binary segmentation of the μ-CT image (white regions denote pore phase, and black areas denote solid matrix).

Although useful insight into heterogeneous media can be derived from modern microscopy imaging techniques, it remains a critical challenge to effectively link microstructural characteristics to macroscopic material properties. Such knowledge is of great value for predicting material properties from the observable microstructural informatics (Torquato, 2013). Additionally, microstructural characteristics can also be linked to material formation processes, which can help to design and manufacture materials to possess desirable properties. In practice, material properties are usually measured by experiments on physical samples or estimated from numerical simulations on digital microstructure samples. However, physical sample acquisition, such as rock sample drilling, is usually expensive, and microscopy imaging is also of high-cost (Wu et al., 2006). Without sufficient physical (or digital) samples, direct experiments (or numerical simulations) are unable to correctly assess heterogeneous media due to their stochastic nature. The complete computational model of a heterogeneous

medium is far beyond a small number of microstructure samples, and it should cover representative samples with all possible configurations. In other words, the complete computational model is an ensemble of representative/statistical volume elements that share the same averaged characteristics. This situation promotes researchers to find an alternative approach to understanding the microstructure-property relationships with limited availability of physical (or digital) raw samples (Ju et al., 2014).

A reliable way to quantitatively characterize the complex microstructure is an essential prerequisite to establishing effective microstructure-property relationships. Stochastic characterization provides a statistical description of microstructural features, where the stochastic morphology of complex microstructures is reduced to a small set of descriptors related to material properties (Torquato, 2013). As mentioned above, the number of raw microstructures is usually limited, due to the high cost of advanced microscopy imaging techniques, which brings a practical difficulty for accurate characterization. Therefore, stochastic microstructure reconstruction was introduced to generate microstructural samples with statistical equivalence by using limited microstructural information (Yeong and Torquato, 1998), and it has been proved to be an effective and economical way to tackle this problem.

Various methods have been developed to reconstruct heterogeneous microstructures over the past few decades. Among them, the stochastic optimization method (Yeong and Torquato, 1998) and Gaussian random field method (Quiblier, 1984; Feng et al., 2014, 2016) are the most commonly-used approaches. The stochastic optimization method generates a microstructure sample through an optimization procedure by setting up the optimization objective with microstructural descriptors (Yeong and Torquato, 1998). The reconstructed trial sample is iteratively adjusted to minimize the statistical difference (measured by the selected descriptors) between it and the original microstructure, in order to produce a microstructural sample that is statistically equivalent to the original one. The simulated annealing algorithm is often used to drive the optimization procedure. This reconstruction method is flexible and robust, and the accuracy of realizations can be controlled by setting an appropriate termination criterion for the optimization process. However, this method is very computation-intensive, due to its iterative nature of the optimization procedure and the repetitive calculation of statistical descriptors. Gaussian random field method (Quiblier, 1984; Feng et al., 2014, 2016) models a microstructure as a Gaussian random field, and it produces a microstructure sample by truncating a manually generated realization of random filed. Generally, this type of method is much faster than the stochastic optimization method to generate microstructure samples, but it only captures the first- and second-order statistical information, which is inadequate to preserve the morphological complexity of heterogeneous microstructures. Moreover, some other reconstruction methods, including process-based method (Øren and Bakke, 2002), Markov Chain Monte Carlo simulation (Wu et al., 2006), multiple-point statistics method (Okabe and Blunt, 2005), patch-based method (Tahmasebi and Sahimi, 2012) and texture synthesis method (Liu and Shapiro, 2015), are also used frequently.

Effective reconstruction methods should allow accurate and efficient reconstruction of realistic microstructures, such that geometrical measurement and numerical simulation can be performed on the reconstructed samples to investigate the impact of microstructures on material properties. Recently, machine learning (Ramprasad et al., 2017) and deep learning (Agrawal and Choudhary, 2019) have been successfully applied to characterize and reconstruct complex heterogeneous media, because of their powerful capacities to analyze complex data and explore hidden connections.

Caers (2001) first used the artificial neural network (ANN) to statistically model connected reservoirs. Latter, Sundararaghavan and Zabaras (2005) considered microstructure reconstruction as a pattern recognition problem, and they adopted the support vector machine (SVM) to reconstruct 3D microstructures by using statistical information extracted from planar images. The SVM-based reconstruction method can generate microstructure samples with multiple features closely matching the original microstructure, but *a prior* database of 3D microstructures is required for feature extraction and SVM model training, which is often unavailable in practice. In 2016, Bostanabad et al. (2016a,b) adopted the decision tree (supervised machine learning) model to learn the stochastic feature of heterogeneous microstructures, and then statistically equivalent microstructure samples are generated based on this characterization. As a simple supervised machine learning model, the decision tree may not fully capture the complex morphology of heterogeneous microstructures. As Markov random field assumption was used for microstructure characterization, this method may have limitations in capturing long-distance correlations. Cang et al. (2017) developed a method to characterize and reconstruct heterogeneous microstruc-

tures via the convolution deep belief network (CDBN). This method is able to generate samples that statistically preserve the long-distance morphological features as well as the critical fracture strength values. Currently, this method can only be used for 2D microstructure reconstruction, and it cannot guarantee the accuracy, because the key parameters such as threshold are determined empirically. Mosser et al. (2017) proposed a method to reconstruct 3D porous media by using the generative neural network (GAN), where the GAN serves as an implicit description of the probability distribution of the 3D image dataset. The reconstructed samples closely match the reference porous media in terms of pore morphology and single-phase permeability, but this method requires *a prior* dataset of 3D digital images to train the GAN model. More recently, the convolutional neural network (CNN) (Li et al., 2018; Wang et al., 2018) has been used to characterize and reconstruct heterogeneous microstructures. GAN and CNN are powerful tools to characterize microstructures, but the efficiency to reconstruct new samples still needs to improve, especially for large 3D microstructures.

Based on deep neural network (DNN), this paper presents a new method for statistical characterization and reconstruction of heterogeneous microstructures in both 2D and 3D formats. To examine its performance, a series of reconstruction tests and comparisons involving many different types of microstructures are conducted, where the reconstruction quality is assessed by comparing both morphological and physical properties of reconstructed and target microstructures. The main technical advantage includes: (1) using stacked sparse autoencoder (SSAE) and Softmax classifier, the proposed method is accurate, efficient, and applicable to a wide range of heterogeneous materials with various morphological features; (2) using a multi-level approach to statistically characterize and reconstruct microstructures in a hierarchical manner, the new method can accurately capture long-distance morphological features, making it particularly suitable for the application in porous media with long-range connectivity. The remainder of the paper is organized as follows. In § 2, statistical microstructure characterization using deep neural network is presented in detail through a 2D microstructure, including basic assumption, training data collection and process, and 'SSAE+Softmax' model training. § 3 explains the procedure of stochastic reconstruction based on the microstructure characterization from the trained 'SSAE+Softmax' model. § 4 examines the performance on 2D microstructures with distinct morphologies, where the proposed method is systematically compared with the recent machine learning-based method proposed by Bostanabad et al. (2016a,b). In § 5, the method is extended for 3D microstructures, and the multi-level approach is developed for microstructures with long-distance morphological features. Finally, conclusions are drawn in § 6.

2 Microstructure Characterization

The raw SEM or μ-CT images of heterogeneous microstructures are in grayscale, so they are usually segmented to represent different phases of the microstructures with distinct intensity values (Schlüter et al., 2014). Statistically characterizing the distributions of different phases within the microstructures is critical, which is the first step to reconstruct new microstructure samples with statistical equivalence. In this section, we take the 2D heterogeneous microstructure with two phases as the example to illustrate the proposed microstructure characterization method using deep neural network, and this method can be readily extended to 3D microstructures with multiple phases.

2.1 Markov random field

An example of 2D digital microstructure with two phases is shown in Figure 1b, where its two phases are represented by white and black regions respectively. In computer vision, a digital image is a dot matrix data structure that represents a typically rectangular grid of pixels. Let the matrix $\boldsymbol{X}_{(s_1 \times s_2)}$ denote the set of pixels in a 2D digital microstructure with size $s_1 \times s_2$:

$$\boldsymbol{X} = \{X_{ij} \mid 1 \leq i \leq s_1, 1 \leq j \leq s_2\} \tag{1}$$

where X_{ij} denotes the element in the matrix $\boldsymbol{X}$, representing the intensity value of the pixel located at row i and column j. For the 2-phase microstructure, X_{ij} is a binary variable described by the phase function as follows:

$$X_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is located at white phase} \\ 0 & \text{if } (i,j) \text{ is located at black phase} \end{cases} \tag{2}$$

4

2.1.1 Markov random field assumption

Due to the stochastic nature of heterogeneous media, the matrix $\boldsymbol{X}$ corresponding to a microstructure sample can be described by the joint probability distribution function (PDF) $p(X_{11}, X_{12}, \ldots, X_{ij}, \ldots, X_{s_1,s_2})$ or $p(\boldsymbol{X})$, which is the perfect characterization of the microstructure $\boldsymbol{X}$. If $p(\boldsymbol{X})$ could be learned, it can be directly used to reconstruct statistically equivalent microstructure samples. However, $p(\boldsymbol{X})$ usually has an extremely high dimension equal to the number of pixels in the image, making it impracticable to estimate $p(\boldsymbol{X})$ from limited number of digital images $\boldsymbol{X}$. To simplify the microstructure characterization problem, the digital image of a heterogeneous medium is assumed to be a Markov random filed, which can greatly reduce the computing complexity of $p(\boldsymbol{X})$.

MRF models a texture as a local and stationary random process, and it has been proven to be an effective method for a wide range of textures in computer vision applications (Li, 2009). Specifically, MRF has two assumptions (Wei and Levoy, 2000):

- **Locality**: each pixel X_{ij} in the digital image is only related to its neighboring pixels $\boldsymbol{N}_{ij}$ within a sufficiently large area:

$$p(X_{ij}|\boldsymbol{X}^{(-ij)}) = p(X_{ij}|\boldsymbol{N}_{ij}) \tag{3}$$

  where $\boldsymbol{X}^{(-ij)}$ denotes the rest of pixels in $\boldsymbol{X}$ excluding pixel X_{ij}.

- **Stationarity**: $p(X_{ij}|\boldsymbol{N}_{ij})$ is the same for all pixels regardless of the pixel location (i, j), which signifies that different regions in the image always share the same statistical characteristics.

2.1.2 Data template and extension of MRF assumption

The geometry and size of spatial region covering the neighboring pixels $\boldsymbol{N}_{ij}$ are important for the application of MRF. In this study, the neighboring pixels $\boldsymbol{N}_{ij}$ within the L-shaped region are adopted to characterize the central pixel X_{ij}, and the whole area covering both the central pixel and its neighborhood is called a 'data template', as shown in Figure 2. The reason to choose the L-shaped data template is that the neighboring pixels $\boldsymbol{N}_{ij}$ are prior to the central pixel X_{ij} in the raster scan order, which meets the requirement of microstructure reconstruction in our method (more details will be discussed in § 3). The size of the data template is measured by the radius r, and it should be large enough to capture the fundamental morphology features in the image.

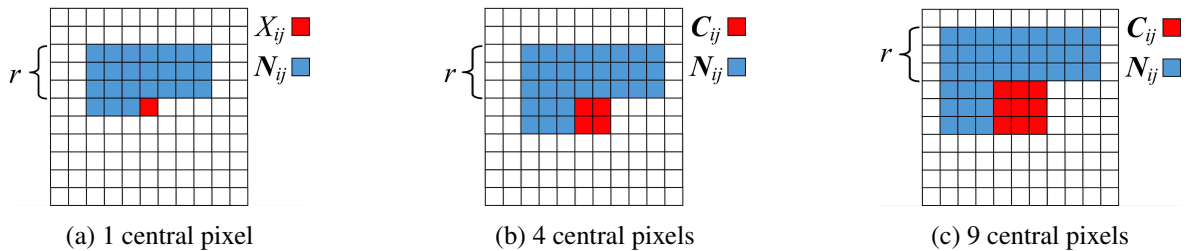| (a) 1 central pixel | (b) 4 central pixels | (c) 9 central pixels |

Figure 2: Data templates with central pixel(s) and neighboring pixels (neighborhood radius $r = 3$ pixels).

The classical MRF model only focuses on the relationship between one central pixel X_{ij} and its surrounding neighboring pixels $\boldsymbol{N}_{ij}$ (Li, 2009). In this study, we extend the classical MRF assumption into more general forms, where correlation between multiple central pixels $\boldsymbol{C}_{ij}$ and their surrounding neighboring pixels $\boldsymbol{N}_{ij}$ can also be considered, and the corresponding data templates are shown in Figures 2b and 2c. The locality and stationarity assumptions are both applied equally to the cases with multiple central pixels, and Eq. (3) can be rewritten as:

$$p(\boldsymbol{C}_{ij}|\boldsymbol{X}^{(-ij)}) = p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij}) \tag{4}$$

where $\boldsymbol{X}^{(-ij)}$ denotes the set of pixels in $\boldsymbol{X}$ excluding the central pixels $\boldsymbol{C}_{ij}$.

The possible configurations of multiple central pixels are much more diverse than the case of single central pixel. Take the 4-central-pixel case as an example, there are 2^4 possible patterns for these 4 central pixels in a 2-phase microstructure. The purpose of MRF extension is two-fold: (1) to capture the complexity (long-distance

correlations) of microstructures by using the proposed multi-level approach (as described in § 5); and (2) to speed up the microstructure reconstruction process by updating multiple pixels at each iteration step (as discussed in § 3). One can choose the most suitable template for microstructure characterization and reconstruction to balance efficiency and accuracy.

2.2 Data event collection

The data template is used to collect primary local patterns from the microstructure, so it has a significant influence on microstructure characterization, thereby determining the quality of microstructure reconstruction. The selected data template scans over the entire digital image X, termed training image here, to collect data events (C_{ij}, N_{ij}) in a raster scan order. Each data event (C_{ij}, N_{ij}) is a paired observation, and its configuration varies with the location (i, j) to represent the variation of local spatial pattern in the training image X. The scanning process starts from the top-left pixels with enough neighboring pixels to fit the data template. The border pixels are not scanned as the central pixel, because they do not have sufficient neighborhoods to accommodate the data template. Figure 3 illustrates the data event collection by using the selected data template to scan over the entire training image.
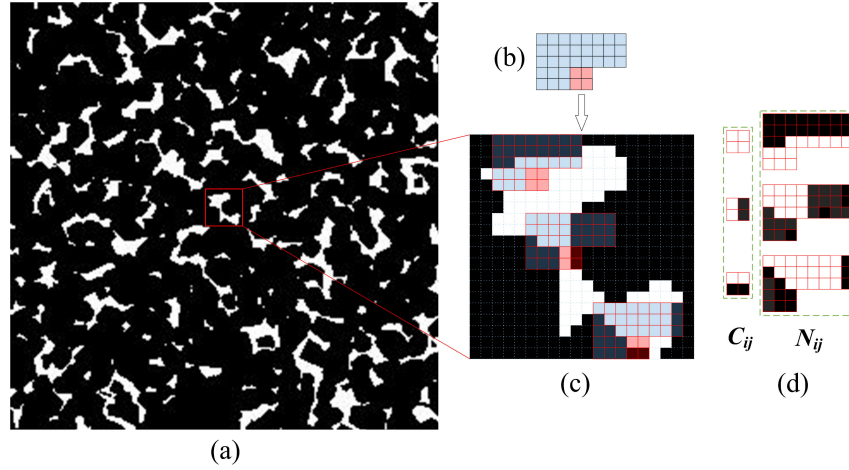

Figure 3: Illustration of data event collection from the training image: (a) The training image of a 2D microstructure; (b) Data template with 4 central pixels ($r = 3$ pixels); (c) Partial enlarge detail of the training image with data template scanning on it; (d) Three examples of data event (C_{ij}, N_{ij}).

The conditional probability distribution function (CPDF) $p(C_{ij}|N_{ij})$ associated with the data template can be estimated from the occurrence frequencies of data events collected from the training image:

$$p(C_{ij}, N_{ij}) = \frac{O(C_{ij}, N_{ij})}{\sum O(C_{ij}, N_{ij})} = \frac{O(C_{ij}, N_{ij})}{O_{\text{all}}} \tag{5}$$

$$p(N_{ij}) = \frac{O(N_{ij})}{\sum O(N_{ij})} = \frac{O(N_{ij})}{O_{\text{all}}} \tag{6}$$

$$p(C_{ij}|N_{ij}) = \frac{p(C_{ij}, N_{ij})}{p(N_{ij})} = \frac{O(C_{ij}, N_{ij})}{O(N_{ij})} \tag{7}$$

where $O(C_{ij}, N_{ij})$ and $O(N_{ij})$ are the occurrence numbers of data event (C_{ij}, N_{ij}) and (N_{ij}) respectively, and O_{all} is the total number of all data events collected from the training image.

2.3 Microstructure characterization via deep neural network

The core of microstructure characterization is to determine the CPDF $p(X_{ij}|N_{ij})$ or $p(C_{ij}|N_{ij})$ of the training image X. However, explicit construction of the CPDF is complicated and computationally intensive (Wei and

Levoy, 2000), especially for heterogeneous microstructures with complicated geometries, sometimes even impossible. As mentioned above, the phase value of each pixel X_{ij} can only be 1 (white) or 0 (black) in the binary representation of a microstructure. Therefore, the conditional probability $p(X_{ij}|\boldsymbol{N}_{ij})$ associated with 1-central-pixel data template can be considered as a class probability, which provides the probability of the central pixel X_{ij} to be white or black when given its neighboring pixels $\boldsymbol{N}_{ij}$. It is a binary classification problem, where the phase value of X_{ij} is the classification category and the configuration of $\boldsymbol{N}_{ij}$ is the classification feature. In addition, the data events $(X_{ij}, \boldsymbol{N}_{ij})$ collected from the training image are paired observations, which are perfect training dataset to train a machine learning model for classification (Bishop, 2006). Taking the collected data events $(X_{ij}, \boldsymbol{N}_{ij})$ as the training dataset, the machine learning classifier can be built to minimize the probability of misclassification for any input features $\boldsymbol{N}_{ij}$. The class probability stored in the fitted machine learning model will be an accurate estimation of the CPDF $p(X_{ij}|\boldsymbol{N}_{ij})$. As to the CPDF $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$ associated with data template having multiple central pixels, it can also be estimated in the same way. But data conversion is required to assign a label L_{ij} to each configuration of the multiple central pixels $\boldsymbol{C}_{ij}$, and then the label L_{ij} together with the neighboring pixels $\boldsymbol{N}_{ij}$ will serve as a pair of observation to train the machine learning model. In other words, the binary classification problem discussed above becomes a multi-class classification problem here, but the basic principles remain. More details about data conversion will be explained in § 2.3.1.

An appropriate machine learning model is critical for the above classification (pattern recognition) task. Considering the high dimension of feature data (feature data $\boldsymbol{N}_{ij}$ usually contain hundreds of pixels, and the number of voxels will exceed 1000 for 3D microstructure cases) and the large number of classification categories (the number of possible classification categories for the n-central-pixel case is 2^n, and it becomes even larger for multi-phase microstructures), the deep neural network (DNN) appears to be the most attractive option. DNN models (Schmidhuber, 2015) with multiple layers are effective for solving classification problems with complex data, and each layer can learn/capture features at a different level.

2.3.1 Data conversion for the central pixels in data template

For the 1-central-pixel data template, the central pixel has two possible configurations, and each configuration can be labeled by the phase value. However, for the 4- or 9-central-pixel data template, there are 2^4 or 2^9 possible configurations for the central pixels. An index label is required for each possible configuration, so that the machine learning model can be trained to map each label to the corresponding feature data.
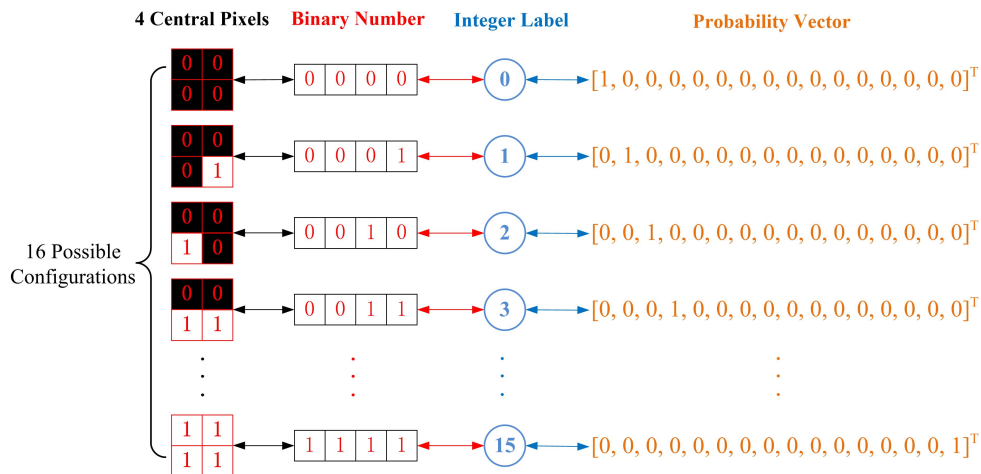

Figure 4: Flow chart of data conversion (taking 4-central-pixel data template as the example).

As shown in Figure 4, a 4-central-pixel data template is used to explain the data conversion procedure. Given 4 central pixels $\boldsymbol{C}_{ij}$, there are 16 possible pattern configurations in total, and for each configuration, the phase values are first recorded as a binary number, which is then translated into an integer label, denoted by L_{ij}. Finally, the integer label is translated into a probability vector so that the class probabilities can be directly outputted from the trained DNN model upon new input data. The data conversion procedure remains the

same for a n-central-pixel data template with two phases. As to multi-phase microstructures, such as 3-phase microstructures, the data conversion still works by changing the binary number to a ternary number. It should be noted that the data conversion is reversible, which is important for the stochastic reconstruction of microstructure samples through probability sampling.

2.3.2 Data compression for the neighboring pixels in data template

Although the MRF assumption has been adopted to reduce dimensionality for microstructure characterization, the number of neighboring pixels in the data template is usually large, especially for 3D microstructures with long-distance morphological characteristics. This section explains how to extract features from the raw pixel inputs through the stacked sparse autoencoder (SSAE) (Hinton and Salakhutdinov, 2006; Xu et al., 2015; Wang et al., 2016).

Autoencoder is an unsupervised machine learning algorithm that learns the representation of high-dimension input data by finding the correlation between the data, and is especially suitable for dimensionality reduction. Consisting of an encoder and a decoder, the autoencoder is a multi-layer feed-forward neural network trained to represent the input through back-propagation. It compresses the input data into lower dimensional code and then reconstructs the output from this latent-space representation. During the training process, the discrepancy between input and reconstruction is minimized by applying back-propagation, and a set of parameters including weights W and biases b are optimized, as illustrated in Figure 5.
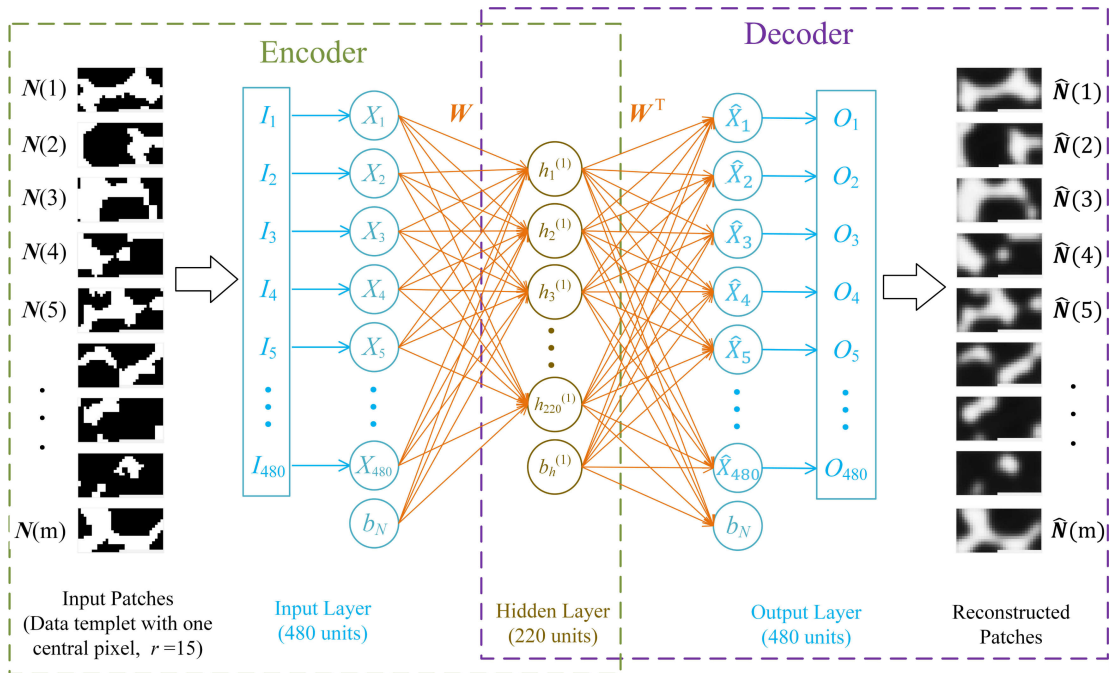

Figure 5: The architecture of basic sparse autoencoder to compress local morphology patterns of heterogeneous microstructures (each input patch contains 480 pixels).

A set of neighboring pixels $\boldsymbol{N}_{ij}$ is called a 'training patch' here. Let $\boldsymbol{Pa} = (\boldsymbol{N}(1), \boldsymbol{N}(2), \ldots, \boldsymbol{N}(m))$ denote the entire training patches collected from the training image, where $\boldsymbol{N}(k) \in \mathcal{R}^{d_N}$, m and d_N are the number of training patches and the number of pixels in each patch, respectively. Let $\boldsymbol{h}^{(l)}(k) = (h_1^{(l)}(k), h_2^{(l)}(k), \ldots, h_{d_h}^{(l)}(k))^T$ denote the learned high-level feature at layer l for the k-th patch, where d_h is the number of hidden units in the current layer l. Throughout this section, the superscript and subscript on a notation are used to define the hidden layer and unit in this layer, respectively. For example, the $h_i^{(1)}$ in Figure 5 represents the i-th unit in the 1st hidden layer. For simplicity, $\boldsymbol{N}$ and $\boldsymbol{h}^{(l)}$ are used to denote an input patch and its representation at hidden layer l, respectively.

Figure 5 shows the architecture of a sparse autoencoder. Generally, the input layer of the autoencoder consists in an encoder, where input patches $\boldsymbol{N}$ are compressed into lower dimensional representations $\boldsymbol{h}$. The

output layer is a decoder which is trained to reconstruct approximations $\widehat{\mathbf{N}}$ of the input patches $\mathbf{N}$ from the hidden representation $\mathbf{h}$. In essence, training an autoencoder is to optimize the parameters by setting the minimization of the discrepancy between input $\mathbf{N}$ and output $\widehat{\mathbf{N}}$ as the objective. This discrepancy is represented by a cost function with three terms as follows:

$$\mathcal{L}(\theta) = \frac{1}{m} \sum_{k=1}^{m} L(\mathbf{N}(k), d_{\hat{\theta}}(e_{\check{\theta}}(\mathbf{N}(k)))) + \alpha \sum_{j=1}^{n} KL(\rho||\hat{\rho}_j) + \beta \|W\|_2^2 \tag{8}$$

The first term in Eq. (8) is the mean sum-of-squared error that accounts for the discrepancy between input $\mathbf{N}$ and output $\widehat{\mathbf{N}}$ over the entire set of data. Encoder $e_{\check{\theta}}(\cdot)$ compresses the input $\mathbf{N} \in \mathcal{R}^{d_N}$ into the hidden representation $\mathbf{h} \in \mathcal{R}^{d_h}$ through the following equation:

$$\mathbf{h} = e_{\check{\theta}}(\mathbf{N}) = f(\mathbf{W}\mathbf{N} + \mathbf{b}_h) \tag{9}$$

where $\mathbf{W}$ is a $d_h \times d_N$ weight matrix, and $\mathbf{b}_h \in \mathcal{R}^{d_h}$ is a bias vector. The encoder is parametrized by $\check{\theta} = (W, b_h)$. Decoder $d_{\hat{\theta}}(\cdot)$ transforms the hidden representation $\mathbf{h}$ into the input space $\widehat{\mathbf{N}}$ by:

$$\widehat{\mathbf{N}} = d_{\hat{\theta}}(\mathbf{h}) = f(\mathbf{W}^T \mathbf{h} + \mathbf{b}_N) \tag{10}$$

where $\mathbf{W}^T$ is a $d_N \times d_h$ weight matrix, and $\mathbf{b}_N \in \mathcal{R}^{d_N}$ is a bias vector. Hence, the decoder is parameterized by $\hat{\theta} = (W^T, b_N)$. The activation function $f(\cdot)$ is a sigmoid logistic function:

$$f(z) = \frac{1}{1 + e^{-z}} \tag{11}$$

where z is the pre-activation of a neuron. The cost function $L(\mathbf{N}, \widehat{\mathbf{N}})$ quantifies the discrepancy between input $\mathbf{N}$ and the reconstruction $\widehat{\mathbf{N}}$, given by:

$$L(\mathbf{N}, \widehat{\mathbf{N}}) = (\mathbf{N} - \widehat{\mathbf{N}})^2 \tag{12}$$

The second term in Eq. (8) describes the sparsity cost, where n is the number of units in the hidden layer, and the index j scans across all hidden units in the network. $\hat{\rho}_j$ is the average activation value of hidden unite j over the entire data set, and ρ is the desired activation value. $KL(\rho||\hat{\rho}_j)$ denotes the Kullback-Leibler (KL) divergence between ρ and $\hat{\rho}_j$:

$$KL(\rho||\hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j} \tag{13}$$

The third term in Eq. (8) defines the regularization cost (also called weight decay term), which tends to decrease the magnitude of the weight to prevent overfitting.

$$\|W\|_2^2 = tr(\mathbf{W}^T \mathbf{W}) \tag{14}$$

The basic sparse autoencoders can also be stacked to form a hierarchical sparse deep model, where high-order features are compressed, organized and extracted. A stacked sparse autoencoder (SSAE) is a deep neural network consisting of multiple sparse autoencoders, where the output of each autoencoder layer is fed into the inputs of the successive autoencoder layer. The architecture of an SSAE that consists of two autoencoder layers is shown in Figure 6. The figure illustrates only the encoder part of each basic autoencoder, because the decoder part is not involved in the feature dimensionality reduction when estimating the class probability for an input patch (further explanation is provided in the next subsection). The SSAE yields a function $S: \mathcal{R}^{d_N} \to \mathcal{R}^{d_{h^{(2)}}}$, which depresses an input raw pixels of a patch to a new feature representation $\mathbf{h}^{(2)} = S(\mathbf{N})$. Due to large numbers of initial parameters (weights and bias) in SSAE, it is challenging to effectively optimize the parameters of these autoencoders. Pretraining the network of a deep autoencoder to make the initial parameters close to the good solutions is a practical strategy to overcome this problem. The greedy layer-wise learning algorithm is an effective method to pretrain a deep network by training each layer in turn.

2.3.3 SSAE+Softmax classifier: An implicit form of CPDF

Softmax classifier (Bishop, 2006) is a multi-class neural network model using supervised learning algorithm, which is used to classify the final output from the SSAE into multiple classes and calculates a probability for each possible class. The architecture of SSAE with two sparse autoencoders followed by a Softmax classifier is shown in Figure 6. The first autoencoder layer of the SSAE model learns primary features $\boldsymbol{h}^{(1)}$ from the raw pixel input $\boldsymbol{N}$. Then, these primary features are used as the input for the second autoencoder layer of the SSAE model to learn secondary features $\boldsymbol{h}^{(2)}$ from the primary features. Next, these secondary features are served as the input for a Softmax classifier layer, which is trained to map the secondary features $\boldsymbol{h}^{(2)}$ to the corresponding labels $\boldsymbol{L}$. These two sparse autoencoders and the Softmax classifier are combined together to form an 'SSAE+Softmax' model with two hidden layers and a classification layer, as illustrated in Figure 6. After the 'SSAE+Softmax' classifier is properly trained, it can provide a probability vector for a new testing patch. This probability vector contains all probabilities of the testing patch belonging to corresponding classes.
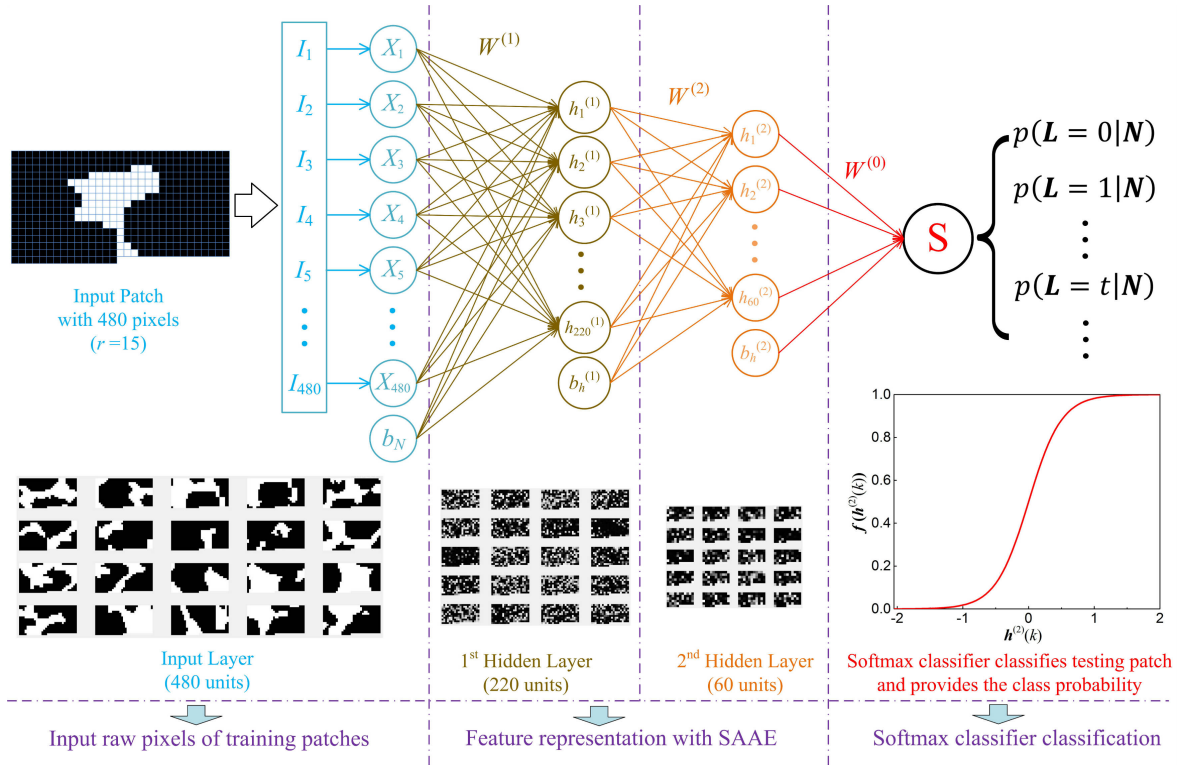

Figure 6: Schematic diagram of the 'SSAE+Softmax' model consisting of two autoencoder layers and a Softmax classifier layer to classify the testing patches (local morphology patterns) of a microstructure. Only the encoder part of each sparse autoencoder is drawn here, because the decoder part is not involved in the feature dimensionality reduction for estimating the class probability.

Mathematically, the class probability of an input patch $\boldsymbol{N}$ belonging to class t can be written as:

$$p(\boldsymbol{L} = t|\boldsymbol{N}) = p(\boldsymbol{L} = t|\boldsymbol{h}^{(2)}) = Softmax(\boldsymbol{W}^{(0)}\boldsymbol{h}^{(2)} + \boldsymbol{b}^{(0)}) = \frac{e^{W_t^{(0)}h^{(2)}+b_t^{(0)}}}{\sum_k^K e^{W_k^{(0)}h^{(2)}+b_k^{(0)}}} \tag{15}$$

where $\boldsymbol{W}^{(0)}$ and $\boldsymbol{b}^{(0)}$ are the weight matrix and bias vector of the Softmax classifier, respectively, and K is the number of classification categories. Considering the mapping relationship between the label $\boldsymbol{L}$ and the configuration of central pixels $\boldsymbol{C}$, the class probability $p(\boldsymbol{L}|\boldsymbol{N})$ from the 'SSAE+Softmax' classifier is equivalent to the conditional probability $p(\boldsymbol{C}|\boldsymbol{N})$. From this perspective, the 'SSAE+Softmax' classifier model trained by the training data $(\boldsymbol{X}_{ij}, \boldsymbol{N}_{ij})$ or $(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$ is an implicit form of the conditional probability distribution (CPDF) $p(\boldsymbol{X}_{ij}|\boldsymbol{N}_{ij})$ or $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$. In practice, cross-validation should be conducted to prevent overfitting the 'SSAE+Softmax' classifier model, and thereby to improve its predictive performance on new observations.

3 Microstructure Reconstruction

In the previous section, statistical microstructure characterization is obtained by training an 'SSAE+Softmax' classifier model to learn the CPDF $p(X_{ij}|\mathbf{N}_{ij})$ or $p(\mathbf{C}_{ij}|\mathbf{N}_{ij})$ of the training image $\mathbf{X}$. This section describes how to generate statistically equivalent microstructure samples based on the obtained characterization.

3.1 Microstructure reconstruction procedure

The joint PDF $p(\mathbf{Y})$ of the reconstructed microstructure sample $\mathbf{Y}$ can be expressed as:

$$
\begin{aligned}
p(\mathbf{Y}) &= p(Y_{11})p(Y_{12}|Y_{11})p(Y_{13}|Y_{11}, Y_{12}) \cdots p(Y_{ij}|Y_{11}, Y_{12}, \cdots Y_{i(j-1)}) \cdots \\
&= p(Y_{11})p(Y_{12}|\mathbf{Y}_{(<12)})p(Y_{13}|\mathbf{Y}_{(<13)}) \cdots p(Y_{ij}|\mathbf{Y}_{(<ij)}) \cdots
\end{aligned}
\tag{16}
$$

where $\mathbf{Y}_{(<ij)}$ represents the pixel sequence ahead of the pixel Y_{ij}. If all conditional probabilities $p(Y_{ij}|\mathbf{Y}_{(<ij)})$ are known, it is feasible to sequentially generate pixel Y_{ij} by sampling from the $p(Y_{ij}|\mathbf{Y}_{(<ij)})$. As shown in Figure 2a, the neighboring pixels $\mathbf{N}_{ij}$ in the data template are all in front of the central pixel X_{ij} in the raster scan order. With the MRF assumption, the following relation holds:

$$
p(Y_{ij}|\mathbf{Y}_{(<ij)}) = p(Y_{ij}|\mathbf{N}_{ij}^{(Y)})
\tag{17}
$$

where $\mathbf{N}_{ij}^{(Y)}$ is the set of neighboring pixels surrounding the central pixel Y_{ij}. Substituting Eq. (17) into Eq. (16), the joint PDF $p(\mathbf{Y})$ is rewritten as:

$$
p(\mathbf{Y}) = p(Y_{11})p(Y_{12}|\mathbf{N}_{12}^{(Y)})p(Y_{13}|\mathbf{N}_{13}^{(Y)}) \cdots p(Y_{ij}|\mathbf{N}_{ij}^{(Y)}) \cdots
\tag{18}
$$

where each conditional probability $p(Y_{ij}|\mathbf{N}_{ij}^{(Y)})$ can be obtained by inputting $\mathbf{N}_{ij}^{(Y)}$ into the trained 'SSAE+Softmax' classifier. The microstructure sample $\mathbf{Y}$ can be reconstructed through sequential probability sampling by giving an initial guess of $p(Y_{11})$. The effect of initialization can be ignored when the pixel sequence is long enough, therefore, the reconstructed microstructure sample $\mathbf{Y}$ is statistically equivalent to the training image $\mathbf{X}$ in terms of the joint PDF.

Basically, the above procedure is to reconstruct a microstructure sample through generating pixels one by one. By using the data template with multiple central pixels (as shown in Figures 2b and 2c), multiple pixels can be generated at each step, which can greatly speed up the reconstruction process. In these cases, the joint PDF $p(\mathbf{Y})$ can be calculated as:

$$
p(\mathbf{Y}) = p(\mathbf{C}_{11}^{(Y)})p(\mathbf{C}_{12}^{(Y)}|\mathbf{N}_{12}^{(Y)})p(\mathbf{C}_{13}^{(Y)}|\mathbf{N}_{13}^{(Y)}) \cdots p(\mathbf{C}_{ij}^{(Y)}|\mathbf{N}_{ij}^{(Y)}) \cdots
\tag{19}
$$

where $\mathbf{C}_{ij}^{(Y)}$ denotes the multiple central pixels in the image $\mathbf{Y}$, and $p(\mathbf{C}_{ij}^{(Y)}|\mathbf{N}_{ij}^{(Y)})$ can be obtained from the trained 'SSAE+Softmax' classifier by inputting $\mathbf{N}_{ij}^{(Y)}$.

The algorithmic workflow of the proposed characterization and reconstruction for 2D microstructures is summarized in Algorithm 1. The 4-central-pixel data template is used as the example to describe the procedure of microstructure characterization and reconstruction, and a schematic illustration is provided in Figure 7.

3.2 Boundary effect

As shown in Figure 7b, the boundary pixels of the initial image $\mathbf{Y}_0$ do not have enough neighboring pixels as the inputs for the trained 'SSAE+Softmax' classifier model, so their phase values will not be updated during the reconstruction process. However, these boundary pixels still play as neighboring pixels for the corresponding inner pixels, which will cause a negative effect on microstructure reconstruction, as shown in Figure 8b. To tackle the boundary effect problem, a simple method is to remove the boundary and affected area from the reconstructed image and only pick the central part as the final result, as shown in Figure 8b. In this study, a more rational approach using periodic/reflected boundary conditions is proposed to overcome the boundary

---
Algorithm 1: Stochastic characterization and reconstruction of 2D heterogeneous microstructures
---

Microstructure Characterization:

Data: The training image X of size $s_1 \times s_2$;

Data: The selected 4-central-pixel data template with neighboring radius equal to r.

while $r + 1 \leq i \leq s_1 - 1$ and $r + 1 \leq j \leq s_2 - r - 1$ **do**

> Collect the data event (C_{ij}, N_{ij}) from the traing image X by using the selected data template;
>
> Convert the multiple central pixels C_{ij} to corresponding label data L_{ij};
>
> Move the data template to the next step in raster scan order.

end

Train the 'SSAE+Softmax' classifier model M by using the training data (L, N);

Result: The trained 'SSAE+Softmax' classifier model M.

Microstructure Reconstruction:

Data: The trained 'SSAE+Softmax' classifier model M;

Data: The selected 4-central-pixel data template with neighboring radius equal to r;

Data: A 2D lattice gride of size $s_3 \times s_4$.

Initialization: Assign binary white noise to the 2D lattice grid as the initial image Y_0;

while $r + 1 \leq i \leq s_3 - 1$ and $r + 1 \leq j \leq s_4 - r - 1$ **do**

> Extract the data event $(C_{ij}^{(Y_0)}, N_{ij}^{(Y_0)})$ from the initial image Y_0 by using the selected data template;
>
> Input the $N_{ij}^{(Y_0)}$ to the trained 'SSAE+Softmax' classifier model M to obtained the class probability vector $p(L|N_{ij}^{(Y_0)})$;
>
> Genearte a label data L_{ij} by probabilty sampling from the class probability vector $p(L|N_{ij}^{(Y_0)})$;
>
> Convert the label data L_{ij} to the corresponding configuration of multiple pixles $C_{ij}^{(Y)}$;
>
> Update corresponding pixels $C_{ij}^{(Y_0)}$ by using $C_{ij}^{(Y)}$;
>
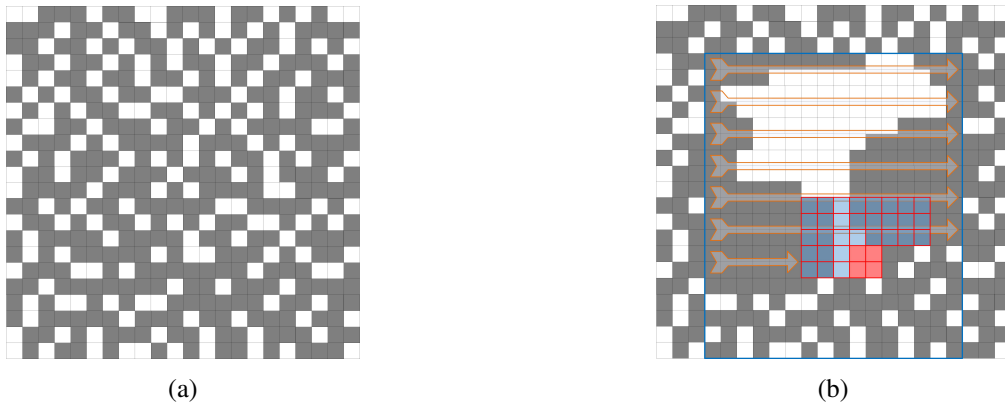> Move the data template to the next step in raster scan order.

end

Result: A statistically equivalent reconstructed microstructure sample Y.

(a)   (b)

Figure 7: Schematic illustration of microstructure reconstruction procedure: (a) A binary white noise image as the initial state; (b) Pixels are updated by using the 4-central-pixel data template to scan the initial image in raster scan order (it should be noted that the boundary pixels outside the blue square will not be updated, because they do not have enough neighboring pixels).

effect problem. Specifically, the boundary of the initial reconstruction is first removed to obtain a defect-free reconstruction (as shown in Figure 8b); then the defect-free sample is used as a tile to build an enlarged copy with periodical/reflected boundaries; and finally the intermediate sample with periodical/reflected boundaries is used as the initialization to reconstruct the final sample, as shown in Figure 8c. This periodical/reflected boundary approach can effectively eliminate the boundary effect problem in the reconstruction.
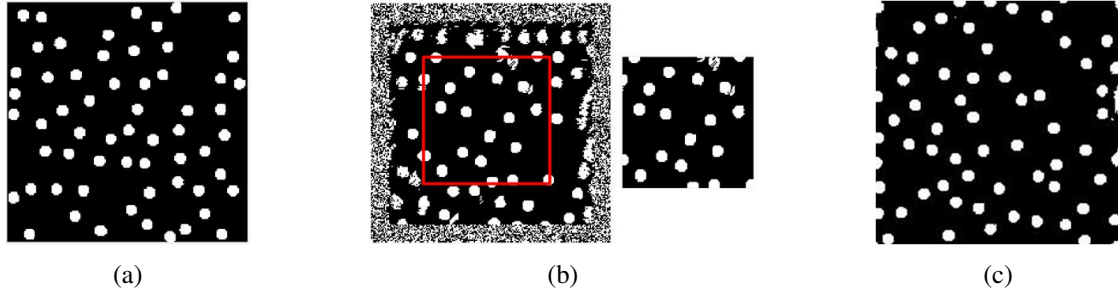

(a) (b) (c)

Figure 8: (a) The training image. (b) The reconstructed image without setting periodical boundary. (c) The reconstructed image with setting periodical boundary.

3.3 Size of data template

The radius r of neighboring pixel zone in the data template is an important parameter affecting the reconstruction quality. A digital image of heterogeneous microstructure can be considered as a spatial signal, so the size of the data template should be large enough to cover the local morphological features. Using the image in Figure 8a as the training image, new samples are reconstructed by using 1-central-pixel data templates with different radii r, and the reconstructed results are shown in Figure 9. The microstructure reconstruction quality is continuously improved before r increasing to a threshold value, and no significant improvement is observed after that. This test confirms that the selection of r has a significant influence on microstructure reconstruction quality, and it needs to be greater than a minimum value to achieve an accurate reconstruction.
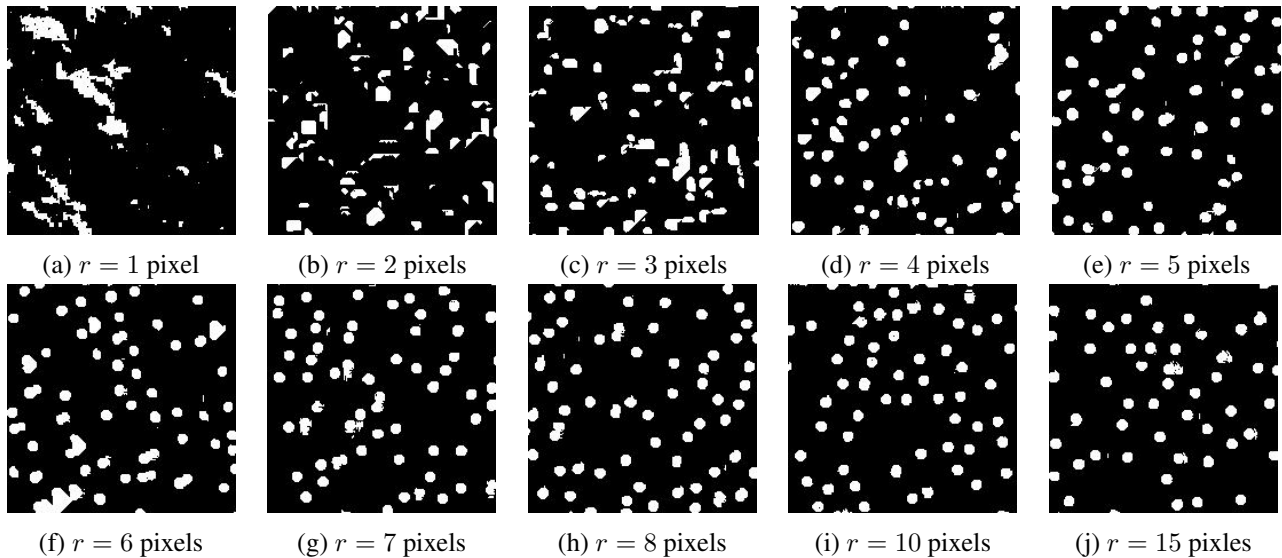

(a) $r = 1$ pixel (b) $r = 2$ pixels (c) $r = 3$ pixels (d) $r = 4$ pixels (e) $r = 5$ pixels

(f) $r = 6$ pixels (g) $r = 7$ pixels (h) $r = 8$ pixels (i) $r = 10$ pixels (j) $r = 15$ pixles

Figure 9: Microstructure reconstruction by using data templates with different radii r.

In our experience, the minimum value of r can be estimated from the two-point correlation function (TPCF) (Torquato, 2013) of the training image. The correlation length a provides a reference value for r, and it can be directly measured from TPCF, as shown in Figure 10. The correlation length a can also be computed by fitting
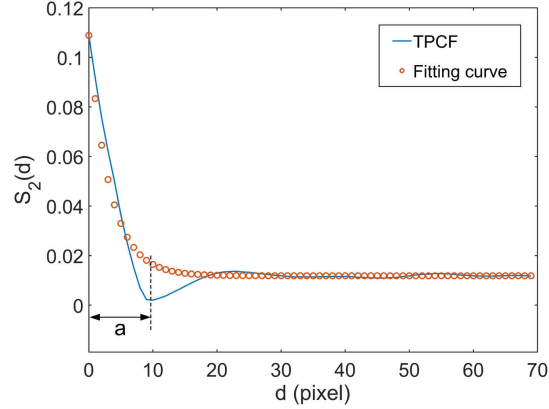
Figure 10: Two-point correlation function and its fitting curve.

the TPCF with an exponential function as defined below (Corson, 1974):

$$S_2(d) = (\phi - \phi^2) \exp(-\frac{3d}{a}) + \phi^2 \tag{20}$$

where $S_2(d)$ denotes TPCF, ϕ is volume fraction of calculated phase, and d is the distance between two points. The correlation length a of the training image in Figure 8a is 9.83 pixels, and the reconstructed result by setting $r = 10$ is visually consistent with the training image, as shown in Figure 9i. As to the 4-central-pixel and 9-central-pixel data template, the minimum sizes of r can also be determined from the correlation length a, which are equal to $a - 0.5$ and $a - 1$, respectively.

4   Results and Discussion

In this section, two groups of microstructures with distinct morphologies are used to examine the proposed statistical characterization and reconstruction method. The new method is also compared with a recent machine learning-based approach (Bostanabad et al., 2016a,b) that uses the decision tree model to learn the stochastic features of heterogeneous media. Different statistical descriptors (Torquato, 2013), including two-point correlation function (TPCF), two-point cluster correlation function (TPCCF) and lineal path function (LPF), are used to assess statistical equivalence between reconstructed and target microstructures. In addition, the time costs for microstructure characterization and reconstruction by using different data templates are also compared. Microstructure characterization and reconstruction methods used in this study are all implemented with MATLAB, and all tests are performed on a PC with an Intel Xeon 3.5GHZ CPU and 32GB memory.

4.1   Microstructures with short-distance correlations

4.1.1   Original microstructures

Representative images of three different microstructures are chosen as the training images, as shown in Figure 11. Each of these microstructures has unique morphology with relatively small size, and the correlation lengths of them are 10, 10, 12 pixels respectively, which can be measured from the TPCFs in Figure 13. The first one is an isotropic microstructure with circular inclusions embedded in the matrix; the second one is an anisotropic microstructure with irregular-shaped inclusions; and the last microstructure is a porous medium with strong randomness in pore shapes and sizes.

For each microstructure, we use all three data templates (as shown in Figure 2) to collect data events from the training image, and then to train 'SSAE+Softmax' classifier models in order to learn the CPDFs of the training image. Finally, each trained 'SSAE+Softmax' classifier model is used to generate a set of microstructures with 50 samples, and these samples are compared with the target microstructures by measuring the statistical

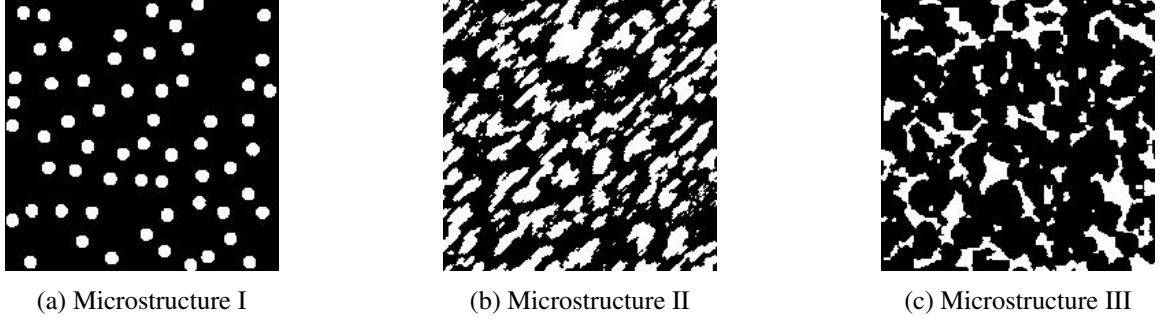|		
(a) Microstructure I	(b) Microstructure II	(c) Microstructure III

Figure 11: The 2D training images of three microstructures with short-distance correlations (image size: 200×200 pixels).

descriptors. The size of the data template is determined from the correlation length of the original microstructure, as described in § 3.3. The parameters of 'SSAE+Softmax' classifier model training are summarized in Table 1. For comparison, Bostanabad's method (Bostanabad et al., 2016a,b) is also used to reconstruct these three microstructures by using data templates with the same sizes.

Table 1: The parameters for 'SSAE+Softmax' classifier model training

Microstructure	Data template	Template size r	Unite number of each layer		
			Input layer	Hidden layer 1	Hidden layer 2
I	1-central-pixel	10 pixels	220	80	40
I	4-central-pixel	10 pixels	240	80	40
I	9-central-pixel	10 pixels	260	120	40
II	1-central-pixel	10 pixels	220	80	40
II	4-central-pixel	10 pixels	240	120	40
II	9-central-pixel	10 pixels	260	140	60
III	1-central-pixel	12 pixels	312	120	50
III	4-central-pixel	12 pixels	336	140	60
III	9-central-pixel	11 pixels	308	140	60

4.1.2 Reconstruction results

As shown in Figure 12, the three columns of images are the representative reconstruction results corresponding to the three target microstructures in Figure 11, by using the proposed method and Bostanabad's method. The first three rows (from top to bottom) of images are reconstructed from the proposed method by using the three data templates in Figure 2, respectively. The last row provides the reconstruction results of Bostanabad's method. Visual inspection of these reconstructed samples implies that the isotropy, anisotropy, and randomness of target microstructures are well preserved, and no obvious discrepancies are observed between the samples reconstructed from the proposed method and Bostanabad's method.

To assess whether these reconstructed samples are statistically equivalent to the original microstructures, we compared the TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and the LPF ($L(d)$) between them. All three morphological descriptors are computed for the white phase in the images, and the results are shown in Figure 13, 14 and 15. Generally, the curves of $S_2(d)$, $C_2(d)$ and $L(d)$ measured from the original microstructures are located inside the red intervals which are the ranges of the statistical descriptors extracted from each set of reconstructed microstructures with 50 samples. The relative errors (quantified by L_2 norm error) between the averages of morphological descriptors from reconstructed samples and that of the original microstructure are listed in Table 2. Each averaged curve calculated from the 50 reconstructed samples is very closed to the curve measured from the original microstructures, with a relative error within 5.00%, and the greatest variation of the statistical descriptor extracted from a particular reconstructed sample is usually less than 10.00%.

More specifically, the reconstructed samples of the first microstructure not only preserve the shape of circular inclusions, but also maintain the spatial distribution of the circular inclusions as evidenced by the good
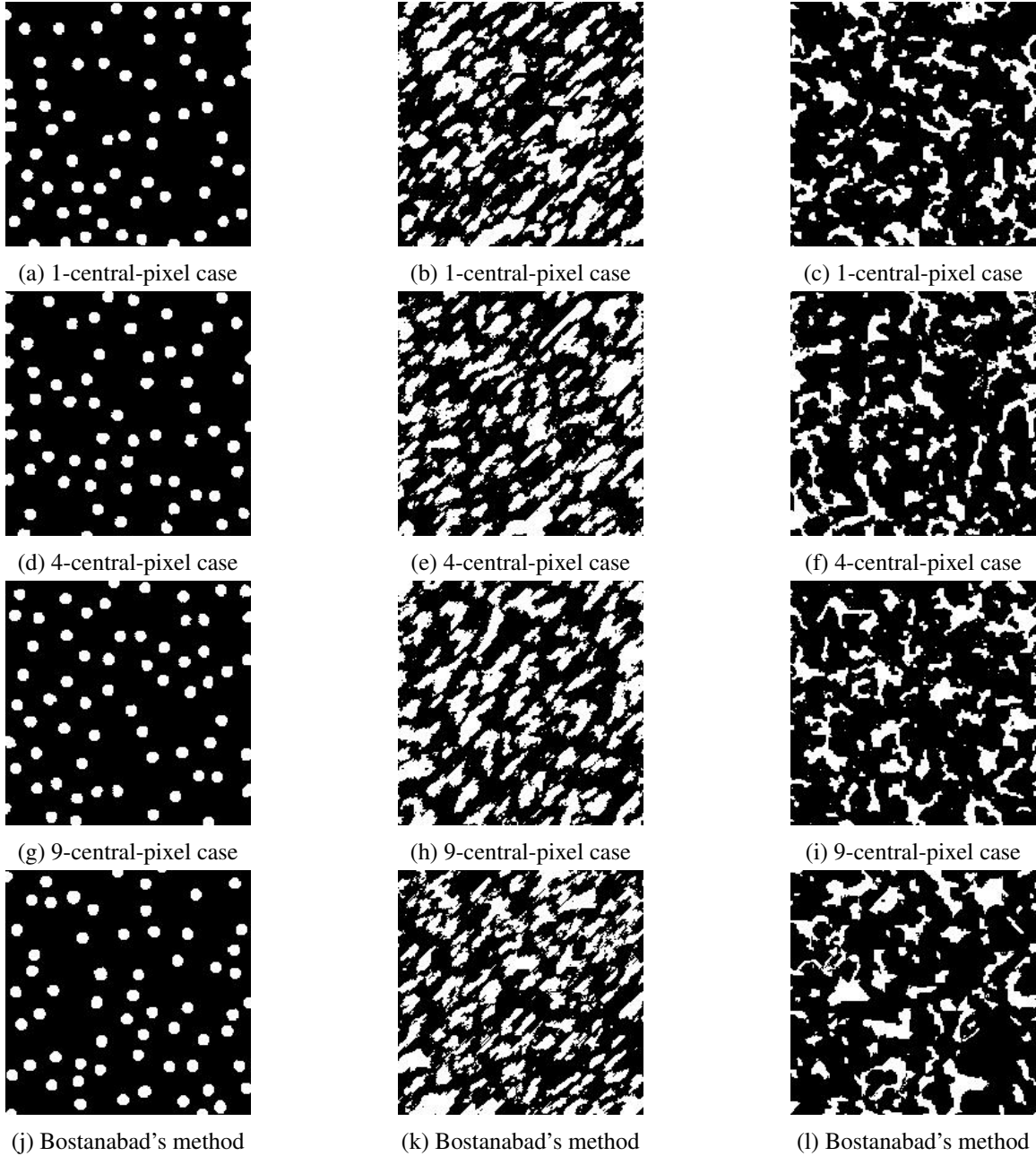
(a) 1-central-pixel case	(b) 1-central-pixel case	(c) 1-central-pixel case
(d) 4-central-pixel case	(e) 4-central-pixel case	(f) 4-central-pixel case
(g) 9-central-pixel case	(h) 9-central-pixel case	(i) 9-central-pixel case
(j) Bostanabad's method	(k) Bostanabad's method	(l) Bostanabad's method

Figure 12: Representative reconstruction results (image size: 200×200 pixels) of microstructure I, II and III by using the proposed method and Bandstand's method.

agreements of statistical descriptors. As to the second microstructure, the reconstructed samples preserve well the anisotropic property and the irregularity of inclusion shapes, which indicates the raster scan order of pixel generation during the reconstruction process does not affect the reconstruction quality. The third microstructure shows the greatest stochasticity in terms of pore shape and pore size, and the trained 'SSAE+Softmax' classifier models are again able to capture stochasticity and complexity of pore geometry and reconstruct statistically equivalent samples.

The reconstructed samples from the proposed method by using different data templates do not differ much visually, and the statistical equivalence shows a similar accuracy level (as shown in Figure 12 and Table 2). However, the efficiency of microstructure characterization and reconstruction by using different data templates varies greatly, as illustrated in Table 2. The time cost of 'SSAE+Softmax' classifier model training (microstructure characterization) increases significantly from using 1-central-pixel template to using 9-central-pixel tem-

plate, while the corresponding time consumption of reconstruction declines notably. Besides the selection of data template, the complexity of the target microstructure also influences the computational cost, because large neighborhood and complicated classifier model are required for such microstructures. Compared with the 1-central-pixel case, the reconstruction using 4- or 9-central-pixel template requires more computer memory, but the reconstruction speed can be enhanced remarkably. In addition, the trained 'SSAE+Softmax' classifier model can be stored and used repeatedly to reconstruct microstructure samples with any size.
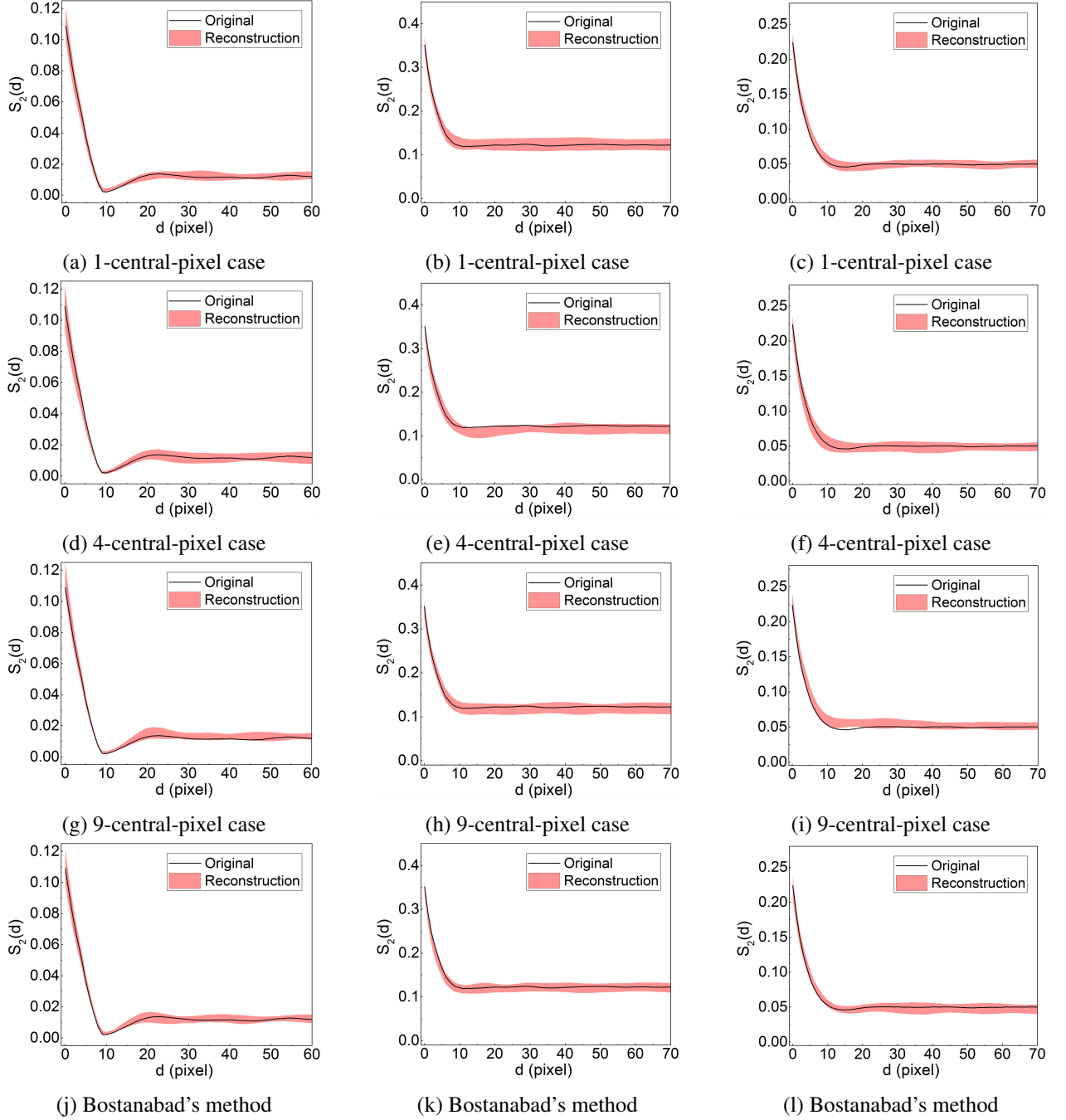

Figure 13: Comparison of TPCF $S_2(d)$ between the original microstructures and the corresponding reconstructed microstructures from the proposed method and Bandstand's method.

For the above three microstructures with short-distance correlations, the proposed method and Bostanabad's method have similar reconstruction accuracy, as recorded in Table 2. Bostanabad's method is more efficient in microstructure characterization, because the decision tree model is much simpler than the 'SSAE+Softmax'
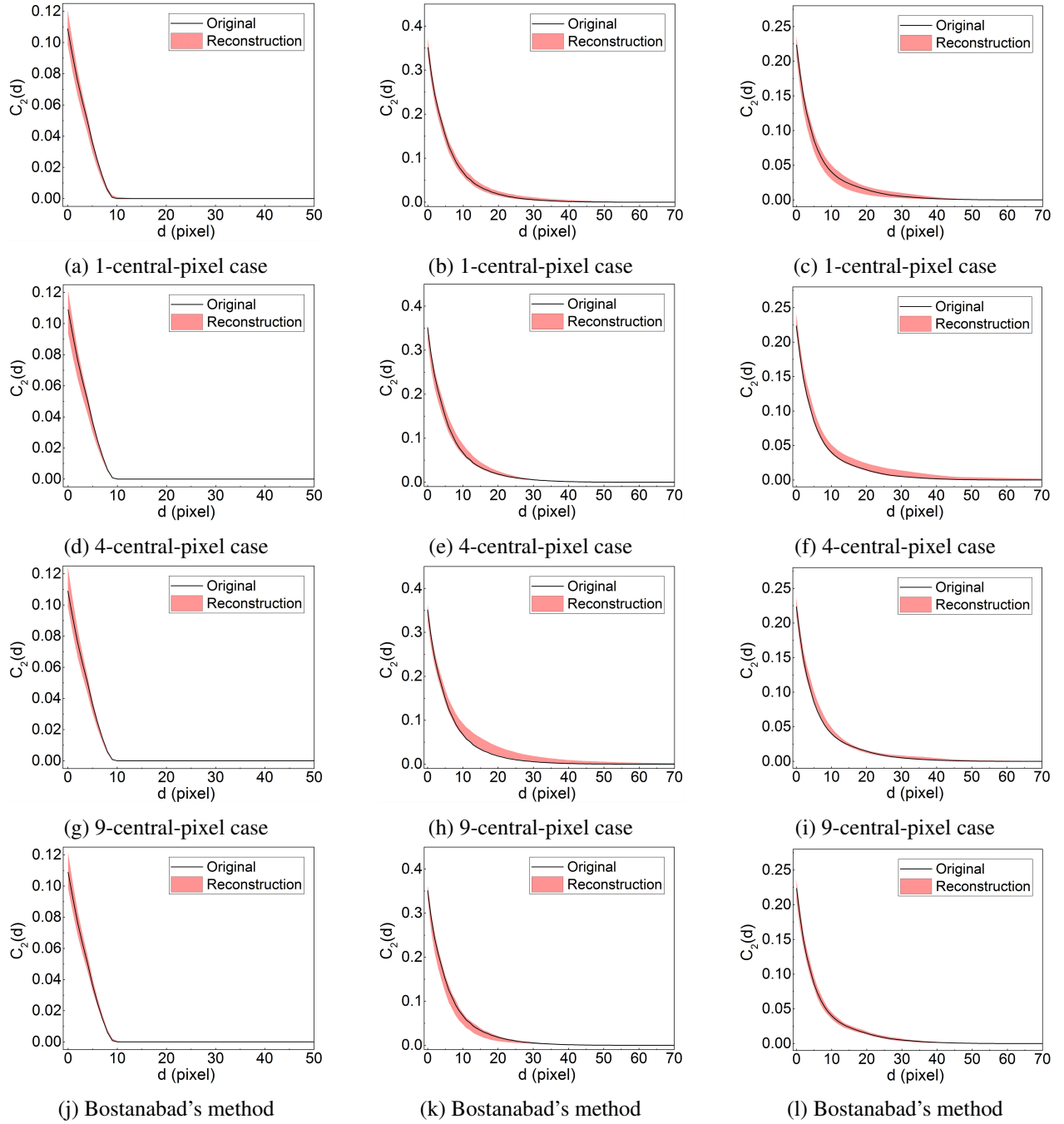
17

Figure 14: Comparison of TPCCF $C_2(d)$ between the original microstructures and the corresponding reconstructed microstructures from the proposed method and Bandstand's method.

classifier model. The proposed method can reconstruct samples much faster than Bandstand's method when the 9-central-pixel template is adapted. Therefore, for the reconstruction of multiple samples, the proposed method will achieve better overall efficiency than Bostanabad's method, and the more samples reconstructed the more time saved. However, the key advantage of the new method is to accurately generate microstructures with long-distance morphological features, which will be discussed in the following section.

4.2 Microstructures with long-distance correlations

This section considers microstructures with long-distance correlations. Again, both the proposed reconstruction method and Bostanabad's method are tested on three microstructures with distinct morphological features.
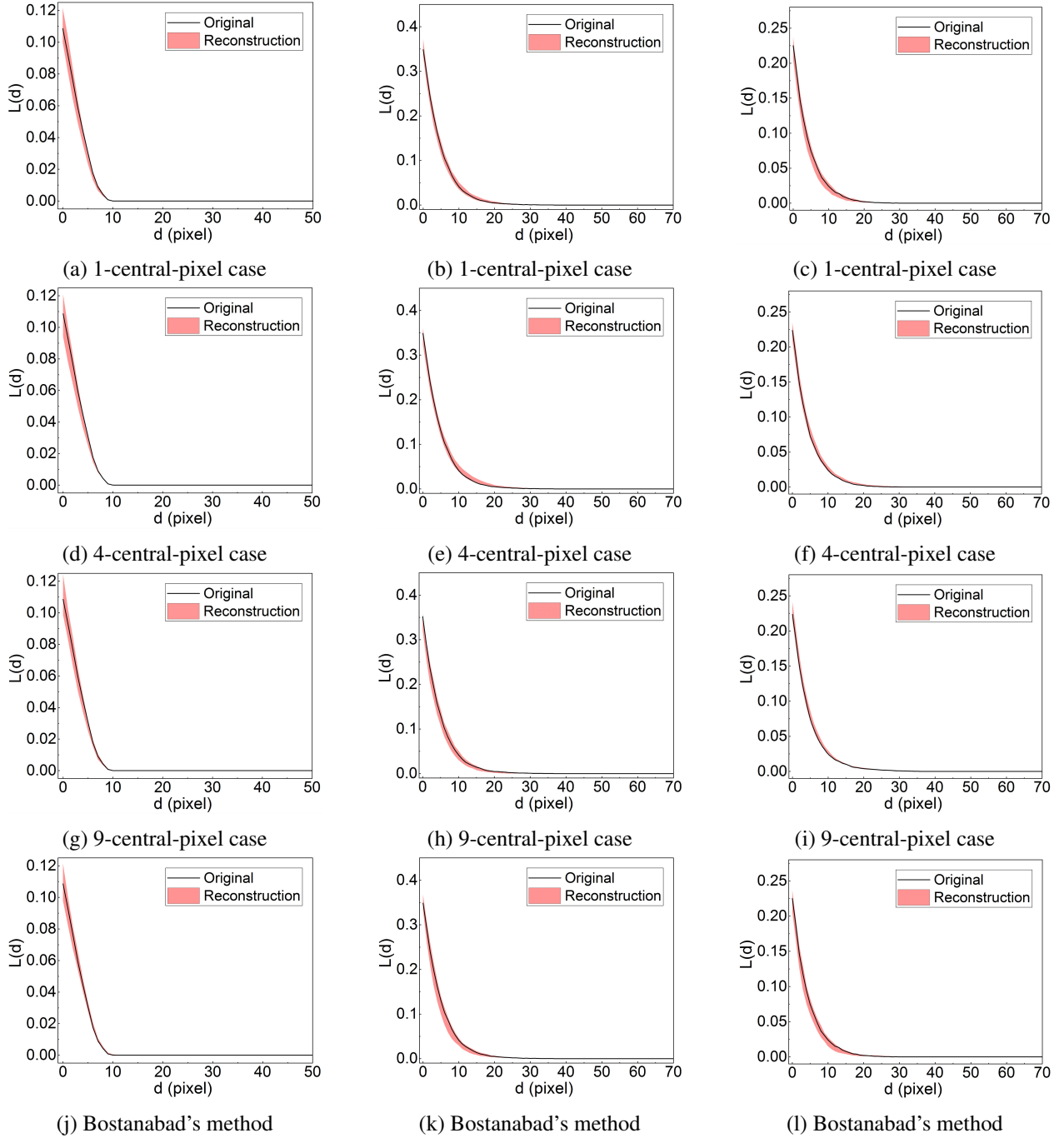
18

(a) 1-central-pixel case   (b) 1-central-pixel case   (c) 1-central-pixel case

(d) 4-central-pixel case   (e) 4-central-pixel case   (f) 4-central-pixel case

(g) 9-central-pixel case   (h) 9-central-pixel case   (i) 9-central-pixel case

(j) Bostanabad's method   (k) Bostanabad's method   (l) Bostanabad's method

Figure 15: Comparison of LPF $L(d)$ between the original microstructures and the corresponding reconstructed microstructures from the proposed method and Bandstand's method.

Table 2: Average reconstruction errors in volume fraction (ϕ), TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and LPF ($L(d)$), as well as the average computational costs of microstructure characterization and reconstruction.

Microstructure	Data template/ Method	$|\Delta\phi|$	$|\Delta S_2(d)|$	$|\Delta C_2(d)|$	$|\Delta L(d)|$	Average CPU time (s)	
						Characterization	Reconstruction (1 sample)
I	1-central-pixel	2.55%	2.83%	1.34%	1.17%	161.5	198.9
I	4-central-pixel	3.15%	2.53%	1.25%	0.90%	336.1	52.6
I	9-central-pixel	2.76%	3.21%	1.51%	1.52%	529.2	26.9
I	Bandstand's method	2.69%	2.74%	1.36%	1.28%	15.2	41.4
II	1-central-pixel	3.90%	2.38%	1.85%	1.21%	139.7	166.4
II	4-central-pixel	3.58%	3.10%	2.92%	1.73%	290.6	53.7
II	9-central-pixel	3.67%	3.45%	2.31%	1.82%	1293.5	24.7
II	Bandstand's method	3.49%	2.46%	2.11%	1.32%	19.1	62.8
III	1-central-pixel	3.43%	2.90%	2.84%	1.41%	116.2	169.1
III	4-central-pixel	3.07%	2.38%	1.67%	1.56%	270.2	55.4
III	9-central-pixel	2.58%	3.41%	1.07%	1.77%	848.8	29.2
III	Bandstand's method	2.97%	3.06%	1.19%	1.84%	18.7	54.1

4.2.1 Original microstructures

As shown in Figure 16, three distinct microstructures with large-sized morphological features are tested. The largest inclusions in microstructure IV and microstructure V exceed the half size of training images, and the wavy stripes in microstructure VI even penetrate through the whole training image. For microstructure IV and microstructure V, both the proposed method with three different data templates and Bostanabad's method are used for statistical characterization and reconstruction. The parameters of 'SSAE+Softmax' classifier model training are listed in Table 3. The data template sizes are determined by the correlation lengths, and for a fair comparison, data templates of the same sizes are used by Bostanabad's method.
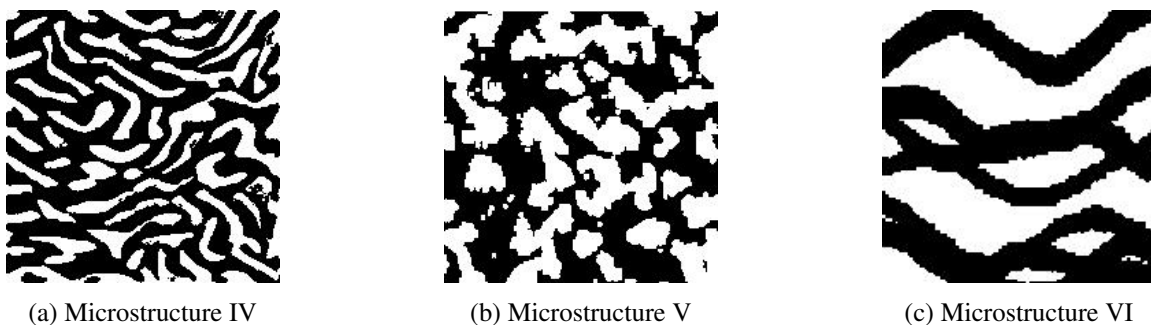

(a) Microstructure IV   (b) Microstructure V   (c) Microstructure VI

Figure 16: The 2D training images of three microstructures with long-distance correlations (image size: 200 $\times$ 200 pixels).

Table 3: The parameters for 'SSAE+Softmax' classifier model training

Microstructure	Data template	Template size r	Unite number of each layer		
			Input layer	Hidden layer 1	Hidden layer 2
IV	1-central-pixel	10 pixels	220	80	40
IV	4-central-pixel	10 pixels	240	80	40
IV	9-central-pixel	10 pixels	260	120	60
V	1-central-pixel	20 pixels	840	200	80
V	4-central-pixel	19 pixels	798	200	80
V	9-central-pixel	19 pixels	836	240	80

As to microstructure VI, the multi-level version of the proposed method is adopted to capture the super large morphological features. The multi-level approach in this study is specifically developed to generate 3D porous media samples that preserve long-distance connectivity of pore networks, and its reconstruction procedures are described in detail in §5.2. Microstructure samples are reconstructed in three levels, and 1-, 4- and 4-central-

pixel templates are used in the first, second and third level of reconstruction respectively, where the template sizes are 4, 8, and 14 pixels respectively.

4.2.2 Reconstruction results

Figure 17 shows the representative reconstruction results corresponding to the three target microstructures in Figure 16. Visually, both the local morphologies and long-distance connectivities of the original microstructures are well preserved by the reconstructed samples from the proposed method, while the performance of Bostanabad's method is poorer, especially for microstructure VI.

(a) 1-central-pixel case   (b) 1-central-pixel case   (c) Multi-level approach

(d) 4-central-pixel case   (e) 4-central-pixel case   (f) Multi-level approach

(g) 9-central-pixel case   (h) 9-central-pixel case   (i) Multi-level approach

(j) Bostanabad's method   (k) Bostanabad's method   (l) Bostanabad's method
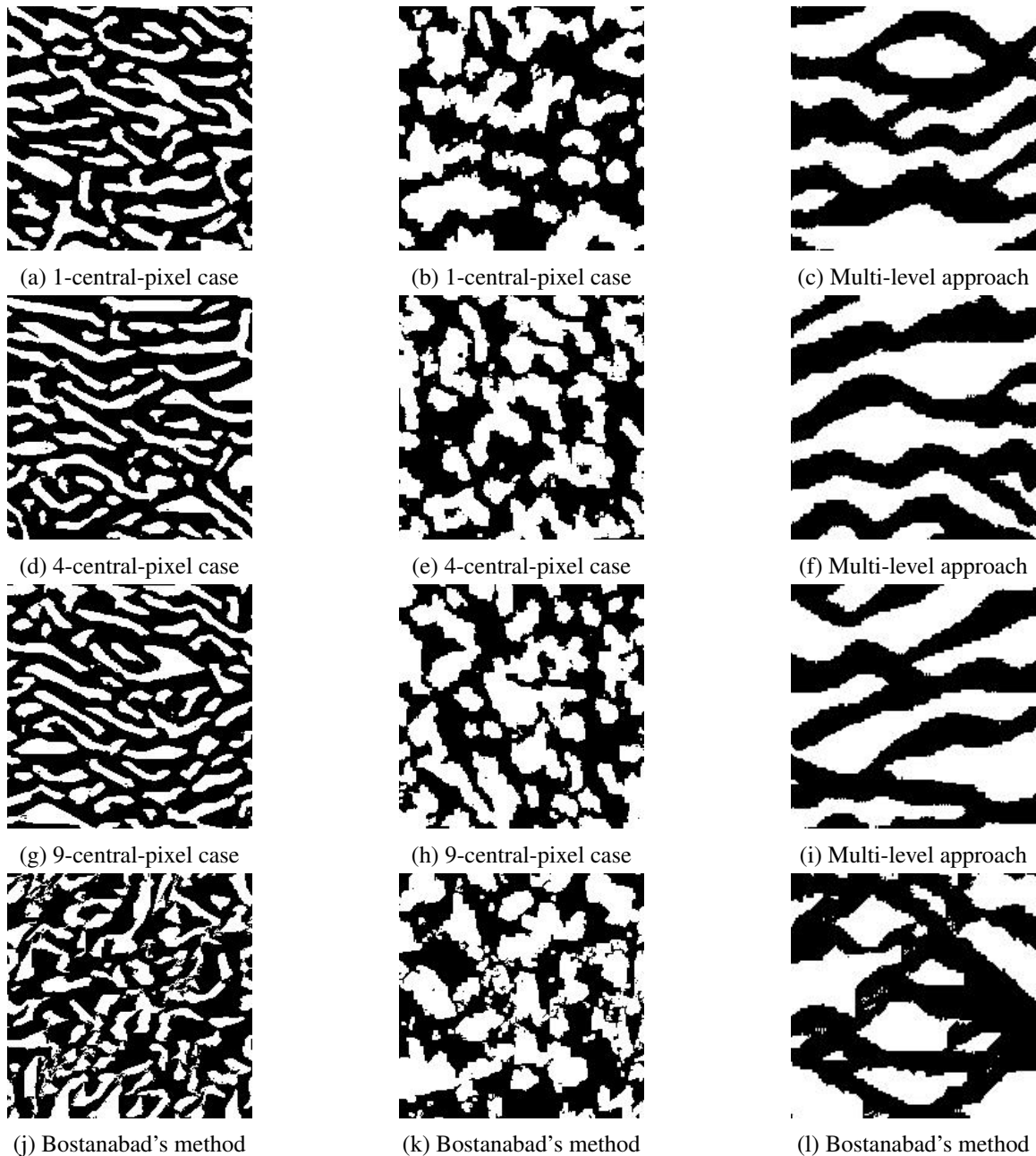
Figure 17: Representative reconstruction results (image size: 200×200 pixels) by using the proposed method and Bandstand's method.

The proposed method well captures the inclusion pattern in microstructure IV as well as its long-distance correlation. The reconstructed samples are also statistically equivalent to the original microstructure measured by TPCF, TPCCF and LPF, as illustrated in Figure 18, Figure 19 and Figure 20, respectively. As to Bostanabad's

method, it approximately captures the long-distance morphology of microstructure IV by adopting a sufficiently large data template, but it loses local morphology information of inclusions. As shown in Figure 17, the inclusions in the reconstructed sample are disordered and very different from the layered pattern in the original sample. This is because the simple decision tree model used by Bostanabad's method cannot capture both the local details and the long-distance morphological features at the same time. The 'SSAE+Softmax' classifier model adopted by the proposed method is capable of preserving more comprehensively morphological features of the microstructure.

(a) 1-central-pixel case

(b) 1-central-pixel case

(c) 4-central-pixel case

(d) 4-central-pixel case

(e) Multi-level approach

(f) 9-central-pixel case

(g) 9-central-pixel case

(h) Bostanabad's method

(i) Bostanabad's method

(j) Bostanabad's method
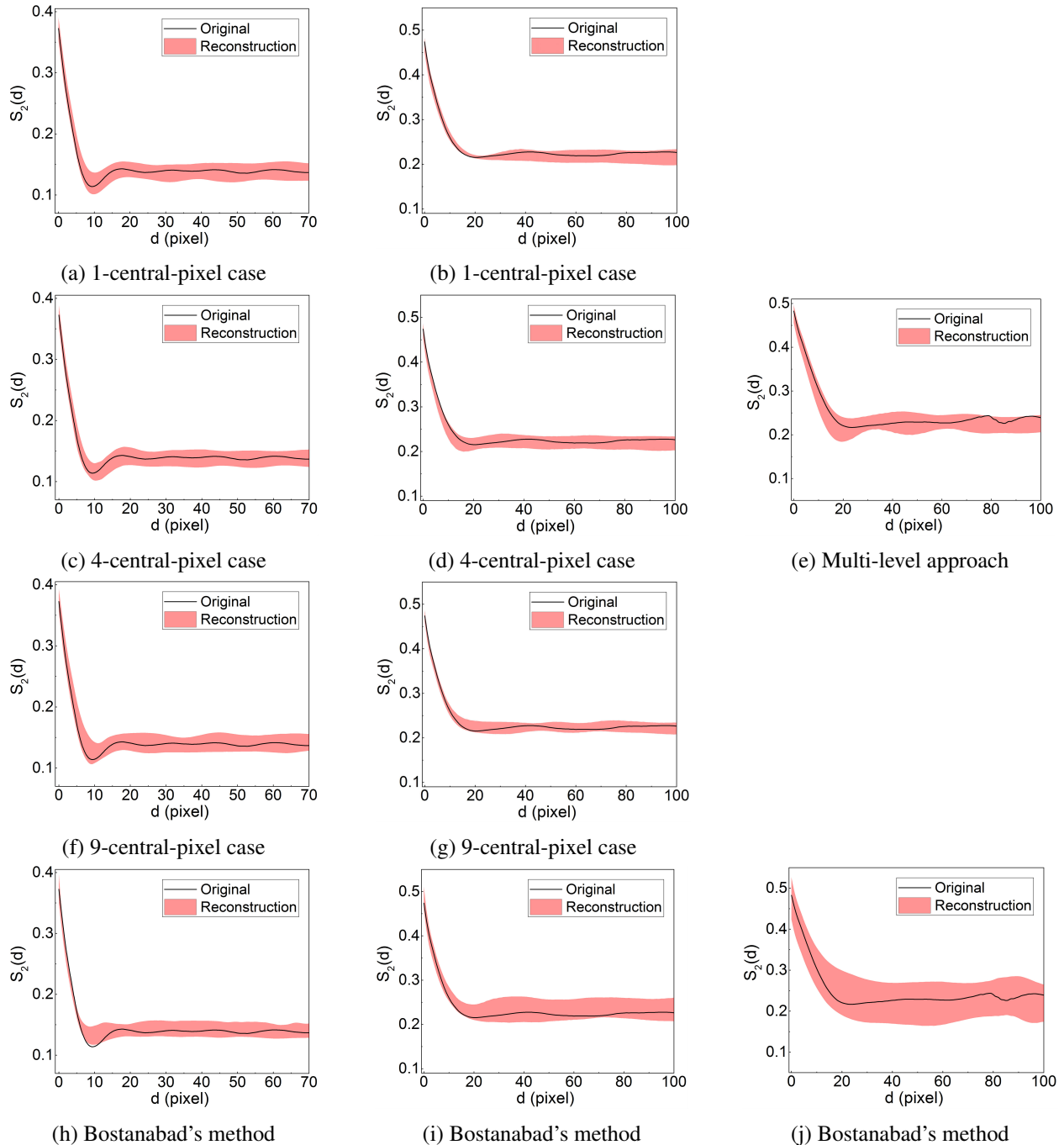
Figure 18: Comparison of TPCF $S_2(d)$ between the original microstructures and the corresponding reconstructed microstructures from the proposed method and Bandstand's method.

The irregularly shaped inclusions in microstructure V are of different sizes, while its morphology exhibits long-distance features. Both the stochastic nature and long-distance morphological features of microstructure V
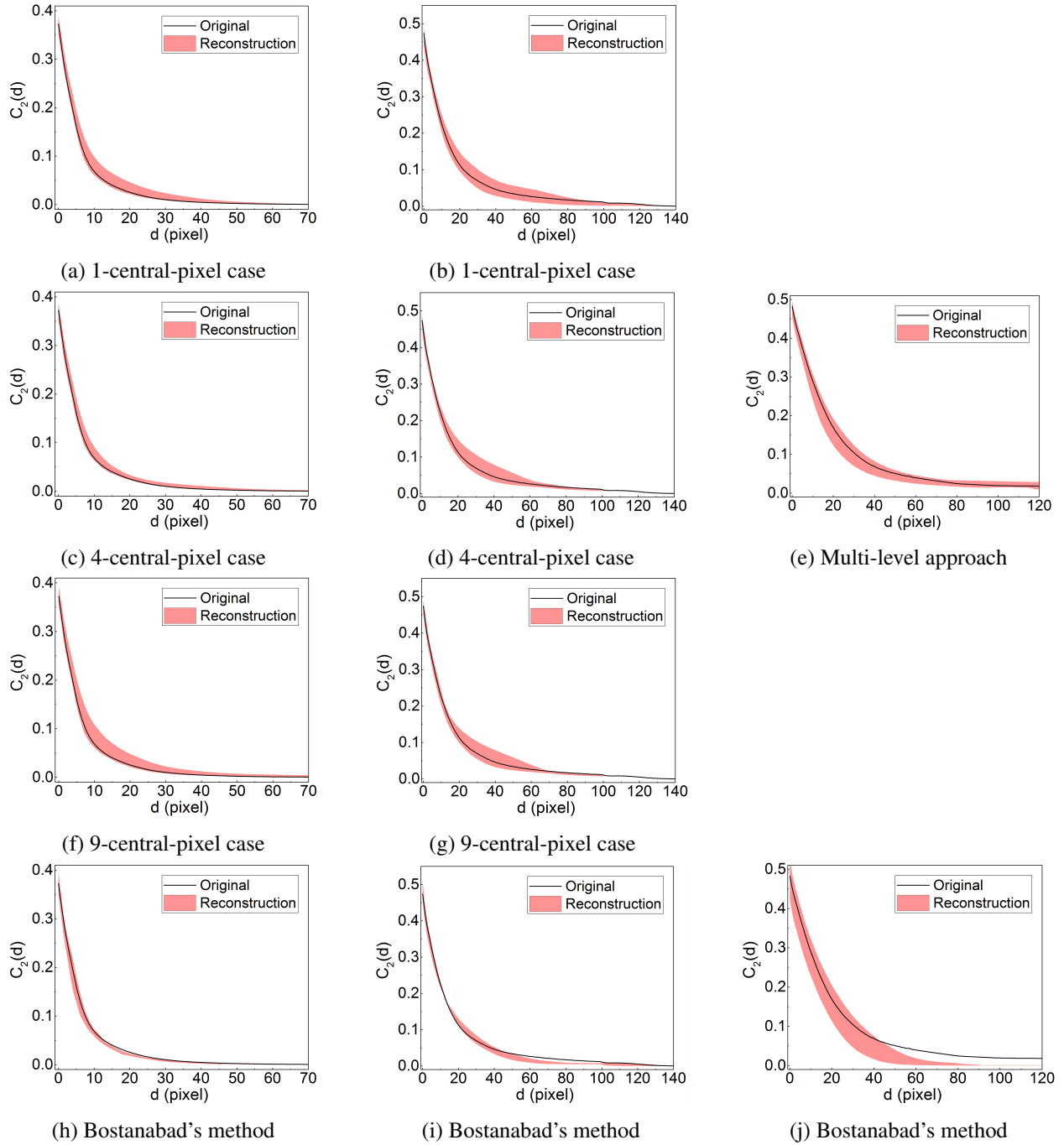
22

(a) 1-central-pixel case

(b) 1-central-pixel case

(c) 4-central-pixel case

(d) 4-central-pixel case

(e) Multi-level approach

(f) 9-central-pixel case

(g) 9-central-pixel case

(h) Bostanabad's method

(i) Bostanabad's method

(j) Bostanabad's method

Figure 19: Comparison of TPCCF $C_2(d)$ between the original microstructures and the corresponding reconstructed microstructures from the proposed method and Bandstand's method.

are well preserved in the samples reconstructed from the proposed method, as shown in Figure 17, Figure 18, Figure 19 and Figure 20. By contrast, the reconstructed samples from Bostanabad's method cannot accurately preserve the long-distance morphological features and as shown in Figure 19 and Figure 20, the TPCCF and LPF of reconstructed samples are below that of the original microstructure. Compared to the 'SSAE+Softmax' classifier model, the decision tree model adopted by Bostanabad's method is less effective to learn and express the complex morphology of microstructure V.

Microstructure VI contains long wavy stripes penetrating through the whole medium, so the multi-level version of the proposed method is used to characterize and reconstruct the microstructure in a hierarchical way. The low-frequency (long-distance) morphology is maintained at the low level of reconstruction, and high-frequency

(a) 1-central-pixel case   (b) 1-central-pixel case

(c) 4-central-pixel case   (d) 4-central-pixel case   (e) Multi-level approach

(f) 9-central-pixel case   (g) 9-central-pixel case

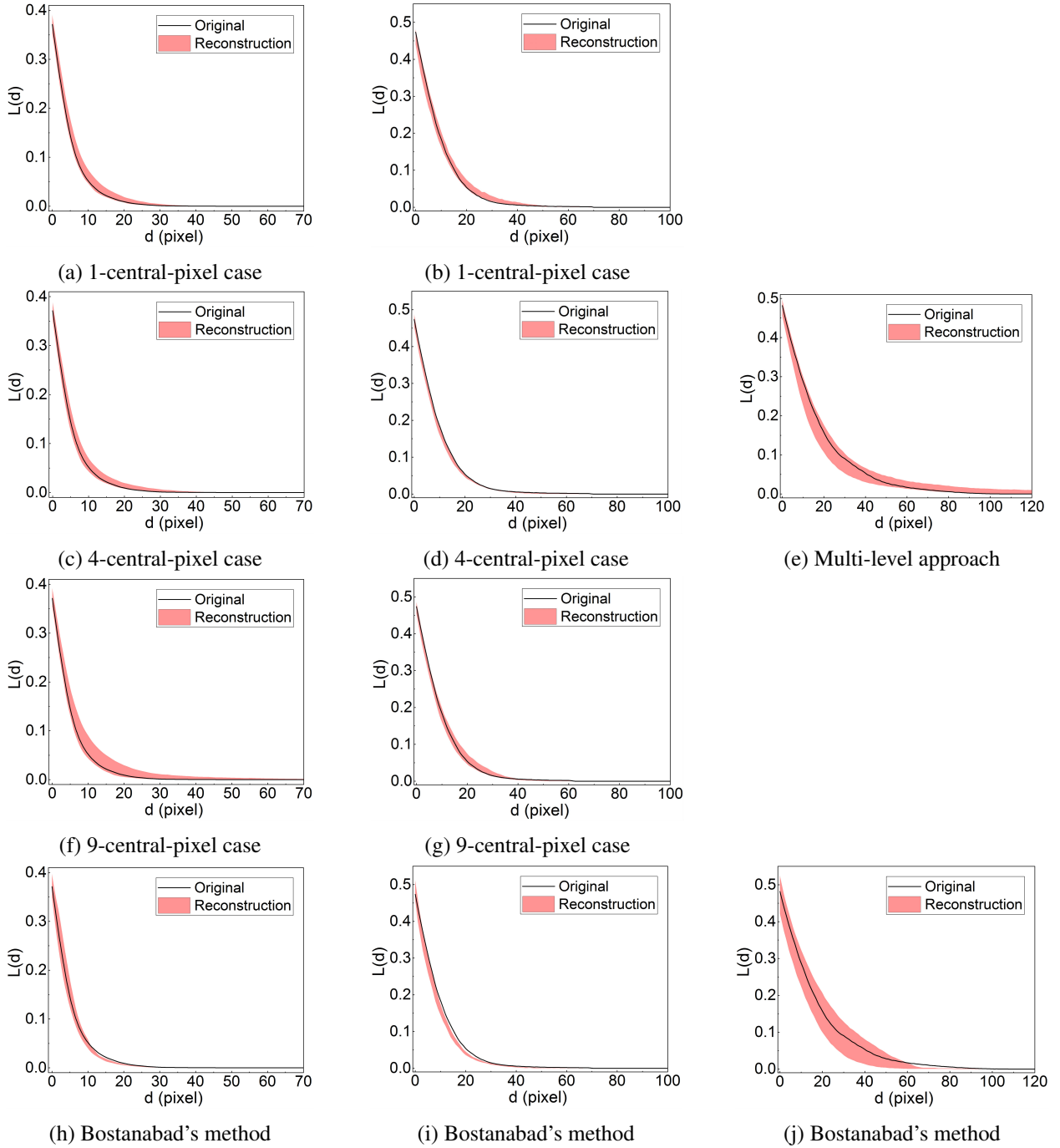(h) Bostanabad's method   (i) Bostanabad's method   (j) Bostanabad's method

Figure 20: Comparison of LPF $L(d)$ between the original microstructures and the corresponding reconstructed microstructures from the proposed method and Bandstand's method.

details are added at the following higher levels of reconstruction. Three representative reconstructed samples are provided in Figure 17, from which one can observe that the long-distance morphological features of microstructure VI are well preserved. The good agreements of TPCF, TPCCF and LPF between the original and reconstructed microstructures also confirm the statistical equivalence between them, as shown in Figure 18, Figure 19 and Figure 20, respectively. As to Bostanabad's method, it fails to capture the long-distance morphology that is far beyond the size of the selected data template, although date templates of different sizes were tried in this study. The disagreements of TPCCF and LPF between the original and reconstructed microstructures also demonstrate the limitation of Bostanabad's method in reconstructing microstructures with long-distance correlations.

Table 4: Average reconstruction errors in volume fraction (ϕ), TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and LPF ($L(d)$), as well as the average computational costs of microstructure characterization and reconstruction.

Microstructure	Data template/ Method	$|\Delta\phi|$	$|\Delta S_2(d)|$	$|\Delta C_2(d)|$	$|\Delta L(d)|$	Average CPU time (s)	
						Characterization	Reconstruction (1 sample)
IV	1-central-pixel	3.68%	3.15%	3.56%	2.40%	127.5	162.8
IV	4-central-pixel	3.35%	3.23%	2.27%	1.73%	284.0	49.6
IV	9-central-pixel	3.72%	4.19%	4.21%	3.62%	930.2	25.1
IV	Bandstand's method	3.59%	5.12%	3.94%	3.48%	17.9	46.5
V	1-central-pixel	4.28%	4.35%	5.56%	2.67%	145.7	205.4
V	4-central-pixel	4.09%	4.43%	4.69%	1.39%	255.3	52.8
V	9-central-pixel	3.84%	3.59%	4.14%	2.36%	1032.7	27.5
V	Bandstand's method	7.71%	8.04%	8.15%	4.25%	18.3	49.3
VI	Multi-level approach	4.82%	5.17%	3.71%	4.51%	465.1	90.2
VI	Bandstand's method	10.31%	12.65%	32.46%	8.74%	19.2	51.4

The average reconstruction errors measured by statistical descriptors are recorded in Table 4, as well as the average time costs of microstructure characterization and reconstruction by using different methods. The training of the 'SSAE+Softmax' classifier model is generally more time-consuming than the training of the decision tree model, but the former model adopted by the proposed methods can better learn and express the complexity of microstructures with long-distance morphologies. Once the 'SSAE+Softmax' classifier model corresponding to a microstructure is properly trained, it can be saved for reuse. Using the 4-central-pixel data template, the proposed method has a similar reconstruction efficiency as Bostanabad's method, while using the 9-central-pixel data template, the reconstruction efficiency of the proposed method gets doubled. In summary, the proposed method has unique advantages in accurately characterizing and efficiently reconstructing heterogeneous microstructures with long-distance correlations.

5 Extension for 3D Microstructures

The methodology presented in § 2 and § 3 can be readily extended to statistically characterize and reconstruct 3D microstructures. The procedures of microstructure characterization and reconstruction are similar to that of 2D microstructures, while the main difference is on the data template. We introduce four data templates for 3D microstructures, as shown in Figure 21. The first two data templates have full neighboring voxels in cubic space, which cover the neighboring voxels that are prior to the central voxel(s) in the raster scan order, and the numbers of central voxels of these two templates are 1 and 8, respectively. Considering memory constraint and computational complexity, the number of neighboring voxels in the 3D data template can be reduced. Another two data templates with partial neighboring voxels are therefore created, and they only cover the neighboring voxels in three orthogonal planes, as shown in Figure 21c and 21d.

Although the partial templates are used for 3D microstructure characterization and reconstruction to reduce the memory requirement, sometimes the size of data template has to be very large in order to capture the largest morphological features in the training image, such as the long-distance connectivity inside the porous microstructures, which may exceed the computational capacity of ordinary computers. To overcome this limitation and efficiently capture long-distance correlations, we propose a multiple-level reconstruction strategy, which is explained in § 5.2.

5.1 Single-level approach

The single-level approach characterizes and reconstructs 3D microstructures in the same procedure as that of 2D microstructures, as described in Algorithm 1. Firstly, the selected 3D data template is used to scan the entire 3D training image $\boldsymbol{X}$ to collect data events $(X_{ijk}, \boldsymbol{N}_{ijk})$ or $(\boldsymbol{C}_{ijk}, \boldsymbol{N}_{ijk})$, and then the 'SSAE+Softmax' classifier model is trained to estimate the CPDF $p(X_{ijk}|\boldsymbol{N}_{ijk})$ or $p(\boldsymbol{C}_{ijk}|\boldsymbol{N}_{ijk})$ of the 3D training image. Finally, 3D statistically equivalent samples can be generated by using the trained 'SSAE+Softmax' classifier model which is considered as an implicit form of CPDF $p(X_{ijk}|\boldsymbol{N}_{ijk})$ or $p(\boldsymbol{C}_{ijk}|\boldsymbol{N}_{ijk})$. A 3D microstructure samples is
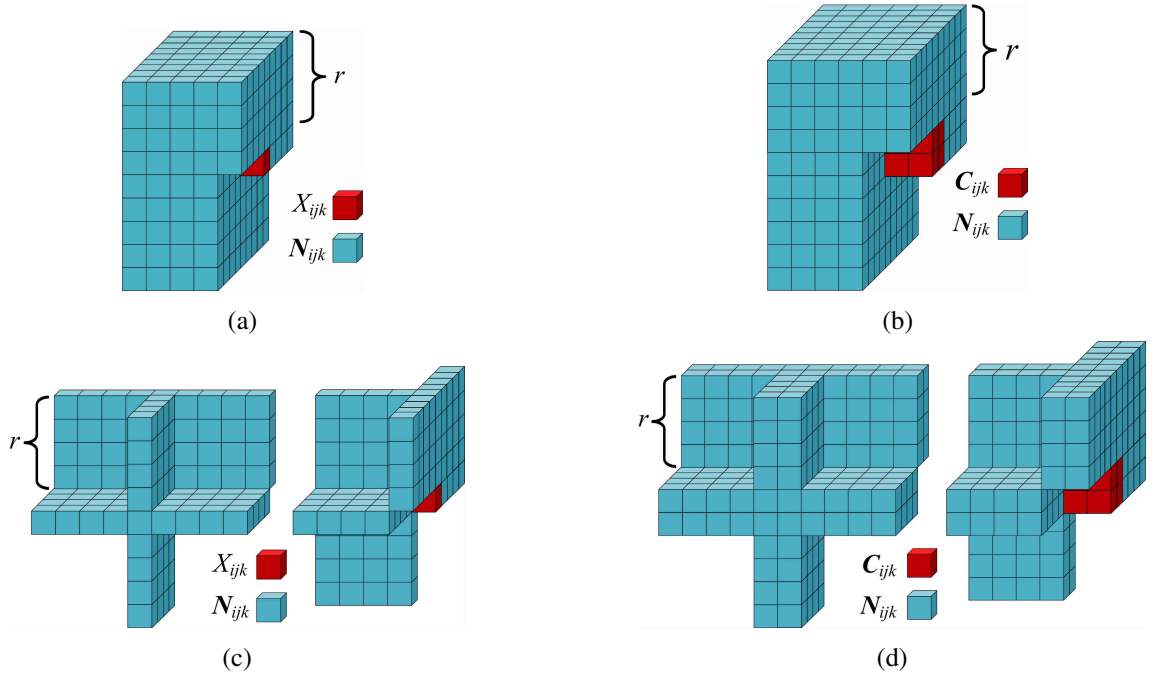
Figure 21: The 3D data templates (neighborhood radius $r = 4$ voxels): (a) 1 central voxel with full neighboring voxels; (b) 8 central voxels with full neighboring voxels; (c) 1 central pixel with partial neighboring voxels; (d) 8 central voxels with partial neighboring voxels.

reconstructed slice by slice, and the voxels on each slice are generated in the raster scan order.

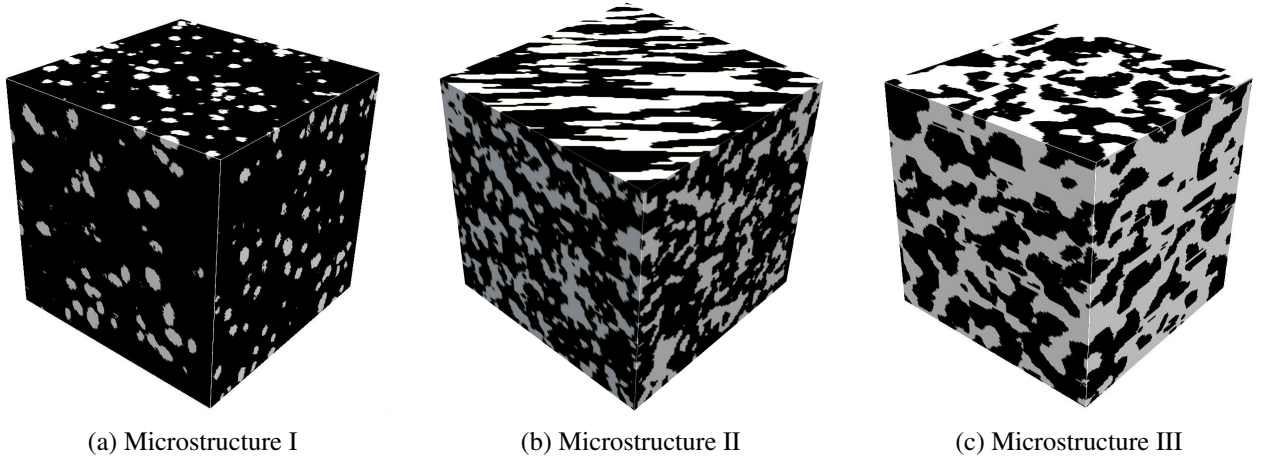### 5.1.1 Original microstructures

(a) Microstructure I (b) Microstructure II (c) Microstructure III

Figure 22: The 3D training images of three microstructures with different morphological features.

In this section, three 3D microstructures with different morphological features are selected to demonstrate the single-level approach, as shown in Figure 22. The first microstructure is a clustered isotropic nanocomposite with silica inclusions randomly embedded in the rubber matrix. The volume fraction of silica inclusion is 9.55%, and the size of the training image is $150 \times 150 \times 150$ voxels. The second microstructure is anisotropic, and its morphological features in orthogonal directions are visibly different in terms of size and shape of the secondary phase. The volume fraction of the secondary phase is 39.16%, and the image size is $100 \times 100 \times 100$ voxels. The last image is a porous microstructure of synthesis silica with pore space (white phase) occupying 42.73% of the bulk volume, which has long-distance morphological features and the image size is 150×150

$\times$ 150 voxels. Different data templates are used to characterize and reconstruct these microstructures, and a set of 50 samples are generated for each case. The statistical equivalence between the original microstructures and the reconstructed samples is evaluated by comparing different morphological descriptors, including TPCF $S_2(d)$, TPCCF $C_2(d)$, and LPF $L(d)$, which are all computed from the white phase in each microstructure. The neighborhood radius of the data template is determined by the correlation length measured from the TPCF, which can be seen in Figure 18. The parameters of the 'SSAE+Softmax' classifier model are listed in Table 5.

Table 5: The parameters for 'SSAE+Softmax' classifier model training

Microstructure	Data template	Template size r	Unite number of each layer		
			Input layer	Hidden layer 1	Hidden layer 2
I	1, full	7 voxels	1687	220	80
I	1, partial	7 voxels	315	100	50
II	1, partial	11 voxels	759	160	70
II	8, partial	10 voxels	1320	220	100
III	1, partial	13 voxels	1053	200	85
III	8, partial	12 voxels	1872	300	120

5.1.2 Reconstruction results

As shown in Figure 23, the three columns of images are the representative reconstructions for the three original microstructures in Figure 22, respectively. These reconstructed samples have the same sizes as the corresponding training images, and they are reconstructed by using different data templates to study the influence of data template on microstructure reconstruction. Visual comparisons between the original microstructures and reconstructed samples imply that the morphological features of the original microstructures, e.g. stochasticity, anisotropy and connectivity, can be well preserved by the reconstructed samples. To evaluate the statistical equivalence between the original microstructure and corresponding reconstructed samples, as well as to quantify the variation of reconstructed samples, morphological descriptors are compared and the results are shown in Figures 24, 25 and 26. The red intervals in Figures 24, 25 and 26 are $S_2(d)$, $C_2(d)$ and $L(d)$ computed from the reconstructed microstructure samples. The black solid curves are the descriptors extracted from original microstructures, and they are all located inside the red intervals. The L_2 norm error between the averages of morphological descriptors extracted from reconstructed samples and that of the original microstructure are usually less than 5.00% (as listed in Table 6), which confirms statistical equivalence is well preserved.

Two data templates with full and partial neighborhoods are used to reconstruct samples for the first microstructure, whose morphological features are relatively simple and in small size. Although the reconstructed sample by using these two data templates are visually similar, the data template covering full neighborhood generates samples with smaller errors and less variation, which can be observed from Figures 24, 25 and 26. The data template covering partial neighborhoods can also produce samples with acceptable error and variation, and it significantly reduces the memory requirement and time cost of microstructure characterization and reconstruction. As to the second and third microstructures, they have long-distance correlation, so the sizes of data templates need to be large enough to capture the main morphological features. The data templates covering partial neighborhoods are adopted to reconstruct samples for them, because these data templates require less memory to operate. No significant difference is observed between the reconstructed samples using partial data templates with 1 and 8 central voxel(s). The reconstructed samples are all visually similar, and the statistical evaluations show similar accuracy level, as shown in Figures 24, 25, 26. According to Table 6, the efficiencies of characterization and reconstruction by using different data templates show great differences. The time cost of microstructure characterization by using 8-central-voxel template is higher than using 1-central-voxel template, while the cost comparison for microstructure reconstruction is opposite. The trained 'SSAE+Softmax' classifier model can be stored for reuse, so microstructure reconstruction by using 8-central-voxel template is time-saving in the long run.
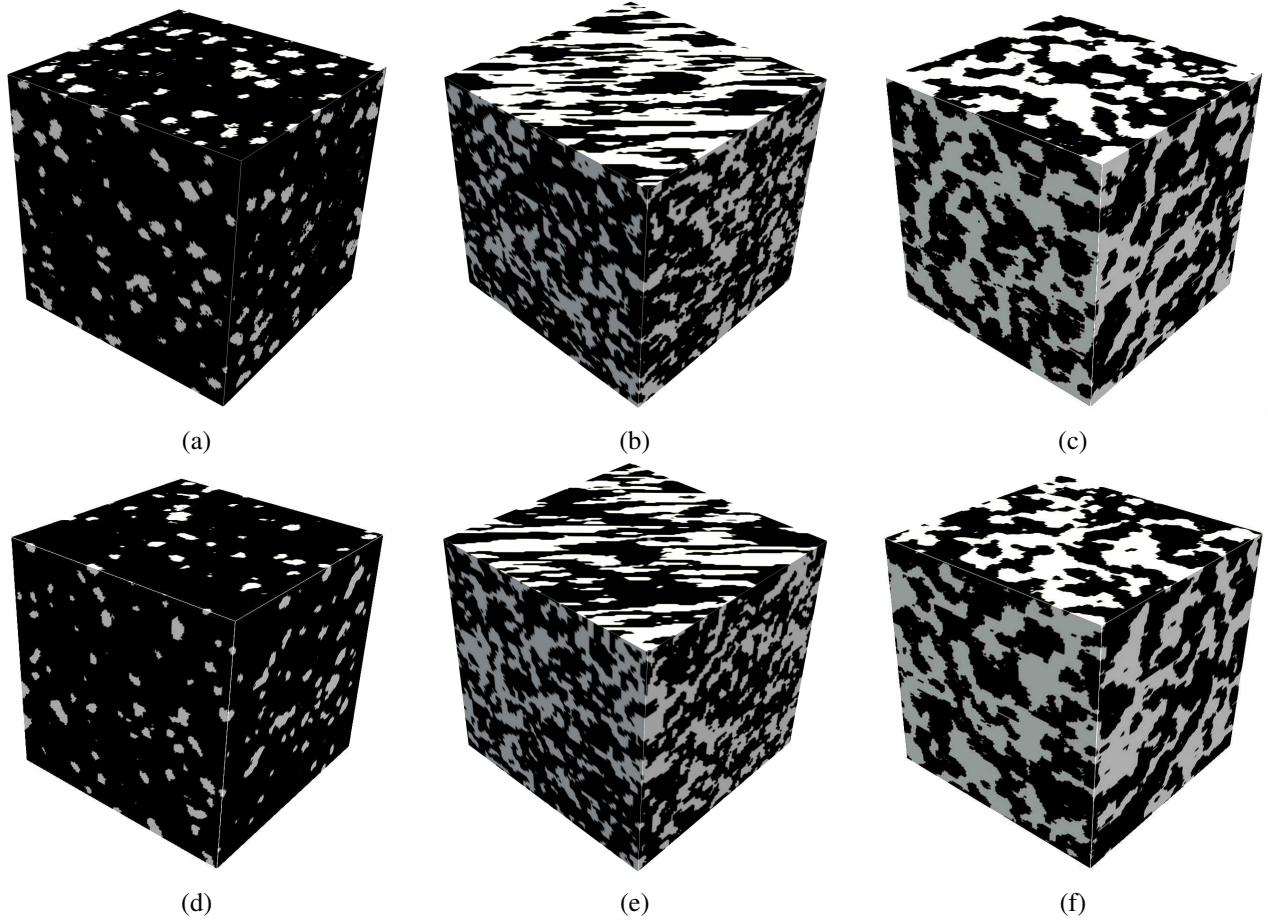
Figure 23: Representative results of microstructure reconstruction by using different data templates: (a) 1 central voxel with full neighborhoods; (b) 1 central voxel with partial neighborhoods; (c) 1 central voxel with partial neighborhoods; (d) 1 central voxel with partial neighborhoods; (e) 8 central voxels with partial neighborhoods; (f) 8 central voxels with partial neighborhoods.

Table 6: Average reconstruction errors of volume fraction (ϕ), TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and LPF ($L(d)$), as well as the average computational costs of microstructure characterization and reconstruction.

Microstructure (Data template)	$|\Delta\phi|$	$|\Delta S_2(d)|$	$|\Delta C_2(r)|$	$|\Delta L(d)|$	Average CPU time (s)	
					Characterization	Reconstruction (1 sample)
I (1, full)	2.25%	2.02%	1.85%	1.21%	10142.2	8029.4
I (1, partial)	2.83%	2.89%	2.49%	2.31%	3669.0	7783.9
II (1, partial)	4.35%	4.77%	4.77%	4.59%	4202.7	3038.4
II (8, partial)	4.07%	4.07%	4.06%	3.88%	6329.4	819.8
III (1, partial)	4.65%	4.58%	4.58%	4.68%	5068.8	8077.4
III (8, partial)	4.88%	4.45%	4.46%	4.23%	12142.5	1809.4

5.2 Multi-level approach

The single-level approach captures the morphological features of microstructures by using data templates of sufficiently large sizes. However, due to limitations in computer memory and computing power, it may not be adequate to capture the long-distance connectivity of the complicated pore network within such pore microstructure as sandstone. In this section, we solve this problem by using a multi-level image pyramid, where the large-scale morphological features can be represented more compactly by using a data template with a smaller size at the lower image pyramid level. The multi-level approach is particularly developed to characterize and reconstruct porous microstructures with long-distance connectivity.
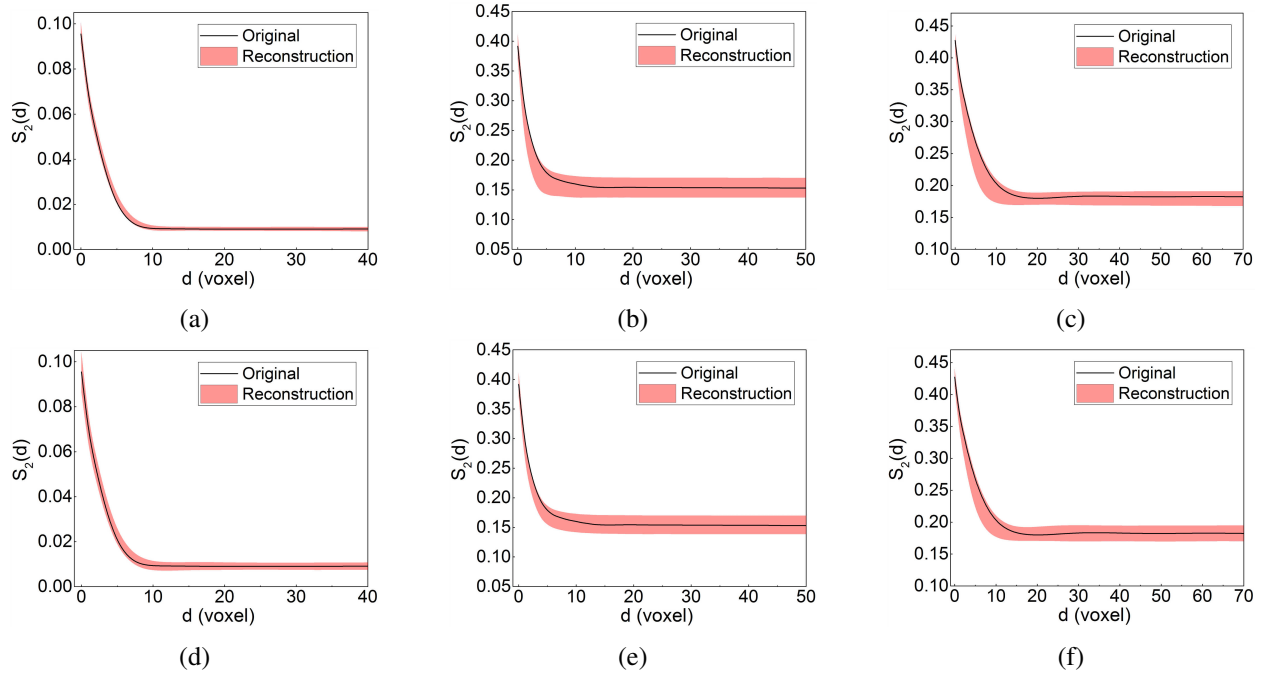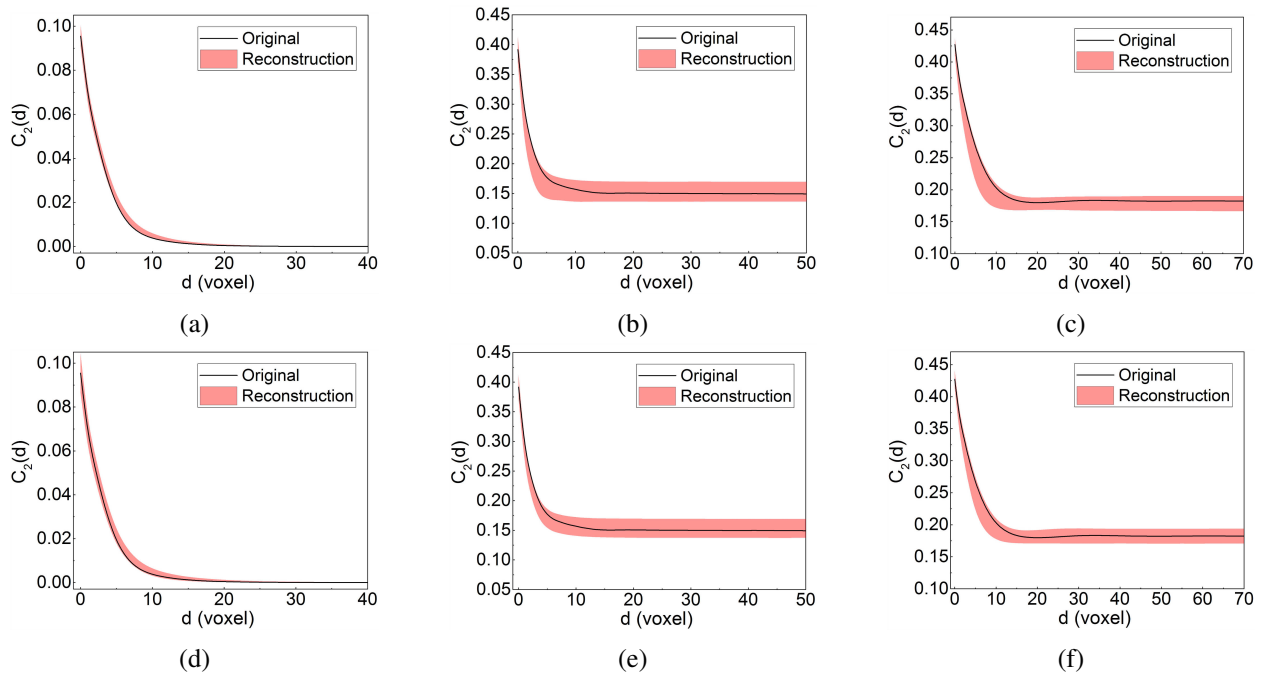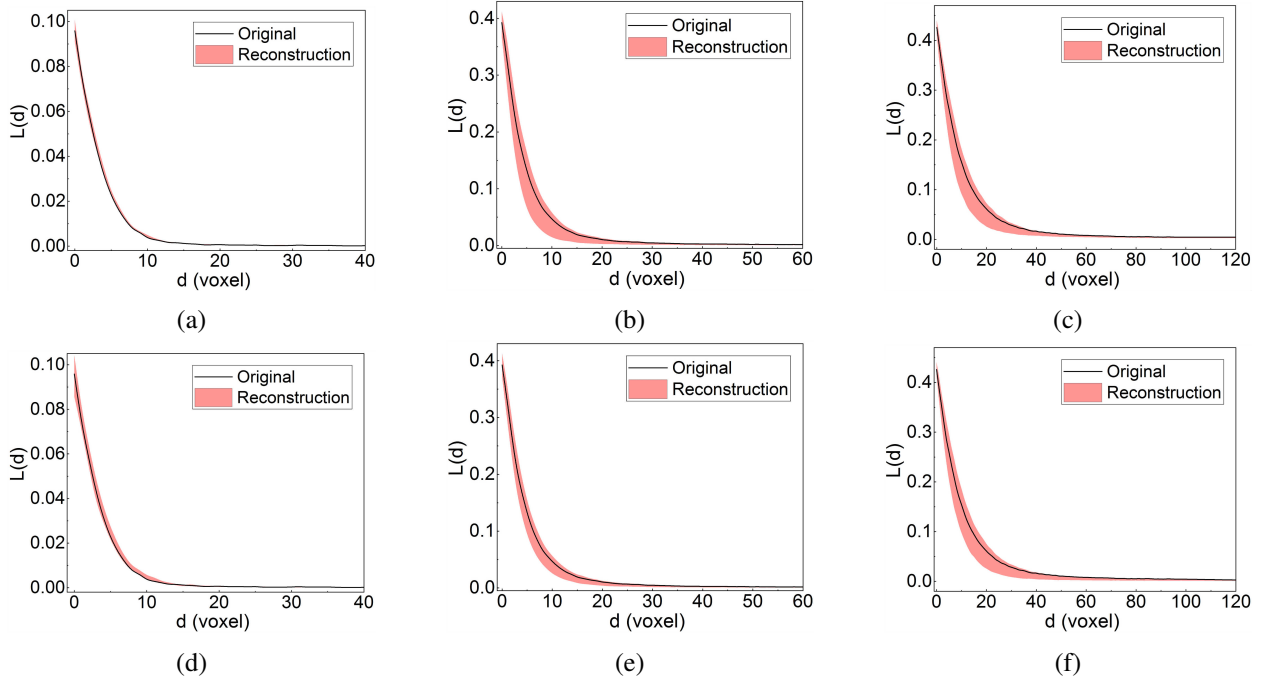
Figure 24: Comparison of TPCF $S_2(d)$ between the original and reconstructed microstructures: (a) 1 central voxel with full neighborhoods; (b) 1 central voxel with partial neighborhoods; (c) 1 central voxel with partial neighborhoods; (d) 1 central voxel with partial neighborhoods; (e) 8 central voxels with partial neighborhoods; (f) 8 central voxels with partial neighborhoods.

Figure 25: Comparison of TPCCF $C_2(d)$ between the original and reconstructed microstructures: (a) 1 central voxel with full neighborhoods; (b) 1 central voxel with partial neighborhoods; (c) 1 central voxel with partial neighborhoods; (d) 1 central voxel with partial neighborhoods; (e) 8 central voxels with partial neighborhoods; (f) 8 central voxels with partial neighborhoods.

Figure 26: Comparison of LPF $L(d)$ between the original and reconstructed microstructures: (a) 1 central voxel with full neighborhoods; (b) 1 central voxel with partial neighborhoods; (c) 1 central voxel with partial neighborhoods; (d) 1 central voxel with partial neighborhoods; (e) 8 central voxels with partial neighborhoods; (f) 8 central voxels with partial neighborhoods.

The three-level approach is taken as an example to explain the procedures to statistically characterize and reconstruct 3D microstructures in a multi-level manner. First, a 3D Gaussian pyramid (Burt and Adelson, 1983) is built in three levels for the original microstructure. The three-level image pyramid plays as the training images for microstructure characterization and reconstruction at different levels, which can be denoted as X_{high}, X_{middle}, and X_{low}, respectively. Then, for the first level, X_{low} is used as the training image, and an 'SSAE+Softmax' classifier model M_{low} is trained to learn the morphological features of X_{low} by using a selected data template. Based on the fitted 'SSAE+Softmax' classifier model M_{low}, a microstructure sample Y_{low} can be generated at the lowest level. Basically, the microstructure characterization and reconstruction at the first level is the same as the single-level approach. Next, for the second level, X_{middle} and X_{low} are used together to serve as the training images. The only modification is that each set of neighboring voxels N not only contains the neighboring voxels in current level but also covers the neighboring voxels and central voxel(s) in the lower level, in order to maintain the morphological consistency between X_{middle} and X_{low}. The second 'SSAE+Softmax' classifier model M_{middle} is trained by using the collected training data, based on which a microstructure sample Y_{middle} can be produced conditional to the reconstructed sample Y_{low} at the lower level. Finally, as to the third level, the procedures are the same as that of the second level, where an 'SSAE+Softmax' classifier model M_{high} at the highest level is trained based on X_{high} and X_{middle}, after which the final microstructure sample Y_{high} can be reconstructed conditional to Y_{middle}.

In summary, the multi-level approach statistically characterizes and reconstructs microstructures in a hierarchical way, where the low-frequency (long-distance) morphological features are captured at a lower level, conditional to which the high-frequency details are added at the higher levels. The same procedures can also be used to characterize and reconstruct 2D microstructures with long-distance correlations, as discussed in §4.2. The algorithmic workflow of the multi-level approach is summarized in Algorithm 2, and a schematic illustration is provided in Figure 28.

Algorithm 2: Stochastic characterization and reconstruction of 3D heterogeneous microstructures using the three-level approach

---

Microstructure Characterization:

Data: The Gaussian pyramid with three levels playing as the training images X_{low}, X_{middle}, and X_{high};

Data: The selected data templates DT_{low}, DT_{middle}, and DT_{high} for different levels.

Collect the data events $(C_{\text{low}}, N_{\text{low}})$ from the training image X_{low} by using the data template DT_{low}, and then train the 'SSAE+Softmax' classifier model M_{low};

Collect the data events $(C_{\text{middle}}, N_{\text{middle}} + N_{\text{low}} + C_{\text{low}})$ from the training image X_{middle} and X_{low} by using the data templates DT_{middle} and DT_{low}, and then train the 'SSAE+Softmax' classifier model M_{middle};

Collect the data events $(C_{\text{high}}, N_{\text{high}} + N_{\text{middle}} + C_{\text{middle}})$ from the training image X_{high} and X_{middle} by using the data templates DT_{high} and DT_{middle}, and then train the 'SSAE+Softmax' classifier model M_{high};

Result: The trained 'SSAE+Softmax' classifier models M_{low}, M_{middle} and M_{high}.

Microstructure Reconstruction:

Data: The trained 'SSAE+Softmax' classifier models M_{low}, M_{middle} and M_{high};

Data: The selected data templates DT_{low}, DT_{middle}, and DT_{high} for different levels.

Reconstruct the first-level microstructure Y_{low} based on the trained 'SSAE+Softmax' classifier model M_{low};

Conditional to Y_{low}, reconstruct the second-level microstructure Y_{middle} based on the trained 'SSAE+Softmax' classifier model M_{middle};

Conditional to Y_{middle}, reconstruct the third-level microstructure Y_{high} based on the trained 'SSAE+Softmax' classifier model M_{high};

Result: A statistically equivalent reconstructed microstructure sample Y_{high}.

5.2.1 Original microstructures

Fontainebleau sandstone is chosen to test the performance of the proposed multi-level approach to statistically characterize and reconstruct pore microstructures, and its reconstruction results are also compared with that of Bostanabad's method (Bostanabad et al., 2016a,b). Fontainebleau sandstone (Øren and Bakke, 2002) is an isotropic porous medium, which is often used as the reference standard for validating models of porous media due to its special properties. It consists of mono-crystalline quartz grains with an average size of about 200 μm and it does not contain clay, inside which only inter-granular porosity exists. A group of 3D digital microstructures (resolution: 5.7μm) of three Fontainebleau sandstone samples are used, whose porosities are 15.6%, 20.9%, and 24.5%, respectively (Berg, 2014). The digital microstructures are shown in Figure 27, where the white region denotes the pore space and the black region denotes the solid matrix. These digital microstructures are used as training images for stochastic characterization and reconstruction.

The three-level approach is used to characterize and reconstruct the microstructures of Fontainebleau sandstones, to capture the long-distance connectivity as well as the geometrical complexity of the pore networks. For the first level, the data template DT_{low} covering 1 central voxel and partial neighboring voxels is selected, and the radius of DT_{low} is 7 voxels which can be measured from the TPCFs $S_2(d)$ of the training images X_{low}. As to the second level, the training images X_{middle} become finner than X_{low}, and the data template DT_{middle} covering 8 central voxels and partial neighboring voxels is chosen, because 8 voxels in the 3D image of current level covers the same domain of 1 voxel in the 3D image of the lower level. Limited by the computing capacity, the radius of DT_{middle} is also 7 voxels in this study. For the third level, the data template DT_{high} covering 8 central voxels and partial neighboring voxels is used, and the radius of DT_{high} is set to be 6 voxels. In additional, the data template with partial neighborhoods is use for Bostanabad's method, and the size of data template is determined by the correlation length.
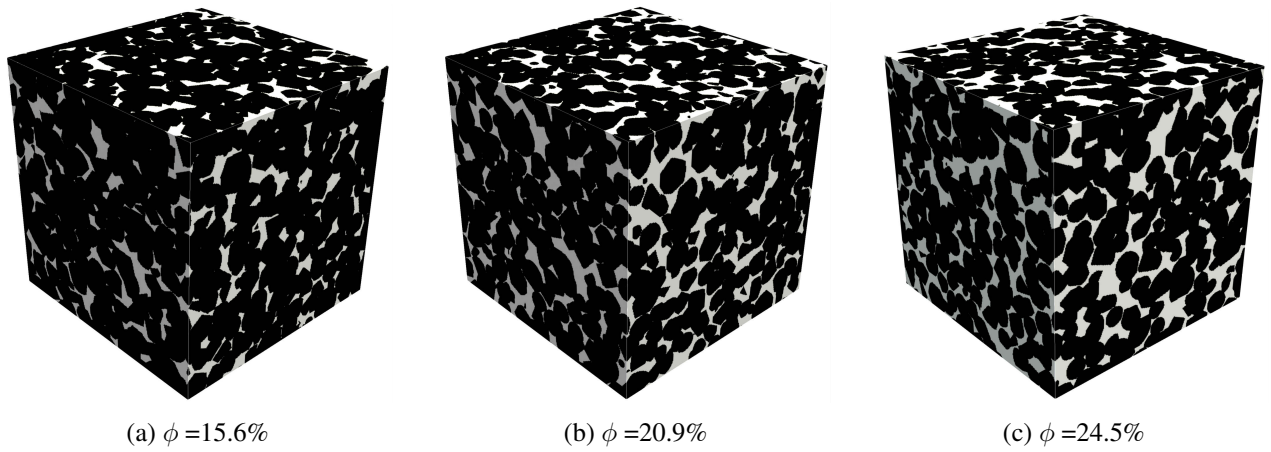
(a) ϕ =15.6%   (b) ϕ =20.9%   (c) ϕ =24.5%

Figure 27: The 3D training images (320×320×320) of Fontainebleau sandstone microstructures with different porosities ϕ.
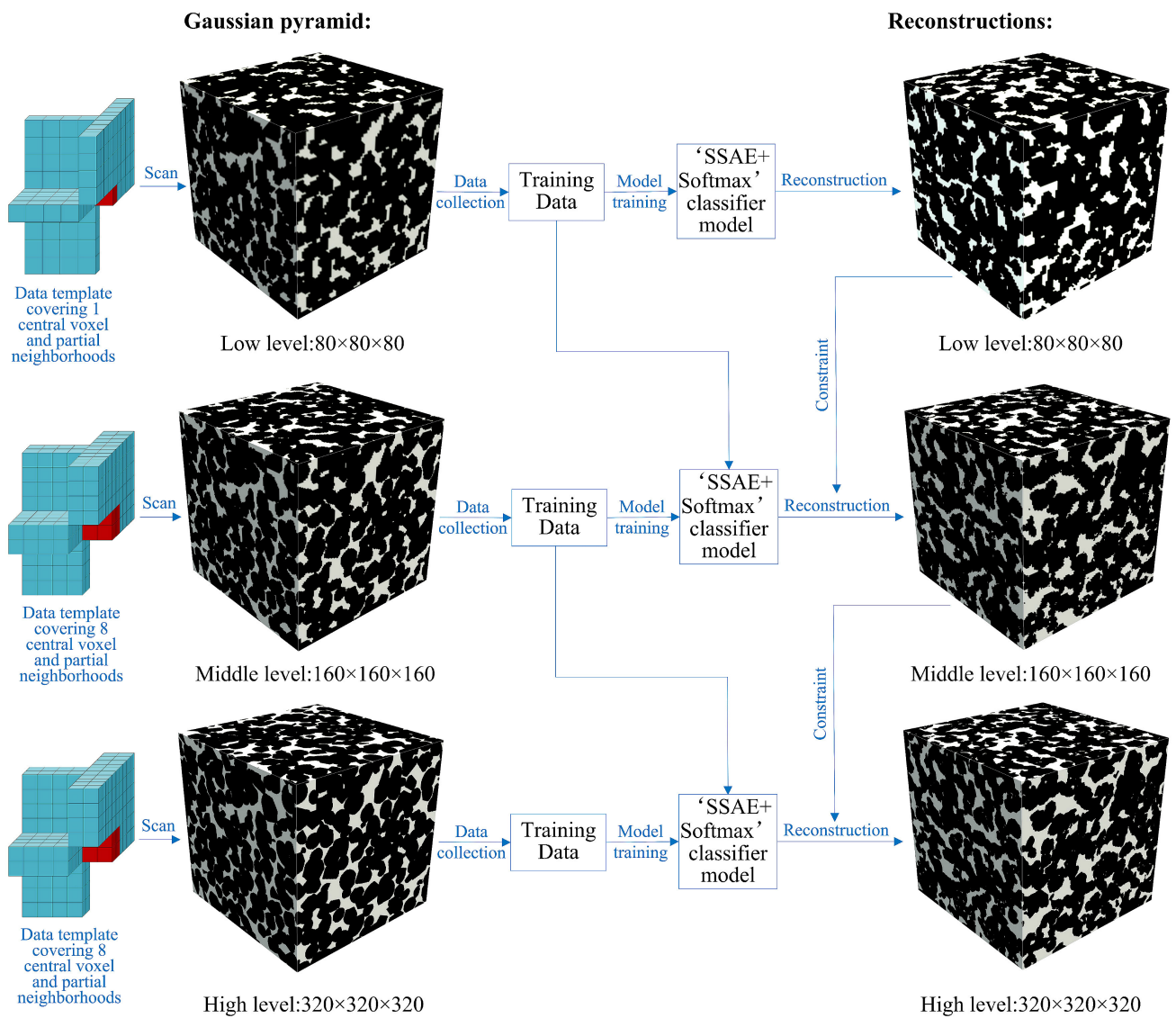

Figure 28: Schematic diagram of microstructure characterization and reconstruction using three-level approach.

5.2.2 Reconstruction results

For each Fontainebleau sandstone sample, a group of 30 microstructure samples are reconstructed by the proposed multi-level approach and Bostanabad's method individually, and representative reconstruction results are shown in Figure 29. The Fontainebleau sandstone sample with the porosity of 24.5% is taken as the example to illustrate the characterization and reconstruction procedure in Figure 28. The Gaussian pyramid with three levels is used as the training images, where the training image is half sized from the high level to lower levels with the main morphology maintained. At the low level, the long-distance morphology is captured compactly by using a data template of a relatively small size. Conditional to the reconstruction result at the low level, morphological details are added to the reconstruction at the middle level, which makes the reconstruction result much closer to the original microstructure, especially in terms of the local connectivity between isolated pores. After adding more high-frequency morphology features at the highest level, the final reconstruction result becomes visually indistinguishable compared to the original microstructure. By contrast, the reconstructed samples from Bostanabad's method have smaller pores and less pore connectivity, as shown in Figure 29.
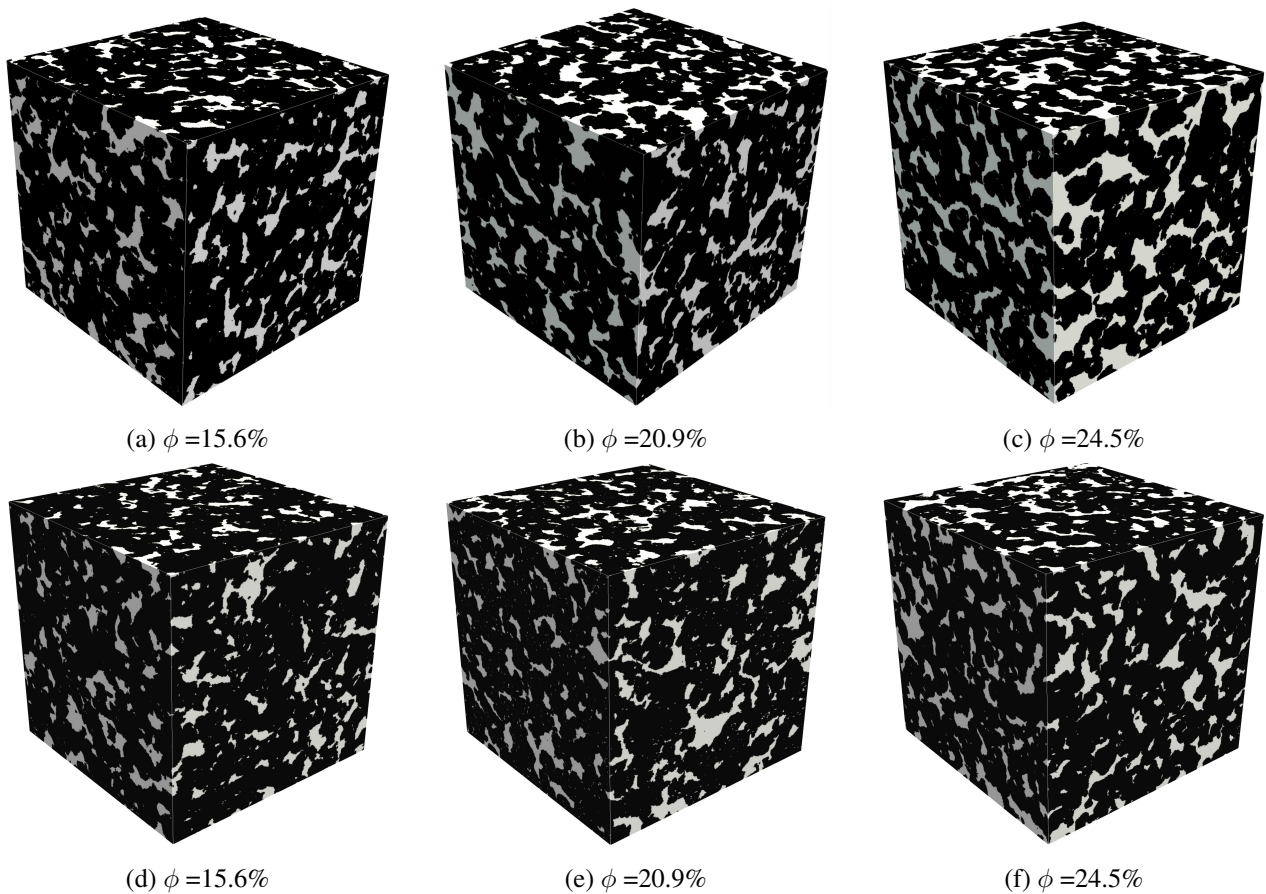

(a) ϕ =15.6%   (b) ϕ =20.9%   (c) ϕ =24.5%

(d) ϕ =15.6%   (e) ϕ =20.9%   (f) ϕ =24.5%

Figure 29: The representative reconstruction results ($320 \times 320 \times 320$) for Fontainebleau sandstone samples by using different methods: (a)∼(c) the proposed multi-level approach and (d)∼(f) Bostanabad's method.

Statistical equivalence between the original and reconstructed microstructures is accessed by TPCF, as shown in Figure 30. Both the proposed multi-level approach and Bostanabad's method show good performance to maintain the stochastic features of original microstructures, but the multi-level approach has higher accuracy in reconstructing samples in terms of TPCF. Pore size distribution (PSD) $P(D)$ (Lindquist et al., 2000) is used to evaluate the reconstruction results from the angle of pore geometry, and it measures the distribution of equivalent diameters D of pore bodies. In Figure 31, the black solid curve is the PSD extracted from the original microstructure, while the red and blue dashed curves are the average results of PSDs calculated from 30 reconstructed samples. Bostanabad's method tends to generate smaller pores inside the reconstructed samples compared to the pores inside the original microstructures, and pore geometry information at large scale are not

fully captured. The proposed multi-level approach is however able to accurately capture the geometry information of both small and large pores, which is confirmed by the good agreements of PSDs extracted from the reconstructed and original microstructures. Total fraction of percolating cells (TFPC) $T_3(L)$ (Øren and Bakke, 2002) measures topology information of pore network, which reflects the connectivity characteristics of porous media. As shown in Figure 32, the average TFPCs extracted from the 30 samples reconstructed by Bostanabad's method are just below the reference TFPCs, especially at long distance, which means the reconstructed samples have worse pore connectivity compared to the original microstructure. However, the long-distance connectivity of pore network is well preserved in the samples reconstructed by the proposed multi-level approach, which is confirmed by the good matches of TFPCs between the original and reconstructed microstructures. In general, the proposed multi-level approach is very powerful to capture the morphological features of porous microstructures with long-distance connectivity.
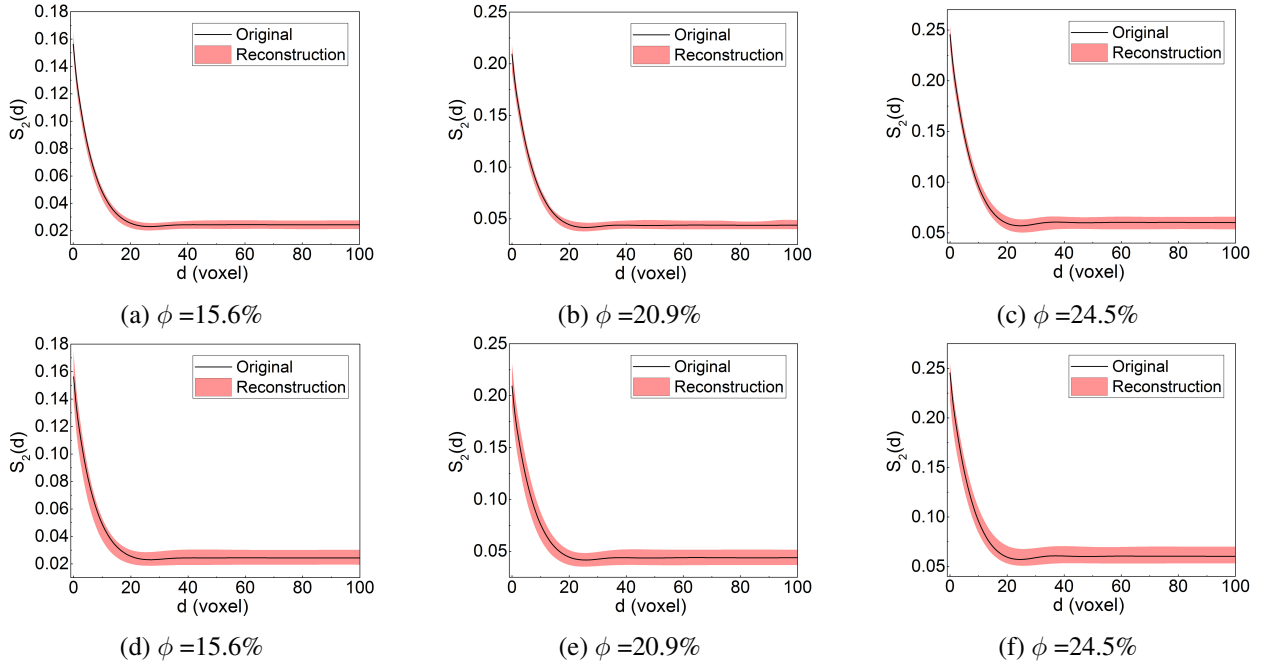

Figure 30: Comparison of TPCF $S_2(d)$ between the original and reconstructed microstructures from different methods: (a)∼(c) the proposed multi-level approach and (d)∼(f) Bostanabad's method.
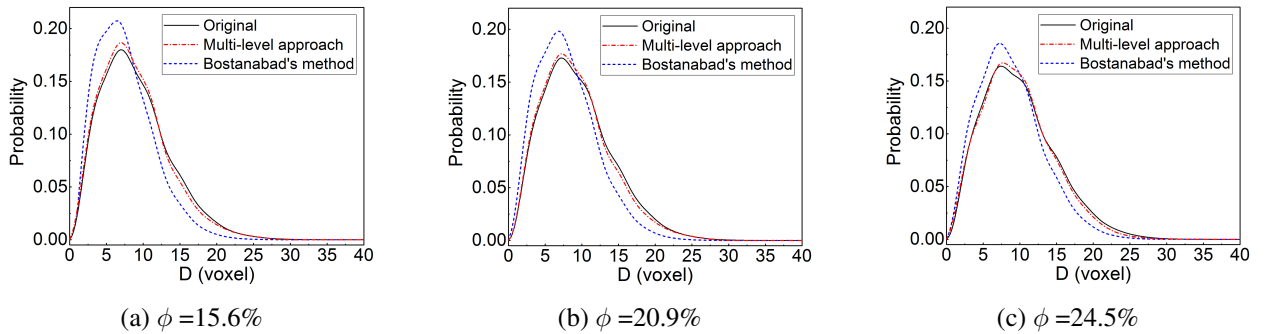

Figure 31: Comparison of PSD $P(D)$ between the original and reconstructed microstructures.

5.2.3 Transport properties

In this section, transport properties, including intrinsic permeability and diffusional tortuosity factor, are compared between the original and reconstructed microstructures to assess the reconstruction quality. Both intrinsic

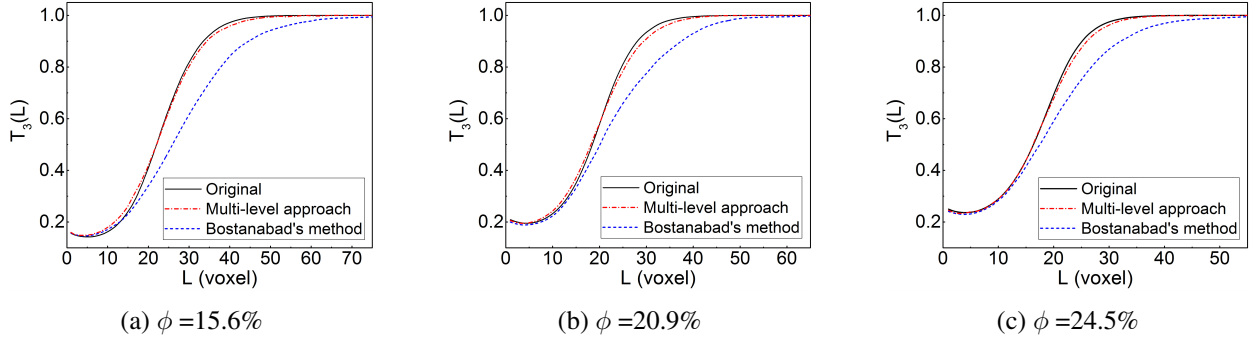|		
(a) $\phi = 15.6\%$	(b) $\phi = 20.9\%$	(c) $\phi = 24.5\%$

Figure 32: Comparison of TFPC $T_3(L)$ between the original and reconstructed microstructures.

permeability and diffusional tortuosity factor are directly related to geometrical attributes of the porous microstructure, such as porosity, specific surface area, and pore/throat size distribution (Berg, 2014; Cooper et al., 2016). Lattice Boltzmann Method (LBM) (Krüger et al., 2017; Fu et al., 2020) and Finite Difference Method (FDM) (LeVeque, 2007) are adopted to evaluate the permeability and tortuosity factor through open-source solvers: $Palabos$ (Degruyter et al., 2010) and $TauFactor$ (Cooper et al., 2016), respectively. These two open-source solvers are widely used by the research community of porous media, where pore-scale simulations of fluid flow and molecular diffusion can be directly performed on the digital microstructures by using the voxels as the mesh elements, and they have good flexibility to work with complex geometries.

The LBM is used to simulate a single-phase fluid flow with low Reynolds number ($Re \ll 1$) passing through a porous medium under a constant pressure gradient. When the fluid flow reaches a steady state, it can be described by Darcy's law, and the intrinsic permeability κ of the porous medium is quantified by the following equation:

$$\kappa = -\frac{\mu}{\nabla p}\langle \mathbf{u} \rangle \tag{21}$$

where ∇p is the pressure gradient along the direction of fluid flow, μ is the dynamic viscosity of fluid, and $\langle \mathbf{u} \rangle$ denotes the average fluid velocity across the simulation domain.

The FDM is adopted to simulate the steady diffusive flow through a porous medium under a constant concentration gradient. Compared to the free diffusion without geometry constraint, the diffusive transport through the porous medium is remarkably decreased because of the convolutions of the flow paths inside the porous medium. The steady-state diffusive flow is described by the Fick's first law, and the decrease of diffusive transport in a porous medium is quantified by tortuosity factor τ (Grathwohl, 2012):

$$\mathbf{J} = -D_{\mathrm{e}}\nabla C \tag{22}$$

$$D_{\mathrm{e}} = D\frac{\phi}{\tau} \tag{23}$$

where $\mathbf{J}$ denotes the diffusion flux vector, D_{e} is the effective diffusivity, ∇C is the concentration gradient, D is the intrinsic diffusivity of the conductive phase (gas or liquid) filling the pores, and ϕ is porosity.

Performing LBM simulations on both the original and reconstructed microstructures of Fontainebleau sandstones, we obtain the velocity fields of fluid flow at steady states, as shown in Figure 33. Driven by the same pressure gradient, the flow velocity magnitude range inside the microstructure reconstructed from the multi-level approach is close to that inside the original microstructure. As to the microstructure reconstructed from Bostanabad's method, it has a smaller velocity magnitude and fewer flow streamlines. The permeability results evaluated from LBM simulations are recorded in Figure 35a. It can be observed that the permeabilities of the samples reconstructed from the multi-level approach fluctuate around the reference values (red stars). The average permeability value of each group of 30 reconstructed samples is close to the corresponding reference permeability, with the average error less than 5.00%. The presence of permeability variation in reconstructed samples is expected, because these reconstructed samples are generated from probability sampling and they statistically represent the heterogeneous nature of the concerned microstructures. However, Bostanabad's method
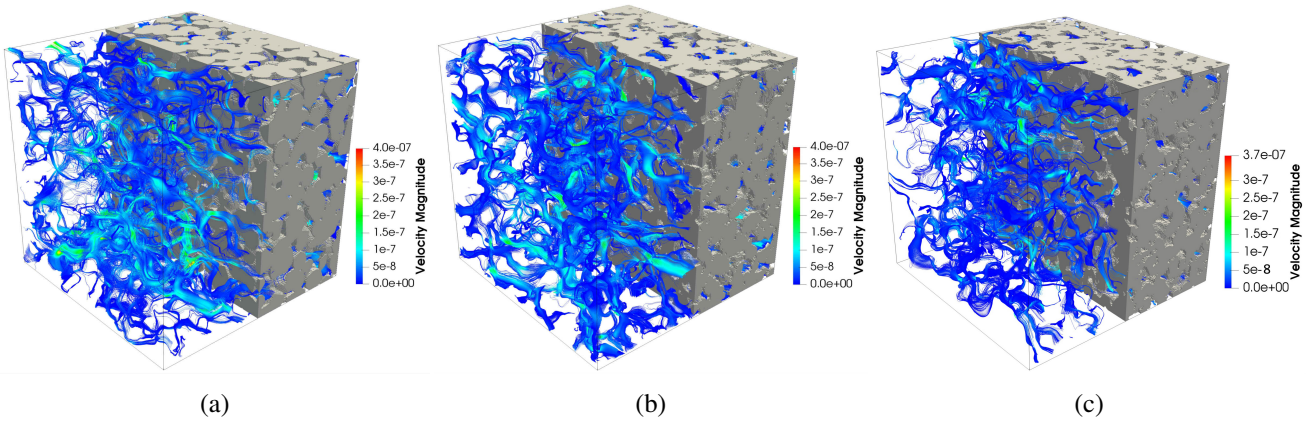
Figure 33: The fluid velocity fields at steady state obtained from LBM simulations: (a) Original microstructure ($\phi = 24.5\%$); (b) The representative microstructure ($\phi = 24.5\%$) reconstructed from the multi-level approach; (c) The representative microstructure ($\phi = 24.5\%$) reconstructed from Bostanabad's method.

fails to preserve the intrinsic permeabilities of original microstructures, and the reconstructed samples generally have significantly smaller permeabilities compared to the reference values, with errors ranging from -79.70% to -36.10%.
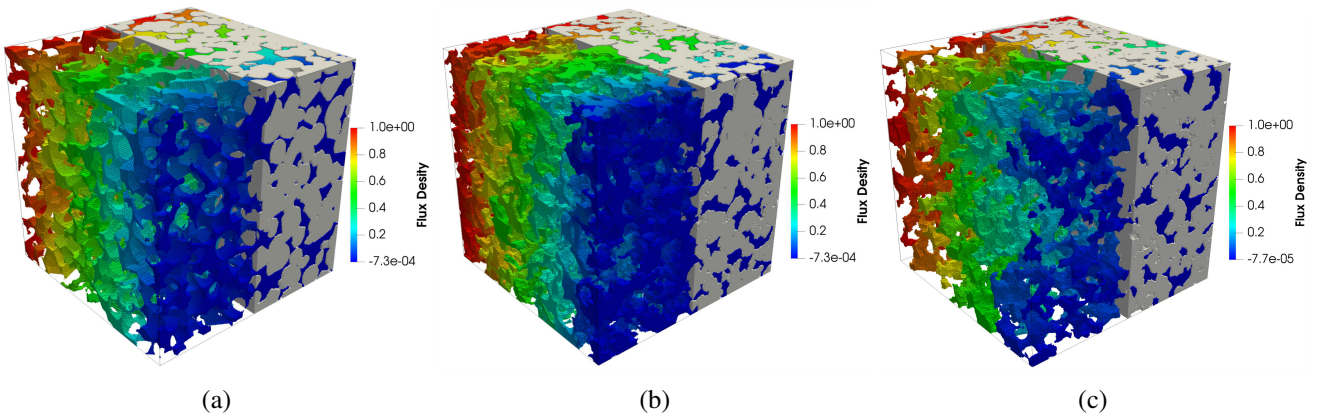

Figure 34: The flux density fields at steady state obtained from FDM simulations: (a) Original microstructure ($\phi = 24.5\%$); (b) The representative microstructure ($\phi = 24.5\%$) reconstructed from the multi-level approach; (c) The representative microstructure ($\phi = 24.5\%$) reconstructed from Bostanabad's method.

(a) Intrisinc permeability
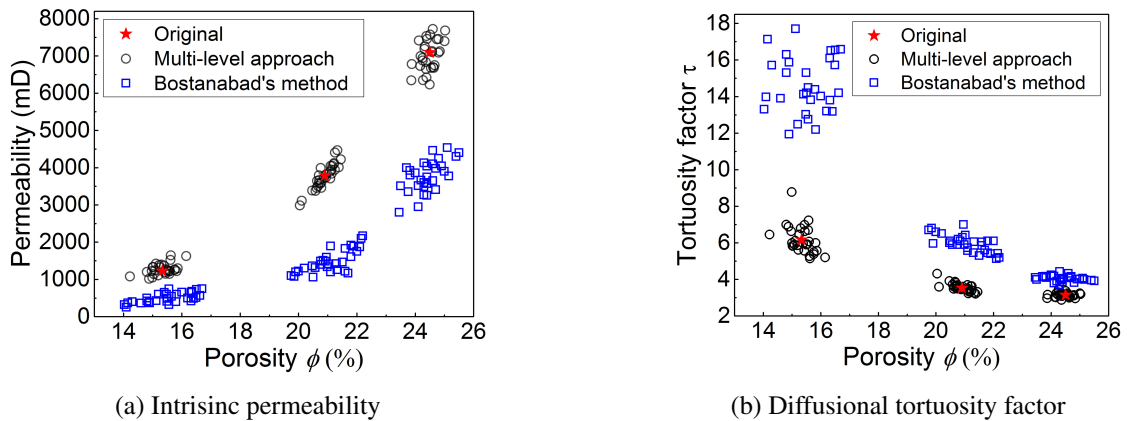
(b) Diffusional tortuosity factor

Figure 35: The comparison of transport properties between the original and reconstructed microstructures.

FDM simulations are applied to both the original and reconstructed microstructures, from which the steady-state flux density fields of diffusion are acquired. As shown in Figure 34, the output flux from the sample reconstructed by the multi-level method is close to that from the original microstructure, under the same concentrate gradient. But the reconstructed microstructure from Bostanabad's method outputs much less diffusional flux, compared to the original one. Tortuosity factors are evaluated from the steady-state flux density fields (as plotted in Figure 35b), which quantify the retarding effects of microstructures on diffusion. Again, the tortuosity factors of reconstructed samples from the multi-level method are located around the reference values, with the average errors smaller than 5.00%. However, Bostanabad's method is not able to maintain the tortuosity factors of original microstructures, and the tortuosity factors of its reconstructed sample are 25.80% to 186.70% larger than the reference values. In summary, the proposed multi-level approach shows excellent performance to statistically preserve the transport properties of porous media, and the reconstructed samples can be used for relevant researches.

6 Conclusions

The main contribution of this study is to present a versatile method that can accurately characterize and efficiently reconstruct heterogeneous microstructures using deep neural network (DNN), and it only requires one (or a few) real microstructure(s) to serve as the training image(s). The most prominent advantage of this method is its powerful capacity to capture long-distance morphological features, which overcomes the short-distance limitation of many MRF-based methods.

Specifically, the digital microstructure image is treated as an extended Markov random field (MRF), in order to improve microstructure characterization quality and reconstruction speed. Different data templates are designed to collect data events (local morphology patterns) from the target image, and these data events are used as training data to fit an 'SSAE+Softmax' model for the microstructure characterization purpose. In essence, the fitted 'SSAE+Softmax' model is an implicit representation of the conditional probability distribution function of pixels/voxels in the target image, from which statistically equivalent microstructure samples can be generated through a specific reconstruction procedure. Moreover, a multi-level approach is developed for microstructures with long-distance correlations, where microstructure characterization and reconstruction are carried out in a hierarchical manner, to avoid the short-distance limitation inherent in the MRF assumption.

The new method is tested on both 2D and 3D microstructures with distinct morphologies, and reconstruction qualities are evaluated through comparisons between target and reconstructed microstructures, in terms of statistical information, geometrical attribute, topological feature and transport properties. The comparisons show excellent consistency, and the statistical equivalence, morphological similarity and transport properties are all well preserved by the reconstructed microstructure samples. In general, the proposed method is applicable to a variety of microstructures, especially suitable for microstructures with long-distance morphological features. It is also flexible to balance the computation memory requirement and reconstruction speed by choosing different data templates or switching between single-level and multi-level approach. It is noted that the proposed method can be readily extended for heterogeneous media with multiple phases, although only two-phase microstructures are involved in this work.

Acknowledgments

References

Agrawal, A. and Choudhary, A. (2019). Deep materials informatics: Applications of deep learning in materials science. *MRS Communications*, pages 1–14.

Berg, C. F. (2014). Permeability description by characteristic length, tortuosity, constriction and porosity. *Transport in Porous Media*, 103(3):381–400.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Bostanabad, R., Bui, A. T., Xie, W., Apley, D. W., and Chen, W. (2016a). Stochastic microstructure characterization and reconstruction via supervised learning. *Acta Materialia*, 103:89–102.

Bostanabad, R., Chen, W., and Apley, D. (2016b). Characterization and reconstruction of 3d stochastic microstructures via supervised learning. *Journal of Microscopy*, 264(3):282–297.

Brandon, D. and Kaplan, W. D. (2013). *Microstructural characterization of materials*. John Wiley & Sons.

Burt, P. and Adelson, E. (1983). The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540.

Caers, J. (2001). Geostatistical reservoir modelling using statistical pattern recognition. *Journal of Petroleum Science and Engineering*, 29(3-4):177–188.

Cang, R., Xu, Y., Chen, S., Liu, Y., Jiao, Y., and Yi Ren, M. (2017). Microstructure representation and reconstruction of heterogeneous materials via deep belief network for computational material design. *Journal of Mechanical Design*, 139(7).

Cooper, S., Bertei, A., Shearing, P., Kilner, J., and Brandon, N. (2016). Taufactor: An open-source application for calculating tortuosity factors from tomographic data. *SoftwareX*, 5:203–210.

Corson, P. B. (1974). Correlation functions for predicting properties of heterogeneous materials. ii. empirical construction of spatial correlation functions for two-phase solids. *Journal of Applied Physics*, 45(7):3165–3170.

Degruyter, W., Burgisser, A., Bachmann, O., and Malaspinas, O. (2010). Synchrotron x-ray microtomography and lattice boltzmann simulations of gas flow through volcanic pumices. *Geosphere*, 6(5):470–481.

Feng, J., Cen, S., Li, C., and Owen, D. (2016). Statistical reconstruction and karhunen–loève expansion for multiphase random media. *International Journal for Numerical Methods in Engineering*, 105(1):3–32.

Feng, J., Li, C., Cen, S., and Owen, D. (2014). Statistical reconstruction of two-phase random media. *Computers & Structures*, 137:78–92.

Fu, J., Dong, J., Wang, Y., Ju, Y., Owen, D., and Li, C. (2020). Resolution effect: An error correction model for intrinsic permeability of porous media estimated from lattice boltzmann method. *Transport in Porous Media*, 132(3):627–656.

Grathwohl, P. (2012). *Diffusion in natural porous media: contaminant transport, sorption/desorption and dissolution kinetics*, volume 1. Springer Science & Business Media.

Gupta, A., Cecen, A., Goyal, S., Singh, A. K., and Kalidindi, S. R. (2015). Structure–property linkages using a data science approach: application to a non-metallic inclusion/steel composite system. *Acta Materialia*, 91:239–254.

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

Ju, Y., Zheng, J., Epstein, M., Sudak, L., Wang, J., and Zhao, X. (2014). 3d numerical reconstruction of well-connected porous structure of rock using fractal algorithms. *Computer Methods in Applied Mechanics and Engineering*, 279:212–226.

Krüger, T., Kusumaatmaja, H., Kuzmin, A., Shardt, O., Silva, G., and Viggen, E. M. (2017). The lattice boltzmann method. *Springer International Publishing*, 10:978–3.

LeVeque, R. J. (2007). *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*, volume 98. Siam.

Li, S. Z. (2009). *Markov random field modeling in image analysis*. Springer Science & Business Media.

Li, X., Zhang, Y., Zhao, H., Burkhart, C., Brinson, L. C., and Chen, W. (2018). A transfer learning approach for microstructure reconstruction and structure-property predictions. *Scientific Reports*, 8.

Lindquist, W. B., Venkatarangan, A., Dunsmuir, J., and Wong, T.-f. (2000). Pore and throat size distributions measured from synchrotron x-ray tomographic images of fontainebleau sandstones. *Journal of Geophysical Research: Solid Earth*, 105(B9):21509–21527.

Liu, X. and Shapiro, V. (2015). Random heterogeneous materials via texture synthesis. *Computational Materials Science*, 99:177–189.

Mosser, L., Dubrule, O., and Blunt, M. J. (2017). Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Physical Review E*, 96(4):043309.

Okabe, H. and Blunt, M. J. (2005). Pore space reconstruction using multiple-point statistics. *Journal of Petroleum Science and Engineering*, 46(1-2):121–137.

Øren, P.-E. and Bakke, S. (2002). Process based reconstruction of sandstones and prediction of transport properties. *Transport in Porous Media*, 46(2-3):311–343.

Quiblier, J. A. (1984). A new three-dimensional modeling technique for studying porous media. *Journal of Colloid and Interface Science*, 98(1):84–102.

Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., and Kim, C. (2017). Machine learning in materials informatics: recent applications and prospects. *NPJ Computational Materials*, 3(1):54.

Schlüter, S., Sheppard, A., Brown, K., and Wildenschild, D. (2014). Image processing of multiphase images obtained via x-ray microtomography: a review. *Water Resources Research*, 50(4):3615–3639.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117.

Stutzman, P. (2004). Scanning electron microscopy imaging of hydraulic cement microstructure. *Cement and Concrete Composites*, 26(8):957–966.

Sundararaghavan, V. and Zabaras, N. (2005). Classification and reconstruction of three-dimensional microstructures using support vector machines. *Computational Materials Science*, 32(2):223–239.

Tahmasebi, P. and Sahimi, M. (2012). Reconstruction of three-dimensional porous media using a single thin section. *Physical Review E*, 85(6):066709.

Torquato, S. (2013). *Random heterogeneous materials: microstructure and macroscopic properties*, volume 16. Springer Science & Business Media.

Vogel, H.-J., Weller, U., and Schlüter, S. (2010). Quantification of soil structure based on minkowski functions. *Computers & Geosciences*, 36(10):1236–1245.

Wang, Y., Arns, C. H., Rahman, S. S., and Arns, J.-Y. (2018). Porous structure reconstruction using convolutional neural networks. *Mathematical Geosciences*, 50(7):781–799.

Wang, Y., Yao, H., and Zhao, S. (2016). Auto-encoder based dimensionality reduction. *Neurocomputing*, 184:232–242.

Wei, L.-Y. and Levoy, M. (2000). Fast texture synthesis using tree-structured vector quantization. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 479–488. ACM Press/Addison-Wesley Publishing Co.

Wu, K., Van Dijke, M. I., Couples, G. D., Jiang, Z., Ma, J., Sorbie, K. S., Crawford, J., Young, I., and Zhang, X. (2006). 3d stochastic modelling of heterogeneous porous media–applications to reservoir rocks. *Transport in Porous Media*, 65(3):443–467.

Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., and Madabhushi, A. (2015). Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE Transactions on Medical Imaging*, 35(1):119–130.

Yang, M., Nagarajan, A., Liang, B., and Soghrati, S. (2018). New algorithms for virtual reconstruction of heterogeneous microstructures. *Computer Methods in Applied Mechanics and Engineering*, 338:275–298.

Yang, X.-Y., Chen, L.-H., Li, Y., Rooke, J. C., Sanchez, C., and Su, B.-L. (2017). Hierarchically porous materials: synthesis strategies and structure design. *Chemical Society Reviews*, 46(2):481–558.

Yeong, C. and Torquato, S. (1998). Reconstructing random media. *Physical Review E*, 57(1):495.