



Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.jrmge.cn

Full Length Article

Real-time arrival picking of rock microfracture signals based on convolutional-recurrent neural network and its engineering application

Bing-Rui Chen^{a,b,*}, Xu Wang^{a,b,**}, Xinhao Zhu^{a,b}, Qing Wang^{a,b}, Houlin Xie^{a,b}

^a State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, 430071, China

^b University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Article history:

Received 30 January 2023

Received in revised form

4 June 2023

Accepted 9 July 2023

Available online 5 August 2023

Keywords:

Rock mass failure

Microseismic event

P-wave arrival

S-wave arrival

Deep learning

ABSTRACT

Accurately picking P- and S-wave arrivals of microseismic (MS) signals in real-time directly influences the early warning of rock mass failure. A common contradiction between accuracy and computation exists in the current arrival picking methods. Thus, a real-time arrival picking method of MS signals is constructed based on a convolutional-recurrent neural network (CRNN). This method fully utilizes the advantages of convolutional layers and gated recurrent units (GRU) in extracting short- and long-term features, in order to create a precise and lightweight arrival picking structure. Then, the synthetic signals with field noises are used to evaluate the hyperparameters of the CRNN model and obtain an optimal CRNN model. The actual operation on various devices indicates that compared with the U-Net method, the CRNN method achieves faster arrival picking with less performance consumption. An application of large underground caverns in the Yebatan hydropower station (YBT) project shows that compared with the short-term average/long-term average (STA/LTA), Akaike information criterion (AIC) and U-Net methods, the CRNN method has the highest accuracy within four sampling points, which is 87.44% for P-wave and 91.29% for S-wave, respectively. The sum of mean absolute errors (MAE_{SUM}) of the CRNN method is 4.22 sampling points, which is lower than that of the other methods. Among the four methods, the MS sources location calculated based on the CRNN method shows the best consistency with the actual failure, which occurs at the junction of the shaft and the second gallery. Thus, the proposed method can pick up P- and S-arrival accurately and rapidly, providing a reference for rock failure analysis and evaluation in engineering applications.

© 2024 Institute of Rock and Soil Mechanics, Chinese Academy of Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Microseismic (MS) monitoring is an effective monitoring and early warning technology for rock mass failure (Cook, 1976; Zhao et al., 2022; Li et al., 2023b), which has been widely used in deep rock engineering (Young and Collins, 2001; Durrheim et al., 2005; Feng et al., 2019; Li et al., 2023a; Mao et al., 2023). MS monitoring

analyzes the spatiotemporal evolution mechanism of rock microfractures and achieves rock failure prediction through waveform recognition, arrival picking, source positioning, and source parameter inversion in sequence (Xiao et al., 2016; Zhang et al., 2021a). Currently, the manual processing of MS data is usually considered to have the highest accuracy and the drawbacks of high cost and latency, and the automatic processing still has a certain gap in accuracy compared to manual processing. The MS P- and S-arrivals are the calculation bases of source positioning and source parameter inversion, and the deviation of the arrivals will be quickly amplified by these subsequent steps (Akram and Eaton, 2016). Thus, the accuracy and efficiency of MS arrival picking have a direct impact on the effect of rock failure monitoring. For this, various arrival-picking methods for MS P- and S-waves have been proposed, which are generally divided into traditional and machine learning methods.

The traditional methods mainly include the short-term average/long-term average (STA/LTA) method (Allen, 1978, 1982), the Akaike

* Corresponding author. State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, 430071, China.

** Corresponding author. State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, 430071, China.

E-mail addresses: brchen@whrsm.ac.cn (B.-R. Chen), wangxu19@mails.ucas.ac.cn (X. Wang).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

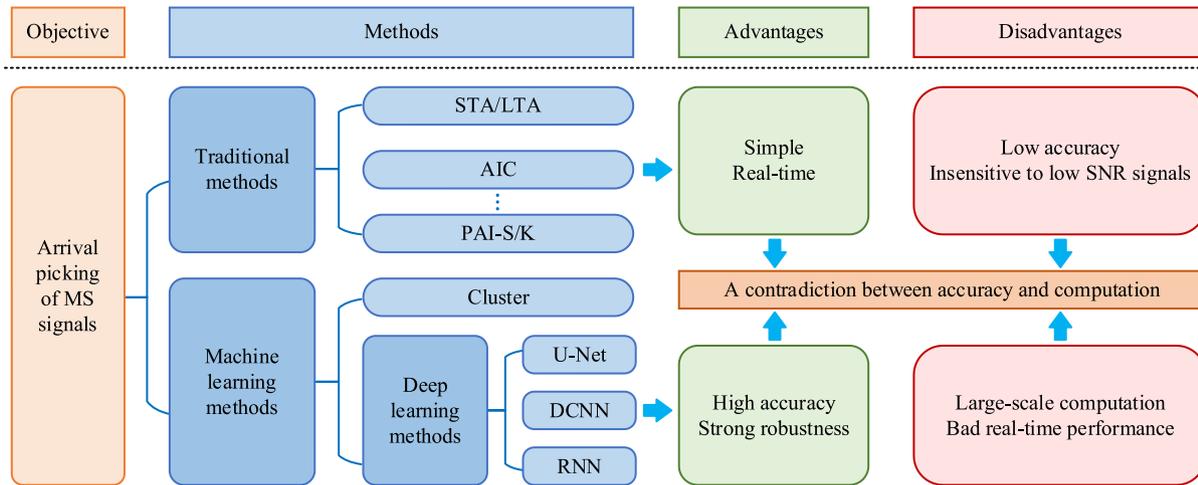


Fig. 1. A contradiction between accuracy and computation for existing methods.

information criterion (AIC) method (Maeda, 1985; Li et al., 2017), the phase arrival identification–skewness/kurtosis (PAI-S/K) method (Saragiotis et al., 2002; Nippres et al., 2010) and their variants. The STA/LTA method is widely used in practical MS monitoring due to its simplicity and rapidity. However, it has the following weaknesses: relatively low accuracy, instability due to dependence on the window length (Akram and Eaton, 2016) and insensitivity to signals with a low signal-to-noise ratio (SNR). The AIC and PAI-S/K methods achieve high-accuracy arrival-picking based on changes in information or statistics. However, these methods also perform poorly on low-SNR signals (Dong et al., 2018).

Machine learning methods mainly include clustering (Zhu et al., 2016, 2022; Ma et al., 2018; Guo et al., 2021; Lan et al., 2022) and deep learning methods for MS arrival picking. Deep learning methods are able to pick P- and S-arrivals simultaneously by extracting hidden features from the original MS signals. Thus, these methods are a popular focus of research because of their strong robustness and high accuracy. A few approaches use the regression method to determine the arrival sampling point directly. For example, Ross et al. (2018) used a convolutional neural network (CNN) to find the regression between waveforms and arrivals. A convolutional network named Cospy uses a two-stage structure, consisting of rough positioning followed by precise regression (Pardo et al., 2019).

In most deep learning methods, arrival picking is converted into a classification task for each sampling point, mainly using U-Net or long short-term memory (LSTM) as the backbone. The U-Net model (Ronneberger et al., 2015) is a typical encoder-decoder structure. Zhang et al. (2021b) proposed a control model for S-arrival based on the outputs of U-Net, which can be used to improve the early warning effect for rockburst. Moreover, data augmentation methods were also used for U-Net training (Zhang and Sheng, 2020; Zhang et al., 2021c). The LSTM model (Hochreiter and Schmidhuber, 1997) is a representative form of recurrent neural network (RNN), and is naturally suited for processing time series. Zheng et al. (2018) established a stacked model consisting of 7 LSTM layers with 1024 hidden units, which shows high accuracy for acoustic emission waveforms but implies an enormous model scale. EQTransformer (Mousavi et al., 2020) can simultaneously realize high-accuracy seismic signal detection and arrival picking. This model combines 1D (one-dimensional) convolution, residual connections, bidirectional LSTM and attention mechanism, implying a very deep encoder. Such large-scale computations may lead to

performance bottlenecks when attempting to process MS signals with a high sampling frequency in real-time. Subsequently, LEQNet (Lim et al., 2022), optimized from EQTransformer and reduced the scale of the model at the cost of a slight decline in accuracy. Xu et al. (2022) used multi-channel singular spectrum analysis (MSSA) and short-time Fourier transform (STFT) to extract features from MS signals, and then input these features into the arrival-picking model containing 34 LSTM cells with 600 hidden units. It has been proven to have good accuracy, while many recursive calculations limit its application in real-time processing. In summary, high accuracy and large-scale computations are the typical characteristics of the existing deep learning methods, making these methods applicable only for post-processing. The existing methods generally have a contradiction between accuracy and computation, as shown in Fig. 1. Thus, an arrival picking method with both high accuracy and low computations is of great significance for realizing real-time arrival picking and early warning of rock mass fracture.

2. Method

MS signals are the typical time series superposed by rock mass microfracture signals and environmental noise, implying changes in waveform characteristics such as frequency and amplitude. To better determine the arrivals of P- and S-waves, it is necessary to reasonably express both the short-term features (STF) and long-term features (LTF) of MS signals, which describe the instantaneous changes on the local scale and the macroscopic trend on the global scale, respectively. Thus, we first introduced the structure of the convolutional-recurrent neural network (CRNN), which is constructed from the perspective of STF and LTF extraction. Then, two targeted optimizations are introduced to ensure applicability to MS signals. Subsequently, we established a complete real-time arrival picking method based on CRNN.

2.1. CRNN model

The proposed CRNN model consists of an input layer, hidden layers, and an output layer, which are responsible for accepting rock mass microfracture signals as input, extracting the features of P- and S-arrivals, and enabling the expression and evaluation of P- and S-arrivals, respectively. The structure of the CRNN model is shown in Fig. 2.

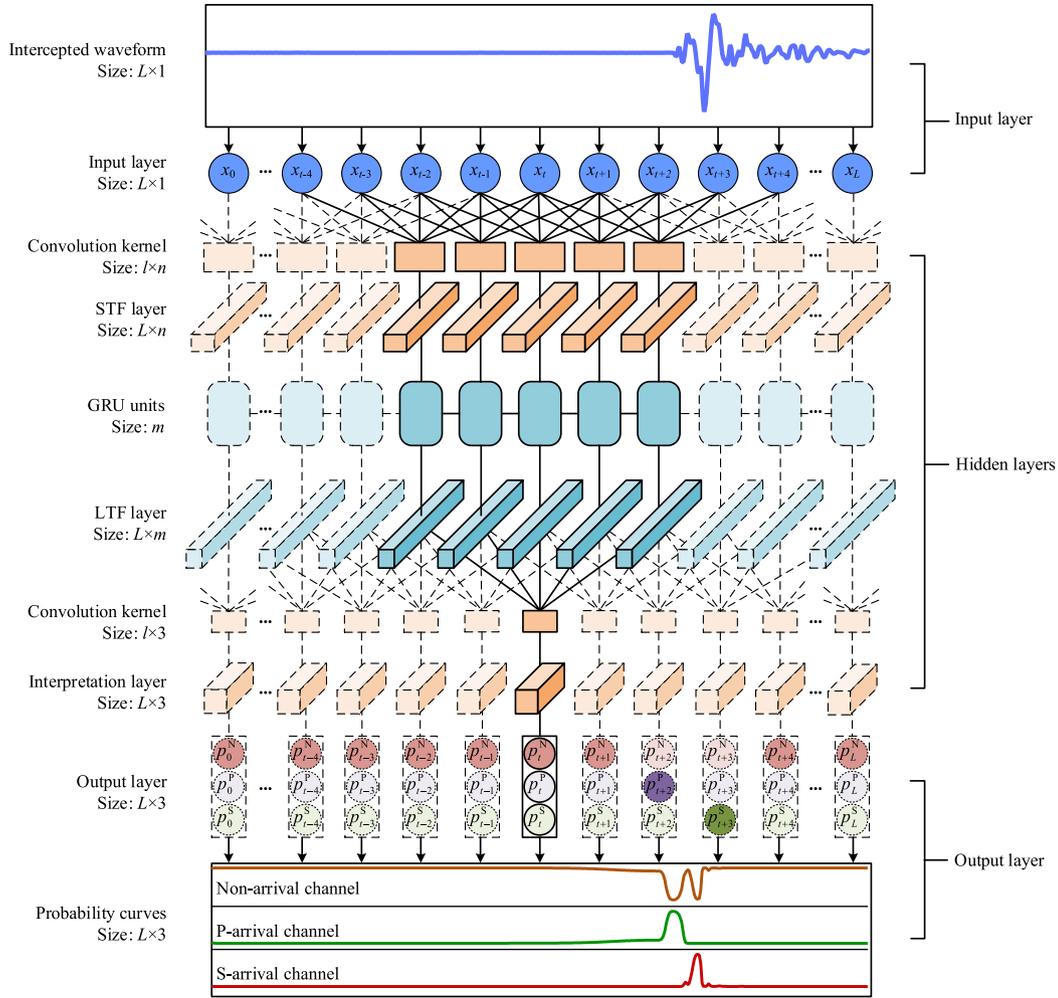


Fig. 2. Structure of the CRNN model. L is the input waveform length, l is the convolution kernel width, n is the convolution channels, m is the GRU units, and the solid line area is the calculation process of the t -th sampling point (taking $l = 5$ as an example in this context).

2.1.1. Input layer

The input of the CRNN model is original MS waveforms of rock mass microfracture signals, considering that deep learning models have a robust feature extraction ability. The need for additional filtering or transformation is non-essential, which would affect the real-time performance. However, it is necessary to normalize the input waveforms, considering that variations in rock fracture size and signal propagation distance can lead to considerable differences in the amplitudes of MS signals, which can even be as high as a factor of 10^3 . The maximum absolute scaling method is used to linearly scale the MS signals to the range of $[-1, 1]$ to eliminate the influence of amplitude differences (Galli, 2022):

$$\mathbf{X}' = \frac{\mathbf{X}}{\max[\text{abs}(\mathbf{X})]} \quad (1)$$

where \mathbf{X}' is the normalized input waveform, \mathbf{X} is the original input waveform, $\max(\cdot)$ is a function that outputs the maximum value of the argument, and $\text{abs}(\cdot)$ is a function that outputs the absolute value of the argument.

2.1.2. Hidden layers

The hidden layers are the key to the CRNN model. Three novel layers, namely the STF layer, the LTF layer and the interpretation

layer, are designed to perfectly extract the features of P- and S-arrivals from rock mass microfracture signals, as shown in Fig. 2. The STF layer uses a convolutional layer to extract the STF of MS signals, where the pooling layer is abandoned to maintain the length consistency of the inputs and outputs. The convolution kernel width (l) and convolution channels (n) are both important hyper-parameters that affect the extraction of STF, and their values are determined through orthogonal tests. The calculation process of the convolutional layer is written as (Goodfellow et al., 2016).

$$c_t^i = \text{ReLU}(\mathbf{X}' * \mathbf{K}^i + b^i) = \max\left(0, \sum_l X'_{t-l} K_l^i + b^i\right) \quad (2)$$

where c_t^i is the t -th output of the i -th convolutional layer, and $\mathbf{c}_t = [c_t^1 \ c_t^2 \ \dots \ c_t^n]$; $\text{ReLU}(\cdot)$ is the rectified linear unit; $*$ is the convolutional operator; \mathbf{K}^i is the i -th convolutional kernel with length l ; and b^i is the i -th bias.

The LTF layer uses a gated recurrent unit (GRU) (Bahdanau et al., 2014), which is a representative structure of RNN, to extract the LTF of MS signals. The GRU uses a reset gate and an update gate to realize information forgetting and updating, and uses state vector (\mathbf{h}_t) to store past information of MS signals before each sampling point, thus realizing the extraction of LTF with a small amount of computation, as shown in Eq. (3) (Goodfellow et al., 2016). The size

of \mathbf{h}_t is determined by the GRU units (m), which is also determined through orthogonal tests.

$$h_t^i = u_{t-1}^i h_{t-1}^i + (1 - u_{t-1}^i) \tanh \left(\sum_j W^{ij} r_{t-1}^j h_{t-1}^j + \sum_j U^{ij} c_t^j \right) \quad (3)$$

where h_t^i is the t -th output of the i -th GRU unit, and $\mathbf{h}_t = [h_t^1 \ h_t^2 \ \dots \ h_t^m]^T$; u_t^i and r_t^i are the t -th value of the i -th update and reset gate, respectively; $\tanh(\cdot)$ is the hyperbolic tangent function; W^{ij} and U^{ij} are weights of historical information and current input, respectively.

The interpretation layer again uses a convolutional layer to process the features obtained in the LTF layer. On one hand, \mathbf{h}_t is compressed into a three-dimensional vector $\mathbf{a}_t = [a_t^N \ a_t^P \ a_t^S]^T$ to facilitate subsequent processing in the output layer, where a_t^N , a_t^P and a_t^S are the features of the non-arrival, P-arrival and S-arrival, respectively, corresponding to sampling point t . On the other hand, the local connections of the convolutional layer can further analyze the local changes of LTF, improving the accuracy of P- and S-arrivals picking for rock mass microfracture signals. To enhance the generalization ability of the CRNN model, the dropout function is added to each hidden layer, which inhibits over-fitting by randomly dropping units during the training process.

2.1.3. Output layer

Since the outputs of the interpretation layer range from 0 to positive infinity, there is a lack of contrast between the pre- and post-sampling points. The softmax function normalizes the interpretation layer outputs to a probability matrix $\mathbf{p}_t = [p_t^N \ p_t^P \ p_t^S]^T$ (Peterson and Söderberg, 1989):

$$p_t^k = \frac{e^{a_t^k}}{\sum_{i=1}^3 e^{a_t^i}} \quad (4)$$

where p_t^k is the predicted probability of the k -th class at the t -th sampling point of the waveform, and k takes N, P and S, corresponding to non-, P- and S-arrival, respectively.

2.2. Optimization of the CRNN model

The arrival-picking task is essentially transformed into the classification of each sampling point using the CRNN model. However, because of the typically high sampling frequency (several kHz), MS waveforms are generally long time series, leading to extreme class imbalance and the loss of effective information. Thus, two corresponding optimizations of the CRNN model are proposed to accurately pick P- and S-arrivals of rock mass microfracture signals in real-time.

2.2.1. Waveform interception

Cho et al. (2014) showed that too long input sequences would reduce the learning performance of a GRU model and significantly increase the computational complexity, while too short input sequences will easily omit P- or S-arrivals and would be disturbed by changes in noise. When all other parameters remain unchanged, the effect of the input waveform length is studied to determine the appropriate value. The performance of the CRNN model is evaluated by the sum of the mean absolute errors for P- and S-arrivals (MAE_{SUM}):

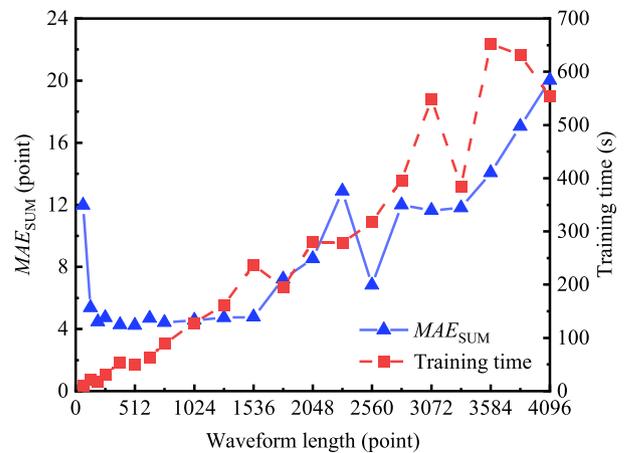
$$MAE_{SUM} = \frac{1}{2N} \sum_{i=1}^N [\text{abs}(T_{i,P} - T_{i,P}^*) + \text{abs}(T_{i,S} - T_{i,S}^*)] \quad (5)$$

where $T_{i,P}$ and $T_{i,S}$ are the automatically picked P- and S-arrivals of the i -th waveform, respectively; $T_{i,P}^*$ and $T_{i,S}^*$ are the real P- and S-arrivals of the i -th waveform, respectively; and N is the total number of waveforms.

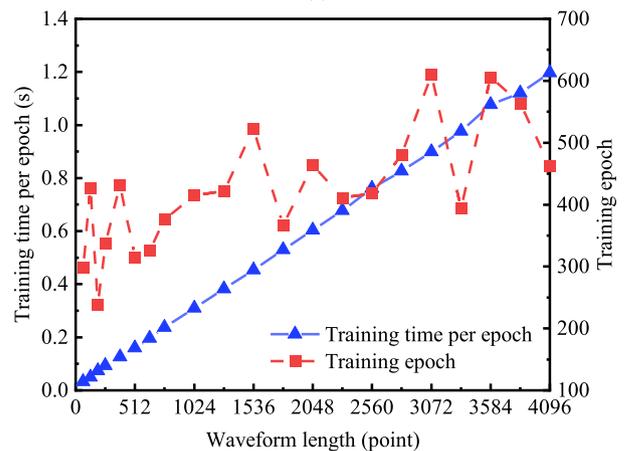
As shown in Fig. 3, with the increase of input waveform length L , the training time per epoch, the number of training epochs and the total training time show continuous increasing trends. More specifically, when the input waveform length L is in each of the three intervals (0, 256], [256, 1536] and [1536, 4096], MAE_{SUM} shows a high value, a stable low value and a rapidly increasing value, respectively. This implies that the optimal value of the input waveform length is between 256 and 1536 sampling points.

In addition, two properties of the input waveform should be guaranteed: it should contain the arrival points of the P- and S-waves, and it should contain background noise of sufficient length. Thus, the waveform interception method is adopted:

$$T_b = T_{\max} - a\Delta T_{\max} \quad (6)$$



(a)



(b)

Fig. 3. Influences of the input waveform length on the model training and picking accuracy: (a) MAE_{SUM} and total training time; and (b) Training time per epoch and training epochs.

$$T_e = T_{\max} + b\Delta T_{\max} \quad (7)$$

where T_b and T_e are the beginning and ending sampling points of the intercepted waveform, respectively; T_{\max} is the sampling point corresponding to the maximum absolute amplitude, and $T_{\max} = \text{argmax}[\text{abs}(\mathbf{X}_{\text{ms}})]$; \mathbf{X}_{ms} is the time series of the MS waveform, $\mathbf{X}_{\text{ms}} = [x_1 \ x_2 \ \dots \ x_i]$, with x_i being the amplitude value of the i -th sampling point; a and b are the front and post amplification factors, respectively; and ΔT_{\max} is the maximum time interval between the P- and S-waves within the monitoring range, which can be obtained by (Zhu et al., 2022).

$$\Delta T_{\max} = f_{\text{sys}} l_{\max} \left(\frac{1}{v_S} - \frac{1}{v_P} \right) \quad (8)$$

where f_{sys} is the sampling frequency of the MS monitoring system, l_{\max} is the maximum monitoring distance, and v_P, v_S are the wave velocities of the P- and S-waves, respectively.

Considering the Yebatan hydropower station project (YBT Project) as an example (see Section 4.1), the sampling frequency is 4000 Hz, the distance from each sensor to the edge of the monitoring area is between 200 m and 400 m, and the rock class in the project is mainly of grade III. Thus, it is assumed that $f_{\text{sys}} = 4000$ Hz, $l_{\max} = 300$ m, $v_P = 5500$ m/s and $v_S = 3500$ m/s (Zhu et al., 2019; Li et al., 2021). We set $a = 3$, where the length of $2\Delta T_{\max}$ is used to learn the background noise information, and the remaining length is used to intercept the arrival information of the P- and S-waves. We set $b = 1$ to prevent missing P- and S-arrivals. According to Eqs. (6)–(8), the input waveform length L is calculated to be 499, and is adjusted to $2^9 = 512$ sampling points for convenience in deep learning processing.

2.2.2. Loss function

The loss function directly affects the gradient direction during model training (Dickson et al., 2022). The relative proportions of non-arrival, P-arrival and S-arrival sampling points are seriously imbalanced for MS arrival picking. To prevent the class imbalance from affecting the learning of minority classes (P- and S-arrivals), we introduced a class weight w for P- and S-arrivals, and decomposed the loss into two parts: the average loss value at all sampling points ($Loss_{\text{all}}$) (Zhu and Beroza, 2019) and the average loss value at arrival points ($Loss_{\text{arr}}$):

$$Loss = Loss_{\text{all}} + \frac{2(w-1)}{L} Loss_{\text{arr}} \quad (9)$$

$$Loss_{\text{all}} = -\frac{1}{NL} \sum_{i=1}^N \sum_{j=1}^L \sum_{k=1}^3 y_{ij}^k \log_{10} p_{ij}^k \quad (10)$$

$$Loss_{\text{arr}} = -\frac{1}{2N} \sum_{i=1}^N \sum_{k=1}^3 (y_{i,P}^k \log_{10} p_{i,P}^k + y_{i,S}^k \log_{10} p_{i,S}^k) \quad (11)$$

where y_{ij}^k and p_{ij}^k are respectively the actual and predicted values of the k -th class for the j -th sampling point of the i -th waveform, $y_{i,P}^k, y_{i,S}^k, p_{i,P}^k$ and $p_{i,S}^k$ can be correspondingly obtained when j is taken as the actual arrival point of the P- and S-waves, respectively.

Through a comprehensive evaluation of $Loss_{\text{all}}$ and $Loss_{\text{arr}}$, it can be readily judged whether the model training direction is correct. The former evaluates the learning effect for the whole waveform without considering the class weight, whereas the latter evaluates the error between the predicted and actual values only at the actual arrival points of the P- and S-waves. Taking the waveform shown in Fig. 4a as an example, Fig. 4b shows that the model is too insensitive

to respond effectively to the arrivals due to class imbalance, as indicated by the small $Loss_{\text{all}}$ and large $Loss_{\text{arr}}$. In contrast, Fig. 4c shows the case of a large $Loss_{\text{all}}$ and a small $Loss_{\text{arr}}$, implying that the model is too sensitive and will incorrectly judge small waveform changes as arrivals. Fig. 4d shows a divergent model with both a large $Loss_{\text{all}}$ and a large $Loss_{\text{arr}}$. Only when $Loss_{\text{all}}$ and $Loss_{\text{arr}}$ are small, as shown in Fig. 4e, the model can achieve good performance. At this time, the model achieves a balance between robustness and sensitivity, picking the P- and S-arrivals correctly without the influence of noise.

Thus, during model training, $Loss$ is used to guide gradient descent, and $Loss_{\text{all}}$ and $Loss_{\text{arr}}$ are taken as early stop criteria to avoid over-fitting issue: (1) the minimum value of $Loss_{\text{all}}$ is not updated for 20 consecutive epochs, and (2) $Loss_{\text{arr}} < 0.1$.

2.3. A real-time arrival picking method based on CRNN

The proposed real-time arrival picking method for rock mass microfracture signals consists of three modules: CRNN model training, real-time arrival picking, and dynamic feedback, as shown in Fig. 5.

2.3.1. CRNN model training

The steps for CRNN model training are as follows:

- (1) Step 1: Select the training waveforms and mark the arrival points manually. The selected training waveforms should encompass rock mass microfracture signals with various amplitudes, frequencies and SNRs to enable comprehensive learning of the arrival features of various signals.

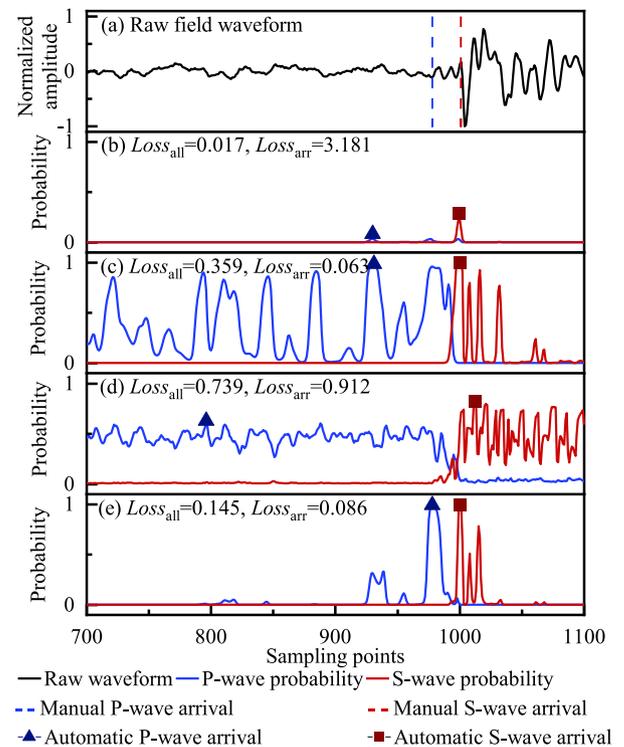


Fig. 4. Arrival picking results of the CRNN models with different $Loss_{\text{all}}$ and $Loss_{\text{arr}}$: (a) The original waveform; (b) A small $Loss_{\text{all}}$ and a large $Loss_{\text{arr}}$; (c) A large $Loss_{\text{all}}$ and a small $Loss_{\text{arr}}$; (d) A large $Loss_{\text{all}}$ and a large $Loss_{\text{arr}}$; and (e) A small $Loss_{\text{all}}$ and a small $Loss_{\text{arr}}$.

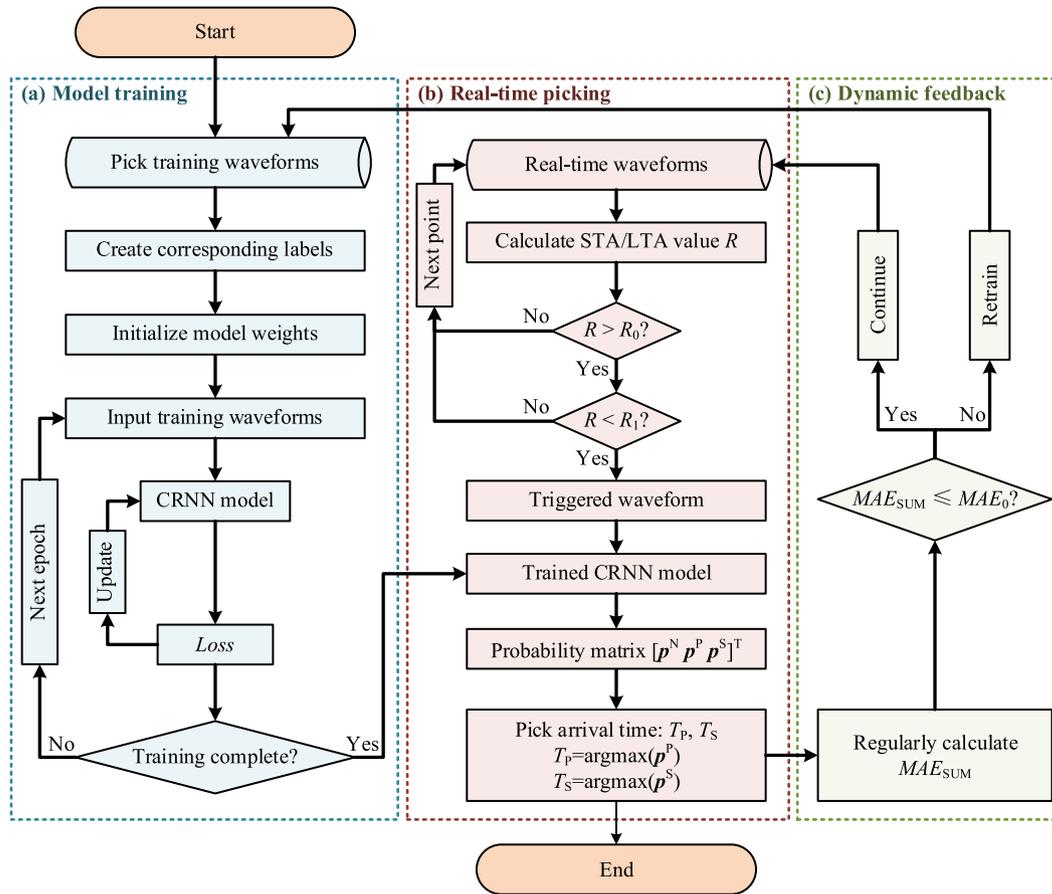


Fig. 5. The flow chart of real-time arrival picking method based on CRNN.

- (2) Step 2: Label each sampling point of the waveforms with one-hot encoding. Non-arrival, P-arrival and S-arrival sampling points are marked as $[1, 0, 0]$, $[0, 1, 0]$ and $[0, 0, 1]$, respectively. Each waveform corresponds to a sparse label matrix in chronological order, with a size of $L \times 3$.
- (3) Step 3: Randomly initialize the weights of the CRNN model.
- (4) Step 4: Input a training waveform into the CRNN model and output the corresponding probability matrix $\mathbf{p} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_t \ \dots \ \mathbf{p}_L]$.
- (5) Step 5: Calculate $Loss$, $Loss_{all}$ and $Loss_{arr}$ according to Eqs. (9)–(11).
- (6) Step 6: Use the Adam optimizer to update the CRNN model weights based on $Loss$.
- (7) Step 7: Judge whether training is completed based on $Loss_{all}$ and $Loss_{arr}$. End training if the conditions are met; otherwise, return to Step 4.

2.3.2. Real-time arrival picking

The steps for real-time arrival picking are as follows:

- (1) Step 1: Monitor rock mass microfracture signals in real-time.
- (2) Step 2: Calculate the recursive STA/LTA value R of the current sampling point.
- (3) Step 3: If $R > R_0$, where R_0 is the trigger threshold, proceed to Step 4; otherwise, return to Step 1.
- (4) Step 4: Record the MS waveform until $R < R_1$, where R_1 is the end threshold. Form a complete triggered waveform by Eqs. (6) and (7).

- (5) Step 5: Input the triggered waveform into the trained CRNN model and obtain the probability matrix.
- (6) Step 6: Take the sampling points corresponding to the maximum probability values of \mathbf{p}^P and \mathbf{p}^S as the P- and S-arrivals of the rock mass microfracture signal.

2.3.3. Dynamic feedback

The steps for dynamic feedback are as follows:

- (1) Step 1: Regularly spot-check and calculate the MAE_{SUM} of the CRNN model. In particular, when the field environment, geological conditions or SNR changes significantly, it is necessary to evaluate the arrival picking accuracy of the CRNN model and calculate its MAE_{SUM} .
- (2) Step 2: If $MAE_{SUM} \leq MAE_0$, where MAE_0 is an error threshold meeting the engineering requirements, the CRNN model is qualified and can be used to pick up arrivals of waveforms; otherwise, retrain the CRNN model with new MS waveforms under the current working conditions.

3. Performance analysis

3.1. Impact of hyperparameters on model performance

Hyperparameters play an important role in the performance of deep learning models (Goodfellow et al., 2016). To optimize the performance of the CRNN model, four main hyperparameters were studied by combining simulation and orthogonal experiments:

Table 1
Levels and values of hyperparameters in the CRNN model.

Hyperparameters	Levels	Fixed hyperparameters	Value
Class weight	[1, 128, 256, 384, 512]	Convolution strides	1
Convolution kernel width	[3, 7, 11, 15, 19]	Learning rate	0.001
Convolution channels	[4, 8, 12, 16, 20]	Activation function	ReLU

class weight, convolution kernel width, convolution channels and GRU units.

3.1.1. Experimental scheme

Each hyperparameter is set with five levels, and the values corresponding to these levels are taken at equal intervals according to experience, as shown in Table 1. Thus, 25 cases are designed in accordance with the orthogonal table $L_{25}(5^4)$ (Hedayat et al., 1999), as shown in Table 2. Four evaluation indexes are designed: MAE_{SUM} , $Loss_{all}$, $Loss_{arr}$ and ATT . MAE_{SUM} is used to evaluate the picking ability of the CRNN model, calculated by Eq. (3), $Loss_{all}$ and $Loss_{arr}$ are used to evaluate the stability and sensitivity of the CRNN model, calculated by Eqs. (8) and (9), and ATT is the time required for each training epoch, which is mainly used to evaluate the calculation speed of the CRNN model. Other hyperparameters of the CRNN model are also listed in Table 1.

3.1.2. Data preparation

In view of the actual arrivals and controllable SNRs, the synthetic MS signals are the best choice to accurately evaluate the performance of the CRNN model. Different from other synthesis methods, we take real noises from the YBT Project as the background noise, and fully consider the characteristic distributions of the field MS signals such as the frequency, duration, and P- and S-wave amplitude ratio. Then we use a Gaussian function to modulate the sine-basis function, calculated by Eq. (12). To ensure that the

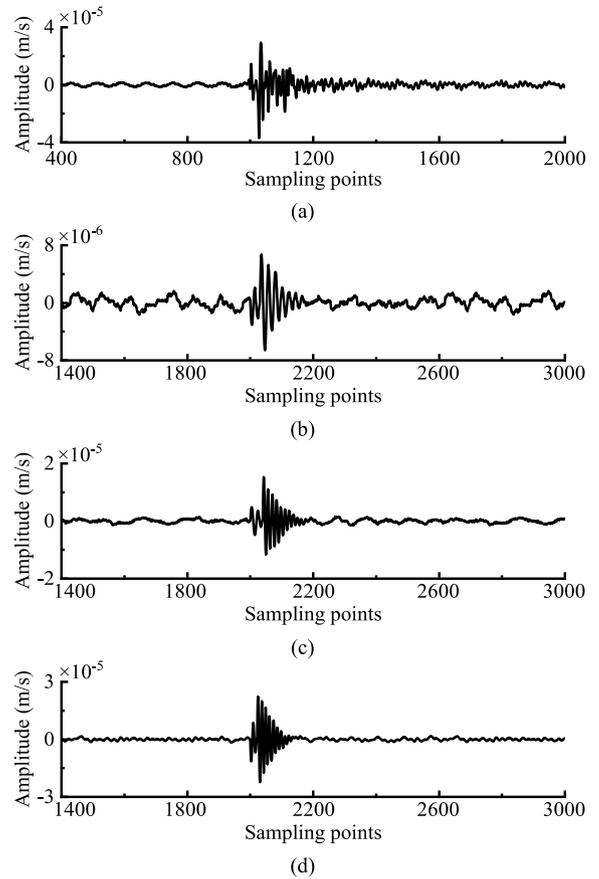


Fig. 6. Typical actual and synthetic MS signals of rock mass microfracture: (a) Real rock mass microfracture signal; (b) Synthetic signal with SNR = 10 dB; (c) Synthetic signal with SNR = 15 dB; and (d) Synthetic signal with SNR = 20 dB.

Table 2
Orthogonal test scheme and corresponding responses.

Case No.	Hyperparameters				Responses			
	Class weight	Convolution kernel width	Convolution channels	GRU units	MAE_{SUM} (point)	$Loss_{all}$	$Loss_{arr}$	ATT (s)
1	1	3	4	8	14.084	0.0168	3.347	0.164
2	1	7	8	16	2.966	0.0055	0.933	0.173
3	1	11	12	24	2.186	0.0041	0.659	0.197
4	1	15	16	32	2.187	0.0030	0.484	0.207
5	1	19	20	40	2.301	0.0011	0.179	0.216
6	128	3	8	24	6.467	0.1553	0.255	0.184
7	128	7	12	32	2.973	0.0398	0.045	0.195
8	128	11	16	40	2.401	0.0302	0.026	0.209
9	128	15	20	8	2.261	0.0487	0.064	0.169
10	128	19	4	16	2.762	0.0475	0.051	0.174
11	256	3	12	40	3.977	0.1176	0.084	0.199
12	256	7	16	8	3.338	0.0891	0.085	0.166
13	256	11	20	16	2.175	0.0537	0.028	0.174
14	256	15	4	24	3.048	0.0632	0.026	0.188
15	256	19	8	32	2.433	0.0393	0.015	0.204
16	384	3	16	16	4.243	0.1368	0.104	0.173
17	384	7	20	24	3.053	0.0833	0.030	0.185
18	384	11	4	32	3.024	0.0760	0.028	0.197
19	384	15	8	40	2.339	0.0632	0.018	0.209
20	384	19	12	8	2.312	0.0693	0.019	0.169
21	512	3	20	32	4.249	0.1152	0.051	0.192
22	512	7	4	40	4.587	0.1396	0.052	0.201
23	512	11	8	8	4.124	0.1252	0.040	0.162
24	512	15	12	16	2.481	0.1034	0.029	0.177
25	512	19	16	24	2.267	0.0658	0.016	0.197

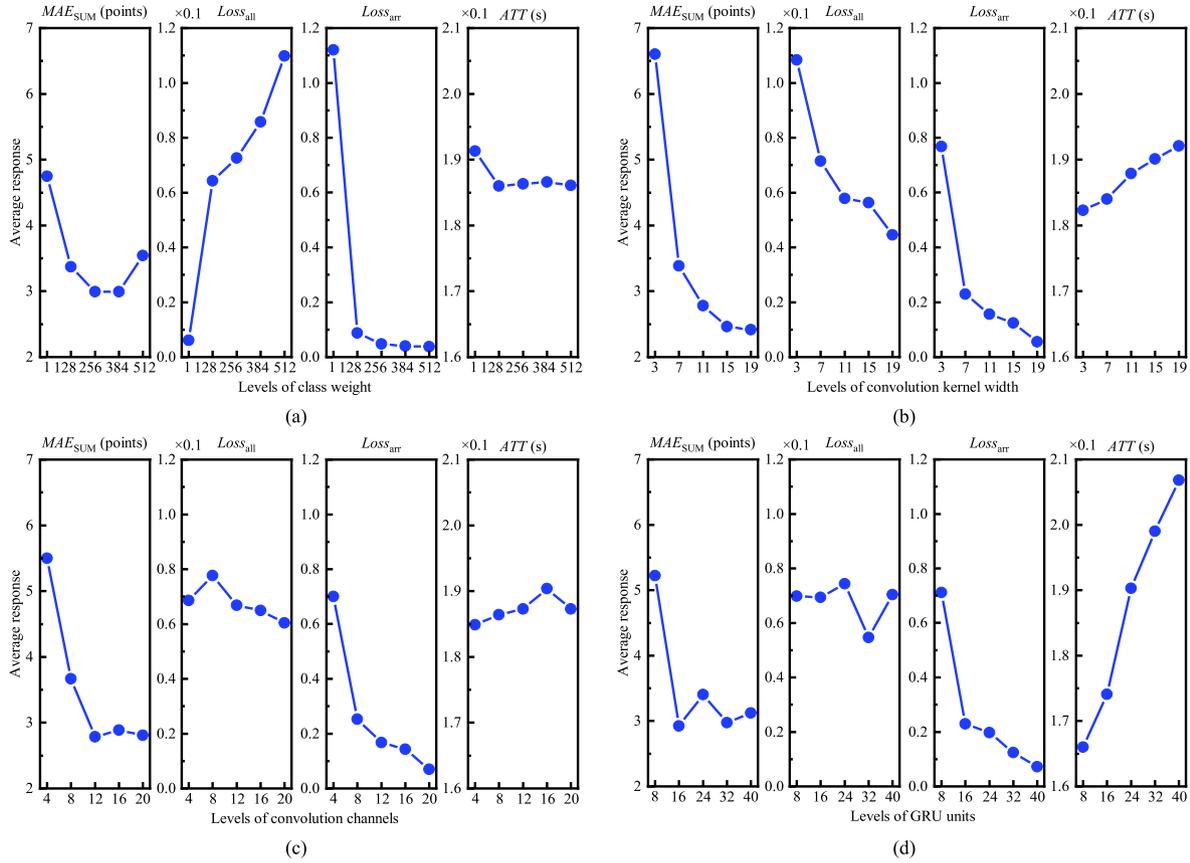


Fig. 7. Average responses of evaluation indexes for each factor: (a) Class weight; (b) Convolution kernel width; (c) Convolution channels; and (d) GRU units.

input noise does not contain potential valid signals, the real noise is defined as a continuous signal segment with STA/LTA values consistently less than 2. To prevent information leakage, the real noises do not include any data used in the engineering applications below. Typical synthetic MS signals are shown in Fig. 6, which are highly similar to actual field signals. A total of 2500 MS signals were synthesized, including 1000 signals for training and 1500 signals for testing.

$$s(t) = A_P g_P(t) + A_S g_S(t) + A_N g_N(t) \quad (12)$$

where $g_N(t)$ is the real noise, A_P , A_S and A_N are the amplitude coefficients of the P-wave, S-wave and noise, respectively, and $g_P(t)$ and $g_S(t)$ are the sine functions of the P- and S-waves, respectively, modulated by a Gaussian function and calculated by

$$g_i(t) = \begin{cases} 0 & (t < T_i) \\ e^{-\frac{9}{2} \left(\frac{t-T_i}{l_i} \right)^2} \sin[2\pi f_i(t - T_i)] & (t \geq T_i) \end{cases} \quad (13)$$

where T_i , f_i and l_i ($i = P, S$) are the arrival time, frequency and duration of the P- and S-waves, respectively.

3.1.3. Experiment results analysis

To reduce the impact of the random initial weights on the training results, each case conducts 3 independent trainings with 3000 epochs. During training, we recorded the MAE_{SUM} of the test waveforms every 50 epochs, and finally selected the model corresponding to the minimum MAE_{SUM} among the three independent

trainings as the best model for each case. The test results of the 25 cases are shown in Table 2.

Fig. 7 shows a line graph of the responses to each hyperparameter based on intuitive analysis. The results show that:

- (1) In Fig. 7a, with the increase of class weight, MAE_{SUM} first rapidly decreases to a minimum value and then increases; $Loss_{all}$ increases continuously; $Loss_{arr}$ first rapidly decreases to a low level and then remains stable; and ATT is basically stable throughout the process. Based on the discussion in Section 2.2, it can be found that the class weight is helpful to reduce the impact of class imbalance, but too large class weight will make the model too sensitive to STF while ignoring LTF, leading to incorrect arrival picking. In addition, the training time is not sensitive to the class weight. The suggested value of the class weight is 256.
- (2) In Fig. 7b, as the convolution kernel width increases, MAE_{SUM} , $Loss_{all}$ and $Loss_{arr}$ first decrease and then tend to stabilize, while ATT increases significantly. Considering the arrival picking accuracy and computational cost, the convolution kernel width is taken to be 15 in this context.
- (3) In Fig. 7c, as the convolution channels increase, the trends of responses of MAE_{SUM} , $Loss_{all}$ and $Loss_{arr}$ are similar to those of the convolution kernel width, whereas ATT increases slowly. The number of convolution channels is taken to be 12 in this paper.
- (4) In Fig. 7d, with the increase of GRU units, MAE_{SUM} first quickly drops to a minimum and then begins to fluctuate; $Loss_{all}$ fluctuates within a consistent level; $Loss_{arr}$ continuously declines; and ATT increases rapidly. Once the value of

GRU units exceeds a certain threshold, the arrival picking ability of the model is not significantly improved, but the computational cost greatly increases. Therefore, the number of GRUs units is set to 16 in this study.

In general, the arrival picking ability of the CRNN model is significantly impacted by the convolution kernel width, followed by the convolution channels, GRU units, and class weight. GRU units have the greatest impact on the computational cost, followed by the convolution kernel width, which is consistent with the theory of CNN and RNN. An increase of hyperparameters can effectively improve the arrival picking accuracy at the cost of a greater computation. However, blindly increasing the hyperparameters will improve the CRNN model only to a limited extent, and may even cause over-fitting issue.

3.1.4. Performance verification

Considering that the optimal hyperparameter configuration is not in the case of orthogonal test, we trained the CRNN model with the optimal hyperparameters to verify the optimization effect. The class weight, convolution kernel width, convolution channels and GRU units are 256, 15, 12 and 16, respectively. The training and test waveforms are the same as those in the orthogonal experiment, and the results are shown in Table 3. The MAE_{SUM} of the optimized CRNN model is 2.26 points, which is smaller than that of any case in Table 2, and the arrival picking accuracy within 4 sampling points is excellent, which is 91.80% for P-wave and 98.73% for S-wave, respectively. Therefore, the orthogonal test correctly reveals the influence of each hyperparameter on the model performance, and the optimized model is defined as the standard CRNN model used below.

3.2. Computational performance analysis

Computational performance is an important index to measure the ability of the CRNN model to pick MS signals in real-time. We compared the computational performance of the standard CRNN model and the typical U-Net model. In the U-Net model (Zhang et al., 2021b, 2021c), the number of convolutional layers, convolution kernel width, convolution channels and the up-sampling rate are 5, 7, [8, 16, 32, 64, 128] and 2, respectively. We used the trained CRNN model and U-Net model to pick arrivals of 100 batches (each batch has 100 waveforms), recorded the CPU usage of the models every 0.1 s, and calculated the average time per batch.

Three multi-core computers with reduced performance and one single-core computer were selected to analyze the universality of the models. The computing performances of the CRNN model and U-Net model on each device are shown in Table 4. On the multi-core CPU devices, both the CPU usage and average time per batch of the CRNN model are lower than those of the U-Net model, which are reduced by 30% and 25% approximately on average, respectively. With the decline in device performance, the gap in the CPU usage between the CRNN method and U-Net method decreases gradually, from 34.4% to 25.3%, and the gap in the average time per batch correspondingly increases, from 21.1% to 28.6%. In the

Table 3 Results of arrival picking based on the optimized CRNN model.

Wave	Absolute error range (%)						MAE _{SUM} (point)
	[0, 4]	[4, 8]	[8, 16]	[16, 32]	[32, 64]	[64, 512]	
P	91.80	4.00	2.47	1.20	0.47	0.07	1.51
S	98.73	0.13	0.87	0.27	0.00	0.00	0.65
Total	95.27	2.07	1.67	0.74	0.24	0.04	2.16

extreme case of the single-core in device 4, both models fully occupy the thread, and the difference in the average time per batch is as high as 43.5%.

Taking device 3 with low performance as an example, the CPU usage curve during arrival picking is shown in Fig. 8. The U-Net model takes 10.50 s to pick arrivals, and the CPU usage is recorded 104 times, with values mainly between 68.46% and 82.30%. Taking 90% usage as the alert threshold, the U-Net model triggers the alert threshold 13 times, accounting for 12.5% of the total. It means that the U-Net method has almost reached the performance limit of device 3, which may affect the real-time acquisition and processing of MS signals, hindering the stable operation of the MS monitoring system. In contrast, the CRNN method takes 7.54 s to pick arrivals, where the CPU usage is mainly between 50.09% and 62.61%, and the maximum usage is only 74.69%. It shows that the CRNN method can work perfectly even on devices with limited performance, without affecting the normal operation of the MS monitoring system. Thus, compared with the U-Net model, the CRNN model has lower CPU usage, less computational cost and broader applicability.

4. Application

4.1. Engineering background

The YBT Project is located on the main upstream of the Jinsha River, at the junction of Baiyu County, Sichuan Province and Gongjue County, Tibet Autonomous Region, China. The underground caverns of this project are located in the Jinshajiang fault zone, with complex geological structures and developed faults. The surrounding rock is mainly of grade III quartz diorite. The uniaxial compressive strength is 120–180 MPa, and the measured maximum principal stress is 37.6 MPa. Relevant analysis shows the risks of rockburst, wall spalling, collapse and fault sliding during construction. Therefore, the SinoSeism (SSS) MS monitoring system, developed by Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, and Hubei Seaquake Technology Co., Ltd., was deployed to monitor MS activity during the whole construction process. The monitoring system consists of 1 time server, 1 data server, 4 acquisition units, 23 mono-component sensors and 3 three-component sensors. The sampling frequency is 4 kHz, the signal trigger algorithm is the improved STA/LTA algorithm, and the trigger threshold value is set to 4, instead of the more generally applied value of 6, to obtain more MS information of rock mass microfracture.

A total of 1607 rock mass microfracture waveforms were collected from November 27, 2021 to December 5, 2021, 803 of which were randomly selected for training, and the remaining 804 waveforms used for testing. The SNR and amplitude distributions of

Table 4 Computational performances of the two methods on different computers.

Cores	Order	Configuration	Average CPU usage (%)			Average time per batch (ms)		
			U-Net	CRNN	D-value (%)	U-Net	CRNN	D-value (%)
Multi-core	1	CPU: i7-11700F RAM: 16G 3200 MHz	59.85	39.27	34.4	57	45	21.1
	2	CPU: R7-4800H RAM: 16G 3200 MHz	55.39	38.33	30.8	108	81	25.0
	3	CPU: E5-1620 RAM: 8G 2400 MHz	75.25	56.20	25.3	105	75	28.6
Single-core	4	CPU: E5-1620 RAM: 8G 2400 MHz	99.60	99.68	0.1	170	96	43.5

Note: D-value is the percentage difference between the CRNN and U-Net models.

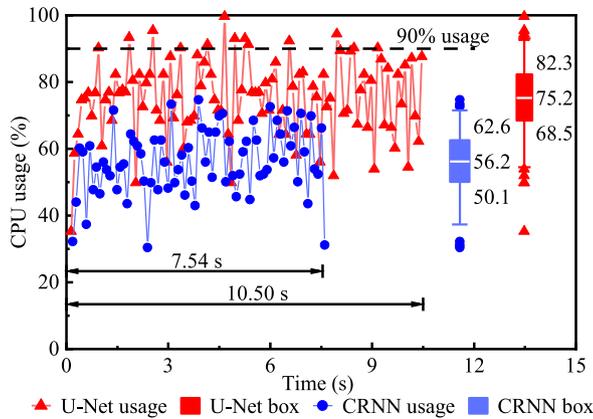
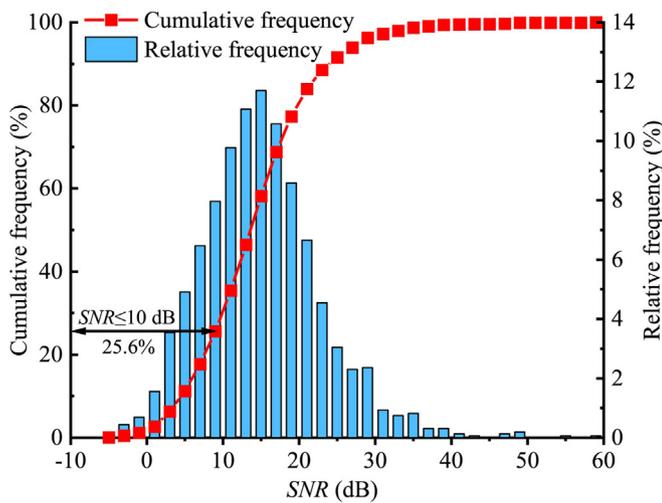
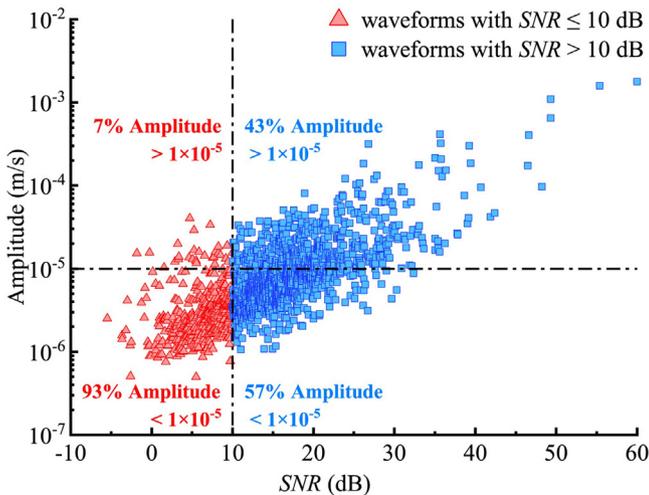


Fig. 8. CPU usage curve during arrival picking on device 3.



(a)



(b)

Fig. 9. Characteristic statistics of rock mass microfracture signals from the YBT Project: (a) Distribution histogram of SNR, and (b) Scatter distribution of amplitude-SNR.

all signals are shown in Fig. 9. The SNR can be roughly estimated by Eq. (14). Fig. 9a shows that the proportion of signals with $SNR < 10$ dB among all signals is 25%. Fig. 9b shows that the amplitude of

66% of the rock mass microfracture signals is lower than 10^{-5} m/s, and this proportion increases to 93% for signals with $SNR < 10$ dB. Thus, the waveforms from the YBT Project are typical weak rock mass microfracture signals, which are not conducive to picking P- and S-arrivals via traditional methods, such as the STA/LTA and AIC methods (Leonard and Kennett, 1999; Dong et al., 2018).

$$SNR \approx 10 \log_{10} \left\{ \frac{\sigma^2[s(t)] - \sigma^2[n(t)]}{\sigma^2[n(t)]} \right\} \quad (14)$$

where $s(t)$ and $n(t)$ are the waveform data with 800 sampling points after and before P-arrival, respectively.

4.2. Model training

To train a CRNN model suitable for picking up P- and S-arrivals of weak rock mass microfracture signals from the YBT Project, we first built the CRNN model structure as described in Section 2.1. Then, the synthetic waveforms were generated to obtain the optimal hyperparameters as described in Section 3.1. Finally the training field waveforms were input (see Section 4.1) to train the model weights via the Adam optimizer. The curves of the losses and time consumed during model training are shown in Fig. 10. Although $Loss_{arr}$ starts at a high level of 1.56 at the beginning of training, it decreases rapidly with the training and drops below $Loss_{all}$ after only 45 epochs. At the same time, $Loss_{all}$ declined steadily and finally stabilized at approximately 0.14, without a trend of first decreasing and then increasing. The trends of $Loss_{all}$ and $Loss_{arr}$ indicate that the class weight value is reasonable for highlighting arrival points, while also maintaining robustness for non-arrival points.

The CRNN model completes its training at 314 epochs, with a $Loss_{all}$ value of 0.14 and a $Loss_{arr}$ value of 0.09. With NVIDIA's CUDNN parallel computing platform, the average training time per epoch is only 0.16 s, and the total training time is 50.28 s. The results show that the CRNN model has an outstanding convergence speed and low training cost, meeting the demands of on-site monitoring.

4.3. Picking results analysis

Considering that the actual arrivals of field MS signals are theoretically unobservable, the manually picked arrivals are approximately considered as the actual arrivals (Tan and He, 2016). We define A_1 as the picking accuracy of P- and S-arrivals within 4 sampling points, compared with the manually picked arrivals, and A_2 as the picking accuracy within 16 sampling points. The improved STA/LTA, AIC, trained CRNN and trained U-Net methods were used to pick the P- and S-arrivals of 804 test waveforms. The short- and long-term windows of the STA/LTA are 20 and 200 sampling points, respectively. The threshold is 5 for P-arrival and half of the maximum value for S-arrival, and the characteristic functions for P- and S-arrival (PCF(t) and SCF(t), respectively) are given in Eqs. (15) and (16) (Zhang et al., 2021b). The time window of the AIC is set to $[T_b, T_{max}]$ for P-arrival, and $[T_p, T_p + \alpha\Delta T]$ for S-arrival, where ΔT is the number of sampling points from T_p to T_{max} , and α is a scale factor with a value of 1.2 (Zhu et al., 2022). The hyperparameters of the U-Net model are shown in Section 3.2.

$$PCF(t) = x_t^2 - x_{t-1}x_{t+1} \quad (15)$$

$$SCF(t) = f_{sys}(x_t - x_{t-1})^2 \quad (16)$$

The A_1 accuracy, A_2 accuracy and MAE values for P- and S-waves are shown in Table 5. In terms of the A_1 accuracy, the CRNN method

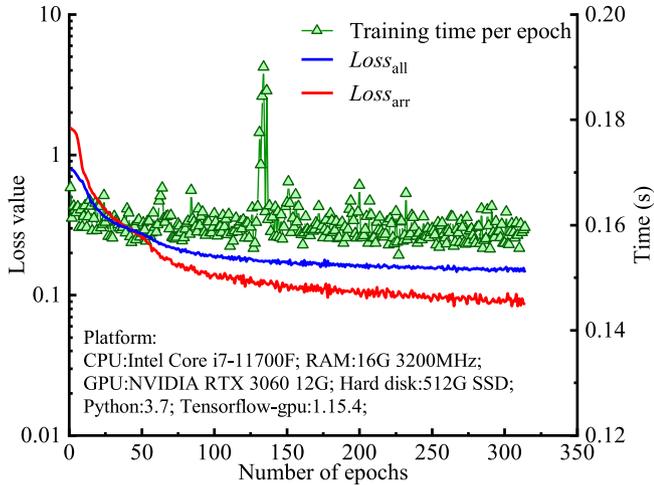


Fig. 10. Training losses and training time.

performs best, with 87.44% for P-wave and 91.29% for S-wave, followed by the U-Net method. The AIC and STA/LTA methods have a poor accuracy, especially the STA/LTA method, with 30.47% for P-wave and 39.18% for S-wave. For the A_2 accuracy, the CRNN and U-Net methods have a similar performance, both exceeding 96% for P- and S-waves. The average accuracy of the AIC and STA/LTA methods increases to 82.34%, which is still generally lower than that of deep learning methods. In terms of MAE, all MAE values of the CRNN are the lowest, with 2.57 sampling points for P-wave, 1.66 sampling points for S-wave and 4.22 sampling points in total, followed by the U-Net method. The AIC method is slightly better than the STA/LTA method and inferior to the CRNN and U-Net methods. Additionally, the S-arrival picking accuracies of the four methods are generally higher than that of the P-arrival. This is because the low SNR of MS signals in the YBT project results in more noise interference on P-waves than S-waves, which is consistent with the relevant study (Zhou et al., 2016; Fu et al., 2020).

To compare the differences in the arrival picking results based on the four methods, the distributions of the arrival picking error and the change in the arrival picking error with SNR are calculated and shown in Figs. 11 and 12, respectively. In Fig. 11, the P- and S-arrivals picking errors of the CRNN and U-Net methods are characterized by ‘symmetrical concentration’, that is, the picked arrivals are distributed evenly and symmetrically on both sides of the manually picked arrivals, mainly concentrated on the range of -4 to 4 sampling points. The AIC method has the characteristic of ‘slight delay for the majority and extreme advance for the minority’ for P-waves, mainly concentrated within 8 sampling points. However, as the noise increases, arrivals can be easily triggered many sampling points in advance by the incorrect minimum points on the AIC curve, resulting in poor picking stability. For S-waves, the AIC method shows the characteristic of ‘scattered advance’, being mainly distributed between -16 and 4 sampling points. The STA/LTA method shows the characteristics of ‘general delay’ for P-waves and ‘delay for the majority and scattered advance for the minority’ for S-waves, mainly caused by its mechanisms of the sliding window and threshold-based trigger.

Fig. 12 shows that as the SNR decreases, the P- and S-arrivals picking errors of all four methods tend to increase, but the CRNN and U-Net methods are less affected. In contrast, the AIC method performs well for $SNR \geq 20$ dB but sharply weakly for $SNR < 20$ dB. The STA/LTA method shows high sensitivity to the SNR for P-arrival picking, and passivation for S-arrival picking.

Table 5
Arrivals picking accuracy of each method.

Method	P-arrival			S-arrival			MAE_{SUM} (point)
	A_1 (%)	A_2 (%)	MAE_P (point)	A_1 (%)	A_2 (%)	MAE_S (point)	
STA/LTA	30.47	74.00	15.22	39.18	84.58	9.59	24.81
AIC	68.41	85.70	13.64	66.29	85.07	7.38	21.02
U-Net	84.33	96.39	3.32	90.80	98.51	1.86	5.18
CRNN	87.44	97.64	2.57	91.29	98.38	1.66	4.22

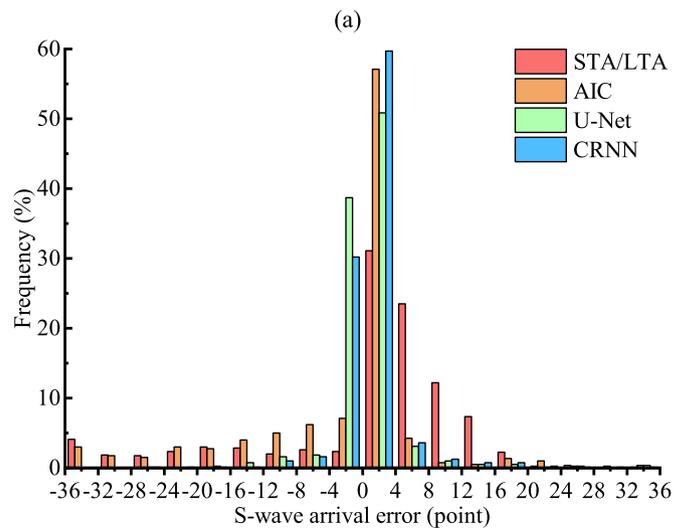
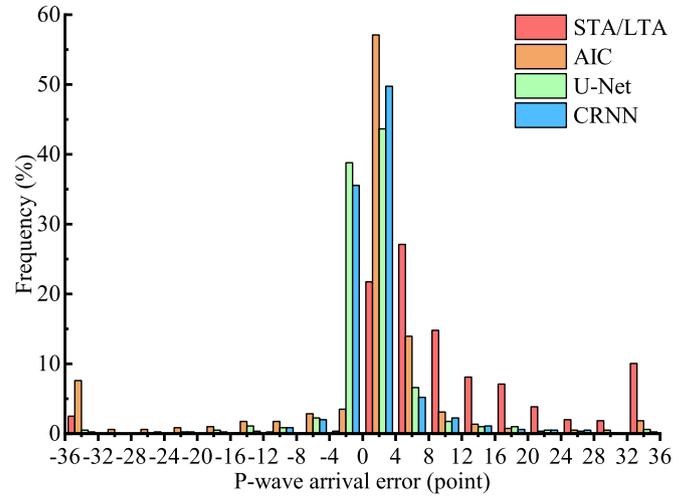


Fig. 11. Distributions of arrival picking errors based on the four methods: (a) P-arrival picking error, and (b) S-arrival picking error. The negative and positive errors mean that the picked arrival point is earlier and later than the manually picked point, respectively. Errors exceeding $+32$ and -32 sampling points are classified into $[32, 36]$ and $[-36, -32]$ intervals, respectively.

Fig. 13 presents the picking details for rock mass microfracture signals with 6 different SNRs. There is a general delay in P- and S-arrivals picking due to the average window of the STA/LTA method, and the degree of delay increases as the SNR decreases. When $SNR < 10$ dB, as shown in Fig. 13d, the P-arrival picked by the STA/LTA method lags behind the S-arrival, which indicates that the method can no longer distinguish the P-wave signal from the strong

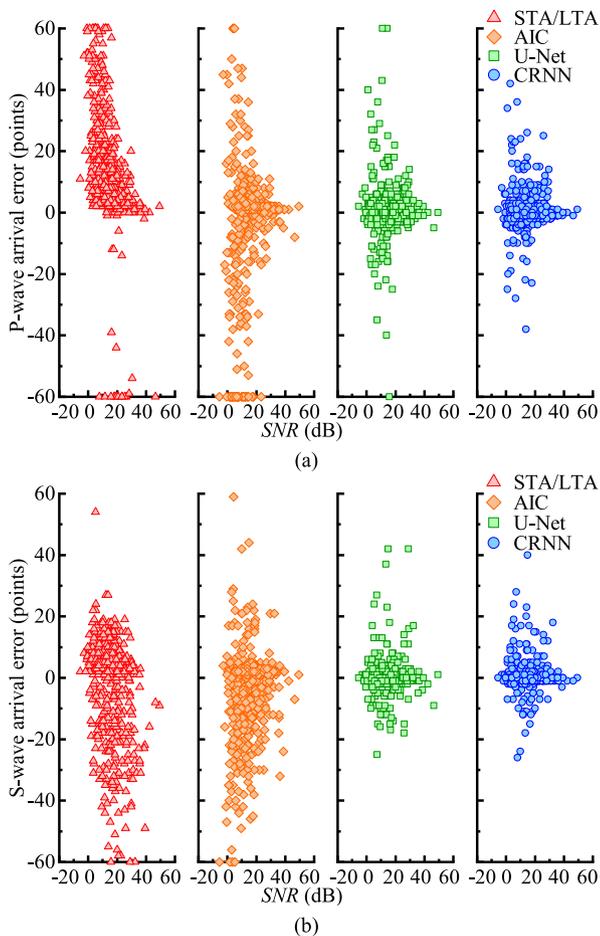


Fig. 12. Scatter distributions of arrival picking errors with SNR based on the four methods: (a) P-arrival picking error, and (b) S-arrival picking error. Errors exceeding ± 60 sampling points are counted as ± 60 , respectively.

background noise. The AIC method has good picking accuracy, but its picking reliability depends on the quality of the input signals. Fig. 13f shows the picking results when the input signal has a low SNR value and high background noise. The AIC method picks the P-arrival in advance under the influence of the noise, and then incorrectly judges the real P-arrival as the S-arrival. The U-Net and CRNN methods can widely adapt to arrival picking under different SNR conditions. Nevertheless, it is worth mentioning that deep learning methods have a data-driven nature, which may lead to unexpected impacts of slight differences in the input signals. As shown in Fig. 13d, the slight change in the background noise around the 950th sampling point is obviously reflected in the outputs of both the U-Net and CRNN methods. The difference is that the U-Net method misjudges it, while the CRNN method distinguishes the disturbance promptly and picks the real P-arrival. This is because the recurrent structure of the CRNN can naturally store all historical waveform information from the beginning, which can be used to combat noise. In contrast, the U-Net method only obtains the waveform information in a certain period before and after, through down- and up-sampling, lacking the ability to identify abrupt local changes.

In summary, it can be found that the CRNN method has the highest accuracy, the lowest MAE values, and a strong anti-interference ability for both P- and S-waves. The U-Net method is inferior to the CRNN method slightly in various indicators, which is relatively similar but superior to the STA/LTA and AIC methods. For the two traditional methods, the AIC method has higher accuracy but poorer stability, and is sensitive to noise. The STA/LTA method has lower picking accuracy but better stability, which is consistent with the existing research results (Shang et al., 2018).

4.4. Application in field failure monitoring

In view of the analysis in Section 4.3, during the MS monitoring for rock-mass stability of the YBT Project construction, the CRNN method has been used to pick P- and S-arrivals in real-time. Then the Newton second order, Newton downhill method (Li and Chen, 2013), was used to locate the rock mass microfracture in order to evaluate the stability of the rock mass.

During the monitoring process, rock mass microfracture signals occurred in the junction area between the shaft and the second gallery on March 19, 2022. As of March 31, 183 rock mass microfracture events had been accumulated in this area. A cloud diagram of the spatial distribution density is shown in Fig. 14. Fig. 15 shows the evolution of the microfracture events with time, implying that rock mass microfractures with moment magnitudes (M_w) mainly ranged from -1 to -0.5 , and continued to occur in this area, especially on March 28. According to on-site feedback, three faults (F85, F88 and F21) pass through this area, leading to poor integrity of the rock mass in the faults and the adjacent areas. The maximum in situ stress in the fault area is 12.5–18.8 MPa, and the maximum in situ stress between the faults with good rock mass integrity is 37.6 MPa. On March 30 and 31, 2022, multiple failures occurred in the area between the faults, which were mainly stress-type spalling failures. The spalled rocks are mainly thin blocky or thin flaky rocks, with fresh and unfilled failure surfaces, as shown in Fig. 16. It shows that, in terms of weak microfracture events with $M_w < -0.5$, the positioning results based on the arrivals picked by the CRNN method have a good consistency with the actual failures, providing a good support for engineering construction.

To analyze the positioning differences of the various methods, the CRNN, U-Net, AIC and STA/LTA methods are used to pick P- and S-arrivals, and 183 rock mass microfracture events in the shaft area are recorded. The measured coordinates of the surface failure are used to approximate the actual location of the failure occurring within the rock mass. The deviation vectors are defined as directed line segments that start from the center of the failure area and end at the MS positioning results, as shown in Fig. 17. Fig. 18 shows the sets of deviation vectors calculated by each method. In terms of the macroscopic distribution of MS sources, all four methods show a trend of roughly distributing along the XOZ plane of the failure area. Meanwhile, most microseismic events are distributed above the second gallery, rather than being symmetrically distributed along the failure center. This is because that this area was undergoing pilot drilling upwards before failure occurred, and the incubation of microfracture inside the rock mass promoted the occurrence of apparent failure. Among these four methods, the deviation vectors of the CRNN method are the most concentrated and closest to the failure center, followed by the U-Net method. Although the deviation vectors of the AIC and STA/LTA methods are also concentrated in the failure area overall, they are more scattered than those of the

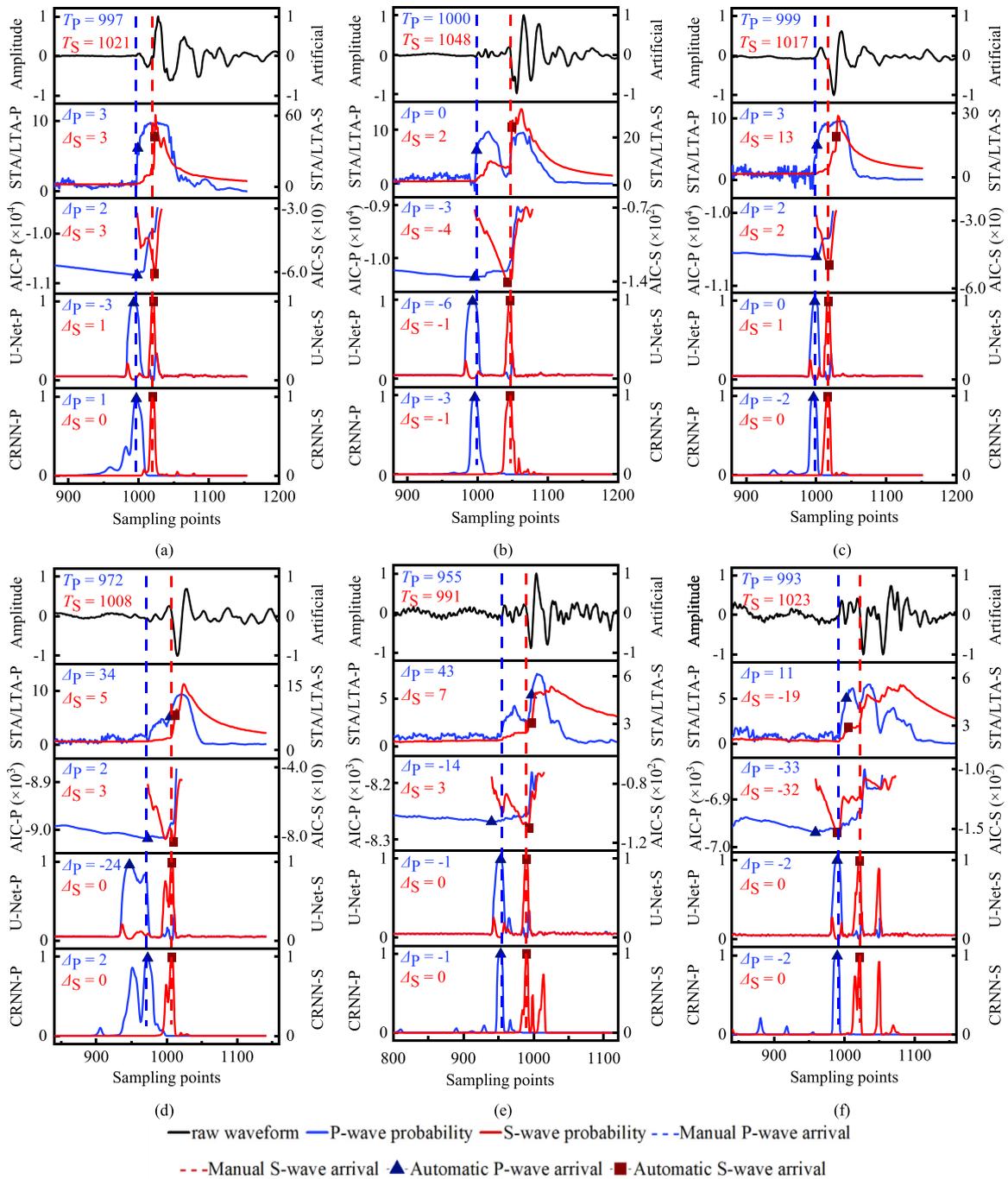


Fig. 13. The arrival picking results of rock mass microfracture signals with different SNRs: (a) SNR = 25.4 dB; (b) SNR = 19.1 dB; (c) SNR = 15.1 dB; (d) SNR = 8.0 dB; (e) SNR = 5.1 dB; and (f) SNR = 2.2 dB. Δ_p and Δ_s are the difference between the picked P- and S-arrivals, respectively, and the manually picked arrivals. Negative numbers represent an advance, and positive numbers represent a delay.

two deep learning methods. The positioning results of the AIC method are the most divergent, which is significantly affected by SNR. As shown in Fig. 19, the deep learning methods pick up the arrivals at the same precision regardless of the SNR. The STA/LTA and AIC methods have a good performance with high SNR, while have an unstable trend with low SNR. The difference is that the STA/LTA method tends to pick up P-waves with hysteresis so that the P-

and S-arrivals overlap, as shown in Fig. 19c and d. The overlapping results in positioning algorithms based on the arrival time differences positioning the source near the sensor array. However, the AIC method tends to pick up P-waves in advance, as shown in Fig. 19a, leading to significant arrival time differences, seriously affecting positioning accuracy and even causing positioning divergence.

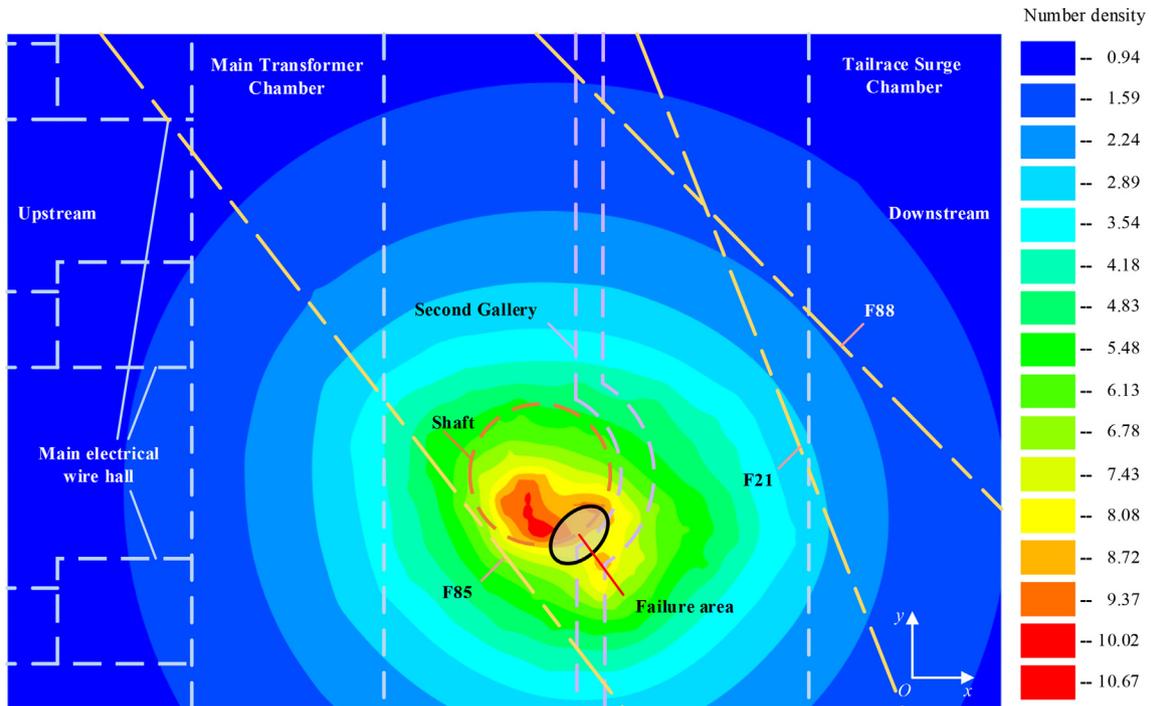


Fig. 14. Density nephogram of microfracture events based on the CRNN method (March 19–31, 2022).

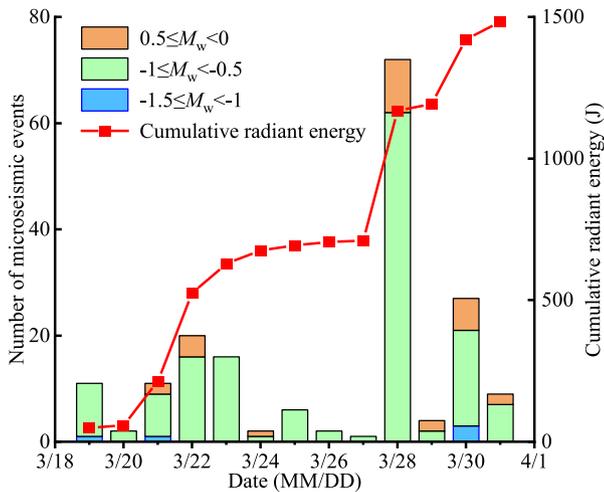


Fig. 15. Evolution trend of microfracture events with time based on the CRNN method (March 19–31, 2022).



Fig. 16. Rock mass failure at the junction area between the shaft and the second gallery.

Table 6 presents the average distances of the 183 microfracture events from the calculated location to the failure center. The average distance of the CRNN method is 28.79 m, which is the smallest among the four methods, indicating that the CRNN method picks arrivals stably and reflects the on-site failures well. The U-Net method is slightly worse than the CRNN method. The average distances of the AIC and STA/LTA methods are greater than 50 m, larger than that of the deep learning methods. In addition, the average distances of all four methods show a trend that can be characterized as the order of Z-axis > X-axis > Y-axis. This is because the MS sensors were mainly distributed in the first gallery

due to the spatial structure and construction progress when failure occurred. The microfracture events were mostly outside the sensor array in Z-axis direction, leading to greater positioning error in Z-axis direction (Feng et al., 2015; Chen et al., 2021).

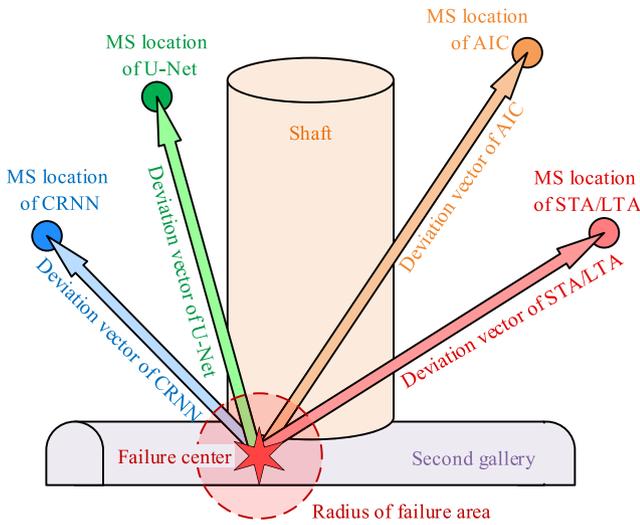


Fig. 17. Schematic of deviation vectors.

5. Conclusions

To address the common contradiction between accuracy and computation in the current arrival picking methods, a real-time arrival picking method based on CRNN is proposed. The hyperparameter sensitivities, computational performance and arrival picking accuracy of the proposed method are analyzed based on synthetic and field MS data, and the proposed method has been applied to the YBT Project for MS monitoring. The conclusions are drawn as follows:

- (1) The hyperparameters have a significant impact on the performance of the CRNN model. An increase of hyperparameters can effectively improve the arrival picking accuracy at the cost of more computations. However, blindly increasing the hyperparameters only shows a limited improvement and may even cause over-fitting issue. The hyperparameters with the most significant impacts on arrival picking ability and computational cost are the convolution kernel width and GRU units, respectively.
- (2) The computational performance of the CRNN method is better than that of the U-Net method, showing average decreases of 30% and 25% in CPU usage and average time per

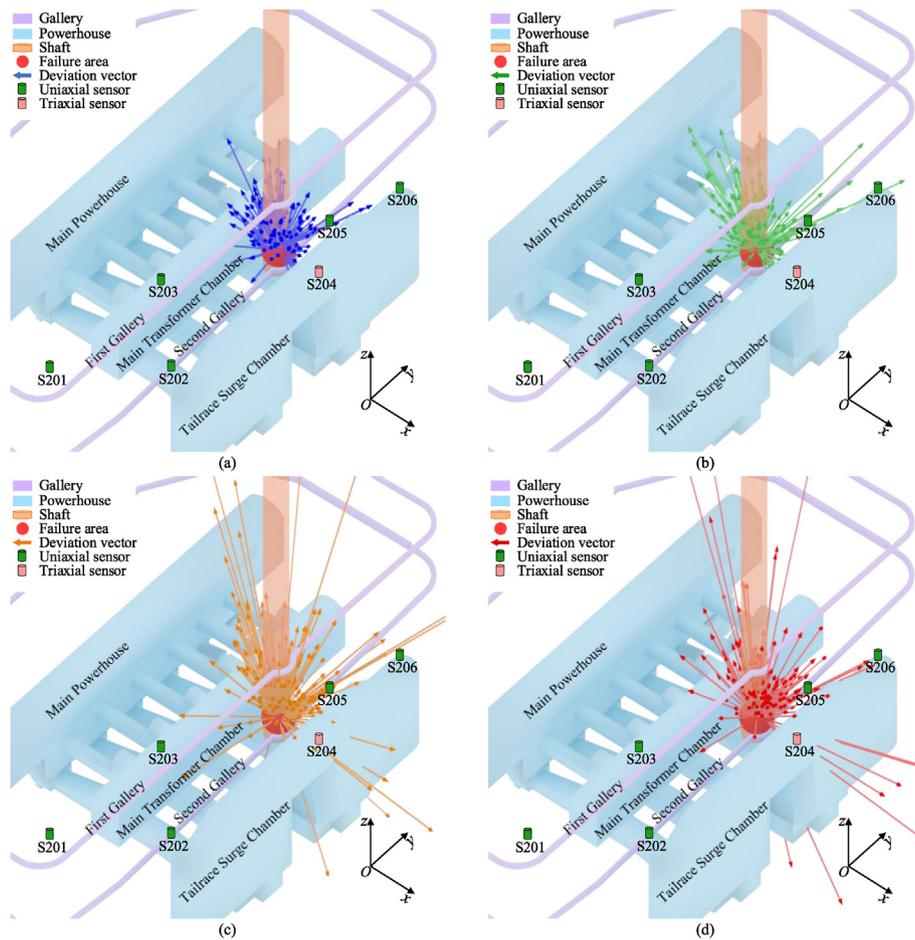


Fig. 18. Deviation vectors from the failure center to the positioning results: (a) CRNN method; (b) U-Net method; (c) AIC method; and (d) STA/LTA method. S201–S206 are the MS sensors arranged near the failure area.

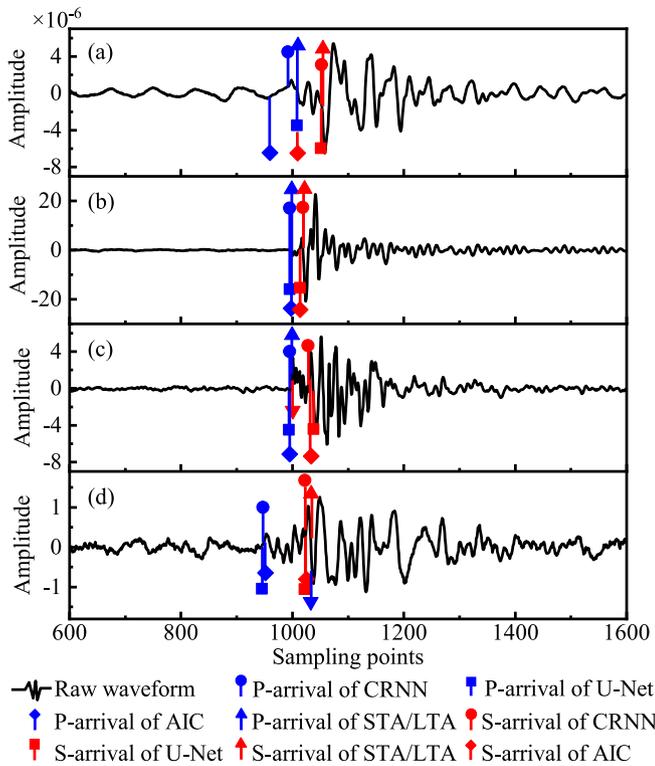


Fig. 19. Arrival picking results for different waveforms of the same microfracture: (a)–(d) are the waveforms of the same microfracture monitored by different MS sensors.

Table 6
The average distances from the calculated location results to the failure center.

Method	Straight-line distance (m)	X-axis distance (m)	Y-axis distance (m)	Z-axis distance (m)
CRNN	28.79	12.31	8.87	21.34
U-Net	30.51	14.02	9.34	21.73
AIC	52.19	23.60	21.78	32.02
STA/LTA	50.66	25.61	13.54	33.62

batch, respectively, implying a broader applicability of the CRNN method.

- Compared with the STA/LTA, AIC and U-Net methods, the CRNN method has the highest picking accuracy, the lowest MAE values, and a strong anti-interference ability for both P- and S-waves. Within 4 sampling points, the accuracy of the CRNN method is 87.44% for P-arrival and 91.29% for S-arrival, both of which exceed 96% within 16 sampling points, and the MAE_{SUM} is only 4.22 points, followed by the U-Net, AIC and STA/LTA methods.
- In the application of YBT Project, the positioning results based on the arrivals picked by the CRNN method show the best consistency with the actual failure zone, followed by those of the U-Net, STA/LTA and AIC methods, thus proving the reliability of the CRNN method in engineering applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Acknowledgments

We acknowledge the funding support from National Natural Science Foundation of China (Grant No. 42077263).

References

- Akram, J., Eaton, D.W., 2016. A review and appraisal of arrival-time picking methods for downhole microseismic data. *Geophysics* 81 (2), Ks71–Ks91.
- Allen, R., 1982. Automatic phase pickers - their present use and future-prospects. *Bull. Seismol. Soc. Am.* 72 (6), S225–S242.
- Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.* 68 (5), 1521–1532.
- Bahdanau, D., Cho, K., Bengio, Y., 2014. Neural Machine Translation by Jointly Learning to Align and Translate arXiv:1409.0473.
- Chen, B.R., Li, T., Zhu, X.H., Wei, F.B., Wang, X., Xie, M.X., 2021. Microseismic source location method based on a velocity model database and statistical analysis. *Arabian J. Geosci.* 14, 2017.
- Cho, K., van Merriënboer, B., Bahdanau, D., Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches arXiv:1409.1259.
- Cook, N.G.W., 1976. Seismicity associated with mining. *Eng. Geol.* 10 (2–4), 99–122.
- Dickson, M.C., Bosman, A.S., Malan, K.M., 2022. Hybridised Loss Functions for Improved Neural Network Generalisation arXiv:2204.12244.
- Dong, X.T., Li, Y., Wu, N., Tian, Y.A., Yu, P.J., 2018. The S-STK/LTK algorithm for arrival time picking of microseismic signals. *J. Geophys. Eng.* 15 (4), 1484–1491.
- Durrheim, R.J., Spottiswoode, S.M., Roberts, M.K.C., Brink, A.V., 2005. Comparative seismology of the Witwatersrand Basin and Bushveld Complex and emerging technologies to manage the risk of rockbursting. *J. South. Afr. Inst. Metall.* 105 (6), 409–416.
- Feng, G.L., Feng, X.T., Chen, B.R., Xiao, Y.X., Jiang, Q., 2015. Sectional velocity model for microseismic source location in tunnels. *Tunn. Undergr. Space Technol.* 45, 73–83.
- Feng, G.L., Feng, X.T., Xiao, Y.X., et al., 2019. Characteristic microseismicity during the development process of intermittent rockburst in a deep railway tunnel. *Int. J. Rock Mech. Min. Sci.* 124, 104135.
- Fu, J., Wang, X., Li, Z., et al., 2020. Automatic phase-picking method for detecting earthquakes based on the signal-to-noise-ratio concept. *Seismol. Res. Lett.* 91 (1), 334–342.
- Galli, S., 2022. Python Feature Engineering Cookbook: over 70 Recipes for Creating, Engineering, and Transforming Features to Build Machine Learning Models. Packt Publishing Ltd., Birmingham, UK.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press, Cambridge, USA.
- Guo, C., Zhu, T.Y., Gao, Y.T., Wu, S.C., Sun, J., 2021. AEnet: automatic picking of P-wave first arrivals using deep learning. *IEEE Trans. Geosci. Rem. Sens.* 59 (6), 5293–5303.
- Hedayat, A.S., Sloane, N.J.A., Stufken, J., 1999. *Orthogonal Arrays: Theory and Applications*. Springer Science & Business Media, New York, USA.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780.
- Lan, Z.Q., Gao, P., Wang, P., Wang, Y.J., Liang, J.D., Hu, G.M., 2022. Automatic first arrival time identification using fuzzy C-Means and AIC. *IEEE Trans. Geosci. Rem. Sens.* 60, 5907613.
- Leonard, M., Kennett, B.L.N., 1999. Multi-component autoregressive techniques for the analysis of seismograms. *Phys. Earth Planet. In.* 113 (1–4), 247–263.
- Li, B., Xu, N.W., Xiao, P.W., et al., 2023a. Microseismic monitoring and forecasting of dynamic disasters in underground hydropower projects in southwest China: a review. *J. Rock Mech. Geotech. Eng.* <https://doi.org/10.1016/j.jrmge.2022.10.017>.
- Li, P.X., Chen, B.R., Xiao, Y.X., Feng, G.L., Zhou, Y.Y., Zhao, J.S., 2023b. Rockburst and microseismic activity in a lagging tunnel as the spacing between twin TBM excavated tunnels changes: a case from the Neelum-Jhelum hydropower project. *Tunn. Undergr. Space Technol.* 132, 104884.
- Li, Q.P., Chen, B.R., 2013. Research on micro-seismic source location during linear excavation process of deep tunnel. In: *Rock Characterisation, Modelling and Engineering Design Methods—The 3rd ISRM SINOROCK Symposium*, pp. 690–695. Shanghai, China.

- Li, T., Chen, B.R., Zhu, X.H., 2021. Research on the propagation mechanism of the arrival time error and wave velocity error under sensor array. *IOP Conf. Ser. Earth Environ. Sci.* 861, 052056.
- Li, X.B., Shang, X.Y., Morales-Esteban, A., Wang, Z.W., 2017. Identifying P phase arrival of weak events: the Akaike Information Criterion picking application based on the Empirical Mode Decomposition. *Comput. Geosci.* 100, 57–66.
- Lim, J., Jung, S., JeGal, C., et al., 2022. LEQNet: light earthquake deep neural network for earthquake detection and phase picking. *Front. Earth Sci.* 10, 848237.
- Ma, H.T., Wang, T., Li, Y., Meng, Y.Q., 2018. A time picking method for microseismic data based on LLE and improved PSO clustering algorithm. *Geosci. Rem. Sens. Lett. IEEE* 15 (11), 1677–1681.
- Maeda, N., 1985. A method for reading and checking phase time in auto-processing system of seismic wave data. *J. Seismol. Soc. Jpn.* 38, 365–379.
- Mao, H.Y., Xu, N.W., Li, X., et al., 2023. Analysis of rockburst mechanism and warning based on microseismic moment tensors and dynamic Bayesian networks. *J. Rock Mech. Geotech. Eng.* <https://doi.org/10.1016/j.jrmge.2022.12.005>.
- Mousavi, S.M., Ellsworth, W.L., Zhu, W.Q., Chuang, L.Y., Beroza, G.C., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* 11 (1), 3952.
- Nippres, S.E.J., Rietbrock, A., Heath, A.E., 2010. Optimized automatic pickers: application to the ANCORP data set. *Geophys. J. Int.* 181 (2), 911–925.
- Pardo, E., Garfias, C., Malpica, N., 2019. Seismic phase picking using convolutional networks. *IEEE Trans. Geosci. Rem. Sens.* 57 (9), 7086–7092.
- Peterson, C., Söderberg, B., 1989. A new method for mapping optimization problems onto neural networks. *Int. J. Neural Syst.* 1 (1), 3–22.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Munich, Germany, pp. 234–241.
- Ross, Z.E., Meier, M.A., Hauksson, E., 2018. P wave arrival picking and first-motion polarity determination with deep learning. *J. Geophys. Res. Solid Earth* 123 (6), 5120–5129.
- Saragiotis, C.D., Hadjileontiadis, L.J., Panas, S.M., 2002. PAI-S/K: a robust automatic seismic P phase arrival identification scheme. *IEEE Trans. Geosci. Rem. Sens.* 40 (6), 1395–1404.
- Shang, X.Y., Li, X.B., Morales-Esteban, A., Dong, L.J., 2018. An improved P-phase arrival picking method S/L-K-A with an application to the Yongshaba Mine in China. *Pure Appl. Geophys.* 175 (6), 2121–2139.
- Tan, Y.Y., He, C., 2016. Improved methods for detection and arrival picking of microseismic events with low signal-to-noise ratios. *Geophysics* 81 (2), KS93–KS111.
- Xiao, Y.X., Feng, X.T., Hudson, J.A., Chen, B.R., Feng, G.L., Liu, J.P., 2016. ISRM suggested method for in situ microseismic monitoring of the fracturing process in rock masses. *Rock Mech. Rock Eng.* 49 (1), 343–369.
- Xu, H.Y., Zhao, Y., Yang, T.H., Wang, S.H., Chang, Y.Q., Jia, P., 2022. An automatic P-wave onset time picking method for mining-induced microseismic data based on long short-term memory deep neural network. *Geomatics, Nat. Hazards Risk* 13 (1), 908–933.
- Young, R.P., Collins, D.S., 2001. Seismic studies of rock fracture at the underground research laboratory, Canada. *Int. J. Rock Mech. Min. Sci.* 38 (6), 787–799.
- Zhang, J.L., Sheng, G.Q., 2020. First arrival picking of microseismic signals based on nested U-Net and Wasserstein Generative Adversarial Network. *J. Pet. Sci. Eng.* 195, 107527.
- Zhang, S.C., Tang, C.A., Wang, Y.C., Li, J.M., Ma, T.H., Wang, K.K., 2021a. Review on early warning methods for rockbursts in tunnel engineering based on microseismic monitoring. *Appl. Sci.* 11 (22), 10965.
- Zhang, W., Feng, X.T., Bi, X., et al., 2021b. An arrival time picker for microseismic rock fracturing waveforms and its quality control for automatic localization in tunnels. *Comput. Geotech.* 135, 104175.
- Zhang, X., Chen, H.H., Zhang, W., Tian, X., Chu, F.D., 2021c. Generalized neural network trained with a small amount of base samples: application to event detection and phase picking in downhole microseismic monitoring. *Geophysics* 86 (5), KS95–KS108.
- Zhao, J.S., Chen, B.R., Jiang, Q., et al., 2022. Microseismic monitoring of rock mass fracture response to blasting excavation of large underground caverns under high geostress. *Rock Mech. Rock Eng.* 55 (2), 733–750.
- Zheng, J., Lu, J., Peng, S.P., Jiang, T.Q., 2018. An automatic microseismic or acoustic emission arrival identification scheme with deep recurrent neural networks. *Geophys. J. Int.* 212 (2), 1389–1397.
- Zhou, H., Zhang, W., Zhang, J., 2016. Downhole microseismic monitoring for low signal-to-noise ratio events. *J. Geophys. Eng.* 13 (5), 805–816.
- Zhu, D., Li, Y., Zhang, C., 2016. Automatic time picking for microseismic data based on a fuzzy C-Means clustering algorithm. *Geosci. Rem. Sens. Lett. IEEE* 13 (12), 1900–1904.
- Zhu, M.B., Cheng, J.Y., Wang, L.G., 2019. The effect of P- and S-phase arrivals on the accuracy of microseismic source location. *J. Appl. Geophys.* 164, 130–139.
- Zhu, W.Q., Beroza, G.C., 2019. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.* 216 (1), 261–273.
- Zhu, X.H., Chen, B.R., Wang, X., Li, T., 2022. Time series segmentation clustering: a new method for S-Phase picking in microseismic data. *Geosci. Rem. Sens. Lett. IEEE* 19. <https://doi.org/10.1109/LGRS.2021.3098809>.



Dr. Bing-Rui Chen is a Professor of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, and University of Chinese Academy of Sciences. He obtained his BSc degree and PhD from Northeastern University, China. He is a Vice Chairman of Crustal Stress and Earthquake Committee of Chinese Society for Rock Mechanics and Engineering (CSRME). He is also an Editorial Board Member of *Gold Science and Technology*, and Young Editorial Board Member of *Tunnelling Construction*. He won one Second Prize of National Science and Technology Progress Award (Ranking 3), one First Prize of Technological Invention, one Patent Gold Award and one Patent Silver Award in Hubei Province and one Patent Excellence Award in China (Ranking 1). His research interests are microseismic monitoring technology, analysis and early warning methods, and mitigation of rock engineering disasters.