A Deep Learning Approach Replacing the Finite Difference Method for In situ Stress Prediction

WENLI GAO[ID]1, XINMING LU1, YANJUN PENG1, AND LIANG WU2

1College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China
2School of Control Science and Engineering, Shandong University, Jinan 250000, China

Corresponding author: Xinming Lu (13853803297@139.com)

ABSTRACT In the domain of geotechnical engineering analysis, fast Lagrangian analysis based on the finite difference method is the most commonly used numerical analysis method. It is a classical numerical algorithm applied to calculate the in situ stress. In this paper, we propose a deep learning (DL) architecture called the enhance-and-split feature capsule network embedded in fully convolutional neural networks (ES-Caps-FCN) to predict the in situ stress for a strain-softening model when using the finite difference method for the numerical computation. Experiments indicate that this novel approach is stable and convergent. Compared with some classical prediction methods including linear regression analysis and a deep neural network, the mean squared error of our proposed algorithm is as low as 0.059866%, which is lower than the 0.616676% of the deep neural network prediction algorithm and the 0.978495% of the conventional machine learning algorithm. Additionally, the calculation efficiency of fully trained deep learning models is superior to that of the conventional finite difference method. Therefore, DL is a feasible and promising fast and accurate surrogate for the finite difference method for solving the in situ stress.

INDEX TERMS Deep learning, in situ stress field, DNN, ES-Caps-FCN.

I. INTRODUCTION

In situ stress [1] refers to the intrinsic stress of the crust and rock formations in their original state without being disturbed by artificial engineering. In situ stress is the source of the force causing various mine pressures to appear and rock mass deformation and displacement. It is also a mechanical circumstance that plays a vital role in the coal mining process, rock-soil excavation and oil exploration and development. As a result, understanding its distribution situation plays a major role in reasonably determining the layout of a stope, mining procedures, roadway support forms and parameters, etc. The current measurement techniques of in situ stress are relatively mature, and include the ground stress relief method, hydraulic fracturing method [2], [3], anelastic strain recovery method [4], Kaiser effect method [5], deformation rate analysis method [6], [7], differential strain curve analysis method [8] and the latest Ultrasonic testing method [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Jenny Mahoney.

Although directly measuring the in situ stress can accurately obtain the in situ stress at a certain point, there are some drawbacks including high measurement costs, high time consumption, the local characteristics of the results and the problems from applying multipoint stress analysis to perform direct measurements. To overcome these shortcomings, geologic body modeling by computer and numerical simulations are effective methods to obtain more and broader geostress data. Common numerical calculation and analysis methods for the stress deformed state of rocks include the boundary element method [10], discrete element method [11], finite element method (FEM) [12], finite difference method (FDM), combined finite-discrete element method [13], displacement inverse analysis method, etc. In addition, there are mechanical engineering numerical simulations for calculating the complexity of time-dependent models in various engineering applications. Based on those numerical calculation methods, abundant approaches for in situ stress computation and analysis have been developed, and the typical in situ stress inference methods are the multivariate regression [14], [15]

and neural network. Generally, numerical simulation analysis transforms basic equations and boundary conditions into linear or nonlinear equations, high-order algebraic equations or integral equations via different processing methods. The properties of rock masses are very complex (such as discontinuous and nonloose media, inelastic, nonplastic, viscous, etc.), and the influence of temperature fields and complex boundary conditions also should be considered. Despite simplifying the model, this process is still very complicated and time-consuming. In the field of mechanical analysis, a great amount of efforts have been made to directly determine stress distributions according to material properties and boundary and loading conditions, thus bypassing the numerical simulation calculation process.

Machine learning [16] is dedicated to mining the implicit rules of data sets from large amounts of data to predict or classify the expected results and hopes that the resulting models possess excellent generalization capabilities. In a deep learning model [17], [18], each layer of networks reasonably composes and processes these features through perceiving data via human brain-like neurons; therefore, the high-layer can continually get more abstract and implicit features. Machine learning and deep learning have become increasingly more widespread in computational mechanics and mechanical analysis [19]–[21]. For instance, Liang *et al.* [22] developed a deep learning model to directly estimate the stress distributions of the aorta. In addition, in [23], they also used machine learning to estimate the zero-pressure geometry for two given pressurized geometries of the same patient at two different blood pressure levels for the human thoracic aorta. Roewer-Despres *et al.* [20] approximated the large deformations of a nonlinear, muscle actuated beam by using a deep-autoencoder to simulate soft tissue biomechanics. Kononenko and Kononenko [24] demonstrated that machine learning can be used as a supplement to the finite element analysis of physical system modeling, and it can efficiently address some corresponding problems. Moreover, neural network algorithms have also been used to replace the conventional constitutive material model to improve finite element analysis [25]. Machine learning and neural networks also perform well in other civil and geotechnical engineering research areas, such as the prediction and classification of rockburst [26], [27].

All the above studies have focused on replacing of the finite element method to solve the mechanical component, while fast Lagrangian analysis based on FDM is a more professional and common numerical analysis method to calculate the in situ stress in geotechnical engineering. Therefore, this paper attempts to imitate it to calculate the in situ stress and discusses the feasibility of substituting it via deep learning. The main objective of this paper is to verify feasibility of using DL as an alternative method for fast Lagrangian analysis numerical simulation. We focus on predicting the three dimensional stress components depending on the relationship between the stress values and loads, boundary conditions, and geometric figures by training some critical features.

Previous classical prediction and calculation methods for the in situ stress based on neural networks simply find numerical nonlinear relations between stress components and the boundary conditions, which ignores the influence of the direction and location of the stress component and boundary force. Furthermore, they cannot explicitly express the geometrical information of the model. These factors are significant for the training of neural networks, especially when the external force value of each boundary surface is not constant, which necessitates setting the variational external load. To address these problems, a deep learning architecture is proposed. We explicitly describe the geometric shape of the geologic body model using a three dimensional matrix and the boundary constraints and external force values using the external boundary coordinates. We directly process the three-dimensional matrix data to maintain the positional relationship between the external forces and stress components via a three dimensional convolutional neural network, and take full advantage of the orientation and position sensitivity of the capsule network. Since the external force only acts on the outer boundary, the resulting matrix is quite sparse. We apply feature enhancement to reduce the impact of the sparse matrix on the prediction results.

The remaining parts of this paper are structured as follows. In section II, we review the most closely related literature on 1) the finite difference method and the general modeling process of the FLAC3D, and 2) the in situ stress inference and prediction methods. Then, in Section III, we present our proposed DL architecture and provide the details of the networks used for the in situ stress prediction. We describe the dataset and design comparison algorithms in Section IV, which is followed by the experimental results and discussion. Finally, some concluding remarks are made in Section V.

II. RELATED WORK
A. FAST LAGRANGIAN ANALYSIS METHOD BASED ON THE FDM

The FDM [28] directly turns the differential problem into an algebraic problem. The key is to establish unknown algebraic equations for the grid node values. First, the solution area is divided into a difference mesh, and a finite number of mesh nodes is used to replace a continuous solution domain. There are many approaches to construct a difference scheme in a mesh region, such as the classical Taylor series expansion method. The difference between the function values on the grid nodes is used to replace the derivative for discretizing this equation, and thus the algebraic equations of the grid node values is established.

Using the principle of the finite difference method, Cundall proposed the fast Lagrangian analysis method, which adopted the mixed discrete method, dynamic relaxation method and display difference method. As a typical numerical simulation analysis method for geotechnical engineering problems, the fast Lagrangian analysis method based on FDM contributes to large deformations and finite deformations. When analyzing the mechanical structure of geotechnical

engineering, we usually obtain several basic relationships including the Equilibrium equation, Geometric equation, Constitutive equation and Boundary conditions. Since these relationships are mostly differential equations that can be approximated by difference equations, then we will solve the algebraic equations instead of solving the differential equations for these problems.

Some large-scale numerical analysis software including ANSYS, FLAC3D and ABAQUS have been developed for the easy and convenient application of numerical analysis methods in multiple fields. In particular, FLAC3D is the most commonly used large-scale simulation software in geotechnical engineering and it can well simulate the force characteristics and plastic flow of three-dimensional structures of specified characteristic materials. Finite difference modeling via FLAC3D forms the sets of data in this paper. The general processes for analyzing the in situ stress by means of FLAC3D [29] are as follows:

- Build the geometry.
- Divide grid, input the material property parameters as needed and select the constitutive model. The evenly distributed external forces are applied in both the horizontal and vertical directions to simulate in situ stress and the boundary constraints are fixed.
- Perform balance calculations, and sometimes add the initial stress values to increase the computational efficiency.

B. IN-SITU STRESS INFERENCE AND PREDICTION METHODS

Multiple regression analysis and neural networks are the key technologies for in situ stress calculation and analysis. A majority of the regression methods determine the coefficients according to the data of the measuring points and the calculated values of a numerical simulation, then calculate in situ stress of other points in the simulation region via a regression equation. Junling Hou et al. used numerical FDM for regression analysis to calculate the in situ stress for an entire geological body. Li *et al.* [30] used a 3D displacement discontinuity method (DDM) program to numerically simulate the nonlinear deformation for a fault.

Wei-Feng *et al.* [31] chose the density, elastic modulus, triaxial compressive strength, fracture rate of the core and stress as the indexes to train neural networks. Zhong-Ming *et al.* [32] established a learning function between parameters and stresses via neural networks, and then calculated simulation values of the initial in situ stress field using the finite element method for the positive sequence calculation. Yong-Song *et al.* [33] gave an intelligent expression method for the in situ stress field using a neural network method, which established a relationship between the X, Y, and Z coordinates of a node and the six corresponding stress components calculated via the finite element calculation. Ri-Qing *et al.* [34] determined the mechanics parameters of soils and retaining walls during the excavation phase by combining artificial neural networks and genetic arithmetic.

Zhang and Yin [35] used an artificial neural network to find the relationship between the in situ stress and the shape of borehole breakouts via inverse analysis.

The boundary load adjustment method and boundary displacement adjustment method respectively adjust load and displacement to obtain the optimal fitting between measured values and calculated values, which are often obtained through a numerical simulation analysis method. In addition, there are other in situ stress field calculation methods, such as the trend analysis of the geostress function, the two-stage optimization algorithm, the genetic algorithm and so on.

The in situ stress field is regarded as a linear combination of the gravity stress and tectonic stress. Based on the above viewpoint, the in situ stress component $\widehat{\sigma}_k$ is represented as a linear superposition of the computational stress σ_k^i calculated by the FEM, FDM or DDM for the first k measuring points. Then, the regression equation of $\widehat{\sigma}_k$ can be expressed as follows:

$$\widehat{\sigma}_k = \sum_{i=0}^{n} L_i \sigma_k^i \qquad (\text{II.1})$$

where L_i is the multivariable regression weight coefficient corresponding to the stress components, and n is the number of loads for the computer modeling. When the number of measuring points is sufficiently abundant, the objective function can be defined as the following formula via the least squares:

$$S_r = \sum_{k=1}^{m} \sum_{j=1}^{6} \left(\sigma_{jk}^* - L_i \sigma_{jk}^i \right)^2 \qquad (\text{II.2})$$

All regression coefficients $L = (L_1, L_2, \cdots, L_n)$ can be obtained after a set of calculations, and then the j-th stress component of the in situ stress of any point p in the calculation region can be determined by the following regression equation:

$$\sigma_{jp} = L_i \sigma_{jp}^i \qquad (\text{II.3})$$

where $j = 1, 2, \cdots, 6$ represents the 6 respective in-situ stress components.

III. MY METHODS
A. DEFINITION OF THE PHYSICAL PROBLEM
Cubes are adopted as the initial information carriers to express the geometric characteristics of solid materials. As the input of our DL model, we define a three-dimensional matrix including the geometry, boundary condition and load position information. The three dimensions of the matrix represent the x-axis, y-axis, and z-axis directions, and the length of the x-axis, y-axis, and z-axis of the three-dimensional matrix are determined according to the proportion and the geometry of the problem. Actually, we can simulate a real and complex model shape by using three dimensional matrix points, and then constrain the surface and geometric shape via the external force value. Herein, we use the (x_{min}, y_i, z_j) plane, (x_{max}, y_i, z_j) plane, (x_i, y_{min}, z_j) plane, (x_i, y_{max}, z_j) plane and

(x_i, y_j, z_{max}) plane to set the boundary conditions and external force values, which are allowed to gradually change, where $i, j = 0, 4, 8, 12 \ldots 60$. A negative value represents the compressive force, a positive value represents the tensile force, a value of 0 denotes a free boundary surface without restraint or force and a nonzero value indicates a constraint at the boundary. In many cases [36], [37], only horizontal and vertical forces are applied, and here we only study the stresses in the X, Y and Z directions after the balance calculation.

B. THE DESCRIPTION OF THE NEURAL NETWORK STRUCTURE USED IN OUR DL ARCHITECTURE

1) THREE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK(3D CNN)

A convolutional neural network [38] is a standard solution for regression and classification problems. The 3D CNN [39] is a great way to extract features via the relationships between slices. A 3D convolution can be seen as an operation between stacked two-dimensional data with a 3D convolution kernel. The value at position (x, y, z) of the jth feature map in the ith layer is as follows:

$$v_{ij}^{xyz} = \tanh\left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}\right)$$
(III.1)

where "tanh" is the activation function for the network. R_i represents the size of the 3D kernel in the time dimension, and w_{ijm}^{pqr} is the (p, q, r) of the kernel connected to the m feature graph in the previous layer.

The feature information extracted by the convolution operation in the convolution layer may be too large to further improve the learning performance. The convolution layer is followed by the pooling layer [40], which can decrease the interference of redundant information and avoid overfitting, so that the network model has strong robustness. The maximum pooling formula is as follows:

$$O(x, y, z) = \max_{(0 \le i \le s_1, 0 \le j \le s_2, 0 \le k \le s_3)} \left(I_{x*p+i, y*q+j, z*r+k}\right) \quad \text{(III.2)}$$

where I is the input of the pooling layer; O is the output; and p, q and r are the sampling steps in the three directions.

2) CAPSULE NETWORK(CapsNet)

Hinton [41] replaced the scalar of each neuron in the original neural network with a vector named a capsule. The modulus of each capsule represents the probability of occurrence of the corresponding feature. The direction of the capsule typifies some of the instantiation parameters of the corresponding feature, such as the pose (position, size, and orientation), deformation, velocity, albedo, hue, and texture. CapsNet can detect the relationship between the orientation and relative spatial features that are difficult to identify using traditional CNNs. Simultaneously, CapsNet is competitive with many state-of-the-art techniques on current datasets with less training data and fewer layers/network parameters. In fact, this is

beneficial to the matter that we are addressing. In terms of the size and direction, the force we apply and the stress values we get are vectors. Then, we add them into our model.

C. ENHANCE - AND - SPLIT FEATURE CAPSULE NETWORK EMBEDDED IN FULLY CONVOLUTIONAL NEURAL NETWORKS (ES-CAPS-FCN)

The input data contain several types of information including direction, size and geometry of the solid material. The DL model uses three network model links with the same structure to train the stress values of the three different directions for each point in parallel. The MSE loss function and the Adam optimization function are applied in this model and a deep neural network is also applied in the comparison experiment in section IV. Meanwhile, the all results of three networks are concatenated into a 4096×3 matrix. The top six layers of the DL model except the fourth layer are three-dimensional convolutional neural networks [42] with a kernel size of $2 \times 2 \times 2$ and a stride of $1 \times 1 \times 1$, and they have 9, 4, 8, 32, and 48 filters, respectively. The fourth layer is the pooling layer with maxpooling of $2 \times 2 \times 2$ and a stride of $1 \times 1 \times 1$. The seventh layer of the DL model is CapsNet with 48 capsules. Fig. 1 depicts the whole structure of the ES-Caps-FCN.

We add a feature enhancement operation after the first layer, split the extracted features into three links, join the results with the original input, and then repeat the convolution. The extracted features are divided into three equal parts and concatenated with the input in sequence. After the sixth layer, the same data feature enhancement processes are carried out. After the input and feature are concatenated and flattened, the training continues in the last two layers. The last two layers are fully connected neural networks containing 128 neurons.

IV. EXPERIMENTS AND DISCUSSIONS

A. DATASETS

Academician Kang *et al.* [43], [44] established a database of the underground in situ stress in Chinese coal mines. The distribution characteristics and main influencing factors of the underground in situ stress for coal mines in China were analyzed using this database. Then, some results revealed that the vertical stress and maximum and minimum horizontal principal stresses are generally positively correlated with the buried depth, but the data have a certain degree of dispersion. The in situ stress type is primarily the normal type in coal mines more than a thousand meters deep, and $\sigma_v > \sigma_H > \sigma_h$ repreats the vertical stress and maximum and minimum horizontal principal stresses, respectively. The type of in situ stress is primarily the thrust type stress in shallow coal mines, and $\sigma_H > \sigma_h > \sigma_v$. The rock bulk density is basically $0.025 - 0.033 \text{MPa/m}^3$.

Consequently, according to the buried depth, we randomly set the lateral pressure coefficient at $0.45 - 0.7$ and $0.8 - 1.3$ to better serve engineering application. To simplify the models' analysis, the material parameters remained unchanged for
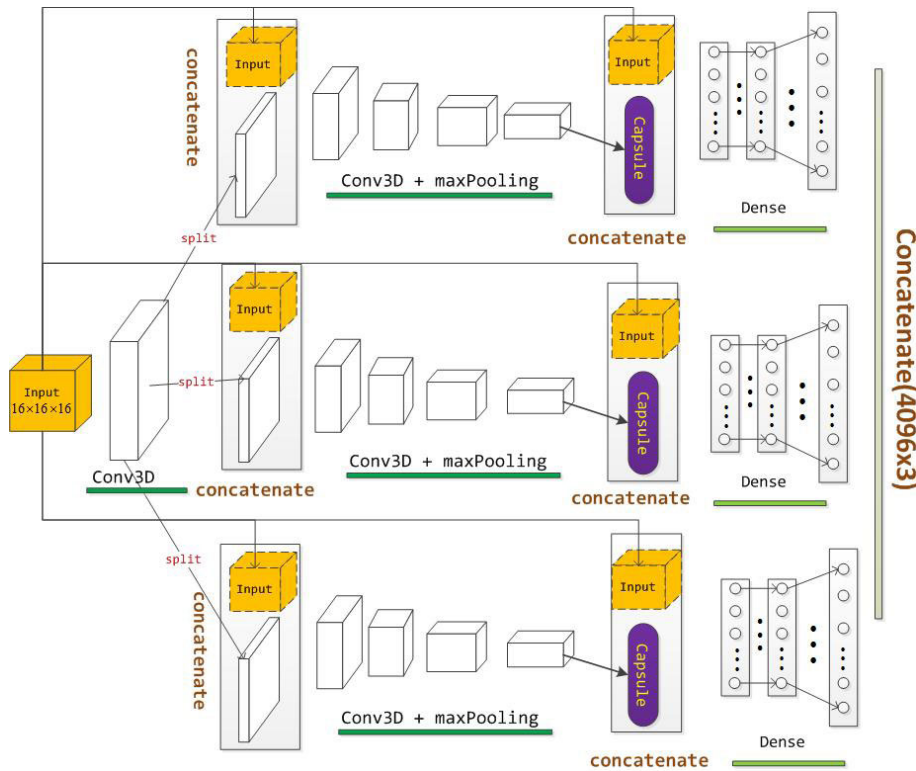
FIGURE 1. The architecture of the ES-Caps-FCN.

all cases, and the FLAC3D software is used to generate the training and testing data. We got 877 cases and a total of 3592192×3 stress values, and some of the examples are randomly selected from the total cases, as shown in Fig. 7. Approximately 80% of the data set for the model was divided into training set and the remaining 20% was used for the testing set. The iteration of the cases calculation was stopped once the error value is less than $10E-6$, which was deemed as an equilibrium state, and then we extracted the stress values of the three directions for each point at this time.

In the case of a small volume of $60 \times 60 \times 60\text{m}^3$, it takes more than 200 iterations on average, and it takes roughly an average of 1 minute to extract and record the stress values in the three directions. Some of the stress chromatograms are show in Fig. 2. While the actual input is a three-dimensional matrix, the output is a flattened one-dimensional matrix for convenience. An intuitive display of the data examples with one input and three output stresses is given in Fig. 2. The first image uses the boundary constraints, force surfaces and force vectors as the input. Figs. 2(b), 2(c) and 2(d) show the stress component values in the X, Y and Z directions of the model, respectively. The last two images show the meshing of the hexahedrons for finite difference calculations. The units of values in the graphs are Pa.

B. COMPARISON ALGORITHM
In response to this problem, to evaluate and analyze the generalization performance of the ES-Caps-FCN of this study, some comparative experiments with machine learning model

are supplied, such as the linear regression and the neural network model designed by us.
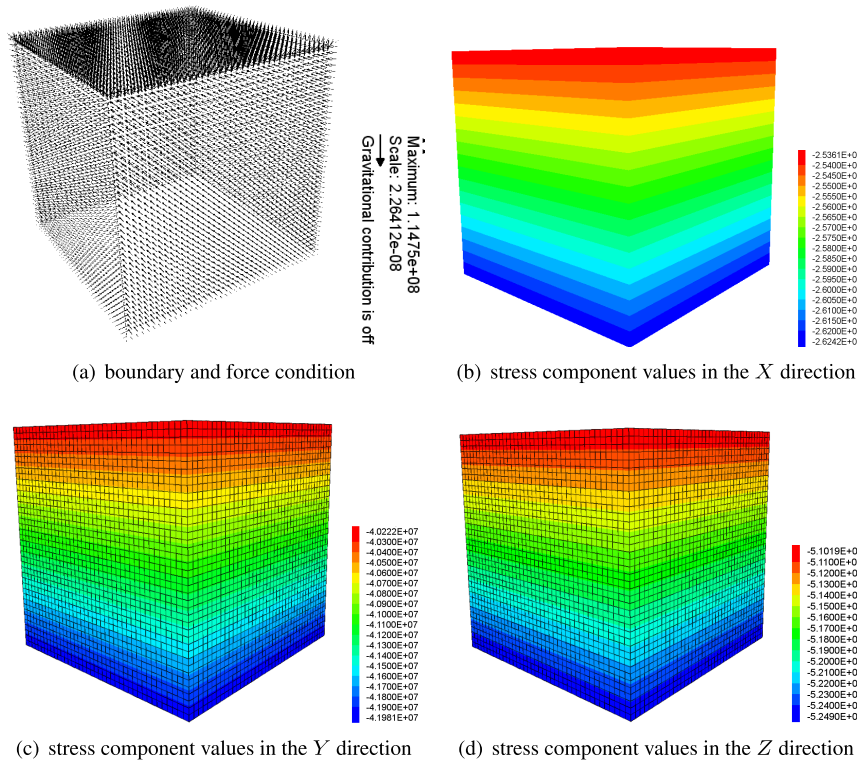
1) LINEAR REGRESSION
The linear regression is one of the most fundamental machine learning algorithms to predict problems. A linear regression model explores the linear relationship between a dependent variable and one or more independent variables. It can take the form of a unary linear regression model or a multiple linear regression model.

Furthermore, the polynomial regression model is also a linear regression model. Polynomials can be used to approximate any function, so their application background is very broad. A polynomial regression [45] can be either a unary polynomial regression or a multivariate polynomial regression according to the number of independent variables. The orders of a polynomial can also be freely selected and determined according to the data distribution, and we can choose appropriate models to solve different problems.

The fundamentals of the FEM regression, FDM regression and DDM regression are same as the stress function method. Based on the principle of the in-stress regression method, the linear superposition of external forces via multiple regression analysis is used to solve the stress components as one of the comparative experiments in this paper.

2) DEEP NEURAL NETWORK(DNN)
The in situ stress prediction based on neural networks methods can explain the nonlinear relation between stress and

(a) boundary and force condition

(b) stress component values in the X direction

(c) stress component values in the Y direction

(d) stress component values in the Z direction

FIGURE 2. Intuitive display of the data examples with one input and three output stresses. The units of the values in the graph are Pa.

boundary conditions, force, and physical properties of rocks via a complex function, and then determine stress values via numerical simulation or direct calculation according to the measuring values and these relations. By employing the DNN model, the authors in [22] predicted the Von Mises stress distribution and peak Von Misesstress in experiments, and they achieved excellent results with average errors of 0.492% and 0.891%, respectively. Meanwhile, neural network modeling performs better than the regression model at predicting Vp [46]. As a consequence, we use this three-layer neural network structure to predict the in situ stress at some specified point, as illustrated in Fig. 3.

Our network is, in fact, a fully connected neural network structure with three parallel links. Its input is the same as the DL structure (ES-Caps-FCN) we designed and will enter the neural network model after flattening. The first layer and the second layer of each neural network model contain 128 neurons and the activation function is the ReLU. The output layer of the model contains 4096 neurons without an activation function. In the end, all outputs are concatenated to a 4096 × 3 matrix, which is similar to the ES-Caps-FCN.

3) EVALUATION METHODS

The mean squared error (MSE) [47] is not only a loss function but also a critical evaluation index for fitting and prediction. Other than this, the mean absolute error (MAE) and R²-Score are also two of the most common metrics used to evaluate the
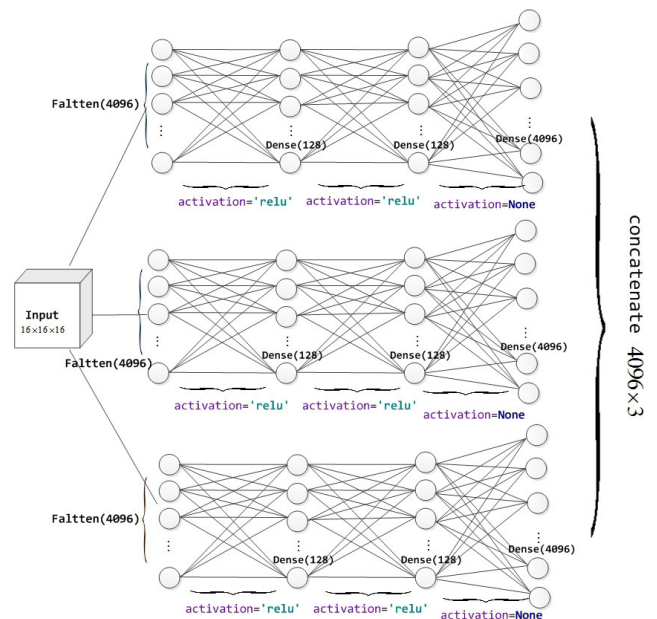
FIGURE 3. Deep neural network used to deal with our problem.

model accuracy for continuous variables [48], [49]:

1. mean-squared-error

$$\text{MSE}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} (y_i - \hat{y}_i)^2 \qquad \text{(IV.1)}$$

2.mean-absolute-error

$$\text{MAE}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}-1} |y_i - \hat{y}_i| \qquad (IV.2)$$

3.R^2-Score

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=0}^{n_{\text{samples}}-1}(y_i - \hat{y})^2}{\sum_{i=0}^{n_{\text{samples}}-1}(y_i - \bar{y})^2} \qquad (IV.3)$$

where n_{samples} is the number of test sample data sets. Our prediction y and ground truth $\hat{y}$ in this paper are both one-dimensional vectors with a length of 4096. For the above three criteria, the closer the first two get to 0 and the third gets to 1, the better the performance of the model selection and fit.

In fact, there are other evaluation criteria, such as the explained variance and root mean squared error (RMSE). Here, we only use the three stated measures, which are representative enough to measure the quality of our models and the prediction results.

C. ANALYSIS RESULTS

Utilizing the above indicators, we evaluate and analyze our model and their forecast results in this section. The three algorithms used in this paper have their own respective characteristics. In this regard, linear regression analysis is the preferred prediction technique and can supply a workable plan for many companies' forecasting analysis. The neural network and deep learning are special methods that can be used as the predominant strategy since they have the potential to analyze more complex nonlinear relationships.

A total of 701 cases are chosen as the training set, and the remaining 176 cases are used as the testing set to validate the models' performance. For the three algorithms, we train and validate them on a personal computer with an NVidia GeForce GTX 1060. In terms of the ES-Caps-FCN and DNN algorithms, we run both on all data sets six times for each algorithm and average their respective evaluation indexes. We display the percentage of each value to six decimal places. All data presented in this article omit excess decimals by rounding.

When the loss function reaches a minimum and does not decrease in the next 500 epochs, the training process of our model is stopped. In other words, we think that the model has converged at this time. Fig. 4 displays the loss function values in one of the training and validating processes of our models. Fig. 4(a) and fig. 4(b) illustrate that all four loss curves rapidly decrease in the first twenty epochs, and then tend to be flat. From the loss curves of the ES-Caps-FCN and the DNN, we know that they all have an acceptable convergence speed. Fig. 5 illustrates the MAEs of the training and validation processes of our models (randomly selected results).

In terms of processing the 176 FDM cases, it takes only 0.750464 seconds for the ES-Caps-FCN, 0.044518 seconds for the DNN and 0.233359 seconds for the linear regression on average to compute the stress fields. The times above are the means of the six experiments, as shown in table 1. The
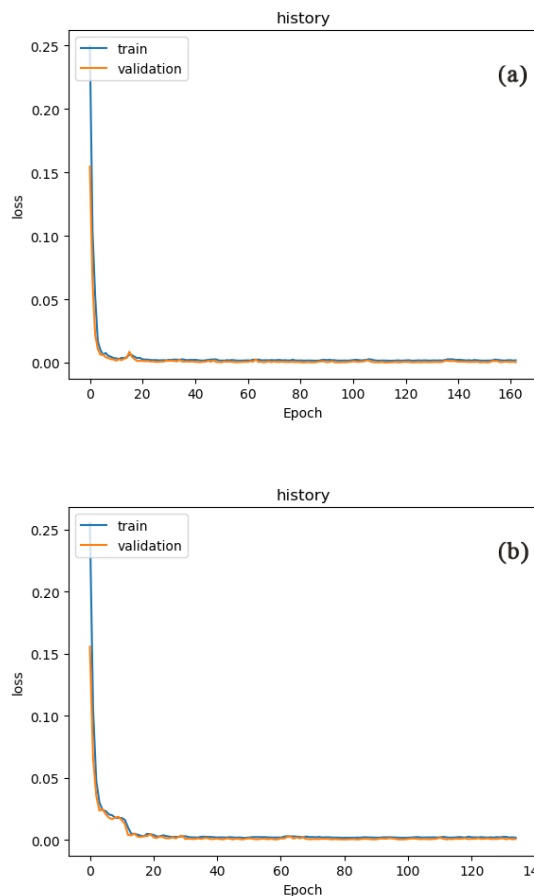

FIGURE 4. MSE curves for the training and validation of the ES-Caps-FCN (a) and the DNN (b).

TABLE 1. Time for the validation of the ES-Caps-FCN, the deep neural network and the linear regression.
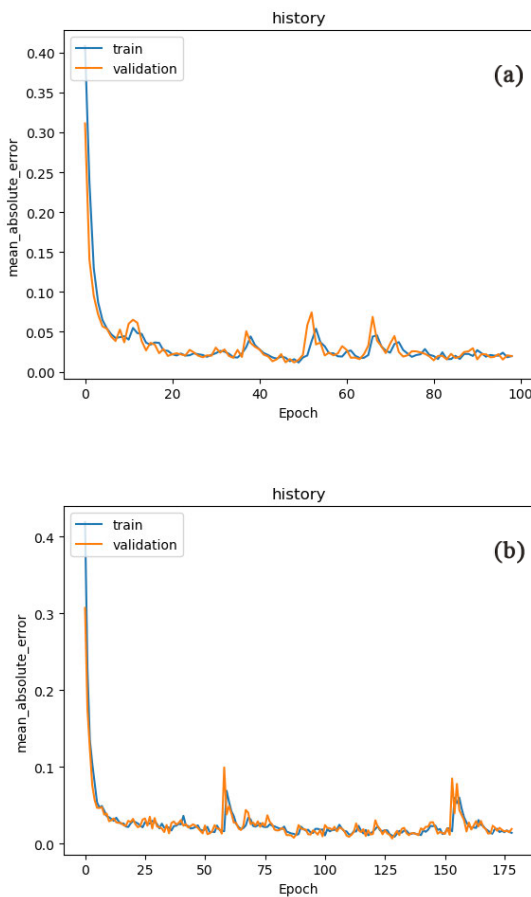
Data	ES-Caps-FCN	deep neural network	linear regression
time	0.750464	0.044518	0.233359

FLAC3D software, by contrast, spends almost three hours to accomplish the same amount of FDM computations and record the stress values of points. This shows that fully trained deep learning models have higher computational efficiency over conventional FDM models for calculating the in stress field. Therefore, the three algorithms are fast enough to calculate the in situ stress field of the real stratum, and it is valuable and significant to investigate the application of deep learning in geotechnical engineering and coal mining.

After the aforementioned models converge, we derive the MSE, MAE and R^2-Score values, and then evaluate the models and prediction results. The MSE and MAE are error rates that measure how close the predictions are to the truth. When the MSE or MAE is zero, it means that the prediction is absolutely the same as the truth.
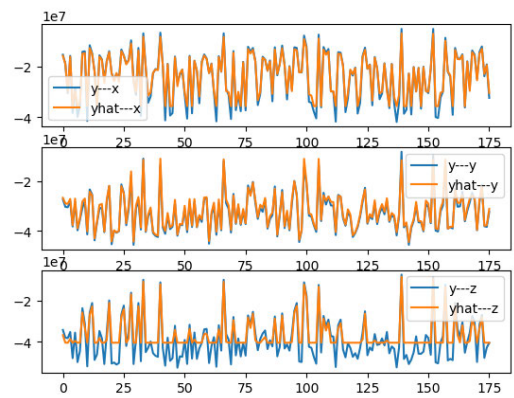
TABLE 2. Error indicator results. E means the ES-Caps-FCN, L means the Linear regression and D means the deep neural network.

Data	E-MSE	D-MSE	L-MSE	E-MAE	D-MAE	L-MAE	E-R^2	D-R^2	L-R^2
Testing set	0.059866%	0.616676%	0.978495%	1.806252%	4.898292%	8.432438%	99.813368%	97.650430%	95.784082%
Standard Deviation	0.038690%	0.848321%	—	0.583082%	3.634844%	—	0.113572%	3.274158%	—

FIGURE 5. MAE curves for the training and validation of the ES-Caps-FCN (a) and the DNN (b).

FIGURE 6. One of the verification results of the DNN. From the top to bottom are the following: the *X* direction, the *Y* direction, and the *Z* direction. The *x*-axis denotes the test sample, and the *y*-axis denotes the stress components.

The experimental results are shown in table 2, which demonstrates that the MSE of the ES-Caps-FCN of 0.059866% is the lowest, and it is followed by the DNN and linear regression. The MAE of the ES-Caps-FCN of 1.806252% is the lowest, followed by the DNN and linear regression. We can draw from table 2 that the corresponding R^2-Score of the ES-Caps-FCN, DNN and linear regression are 99.813368%, 97.650430% and 95.784082%, respectively. In summary, the generalization ability and prediction accuracy of the ES-Caps-FCN are the best one among the three algorithms in experiments.

The standard deviation can reflect the degree of dispersion between individuals within their group. From the standard deviation of the three indicators, as shown in table 2, we know that the ES-Caps-FCN is generally stable for these six results, and the DNN appears to be slightly unstable. Fig. 6 shows that the DNN has great fluctuation for stresses prediction in the *Z* direction in one experiment.

Zhenguo Nie and their coworkers [50] used a deep learning method to simulate the finite element method to analyze relationships between the stress values and the loads, boundary conditions, and geometric figures in a two-dimensional plane. In this similar two-dimensional mechanics problem, this article defines a criterion called the mean relative error (MRE) by calculating the ratio of the MAE to the mean of the truth. The specific calculation formula is as follows:
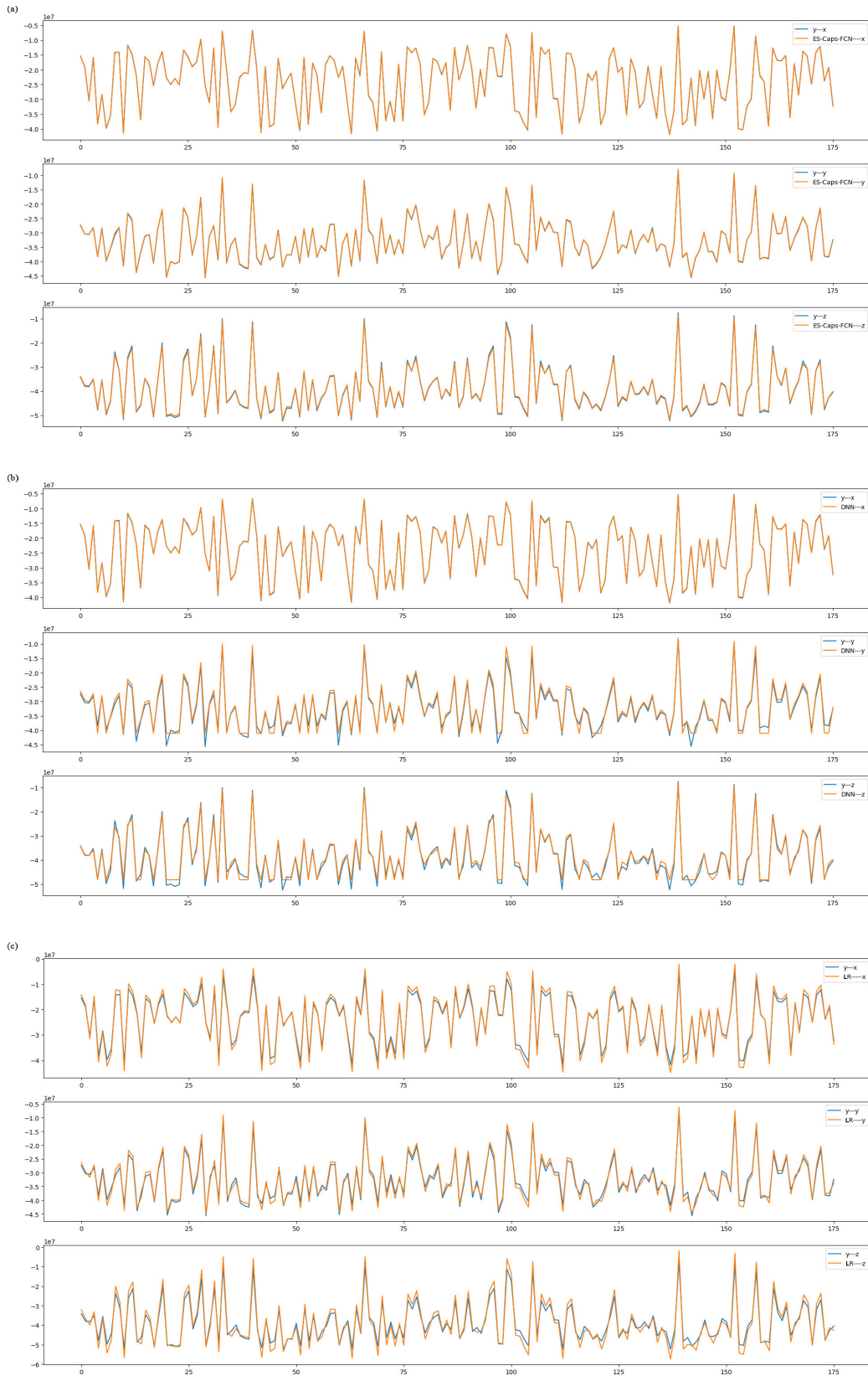
$$\text{MRE} = \frac{\text{MAE}}{\bar{y}} \times 100\% \qquad \text{(IV.4)}$$

where the MAE is calculated in the same way as in IV-B3 and $\bar{y}$ is as follows:

$$\bar{y} = \frac{1}{n} \sum_{i=0}^{n} y_i \qquad \text{(IV.5)}$$

In this paper, we also calculate the MREs to evaluate the ES-Caps-FCN, DNN and linear regression. Table 3 lists the results of the MREs for the aforementioned three algorithms.

FIGURE 7. A part of the verification results of the three models are randomly selected for display: the ground truth and predictions in the *X* direction, *Y* direction, *Z* direction. From top to bottom are the following: the ES-Caps-FCN (a), DNN (b) and linear regression (c). The *x*-axis denotes the test sample, and the *y*-axis denotes the stress components.

TABLE 3. The MREs of the ES-Caps-FCN, deep neural network and linear regression. E means the ES-Caps-FCN, L means the linear regression and D means the deep neural network.

Data	E-MRE	D-MRE	L-MRE
Testing set	0.011478%	0.032781%	0.363551%
Standard Deviation	0.354337%	2.450832%	—

We show our data using percentages up to six decimal places. Table 3 shows that the MREs of ES-Caps-FCN, DNN and linear regression are 0.011478%, 0.032781% and 0.363551%, respectively, and the corresponding standard deviations of the ES-Caps-FCN and DNN are 0.354337% and 2.450832%, respectively. The MREs and thier standard deviations of the ES-Caps-FCN are still lower than those of the DNN. Therefore, the architecture designed by this paper performs the best both in precision and stability. In fact, the accuracies of these three methods meet the requirements of engineering applications.

With regard to the prediction results for the different models, we display the true and predicted values of the random testing samples that are extracted by the models. The fitting curves are shown in Fig. 7. It is particularly necessary to point out that the blue curves stand for the true stress values, and the orange curves are the values predicted by the ES-Caps-FCN, DNN, and linear regression, respectively. It can be seen from the Fig. 7 (b) and (c) that the predicted trends of the DNN and linear regression are the same as the true values, and the curves of the two colors in the Fig. 7 (a) almost completely overlap. So the predicted performances of the ES-Caps-FCN are significantly better than those of the DNN and linear regression for the X direction, Y direction, and Z direction.

V. CONCLUSION

We propose a DL framework to calculate the geostress in a way of simulating the processing of the fast Lagrangian analysis method based on the FDM, which can capture more information about a three-dimensional geological body to train the model for stress analysis. The experiments prove that its prediction accuracy is better than those of the traditional neural network and linear regression analysis, and it has better expansibility. Compared with linear regression, both neural network and deep learning require additional training time, but the training convergence speed of the DL framework and DNN is efficient and the running time required to predict the stress component values is distinctly shortened. This indicates that the DL framework can be considered as an alternative approach to efficiently calculate the in situ stress component of the fast Lagrangian analysis method.

Although the features used in this paper are incomprehensive, we believe that when there are enough features information involved in the training, a deep learning model designed and trained for a specific problem will be an excellent substitute for the fast Lagrangian analysis method, thereby avoiding complex calculations when predicting the in situ stress components via numerical simulation calculations. This method, therefore, has further research potential. In the future, more efforts should be made to predict the distribution of the in situ stress using more or even all of the input characteristics and material behavior, or to evaluate the network model to determine which parameters are the most important during the procedure. This can possibly allow one to predict the stress values in all directions.

REFERENCES

[1] M. D. Zoback, D. Moos, L. Mastin, and R. N. Anderson, "Well bore breakouts and *in situ* stress," *J. Geophys. Res. Solid Earth*, vol. 90, no. B7, pp. 5523–5530, 1985.

[2] B. Haimson and C. Fairhurst, "*In-situ* stress determination at great depth by means of hydraulic fracturing," in *Proc. 11th US Symp. Rock Mech. (USRMS)*. Golden, CO, USA: American Rock Mechanics Association, 1969, pp. 559–584.

[3] B. C. Haimson, "Deep *in-situ* stress measurements by hydrofracturing," *Tectonophysics*, vol. 29, no. 1, pp. 41–47, 1975.

[4] T. Belytschko and T. Black, "Elastic crack growth in finite elements with minimal remeshing," *Int. J. Numer. Methods Eng.*, vol. 45, no. 5, pp. 601–620, 2015.

[5] T. Belytschko, N. Moës, S. Usui, and C. Parimi, "Arbitrary discontinuities in finite elements," *Int. J. Numer. Methods Eng.*, vol. 50, no. 4, pp. 993–1013, 2001.

[6] K. Yamamoto, Y. Kuwahara, N. Kato, and T. Hirasawa, "Deformation rate analysis: A new method for *in situ* stress estimation from inelastic deformation of rock samples under uni-axial compressions," *Tohoku Geophys. J.*, vol. 33, no. 2, pp. 127–147, 1990.

[7] M. Seto, D. K. Nag, and V. S. Vutukuri, "*In-situ* rock stress measurement from rock cores using the acoustic emission method and deformation rate analysis," *Geotech. Geolog. Eng.*, vol. 17, nos. 3–4, pp. 241–266, 1999.

[8] N. Ren and J.-C. Roegiers, "Differential strain curve analysis-a new method for determining the pre-existing *in-situ* stress state from rock core measurements," in *Proc. 5th Int. Soc. Rock Mech. Rock Eng. Congr. (ISRM)*, 1983, pp. 10–15.

[9] S. D. McKinnon, "Analysis of stress measurements using a numerical model methodology," *Int. J. Rock Mech. Mining Sci.*, vol. 38, no. 5, pp. 699–709, Jul. 2001.

[10] J. T. Katsikadelis, *The Boundary Element Method for Engineers and Scientists: Theory and Applications*. New York, NY, USA: Academic, 2016.

[11] X. Qiu, "DEM simulations in mining and mineral processing," in *Proc. 7th Int. Conf. Discrete Element Methods*. Singapore: Springer, 2016, pp. 37–43.

[12] B. Vazhbakht and A. M. Zsaki, "A finite element mesh optimization method incorporating geologic features for stress analysis of underground excavations," *Int. J. Rock Mech. Mining Sci.*, vol. 59, pp. 111–119, Apr. 2013.

[13] W. Zhou, W. Yuan, and X. Chang, "Combined finite-discrete element method modeling of rock failure problems," in *Proc. 7th Int. Conf. Discrete Element Methods*. Singapore: Springer, 2016, pp. 301–309.

[14] X. L. Yue, Y. Li, and H. P. Wang, "Inversion of initial *in situ* stress field by the method of multivariate analysis and engineering application," *Appl. Mech. Mater.*, vols. 44–47, pp. 1203–1206, 2011.

[15] Y. Liang, D. Zhou, T. Yang, and G. Zhao, "Application of *in situ* stress regression method in rock slope stability analysis," in *Geotechnical Engineering for Disaster Mitigation and Rehabilitation*. Berlin, Germany: Springer, 2008, pp. 540–545.

[16] B. Ratner, *Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data*. London, U.K.: Chapman & Hall, 2017.

[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[18] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[19] M. Raissi, P. Perdikaris, and G. Em Karniadakis, "Multistep neural networks for data-driven discovery of nonlinear dynamical systems," 2018, *arXiv:1801.01236*. [Online]. Available: http://arxiv.org/abs/1801.01236

[20] F. Roewer-Despres, N. Khan, and I. Stavness, "Towards finite element simulation using deep learning," in *Proc. 15th Int. Symp. Comput. Methods Biomechan. Biomed. Eng.*, 2018, pp. 1–20.

[21] J. Ghaboussi, "Neural networks in computational mechanics," *Arch. Comput. Methods Eng.*, vol. 3, no. 4, p. 435, 1996.

[22] L. Liang, M. Liu, C. Martin, and W. Sun, "A deep learning approach to estimate stress distribution: A fast and accurate surrogate of finite-element analysis," *J. Roy. Soc. Interface*, vol. 15, no. 138, Jan. 2018, Art. no. 20170844.

[23] L. Liang, M. Liu, C. Martin, and W. Sun, "A machine learning approach as a surrogate of finite element analysis-based inverse method to estimate the zero-pressure geometry of human thoracic aorta," *Int. J. Numer. Methods Biomed. Eng.*, vol. 34, no. 8, p. e3103, May 2018.

[24] O. Kononenko and I. Kononenko, "Machine learning and finite element method for physical systems modeling," 2018, *arXiv:1801.07337*. [Online]. Available: http://arxiv.org/abs/1801.07337

[25] A. Javadi, T. Tan, and M. Zhang, "Neural network for constitutive modelling in finite element analysis," *Comput. Assist. Mech. Eng. Sci.*, vol. 10, no. 4, pp. 523–530, 2003.

[26] J. Zhou, X. Li, and X. Shi, "Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines," *Saf. Sci.*, vol. 50, no. 4, pp. 629–644, Apr. 2012.

[27] J. Zhou, X. Li, and H. S. Mitri, "Classification of rockburst in underground projects: Comparison of ten supervised learning methods," *J. Comput. Civil Eng.*, vol. 30, no. 5, p. 04016003, 2016.

[28] J. W. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*, vol. 22. Berlin, Germany: Springer, 2013.

[29] *FLAC (Fast Lagrangian Analysis of Continua) Version 2. 20*, Software Summary, Minneapolis, MN, USA, 1989.

[30] K. Li, Y.-Y. Wang, and X.-C. Huang, "DDM regression analysis of the *in-situ* stress field in a non-linear fault zone," *Int. J. Minerals, Metall., Mater.*, vol. 19, no. 7, pp. 567–573, Jul. 2012.

[31] S. Wei-feng, T. Cheng-xuan, W. Zhi-ming, Z. Chun-shan, and W. Shu-ren, "Prediction of crustal stress of deep-buried tunnels based on bp artificial neural network," *J. Geomechan.*, vol. 13, no. 3, pp. 227–232, 2007.

[32] J. Zhong-ming, X. Wei-ya, and S. Jian-fu, "Ann-based 3-d back analysis of initial stress in rock masses," *J. Hohai Univ. (Natural Sci.)*, vol. 30, no. 3, pp. 52–56, 2002.

[33] L. Yong-song, Y. Jian-min, C. Jian-ping, and X. Jing, "Analysis of 3D *in-situ* stress field and query system's development based on visual BP neural network," *Procedia Earth Planet. Sci.*, vol. 5, pp. 64–69, Jan. 2012.

[34] X. Ri-qing, "The GA-ANN method for determining calculation parameters for deep excavation," *J. Zhejiang Univ.-Sci. A*, vol. 1, no. 4, pp. 408–413, Oct. 2000.

[35] H. Zhang and S. Yin, "Inference of *in situ* stress from thermoporoelastic borehole breakouts based on artificial neural network," *Int. J. Numer. Anal. Methods Geomechan.*, vol. 43, no. 16, pp. 2493–2511, Aug. 2019.

[36] L. Jiang, P. Kong, J. Shu, and K. Fan, "Numerical analysis of support designs based on a case study of a longwall entry," *Rock Mech. Rock Eng.*, vol. 52, no. 9, pp. 3373–3384, Jan. 2019.

[37] P. Kong, L. Jiang, J. Shu, A. Sainoki, and Q. Wang, "Effect of fracture heterogeneity on rock mass stability in a highly heterogeneous underground roadway," *Rock Mech. Rock Eng.*, vol. 52, no. 11, pp. 4547–4564, Jun. 2019.

[38] J. Bouvrie, "Notes on convolutional neural networks," Eprints, Tech. Rep., 2006. [Online]. Available: http://cogprints.org/5869/

[39] Y. Li, H. Zhang, and Q. Shen, "Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, p. 67, Jan. 2017.

[40] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," 2013, *arXiv:1301.3557*. [Online]. Available: http://arxiv.org/abs/1301.3557

[41] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3856–3866.

[42] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[43] H. P. Kang, B. D. Yi, F. Q. Gao, and H. W. Lv, "Database and characteristics of underground *in-situ* stress distribution in Chinese coal mines," *J. China Coal Soc.*, vol. 1, no. 1, pp. 23–33, 2019.

[44] H. Kang, X. Zhang, L. Si, Y. Wu, and F. Gao, "*in-situ* stress measurements and stress distribution characteristics in underground coal mines in china," *Eng. Geol.*, vol. 116, nos. 3–4, pp. 333–345, Nov. 2010.

[45] B. Chen. *Polynomial Regression*. New York, NY, USA: Springer, 1986, pp. 235–268.

[46] S. Paul, M. Ali, and R. Chatterjee, "Prediction of compressional wave velocity using regression and neural network modeling and estimation of stress orientation in bokaro coalfield, india," *Pure Appl. Geophys.*, vol. 175, no. 1, pp. 375–388, 2018.

[47] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geosci. Model Develop.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.

[48] G. Sprint, D. J. Cook, D. L. Weeks, and V. Borisov, "Predicting functional independence measure scores during rehabilitation with wearable inertial sensors," *IEEE Access*, vol. 3, pp. 1350–1366, 2015.

[49] Q. Zhang, X. Zeng, W. Hu, and D. Zhou, "A machine learning-empowered system for long-term motion-tolerant wearable monitoring of blood pressure and heart rate with ear-ECG/PPG," *IEEE Access*, vol. 5, pp. 10547–10561, 2017.

[50] Z. Nie, H. Jiang, and L. Burak Kara, "Stress field prediction in cantilevered structures using convolutional neural networks," 2018, *arXiv:1808.08914*. [Online]. Available: http://arxiv.org/abs/1808.08914
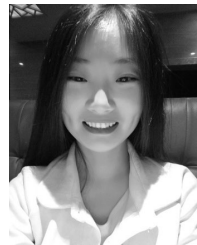
WENLI GAO was born in Jining, Shandong, China, in 1991. She received the M.S. degree in computer technology from Central China Normal University, in 2017. She is currently pursuing the Ph.D. degree with the Shandong University of Science and Technology. Her research interests are 3D modeling of geological bodies, artificial intelligence, virtual reality, and so on.

XINMING LU was born in 1961. He received the Ph.D. degree from the Institute of Applied Mathematics, Chinese Academy of Sciences, in 1994. He joined the work in July 1982. He is currently with the Key Laboratory of Intelligent Mine Information Technology of Shandong Province and the Director of Shandong Smart Mine Engineering Technology Research Center. He is also a Professor with the Shandong University of Science and Technology. In view of the current situation of mine informationization in China, the team led the team to carry out research work on intelligent mine information technology. His researches interests include intelligent mine information technology, computer-aided design, geographic information systems, computer graphics, and so on. He has completed and undertaken ten national-level projects such as the national 863 key and theme projects, the National Innovation Fund, the National Electronics and National Industrialization Promotion Project, and the National Natural Science Foundation. In 2012, 2011, and 2010, he received the first prize, second prize, and third prize of Shandong Science and Technology Progress Award, respectively. In 2011 and 2007, he received the second prize and the third prize of the Science and Technology of the China Coal Industry Association, respectively. In 2013, he received the second prize of National Science and Technology Progress Award. In 2014, he was elected as a Distinguished Professor of the Taishan Scholar Climbing Program.
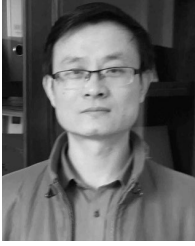
YANJUN PENG received the Ph.D. degree in computer application specialty from Zhejiang University, in 2004. He joined the Department of Computer Science, Shandong University of Science and Technology, Qingdao, China, in 1996, where he was promoted to Professor, in 2010. He is currently the Director of the National Coal Mining Virtual Simulation Experiment Center. He was rewarded young and middle-aged expert with outstanding contributions for Shandong Province, in 2014. His main research interests are visualization in scientific computing, image processing, and virtual reality. He has completed and undertaken five national-level projects such as the national 863 key and theme projects, national science and technology support, and the National Natural Science Foundation. He has completed three provincial and ministerial projects, such as the Natural Science Foundation of Shandong Province and the Science and Technology Plan of Shandong Province.

LIANG WU received the B.S. degree in biomedical engineering from Shandong First Medical University (Taishan Medical University), Taian, China, in 2015, and the M.S. degree in computer science and technology from the Inner Mongolia University of Science and Technology, Baotou, China, in 2018. He is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Shandong University, Jinan, China. His research interests include pattern recognizing and image processing.

• • •