AIMS Environmental Science

AIMS Environmental Science, 12(6): 979–998.
DOI: 10.3934/environsci.2025043
Received: 08 September 2025
Revised: 22 October 2025
Accepted: 19 November 2025
Published: 28 November 2025

http://www.aimspress.com/journal/environmental

Research article

A real-time evaluation method of unstable rock risk level based on Microseismic data and CWT-CNN integrated algorithm

Yan Du[1,2], Renjian Li[1], Mowen Xie[1], Yujing Jiang[3], Santos Daniel CHICAS[4], Jingnan Liu[1,3,*] and Weikang Lu

[1] Beijing Key Laboratory of Urban Underground Space Engineering, University of Science and Technology Beijing, Beijing 100083, China
[2] School of Future Cities, University of Science and Technology Beijing, Beijing 100083, China
[3] Graduate School of Engineering, Nagasaki University, Nagasaki 852-8521, Japan
[4] Department of Agro-environmental Science, Faculty of Agriculture, Kyushu University, 744 Motooka, Fukuoka 819-0395, Japan

* **Correspondence:** Jingnan Liu, Email: ljn1203846583@163.com; Tel: +8617732625563.

Abstract: Geological hazards caused by unstable rocks, including rock collapse and fall, pose significant threats to both engineering productivity and the safety of residents, resulting in substantial economic losses. This study proposed a comprehensive CNN (Convolutional Neural Networks) classification recognition algorithm based on continuous wavelet analysis of microseismic monitoring data, which establishes a CNN recognition-based method for identifying risk levels of unstable rocks to predict their real-time collapse. To collect training data, fixed-base tests were conducted, and validation was performed using vibration data from freeze-thaw tests. Results of the study showed that the accuracy of the CNN classification recognition model trained with fixed-base test data reached 97.6%, and the per-second classification accuracy of vibration segments from freeze-thaw tests was above 86%. Furthermore, the study discussed the correlation between the calculated risk-level eigenvalues and safety coefficients. The CNN risk-level eigenvalues were found to be highly negatively correlated with the safety coefficients, with correlation coefficients as high as 0.8703. Finally, the study verified the superiority of the precursor identification of the risk-level evaluation method by comparing it with the safety coefficients.

Keywords: unstable rock; continuous wavelet transform; convolutional neural networks; risk assessment

1. Introduction

Unstable rock is a global disaster that seriously threatens the safety of mountain highway transportation, people's lives and property, and engineering work [1–3]. Therefore, identifying unstable rock is a critical problem that engineering must solve. Current research shows that the main cause of rock failure is dynamic failure caused by rock system instability [4]. Thus, the dynamic detection index to identify rock system damage is key to early warning of a collapse disaster [5]. Ma et al. [6,7] proposed a new method for evaluating the stability of USR (Unstable Rock) using a remotely positioned laser Doppler vibrometer (LDV). Du et al. [8–10], based on a model test and theoretical derivations, established the relationship between the intrinsic vibration frequency, bond area, elastic modulus, and rock mass. Jia et al. [11,12] analyzed mechanical parameters, such as bond force and friction force of unstable rock, under natural frequency and verified the effectiveness of dynamic monitoring in practical engineering.

However, vibration monitoring equipment is highly susceptible to environmental noise in engineering and equipment installation. External factors interfere with the vibration signals of rocks, and vibration signals are often covered by noise, being difficult to analyze [13,14]. Furthermore, in traditional dynamic signal analysis, the vibration signal of the unstable rock mass has been analyzed by extracting characteristic parameters, such as the maximum, minimum, average, variance, root mean square, peak, and natural frequency. While these may represent a portion of signature features, they require significant time and effort, as well as a great deal of prior knowledge. The accuracy of identifying unstable rocks can be severely affected by the validity of feature extraction [15,16]. For this reason, machine learning methods do not inherently eliminate noise or interference; instead, they learn robust features that remain distinguishable under noisy conditions. Thus, while noise can still affect prediction accuracy, the proposed CNN–CWT framework improves resilience compared with purely empirical approaches.

In recent years, machine learning has attracted extensive interest in geotechnical engineering [17–19]. Among them, a convolutional neural network (CNN) has been used to reveal the vibration signal feature extraction and classification of structural damage due to its ability to integrate feature extraction and classification into a framework [20–22]. Zhu et al. [23] combined phase-based motion estimation (PME) with CNNs, greatly facilitating CNN applications. Zhang et al. [24] presented a classification framework to categorize rock blocks based on the principles of block theory, namely, key blocks, trapped blocks, and stable blocks. Zhang et al. [25] established a CNN-based microseismic detection network (CNN-MDN) model, which performs well on microseismic detection through training, validation, and testing. Yin et al. [26] proposed an integrated CNN-Adam-BO algorithm based on microseismic monitoring data to achieve real-time prediction of rockburst intensity. The algorithm was compared with the continuous wavelet transform (CWT) and cross wavelet transform, verifying the superiority of the dimension enhancement method of the microseismic sequence described. These advances highlight the feasibility of employing CNNs for instability recognition in complex geotechnical environments, forming the foundation of the present study.

In the present research, three stages of the unstable rock process, namely *stabilization stage* (stage 1), *separation stage* (stage 2), and *failure stage* (stage 3) are defined based on previous studies. The CNN model is then used to train and identify the microseismic signals in these three phases using CWT

analysis. In this study, the term “real-time” refers to the near-real-time processing capability of the proposed framework, in which vibration data are classified every 1 s and aggregated over a 100-s moving window. Considering the computation time (less than 3 s for data segmentation, transformation, and inference), the total latency from data acquisition to risk output is within ~10 s, which satisfies engineering requirements for on-site monitoring. Based on the identification results, a CNN-based risk-grade evaluation method for unstable rock collapse is proposed to achieve real-time assessment of unstable rock.

2. Materials and methods

2.1. Continuous wavelet transform

The CWT is a modern analysis method that builds upon the idea of a localized short-time Fourier Transform (SFT). It addresses the limitation of constant window size in SFT by providing a "time-frequency" window that varies with frequency, making it a powerful tool for time-frequency analysis. The CWT highlights specific features of the signal, allows for local analysis of time and frequency, and allows for progressive refinement of the signal on multiple scales. This leads to high-frequency time division and low-frequency frequency division, adapting to the demands of time-frequency signal analysis.

Traditionally, the short-time Fourier Transform (FFT) has been used to analyze unstable rock vibrations, but it lacked sensitivity in the time dimension and couldn’t provide a comprehensive frequency domain analysis. Rock vibration during progressive failure is highly nonstationary, with transient bursts and evolving dominant frequencies. The Morlet wavelet offers a favorable balance between time localization and frequency resolution, enabling clear visualization of energy ridges that are indicative of microcrack initiation and coalescence. In our case, the scale range was selected to cover the dominant frequency band (≈50–5000 Hz) observed in preliminary spectra. The scale–frequency mapping follows $f_a = f_c/(a\Delta t)$, where f_c is the wavelet center frequency, and Δt is the sampling interval. Lightweight checks showed that varying the scale range by ±20% changed the final classification accuracy by <2%, indicating robustness. This is why this work uses the Morlet CWT, which employs a wavelet basis function (Eq 1) to achieve higher temporal resolution compared to FFT (Eq 2). Figure 1 presents a diagram of the wavelet analysis results.

$$\varphi_{u,v}(t) = \exp(i\omega_0 \frac{(t-v)}{u} \exp(-\frac{(t-v)^2}{2u^2}) \tag{1}$$

$$\Psi f(u,v) = \int_{-\infty}^{+\infty} f(t)\varphi_{u,v}(t)dt \tag{2}$$

where u is the scale parameter, v is the panning parameter, and t is the time step.

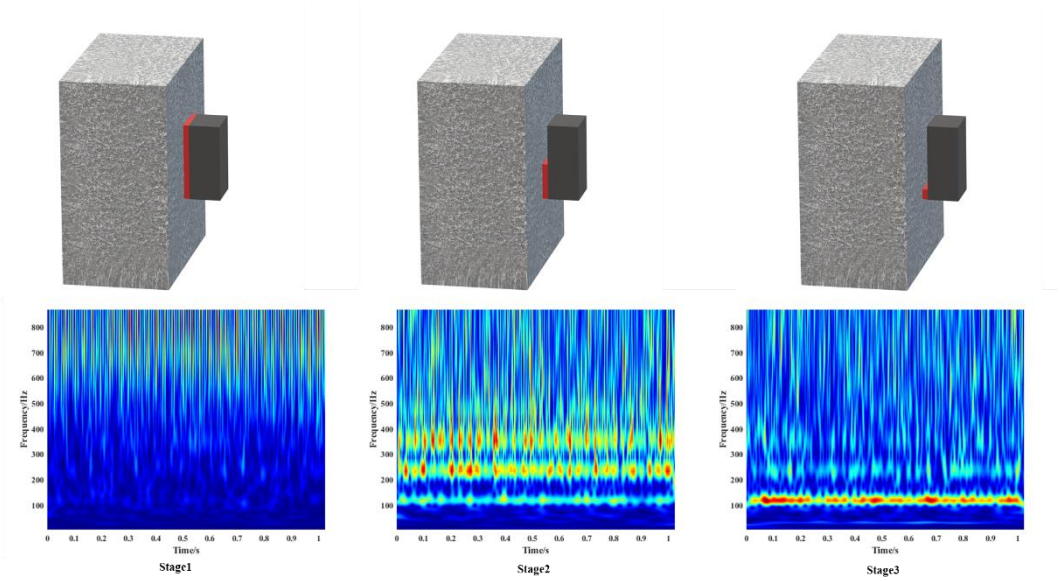

Figure 1. Continuous wavelet analysis diagram of vibration signal.

2.2. Continuous wavelet transform

The analysis of CWT parameters and characteristics is challenging when using the naked eye because the vibration signal is impacted by several factors. However, recent research shows that machine learning, specifically deep learning neural networks, can effectively solve this recognition problem. ANNs, including FCANNs, are a critical branch of AI and have been commonly used in geotechnical engineering. However, FCANNs show three limitations when processing large images: losing spatial information during vector transformation, more parameters causing training difficulties and efficiency issues, and overfitting due to a high number of parameters.

To overcome these limitations, CNNs were introduced. They improve the situation by using a local receptive field and weight sharing to reduce hyperparameters and avoid overfitting. The basic architecture of a CNN includes an input layer, a convolutional layer, an activation function, a pooling layer, a fully connected layer, and an output layer. The layers can be stacked to form a complete CNN, as shown in Figure 2.

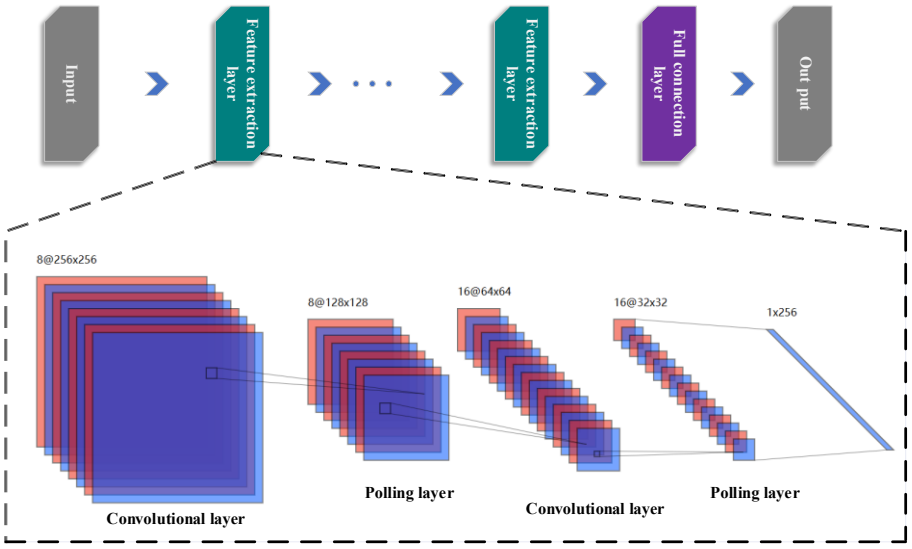

Figure 2. Convolutional neural network structure diagram.

2.2.1. Convolutional layer

The convolutional layer is tasked with feature extraction within images. The layer employs multiple convolution kernels, referred to as filters. These filters are similar to sliding windows, which traverse the numerical matrix in steps, performing convolution operations. The results of these operations are then input into the activation function to produce the image's feature map. The feature extraction process utilizes Eq 3, and the calculation of the convolutional layer is illustrated in Figure 3.

$$a_j^l = f\left(\sum_{i \in M_j^l} a_i^{l-1} \square k_{ij}^l\right) \tag{3}$$

where a is the feature value extracted from the l-th convolutional layer, k_{ij}^l is the connection weight between features in the filter and convolutional kernel size, M_j^l is the input feature, and f is the activation function. To extract depth features, CNNs usually contain multiple convolutional layers.

2.2.2. Pooling layer

The pooling layers follow the convolutional layers, with each convolutional layer having a corresponding pooling layer. The purpose of the pooling layer is to reduce the dimensionality of the feature map and eliminate redundant information, using the formula in Eq 4. Commonly used pooling functions include mean-pooling, max-pooling, min-pooling, and stochastic-pooling. Mean-pooling calculates the average of all feature points, while max-pooling takes the maximum. Stochastic-pooling assigns probabilities to pixels based on their numerical values, subsamples them accordingly, and approximates mean-pooling on average, while locally following the criterion of max-pooling.

$$a_j^l = f\left(b_j^i + \beta_j^l down(a_j^{l-1}, M^l)\right) \tag{4}$$

where b is the bias, β is the multiplicative residual, M is the size of the pooling frame used in the lth layer, and $down(.)$ is the pooling function.

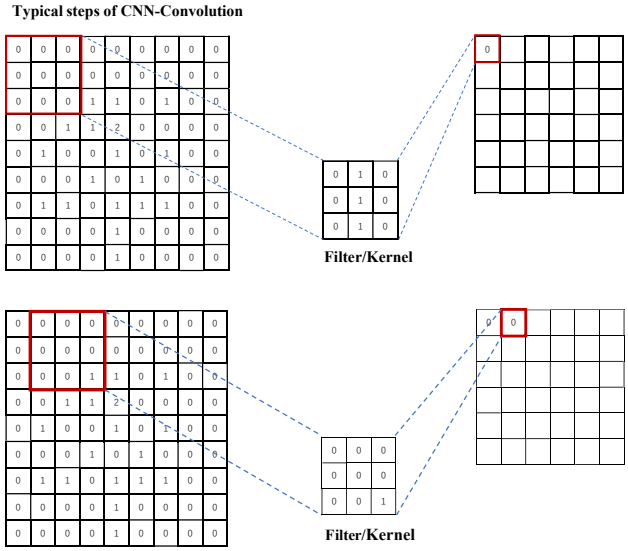

Figure 3. Convolutional neural network convolutional layer schematic.

2.2.3. Fully connected layer

Following multiple convolutional and pooling layers, one or more fully connected layers are connected, where each neuron is fully connected to all neurons in the previous layer. These layers integrate local information from the convolutional or pooling layers for category differentiation. To enhance the performance of CNN networks, the activation function for each neuron in the fully connected layer is commonly implemented as a ReLU function.

2.2.4. CNN architecture

The CNN architecture was designed considering both model complexity and computational efficiency. The number of convolution and pooling layers was determined through a small-scale parameter tuning process to balance feature abstraction and overfitting risk. Specifically, a three-layer convolutional structure achieved optimal performance during preliminary validation, while deeper networks did not significantly improve accuracy but increased training time. The kernel size of 3×3 was adopted to capture localized patterns in the wavelet spectrograms, which has been widely proven effective for time–frequency images. The hyperparameters, such as learning rate (0.001) and batch size (32), were selected based on a grid search within a limited range to ensure stable convergence.

2.3. Unstable rock risk assessment model

According to the latest research, the entire process of rock mass failure can be categorized into three stages based on the dynamic change process of slip resistance, which are the stabilization stage, separation stage, and failure stage.

Stage 1, known as the stabilization stage: The rock mass is bonded firmly to the bedrock and the

slip resistance is provided by the cohesive force in the potential slip plane. As a result, there is no tendency to slide, and the rock mass in this stage is referred to as *stable rock*, rather than *unstable*.

Stage 2, known as the separation stage: The bond strength of the rock mass decreases, causing it to separate from the bedrock when the cohesive force is insufficient to resist the sliding force. The rock body in this stage is still stable but has a tendency to slide, making it defined as *unstable rock*.

Stage 3, known as the damage stage: The rock body is further disturbed by damage, causing the potential slip plane to be damaged by penetration, resulting in instantaneous collapse and slumping.

In this study, we set the classification results of CNN as three stages of unstable rocks, which are collectively referred to as stage 1, stage 2, and stage 3, and then input the analysis results into CNN for the classification model training after performing CWT on the micro-oscillation data of each stage. After several simulations and calculations, it was found that intercepting the micro-vibration data with a time duration of 1 s for CWT can ensure a sufficient amount of analysis samples and maximize the recognition accuracy of the CNN. To shorten the model's learning time, the wavelet coefficient map is compressed to 256 × 256 pixels. As shown in Table 1, the image resolution of 256 × 256 pixels was selected to balance computational efficiency and feature fidelity. Table 1 compares the classification performance and processing time between 256 × 256 and 512 × 512 inputs. The accuracy difference was less than 1.3%, whereas computation time per sample decreased by 67%. Hence, the adopted resolution ensures adequate preservation of high-frequency information while enabling efficient real-time implementation. 80% of the data was selected as the training set, and 20% was selected as the validation set. The CNN classification recognition model for unstable rocks was then trained.

Table 1. Comparison of CNN training data parameters.

Pixels	Accuracy	Computation time (s)
512 × 512	98.9%	9668
256 × 256	97.6%	5768
128 × 128	72.9%	3265

Although a fully trained CNN can achieve high accuracy in recognizing vibration data with a time length of 1 s, the accurate identification by the CNN model is affected by various factors, such as environmental noise and equipment errors. As such, it cannot identify all data with 100% accuracy. To address this, the authors proposed targeted algorithms to enhance the CNN recognition results' interference resistance. The trained model classifies and identifies vibration data types of the target rock over a specific time period (e.g., 100 s), computes the ratio of the three states in the total recognition data volume, and generates the risk evaluation weight proportion of the unstable rock in the current time period. The risk-level evaluation of unstable rocks is then determined based on the risk evaluation weight ratio obtained from the classification identification model, as described in Eq 5. An example of the specific algorithm is presented in Figure 4.

$$RL = 1 \cdot K_1 + 2.5 \cdot K_2 + 3.5 \cdot K_3 \tag{5}$$

where K_1, K_2, K_3 are the percentages of CNN model classification recognition in the stable stage, development stage, and accelerated destruction stage of unstable rocks, respectively.

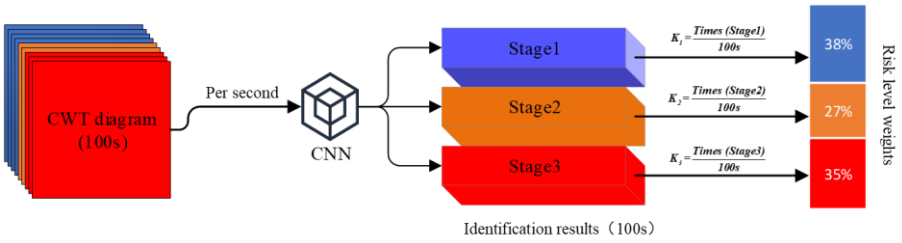

Figure 4. Risk-level weight calculation process.

3. Results

Considering the limited size and class distribution of the CNN training data of CNN neural networks, two experiments were conducted: a fixed-base test and a freeze-thaw dynamic validation test, using rock blocks of identical material, size, and mass. The fixed-base test was used to gather the training dataset of the CNN classification recognition model, and the vibration data from the freeze-thaw test was utilized to initially assess the accuracy of the model. The process of experimental validation is shown in Figure 5.

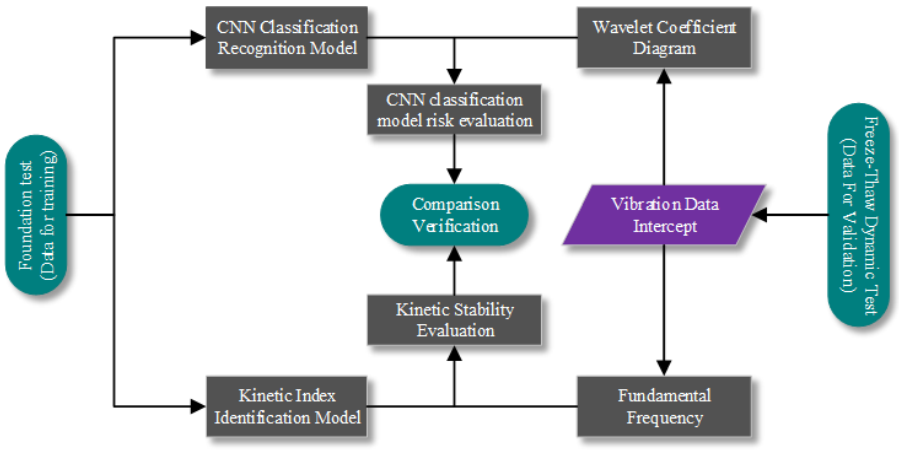

Figure 5. Experimental validation technology roadmap.

3.1. Fixed-base test

In this study, a 12 cm × 3 cm × 3 cm solid block of unstable rock with the potential to collapse and fall was used, along with a 30 cm × 30 cm × 30 cm marble block as the bedrock. A binder was applied to the marble floor to emphasize the role of the bedrock. During the fixed-base tests, tri-axial accelerometers (sensitivity 100 mV/g) were mounted on the top and lateral faces of the rock specimens using epoxy adhesive. The sampling frequency was 10 kHz to ensure adequate time resolution. All tests were conducted in a vibration-isolated environment with ambient noise levels below 0.03 g. The data acquisition system was synchronized across all sensors to prevent timing errors. These measures ensured the reliability and reproducibility of the recorded vibration signals.

The experiment provided three stages of unstable rock for the CNN model, including a “Stage1” fully bonded to the bedrock, a “Stage2” with 50%–40% bonded area, and a “Stage3” using a binder to keep the unstable rock in the ultimate stress stage. Test photos are shown in Figure 6. Vibration data

of the test block was measured using an accelerometer with a sampling frequency of 2000 Hz, and five sets of measurements were performed for each corresponding stage of the unstable rock, each with a measurement time of 600 s. No noise reduction measures were taken during the measurements. Data were intercepted every 2000 data points (i.e., 1 s) after measurements to create a training set of the CNN model. The CWT of the intercepted vibration data was used as the training set, and the training dataset consisted of 500 CWT analysis maps for each corresponding stage.

After several training adjustments, the CNN model with a convolutional structure of 9-9-7-3, Lr = 0.0002, and Epoch = 30 achieved the highest recognition accuracy and the lowest loss function. The architecture (kernel configuration 9-9-7-3) was determined following three practical design principles:

(1) Feature-scale matching: The receptive field of the first two convolutional layers was designed to cover approximately 15%–25% of the CWT image width, sufficient to capture local high-frequency patterns while preserving global continuity.

(2) Model capacity control: According to LeCun et al. [27] and Krizhevsky et al. [28], increasing depth beyond four convolutional layers yields diminishing returns for mid-sized datasets. Hence, a moderate depth was chosen to avoid over-parameterization.

(3) Engineering feasibility: The adopted network can execute inference on an embedded GPU (NVIDIA Jetson TX2) within 3 s per sample, satisfying real-time field requirements.

Empirical tuning confirmed that deeper or wider variants improved accuracy by <1% but doubled computation cost, validating the efficiency of the final configuration. The accuracy variation and loss function are displayed in Figure 7. The recognition accuracy of the CNN model reached 97.6%, and the loss value was 0.041 after sufficient training. Besides accuracy, additional metrics, including precision, recall, and F1-score, were calculated to provide a more comprehensive evaluation of the model's performance. The CNN–CWT model achieved an average precision of 96.2% and recall of 95.6% across all classes, with an overall F1-score of 95.8%. These results, consistent with the confusion matrix analysis, indicate that the proposed method achieves balanced performance between false alarms and missed detections.

(a) Stage1 (b) Stage2 (c) Stage3

Figure 6. Fixed-base test photos.

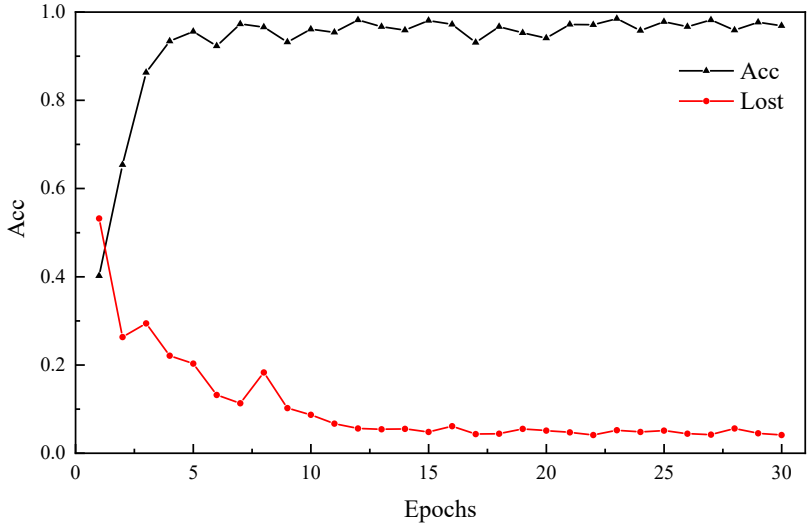

Figure 7. CNN training results.

3.2. Freeze-thaw dynamic validation test

The dynamic validation test used the same test block as the foundation test, using a low-temperature freezing bonding medium to firmly bond the unstable rock to the bedrock. The melting of the low-temperature adhesive was visually monitored and recorded by a high-definition camera. The adhesive was applied in a uniform layer of 2 mm, and its temperature was controlled using a thermostat to ensure homogeneous heating. The recorded images confirmed a consistent melting front within ±5 s across the bonding area, minimizing irregular vibration pulses during the transition period. The unstable rock in the test can undergo the complete stability-to-collapse process at room temperature, as simulated by melting the low-temperature bonding medium, as depicted in Figure 8. The adhesive failure mode differs from micro-crack coalescence in natural rock, which may introduce vibration modes not fully representative of in situ behavior. However, this approach allows repeatable laboratory control of boundary weakening and stress release. The potential discrepancy was mitigated by verifying that the dominant frequency evolution pattern remained consistent with reported field microseismic data. In the experiments, the vibration data of the blocks was measured using accelerometers with the same experimental parameters as in the foundation tests, and it was found that the unstable rock took approximately 2043 s to go from stable to crumbling and falling after several tests were repeated.

(a) Before collapse **(a) After collapse**

Figure 8. Dynamic freeze-thaw test photos.

3.3. Risk assessment validation

In the freeze-thaw test, the vibration data obtained from the CWT analysis were divided into 21 time periods of every 100 s from two randomly selected cases. The CNN model trained in the fixed-base test was applied to identify the CWT analysis maps in each time period, and the risk-level weights of the blocks in the freeze-thaw test were calculated using the method proposed in this paper in Figure 4.

The risk-level eigenvalue thresholds listed in Table 2 were established by cross-validating the results from the fixed-base and freeze–thaw dynamic tests. In the fixed-base tests, the mechanical properties corresponding to three stability states were repeatedly measured, and the safety factor ranges of each state were calculated using the limit equilibrium method. In parallel, dynamic safety factors were estimated from the natural frequency evolution following the procedure of Du et al. [10]. Based on these two safety assessments, the freeze–thaw experiment was divided into three stages, consistent with the calibrated safety factor ranges: Stage 1 (0–1000 s), corresponding to the stable phase, Stage 2 (1000–1800 s), representing the transitional deformation phase, and Stage 3 (>1800 s), corresponding to the progressive failure phase, as presented in Figure 9. The CNN classification results within these time intervals were then averaged to obtain the characteristic eigenvalues for each risk level. This approach ensures that the thresholds in Table 1 are physically grounded, experimentally validated, and statistically consistent between static and dynamic evaluations.

The variation of the CNN risk-evaluation eigenvalues and safety coefficients over time was demonstrated in Figure 10. After several calculations, it was determined that all six cases in the test were in “Stage 1” during 0–1000 seconds, and the test block was in “Stage 2” between 1000 and 1800 seconds and entered “Stage 3” after 1800 seconds.

Table 2. Unstable rock identification criteria.

Calculation indicators	Stage 1	Stage 2	Stage 3
Risk-level eigenvalues	<1.75	[1.75, 2.5)	≥2.5
Safety factor	>2.0	[2.0, 1.2)	≤1.2

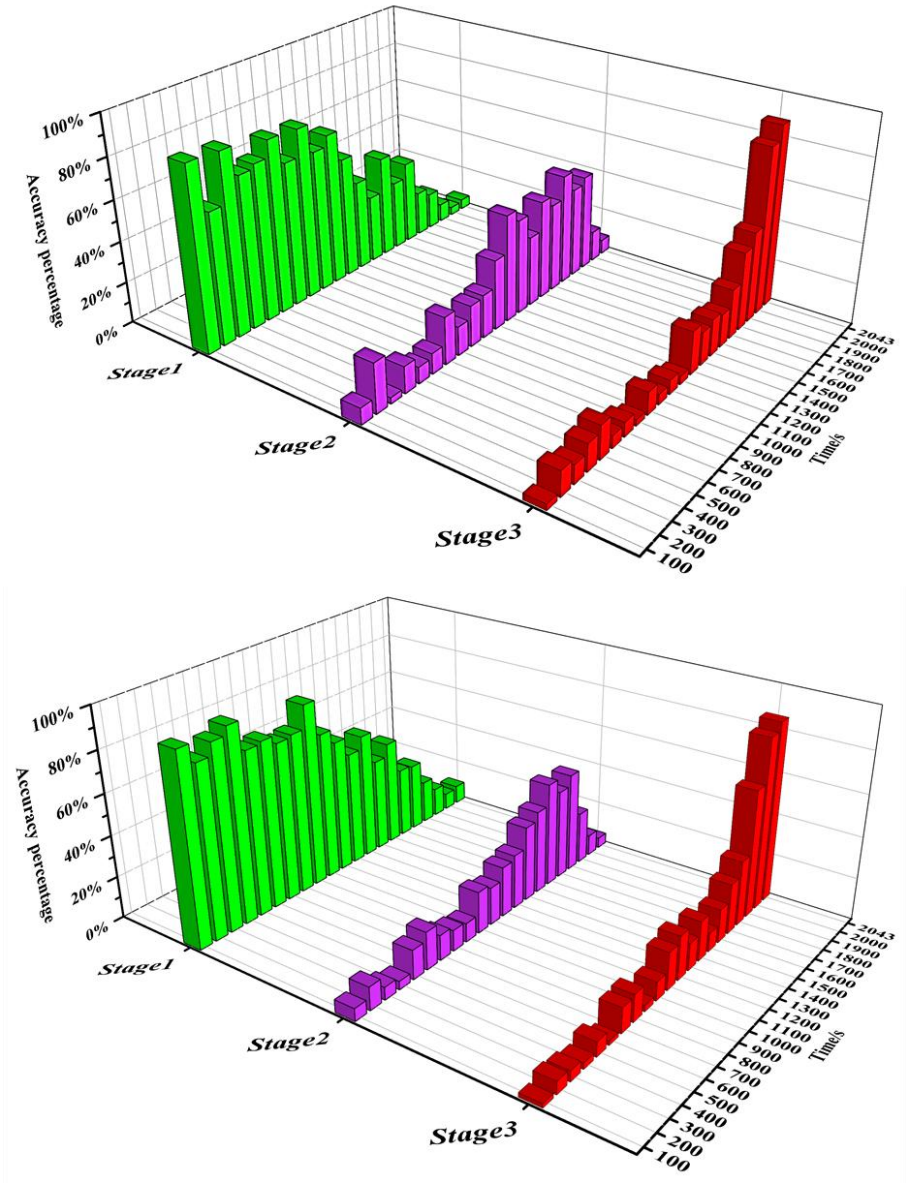

Figure 9. Risk-level weighting ratio based on CNN calculation.

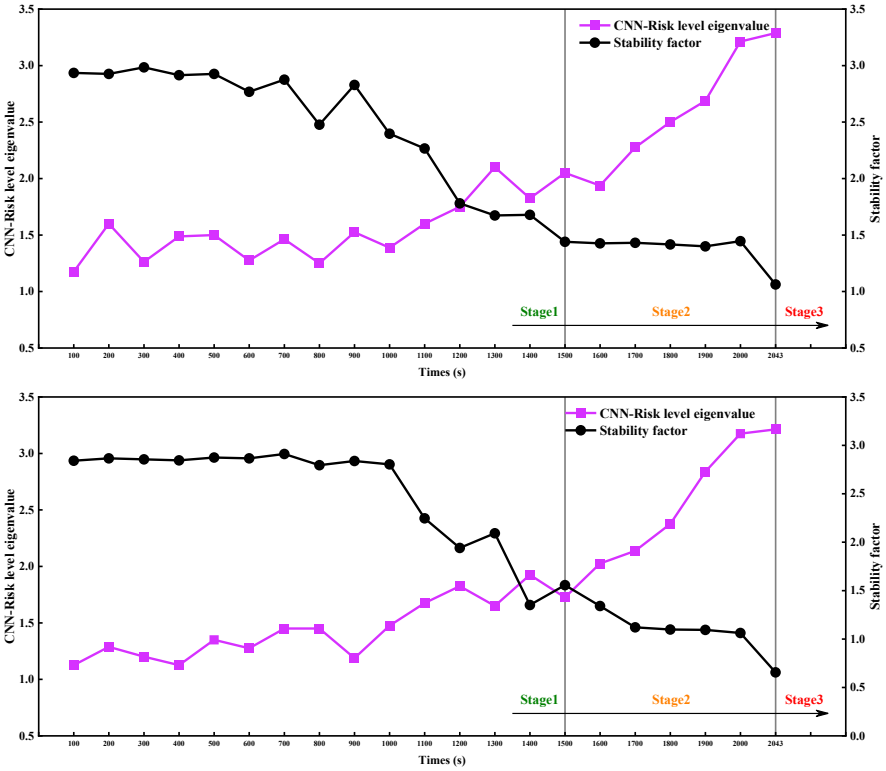

Figure 10. Risk-level eigenvalues and safety factor.

4. Discussion

4.1. Validation of CNN risk-level eigenvalues

Research has revealed that performing CWT-CNN analysis on vibration data at a rate of once per second ensures both the maximum training data volume and recognition accuracy for the CNN model. However, since CNN struggles to achieve 100% recognition accuracy, this analytical approach significantly distorts the calculation of risk-level assessments. Therefore, we have refined the CNN recognition results by grouping and extracting segments from the identified outcomes. Identification results were grouped and intercepted, with the percentage of identifications in each group calculated to enhance the robustness of the CNN risk-level identification method. The risk-level calculation results obtained by grouping data into 1, 10, and 50 s segments are shown in Figure 11. Through multiple trials, it was found that using 100 s segments for recognition results yields the optimal risk-level feature values when calculating risk-level weights.

Previous studies have demonstrated the efficacy of using the kinetic index (fundamental frequency) for the safety evaluation of unstable rocks [29,30]. To further verify the effectiveness of the CNN risk-level evaluation method, a correlation analysis was performed between the CNN results and the safety coefficient in different window lengths (50, 100, 150 s). Fitted curves and analysis results are presented in Figure 12. The 100 s window achieved the best compromise, as the correlation between aggregated risk level and safety factor peaked at R2 = 0.89. Shorter windows produced noisier curves, whereas longer windows delayed the response. Therefore, 100 s was adopted as the optimal

aggregation length. Furthermore, the risk-level values increased significantly as the safety coefficients decreased, demonstrating the effectiveness of the CNN risk-level evaluation method in identifying unstable rock collapse behavior. However, it should be noted that the safety factor derived from fundamental frequency serves as an indirect stability indicator rather than an absolute benchmark. Both the proposed risk eigenvalue and the frequency-based safety factor may share common dependencies on material stiffness and boundary conditions, which could partially explain their strong correlation. This limitation will be addressed in future work through multi-parameter cross-validation.

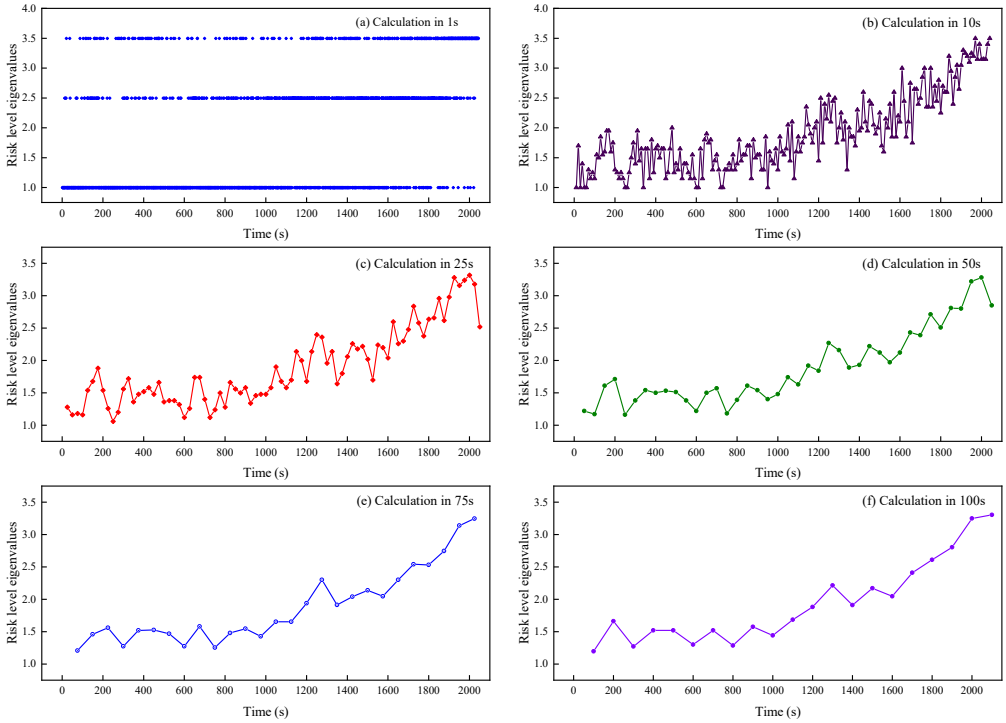

Figure 11. Risk-level characteristics after group-based screening.

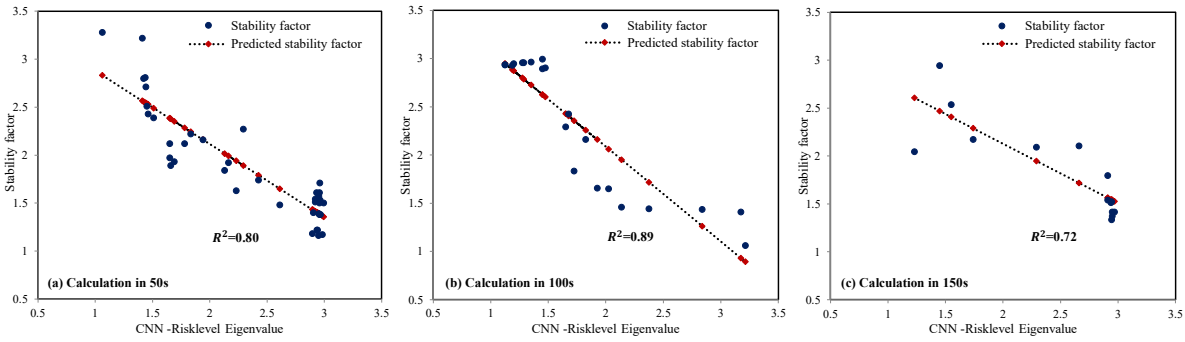

Figure 12. Correlation analysis of CNN risk-level eigenvalues and safety factor.

4.2. Comparative analysis of calculated results

A brief comparison was conducted to evaluate the influence of signal preprocessing. When applying a 100–5000 Hz bandpass filter prior to CWT transformation, the overall classification

accuracy increased marginally from 96.9% to 98.7%, as depicted in Figure 13. This indicates that the CNN–CWT model inherently learns noise-resistant features, and that basic filtering can slightly enhance but is not essential for performance.

According to the test results and risk-level calculations presented in Figure 10, it is evident that the CNN risk-level characteristic values exhibit distinct variation characteristics during the collapse of the test block. In the initial stage of the experiment (1–900s), the risk-level characteristic values were observed to fluctuate without exceeding 1.5, indicating that the test block remained stable for an extended period without any collapse precursors. The safety factor values were also observed to be around 3.0 during this period, leading to the same conclusion. However, between 1200 and 1700 s of the test, the risk-level characteristic values showed a significant increase and fluctuation, ranging from 1.3 to 2.0. The melting of the temperature-sensitive binder at room temperature led to uneven melting. This caused transient separation and locking effects, resulting in slight accelerometer disturbances. To quantitatively verify this explanation, two relative spectral indicators were calculated for the stable (1000–1200 s) and fluctuation (1200–1700 s) intervals, as shown in Table 3. The relative spectral energy variation (RSEV) describes the temporal fluctuation of spectral energy, while the normalized dominant frequency shift (NDFS) reflects transient deviations of the dominant frequency. Both indicators increased by more than twofold during the fluctuation period, confirming that the observed irregularities originated from transient frequency-domain disturbances rather than structural degradation. Despite this, the safety factor recognition rate remained higher than the CNN risk-level eigenvalue, which can be attributed to the sufficient elimination of noise in the safety factor calculations. Toward the end of the test (1800–2043 s), the risk level calculation results exhibited a clear upward trend, rapidly increasing from about 2.5 to more than 3.0. All CNN risk-level eigenvalues were greater than 3 before the block collapsed and fell, and compared with the safety factor, the CNN risk-level eigenvalues continued to increase in the 200 s leading up to the collapse, with obvious collapse precursors.

Table 3. Smoothness parameters of vibration signals in different segments.

Time interval	Relative spectral energy variation	Normalized dominant frequency shift
1000–1200 (s)	4.8%	2.1
1200–1700 (s)	10.2%	4.3

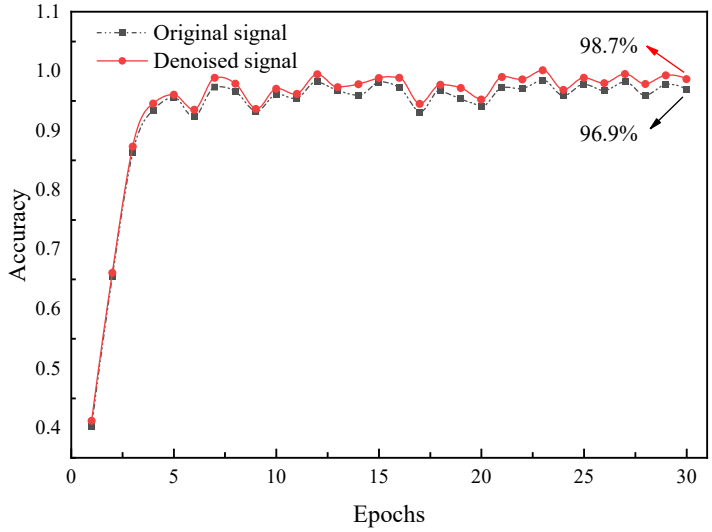

Figure 13. Comparison of trained data with and without noise reduction.

4.3. Engineering application of the CNN risk-level evaluation method

The application of machine learning techniques, specifically neural networks, in geotechnical engineering has garnered considerable attention due to its potential for enhancing the efficiency and accuracy of data analysis and prediction [31,32]. Conventional methods, which primarily rely on manual calculation and expert knowledge, are often labor-intensive and prone to error. On the other hand, neural networks can effectively extract features from data, reducing the dependence on human experience and minimizing the risk of error [33–35]. Furthermore, the vibration data used in this study were not preprocessed to eliminate noise, while kinetic index analysis underwent three stages of data screening, noise elimination, and human-empirical processing [36–38]. The results of this study indicated that the CNN risk-identification method has high anti-interference capability, accurately identifying the relationship between input and output variables even in complex and noisy environments. The accuracy of neural networks is also an important advantage. With the development of hardware and the increase in data, training neural networks has become more and more effective [39–41]. The ability of neural networks to learn from large amounts of data has improved predictive accuracy. In particular, deep neural networks have higher accuracy than traditional methods in many engineering applications [42,43]. The results from experiments and analysis indicate that, with adequate training, the proposed CNN identification model can precisely identify unstable rocks of consistent size, mass, and damage type and efficiently and accurately evaluate the changes and states of unstable rocks through cross-validation with the corresponding CNN risk-level evaluation method and kinetic index evaluation method.

5. Conclusions

In this study, a CNN classification-based identification method was proposed for the assessment of the risk level of unstable rocks using continuous wavelet analysis. The vibration signals from the fixed-base test were utilized as the training set for the CNN, and the vibration signals from the freeze-thaw test were used for validation. The robustness and accuracy of the model were improved through

further stratified sampling. The main findings of this study are summarized as follows:

(1) The CNN was trained with 3000 s of vibration data for each corresponding risk level, obtained from the fixed-base test. After several adjustments, the optimal parameters for the CNN in this study were a convolutional kernel structure of 9-9-7-3, a learning rate of 0.0002, and 30 epochs. The accuracy of the model was 97.6% and the loss value was 0.041 after sufficient training.

(2) The validation accuracy of the CNN models was above 84% after verification with vibration data from the freeze-thaw tests. The validation results were stratified, and feature percentages were extracted separately to obtain the weights of unstable rock risk levels. The results showed that the risk-level eigenvalues were approximately 1.2–1.5 when the test block was stable, and about 3.3–3.5 when it was close to collapse. The correlation results showed that the CNN risk-level eigenvalues were highly negatively correlated with the stability coefficient, with a correlation coefficient no lower than 0.8. This CWT-based CNN risk-level assessment model can effectively reflect the whole process from stability to collapse of unstable rocks with consistent mass, size, and damage type.

(3) The CNN risk-level evaluation method can be combined with existing unstable rock stability assessment methods to achieve a comprehensive evaluation of safety coefficients and risk levels for different types of unstable rocks. It is expected to play a positive role in slope engineering safety monitoring and risk management.

Despite the promising results, the present study has limitations. The model was trained on laboratory-scale specimens with homogeneous lithology, and its generalization to different rock types, structural scales, and field conditions remains to be validated. Future work will focus on integrating multi-sensor data (acoustic, displacement, and stress) and employing transfer-learning techniques to enhance adaptability under complex in situ environments.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The National Key R&D Program of China (2023YFC3081400), the National Natural Science Foundation of China (41702371), the Open Fund Project of State Key Laboratory of Mining Response and Disaster Prevention in Deep Coal Mines (SKLMRDPC22KF13), the Open Fund Research Project Supported by State Key Laboratory of Strata Intelligent Control and Green Mining Co-founded by Shandong Province and the Ministry of Science and Technology (SICGM202503), the Supported by the Fund of Chongqing Key Laboratory of Facility Damage Mechanism and Protection in Highland Mountain Environment (LQ24KFJJ09), the Supported by State key Laboratory of Mining Disaster Prevention and Control (Shandong University of Science and Technology), Ministry of Education (DPEPM202502).

Conflict of interest

The authors declare no conflicts of interest, financial or otherwise, that could influence or bias this work.

References

1. Vick L, Bohme M, Rouyet L, et al. (2020) Structurally controlled rock slope deformation in northern Norway. *Landslides* 17: 1745–1776. https://doi.org/10.1007/s10346-020-01421-7
2. Bolla A, Paronuzzi P (2020) Geomechanical Field Survey to Identify an Unstable Rock Slope: The Passo della Morte Case History (NE Italy). *Rock Mech Rock Eng* 53: 1521–1544. https://doi.org/10.1007/s00603-019-01963-w
3. Bolla A, Paronuzzi P (2020) Numerical Investigation of the Pre-collapse Behavior and Internal Damage of an Unstable Rock Slope. *Rock Mech Rock Eng* 53: 2279–2300. https://doi.org/10.1007/s00603-019-02031-z
4. Du Y, Xie MW, Jiang Y, et al. (2015) Methods for determining early warning indices Based on natural frequency monitoring. *Rock Soil Mech* 36: 2284–2290. https://doi.org/10.16285/j.rsm.2015.08.022L
5. Du Y, Xie MW, Jiang Y, et al. (2019) Research progress on dynamic monitoring index for early warning of rock collapse. *Chin J Eng* 41: 427–435. https://doi.org/10.13374/j.issn2095-9389.2019.04.002
6. Ma G. (2012) Study on Evaluating Rock Block Stability by Using A Remotely Positioned Laser Doppler Vibrometer. *Int J Geomate* 2: 247–252. https://doi.org/10.21660/2012.4.3j
7. Ma G, Sawada K, Yashima A, et al. (2014) Experimental Study of the Applicability of the Remotely Positioned Laser Doppler Vibrometer to Rock-Block Stability Assessment. *Rock Mech Rock Eng* 48: 787–802. https://doi.org/10.1007/s00603-014-0577-x
8. Du Y, Xie MW (2022) Indirect method for the quantitative identification of unstable rock. *Nat Hazards* 112: 1005–1012. https://doi.org/10.1007/s11069-021-05197-4
9. Du Y, Wu Z, Xie MW, et al. (2018) Early-warning method of rock collapse and its experimental verification. *J China Coal Soc* 44: 3069–3075. https://doi.org/10.13225/j.cnki.jccs.2018.1467
10. Du Y, Xie MW, Jiang YJ, et al. (2017) Experimental Rock Stability Assessment Using the Frozen-Thawing Test. *Rock Mech Rock Eng* 50: 1049–1053. https://doi.org/10.1007/s00603-016-1138-2
11. Jia YC, Song GH, Jiang T et al. (2022) Buried depth calculation of the slope of the unstable rock based on natural vibration frequency. *Front Earth Sci* 10. https://doi.org/10.3389/feart.2022.929825
12. Jia B, Wu Z, Du Y (2019) Real-time stability assessment of unstable rocks based on fundamental natural frequency. *Int J Rock Mech Min Sci* 124: 104–134. https://doi.org/10.1016/j.ijrmms.2019.104134
13. Cao JH, Wei ZB (2011) Denoising in Steel Structural Impulsive Vibration Signal Based on Independent Component Analysis. *Adv Mat Res* 219–220: 1337–1341. https://doi.org/10.4028/www.scientific.net/AMR.219-220.1337
14. Fan G, Li J, Hao H. (2020) Vibration signal denoising for structural health monitoring by residual convolutional neural networks. *Measurement* 157. https://doi.org/10.1016/j.measurement.2020.107651
15. Huo L, Du Y, Xie M, et al. (2021) Unstable rock mass identification method based on multi-level dynamic parameters. *Chin J Rock Mech Eng* 41: 3124–3131. https://doi.org/10.13722/j.cnki.jrme.2021.1058
16. Huo L, Du Y, Xie M, et al. (2021) Chinese Journal of Rock Mechanics and Engineering. *Chin J Rock Mech Eng* 40: 3156–3162. https://doi.org/10.13722/j.cnki.jrme.2021.0337

17. Jooshaki M, Nad A, Michaux S. (2021) A Systematic Review on the Application of Machine Learning in Exploiting Mineralogical Data in Mining and Mineral Industry. *Minerals* 11: https://doi.org/10.3390/min11080816
18. Kang J, Ullah Z, Gwak J (2021) MRI-Based Brain Tumor Classification Using Ensemble of Deep Features and Machine Learning Classifiers. *Sensors* 21: https://doi.org/10.3390/s21062222
19. Liao YY, Han L, Wang HY, et al. (2022) Prediction Models for Railway Track Geometry Degradation Using Machine Learning Methods: A Review. *Sensors* https://doi.org/10.3390/s22197275
20. Cheng HD, Chen HX, Li ZF, et al. (2020) Ensemble 1-D CNN diagnosis model for VRF system refrigerant charge faults under heating condition. *Energy Build* 224: 110256. https://doi.org/10.1016/j.enbuild.2020.110256
21. Oh BK, Park Y, Park HS (2022) Seismic response prediction method for building structures using convolutional neural network. *Struct Control Hlth* 27: https://doi.org/10.1002/stc.2519
22. Hu XD, Zhang PL, Zhang Q, et al. (2020) A novel framework of cnn integrated with adaboost for remote sensing scene classification. *IGARSS 2020 - 2020 Ieee International Geoscience and Remote Sensing Symposium* 2643–2646. https://doi.org/10.1109/IGARSS39084.2020.9324261.
23. Zhu HH, Azarafza M, Akgun H (2022) Deep learning-based key-block classification framework for discontinuous rock slopes. *J Rock Mech Geotech Eng* 14: 1131–1139. https://doi.org/10.1016/j.jrmge.2022.06.007
24. Zhang H, Ma C, Pazzi V, et al. (2020) Deep Convolutional Neural Network for Microseismic Signal Detection and Classification. *Pure Appl Geophys* 177: 5781–5797. https://doi.org/10.1007/s00024-020-02617-7
25. Zhang T, Shi D, Wang Z, et al. (2022) Vibration-based structural damage detection via phase-based motion estimation using convolutional neural networks. *Mech Syst Signal Process* 178: 109320. https://doi.org/10.1016/j.ymssp.2022.109320
26. Yin X, Liu Q, Huang X, et al. (2021) Real-time prediction of rockburst intensity using an integrated CNN-Adam-BO algorithm based on microseismic data and its engineering application. *Tunn Undergr Space Technol* 117: 104133. https://doi.org/10.1016/j.tust.2021.104133
27. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553): 436–444. doi: 10.1038/nature14539. PMID: 26017442.
28. Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton. (2017) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM* 60: 84–90.
29. Du Y, Huo L, Zhang H, et al. (2022) Experimental research on remote monitoring and early warning of rock collapse. *J China Univ Min Technol* 51: 1201–1208. https://doi.org/10.13247/j.cnki.jcumt.001427
30. Zhang X, Xie M, Zhang L, et al. (2022) Study on calculation model of stability coefficient of falling dangerous rock mass based on natural frequency. *Chin J Rock Mech Eng* 1–9. https://doi.org/10.13722/j.cnki.jrme.2022.0361
31. Zhang WG, Li HR, Li YQ, et al. (2021) Application of deep learning algorithms in geotechnical engineering: a short critical review. *Artif Intell Rev* 54: 5633–5673. https://doi.org/10.1007/s10462-021-09967-1
32. Lin H, Kumar Singh S, Xiang Z, et al. (2022) An investigation of machine learning techniques to estimate minimum horizontal stress magnitude from borehole breakout. *Int J Min Sci Technol* 32: 1021–1029. https://doi.org/10.1016/j.ijmst.2022.06.005

33. Bai F, Fan M, Yang H, et al. (2021) Fast recognition using convolutional neural network for the coal particle density range based on images captured under multiple light sources. *Int J Min Sci Technol* 31: 1053–1061. https://doi.org/10.1016/j.ijmst.2021.09.004
34. Zhang Y, Tang J, Cheng Y, et al. (2022) Prediction of landslide displacement with dynamic features using intelligent approaches. *Int J Min Sci Technol* 32: 539–549. https://doi.org/10.1016/j.ijmst.2022.02.004
35. Jong SC, Ong DEL, Oh E. (2021) State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. *Tunn Undergr Sp Tech* 113. https://doi.org/10.1016/j.tust.2021.103946
36. Zhong T, Cheng M, Lu SP, et al. (2022) RCEN: A Deep-Learning-Based Background Noise Suppression Method for DAS-VSP Records. *IEEE Geosci Remote Sens Lett* 19: 1–5. https://doi.org/10.1109/LGRS.2021.3127637
37. Tang HH, Cheng SJ, Li WQ, et al. (2022) Simultaneous reconstruction and denoising for DAS-VSP seismic data by RRU-net. *Front Earth Sci* 10. https://doi.org/10.3389/feart.2022.993465
38. Sadeqi A, Moradi S. (2021) A new SVD-based filtering technique for operational modal analysis in the presence of harmonic excitation and noise. *J Sound Vib* 510. https://doi.org/10.1016/j.jsv.2021.116252
39. Temeng VA, Ziggah YY, Arthur CK. (2020) A novel artificial intelligent model for predicting air overpressure using brain inspired emotional neural network. *Int J Min Sci Technol* 30: 683–689. https://doi.org/10.1016/j.ijmst.2020.05.020
40. Soofastaei A, Aminossadati SM, Arefi MM, et al. (2016) Development of a multi-layer perceptron artificial neural network model to determine haul trucks energy consumption. *Int J Min Sci Technol* 26: 285–293. https://doi.org/10.1016/j.ijmst.2015.12.015
41. Luo Z, Zuo H, Jia N, et al. (2013) Instability identification on large scale underground mined-out area in the metal mine based on the improved FRBFNN. *Int J Min Sci Technol* 23: 821–826. https://doi.org/10.1016/j.ijmst.2013.10.007
42. Shrestha A, Mahmood A. (2020) Review of Deep Learning Algorithms and Architectures. *IEEE Access* 7: 53040–53065.10.1109/ACCESS.2019.2912200
43. Monga V, Li YL, Eldar YC. (2020) Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing. *IEEE Signal Process Mag* 38: 18–44. https://doi.org/10.1109/MSP.2020.3016905

AIMS Press

© 2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)