Microseismic Signal Classification Based on Artificial Neural Networks

Chong-wei Xin, Fu-xing Jiang, and Guo-dong Jin

1Civil and Resource Engineering School, University of Science and Technology Beijing, Beijing 100083, China
2Beijing Anke Industrial Technology Co., Ltd., Beijing 100083, China
3Frontiers Science Center for Deep Ocean Multispheres and Earth System, Key Lab of Submarine Geosciences and Prospecting Techniques, MOE and College of Marine Geosciences, Ocean University of China, Qingdao 266100, China

Correspondence should be addressed to Chong-wei Xin; xinchw2000@163.com

Received 22 October 2020; Revised 4 June 2021; Accepted 8 July 2021; Published 28 July 2021

Academic Editor: Salvatore Caddemi

Copyright © 2021 Chong-wei Xin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The classification of multichannel microseismic waveform is essential for real-time monitoring and hazard prediction. The accuracy and efficiency could not be guaranteed by manual identification. Thus, based on 37310 waveform data of Junde Coal Mine, eight features of statistics, spectrum, and waveform were extracted to generate a complete data set. An automatic classification algorithm based on artificial neural networks (ANNs) has been proposed. The model presented an excellent performance in identifying three preclassified signals in the test set. Operated with two hidden layers and the Logistic activation function, the multiclass area under the receiver operating characteristic curve (AUC) reached 98.6%.

1. Introduction

As a real-time, continuous, and dynamic method, microseismic monitoring has been widely used in coal mines at home and abroad in recent years [1–3]. The microseismic monitoring system mainly monitors rock stability by analyzing the microseismic events produced by mining activities [4, 5]. Due to the complex underground environment, the signals collected by microseismic monitoring systems usually include rock fracture signals, blasting vibration signals, mechanical drilling signals, and electromagnetic noise signals [6–9]. Traditional manual classification is greatly influenced by subjective factors, which is not only inefficient but also inaccurate [10]. Therefore, it is necessary to develop an automatic microseismic signal classification system to realize the intelligence and reliability of microseismic monitoring.

Researchers at home and abroad have carried out practical research on the identification of microseismic or seismic signals. The most widely used method was the short-term average/long-term average (STA/LTA) algorithm proposed by Allen [11], which calculated a short-term and a long-term average of the magnitude of the signal to recognize earthquakes from single traces. This method has been significantly improved by many researchers. Song et al. [12] proposed the array-based waveform correlation approach to increase the detectability of small magnitude events near the master event detected by the STA/LTA algorithm. To improve the confidence of arrival-time picking in low signal-to-noise ratio (SNR) condition, Akram and Eaton [13] combined the ratio of the peak eigenvalues (PER) and STA/LTA to enhance signal coherency. However, the problem of STA/LTA algorithm is that it is sensitive to the background noise and window size is crucial for the accuracy of the picking results. Yoon et al. [14] developed a Fingerprint And Similarity Thresholding (FAST) to detect earthquakes using waveform similarity. The performance of FAST was comparable to that of autocorrelation with less detection time. However, the method produced false positives. It is widely accepted that the accuracy of classical algorithms is not satisfactory.
Contrary to the classical algorithms, machine learning based on data driving has been paid more attention. Dai and MacBeth [15] utilized three-component recordings as the input of artificial neural networks (ANNs) and constructed a discriminant function determined from the output of the ANN to pick seismic arrivals. Wang and Teng [16] designed two types of artificial neural detector (AND) for earthquake detection used STA/LTA series and moving window spectrums as input data, respectively, with better performance than the algorithm based on STA/LTA. Ursino et al. [17] established a three-layer ANN to discriminate quarry blasts and earthquakes automatically. Tiira [18] calculated four STA/LTA values for seven frequency bands of 193 teleseismic events, which were used as the input of the network to pick seismic arrivals. Gentili and Michelini [19] designed particular neural trees called IUANT2 to detect P- and S-wave onset time automatically. The results showed that for P- and S-waves, the standard deviation of the difference between manual and automatic picking is 0.064 s and 0.11 s. Madureira and Ruano [20] tested two types of neural networks: multilayer perceptrons (MLPs) and support vector machines (SVMs), with a performance between 98% and 100%. In recent years, deep learning has been gradually used to identify microseismic or seismic signals with the improvement of computer performance. Wiszniowski et al. [21] applied a real-time recurrent neural network (RTRN) to detect seismic signals in the frequency domain as well as in the phase arrival times. Ross et al. [22] applied a convolutional neural network (ConvNet) on labeled data from the Southern California Seismic Network to recognize seismic body wave phases, which is extremely sensitive and robust. Zhu and Beroza [23] proposed a deep-neural-network-based arrival-time picking method called PhaseNet that achieved higher precision and recall rate than traditional methods. The training set of PhaseNet included more than 700,000 waveform samples over 30 years of earthquake recordings. Zhou et al. [24] developed a hybrid algorithm using both convolutional and recurrent neural networks. The convolutional neural network (CNN) was used to detect events while recurrent neural networks (RNNs) were used to pick P and S arrival times. The CNN could detect over 95% of the hand-picked events with a signal-to-noise ratio of greater than 100 and around 80% of the events with a signal-to-noise ratio of 1. Mousavi et al. [25] designed a detector based on deep neural networks named CNN-RNN Earthquake Detector (CRED) with a precision of greater than 96% testing on the held-back data. Zheng et al. [26] developed a recurrent neural network (RNN) to identify experiments regarding AE signals in a laboratory. The accuracy to find fracturing events was between 70% and 80% depending on the signal-to-noise ratio. However, the results showed that CNN overfit the small data set. Wilkins et al. [27] trained a CNN using hand-labeled, multichannel microseismic data from a coal mine to detect events. The results showed that the CNN outperformed the human microseismic expert in picking more true events and in eliminating more false events.

Previous research results are fruitful, but there are still some issues: (1) the algorithms were designed for seismic signal detection in most cases. (2) The SNR of microseismic signals was much smaller than seismic signals. (3) The microseismic signals were polluted by the electromagnetic interference signals, mechanical drilling signals, and blasting vibration signals. (4) The arrivals of all microseismic signals were similar because the sensors were spatially quite close. (5) The algorithms were complex and lacked real-time performance. The above features make the detection of microseismic signals more challenging than seismic signals. Therefore, in this paper, a classification algorithm for microseismic signals based on ANN has been proposed. The statistical features, spectral features, and waveform features were extracted to divide the multi-channel signals into three classes. Class 1 was high-quality microseismic signals, which can be used to pick up arrival-time data accurately of P-wave and S-wave, to determine the source with high reliability; Class 2 is the low-quality microseismic signals, which were difficult to pick up the arrival-time data; Class 3 is interference signals, such as electromagnetic interference signals, mechanical drilling signals, and pure background noise signals. Because explosives were limited in the studied mine, the explosion signals were not included in the study. By classifying the signals collected from the mine, we could find the activity law of microseismic events and identify high-quality microseismic signals, which could provide a more exact arrival time for fine positioning.

2. Methodology

In this section, a method based on ANN is proposed, which can be used to classify mine microseismic signals. The method mainly includes three parts: extraction of discriminant features, construction of classification model, and evaluation of classifier performance. The process of the methodology is shown in Figure 1.

2.1. Discriminant Features. Data and features determine the upper limit of machine learning, and models and algorithms only approach this upper limit. Feature engineering is the cornerstone for building high-performance classification models. Therefore, according to the characteristics of microseismic signals, cluttered signals, and noise, three main aspects of features are extracted, including statistical features, spectral features, and waveform features.

2.1.1. Statistical Features

Skewness. In probability theory and statistics, skewness (Sk) is a measure of the variable’s asymmetry to its mean. It is a shape parameter that characterizes the degree of asymmetry [28, 29]. Skewness can be positive, negative, and zero. Under normal conditions, the microseismic records are approximately symmetrically distributed concerning the y-axis, and skewness is approximately zero. For a complete microseismic record, the skewness can be defined as
Set of inputs → Feature extraction → Training process of ANN → Application

Seismograms → Spectrum features → Training samples → ANN → Classification

Waveform features → Test samples → Model → Prediction

Figure 1: Classification steps of microseismic signals. The green, blue, brown-yellow, and wine red correspond to steps (1), (2), (3), and (4), respectively. (1) Set of inputs; (2) feature extraction, including statistical features, spectrum features, and waveform features; (3) training of ANN; (4) application of the classification model.

\[
Sk = E \left( \frac{(X - \mu)^3}{\sigma^4} \right), \quad (1)
\]

where \(X\) is the microseismic record, \(\mu\) and \(\sigma\) are the mean value and standard deviation of the microseismic record, respectively, and \(E\) is the mathematical expectation. When considering a microseismic signal of \(n\) samples, represented as \(x = \{x_1, x_2, \ldots, x_n\}\), the discrete form of equation (1) is

\[
Ku = \left( \frac{1}{n-1} \right) \sum_{i=1}^{n} \left( \frac{x_i - \mu}{\sqrt{\sum_{i=1}^{n} (x_i - \mu)^2 / (n-1)}} \right)^3. \quad (2)
\]

② Kurtosis. In probability theory and statistics, kurtosis (Ku) is a shape parameter measure of the relative peakedness of a given distribution [28, 30]. In general, the intensity and duration of standard microseismic signals at a coal mine are always in a particular range; the kurtosis is also in a particular range accordingly. The kurtosis of a microseismic signal can be defined as

\[
Ku = E \left( \frac{(X - \mu)^3}{\sigma^4} \right). \quad (3)
\]

For a series of \(n\) samples \(x_i\), Ku is expressed as

\[
Ku = \left( \frac{1}{n-1} \right) \sum_{i=1}^{n} \left( \frac{x_i - \mu}{\sqrt{\sum_{i=1}^{n} (x_i - \mu)^2 / (n-1)}} \right)^4. \quad (4)
\]

2.1.2. Spectral Features

① Peak Area Ratio. A Fourier transform can convert time-domain signals into frequency-domain signals to analyze the characteristics from another aspect. Peak frequency is expressed as the frequency of the peak amplitude in the spectrum. Peak area ratio (PAR) represents the proportion of spectrum window with peak frequency inside, as shown in the pink band in Figure 2. PAR is defined as

\[
PAR = \frac{\int_{f_{ub}}^{f_{ub}} A^2 df}{\int_{f_{lb}}^{f_{ub}} A^2 df}, \quad (5)
\]

where \(A\) is frequency-domain amplitude, \(f_{ub}\) and \(f_{lb}\) are upper and lower bounds of frequency-domain, respectively, and \(f_{ub}\) and \(f_{lb}\) are the upper and lower limits of the window with peak frequency inside, respectively.

② Low Area Ratio. Unlike seismic records, the frequency band of typical microseismic records is broad and generally above 10 Hz [31]. The low-frequency part can be regarded as the characteristic of a standard microseismic signal. Low area ratio (LAR) represents the proportion of low-frequency parts below the limit value. LAR is defined as

\[
LAR = \frac{\int_{f_{ub}}^{f_{ub}} A^2 df}{\int_{f_{lb}}^{f_{ub}} A^2 df}, \quad (6)
\]

where \(f_{ub}\) is the upper-frequency limit.

③ Approximate Signal-to-Noise Ratio. Spectrum estimation is a physically explicit method of calculating approximate signal-to-noise ratio (SNR). The method assumes that the microseismic signal has a specific range below or above, which is judged to be noise; the noise is distributed uniformly in the frequency band, as shown in the yellow band in Figure 2(b). It should be noted that the result calculated by this method is an approximate signal-to-noise ratio, but this error can be tolerated for the fast operation of the program. In specific calculations, the signal is considered to be above a certain threshold (red horizontal line in Figure 2), and the noise is considered below this threshold. Approximate SNR is defined as

\[
SNR = \frac{\sum_{f_{lb}}^{f_{ub}} A^2 df}{\int_{f_{lb}}^{f_{ub}} A^2 df}, \quad (7)
\]

where \(f_{lb}\) and \(f_{ub}\) are lower and upper limits of the signal band, respectively.

2.1.3. Waveform Features

① Envelope Peak. We propose a variable resolution envelope algorithm for microseismic waveforms. When the resolution is high, the envelope can describe the details more accurately; when the resolution is low, the influence of noise or interference on the whole signal can be reduced. The waveform envelope calculated from each signal segment with 20 raw microseismic data points is presented in Figure 3. Envelope peak (EP) is defined as the number of peaks
on a continuous envelope. The gray area in Figure 3(b) is a complete envelope peak.

2) Ratio of Maximum to Average of Envelope. Ratio of the maximum to the average value of envelope (RMAVE) is an important index to characterize the signal strength. In general, the smaller the disturbance and the stronger the vibration, the larger the RMAVE.

3) Correlation Coefficients. For two microseismic records $X_1$ and $X_2$, the covariance is defined as

$$\text{cov}(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)],$$

where $E$ is the mathematical expectation and $\mu_i = E(X_i)$.

Cross-correlation is a standard method to estimate the degree of correlation between two sequences. Consider two microseismic records $X_1$ and $X_2$, and the Correlation coefficient (Corr) is defined as

$$\text{Corr} = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{cov}(X_1, X_1) \text{cov}(X_2, X_2)}}.$$  (9)

2.2. Classification Algorithm. An artificial neural network (ANN), which is inspired by the biological neural network in nature, is a mathematical model composed of a large number of simple interconnected nodes (called neurons) to generate a solution for special issues. Because of its robustness and nonlinearity, ANN is widely used in classification and regression fields. An advantage of a neural net is that it can model various response surfaces given enough hidden nodes. The neural network we build is shown in Figure 4. It contains an input layer, two hidden layers, and an output layer, while each layer is connected with adaptive weights. In this paper, there are eight neurons in input layer corresponding to the eight features proposed in Section 2.1. We divide the signals into three categories, namely, Class 1, Class 2, and Class 3, corresponding to the binarized $[1,0,0]$, $[0,1,0]$, and $[0,0,1]$. For the above reasons, there are three neurons in the output layer.

In this paper, a typical backpropagation (BP) neural network is used to train and classify microseismic signals into Class 1, Class 2, and Class 3. For any hidden node $j_1$, the output can be calculated as

$$H_{j_1} = f\left(\sum_{i} W_{ij_1} I_i - y_{j_1}\right).$$  (10)

where $f$ is the activation function, $N$ is the number of input layer nodes, $W_{ij_1}$ is the weight connecting input node $i$ to hidden node $j_1$, $I_i$ is the output of input node $i$, and $y_{j_1}$ is the threshold of the hidden node $j_1$. The calculation of $H_{j_2}$ and $O_k$ for hidden layer $j_2$ and output layer is the same as equation (10) with their weights, respectively.

Without activation function, the neural networks would act as a linear regression model with limited performance, and it would not be enough to deal with complex nonlinear tasks. Activation functions commonly used are the Logistic, hyperbolic tangent (Tanh), and rectified linear unit (ReLU) [32]. The details of the activation functions are shown in Figure 5.

In the backpropagation to update weights, the error function $E$ is calculated from the nodes in output layers to the nodes in the first hidden layer [16, 18, 33]. As is shown in equation (11), the cumulative square error of the outputs compared to the desired outputs is proportional to the error function. The Logistic activation function is taken as an example to illustrate the specific process of backpropagation.

![Figure 2: Waveform and spectrum of microseismic signal. (a) Waveform with a sampling frequency of 1000 Hz. (b) Spectrum of signal. $f_{uf}$ is the upper-frequency limit in LAR calculation, and the corresponding green band is the spectrum window of low frequency. The pink band is the window with peak frequency inside, with $f_{ubw}$ and $f_{lbw}$, which are the upper and lower limits of the window. The yellow band indicates the signal part judged by the threshold, with $f_{sl}$ and $f_{su}$ which are lower and upper limits of the band.](image)
The weight value $W_{ij}$ is updated with the derivative of error to the weight during the back propagation:

$$E = \frac{1}{2} \sum_{k=1}^{Q} (\hat{O}_K - O_K)^2,$$

$$\Delta W_m^{ij} = -\eta \frac{\delta E}{\delta W_{ij}} = \eta \delta_j O_i^{m-1},$$

where $Q$ is the number of output layer nodes, $\hat{O}_K$ is the predicted value, $O_K$ is the real value, $\Delta W_m^{ij}$ is the change of weight, $\eta$ is the learning rate, $O_i^{m-1}$ is the output of a node in the previous layer, $\delta_j$ is the error term, and $m$ donates the current layer.

After mathematical derivation [18], the error term $\delta_j^m$ for a node in output layer is

$$\delta_j^m = (O_j - \hat{O}_j^m) \hat{O}_j^m (1 - \hat{O}_j^m).$$

The error term $\delta_j^m$ for a node in hidden layer is
2.3. Classifier Performance. For a binary classification problem, the combination of the observed category and the predicted category determined by machine learning can be divided into four categories, as listed in Table 1. True positive (TP) means the predictive value is positive (P) while the observed value is positive (P'), false positive (FP) means the predicted value is positive (P) while the observed value is negative (N'), false negative (FN) means the predicted value is negative (N) while the observed value is positive (N'), and true negative (TN) means the predicted value is negative (N) while the observed value is negative (N'). Once a decision threshold is selected, the true positive ratio (TPR) and false positive ratio (FPR) are selected as metrics to evaluate the classifier with the contingency matrix [34, 35]:

\[
\delta_j^m = \tilde{O}_j^m (1 - \tilde{O}_j^m) \sum_{h=1}^N \delta_h^m W_{hj},
\]

where \( W_{hj} \) is the weight connecting the current layer to the next layer and \( N \) is the number of the nodes of the next layer.

The neural network is trained with random initial weights. In the process of training, each weight will be updated according to the principle of backpropagation. This process will be repeated until the error is less than the set threshold or the maximum number of iterations is reached.

Because AUC is part of the unit square, the value of AUC is always between 0 and 1. AUC is evaluated by the following criteria: 0.9-1.0 excellent, 0.8-0.9 good, 0.7-0.8 fair, 0.6-0.7 poor, and 0.5-0.6 failure [36–38].

As for the multiclassification problem, the contingency matrix is shown in Table 2. When dealing with m-class classification problems, each class corresponds to a ROC curve. If \( C \) is the set of all classes, the ROC curve \( i \) evaluates classification performance using class \( c_i \) as positive class and all other classes as negative class [37]:

\[
P_i = c_i; \\
N_i = \bigcup_{j \neq i} c_j \in C.
\]

Then, the multiclass AUC is calculated as [37, 39]

\[
AUC_{\text{total}} = \sum_{c_i \in C} AUC(c_i) \cdot p(c_i),
\]

where \( AUC(c_i) \) is calculated from ROC curve \( i \) and \( p(c_i) \) is the weight of the reference class’s prevalence in the data, and it can be determined by calculating the probability of the occurrence of class \( i \) observed [37, 40].

3. Experimental Study

3.1. Geological Conditions and Monitoring Schemes. No.1702 working face is located in the 3-4 district of third horizontal level at Junde Coal Mine. Comprehensive light-duty coal mining technology is used to extracted mine underground. The effective strike and sloping lengths of the working face are 1080 m and 176 m, respectively. The maximum buried depth is ~412 m, and the thickness of coal seam is 10.08 m–14.93 m (the average is 11.93 m). The dip angle is from 28° to 34°, with an average of 32°.

The KJ551 microseismic monitoring system developed by Beijing Anke Technology Co., Ltd. was installed at Junde Coal Mine. The system could collect and filter signals continuously, locate automatically, and calculate energy. The bandwidth of a single uniaxial sensor was 0.1 to 600 Hz, the sampling rate was 1000 Hz, and the collected signals were transmitted by cable. It should be noted that the trigger

![Figure 5: Functions and plots of activation functions: (a) Logistic; (b) Tanh; (c) ReLU.](image-url)

| Name   | Function          | Plot |
|--------|-------------------|------|
| Logistic | \( f(x) = (1/1 + e^{-x}) \) | ![Graph](image-url) |
| Tanh   | \( f(x) = (e^x - e^{-x})/(e^x + e^{-x}) \) | ![Graph](image-url) |
| ReLU   | \( f(x) = \max(0, x) \) | ![Graph](image-url) |
mode was used to record events by the monitoring system. Therefore, the electrical interference signals and mechanical interference signals would be recorded inevitably. As shown in Figure 6, No. 1702 working face was surrounded by ten sensors, of which five were in the headentry, and the other five were in the tailentry. The spacings of adjacent sensors range from 50 to 70 m.

3.2. Dataset Description. The seismic signals (from 09/01/2019 to 09/30/2019) recorded by the KJ551 microseismic monitoring system contain a total of 37,310 waveforms with all features evaluated. All waveforms are manually labeled as Class 1 for high-quality microseismic signals, Class 2 for low-quality microseismic signals, and Class 3 for noise and other interfering signals, of which 17,662 belong to Class 1, 11,211 belong to Class 2, and the rest 8,437 belong to Class 3. After the binarization of labels, Class 1, Class 2, and Class 3 are defined as [1,0,0], [0,1,0], and [0,0,1], respectively. Figures 7–10 are typical representatives of the three types of signals, while subgraph Figure 7(a) is a waveform, subgraph Figure 7(b) is a spectrum, and subgraph Figure 7(c) is a time-frequency spectrum.

We utilize MATLAB to obtain the original waveform data features, which are received by the sensors. The values of mean, standard deviation (std.), minimum, first quartile (Q1), median, third quartile (Q3), and maximum for features are listed in Table 3. The distribution of features varies widely. For instance, the standard deviation ranges from 0.007 to 48.762, and the latter is 6966 times the former. Z score standardization is select to pretreat the input data to avoid the asymptotic effect:

$$x_{ij}' = \frac{x_{ij} - \bar{x}_i}{\sigma_i}$$  \hspace{1cm} (19)

where $x_{ij}$ is the j-th value of the i-th variable and $\bar{x}_i$ and $\sigma_i$ are the mean and standard deviation of the i-th variable, respectively.

Figure 11 shows the distribution of features by class after normalization. Obviously, there are significant differences in the distribution, with Sk, EP, and RMAVE the most obvious. However, it is still challenging to classify signals based on these characteristics linearly. Therefore, there is a high need for ANN for classification accurately.

3.3. Model Building. We divided all the samples into two categories, 70% of the training set and 30% of the test set. The former is used to train the model, and the latter is mainly used to test the generalization ability of the model. The initial learning rate is set to 0.01, the maximum number of iterations is set to 500, and the tolerance for the optimization is set to 0.0001, which means when the training loss does not improve more than tolerance for two consecutive epochs, the training stops. The number of hidden layers is designed as double the number of input nodes plus 1 [41]. Because eight microseismic features are extracted, the number of nodes in the hidden layer is set to 17. Generally speaking, when the number of hidden layers is 2, any decision boundary with arbitrary precision can be represented with an appropriate activation function, and any smooth mapping with any precision can be fitted. Therefore, we study the influence of the combination of activation functions and hidden layer numbers on training loss.

3.4. Results and Discussions. In the case of the Junde Coal Mine dataset, the main objective of the model is to identify high-quality microseismic signals by dividing the multichannel signals into three classes. By eliminating the pollution of electromagnetic interference signals, mechanical drilling signals, and blasting vibration signals, we can determine the occurrence law of pure microseismic events in the mine and identify high-quality microseismic signals for fine positioning. The relationship between the loss and the training iterations is shown in Figure 12 under different combinations of hidden layer numbers and activation functions. It can be observed that the loss of a neural network with two hidden layers is less than that with one hidden layer. The best combination is two hidden layers and the Logistic activation function. Under this operating condition,
Figure 6: The plane sketch of No. 1702 working face and sensor layout.

Figure 7: Class 1 signal. (a) Waveform. (b) Spectrum. (c) Time-frequency spectrum.
the loss is 0.299. The reason why we did not study ANN with three hidden layers and above is that the deeper the number of layers, the stronger the ability of classification, and the better the effect is in theory. However, the deeper the number of layers may lead to overfitting, and at the same time, it will increase the difficulty of training and make the model difficult to converge. As we can see from Figure 12, the loss of one hidden layer and two hidden layers is 0.343 and 0.299 with Logistic activation function, 0.333 and 0.311 with Tanh activation function, and 0.330 and 0.315 with ReLU activation function, respectively. We can find that the loss reduction from one hidden layer to two hidden layers is limited.

Through the analysis above, the final model is built with two hidden layers and Logistic activation function. The contingency matrix and performance matrix for testing samples are listed in Tables 4 and 5. Furthermore, the ROC curve is presented in Figure 13. The precision (TP/(TP + FP)), recall (TP/(TP + FN)), and $F_1((1/\text{precision}+1/\text{recall})/2)$ of three classes are all greater than 0.95. The AUC of Class 1, Class 2, and Class 3 is 0.987, 0.986, and 0.982, respectively, and the AUC$_{\text{total}}$ is 0.986.

Figure 8: Class 2 signal. (a) Waveform. (b) Spectrum. (c) Time-frequency spectrum.
Inspection of Tables 4 and 5 and Figure 13 provides the following:

1. The ANN model presents an excellent classification of signals during the test using mine data.
2. According to the results of the test set, it can be deduced that the ANN model generalizes extremely well.
3. There are still a small number of Class 3 signals identified as Class 1 signals.
4. The performance of correctly identifying each type of signal is similar.

The several advantages of the proposed ANN model are as follows:

1. The model can classify the collected signals in real-time.
2. The model minimizes classification errors, which can provide a solid foundation for the analysis and positioning of mine personnel.
3. The model is convenient to train and deploy, and it can serve the actual monitoring work of mine to the maximum.

Figure 9: Class 3 signal. (a) Waveform. (b) Spectrum. (c) Time-frequency spectrum.
Figure 10: Class 3 signal. (a) Waveform. (b) Spectrum. (c) Time-frequency spectrum.

Table 3: Values of mean, standard deviation (std.), minimum, first quartile (Q1), median, third quartile (Q3), and maximum for features.

| Variables | Mean   | Std.  | Min.    | Q1     | Median  | Q3     | Max.   |
|-----------|--------|-------|---------|--------|---------|--------|--------|
| Sk        | −0.019 | 0.833 | −32.937 | −0.227 | 0.032   | 0.230  | 56.835 |
| Ku        | 37.947 | 48.762| 1.273   | 7.355  | 26.095  | 50.573 | 3527.524|
| PAR       | 0.008  | 0.007 | 0.001   | 0.005  | 0.006   | 0.009  | 0.185  |
| LAR       | 0.007  | 0.013 | 0       | 0.003  | 0.005   | 0.007  | 0.659  |
| SNR       | 0.908  | 0.856 | 0.025   | 0.341  | 0.622   | 1.178  | 15.723 |
| EP        | 1.078  | 0.521 | 0       | 1      | 1       | 1      | 9      |
| RMAVE     | 24.328 | 40.042| −47.688 | 4.450  | 10.139  | 24.111 | 625.243|
| Corr      | 0.401  | 0.262 | −0.012  | 0.176  | 0.330   | 0.615  | 0.981  |
Figure 11: Box graph of discriminant features normalized by equation (19).

Table 4: Contingency matrix of 3-class prediction problem on test set.

|       | Predicted Class 1 | Predicted Class 2 | Predicted Class 3 |
|-------|-------------------|-------------------|-------------------|
| Class 1 | 5127              | 63                | 28                |
| Class 2 | 131               | 3197              | 47                |
| Class 3 | 84                | 37                | 2479              |

Table 5: Performance matrix on test set.

|       | Precision | Recall | F1   | Support | AUC  | AUC_total |
|-------|-----------|--------|------|---------|------|-----------|
| Class 1 | 0.96      | 0.98   | 0.97 | 5218    | 0.987|           |
| Class 2 | 0.97      | 0.95   | 0.96 | 3375    | 0.986| 0.986     |
| Class 3 | 0.97      | 0.95   | 0.96 | 2600    | 0.982|           |
The model has great guiding significance for the signal classification of new mines without monitoring records and can help them form their own unique classification system.

4. Conclusions

In this paper, a method based on ANN for microseismic signals classification has been presented. The performance of classifier has been trained and tested in microseismic records for one month of the Junde Mine. The features Sk, Ku, PAR, LAR, SNR, EP, RMAVE, and Corr can be extracted automatically. This method can divide multichannel microseismic signals into three classes. In the case of 2 hidden layers and Logistic activation function, the AUC_total reached 98.6%. The method has the advantages of accuracy, efficiency, and operability. Moreover, this method lays a solid foundation for real-time monitoring and intelligent classification and provides a prerequisite for the high-quality positioning of microseismic events.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are thankful to the following organizations who provided the data basis and technical support for this research: Longmei Mining Group Company Limit and Beijing Anke Technology Co., Ltd. This research was supported by the National Natural Science Foundation of China: 51634001, rock burst monitoring and early warning principles and methods with electromagnetic-microseismic coupling.

References

[1] N. W. Xu, F. Dai, Z. Z. Liang, Z. Zhou, C. Sha, and C. A. Tang, “The dynamic evaluation of rock slope stability considering the effects of microseismic damage,” Rock Mechanics and Rock Engineering, vol. 47, no. 2, pp. 621–642, 2014.

[2] V. Kumar, N. Gopalakrishnan, N. P. Singh, and S. Cherukuri, “Microseismic monitoring application for primary stability evaluation of the powerhouse of the Tapovan Vishnugad hydropower project,” Journal of Earth System Science, vol. 128, no. 6, p. 169, 2019.

[3] S. Qiao, Q. Zhang, and Q. Zhang, “Mine fracturing monitoring analysis based on high-precision distributed wireless microseismic acquisition station,” IEEE Access, vol. 7, pp. 147215–147223, 2019.

[4] P. Wang, L.-s. Jiang, P.-q. Zheng, G.-p. Qin, and C. Zhang, “Inducing mode analysis of rock burst in fault-affected zone with a hard-thick stratum occurrence,” Environmental Earth Sciences, vol. 78, no. 15, p. 467, 2019.

[5] X. Li, E. Wang, Z. Li, Z. Liu, D. Song, and L. Qiu, “Rock burst monitoring by integrated microseismic and electromagnetic radiation methods,” Rock Mechanics and Rock Engineering, vol. 49, no. 11, pp. 4393–4406, 2016.

[6] Z.-q. Yin, Z.-x. Hu, Z.-d. Wei et al., “Assessment of blasting-induced ground vibration in an open-pit mine under different rock properties,” Advances in Civil Engineering, vol. 2018, Article ID 4603687, 10 pages, 2018.

[7] P. Peng, Z. He, and L. Wang, “Automatic classification of microseismic signals based on MFCC and GMM-HMM in
underground mines,” *Shock and Vibration*, vol. 2019, Article ID 5803184, 9 pages, 2019.

[8] Z. Zhou, R. Cheng, X. Cai, D. Ma, and C. Jiang, “Discrimination of rock fracture and blast events based on signal complexity and machine learning,” *Shock and Vibration*, vol. 2018, Article ID 9753028, 10 pages, 2018.

[9] Y. Xu, P. Yuan, and Z. Zheng, “Time-frequency analyses of blasting vibration signals in single-hole blasting model experiments,” *Journal of Vibroengineering*, vol. 19, pp. 363–375, 2017.

[10] C. He, P. Peng, and Z. Liao, “An automatic identification and classification method of complex microseismic signals in mines based on Mel-frequency cepstral coefficients,” *Journal of Safety Science and Technology*, vol. 14, pp. 41–47, 2018.

[11] R. V. Allen, “Automatic earthquake recognition and timing from single traces,” *Bulletin of the Seismological Society of America*, vol. 68, pp. 1521–1532, 1978.

[12] F. Song, H. S. Kuleli, M. N. Toksoz, E. Ay, and H. Zhang, “An improved method for hydrofracture-induced microseismic event detection and phase picking,” *Geophysics*, vol. 75, no. 6, pp. A47–A52, 2010.

[13] J. Akram and D. W. Eaton, “A review and appraisal of arrival-time picking methods for downhole microseismic data,” *Geophysics*, vol. 81, no. 2, pp. K571–K591, 2016.

[14] C. E. Yoon, O. O’Reilly, K. J. Bergen, and G. C. Beroza, “Earthquake detection through computationally efficient similarity search,” *Science Advances*, vol. 1, no. 11, Article ID e1501057, 2015.

[15] H. Dai and C. MacBeth, “Automatic picking of seismic arrivals in local earthquake data using an artificial neural network,” *Geophysical Journal International*, vol. 120, no. 3, pp. 758–774, 1995.

[16] J. Wang and T.-L. Teng, “Artificial neural network-based seismic detector,” *Bulletin of the Seismological Society of America*, vol. 85, pp. 308–319, 1995.

[17] A. Ursino, H. Langer, L. Scarfi, G. Di Grazia, and S. Gresta, “Discrimination of quarry blasts from tectonic microearthquakes in the Hyblean plateau (southeastern Sicily),” *Annals of Geophysics*, vol. 44, 2001.

[18] T. Tiira, “Detecting teleseismic events using artificial neural networks,” *Computers & Geosciences*, vol. 25, no. 8, pp. 929–938, 1999.

[19] S. Gentili and A. Michelini, “Automatic picking of P and S phases using a neural tree,” *Journal of Seismology*, vol. 10, no. 1, pp. 39–63, 2006.

[20] G. Madureira and A. E. Ruano, “A neural network seismic detector,” *IFAC Proceedings Volumes*, vol. 42, no. 19, pp. 304–309, 2009.

[21] J. Wsziwinski, B. M. Pesiewicz, and J. Trojanowski, “Application of real time recurrent neural network for detection of small natural earthquakes in Poland,” *Acta Geophysica*, vol. 62, no. 3, pp. 469–485, 2014.

[22] Z. E. Ross, M. A. Meier, E. Hauksson, and T. H. Heaton, “Generalized seismic phase detection with deep learning,” *Bulletin of the Seismological Society of America*, vol. 108, no. 5A, pp. 2894–2901, 2018.

[23] W. Zhu and G. C. Beroza, “PhaseNet: a deep-neural-network-based seismic arrival-time picking method,” *Geophysical Journal International*, vol. 216, pp. 261–273, 2019.

[24] Y. Zhou, H. Yue, Q. Kong, and S. Zhou, “Hybrid event detection and phase-picking algorithm using convolutional and recurrent neural networks,” *Seismological Research Letters*, vol. 90, no. 3, pp. 1079–1087, 2019.

[25] S. M. Mousavi, W. Zhu, Y. Sheng, and G. C. Beroza, “CRED: a deep residual network of convolutional and recurrent units for earthquake signal detection,” *Scientific Reports*, vol. 9, no. 1, Article ID 10267, 2019.

[26] J. Zheng, J. Lu, S. Peng, and T. Jiang, “An automatic microseismic or acoustic emission arrival identification scheme with deep recurrent neural networks,” *Geophysical Journal International*, vol. 212, no. 2, pp. 1389–1397, 2018.

[27] A. H. Wilkins, A. Strange, Y. Duan, and X. Luo, “Identifying microseismic events in a mining scenario using a convolutional neural network,” *Computers & Geosciences*, vol. 137, 2020.

[28] S. M. Gandhi and B. C. Sarkar, “Conventional and statistical resource/reserve estimation,” in *Essentials of Mineral Exploration and Evaluation*, 2016.

[29] N. Pérez, P. Venegas, D. Benitez, R. Lara-Cueva, and M. Ruiz, “A new volcanic seismic signal descriptor and its application to a data set from the cotopaxi volcano,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, 2020.

[30] C. Baillard, W. C. Crawford, V. Balu, C. Hibert, and A. Mangeney, “An automatic kurtosis-based P- and S-phase picker designed for local seismic networks,” *Bulletin of the Seismological Society of America*, vol. 104, no. 1, pp. 394–409, 2013.

[31] D. W. Eaton, J. Akram, A. St-Onge, and F. Fournide, “Determining microseismic event locations by semblance-weighted stacking,” in *Proceedings of the CSPG CSEG CWLS Convention*, Calgary, Canada, January 2011.

[32] S. Sharma, S. Sharma, and A. Athaiya, “Activation functions in neural networks,” *International Journal of Engineering Applied Sciences and Technology*, vol. 4, no. 12, pp. 310–316, 2020.

[33] J. Li, J.-h. Cheng, J.-y. Shi, and F. Huang, “Brief introduction of back propagation (BP) neural network algorithm and its improvement,” in *Advances in Computer Science and Information Engineering*, Springer, Berlin, Germany, 2012.

[34] B. Zhao, Q. Dai, D. Han, J. Zhang, L. Zhuo, and M. Berti, “Application of hydrological model simulations in landslide predictions,” *Landslides*, vol. 17, no. 4, pp. 877–891, 2019.

[35] J. A. Vallejos and S. D. McKinnon, “Logistic regression and neural network classification of seismic records,” *International Journal of Rock Mechanics and Mining Sciences*, vol. 62, pp. 86–95, 2013.

[36] X. Shang, X. Li, A. Morales-Esteban, and G. Chen, “Improving microseismic event and quarry blast classification using artificial neural networks based on principal component analysis,” *Soil Dynamics and Earthquake Engineering*, vol. 99, pp. 142–149, 2017.

[37] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[38] S. Belciug, “Logistic regression paradigm for training a single-hidden layer feedforward neural network. Application to gene expression datasets for cancer research,” *Journal of Biomedical Informatics*, vol. 102, Article ID 103373, 2020.

[39] B. Van Calster, V. Van Belle, G. Condous, T. Bourne, D. Timmerman, and S. Van Huffel, “Multi-class AUC metrics and weighted alternatives,” in *Proceedings of the 2008 IEEE International Joint Conference on Neural Networks*, pp. 1390–1396, IEEE World Congress on Computational Intelligence, Hong Kong, China, June 2008.

[40] S. J. Mullen and M. S. Wandishin, “Multi-class ROC analysis,” *Weather and Forecasting*, vol. 24, pp. 530–547, 2009.

[41] R. Hecht-Nielsen, “Kolmogorov’s mapping neural network existence theorem,” in *Proceedings of the International Conference on Neural Networks*, pp. 11–14, IEEE Press New York, San Diego, CA, USA, June 1987.