Accepted Manuscript

Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning

Xiang Li, Zhanli Liu, Shaoqing Cui, Chengcheng Luo, Chenfeng Li, Zhuo Zhuang

PII: S0045-7825(19)30009-X
DOI: https://doi.org/10.1016/j.cma.2019.01.005
Reference: CMA 12242

To appear in: *Comput. Methods Appl. Mech. Engrg.*

Received date : 27 September 2018
Revised date : 22 December 2018
Accepted date : 2 January 2019

Please cite this article as: X. Li, Z.L. Liu, S. Cui et al., Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning, *Computer Methods in Applied Mechanics and Engineering* (2019), https://doi.org/10.1016/j.cma.2019.01.005

- Predict the mechanical properties of heterogeneous materials by deep learning
- Generate numerous training samples based on stochastic reconstruction
- Transform sample images to finite element models by image processing
- An artificial neural network to predict mechanical properties by material structure
- The novel method is accurate and efficient in predicting mechanical properties

Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning

Xiang Li[a], Zhanli Liu[a,*], Shaoqing Cui[b], Chengcheng Lu[a]
Chenfeng Li[b], Zhuo Zhuang[a]

[a]Applied Mechanics Lab., Dept. of Engineering Mechanics, School of Aerospace, Tsinghua University, Beijing 100084, China

[b] Zienkiewicz Centre for Computational Engineering, Swansea University, Swansea, United Kingdom

Abstract

In contrast to the composition uniformity of homogeneous materials, heterogeneous materials are normally composed of two or more distinctive constituents. It is usually recognized that the effective material property of a heterogeneous material is related to the mechanical property and the distribution pattern of each forming constituent. However, to establish an explicit relationship between the macroscale mechanical property and the microstructure appears to be complicated. On the other hand, machine learning methods are broadly employed to excavate inherent rules and correlations based on a significant amount of data samples. Specifically, deep neural networks are established to deal with situations where input-output mappings are extensively complex. In this paper, a method is proposed to establish the implicit mapping between the effective mechanical property and the mesoscale structure of heterogeneous materials. Shale is employed in this paper as an example to illustrate the method. At the mesoscale, a shale sample is a complex heterogeneous composite that consists of multiple mineral constituents. The mechanical properties of each mineral constituent vary significantly, and mineral constituents are distributed in an utterly random manner within shale samples. Large quantities of shale samples are generated based on mesoscale scanning electron microscopy images using a stochastic reconstruction algorithm. Image processing techniques are employed to transform the shale sample images to finite element models. Finite element analysis is utilized to evaluate the effective mechanical properties of the shale samples. A convolutional neural network is trained based on the images of stochastic shale samples and their effective moduli. The trained network

* Corresponding author. Tel.: +86- 10-62771114; Fax: +86- 10-62771114.
Email: liuzhanli@tsinghua.edu.cn (ZL Liu)

is validated to be able to predict the effective moduli of real shale samples accurately and efficiently. Not limited to shale, the proposed method can be further extended to predict effective mechanical properties of other heterogeneous materials.

Keywords: Heterogeneous materials; Shale; deep learning; stochastic reconstruction;

1. Introduction

The complex microstructure and various forming constituents of heterogeneous materials have long posed difficulties to the study of their effective mechanical properties. Researchers have studied the effective mechanical properties, such as elastic modulus and thermal conductivity, based on analytical approaches [1-5]. These approaches are generally developed based on simplified models and statistical data. In this paper, a new method is proposed to predict the effective mechanical properties of heterogeneous materials. Unlike the studies mentioned above, the proposed method takes advantage of computational mechanics and deep learning methods. A representative heterogeneous material, shale, is employed in this paper for illustration.

Shale is a multi-phase, multi-scale fine grained sedimentary rock. Shale makes up around 75% of sedimentary basins of the earth and is critical to petrol and natural gas exploitation [6]. Shale gas is normally absorbed onto organic kerogen of shale rock. The recent development of horizontal drilling and hydraulic fracturing techniques has made the large production of shale gas possible. These techniques are closely related to the shale's macroscopic mechanical properties such as effective modulus, hardness, and strength. However, experimental researches reveal that the macroscopic mechanical properties of different shale samples vary drastically [7]. For this reason, the research on mechanical properties of shales is sometimes conducted from a mesoscale point of view [8].

At the mesoscale, shale is considered as a type of complex heterogeneous composites that consists of multiple mineral constituents [9]. These mineral constituents include quartz, illite, feldspar, calcite, kaolinite, pyrite, dolomite, kerogen, etc [10, 11]. The modulus of each mineral constituent is proved to be highly deviated based on nanoindentation experiments [8]. As an example, the modulus of quartz and feldspar are relatively high, and the modulus of kaolinite and illite are considerably lower in comparison. To understand the correlation between shale mesoscale structure and the macroscopic mechanical is vital to provide insights into the engineering design of shale gas extraction. Because of the complex mesoscale structure of shale, to reveal this implicit correlation by means of analytical analysis seems impractical and arduous. However, with the advancement of computer technology, this problem becomes much more accessible with the help of computational mechanics approaches.

The development of machine learning has been highly motivated with the

advancement of computer science. Some of the early machine learning methods include the perceptron [12], genetic programming [13], and the Monte Carlo method [14]. After the 1990s, the proper orthogonal decomposition [15], adaptive boost [16], support vector machine [17, 18], particle swarm optimization method [19] have been established and applied to various engineering fields.

Besides the previously mentioned machine learning methods, the artificial neural network (ANN) has become a significant branch of machine learning. In 1943, McCulloch and Pitts [20] proposed the mathematical framework of the artificial neuron inspired by the characteristics of nervous activity. Rosenblatt [12] established the perceptron in 1958, which is generally recognized as the predecessor of the modern artificial neural network. The prototypes of the artificial neural network were the models named "ADALINE" and "MADALINE" created by Stanford University. The limitation of the early network models is the difficulty to solve nonlinear problems due to their simple linear architectures. Ivakhnenko [21] made the earliest efforts to develop deep learning models. Rumelhart et al. [22] and other researchers came up with the back propagation (BP) algorithm, which later becomes the backbone of deep neural networks. In 2006, Hinton and Salakhutdinov [23] proposed the framework for deep learning based on the concept of Deep Belief Networks. Due to the nonlinear activation function and hidden neurons, deep artificial neural networks are able to extract implicit and complex data mappings based on numerous training data. For this reason, deep artificial neural networks have been adopted in various applications of mechanics and engineering.

Ghaboussi et al. [24], Jung and Ghaboussi [25], Ji et al. [26], Furukawa and Yagawa [27], Hashash et al. [28], Sun et al. [29] implemented artificial neural networks to study the constitutive models of solid materials. Faller and Schreck [30], Wang and Liao [31], Xuhong and Wenxin [32], Butz and Von Stryk [33], Beigzadeh and Rahimi [34], Mi et al. [35] study the fluid characterizations based on numerical approaches assisted by artificial neural networks.

In recent years, researchers have employed machine learning models in the study of heterogeneous materials. Sundararaghavan and Zabaras [36] develops a framework to classify and reconstruct 3-D heterogeneous material based on support vector machines. Liu et al. [37] proposed approaches based on machine learning to predict elastic strain fields in a 3-D microstructure volume element of heterogeneous composite materials. Kondo et al. [38] employ convolutional neural networks to establish the mapping between the microstructure and the ionic conductivity of ceramics. The networks are trained by supervised learning based on cropped microscope scanning images. The data labels are the macroscopic ionic conductivities measured by impedance spectroscopy. Cang et al. [39] propose a method mainly to

generate stochastic microstructure samples based on variational auto-encoder. A predictive model based on convolutional neural networks is also proposed to reveal the data mapping between microstructure and effective properties. The network is trained by microscale two-phase sandstone samples. The labels of the samples are effective material properties calculated based on analytical approximations. Bessa et al. [40] propose a data-driven computational framework to design structures and materials. Sample data that represent microstructures, material properties, and boundary conditions are extracted; a database of material responses is established based on computational analyses; augmented by machine learning algorithms, the mapping between descriptors of sample data and the concerned material properties is constructed, and new designs or response models can be further obtained.

In this paper, a framework for predicting the effective material properties of multi-phase heterogeneous materials is proposed. For a demonstration of this framework, a convolutional neural network is established to exploit the implicit mappings between the mesoscale structures and the effective moduli of shale samples. Scanning electron microscope (SEM) is employed to obtain mesoscale structure images of shale. A simplified model is introduced to transform SEM images to 5-phase heterogeneous shale samples. Large quantities of shale samples are generated based on the 5-phase samples using a stochastic reconstruction algorithm. Finite element method is utilized to calculate the stochastic shale samples' effective moduli, which are further used as labels of training samples. A deep convolutional neural network is trained based on the images of stochastic shale samples and their effective moduli. The trained network is further employed to predict the effective moduli of real shale samples. The workflow of this process is depicted in Fig. 1. Each portion presented in Fig. 1 will be illustrated in details in the following sections.

Fig. 1 The workflow of establishing a deep neural network for calculating the shale modulus.

This paper is organized in the following scheme. The mesoscale structure of shale and the mechanical properties of its forming constituents are illustrated in Section 2. The stochastic reconstruction method and the finite element analysis to calculate the effective moduli of shale samples are discussed in Section 3. In Section 4, the principal theory of deep neural networks is demonstrated. The artificial neural network architecture used for modulus prediction is elaborated. The prediction accuracy of the deep neural network is discussed, and several conclusions are drawn based on the prediction result.

2. Shale mesoscale structure and mechanical properties of constituents

As mentioned in the previous section, the objective of this research is to develop a new approach to predict the moduli of mesoscale shale samples. In this section, we first discuss the mineral constituents and the structures of mesoscale shale. Then, a simplified mesoscale shale model that includes five main constituents is explained. The nanoindentation test, which is utilized to measure the moduli of each main constituent, is introduced in the final section.

2.1 Shale mesoscale structure and a simplified mesoscale model

Shale is a most commonly found sedimentary rock that accounts for about 50% of all sedimentary rocks on earth [41, 42]. From the macroscale point of view, a shale sample normally appears to be in dark gray color. Fig. 2 (a) shows a cylindrical shale sample for laboratory research. The length of the sample is about 5 cm, and the radius is about 1.2 cm. Fig. 2 (b) shows the typical laminated structure of shale formation. The laminated pattern is formed due to the sedimentation process.

Fig. 2 The macroscopic presentation of shale. (a) A laboratory shale sample. (b) The structure of a typical shale formation laminate.

From a mesoscale point of view, a shale sample is usually considered as a complex heterogeneous material that consists of multiple mineral constituents. It is normally comprised of quartz, calcite, smectite, pyrite, clay, organic matter and other minerals [7]. Energy dispersive X-ray spectroscopy (EDX), focused ion beam milling (FIB) and scanning electron microscope (SEM) are often utilized to identify the mineral constituents and characterize the heterogeneity of shale samples [43-47].

Studies reveal that the deformation and damage characteristics of shale are related to the mechanical properties and the distributions of its forming constituents. Therefore, it is important to study the distributions and small-scale mechanical properties of the forming constituents to understand the upscaling mechanical properties of macroscopic shale samples [48-50].

In this research, scanning electron microscope (SEM) is employed to investigate the distribution of each mineral constituent of a shale sample. The sample is scanned for 13.5 hours to generate an SEM image. The scanning voltage is 15 kV. The scanning resolution is 1 m.

The SEM image is shown in Fig. 3(a). It can be observed that the forming constituents of the sample include quartz, feldspar, pyrite, calcite, dolomite, kaolinite, illite, kerogen, etc. The distribution manner of the constituents appears to be random, which brings in strong heterogeneity and anisotropy to the shale samples.

In engineering and geomechanics applications, mineral constituents with similar

material properties are usually grouped for simplicity. A widely accepted shale model classifies various mineral constituents into four categories [51-53]. The first category is abbreviated as QFP. It contains quartz, feldspar, and pyrite, and they are the most commonly found silicate minerals in shale [54]. The modulus and hardness of QFP are generally highest among all constituents. The second category is clay, which contains kaolinite, illite, chlorite, and montmorillonite. The third category is the organic matter, which is also known as kerogen. The modulus and hardness of kerogen are the lowest. Shale gas is normally absorbed on kerogen, and the proportion of kerogen is relatively low. The final category contains all other mineral constituents that are not mentioned above, and it is sometimes referred to as the matrix phase.

In this research, we take the characteristics of the shale formation in southwest China into account. An additional phase, carbonate, is introduced to the aforementioned mesoscale shale model. The carbonate phase contains calcite and dolomite. Hence, a 5-phase mesoscale shale model is employed in this paper. The 5 phases are silicate, carbonate, clay, kerogen, and matrix, respectively. An SEM image of a shale sample and the corresponding simplified 5-phase model is shown in Fig. 3 (b). The forming constituents of each phase are listed in Tab. 1.

Fig. 3 (a) The SEM image of a mesoscale shale sample. (b) The corresponding simplified 5-phase model.

Tab. 1 The forming constituents of each phase of the mesoscale shale model.

2.2 Nanoindentation and modulus measurement

As previously mentioned, the deformation and damage characteristics of shale are related to the mechanical properties of its forming constituents. Therefore, to study the mechanical properties of each primary constituent of shale is important to understand the macroscopic mechanical properties of shale. Instrumented nanoindentations are usually employed to investigate shale's basic mechanical properties, such as modulus and hardness [55-57].

In this research, a series of nanoindentation tests are conducted on different locations of shale samples. The probe indents shale samples at locations where quartz, calcite, clay and organic matter are concentrated, respectively. The indentation depths range from 0.5 to 5 m.

Fig. 4 The moduli of shale constituents measured based on nanoindentation tests.

The measured modulus of each constituent is given in Fig. 4. It can be observed that the measured moduli of all these constituents converge towards an intermediate level as the indentation depth increases. It is mainly because the surrounding constituents and the supporting matrix tend to disturb the measurement as the

indentation depth increases [8]. For this reason, we select the measured modulus when the indentation is initially applied as the effective modulus of each constituent. For simplicity, we use the modulus of quartz and calcite to represent the modulus of silicate and carbonate phases, respectively. The moduli of silicate, carbonate, clay, organic matter and matrix are summarized in Tab. 2. It can be concluded that silicate is the stiffest constituent, and the organic matter is the softest. It is in agreement with the results from other researchers' work [8, 56].

Tab. 2 The measured moduli of the primary constituents in shale.

3. Stochastic reconstruction method and finite element Analysis

The objective of this research is to establish an artificial neural network to predict the effective moduli of shale samples. A large number of shale samples are required to train the network. A stochastic reconstruction algorithm is employed to generate these shale samples. The stochastic sample images are then transferred to finite element models. The effective moduli of shale samples are evaluated based on finite element analysis.

3.1 Stochastic reconstruction method

Due to time-consuming issues, it is not practical to obtain a large number of shale samples by SEM scanning. Hence, the stochastic reconstruction technique is employed to generate stochastic shale samples in this study. The basic idea of the stochastic reconstruction technique is to rapidly reproduce the reference sample based on the statistical information of morphology. In this way, the effective statistical characteristics of the stochastic samples and the reference sample are matched.

Various statistical reconstruction methods have been proposed with different performances and applicabilities. The stochastic reconstruction technique adopted in this study is the stochastic optimization reconstruction algorithm [52].

$$E = \sum_i w_i \left(\sum_x \left[D_0^i(x) - D_s^i(x) \right]^2 \right) \tag{1}$$

In the equation, $D_0^i(x)$ is the i_{th} statistical descriptor measured from the reference sample. $D_s^i(x)$ is the same descriptor measured from the stochastic sample. w_i is the weight parameter of the i_{th} descriptor. Various descriptors have been developed to capture different microscale morphological characteristics. Some of the descriptors are two-point correlation function [53], two-point cluster correlation function [58], lineal-path function [59, 60], etc. In this study, the two-point correlation function and lineal-path function are considered as the target descriptors to capture the basic statistical information from the reference samples.

By optimization, the stochastic sample evolves in a way that its statistical characteristics gradually approach those of the reference sample. The simulated annealing algorithm is adopted in this study to optimize the stochastic samples. The simulated annealing algorithm was introduced into the stochastic reconstruction technique by Yeong and Torquato [52]. Firstly, a random guess is conducted to generate an initial stochastic sample. The initial sample is then iteratively evolved based on a spin-exchange strategy. In other words, two image pixels that represent two different material phases are swapped in each interative step. After that, the new sample is accepted with a probability of $\min\{\exp(-\Delta E/T),1\}$. In the formulation, ΔE is the difference of the value of objective function E between the old and the new sample. T represents the temperature of the simulated annealing process. By gradually decreasing the temperature T in the prescribed annealing algorithm, the evolvement iterations are repeated until the termination criterion is met. Based on the proposed stochastic reconstruction method, a large number of stochastic shale samples are generated.

3.2 Creating finite element models based on mesoscale shale images

The method to generate stochastic shale samples is illustrated in the previous section. Numerous stochastic shale samples are generated based on the proposed method. These stochastic samples are further transformed to corresponding finite element models to calculate the effective moduli. In this section, we briefly discuss the procedure to generate a finite element model based on a stochastic shale sample.

In the research, an application is developed to generate finite element models based on stochastic shale samples. The workflow to generate the finite element model is shown in Fig. 5. As is known, a stochastic shale sample is composed of pixels. Different color on a sample image represents different mineral constituent. The application firstly generates a finite element model that has the same amount of elements as that of pixels on the shale image. Quadrilateral plane strain elements are employed in this study. Then, the application scans the shale sample image pixel by pixel, and it keeps track of the constituent type represented by each pixel. Each pixel is transformed into an element in the finite element model. The location of each pixel of the shale image is utilized to generate the coordinates of each node. Besides, five different node and element sets are generated and updated based on the color of the pixel. These sets are referred to as nodes and elements of silicate, carbonate, clay, kerogen and matrix phase, respectively. Each element set is assigned with a distinctive material property that corresponds to the constituent type that it represents. Finally, initial and boundary conditions are prescribed, and the finite element model of a shale sample is generated.

Fig. 5 The workflow to generate the finite element model based on a mesoscale shale sample image.

Finite element method is then employed to calculate the effective modulus of the sample. The scheme of the compression test is depicted in Fig. 6 (a). The left boundary is fixed in the horizontal direction. A compressional displacement loading is applied along the horizontal direction to the right boundary. As is known, constituents are distributed in a complex manner in a mesoscale shale model. During the loading process, the stress at each constituent is different because of the heterogeneous material property distribution. The contour of the stress component S_{11} is shown in Fig. 6 (b).

Fig. 6 (a) The computational model of the finite element method. (b) The stress component S_{11} contour on the model.

After finite element calculation, the total reaction force on the right boundary is summarized. The reaction force is then divided by the area of the right boundary to obtain the effective modulus. The moduli are further utilized as the labels of training samples of an artificial neural network. It is worth mentioning that it takes about 20 seconds to conduct the finite element analysis for each sample. As will be discussed in Section 4, 12500 shale samples are used in this study. The finite element analysis of all these samples is distributed to 3 desktop computers in parallel. It takes about 23 hours to calculate the modulus of all the training and testing samples.

4. Training and testing of an artificial neural network

In this section, we briefly talk about the basic theory of the artificial neural network. Then, the characteristics of a convolutional neural network are discussed. The training and prediction processes in this research are explained in the final section.

4.1 The basic theory of artificial neural network

The fundamental concept of the artificial neural network is inspired by neuroscience. The typical architecture of a multi-layer artificial neural network is shown in Fig. 7. The first layer of the network is the input layer, and the last layer is the output layer. The layers between the input and output layers are hidden layers [61, 62].

Each circle in the figure represents a neuron. The value of each neuron is known as the activation, which is normally represented by a real number σ that ranges between 0 and 1. The superscript l of neuron σ_j^l represents the layer number; the

subscript j represents the neuron number. In a standard fully-connected network, the lines that interconnect between neurons are weights w. w_{jk}^{l} represents the weight that connects the k^{th} neuron in the $(l-1)^{th}$ layer and the j^{th} neuron in the l^{th} layer.

Fig. 7 The architecture of an artificial neural network.

Fig. 8 shows the process to calculate the activation of neuron σ_j^l. To obtain the value of neuron σ_j^l, a function z is first defined as a linear combination of the weights and biases, as given in Eq. (2) [63].

$$z_j^l = \sum_{i=1}^{n_{l-1}} \left(w_{ji}^l \cdot \sigma_i^{l-1} + b_j^l \right) \tag{2}$$

Fig. 8 The scheme to calculate the activation value of a neuron.

As previously mentioned, a disadvantage of the early network model is the difficulty to extract complex mappings due to their simple linear architectures. In contrast, nonlinearity is introduced to overcome this disadvantage. A nonlinear function f is applied over function z to calculate the activation of that specific neuron. f is also known as the activation function. Some broadly chosen activation functions include the sigmoid function and the rectified linear unit (ReLU) function [64]. The equation of sigmoid function is given by

$$f(z) = \frac{1}{1+e^{-z}} \tag{3}$$

The equation of rectified linear unit function is given by the maximum of 0 and z as

$$f(z) = \max(0, z) \tag{4}$$

Hence, the activation of the j^{th} neuron in layer l is calculated in Eq. (5).

$$\sigma_j^l = f\left(z_j^l\right) = f\left\{ \sum_{i=1}^{n_{l-1}} \left(w_{ji}^l \cdot \sigma_i^{l-1} + b_j^l \right) \right\} \tag{5}$$

Based on the mapping scheme given in Eq. (5), a neural network is able to provide the estimated output from the input. Since the initial values of all the weights and biases are chosen randomly, the output value is initially different from the desired output. The difference is usually referred to as the error or cost of a neural network. The desired output is also known as the label of training data. As an example, a broadly adopted cost function is the mean squared error function [61, 63].

$$C = \frac{1}{2n} \sum_{i=1}^{n} \left\| o_i - l_i \right\|^2 \tag{6}$$

In the equation, n is the number of training samples. o_i is the output value of the i_{th} training sample, while l_i is the label of the i_{th} training sample. In this case, the notation $\|\ \|$ is the L2 norm that measures the distance between vector o_i and l_i. As the labels of samples are fixed, the cost C in Eq. (6) is a function of outputs o. The objective of training an artificial neural network is to determine all the network parameters, in other words, the values of all the weights w and biases b. The parameters are determined in a way so that the cost C should be minimized.

After the cost function is given, the gradient descent algorithm is implemented to approximate the minimum of the cost function C. For simplicity, C is assumed to be a function of tensor v, which represents all the parameters of the artificial neural network. The variation of C can be approximated by the slight variation of v as [65, 66]

$$\Delta C \approx \nabla C \cdot \Delta \mathrm{v} \tag{7}$$

In the equation above, ∇C is the gradient vector of the cost function C. $\Delta \mathrm{v}$ is the slight variation of v. It is chosen as the format in Eq. (8) so that the cost function C is descended and approximated to 0 step by step

$$\Delta \mathrm{v} = -\eta \nabla C \tag{8}$$

η in Eq. (8) is known as the learning rate. The value of η should be cautiously selected so that the approximation in Eq. (7) holds. The values of all the parameters of the network can be iteratively updated by

$$\mathrm{v}^{updated} = \mathrm{v} - \eta \nabla C \tag{9}$$

Therefore, each component of the weights and biases of the network is updated by Eq. (10).

$$\begin{aligned} w_i^{updated} &= w_i - \eta \frac{\partial C}{\partial w_i} \\ b_j^{updated} &= b_j - \eta \frac{\partial C}{\partial b_j} \end{aligned} \tag{10}$$

4.2 Convolutional neural network

The principal theory of artificial neural networks is discussed in the previous section. The specific artificial neural network used in this research is a convolutional neural network. The architecture of the convolutional neural network was first introduced by Fukushima and Miyake [67]. The convolutional neural network shares a lot of similarities with the artificial neural network mentioned in the previous section. However, convolutional neural networks employ some unique features to specialize in image classification applications [68-72]. In this section, we will talk about some of the features and further introduce the network architecture used in this

research.

One feature of the convolutional neural network is called the local receptive field. As is discussed in the previous sections, in an artificial neural network, the neurons from two adjacent layers are normally fully connected. It means that any neurons from two adjacent layers are connected to each other, as shown in Fig. 9 (a). In contrast, it will be helpful to visualize that neurons are placed in a square pattern in a convolutional neural network, as shown in Fig. 9 (b). In the figure, only small and localized regions of neurons are connected to a neuron in the next layer [65].

Fig. 9 (a) The fully-connected architecture of an artificial neural network. (b) Local receptive field of a convolutional neural network.

Another feature is known as the pooling layer. In convolutional neural networks, pooling layers are normally applied after convolutional layers. Maximum pooling and average pooling are widely used to simplify the information of the output neurons from a convolutional layer [65].

With the features mentioned above, the architecture of convolutional neural networks takes into account the spatial structure of images [65]. In recent years, large convolutional networks demonstrate outstanding performance in image classifications [71]. To understand the underlying mechanism, an approach named deconvolutional network is implemented [69, 73, 74]. Using this technique, image patterns that stimulate high activations in a given feature map are reconstructed. In this way, researchers are able to study what information is visualized and learned by various convolutional layers. A convolutional neural network for facial recognition is given here for illustration. The convolutional neural network in Fig. 10 contains several hidden layers. Researchers found that neurons in the lower-level convolutional layers are able to "witness" very detailed facial features, such as an edge or a dot on the face. The intermedium layers can "visualize" relatively larger features like an eye, a nose or a mouth. The high-level layers are able to "see" the overall facial features.

Fig. 10 A convolutional neural network for a facial recognition application.

Inspired by the mechanism mentioned above, we hope to establish a multilayer convolutional neural network to predict the modulus of a shale sample, as given in Fig. 11. The lower-level layers are supposed to extract the features of tiny pieces of the shale sample. The intermedium layers are able to extract the features of medium-size shale pieces. The high-level layers are able to grasp the features of the overall shale sample. Several fully-connected layers are added after the last convolutional layer to obtain the effective modulus. In this way, the network establishes the implicit mapping between the mesoscale structure of a shale sample and its effective modulus. In the following section, we will discuss the training process of this convolutional

neural network.

Fig. 11 A convolutional neural network to establish the implicit mapping between the mesoscale structure of a shale sample and its effective moduli.

4.3 Training the convolution neural network

In this research, 10,000 stochastic mesoscale shale samples are generated to train a convolutional neural network. The training process is conducted on a desktop computer with i7-8700 CPU, 32G RAM, and an Nvidia GTX1080Ti. The training is iterated for 100 cycles, and it takes about 43 minutes to finish the training process.

As previously mentioned, the training process is to determine all the weights and biases of the convolutional neural network. The cost function used for training is the mean squared error function.

$$C = \frac{1}{2n}\sum_{i=1}^{n}\left\| o_i - t_i \right\|^2 \tag{11}$$

Stochastic gradient descent algorithm is employed to update all the weights and biases iteratively. The relation between the cost and the training iteration is plotted in Fig. 12 (a). It can be observed from the figure that the cost at the first iteration is significant because the initial values of weights and biases are randomly assigned. The weights and biases are updated based on the algorithm illustrated in Section 4.1. After several iterations, the cost rapidly descends.

Fig. 12 (a) The training error descends as the training iteration increases. (b) The cross validation error descends as the training iteration increases.

After the training is finished, 2,000 stochastic shale samples are used for cross validation. The moduli distribution of these stochastic samples is shown in Fig. 13. The minimum and maximum modulus of these samples are 51.98 GPa and 69.26 GPa, respectively.

Fig. 13 The moduli distribution of the 2000 stochastic shale samples.

The 2,000 samples are used as input data of the trained convolutional neural network. The network processes the input data and outputs the predicted effective moduli of the 2,000 samples. The predicted effective moduli are then compared with the moduli calculated based on finite element method. The relation between the cross validation error and the training iteration is shown in Fig. 12 (b). From the figure, it can be observed that the cross validation error converges after about 50 iterations. Fig. 14 depicts the distribution of the cross validation errors after 100 training iterations. The majority of the cross validation errors for the 2,000 samples is under 2%. The

average cross validation accuracy is 0.55 %. Hence, no obvious over fitting is observed in the cross validation, and the network can be further employed to predict the effective moduli of real shale samples.

Fig. 14 The distribution of the cross validation errors.

4.4 Predicting the moduli of real shale samples

The objective of establishing this network is to predict the effective moduli of real shale samples. In the final step, we employ the trained convolutional neural network to predict the effective moduli of 500 real shale samples. Some of these real shale samples are shown in Fig. 15 . It can be observed that the percentages and distributions of the forming constituents appear to be extensively random on these images.

Fig. 15 Several real mesoscale shale samples employed for modulus prediction.

The moduli distribution of these real samples is shown in Fig. 16 . The moduli of the real samples vary from 56.26 GPa to 82.10 GPa.

Fig. 16 The moduli distribution of the 500 real mesoscale shale samples.

The 500 real shale samples are used as input data of the trained convolutional neural network. It should be mentioned that none of the 500 real samples has been used in the training process. The network outputs the effective moduli of these 500 samples. The predicted moduli are compared with the labels of these samples to evaluate the prediction errors. The prediction results are given in Fig. 17. From the figure, it can be observed that most of the prediction errors are below 3%. The average prediction error is 0.37 %. We reckon that, based on the limited 10,000 training samples, the trained convolutional neural network exhibits promising performance in predicting the effective moduli of real shale samples.

Fig. 17 The prediction errors distribution of the 500 real shale samples.

5. Conclusion

A new method to predict effective mechanical properties of heterogeneous materials is presented in this paper. In this method, numerous stochastic mesoscale material samples are generated based on scanning image of material samples. The stochastic samples are transformed into corresponding finite element models. The effective mechanical properties of the finite element models are calculated based on finite element analysis. The effective mechanical properties are regarded as labels of the samples. The mesoscale structures of the stochastic samples and their labels are combined as training data to train a convolution neural work. The proposed method

takes advantages of the advanced fitting capability of the deep learning algorithm, and a multiple-layer convolution neural network is trained to excavate the implicit mapping between the mesoscale structure of material samples and their effective mechanical properties. The network is validated by cross validation and then employed to predict the effective mechanical properties of real samples. The prediction accuracy and efficiency of the method are promising.

In this paper, the proposed method combines image processing techniques, stochastic reconstruction approaches, finite element analysis, and deep learning method to predict the effective moduli of mesoscale shale samples. It should be noted that the prediction of the effective moduli of shale samples is used as an example to illustrate the method. The method can be further applied to predict the effective mechanical properties of other heterogeneous materials and even be integrated into the design of new composites with anticipated effective properties.

Acknowledgement

This work is supported by the Science Challenge Project, No. TZ2018001, National Natural Science Foundation of China, under Grant No. 11722218, 11302115 and 11532008, Tsinghua University Initiative Scientific Research Program, Chinese 1000-talents Plan for Young Researchers.

References

[1] Z. Hashin, The elastic moduli of heterogeneous materials, Journal of Applied Mechanics 29(1) (1962) 143-150.

[2] B. Budiansky, On the elastic moduli of some heterogeneous materials, Journal of the Mechanics and Physics of Solids 13(4) (1965) 223-227.

[3] M. Hori, T. Yonezawa, Statistical theory of effective electrical, thermal, and magnetic properties of random heterogeneous materials. IV. Effective− medium theory and cumulant expansion method, Journal of Mathematical Physics 16(2) (1975) 352-364.

[4] E.J. Garboczi, A. Day, An algorithm for computing the effective linear elastic properties of heterogeneous materials: three-dimensional results for composites with equal phase Poisson ratios, Journal of the Mechanics and Physics of Solids 43(9) (1995) 1349-1362.

[5] J. Wang, J.K. Carson, M.F. North, D.J. Cleland, A new approach to modelling the effective thermal conductivity of heterogeneous materials, International journal of heat and mass transfer 49(17-18) (2006) 3075-3083.

[6] C.P. Bobko, Assessing the mechanical microstructure of shale by nanoindentation: The link between mineral composition and mechanical properties, Massachusetts Institute of Technology, 2008.

[7] S. Lee, L. Hyder, P. Alley, Microstructural and mineralogical characterization of selected shales in support of nuclear waste respository studies, Microstructure of Fine-grained Sediments, Springer1991, pp. 545-560.

[8] K.C. Bennett, L.A. Berla, W.D. Nix, R.I. Borja, Instrumented nanoindentation and 3D mechanistic modeling of a shale at multiple scales, Acta Geotechnica 10(1) (2015) 1-14.

[9] V. Kumar, C.H. Sondergeld, C.S. Rai, Nano to macro mechanical characterization of shale, SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, 2012.

[10] C.H. Sondergeld, R.J. Ambrose, C.S. Rai, J. Moncrieff, Micro-structural studies of gas shales, SPE Unconventional Gas Conference, Society of Petroleum Engineers, 2010.

[11] M.E. Curtis, R.J. Ambrose, C.H. Sondergeld, Structural characterization of gas shales on the micro-and nano-scales, Canadian unconventional resources and international petroleum conference, Society of Petroleum Engineers, 2010.

[12] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, Psychological review 65(6) (1958) 386.

[13] W. Banzhaf, P. Nordin, R.E. Keller, F.D. Francone, Genetic programming: an introduction, Morgan Kaufmann San Francisco1998.

[14] N.M. Nasrabadi, Pattern recognition and machine learning, Journal of electronic imaging 16(4) (2007) 049901.

[15] Y. Liang, H. Lee, S. Lim, W. Lin, K. Lee, C. Wu, Proper orthogonal decomposition and its applications—Part I: Theory, Journal of Sound and vibration 252(3) (2002) 527-544.

[16] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, Journal of computer and system sciences 55(1) (1997) 119-139.

[17] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers, Proceedings of the fifth annual workshop on Computational learning theory, ACM, 1992, pp. 144-152.

[18] C. Cortes, V. Vapnik, Support-vector networks, Machine learning 20(3) (1995) 273-297.

[19] J. Kennedy, Particle swarm optimization, Encyclopedia of machine learning, Springer2011, pp. 760-766.

[20] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, The bulletin of mathematical biophysics 5(4) (1943) 115-133.

[21] A.G. Ivakhnenko, Polynomial theory of complex systems, IEEE transactions on Systems, Man, and Cybernetics (4) (1971) 364-378.

[22] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, nature 323(6088) (1986) 533.

[23] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, science 313(5786) (2006) 504-507.

[24] J. Ghaboussi, D.A. Pecknold, M. Zhang, R.M. Haj-Ali, Autoprogressive training of neural network constitutive models, International Journal for Numerical Methods in Engineering 42(1) (1998) 105-126.

[25] S. Jung, J. Ghaboussi, Neural network constitutive model for rate-dependent materials, Computers & Structures 84(15-16) (2006) 955-963.

[26] G. Ji, F. Li, Q. Li, H. Li, Z. Li, A comparative study on Arrhenius-type constitutive model and artificial neural network model to predict high-temperature deformation behaviour in Aermet100 steel, Materials Science and Engineering: A 528(13-14) (2011) 4774-4782.

[27] T. Furukawa, G. Yagawa, Implicit constitutive modelling for viscoplasticity using neural networks, International Journal for Numerical Methods in Engineering 43(2) (1998) 195-219.

[28] Y. Hashash, S. Jung, J. Ghaboussi, Numerical implementation of a neural network based material model in finite element analysis, International Journal for numerical methods in engineering 59(7) (2004) 989-1005.

[29] Y. Sun, W. Zeng, Y. Zhao, Y. Qi, X. Ma, Y. Han, Development of constitutive relationship model of Ti600 alloy using artificial neural network, Computational Materials Science 48(3) (2010) 686-691.

[30] W.E. Faller, S.J. Schreck, Unsteady fluid mechanics applications of neural networks, Journal of aircraft 34(1) (1997) 48-55.

[31] D. Wang, W. Liao, Modeling and control of magnetorheological fluid dampers using neural networks, Smart Materials and Structures 14(1) (2004) 111.

[32] Z. Yuhong, H. Wenxin, Application of artificial neural network to predict the friction factor of open channel flow, Communications in Nonlinear Science and Numerical Simulation 14(5) (2009) 2373-2378.

[33] T. Butz, O. Von Stryk, Modelling and Simulation of Electro- and Magnetorheological Fluid Dampers, ZAMM- Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik: Applied

Mathematics and Mechanics 82(1) (2002) 3-20.

[34] R. Beigzadeh, M. Rahimi, Prediction of heat transfer and flow characteristics in helically coiled tubes using artificial neural networks, International Communications in Heat and Mass Transfer 39(8) (2012) 1279-1285.

[35] Y. Mi, M. Ishii, L. Tsoukalas, Flow regime identification methodology with neural networks and two-phase flow models, Nuclear engineering and design 204(1-3) (2001) 87-100.

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

[37] R. Liu, Y.C. Yabansu, A. Agrawal, S.R. Kalidindi, A.N. Choudhary, Machine learning approaches for elastic localization linkages in high-contrast composite materials, Integrating Materials and Manufacturing Innovation 4(1) (2015) 13.

[38] R. Kondo, S. Yamakawa, Y. Masuoka, S. Tajima, R. Asahi, Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics, Acta Materialia 141 (2017) 29-38.

[39] R. Cang, H. Li, H. Yao, Y. Jiao, Y. Ren, Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model, Computational Materials Science 150 (2018) 212-221.

[40] M. Bessa, R. Bostanabad, Z. Liu, A. Hu, D.W. Apley, C. Brinson, W. Chen, W.K. Liu, A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality, Computer Methods in Applied Mechanics and Engineering 320 (2017) 633-667.

[41] F.J. Pettijohn, Sedimentary rocks, (1957).

[42] S. Boggs, Petrology of sedimentary rocks, Cambridge University Press2009.

[43] S. Bernard, R. Wirth, A. Schreiber, L. Bowen, A. Aplin, E. Mathia, H. Schulz, B. Horsfield, A. Aplin, E. Mathia, FIB-SEM and TEM investigations of an organic-rich shale maturation series from the Lower Toarcian Posidonia Shale, Germany: Nanoscale pore system and fluid-rock interactions, Electron microscopy of shale hydrocarbon reservoirs: AAPG Memoir 102 (2013) 53-66.

[44] N. Ohkouchi, J.i. Kuroda, M. Okada, H. Tokuyama, Why Cretaceous black shales have high C/N ratios: Implications from SEM-EDX observations for Livello Bonarelli black shales at the Cenomanian-Turonian boundary, Frontier research on Earth evolution 1 (2003) 239-241.

[45] S. Abedi, M. Slim, R. Hofmann, T. Bryndzia, F.-J. Ulm, Nanochemo-mechanical signature of organic-rich shales: a coupled indentation–EDX analysis, Acta Geotechnica 11(3) (2016) 559-572.

[46] S. Kelly, H. El-Sobky, C. Torres-Verdín, M.T. Balhoff, Assessing the utility of FIB-SEM images for shale digital rock physics, Advances in water resources 95 (2016) 302-316.

[47] P. Tahmasebi, F. Javadpour, M. Sahimi, Three-dimensional stochastic characterization of shale SEM images, Transport in Porous Media 110(3) (2015) 521-531.

[48] C.D. Foster, T.M. Nejad, Embedded discontinuity finite element modeling of fluid flow in fractured porous media, Acta Geotechnica 8(1) (2013) 49-57.

[49] J.F. Barthélémy, C. Souque, J.M. Daniel, Nonlinear homogenization approach to the friction coefficient of a quartz- clay fault gouge, International Journal for Numerical and Analytical Methods in Geomechanics 37(13) (2013) 1948-1968.

[50] J.A. White, Anisotropic damage of rock joints during cyclic loading: constitutive framework and numerical integration, International Journal for Numerical and Analytical Methods in Geomechanics 38(10) (2014) 1036-1057.

[51] A.H. Kohli, M.D. Zoback, Frictional properties of shale reservoir rocks, Journal of geophysical research: solid earth 118(9) (2013) 5109-5125.

[52] C. Yeong, S. Torquato, Reconstructing random media, Physical Review E 57(1) (1998) 495.

[53] S. Torquato, Statistical description of microstructures, Annual review of materials research 32(1) (2002) 77-111.

[54] D.B. Shaw, C.E. Weaver, The mineralogical composition of shales, Journal of Sedimentary Research 35(1) (1965).

[55] A. Deirieh, J. Ortega, F.J. Ulm, Y. Abousleiman, Nanochemomechanical assessment of shale: a coupled WDS-indentation analysis, Acta Geotechnica 7(4) (2012) 271-295.

[56] V. Kumar, Geomechanical characterization of shale using nano-indentation, University of Oklahoma Norman, OK, USA, 2012.

[57] B. Gathier, Multiscale strength homogenization: application to shale nanoindentation, Massachusetts Institute of Technology, 2008.

[58] D. Cule, S. Torquato, Generating random media from limited microstructural information via stochastic optimization, Journal of applied physics 86(6) (1999) 3428-3437.

[59] B. Lu, S. Torquato, Lineal-path function for random heterogeneous materials, Physical Review A 45(2) (1992) 922.

[60] B. Lu, S. Torquato, Lineal-path function for random heterogeneous materials. II. Effect of polydispersivity, Physical Review A 45(10) (1992) 7292.

[61] B. Yegnanarayana, Artificial neural networks, PHI Learning Pvt. Ltd.2009.

[62] R. Lippmann, An introduction to computing with neural nets, IEEE Assp

magazine 4(2) (1987) 4-22.

[63] R.J. Schalkoff, Artificial neural networks, McGraw-Hill New York1997.

[64] A.K. Jain, J. Mao, K.M. Mohiuddin, Artificial neural networks: A tutorial, Computer 29(3) (1996) 31-44.

[65] M.A. Nielsen, Neural networks and deep learning, Determination Press2015.

[66] M.H. Hassoun, Fundamentals of artificial neural networks, MIT press1995.

[67] K. Fukushima, S. Miyake, Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition, Competition and cooperation in neural nets, Springer1982, pp. 267-285.

[68] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).

[69] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, European conference on computer vision, Springer, 2014, pp. 818-833.

[70] S. Lawrence, C.L. Giles, A.C. Tsoi, A.D. Back, Face recognition: A convolutional neural-network approach, IEEE transactions on neural networks 8(1) (1997) 98-113.

[71] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, 2012, pp. 1097-1105.

[72] D. Ciregan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, Computer vision and pattern recognition (CVPR), 2012 IEEE conference on, IEEE, 2012, pp. 3642-3649.

[73] M.D. Zeiler, G.W. Taylor, R. Fergus, Adaptive deconvolutional networks for mid and high level feature learning, Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE, 2011, pp. 2018-2025.

[74] M.D. Zeiler, D. Krishnan, G.W. Taylor, R. Fergus, Deconvolutional networks, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, 2010, pp. 2528-2535.

[75] K.J. Bathe, Finite element method, Wiley Online Library2008.

[76] X.-c. Wang, Finite element method, Tsinghua University Press, Beijing, 2003.

Appendix

Linear elastic finite element method

In this section, we briefly discussed the linear-elastic finite element model used in this paper [75, 76]. In the context of small strain, the strain tensor is defined by the displacement gradient as

$$\boldsymbol{\varepsilon}_e = \frac{1}{2}\left(\nabla \mathbf{u} + (\nabla \mathbf{u})^{\mathrm{T}}\right) \tag{12}$$

With respect to the linear elastic constitutive model, the stress and strain tensor complies the following linear relation,

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\varepsilon}_e \tag{13}$$

In a quasi-static system, the energy balance of an elastic solid is illustrated as

$$W_{ext} = W_{int} \tag{14}$$

W_{int} is the internal energy defined by the elastic deformation of the solid body.

$$W_{int}\left(\boldsymbol{\varepsilon}_e\right) = \int_{\Omega} \frac{1}{2} \boldsymbol{\varepsilon}_e : \mathbf{C} : \boldsymbol{\varepsilon}_e \mathrm{d}\Omega \tag{15}$$

$\mathbf{C}$ is the stiffness matrix of the solid body, and $\boldsymbol{\varepsilon}_e$ is the strain W_{ext} is the external energy contributed by the body force and the boundary traction.

$$W_{ext} = \int_{\Omega} \mathbf{b} \cdot \mathbf{u} \mathrm{d}\Omega + \int_{\partial\Omega} \mathbf{h} \cdot \mathbf{u} \mathrm{d}\partial\Omega \tag{16}$$

The energy balance should also hold for the variation of the internal and external energy

$$\delta W_{ext} = \delta W_{int} \tag{17}$$

The variation of the internal and external energy are respectively given as

$$\delta W_{int}\left(\boldsymbol{\varepsilon}_e\right) = \frac{\partial W_{int}}{\partial \boldsymbol{\varepsilon}} \delta\boldsymbol{\varepsilon} = \int_{\Omega} \boldsymbol{\varepsilon}_e : \mathbf{C} : \delta\boldsymbol{\varepsilon}_e \mathrm{d}\Omega \tag{18}$$

$$\delta W_{ext} = \int_{\Omega} \mathbf{b} \cdot \delta\mathbf{u} \mathrm{d}\Omega + \int_{\partial\Omega} \mathbf{h} \cdot \delta\mathbf{u} \mathrm{d}\partial\Omega \tag{19}$$

Eq. (18) and Eq. (19) are substituted into Eq. (17) to give the strong form of the equilibrium equation to describe solid deformation.

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma} + \mathsf{b} = 0 & in\,\Omega, \\ \mathsf{u} = \bar{\mathsf{u}} & on\,\Omega_{\mathsf{u}}, \\ \boldsymbol{\sigma} \cdot \mathsf{n} = \bar{\mathsf{h}} & on\,\partial\Omega_{\mathsf{h}}, \end{cases} \tag{20}$$

The Galerkin weak forms are derived based on strong form boundary value equations in Eq. (20).

$$\int_{\Omega} \left[\boldsymbol{\sigma} : \delta\boldsymbol{\varepsilon}_{e} \right] \mathrm{d}\Omega - \int_{\partial\Omega_{\mathsf{h}}} \left[\bar{\mathsf{h}} \cdot \delta\mathsf{u} \right] \mathrm{d}\partial\Omega = 0 \tag{21}$$

In the above equations, $\delta\mathsf{u}$ is variational test functions of displacement. Thus, displacement is expressed using interpolation of nodal variables.

$$\mathsf{u} = \sum_{I=1}^{4} \mathsf{N}_{I} \mathsf{u}_{I} \tag{22}$$

In Eq. (22), $\mathsf{N}_{I} = \begin{bmatrix} N_{I} & 0 \\ 0 & N_{I} \end{bmatrix}$ is shape function matrix for displacement field by introducing Voigt notation. The gradient matrix of shape function matrix N can be defined correspondingly as

$$\mathsf{B}_{I} = \begin{bmatrix} N_{I,x} & 0 \\ 0 & N_{I,y} \\ N_{I,y} & N_{I,x} \end{bmatrix} \tag{23}$$

The gradient of displacement can be discretized using B matrices as

$$\boldsymbol{\varepsilon} = \sum_{I=1}^{4} \mathbf{B}_{I} \mathsf{u}_{I} \tag{24}$$

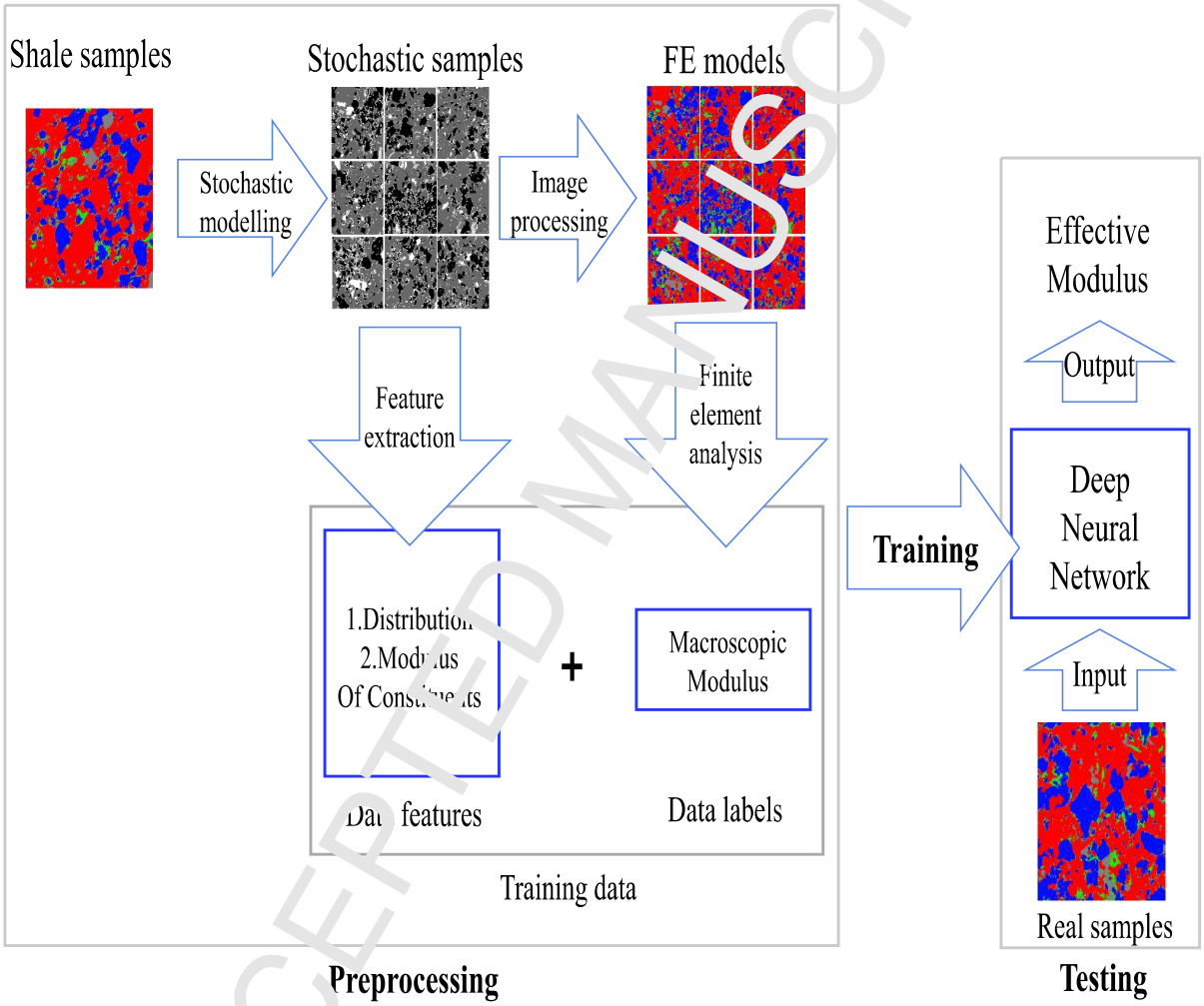
By inserting discretized expressions of primary variables and their gradients into Eq. (21), the discretized format of the residual of the weak form equation is obtained.

$$\mathsf{R}_{I} = \int_{\Omega} \left[\boldsymbol{\sigma} \mathsf{B}_{I} \right] \mathrm{d}\Omega - \int_{\partial\Omega_{\mathsf{h}}} \left[\mathsf{N}_{I} \bar{\mathsf{h}} \right] \mathrm{d}\partial\Omega \tag{25}$$

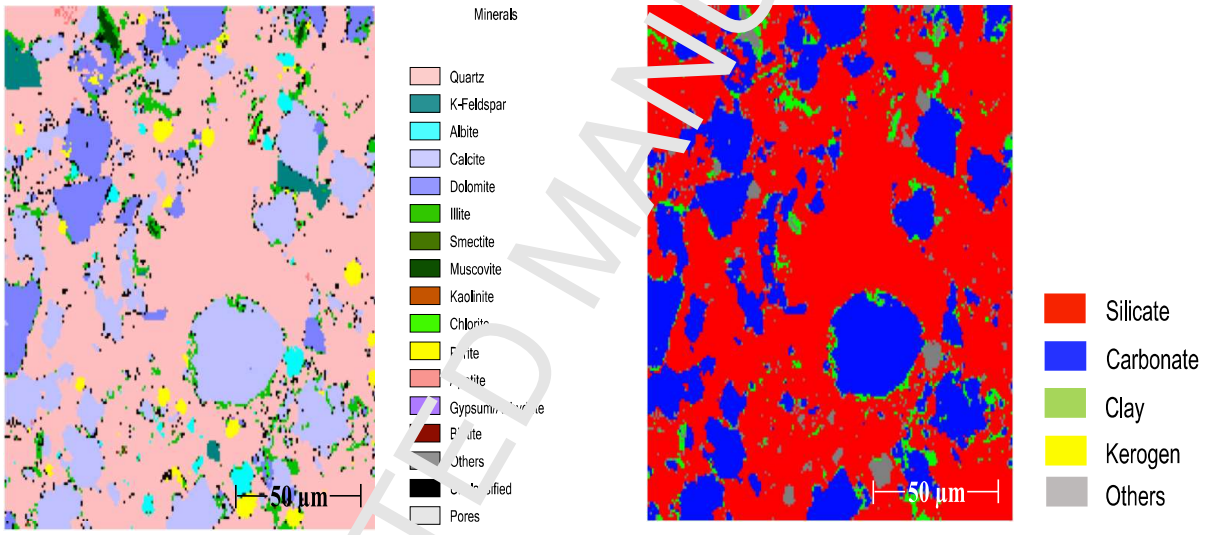
The corresponding tangent stiffness matrices of the above equations are

$$\mathsf{K}_{IJ} = \frac{\partial \mathsf{R}_{I}}{\partial \mathsf{u}_{J}} = \int_{\Omega} \left[\left(\mathsf{B}_{I} \right)^{\mathrm{T}} \mathsf{C} \mathsf{B}_{J} \right] \mathrm{d}\Omega \tag{26}$$

These discretized residual equations are implicitly solved using a linear solver, and the displacements of all nodes are calculated.



(a) (b)



(a) (b)

Figure 4

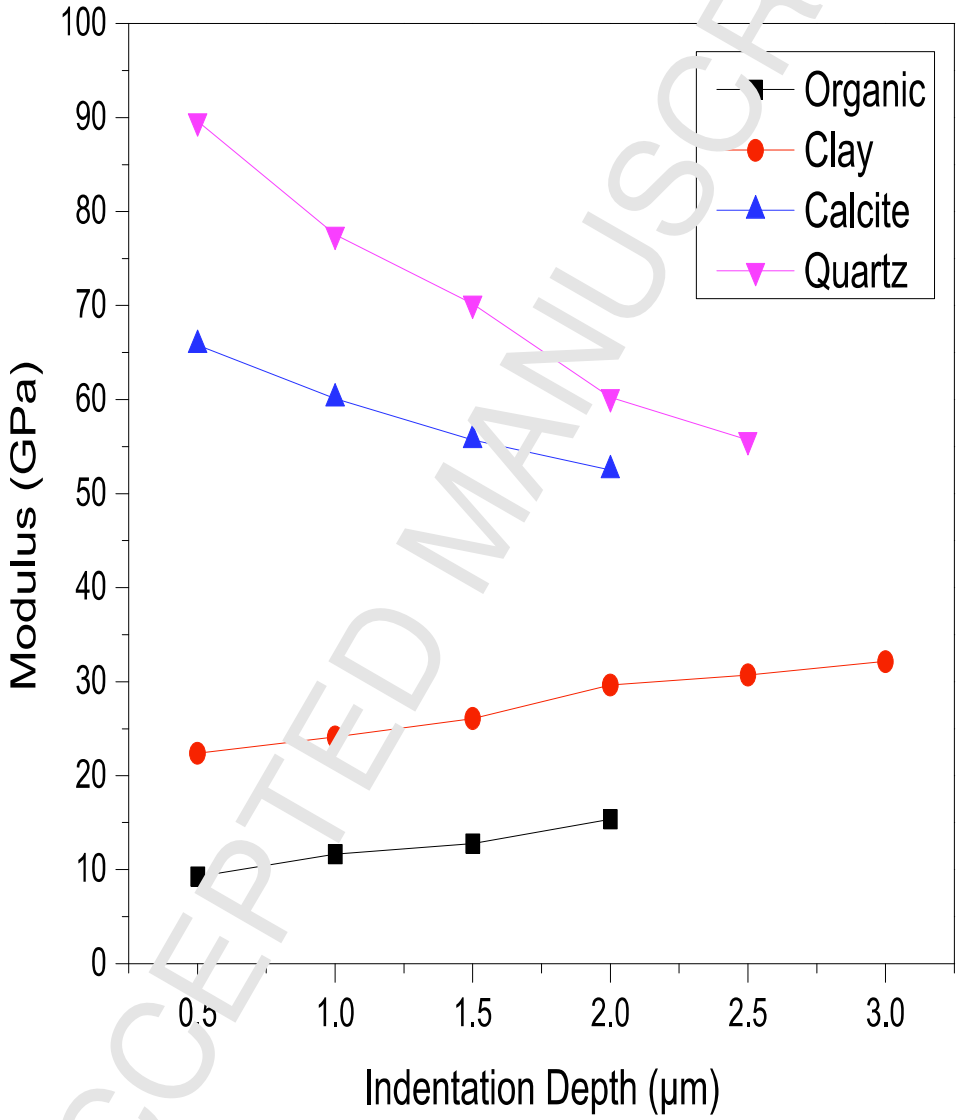
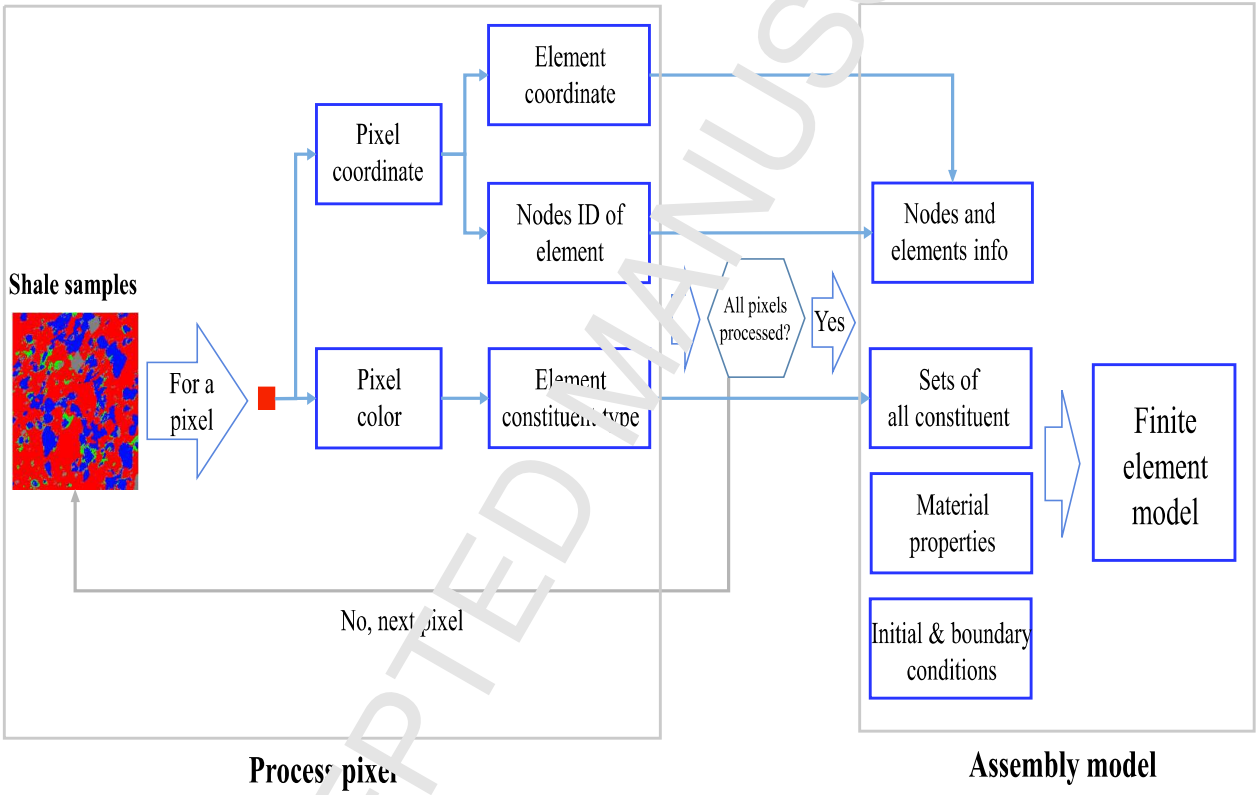
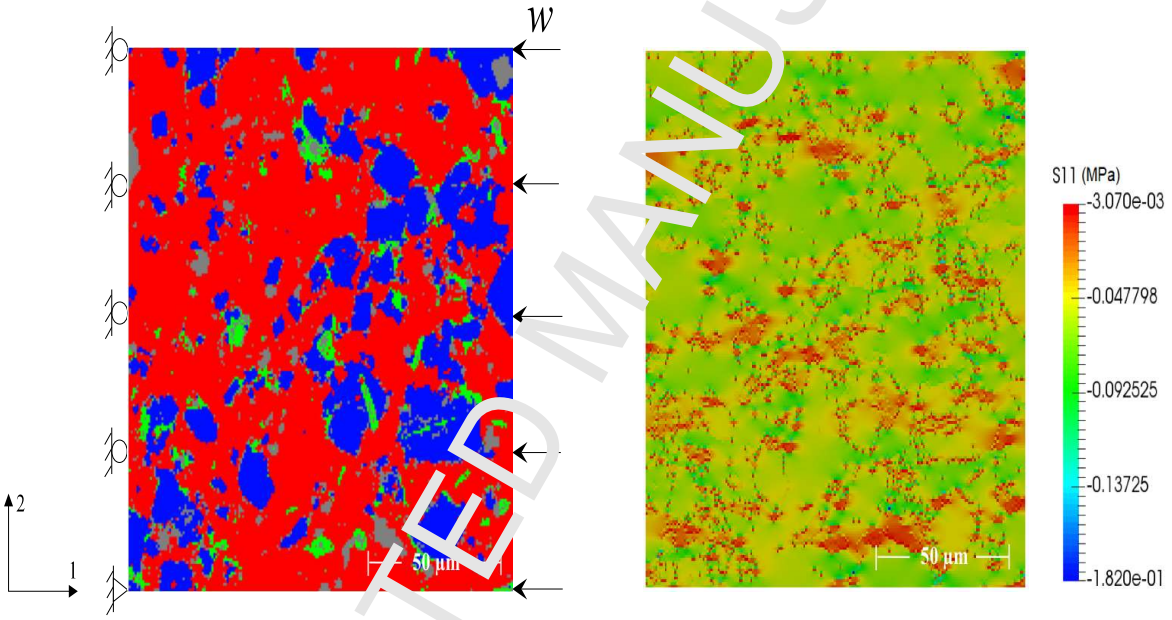


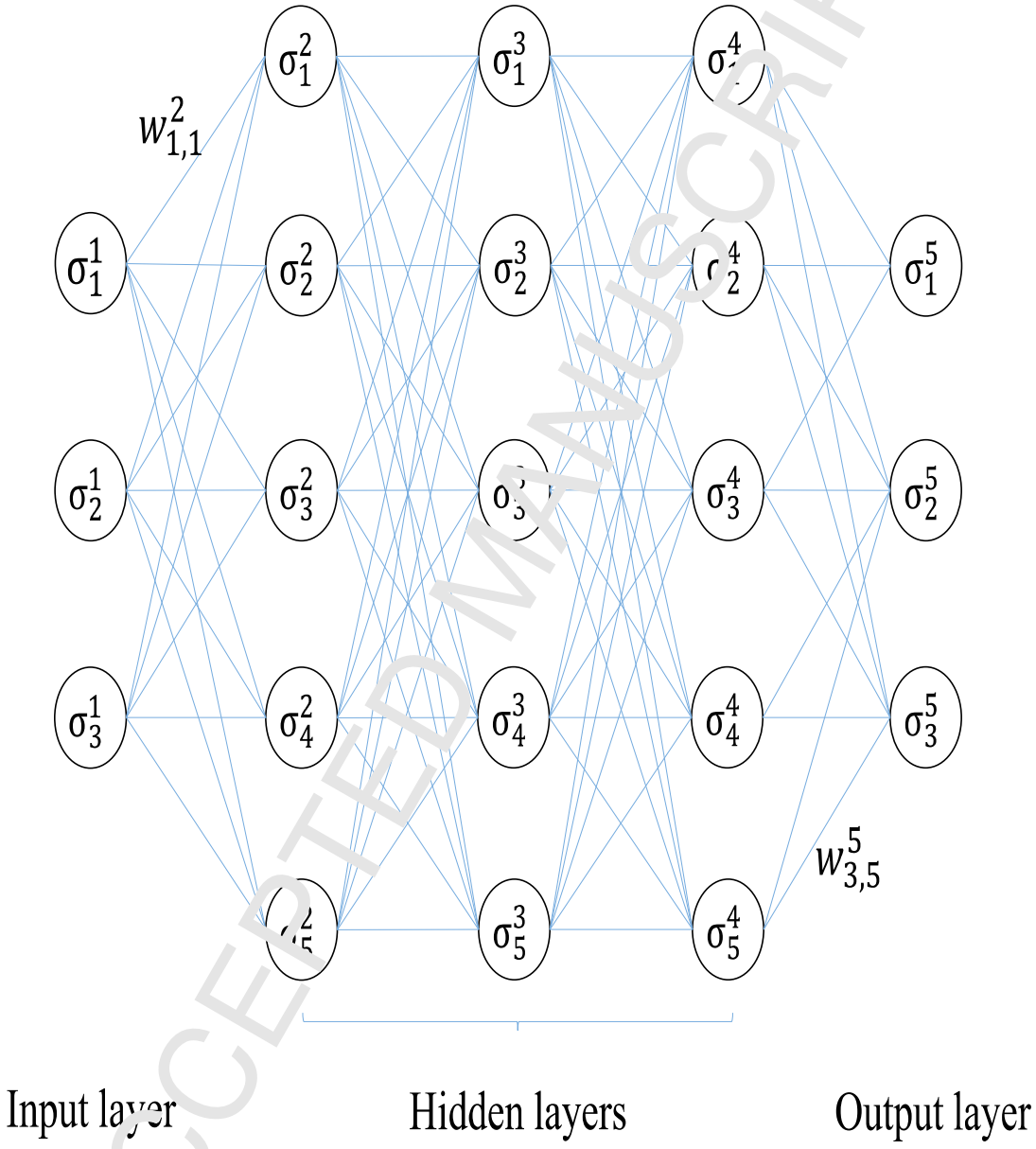
Figure 5

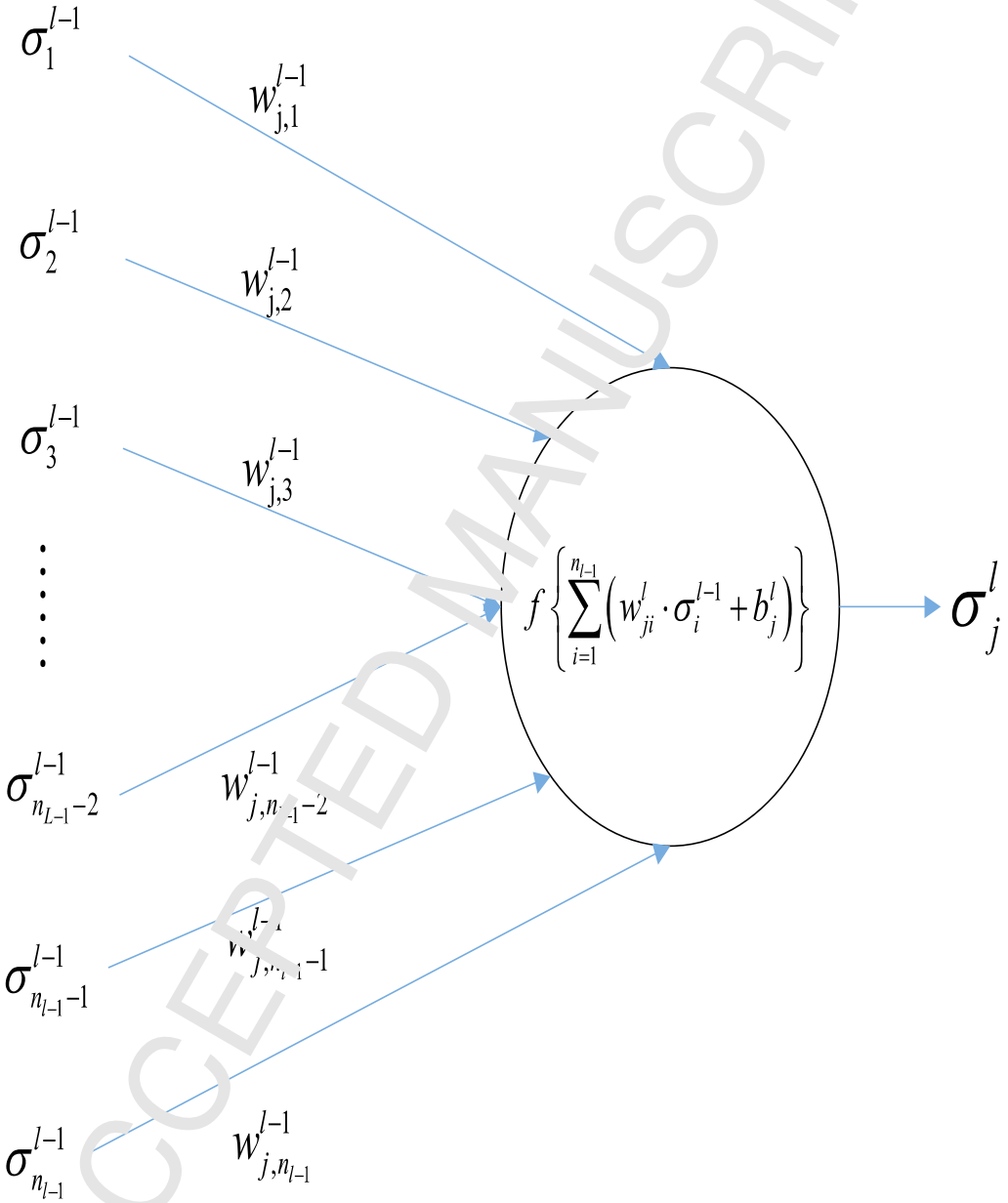


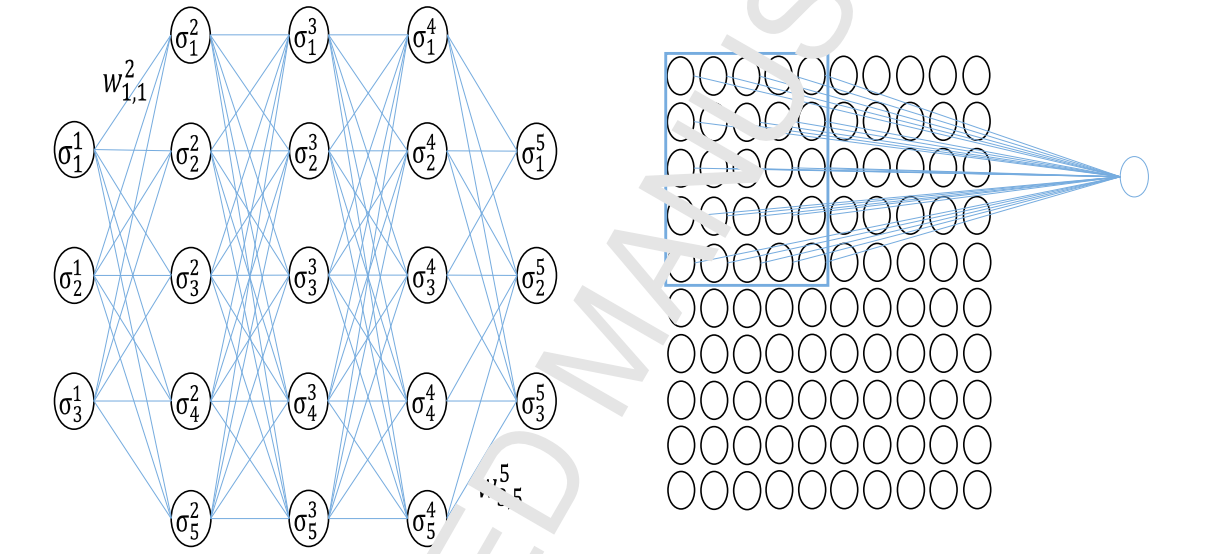


(a) (b)

Figure 7







(a) (b)

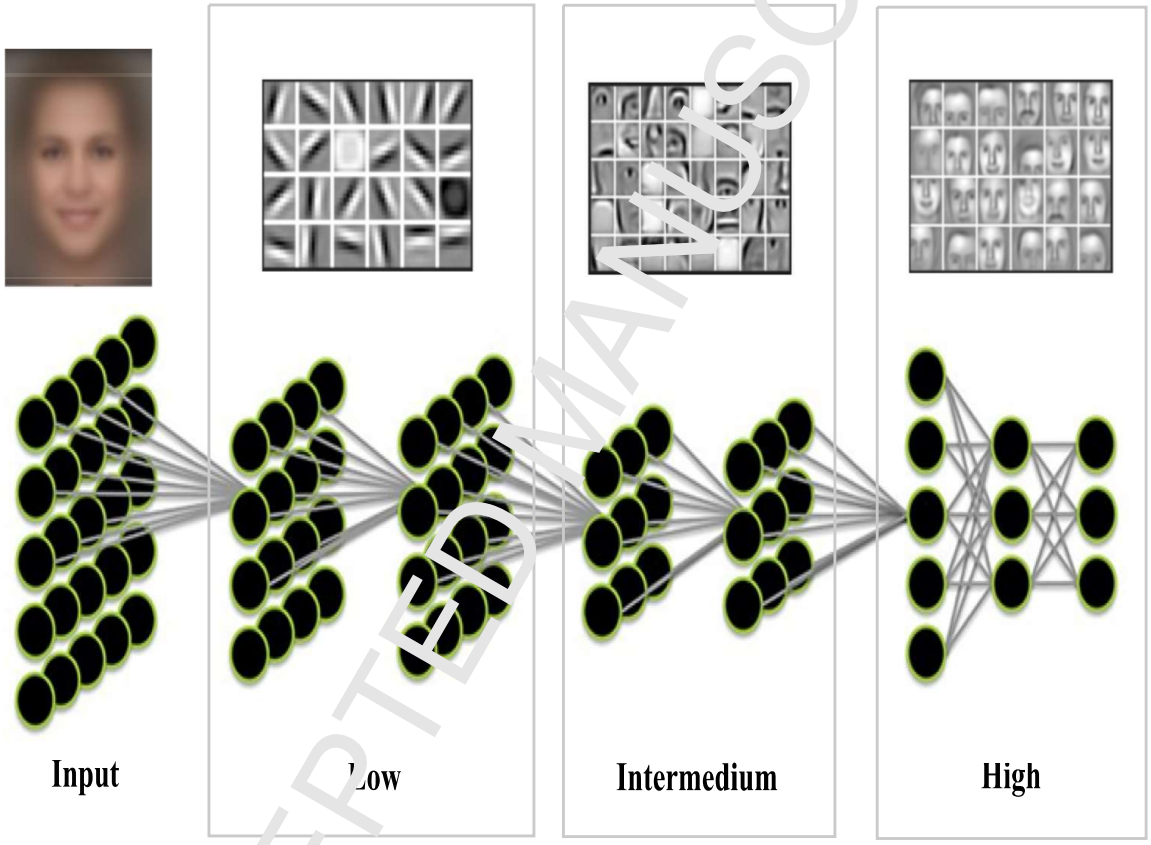
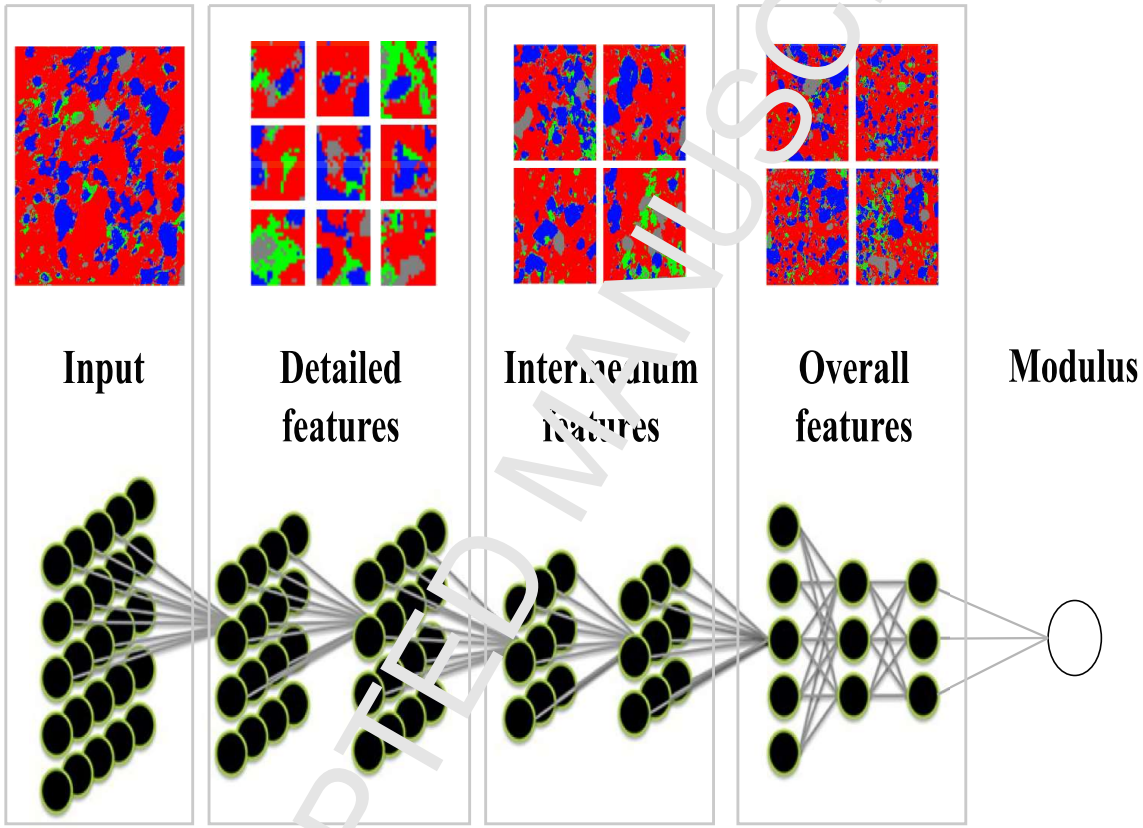


Figure 11



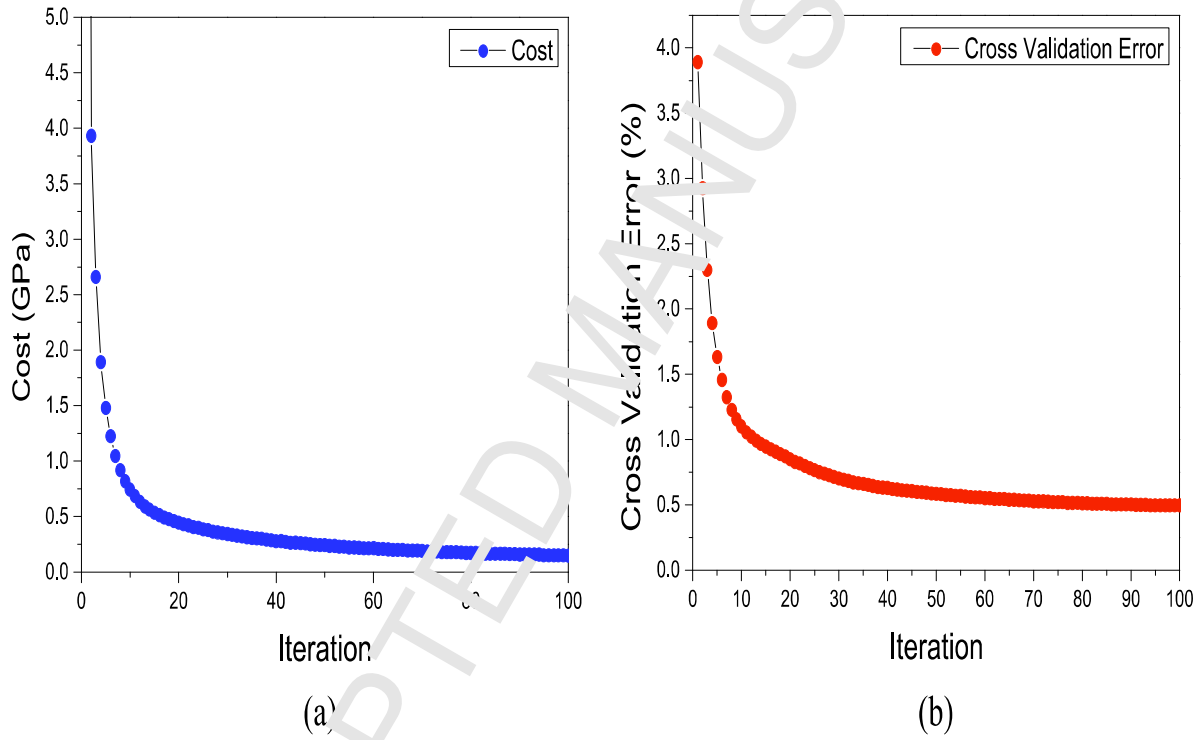


Figure 13

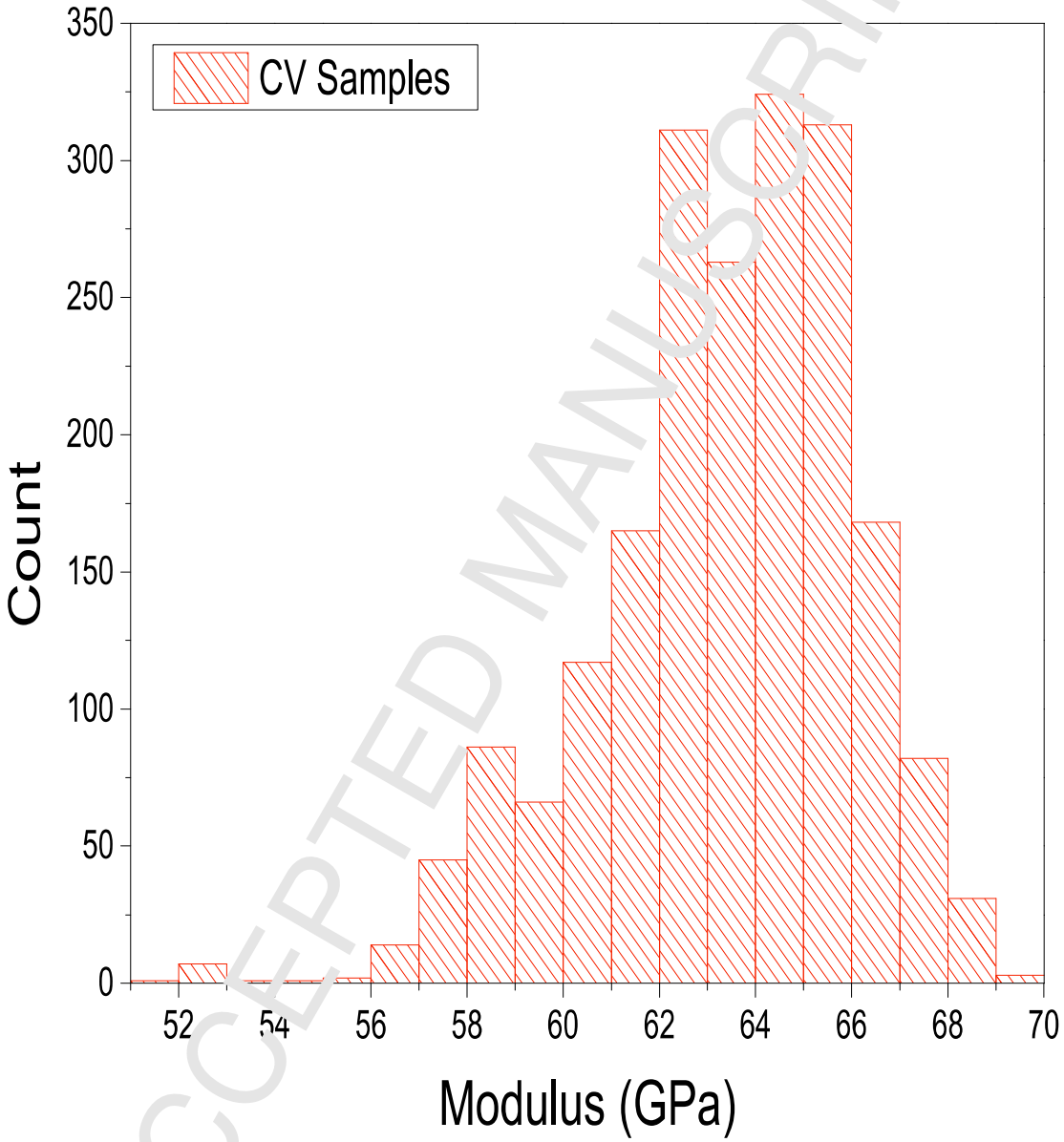


Figure 14

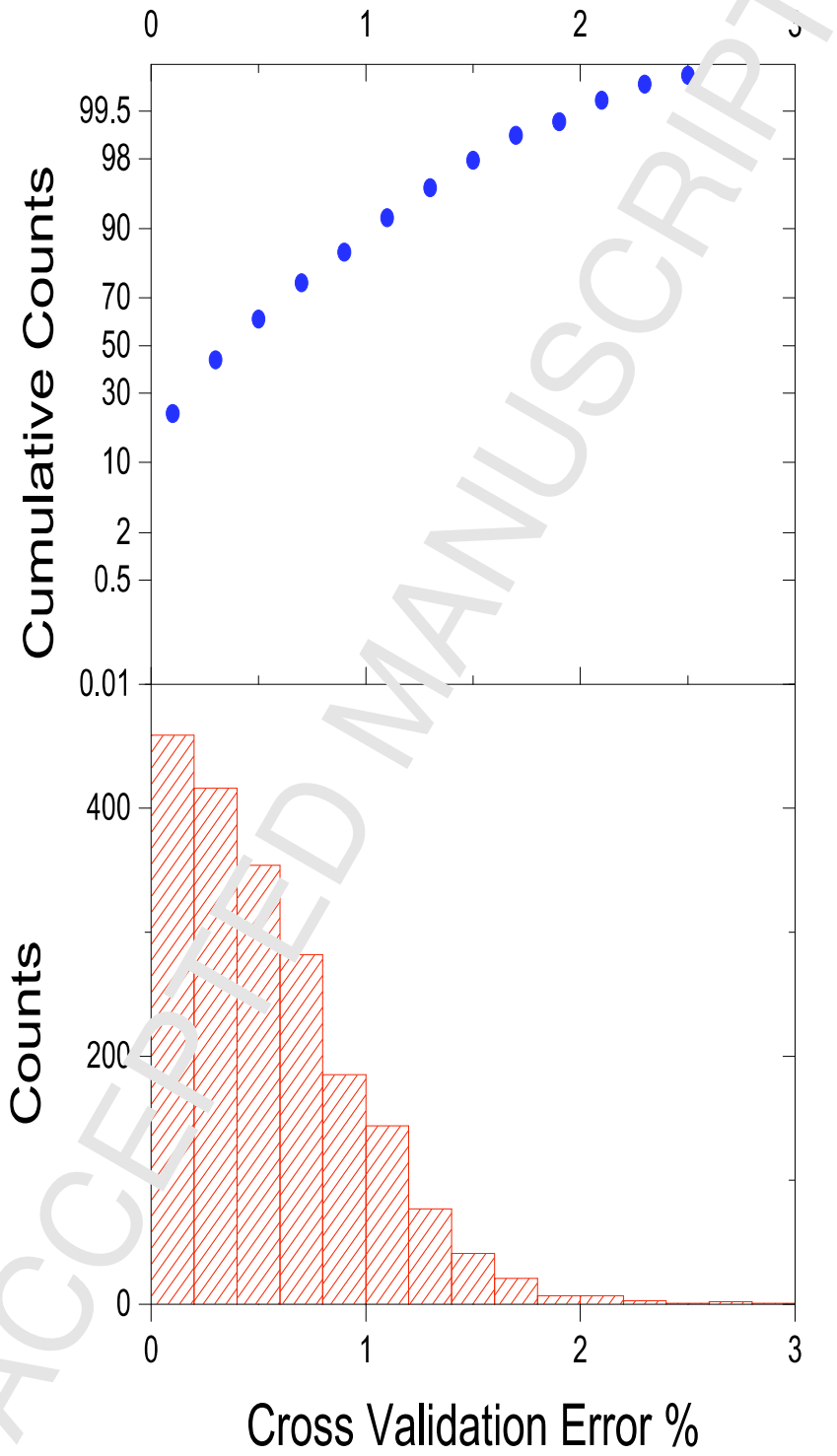
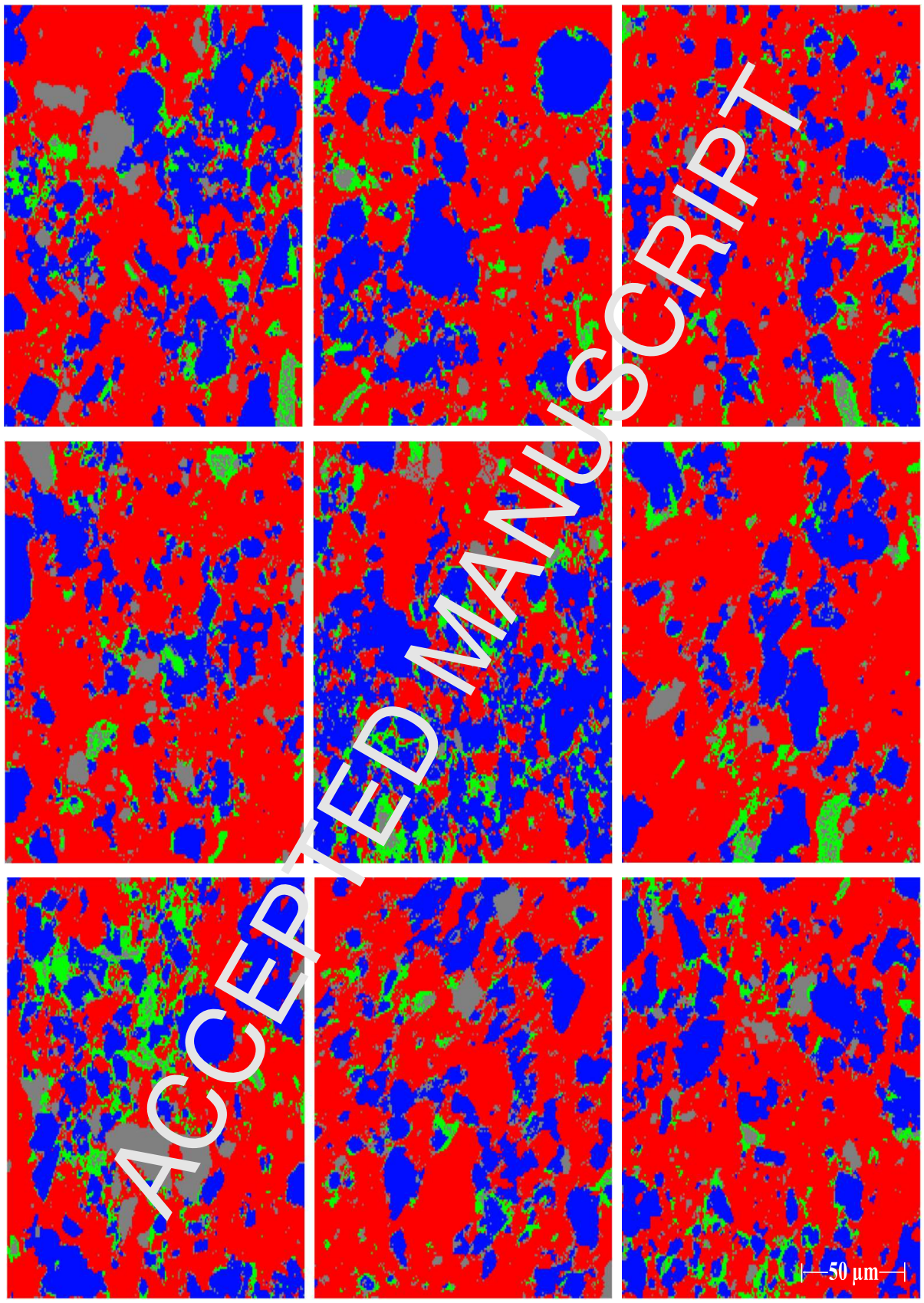
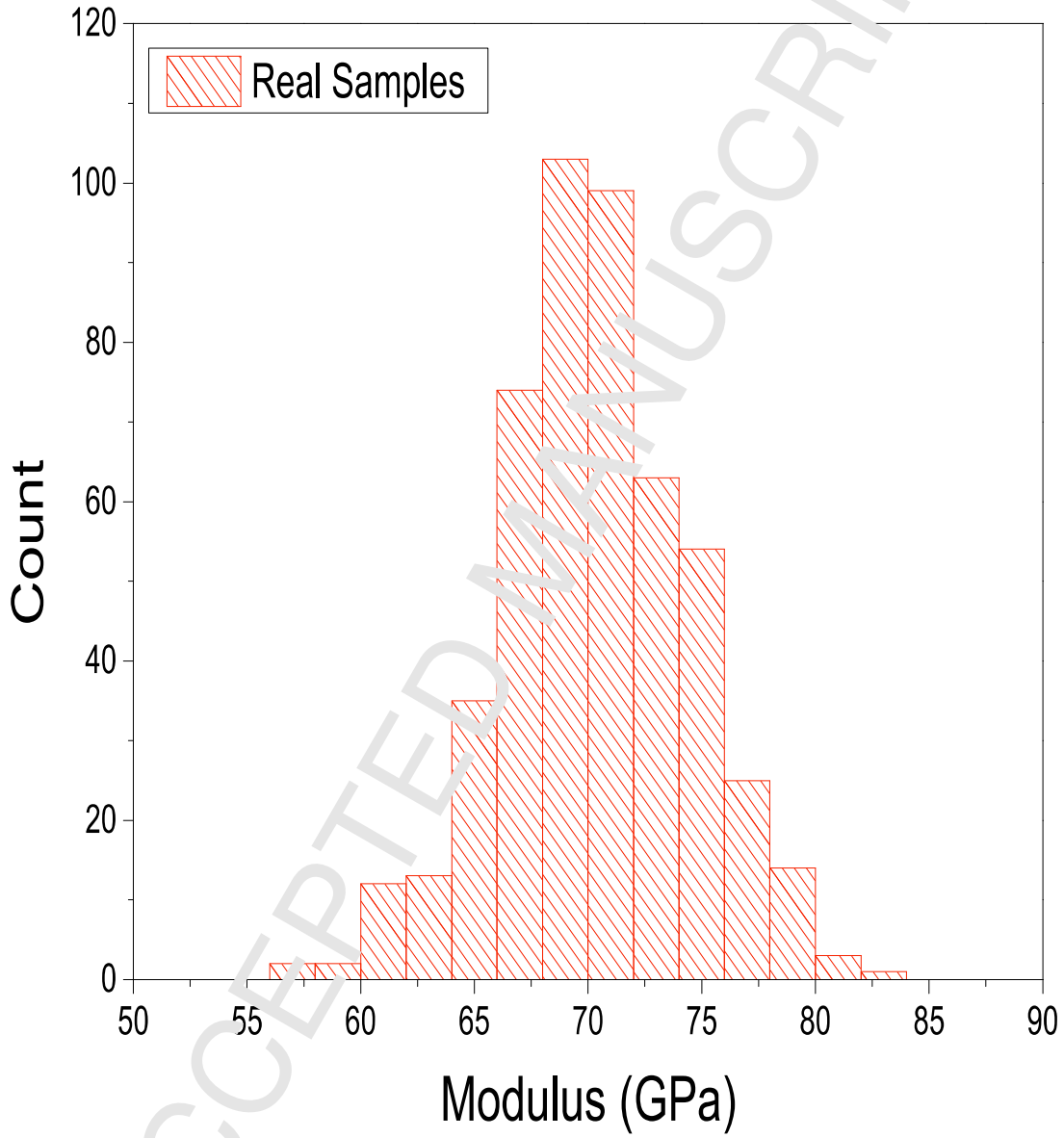
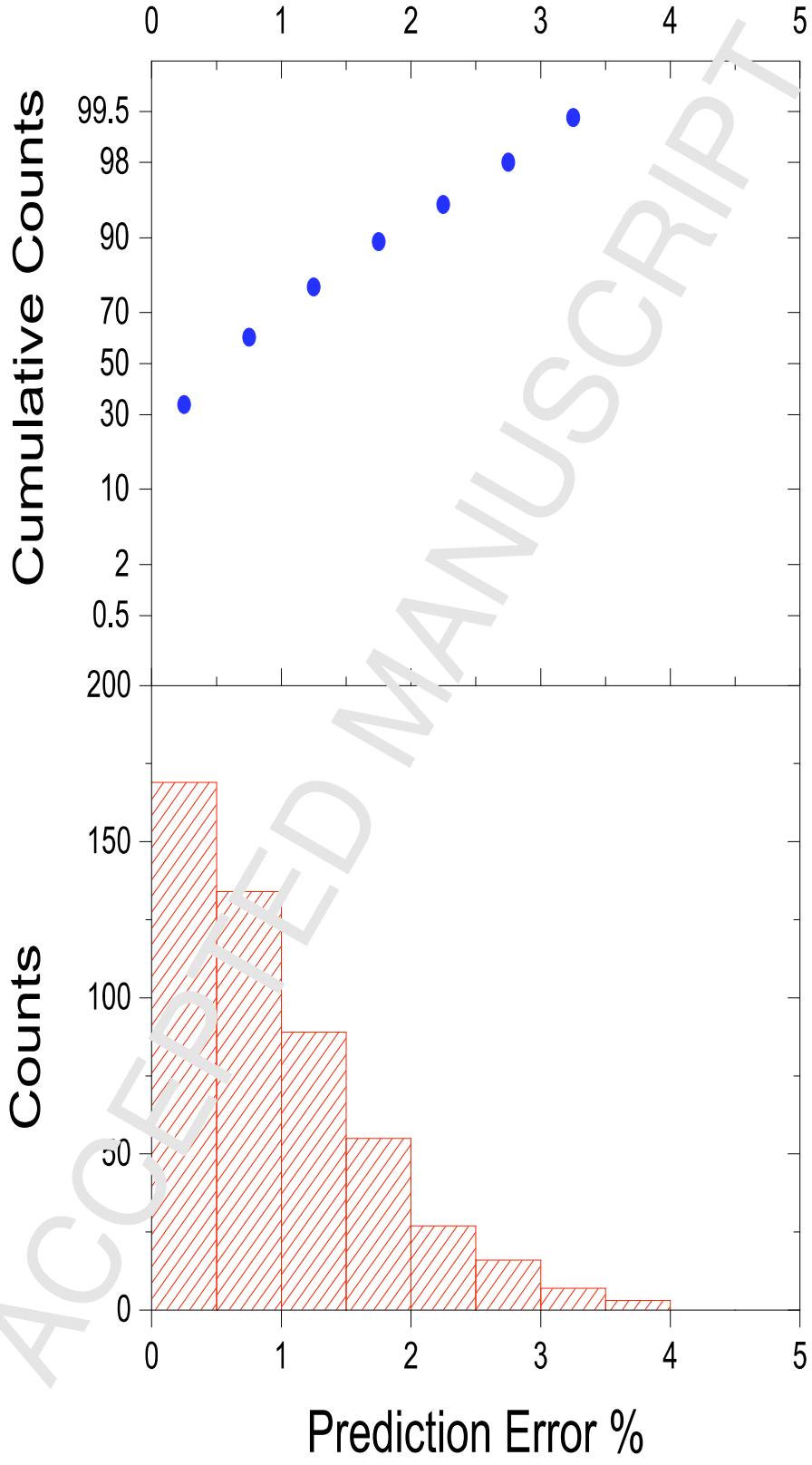


Figure 15







Tab. 1 The forming constituents of each phase of the mesoscale shale model

Phase	Forming constituents
Silicate	Quartz, feldspar, pyrite
Carbonate	Calcite, dolomite
Clay	Kaolinite, illite, chlorite, montmorillonite
Kerogen	Organic matter
Others	Other mineral constituents

Tab. 2 The measured moduli of the primary constituents in shale.

	Silicate	Carbonate	Clay	Kerogen	Matrix
Modulus (GPa)	89.6	65.8	22.3	9.2	12.392