scientific reports

OPEN

Comparative study on convolutional neural network and regression analysis to evaluate uniaxial compressive strength of Sandy Dolomite

Meiqian Wang[1,2], Wenlian Liu[3,4], Haiming Liu[1], Ting Xie[5], Qinghua Wang[6] & Wei Xu[1,2]✉

Sandy Dolomite is a kind of widely distributed rock. The uniaxial compressive strength (UCS) of Sandy Dolomite is an important metric in the application in civil engineering, geotechnical engineering, and underground engineering. Direct measurement of UCS is costly, time-consuming, and even infeasible in some cases. To address this problem, we establish an indirect measuring method based on the convolutional neural network (CNN) and regression analysis (RA). The new method is straightforward and effective for UCS prediction, and has significant practical implications. To evaluate the performance of the new method, 158 dolomite samples of different sandification grades are collected for testing their UCS along and near the Yuxi section of the Central Yunnan Water Diversion (CYWD) Project in Yunnan Province, Southwest of China. Two regression equations with high correlation coefficients are established according to the RA results, to predict the UCS of Sandy Dolomites. Moreover, the minimum thickness of Sandy Dolomite was determined by the Schmidt hammer rebound test. Results show that CNN outperforms RA in terms of prediction the precision of Sandy Dolomite UCS. In addition, CNN can effectively deal with uncertainty in test results, making it one of the most effective tools for predicting the UCS of Sandy Dolomite.

Abbreviations
CYWD project

UCS	Uniaxial compressive strength
RA	Regression analysis
PMP	Petrophysical and mechanical properties
SH	Schmidt hammer rebound index
V_p	Compressional wave velocity
n	Porosity
MSE	Mean square error
MAE	Mean absolute error
RQD	Rock quality designation
$f_{UCS,i}$	Predicted values of UCS

[1]Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming 650500, Yunnan, China. [2]Intelligent Infrastructure Operation and Maintenance Technology Innovation Team of Yunnan Provincial Department of Education, Kunming University of Science and Technology, Kunming 650500, Yunnan, China. [3]Kunming Prospecting Design Institute of China Nonferrous Metals Industry Co., Ltd., Kunming 650051, Yunnan, China. [4]Yunnan Key Laboratory of Geotechnical Engineering and Geohazards, Kunming 650051, Yunnan, China. [5]Faculty of Foreign Languages and Cultures, Kunming University of Science and Technology, Kunming 650500, Yunnan, China. [6]Zhongsheng Civil Engineering (Yunnan) Co., Ltd., Kunming 650000, Yunnan, China. ✉email: xuwei@kust.edu.cn

m	Number of the measured values of UCS
ASTM	American society for testing and materials

Central Yunnan water diversion project

PLT	Point load test
CNN	Convolutional neural network
ANN	Artificial neural network
$I_{s(50)}$	Point load index
ρ	Density
R^2	Correlation coefficient
RMSE	Root mean square error
MAPE	Mean absolute percentage error
f_i	Measured values of UCS
$\bar{f}$	Average of the measured UCS values
ReLU layer	Rectified linear units layer
ISRM	International society for rock mechanics

The Central Yunnan Water Diversion (CYWD) Project is one of the strategic supporting projects for Yunnan Province to build a radiating center facing South and Southeast Asia, the most significant project out of the 172 major water conservancy and supply projects approved by the Ministry of Water Resources of China, and a key water source project of Yunnan's revitalization strategy. The problem of dolomite sandification in the Yuxi section is one of the major challenges faced by the CYWD Project. Dolomite sandification is a unique geological phenomenon in which dolomite, under the influence of a complex geological environment of multistage tectonic movement, is weathered into silty fine sand, gravel, and/or fragment under the combined action of karstification and weathering, which results in a notable reduction in the rock mass strength and quality[1]. The engineering geological problems caused by Sandy Dolomite are primarily found in the underground engineering projects of transportation and water conservancy industries in Southwest of China[1–4], as well as the projects in the United States[5], Canada[6], Iran[7], Germany[8], Egypt[9] and Italy[10], etc.

Rock mass strength is one of the most important mechanical properties of rock[11], which is affected by particle arrangement, discontinuity, saturation, temperature, humidity, and/or weathering[12]. Determining the rock mass strength and deformation is a key field of rock mechanics[13]. To a certain extent, rock mass stability has a substantial effect on human life and property security, and infrastructure construction[14]. It is crucial to evaluate the rock mass strength accurately, efficiently, and reliably, especially in the construction of tunnels and long-distance water diversion projects[15]. As a result, the study of the mechanical characteristics of rock is essential in the field of engineering[16]. The most common method for measuring rock mass strength is the indoor test on the uniaxial compressive strength UCS[17]. The UCS is one of the key parameters of rock mechanics[18] and one of the most important geomechanical parameters for preliminary and final designs of civil engineering, geotechnical engineering, mining engineering, and underground engineering, i.e., dams, rock excavation, tunnels, slope stability, and infrastructure[19,20], and is used for the evaluation of rock hardness[21] and grade classification. Since inaccurate UCS results would cause project budget overruns or even the collapse of related structures[22], precise UCS calculation of the rock mass is of great significance.

Generally, UCS can be measured by direct or indirect test method. The direct test method refers to the testing of standard samples in the laboratory. The indirect test method serves as the prediction for the UCS based on empirical equation[23]. According to the recommendations of the International Society for Rock Mechanics (ISRM) and the American Society for Testing Materials (ASTM), direct test of UCS necessitates high-quality rock core samples and highly skilled operators, which is costly and time-consuming, making the direct test of UCS impossible in some cases[24,25]. As a result, indirect tests are commonly utilized to predict UCS, such as the Schmidt hammer rebound test[26], point load test[26,27], wave velocity test[28], and needle penetration test[29], which are easier to perform due to less or no sample preparation, user-friendly equipment, and strong operability in the field. Therefore, the indirect test is a simpler, faster, and more cost-effective measuring method for UCS compared to direct test[30].

Regression analysis (RA) techniques are frequently used for establishing empirical equations. The researchers have developed many empirical equations for predicting the UCS of rocks. Soft computing-based UCS prediction primarily, i.e., Bayesian analysis[20], ANN[31,32], fuzzy system[33,34], regression trees[35,36], genetic algorithm[23,37], imperialist competitive algorithm[38], and particle swarm optimization technique[39]. So, Soft computing techniques predict UCS more accurately than traditional statistical methods. To overcome the restriction of the prediction precision of RA, this study employed CNN to predict the UCS of Sandy Dolomite. The CNN originated in the 1980s[40], and then has been widely applied in, i.e., civil and mining engineering and detection[41–46], rock properties (physico-mechanical properties, chemical compositions, permeability, porosity, rock mass strength, macro and micro image recognition)[47–60]. The CNN-based analysis method for the prediction of the UCS of Sandy Dolomite has not been reported yet.

In this study, based on the testing values of UCS, *SH*, $I_{s(50)}$, V_p, ρ, and *n*, this study aims to predict the UCS of Sandy Dolomite by RA and CNN, respectively, and compare the corresponding results.

Study area and material

The Yuxi Section of the CYWD Project begins from Muyang Village, south of Xinzhuang, Jinning District of Kunming City. The CYWD Project passing through Jinning District and Yuxi City (Hongta District, Jiangchuan District, and Tonghai County). It is part of the plateau mountain area of the central Yunnan basin, with carbonate rocks widely distributed on the surface of the tunnel area. Strong Karst strata are mainly limestone of the Permian Qixia Formation, Maokou Formation ($P_1q + m$), and Dalongkou Formation (Pt_1d) of the Presinian Kunyang Group. Moderate Karst strata are mainly dolomite and limestone of the upper and middle Carboniferous system (C_{2+3}), middle Devonian Zaige formation (D3zg), Sinian Dengying formation (Z_bd_n), Sinian Doushantuo formation (Z_bd) dolomite, etc. In this study, the Z_bdn and Z_bd Sandy Dolomite along and across the tunnel in the Yuxi section (Fig. 1) are used as the research objects. A total of 158 rock samples were collected from different sites for both field and laboratory tests. RA and CNN are employed in this study to investigate the transformation relation between UCS and SH, $I_{s(50)}$, V_P, ρ, and n based on the PMP of dolomite with different sandification grades, followed by an analysis of the prediction precision and reliability of the two methods.

Methods

Porosity and density properties

According to the level of sanding with the differences in color, rock mass structure, microscopic features, alterations, rock mass integrity indexes, rock quality designation (RQD), wave velocity (V_P), the intactness index of rock mass (k_v), dolomites can be divided into four types of sanding: fierce, strongly, medium and weakly (micronew rock mass), as listed in Table 1 [1,61].
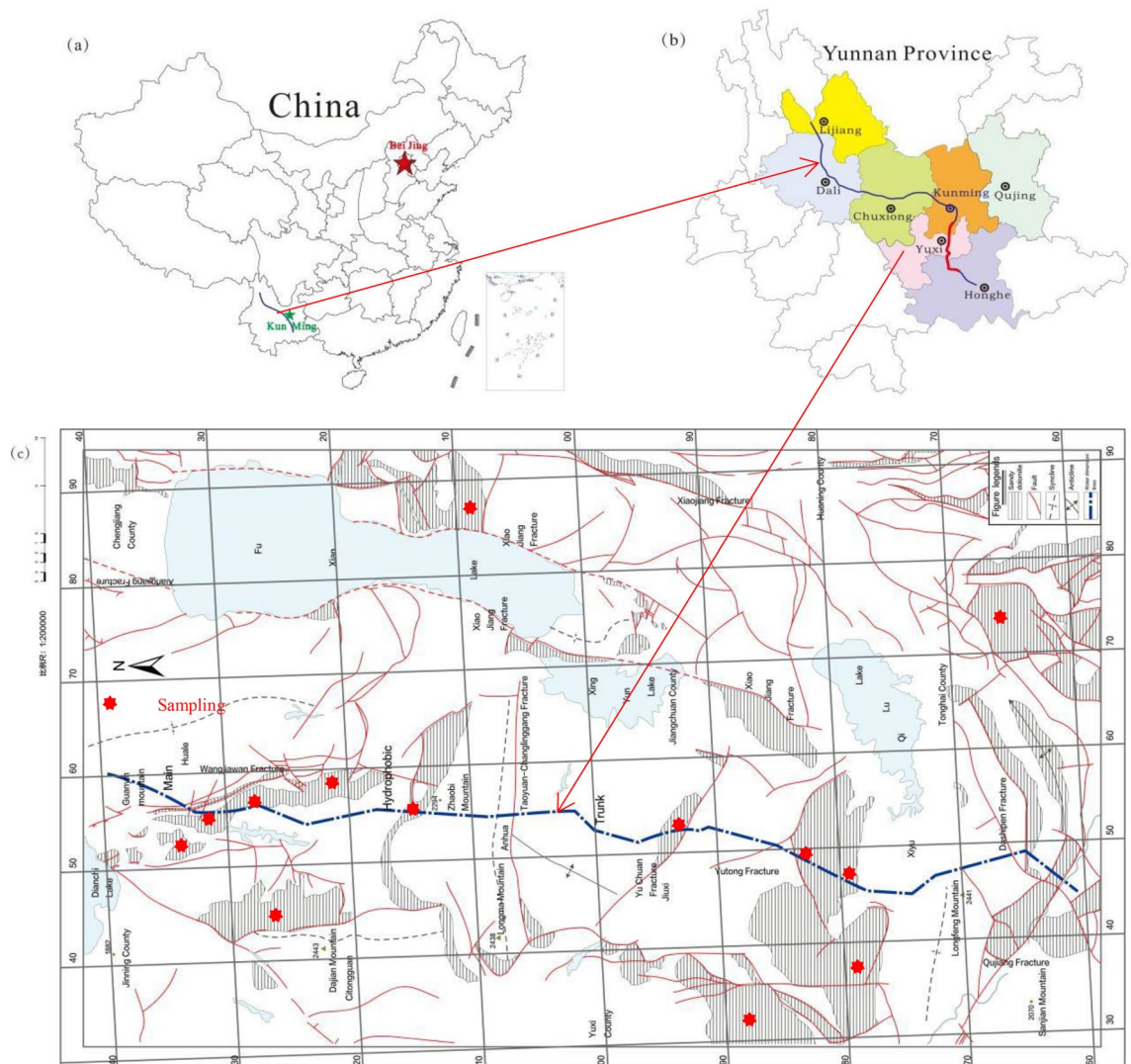

Figure 1. Geological setting of the sampling area and positions of Sandy Dolomites, the construction materials in this study. (**a**) the location of the tunnel of the CYWD Project in China, (**b**) the location of the tunnel of the CYWD Project in Yunnan, (**c**) the location of the tunnel of the CYWD Project in Yuxi section[61].

Sandy degree	Color	Rock organization structure	Volume change	Microscopic feature	Alteration	Rock main characteristics value
Fierce sandy	Uniform discoloration and gloss loss	Totally destroyed and disintegrates and decomposes into loose sand particles	Large	Powder crystal structure	Except for quartz, most of the residual minerals alter to secondary minerals	$K_v < 0.10$ $V_P < 1.0$ km/s
Strongly sandy	Primary discoloration with part of rock blocks retaining their original color	Mostly destroyed and small part disintegrates and decomposes into loose sand particles	Not small	Fine crystal—medium crystal structure	Except for quartz, feldspar, mica, and femic minerals are already weathered	$RQD < 20\%$ $K_v = 0.10 \sim 0.15$ $V_P = 1.0 \sim 2.0$ km/s
Medium sandy	Mostly discoloration and only the fracture of the rock retain slight discoloration as bright color	Mostly appear clear and complete, and the body of rock exhibits fragmentation; small part appears embedded and exhibits fragmental structure	No	Aplitic texture	Femic minerals exhibit oxidation and corrosion; feldspar exhibits opacification	$RQD = 20\% \sim 40\%$ $K_v = 0.15 \sim 0.35$ $V_P = 2.0 \sim 3.5$ km/s
Weakly sandy	Uniform slight discoloration as bright color	All appear original withcomplete organization structure	No	Medium crystal structure	Only the part along with the crack appear the phenomenon of weathering and alteration or/ and the immersion of the corrosion	$RQD > 50\%$ $K_v > 0.40$ $V_P > 4.0$ km/s

Table 1. Sandy degree of dolomite classification[1,61].

In this study, the Sandy Dolomite along and near the Yuxi section of the CYWD Project was taken as the research object. Dolomite with various sandification grades was sampled in field outcrops, boreholes, and tunnels to determine the PMP. The test results are displayed in Table 2, Figs. 2, and 3.

According to Matula's method[62], fierce Sandy Dolomite is a rock with low density and high porosity; strongly Sandy Dolomite is a rock with medium density and porosity; medium and weakly Sandy Dolomite is a rock with high density and low porosity.

According to Figs. 2 and 3, the higher the sandification grade increases, the higher porosity and the lower density of Sandy Dolomite, indicating that the sandification grade increases with the rise of porosity and the decline of density.

Since the fierce Sandy Dolomite can be easily crushed by hand, a complete rock block cannot be obtained for mechanical testing, as shown in Fig. 4a. The thin rock section identification is shown in Fig. 4b. The indoor size distribution test result shows that the fierce Sandy Dolomite is in a state of silty fine sand.

Since the rock mass strength of strongly Sandy Dolomite is obviously weakened, with forming weathering fissures, a complete rock block cannot be obtained for mechanical testing, as shown in Fig. 5a. The thin rock

Sandy Dolomite	Number of tests	ρ (g/m³)			n (%)		
		Min	Max	Mean	Min	Max	Mean
Fierce	19	1.7	1.9	1.82	9.09	23.08	16.63
Strongly	24	2.442	2.608	2.549	5.53	11.62	8.59
Medium	90	2.62	2.761	2.699	0.77	4.45	2.577
Weakly	68	2.601	2.797	2.723	0.34	3.26	1.426

Table 2. Statistics of ρ and n of Sandy Dolomite.
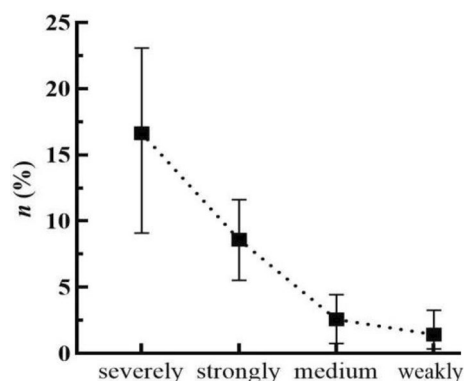

Figure 2. Relationship between Sandy Dolomite and n.
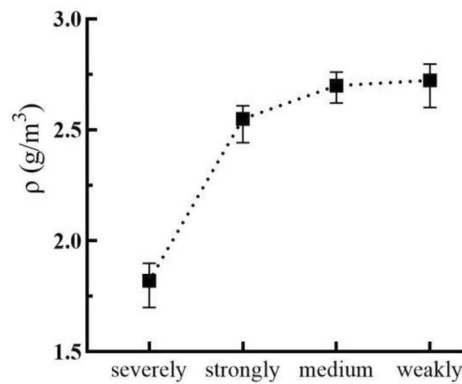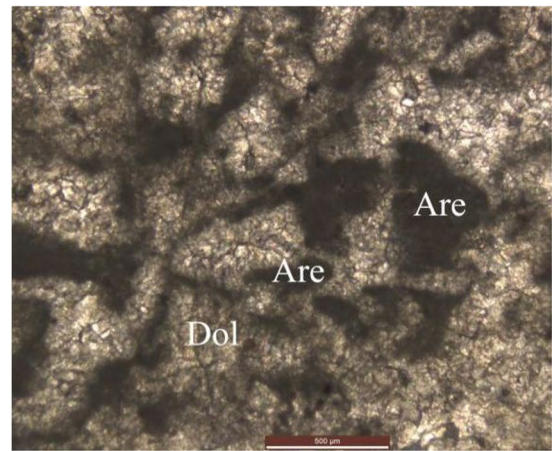
Figure 3. Relationship between Sandy Dolomite and ρ.

(a) Field tests Photos.

(b) Thin Section Identification ("Dol" is the abbreviation for "dolomite", "Are" is the abbreviation for "argillaceous").

Figure 4. Fierce Sandy Dolomite (E 102°41′38.34″, N 24°35′48.82″). (**a**) Field tests Photos. (**b**) Thin Section Identification ("Dol" is the abbreviation for "dolomite", "Are" is the abbreviation for "argillaceous").

section identification is shown in Fig. 5b. Schmidt rebound test on the strongly Sandy Dolomite sample shows that there is no rebound reading, indicating that the rock mass has been damaged.

In view of the aforementioned condition, this study only took the medium and weakly Sandy Dolomite as the research objects and analyzed the relationship between UCS and SH, $I_{s(50)}$, V_p, ρ, and n. The PMP of rock samples were tested using the methods specified by ISRM by establishing at least six sets of samples for each rock sample and calculating their average value.

Uniaxial compressive strength test

Rock core samples were cut using an automatic double-blade rock core cutting machine (SCQ-4A). The diameter of cylindrical core samples ranges from 44.35 to 65.40 mm, and the length ranges from 42.27 to 105.35 mm, with a length/diameter ratio of 0.66 ~ 2.17, as shown in Fig. 6. According to ISRM[63], the dolomite with different sandification grades were tested, the average value was taken as the UCS of the sample, and the length/diameter ratio of UCS should be 2.0 (50 mm × 100 mm)[64], which is not a standard size according to the ratio suggested by the ISRM[63] (> 2.5), Eq. (1) was adopted to convert the tested UCS* to UCS[65]. The sample should be loaded using an electro-hydraulic loading testing machine (HYE-2000), with the loading rate controlled between 1000 N/s and 2000 N/s, and a maximum loading capacity of 2000 kN. The rock core after the loading is shown in Fig. 7.

$$UCS = \frac{0.8668\,UCS^*}{0.778 + \frac{0.222}{L/D}} \tag{1}$$

where L is length, D is diameter, and UCS* is the UCS of the specimen at a ratio of L/D.

(a) Field tests Photos.
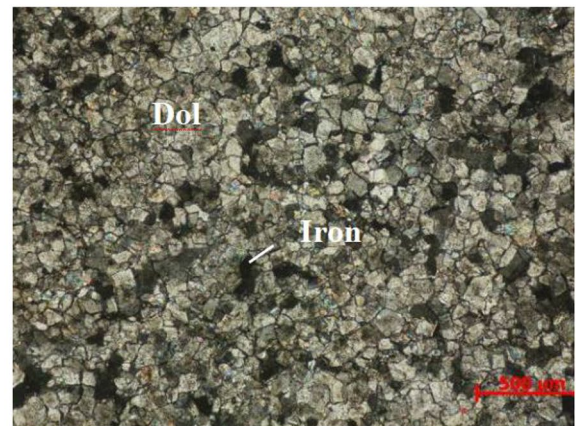
(b) Thin Section Identification ("Dol" is the abbreviation for "dolomite", "Iron" is the abbreviation for "Iron sludge").

Figure 5. Strongly Sandy Dolomite (E 102°39′48.92″, N 24°15′33.16″). (**a**) Field tests Photos. (**b**) Thin Section Identification ("Dol" is the abbreviation for "dolomite", "Iron" is the abbreviation for "Iron sludge").

Figure 6. Processed partially Sandy Dolomite core sample.

Figure 7. Rock core after UCS test.

Schmidt hammer rebound test

Schmidt hammer was originally invented to test the strength of concrete[66] and widely used as an indicator test for characterizing rock mass strength and deformation owing to its rapidity and ease of use, simplicity, portability, low cost and non-destructive nature[67]. Taking into account the impact factors, i.e., grain size of rock mass, anisotropy, sample size, weathering, and water content, its recommended method was improved in ISRM by Aydin[68].

Although the prediction accuracy of UCS could be improved via the Schmidt rebound test, there still are some shortcomings with RA. RA is based on limited experimental data sets and specific rock types, so some empirical formulas couldn't be applied generally in engineering practice. Obviously, the type of empirical formula of RA was chosen by academics objectively, and the outcomes are not appropriate enough. There are also some limitations with ANN. The data processing procedure is not clear or sufficiently transparent due to the hidden layer in the ANN is a black box, and the neural network may misunderstand the purpose of the researchers, so the prediction result should be verified frequently[26].

In this study, the HT-225 Schmidt hammer is utilized to calculate the rebound value, with the rebound test repeated 25 times for each sample. After removing the 5 maximum and 5 minimum values, the average of the remaining values was taken as the final result, and the field tests is shown in Fig. 8. Moreover, the minimum thickness of medium and weakly Sandy Dolomite is 110 mm and 75 mm according to the SH test results, respectively.

Point load test

The point load tester is portable and suitable for all types of rocks, and the sample of PLT need not to be cut and grind in the field or laboratory[27]. Protodyakonov[69] first put forward the idea of PLT with irregular blocks, and then D'Andrea[70] and Franklin[71] studied the transformation between rock's $I_{s(50)}$ and UCS. The point load tester can be used in various working conditions, i.e., outcrops, exploration pits, adits, roadways, and other caverns. The tested sample can be cylindrical, irregular shaped[72], or unprocessed, filling the blanks in soft and broken rock mass strength tests.

Figure 8. Field tests of SH test.

Figure 9. Field tests of PLT.

The PLT was performed only on the irregular block samples in this study (Fig. 9). The irregular sample can be calculated by utilizing the method of equivalent core diameter by ASTM standards, and $I_{s(50)}$ was determined by Eq. (2).

$$D_e^2 = \frac{4 \cdot W \cdot D}{\pi}$$
$$I_{s50} = \frac{4 \cdot P}{D_e^2}$$

(2)

where P is the failure load and De is the equivalent diameter of irregular blocks; D and W are of the maximum length and average width of the failure surface in millimetres.

Compressional wave velocity test

The compressional wave velocity, depends on mineral composition, texture, fabric, weathering degree, water content, and rock density[22,33,39,73], is calculated according to the transmission time between the transmitted and received wave, which is non-destructive and is also feasible in field or laboratory.

Clarify

The person in Fig. 8 is the author of this article, the recorder is the first author Meiqian Wang, and the experimenter is the fifth author Qinghua Wang.

Model analysis
Regression analysis

SPSS26.0 was utilized to investigate the relationship between UCS and $I_{s(50)}$, SH, V_p, ρ, and n based on the RA model. F-test and T-test with 95% confidence intervals were applied to confirm the model's dependability. The results demonstrated that the RA model has a wide variety of applications. The UCS value can be predicted simply by taking the parameters of most frequent rock properties as the input parameters of the RA model.

R^2 refers to the proportion of the regression sum of squares in the entire sum of squares, which is used for assessing the effectiveness of model. R^2 ranges from 0 to 1. The closer R^2 is near to 1, the higher the model's goodness is of fitness. The variance analysis of the entire regression equation is equal to the R^2-based hypothesis testing on the goodness of fit of the regression equation.

Convolutional neural network

The CNN consists of neurons with learnable weights and bias vectors. Compared with ANN, CNN adopts a mathematical operation called convolution, the traditional matrix multiplication of ANN was replaced by the mathematical operation in the network structure. The convolution operation effectively utilizes the two-dimensional structure of the input parameters, thus obtaining superior calculation outcomes.

Similar to conventional neural networks, CNN has two operational states, including the training and testing phases. The parameters are continuously optimized during the training phase in the learning phases of CNN, and the brand new dataset is used to evaluate the learning ability of the completed CNN in the testing phases (Fig. 10).

As shown in Fig. 11, the procedure contains four steps, which are discussed in detail below[54].

Step 1 Take a sample randomly from the data set and record it as X_p and Y_p, and input X_p into the network as the input parameter. The Y_p serves as the reference value of the result.

Step 2 Obtain the corresponding output Q_P through hierarchical calculation.

The calculation of CNN works base on the dot product of the input value and the weight matrix of each layer, and the final output value is obtained after operation layer by layer, as shown in Eq. (3).

$$Q_P = F_n(\cdots F_2(F_1(X_P W^1) W^2) \cdots) W^N$$

(3)

Step 3 Calculate the deviation between the output value Q_P and the corresponding true value Y_P.

Step 4 The weight matrix is adjusted through the back-propagation algorithm by minimizing the error.

In the *K*th layer, the weights of the *i*th neuron($i = 1, 2, \ldots, n$) can be written as $W_{i,1}$, $W_{i,2}$, ..., $W_{i,n}$. First, the weight coefficient $W_{i,j}$ should be set to a random number close to zero, allowing the gradient descent algorithm to find the local optimal solution. The training samples $X = (X_1, X_2, \ldots, X_n)$ are inputted, and the corresponding real outputs $Y = (Y_1, Y_2, \ldots Y_n)$ are obtained. Similarly, the real outputs of each layer can be calculated through weight calculation, as shown in Eq. (4–5):

$$U_i^k = \sum_{j=1}^{n+1} W_{i,j} X_i^{k-1}, \left(X_{n+1}^{k-1} = 1, W_{i,n+1} = \theta_i \right)$$

(4)

$$X_i^k = f(U_i^k)$$

(5)

where X_i^k refers to the output of the *i*th neuron in the *K*th layer; $W_{i,n+1}$ represents the threshold value θ_i. The desired and real outputs can be used to calculate the learning error d_i^k of each layer, as well as the corresponding errors of the hidden and output layers. If the output layer is recorded as *m*, the corresponding expression is shown in Eq. (6):
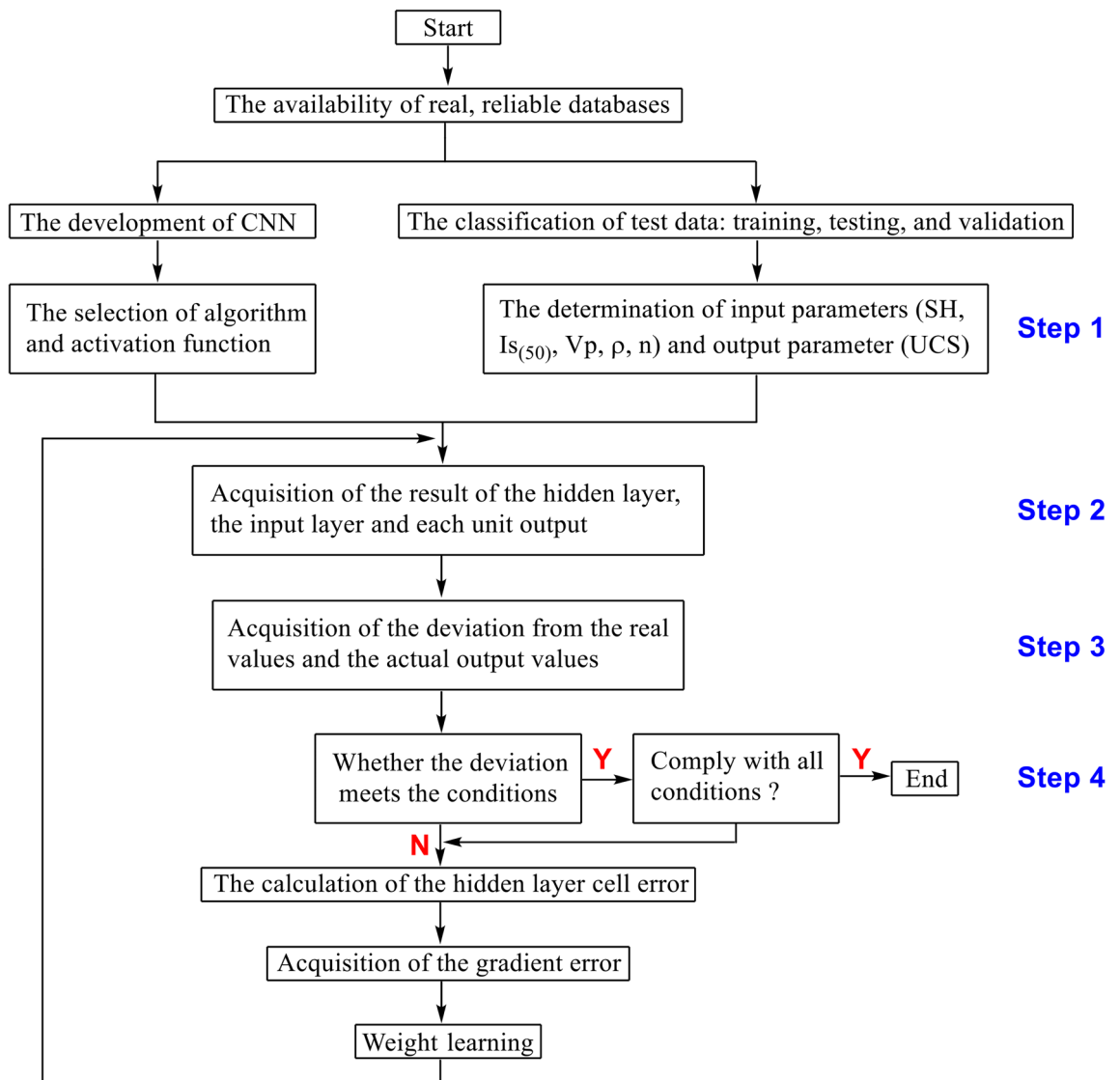
Figure 10. Flow chart for USC prediction based on CNN from index properties of rock.

$$d_i^k = X_i^m (1 - X_i^m)(X_i^m - Y_i) \tag{6}$$

Based on the gradient error calculation, the learning error of other layers can be calculated by Eq. (7):

$$d_i^k = X_i^k (1 - X_i^k) \sum_l W_{lj} d_i^{k+1} \tag{7}$$

Judge whether the error meets the calculation condition. If YES, the algorithm ends. Otherwise, the algorithm continues by modifying the weight coefficients through learning error, as shown in Eq. (8).

$$W_{ij}(t + 1) = W_{ij}(t) - \eta \cdot d_i^k \cdot X_j^{k+1} + \alpha \Delta W_{ij}(t) \tag{8}$$

where $\Delta W_{ij}(t) = -\eta \cdot d_i^k \cdot X_j^{k-1} + \alpha W_{ij}(t + 1) = W_{ij}(t) - W_{ij}(t - 1)$; η refers to the learning rate. The modified weight values will be used to calculate the real outputs until the error meets the conditions.

Advantages and performance evaluation of CNN

Compared with ANN, CNN is superior for the following reasons: (1) the introduction of the receptive field, also known as the local connection. Each neuron only receives connections from a small number of neurons in the upper layer, which only perceives data from local rather than all input dimensions. (2) The introduction of weights. Each neuron functions as a filter, using a fixed convolution kernel for the convolution of the entire set of inputs. (3) The introduction of multiple convolution kernels. By adding the channel dimension, CNN can extract more features. In addition to the global data distribution, multiple convolution kernels are conducive to
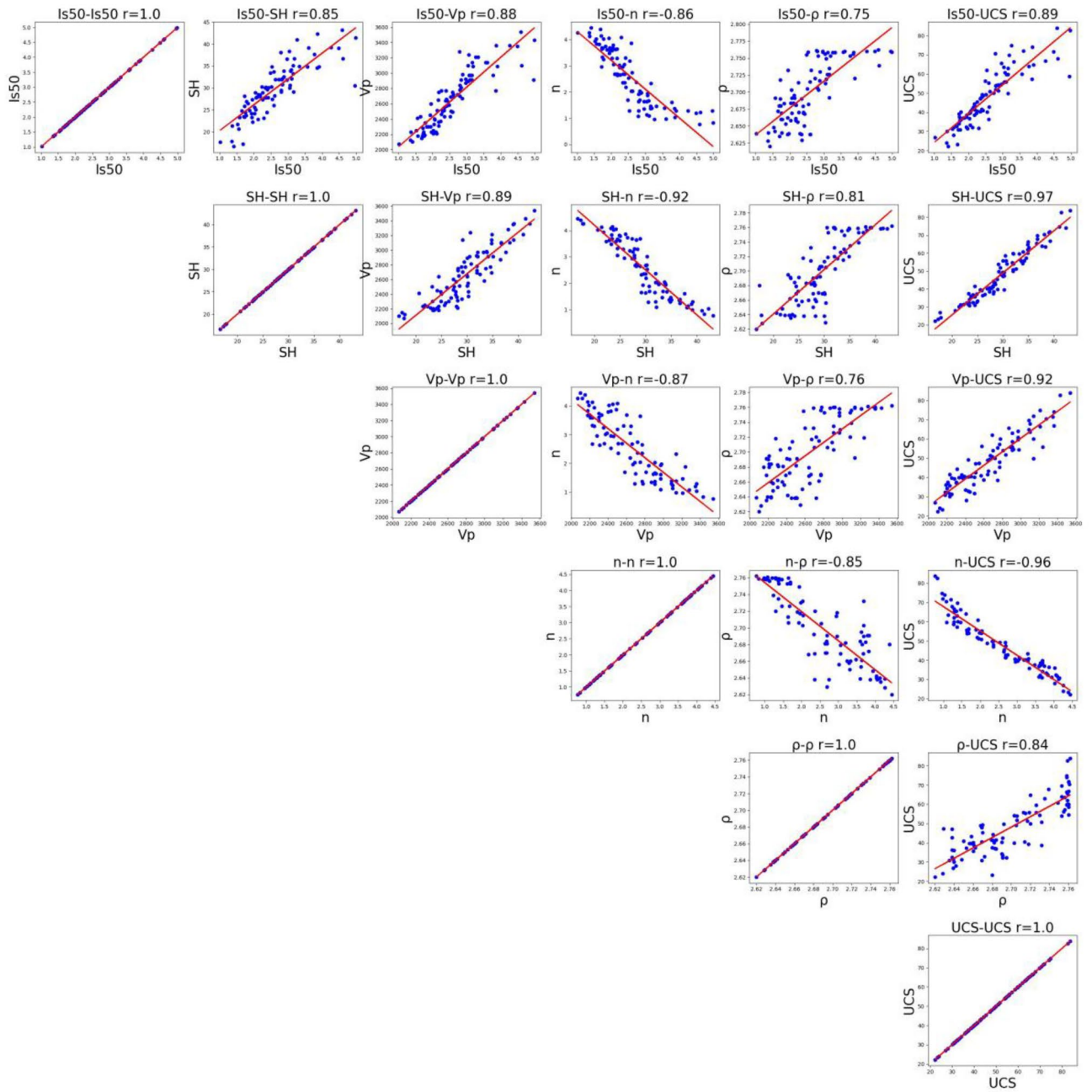
Figure 11. Scatterplot of medium Sandy Dolomite in terms of UCS, *SH*, $I_{s(50)}$, V_p, ρ, and *n*.

better modeling the correlation between local features in the condition of complex data analysis, which could be fitter various nonlinear mappings.

The prediction of UCS is a research topic in the field of RA. Five statistical indices are used for evaluating the performance of the prediction model in RA and CNN training, including MAE, MSE, RMSE, MAPE, and R^2. The corresponding mathematical expressions of all indices are shown in Eq. (9)–(13) [74,75]:

$$MAE = \frac{1}{m} \sum_{i=1}^{m} \left| f_i - f_{UCS,i} \right| \tag{9}$$

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (f_i - f_{UCS,i})^2 \tag{10}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(f_i - f_{UCS,i})^2} \tag{11}$$

$$MAPE = \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{f_i - f_{UCS,i}}{f_{UCS,i}}\right| \tag{12}$$

$$R^2 = 1 - \frac{\sum(f_i - f_{UCS,i})^2}{\sum(f_i - \bar{f})^2} \tag{13}$$

where f_i and $f_{UCS,i}$ refer to the measured and predicted values of UCS, respectively. $\bar{f}$ refers to the average of the measured UCS values. m is the number of the measured values of UCS.

In the CNN-based prediction method, statistical indices are employed to assess the effectiveness of the model. The input parameters are ranked according to the contribution of each parameter in the prediction of UCS. If one input parameter results in high MSE, MAE, MAPE, RMSE, and low R^2, it means that the deleted parameter has a significant influence on CNN.

Results and discussion

Two groups of real UCS data from Sandy Dolomite along and near the Yuxi Section of the CYWD Project were used for research object in this study. A total of 158 rock samples were collected and tests, and 158 groups of UCS and SH, $I_{s(50)}$, V_p, ρ, and n data were obtained.

Analysis results of medium Sandy Dolomite

The UCS of medium Sandy Dolomite was taken as the dependent variable, and SH, $I_{s(50)}$, V_p, ρ, and n were taken as the independent variables. SPSS 26.0 was used for fitting linear RA to obtain the relationship between the dependent and independent variables [Eq. (14)]. This method is obtained through multiple RA, and the used data in this method were considered the relationship between parameters. The equation with the highest R^2 was chosen as the empirical equation after various equations have been taken into account.

$$UCS = 1.21I_{s(50)} + 1.09SH + 0.006Vp - 4.528n + 6.889\rho - 9.471 \quad R^2 = 0.973 \tag{14}$$

Equation (14) shows a good correlation between the PMP (SH, $I_{s(50)}$, Vp, ρ, n and UCS) of medium Sandy Dolomite, with the R^2 of 0.973, indicating the goodness of fitness of the RA.

In addition to the descriptive statistics of UCS, SH, $I_{s(50)}$, V_p, ρ, and n, the Pearson correlation coefficients between each pair of variables were calculated based on the database (Fig. 11). The five parameters are highly correlated with UCS, which implied that they have a significant correlation in predicting UCS, and the correlation matrix is shown in Table 3.

The synergetic coefficient between UCS and SH, $I_{s(50)}$, V_p, and ρ of medium Sandy Dolomite is positive, which demonstrate a positive correlation between UCS and these four parameters (SH, $I_{s(50)}$, Vp, ρ) and that UCS increases with the rise of four parameters. The synergetic coefficient between UCS and n is negative, which demonstrate that there is a negative correlation between the two, and UCS decreases with the rise of n (Fig. 11 and Table 3). The UCS of medium Sandy Dolomite, five parameters exhibits good prediction performance, among which the most optimal indices are SH, $I_{s(50)}$, and V_p [Eq. (14) and Table 2].

Analysis results of weakly Sandy Dolomite

The UCS of medium Sandy Dolomite was taken as the dependent variable, and SH, $I_{s(50)}$, V_p, ρ, and n were taken as the independent variables. SPSS 26.0 was used for fitting linear RA to obtain the relationship between the dependent and independent variables [Eq. (15)].

$$UCS = 6.36I_{s(50)} + 0.631SH + 0.019Vp + 0.566n - 1.405\rho - 50.784 \quad R^2 = 0.795 \tag{15}$$

Equation (15) shows a good correlation between the PMP (SH, $I_{s(50)}$, Vp, ρ, n and UCS) of weakly Sandy Dolomite, with the R^2 of 0.795, indicating the goodness of fitness of the RA.

	$I_{s(50)}$	SH	Vp	n	ρ	UCS
$I_{s(50)}$	1.000000					
SH	0.846322	1.000000				
Vp	0.882991	0.885593	1.000000			
n	− 0.857435	− 0.918388	− 0.874688	1.000000		
ρ	0.751594	0.812292	0.756019	− 0.851650	1.000000	
UCS	0.886514	0.965271	0.919986	− 0.960027	0.841166	1.000000

Table 3. Correlation matrix between UCS and PMP.
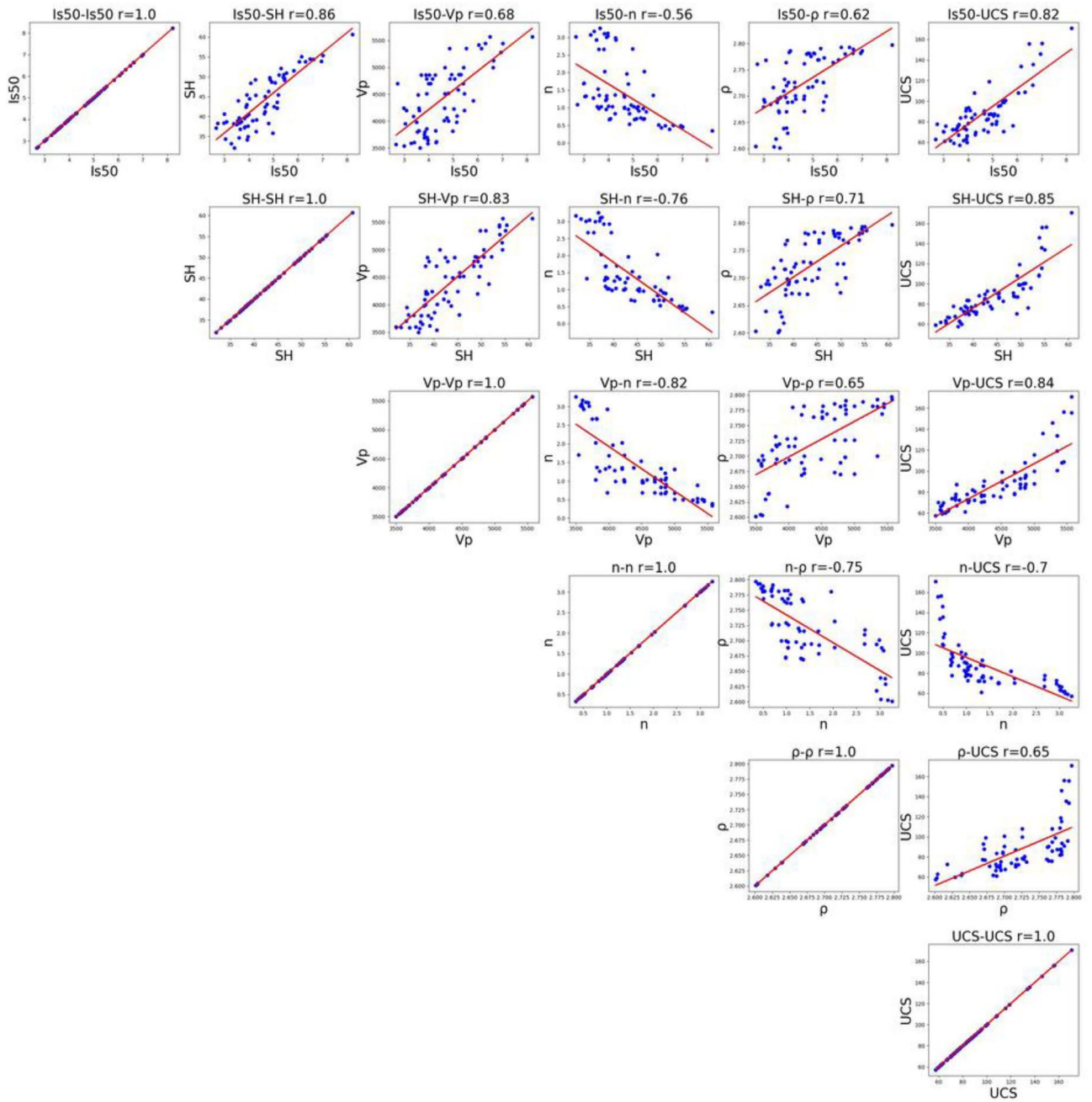
Figure 12. Scatterplot of weakly Sandy Dolomite in terms of UCS, *SH*, $I_{s(50)}$, V_p, *ρ*, and *n*.

	$I_{s(50)}$	SH	Vp	n	ρ	UCS
$I_{s(50)}$	1.000000					
SH	0.855277	1.000000				
Vp	0.680445	0.831487	1.000000			
n	− 0.5577330	− 0.757734	− 0.818705	1.000000		
ρ	0.620034	0.713312	0.653436	− 0.747263	1.000000	
UCS	0.820452	0.852588	0.840153	− 0.699622	0.654727	1.000000

Table 4. Correlation matrix between UCS and PMP.

In addition to the descriptive statistics of UCS, SH, $I_{s(50)}$, V_p, ρ, and n, the Pearson correlation coefficients between each pair of variables were calculated based on the database (Fig. 12). The five parameters are highly correlated with UCS, which implied that they have a significant correlation in predicting UCS, and the correlation matrix is shown in Table 4.

The synergetic coefficient between UCS and SH, $I_{s(50)}$, V_p, and ρ of weakly Sandy Dolomite is positive, which demonstrate a positive correlation between UCS and these four parameters (SH, $I_{s(50)}$, Vp, ρ) and that UCS increases with the rise of four parameters. The synergetic coefficient between UCS and n is negative, which demonstrate that there is a negative correlation between the two, and UCS decreases with the rise of n (Fig. 12 and Table 4). The UCS of weakly Sandy Dolomite, five parameters exhibits good prediction performance, among which the most optimal parameters are SH and $I_{s(50)}$, the effective parameters are V_p and ρ [Eq. (15) and Table 4].

CNN training

MSE was employed as the loss function of CNN training. Both the input and output data must first be normalized so that produce a remarkable convergence effect during training. In fact, the description of MSE is actually the prediction error following normalization. Meanwhile, MAE was also used to describe the prediction error. MAE will not alter the input or output data and can accurately reflect the prediction error of UCS in standard units, and thus MAE is also utilized for final verification.

There are five input feature parameters (SH, $I_{s(50)}$, Vp, ρ, n) in training, and the dimension is 5*1*n, n is the number of input data groups each time. The UCS is the output parameter and its dimension is 1*n. In order to fit the mapping relationship between the input and output parameters, two types of hidden layers were designed to ensure the degree of non-linearity in CNN. The structure of input layer + hidden layer I + hidden layer II + output layer was established, and the Bayesian optimization method was employed to search the number of neurons in the hidden layer, ensuring that the ReLU layer was used for nonlinear operation following each layer.

With different input data units, we performed normalized operations for each input so that the data values are uniformly mapping within the range of [−1, 1]. The standardized predicted value and real value (measured value of UCS) are also used in the loss function during training to calculate MSE.

For medium Sandy Dolomite, the data of 75% ($90 \times 75\% \approx 67$) database with a size of 90 was used for CNN training and selection, and the remaining 25% ($90-67=23$) was used for comparison and verification. For weakly Sandy Dolomite, the data of 75% ($68 \times 75\% = 51$) database with a size of 68 was used for CNN training and selection, and the remaining 25% ($68-51=17$) was used for comparison and verification.

Comparison between RA and CNN

Comparison between RA and CNN of medium Sandy Dolomite

According to the CNN calculation results, the optimal CNN structure for predicting medium Sandy Dolomite data is 5–300–300–1. That is, the input layer with five parameters (SH, $I_{s(50)}$, Vp, ρ, n) + the hidden layer I with 300 neurons + the hidden layer II with 300 neurons + the output layer with one parameter (UCS). The values of UCS in validation data set were predicted and compared with the real measured values for evaluating and comparing the performance of RA and CNN. Similarly, R^2, MSE, RMSE, MAE, and MAPE were calculated by using Eqs. (9)–(13), as shown in Table 5.

The value of MAE, MSE, RMSE, and MAPE in CNN are all smaller than in the RA, while R^2 of CNN is larger than that of the RA. The larger the R^2 is, the smaller the MSE, RMSE, and MAPE are, which implied that the

Model	MAE	MSE	RMSE	MAPE (%)	R^2
RA	1.848	5.692	2.386	4.397	0.973
CNN	1.407	3.232	1.798	3.168	0.979

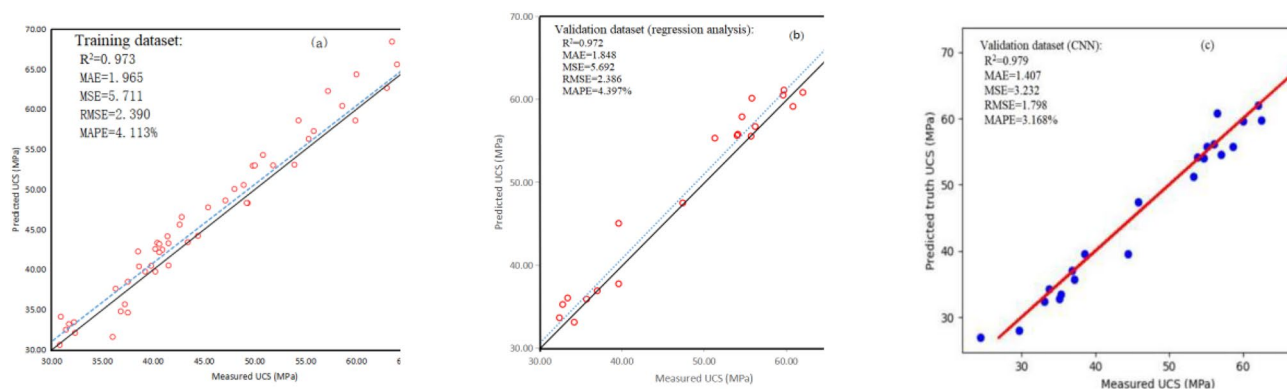**Table 5.** MAE, MSE, RMSE, MAPE, and R^2 calculation results of medium Sandy Dolomite.
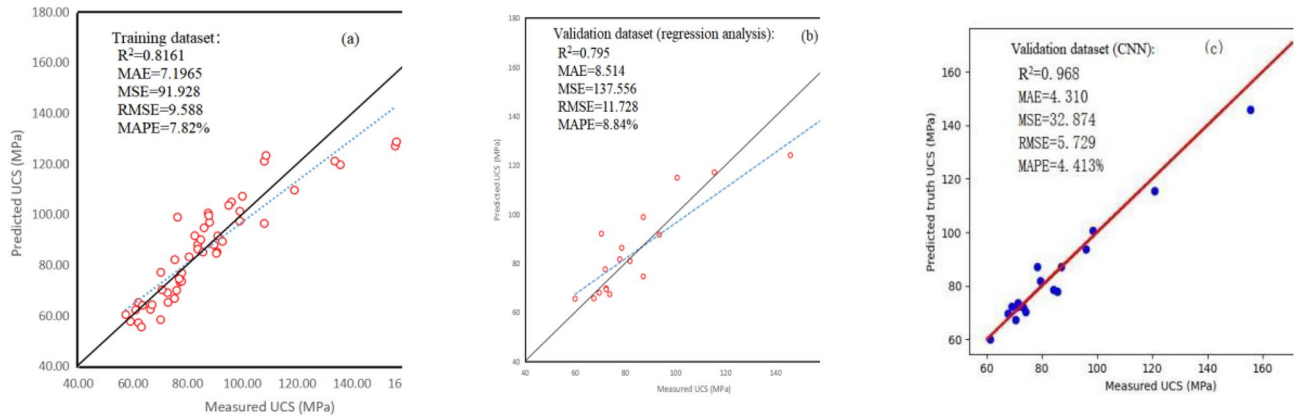

Figure 13. R^2, MSE, RMSE, MAE, MAPE and mean ratio of predicted and measured UCS values. (**a**) training data set; (**b**) validation data set (RA); (**c**) validation data set (CNN).

Model	MAE	MSE	RMSE	MAPE (%)	R²
RA	8.514	137.556	11.728	8.840	0.795
CNN	4.310	32.824	5.729	4.413	0.968

Table 6. MAE, MSE, RMSE, MAPE, and R^2 calculation results of weakly Sandy Dolomite.

Figure 14. R^2, MSE, RMSE, MAE, MAPE and mean ratio of predicted and measured UCS values. (**a**) training data set; (**b**) validation data set (RA); (**c**) validation data set (CNN).

performance of the prediction model is better (Table 5). Therefore, the CNN is more suitable for predicting the UCS of medium Sandy Dolomite than RA due to its more precise results. The predicted and measured UCS values are displayed in Fig. 14a–c to visualize the advantages and disadvantages of RA and CNN in predicting UCS. The dot result usually fluctuates near the 1:1 line (the solid line), indicating that the training and validation data sets are suitable for the prediction of UCS (Fig. 13a,b).

The results base on validation data illustrate that the R^2 of RA is 0.972, while the R^2 of CNN is 0.979. Compared with the R^2 of 0.973 based on training data, the R^2 of CNN is 0.72% higher than that of RA. Similarly, compared with the results based on training data, the MAE, MSE, RMSE, and MAPE values of CNN are 22.45%, 43.07%, 24.60%, and 29.88% lower than those of RA, indicating that CNN has an evidently better performance than RA in terms of the UCS prediction of medium Sandy Dolomite (Fig. 13).

Comparison between RA and CNN on weakly Sandy Dolomite

According to the CNN calculation results, the optimal CNN structure for predicting weakly Sandy Dolomite data is 5–128–224–1. That is, the input layer with five parameters (SH, $I_{s(50)}$, Vp, ρ, n) + the hidden layer I with 128 neurons + the hidden layer II with 224 neurons + the output layer with one parameter (UCS). The values of UCS in validation data set were predicted and compared with the real measured values for evaluating and comparing the performance of RA and CNN. Similarly, R^2, MSE, RMSE, MAE, and MAPE were calculated by using Eqs. (9)–(13), as shown in Table 6.

The value of MAE, MSE, RMSE, and MAPE in CNN are all smaller than in the RA, while R^2 of CNN is larger than that of the RA. The larger the R^2 is, the smaller the MSE, RMSE, and MAPE are, which imply that the performance of the prediction model is better (Table 6). Therefore, the CNN is more suitable for predicting the UCS of weakly Sandy Dolomite than RA due to its more precise results. The predicted and measured UCS values are displayed in Fig. 14a–c to visualize the advantages and disadvantages of RA and CNN in predicting UCS. The dot result usually fluctuates near the 1:1 line (the solid line), indicating that the training and validation data sets are suitable for the prediction of UCS (Fig. 14a,b).

The results based on validation data illustrate that the R^2 of RA is 0.795, while the R^2 of CNN is 0.968. Compared with the R^2 of 0.8161 based on training data, the R^2 of CNN is 21.2% higher than that of RA. Similarly, compared with the results based on training data, the MAE, MSE, RMSE, and MAPE values of CNN are 58.42%, 113.87%, 62.76%, and 56.61% lower than those of RA, indicating that CNN has an evidently better performance than RA in terms of the UCS prediction of weakly Sandy Dolomite (Fig. 14).

Conclusions

In conclusion, SH, $I_{s(50)}$, Vp, ρ, and n were utilized to predict the UCS of Sandy Dolomite, and R^2, MSE, RMSE, MAE, and MAPE were utilized to evaluate and compare the performance of both RA and CNN in this study. The main conclusions are as follows:

(1) Bayesian optimization is utilized in CNN to search the data in the hidden layer, and obtained prediction results were much higher than that of RA, indicating the reliability of the CNN-based prediction model. The prediction precision of the CNN model can be further improved as the number of indicators increases.

(2) Fierce Sandy Dolomite is a rock with low density and high porosity; strongly Sandy Dolomite is a rock with medium density and porosity; medium and weakly Sandy Dolomite is a rock with high density and low porosity.

(3) The sandification grade of dolomite increases with the rise of porosity and decreases with the rise of density.

(4) Schmidt hammer rebound test results demonstrated that the minimum thickness of medium and weakly Sandy Dolomite is 110 mm and 75 mm, respectively.

(5) As for Sandy Dolomite, there is a positive correlation between UCS and SH, $I_{s(50)}$, V_p, and ρ, while there is a negative correlation between UCS and n.

Compared with RA, CCN has the higher accuracy for predicting the UCS of Sandy Dolomite. However, the mathematic expressions between UCS and its relevant indexes are unavailable via CNN analysis. Such impact maybe come from CNN itself, in most situations, the interaction between inputs and results does not provide a deterministic pattern due to the so called "black box". This indicates that some more straightforward and understandable methods within a full system are needed to be further researched.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

1. Li, J., Mu, H. & Mi, J. Preliminary study on engineering geological characteristics of sanding dolomite. in *Application and Development of Hydraulic Tunnel Technology: Survey* (2018).
2. Jiang, Y. *et al.* Failure analysis and control measures for tunnel faces in water-rich Sandy Dolomite formations. *Eng. Fail. Anal.* https://doi.org/10.1016/j.engfailanal.2022.106350 (2022).
3. Wang, M. *et al.* Study on construction and reinforcement technology of dolomite sanding tunnel. *Sustainability* https://doi.org/10.3390/su14159217 (2022).
4. Wang, P., Yao, J. & Jiang, L. Sandification characteristics of guizhou dolomite and the influence on tunnel support structure. *J. Guizhou Univ. (Nat. Sci.)* https://doi.org/10.15958/j.cnki.gdxbzrb.2019.03.08 (2019).
5. Charles, R. F. Subsurface trenton and sub-trenton rocks in Ohio, New York, Pennsylvania, and West Virginia. *AAPG Bull.* **32**, 1457–1492. https://doi.org/10.1306/3d933bff-16b1-11d7-8645000102c1865d (1948).
6. Chown, E. H. & Caty, J. Diagenesis of the Aphebian Mistassini regolith, Quebec, Canada. *Precamb. Res.* **19**, 285–299. https://doi.org/10.1016/0301-9268(83)90017-7 (1983).
7. Maghfouri, S., Hosseinzadeh, M. R., Lentz, D. R. & Choulet, F. Geological and geochemical constraints on the Farahabad vent-proximal sub-seafloor replacement SEDEX-type deposit, Southern Yazd basin, Iran. *J. Geochem. Explor.* https://doi.org/10.1016/j.gexplo.2019.106436 (2020).
8. Richter, D. K., Gillhaus, A. & Neuser, R. D. The alteration and disintegration of dolostones with stoichiometric dolomite crystals to dolomite sand: New insights from the Franconian Alb (Upper Jurassic, SE Germany). *Z. Deutsch. Gesellsch. Geowissensch.* **169**, 27–46. https://doi.org/10.1127/zdgg/2018/0150 (2018).
9. Attia, R. M. & Awny, E. G. Leaching characterisations and recovery of copper and uranium with glycine solution of Sandy Dolomite, Allouga area, South Western Sinai, Egypt. *Int. J. Environ. Anal. Chem/* **1**, 1–14. https://doi.org/10.1080/03067319.2021.2014471 (2021).
10. Bosellini, A. & Hardie, L. A. Depositional theme of a marginal marine evaporite. *Sedimentology* **20**, 5–27. https://doi.org/10.1111/j.1365-3091.1973.tb01604.x (1973).
11. Garrido, M. E. *et al.* Predicting the uniaxial compressive strength of a limestone exposed to high temperatures by point load and leeb rebound hardness testing. *Rock Mech. Rock Eng.* **55**, 1–17. https://doi.org/10.1007/s00603-021-02647-0 (2022).
12. Ji, K. & Arson, C. Tensile strength of calcite/HMWM and silica/HMWM interfaces: A molecular dynamics analysis. *Construct. Build. Mater.* https://doi.org/10.1016/j.conbuildmat.2020.118925 (2020).
13. Walton, G. Initial guidelines for the selection of input parameters for cohesion-weakening-friction-strengthening (CWFS) analysis of excavations in brittle rock. *Tunnel. Underground Space Technol.* **84**, 189–200. https://doi.org/10.1016/j.tust.2018.11.019 (2019).
14. Tahmasbi, S., Giacomini, A., Wendeler, C. & Buzzi, O. On the computational efficiency of the hybrid approach in numerical simulation of rockall flexible chain-link mesh. *Rock Mech. Rock Eng.* **52**, 3849–3866. https://doi.org/10.1007/s00603-019-01795-8 (2019).
15. He, M. *et al.* Deep convolutional neural network for fast determination of the rock strength parameters using drilling data. *Int. J. Rock Mech. Min. Sci.* https://doi.org/10.1016/j.ijrmms.2019.104084 (2019).
16. Cowie, S. & Walton, G. The effect of mineralogical parameters on the mechanical properties of granitic rocks. *Eng. Geol.* **240**, 204–225. https://doi.org/10.1016/j.enggeo.2018.04.021 (2018).
17. Jeffery, M., Huang, J., Fityus, S., Giacomini, A. & Buzzi, O. A rigorous multiscale random field approach to generate large scale rough rock surfaces. *Int. J. Rock Mech. Min. Sci.* https://doi.org/10.1016/j.ijrmms.2021.104716 (2021).
18. Matin, S. S., Farahzadi, L., Makaremi, S., Chelgani, S. C. & Sattari, G. Variable selection and prediction of uniaxial compressive strength and modulus of elasticity by random forest. *Appl. Soft Comput.* **70**, 980–987. https://doi.org/10.1016/j.asoc.2017.06.030 (2018).
19. Murlidhar, B. R., Ahmed, M., Mavaluru, D., Siddiqi, A. F. & Mohamad, E. T. Prediction of rock interlocking by developing two hybrid models based on GA and fuzzy system. *Eng. Comput.* **35**, 1419–1430. https://doi.org/10.1007/s00366-018-0672-9 (2019).
20. Zhao, T., Song, C., Lu, S. & Xu, L. Prediction of uniaxial compressive strength using fully Bayesian gaussian process regression (fB-GPR) with model class selection. *Rock Mech. Rock Eng.* https://doi.org/10.1007/s00603-022-02964-y (2022).
21. Kong, F., Xue, Y., Qiu, D., Gong, H. & Ning, Z. Effect of grain size or anisotropy on the correlation between uniaxial compressive strength and Schmidt hammer test for building stones. *Construct. Build. Mater.* https://doi.org/10.1016/j.conbuildmat.2021.123941 (2021).
22. Ng, I.-T., Yuen, K.-V. & Lau, C.-H. Predictive model for uniaxial compressive strength for Grade III granitic rocks from Macao. *Eng. Geol.* **199**, 28–37. https://doi.org/10.1016/j.enggeo.2015.10.008 (2015).
23. Baykasoglu, A., Gullu, H., Canakci, H. & Ozbakir, L. Prediction of compressive and tensile strength of limestone via genetic programming. *Expert Syst. Appl.* **35**, 111–123. https://doi.org/10.1016/j.eswa.2007.06.006 (2008).

24. Alzabeebee, S., Mohammed, D. A. & Alshkane, Y. M. Experimental study and soft computing modeling of the unconfined compressive strength of limestone rocks considering dry and saturation conditions. *Rock Mech. Rock Eng.* https://doi.org/10.1007/s00603-022-02948-y (2022).

25. Kalantari, S., Hashemolhosseini, H. & Baghbanan, A. Estimating rock strength parameters using drilling data. *Int. J. Rock Mech. Min. Sci.* **104**, 45–52. https://doi.org/10.1016/j.ijrmms.2018.02.013 (2018).

26. Wang, M. & Wan, W. A new empirical formula for evaluating uniaxial compressive strength using the Schmidt hammer test. *Int. J. Rock Mech. Min. Sci.* https://doi.org/10.1016/j.ijrmms.2019.104094 (2019).

27. Wang, M. *et al.* Summary of the transformational relationship between point load strength index and uniaxial compressive strength of rocks. *Sustainability* https://doi.org/10.3390/su141912456 (2022).

28. Rabat, Á., Cano, M. & Tomás, R. Effect of water saturation on strength and deformability of building calcarenite stones: Correlations with their physical properties. *Construct. Build. Mater.* https://doi.org/10.1016/j.conbuildmat.2019.117259 (2020).

29. Rabat, Á., Cano, M., Tomás, R., Tamayo, Á. E. & Alejano, L. R. Evaluation of strength and deformability of soft sedimentary rocks in dry and saturated conditions through needle penetration and point load tests: A comparative study. *Rock Mech. Rock Eng.* **53**, 2707–2726. https://doi.org/10.1007/s00603-020-02067-6 (2020).

30. Kahraman, S. Evaluation of simple methods for assessing the uniaxial compressive strength of rock. *Int. J. Rock Mech. Min. Sci.* **38**, 981–994. https://doi.org/10.1016/s1365-1609(01)00039-9 (2001).

31. Kahraman, S., Altun, H., Tezekici, B. S. & Fener, M. Sawability prediction of carbonate rocks from shear strength parameters using artificial neural networks. *Int. J. Rock Mech. Min. Sci.* **43**, 157–164. https://doi.org/10.1016/j.ijrmms.2005.04.007 (2006).

32. Le, T.-T., Skentou, A. D., Mamou, A. & Asteris, P. G. Correlating the unconfined compressive strength of rock with the compressional wave velocity effective porosity and Schmidt hammer rebound number using artificial neural networks. *Rock Mech. Rock Eng.* https://doi.org/10.1007/s00603-022-02992-8 (2022).

33. Mishra, D. A. & Basu, A. Estimation of uniaxial compressive strength of rock materials by index tests using regression analysis and fuzzy inference system. *Eng. Geol.* **160**, 54–68. https://doi.org/10.1016/j.enggeo.2013.04.004 (2013).

34. Parsajoo, M., Armaghani, D. J. & Asteris, P. G. A precise neuro-fuzzy model enhanced by artificial bee colony techniques for assessment of rock brittleness index. *Neural Comput. Appl.* **34**, 3263–3281. https://doi.org/10.1007/s00521-021-06600-8 (2022).

35. Ghasemi, E., Kalhori, H., Bagherpour, R. & Yagiz, S. Model tree approach for predicting uniaxial compressive strength and Young's modulus of carbonate rocks. *Bull. Eng. Geol. Environ.* **77**, 331–343. https://doi.org/10.1007/s10064-016-0931-1 (2018).

36. Liang, M., Mohamad, E. T., Faradonbeh, R. S., Jahed Armaghani, D. & Ghoraba, S. Rock strength assessment based on regression tree technique. *Eng. Comput.* **32**, 343–354. https://doi.org/10.1007/s00366-015-0429-7 (2016).

37. Alemdag, S., Gurocak, Z., Cevik, A., Cabalar, A. & Gokceoglu, C. Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming. *Eng. Geol.* **203**, 70–82 (2016).

38. Jahed Armaghani, D., Mohd Amin, M. F., Yagiz, S., Faradonbeh, R. S. & Abdullah, R. A. Prediction of the uniaxial compressive strength of sandstone using various modeling techniques. *Int. J. Rock Mech. Min. Sci.* **85**, 174–186. https://doi.org/10.1016/j.ijrmms.2016.03.018 (2016).

39. Momeni, E., Armaghani, D. J., Hajihassani, M. & Amin, M. F. M. Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement* **60**, 50–63 (2015).

40. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536. https://doi.org/10.1038/323533a0 (1986).

41. Huang, H., Li, Q. & Zhang, D. Deep learning based image recognition for crack and leakage defects of metro shield tunnel. *Tunnel. Underground Space Technol.* **77**, 166–176. https://doi.org/10.1016/j.tust.2018.04.002 (2018).

42. Karimpouli, S., Tahmasebi, P. & Saenger, E. H. Ultrasonic prediction of crack density using machine learning: A numerical investigation. *Geosci. Front.* https://doi.org/10.1016/j.gsf.2021.101277 (2022).

43. Zhang, B., Zhou, L. & Zhang, J. A methodology for obtaining spatiotemporal information of the vehicles on bridges based on computer vision. *Comput. Aided Civil Infrastruct. Eng.* **34**, 471–487. https://doi.org/10.1111/mice.12434 (2019).

44. Zhang, Z.-X., Chi, L. Y., Qiao, Y. & Hou, D.-F. Fracture initiation, gas ejection, and strain waves measured on specimen surfaces in model rock blasting. *Rock Mech. Rock Eng.* **54**, 647–663. https://doi.org/10.1007/s00603-020-02300-2 (2020).

45. Öge, İF. Regression analysis and neural network fitting of rock mass classification systems. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi* **20**, 354–368 (2018).

46. Lai, X. *et al.* Research on mechanism of rockburst induced by mined coal-rock linkage of sharply inclined coal seams. *Int. J. Miner. Metall. Mater.* https://doi.org/10.1007/s12613-024-2833-8 (2024).

47. Alzubaidi, F., Mostaghimi, P., Si, G., Swietojanski, P. & Armstrong, R. T. Automated rock quality designation using convolutional neural networks. *Rock Mech. Rock Eng.* **55**, 3719–3734. https://doi.org/10.1007/s00603-022-02805-y (2022).

48. Bergen, K. J., Johnson, P. A., de Hoop, M. V. & Beroza, G. C. Machine learning for data-driven discovery in solid Earth geoscience. *Science* https://doi.org/10.1126/science.aau0323 (2019).

49. Chen, J., Yang, T., Zhang, D., Huang, H. & Tian, Y. Deep learning based classification of rock structure of tunnel face. *Geosci. Front.* **12**, 395–404. https://doi.org/10.1016/j.gsf.2020.04.003 (2021).

50. Chen, J., Zhou, M., Zhang, D., Huang, H. & Zhang, F. Quantification of water inflow in rock tunnel faces via convolutional neural network approach. *Autom. Construct.* https://doi.org/10.1016/j.autcon.2020.103526 (2021).

51. Ferreira, A. & Giraldi, G. Convolutional neural network approaches to granite tiles classification. *Expert Syst. Appl.* **84**, 1–11. https://doi.org/10.1016/j.eswa.2017.04.053 (2017).

52. Huang, L., Li, J., Hao, H. & Li, X. Micro-seismic event detection and location in underground mines by using convolutional neural networks (CNN) and deep learning. *Tunnel. Underground Space Technol.* **81**, 265–276. https://doi.org/10.1016/j.tust.2018.07.006 (2018).

53. Karimpouli, S. & Tahmasebi, P. Image-based velocity estimation of rock using convolutional neural networks. *Neural Netw.* **111**, 89–97. https://doi.org/10.1016/j.neunet.2018.12.006 (2019).

54. Niu, Y., Mostaghimi, P., Shabaninejad, M., Swietojanski, P. & Armstrong, R. T. Digital rock segmentation for petrophysical analysis with reduced user bias using convolutional neural networks. *Water Resour. Res.* https://doi.org/10.1029/2019wr026597 (2020).

55. Sidorenko, M., Orlov, D., Ebadi, M. & Koroteev, D. Deep learning in denoising of micro-computed tomography images of rock samples. *Comput. Geosci.* https://doi.org/10.1016/j.cageo.2021.104716 (2021).

56. Tang, P., Zhang, D. & Li, H. Predicting permeability from 3D rock images based on CNN with physical information. *J. Hydrol.* https://doi.org/10.1016/j.jhydrol.2022.127473 (2022).

57. Tian, J., Qi, C., Sun, Y. & Yaseen, Z. M. Surrogate permeability modelling of low-permeable rocks using convolutional neural networks. *Comput. Methods Appl. Mech. Eng.* https://doi.org/10.1016/j.cma.2020.113103 (2020).

58. Wang, Y. D., Shabaninejad, M., Armstrong, R. T. & Mostaghimi, P. Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images. *Appl. Soft Comput.* https://doi.org/10.1016/j.asoc.2021.107185 (2021).

59. Wu, J., Yin, X. & Xiao, H. Seeing permeability from images: Fast prediction with convolutional neural networks. *Sci. Bull.* **63**, 1215–1222. https://doi.org/10.1016/j.scib.2018.08.006 (2018).

60. Zhou, Y., Wong, L. N. Y. & Tse, K. K. C. Novel rock image classification: The proposal and implementation of RockNet. *Rock Mech. Rock Eng.* https://doi.org/10.1007/s00603-022-03003-6 (2022).

61. Wang, M., Wu, Y., Song, B. & Xu, W. Point load strength test power index of irregular Sandy Dolomite blocks. *Rock Mech. Rock Eng.* https://doi.org/10.1007/s00603-023-03733-1 (2023).
62. Matula, M. Classification of rocks and soils for engineering geological mapping part I: Rock and soil materials. *Bull. Eng. Geol. Environ.* **1**, 1–10 (1979).
63. ISRM. (International Society for Rock Mechanics, 2007).
64. ASTM. Standard test methods for compressive strength and elastic moduli of intact rock core specimens under varying states of stress and temperatures: D7012–14. in *Annual Book of ASTM Standards*. (2014).
65. Liu, Q., Zhao, Y. & Zhang, X. Case study: Using the point load test to estimate rock strength of tunnels constructed by a tunnel boring machine. *Bull. Eng. Geol. Environ.* **78**, 1727–1734. https://doi.org/10.1007/s10064-017-1198-x (2019).
66. Schmidt, E. A non-destructive concrete tester. *Concrete* **59**, 34–35 (1951).
67. Yilmaz, I. A new testing method for indirect determination of the unconfined compressive strength of rocks. *Int. J. Rock Mech. Min. Sci.* **46**, 1349–1357. https://doi.org/10.1016/j.ijrmms.2009.04.009 (2009).
68. Aydin, A. ISRM Suggested method for determination of the Schmidt hammer rebound hardness: Revised version. *Int. J. Rock Mech. Min. Sci.* **46**, 627–634. https://doi.org/10.1016/j.ijrmms.2008.01.020 (2009).
69. Protodyakonov, M. *Proceedings of the International Conference on Strata Control*, 187–195.
70. D'Andrea, D. V., Fischer, R. L. & Fogelson, D. E. *Prediction of Compressive Strength from Other Rock Properties*, vol. 6702 (US Department of the Interior, Bureau of Mines, 1964).
71. ISRM. Suggested method for determining point load strength. *Int. J. Rock Mech. Min. Sci. Geomech. Abstracts* **22**, 51–60. https://doi.org/10.1016/0148-9062(85)92327-7 (1985).
72. Şahin, M., Ulusay, R. & Karakul, H. Point load strength index of half-cut core specimens and correlation with uniaxial compressive strength. *Rock Mech. Rock Eng.* **53**, 3745–3760. https://doi.org/10.1007/s00603-020-02137-9 (2020).
73. Yaşar, E. & Erdoğan, Y. Estimation of rock physicomechanical properties using hardness methods. *Eng. Geol.* **71**, 281–288. https://doi.org/10.1016/s0013-7952(03)00141-8 (2004).
74. Miah, M. I., Ahmed, S., Zendehboudi, S. & Butt, S. Machine learning approach to model rock strength: Prediction and variable selection with aid of log data. *Rock Mech. Rock Eng.* **53**, 4691–4715. https://doi.org/10.1007/s00603-020-02184-2 (2020).
75. Zendehboudi, S., Rezaei, N. & Lohi, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Appl. Energy* **228**, 2539–2566. https://doi.org/10.1016/j.apenergy.2018.06.051 (2018).

Author contributions

Wei Xu conceived the project, supervised the research and guided the idea of writing the paper. Wenlian Liu, Haiming Liu, Ting Xie and Qinghua Wang guided the tests and the analysis of data. Meiqian Wang and Qinghua Wang contributed to the field and indoor tests studies, simulation analysis and data processing. Meiqian Wang and Ting Xie contributed tohave made substantial contributions to the substantively revised it. Meiqian Wang and Wei Xu contributed to original draft of the manuscript which was revised by all authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to W.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.