

## Research Article

# Coal-Rock Recognition in Top Coal Caving Using Bimodal Deep Learning and Hilbert-Huang Transform

Guoxin Zhang,<sup>1,2</sup> Zengcai Wang,<sup>1,2</sup> Lei Zhao,<sup>1,2</sup> Yazhou Qi,<sup>1,2</sup> and Jinshan Wang<sup>1,2</sup>

<sup>1</sup>School of Mechanical Engineering, Shandong University, No. 17923 Jingshi Road, Jinan 250061, China

<sup>2</sup>Key Laboratory of High Efficiency and Clean Mechanical Manufacture, Shandong University, Ministry of Education, No. 17923 Jingshi Road, Jinan 250061, China

Correspondence should be addressed to Zengcai Wang; wangzc@sdu.edu.cn

Received 5 April 2017; Revised 11 June 2017; Accepted 19 June 2017; Published 27 July 2017

Academic Editor: Matteo Filippi

Copyright © 2017 Guoxin Zhang 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.

This study employs the mechanical vibration and acoustic waves of a hydraulic support tail beam for an accurate and fast coal-rock recognition. The study proposes a diagnosis method based on bimodal deep learning and Hilbert-Huang transform. The bimodal deep neural networks (DNN) adopt bimodal learning and transfer learning. The bimodal learning method attempts to learn joint representation by considering acceleration and sound pressure modalities, which both contribute to coal-rock recognition. The transfer learning method solves the problem regarding DNN, in which a large number of labeled training samples are necessary to optimize the parameters while the labeled training sample is limited. A suitable installation location for sensors is determined in recognizing coal-rock. The extraction features of acceleration and sound pressure signals are combined and effective combination features are selected. Bimodal DNN consists of two deep belief networks (DBN), each DBN model is trained with related samples, and the parameters of the pretrained DBNs are transferred to the final recognition model. Then the parameters of the proposed model are continuously optimized by pretraining and fine-tuning. Finally, the comparison of experimental results demonstrates the superiority of the proposed method in terms of recognition accuracy.

## 1. Introduction

Coal is an important source of energy, accounting for approximately 29.21% of primary energy consumption of the world in 2015 according to *BP Statistical Review of World Energy* (June 2016). China both produces and consumes large amount of coal, accounting for 47.70% and 50.01% of global coal production and consumption in the past year, respectively. Approximately 12.84% of coal reserves in the world are distributed in China, of which 44.8% are thick coal seam. Therefore, safe and efficient mining thick coal seam is considerably important. Fully mechanized top coal caving has been widely applied in the mining of thick coal seam due to its safety, high efficiency, high yield, and low production cost.

However, low-level automation and intelligence have always been problems in fully mechanized technology on top coal caving. Particularly, one of the key technologies of caving degrees completely relies on people's judgment.

Relying on artificial vision and auditory in determining the degree of caving is prone to over- and less caving caused by harsh environment, including poor light, coal dust noise, and narrow space. Overcaving and less caving can lead to low recovery rate, decline in coal quality, and increase in cost. In addition, the safety and health of operators are often threatened because they are relatively close to coal-falling areas. Therefore, an accurate and rapid approach of identifying coal-rock is considerably important to control the coal-falling time precisely and improve the automation and intelligence of a fully mechanized top coal caving. These objectives are important in ameliorating the working environment of coal miners, improving coal recovery rate, and reducing production cost.

Since 1960s, more than 20 types of coal-rock interface identification methods and coal-rock recognition methods have been proposed by researchers. The representative coal-rock interface identification methods include detection through artificial  $\gamma$ -ray, natural  $\gamma$ -ray [1–3], and radar [4]. The

artificial  $\gamma$ -ray detection method uses  $\gamma$ -ray backscattering to detect the thickness of the top coal and identify coal-rock interface. However, artificial  $\gamma$ -ray is harmful to human beings, its penetration ability is limited, and it has been gradually abandoned. Natural  $\gamma$ -ray detection measures the intensity of gamma ray in the roof passing through the remaining coal seam and determines the thickness of the top seam according to its attenuation law to identify coal-rock interface. Natural  $\gamma$ -ray detection has been relatively mature for commercialized coal-rock interface recognition. However, when the top coal contains no or low radioactive elements or contains excessive gangue, natural  $\gamma$ -ray detection is no longer applicable. In addition, the cost of this detection is high. Radar detection utilizes the reflection of the electromagnetic wave at the interface of coal gangue to detect the thickness of the top coal. Moreover, the advantages of radar detection include far resolution distance and extensive application range. However, radar detection is no longer applicable when the thickness of the top coal is extremely thick.

The coal-rock identification method recognize coal and rock to determine whether it reached the coal-rock interface. The representative methods include detection through cutting force response [5], image [6, 7], vibration acoustic [8–15], and infrared [16]. Cutting force response detection using drum picks will have different performance cutting coal or rock to identify coal and rock. This method has good adaptability, but it is inapplicable to top coal caving. Image detection uses coal and gangue with different colors, hardness, gloss, and other information to identify coal and rock. However, this method is sensitive to dust and light and needs further research. Infrared detection uses a cutting machine to produce different temperatures when it cuts to coal and rock and determine whether the cut is coal or rock. This method has a quick response and is ideal for real-time application. However, this detection is still immature because of its sensitivity to an ambient temperature.

In recent years, the coal-rock recognition method based on vibration technology has been widely used. Proposed by the US Mining Bureau, the vibration method mainly comprises acoustic method, slot wave seismic method, and mechanical vibration method. Its principle is coal and gangue, which have different frequencies. In [10], the vibration of till beam was analyzed to recognize coal and rock in the top coal caving process. In [11, 12], vibration signals were used to determine whether cutting machine is cutting coal or rock. In [13], the cutting acoustic signal is used to recognize the cutting pattern. In [15], acoustic signals of coal and rock in fully mechanized face were analyzed to rock-coal recognition in top coal caving.

Although significant progress has been made on coal-rock recognition in the past few decades, challenges still remain. Majority of the previous works based on vibration technology independently treated the acoustic or mechanical vibration methods, wherein only the acoustic or mechanical vibration is employed for recognition. Each single modality has been demonstrated to be useful for recognition. However, one single modality alone cannot provide sufficient information on the differences between coal and rock. Therefore,

the method of integrating two modalities to improve the recognition accuracy of rock and coal still needs further investigation.

Several algorithms were proposed to address the representation learning for multiple modalities. In [17], video and audio modality inputs were employed to learn bimodal deep belief networks (DBN). In [18], multimodal deep neural networks (DNN) were proposed to study the correlation between texture and landmark modalities for facial expression reorganization, wherein several stacked autoencoders (AE) were used. In [19], bimodal DNN were used to determine driver fatigue expression. In this study, acoustic and mechanical vibrations are integrated for the first time, and feature statistics and classification are assembled together for coal-rock recognition. Acceleration and sound pressure sensors are employed to detect the vibration and acoustic wave signals. A joint representation layer for recognition is learned from the acoustic and mechanical vibration modalities.

In addition, transfer learning method is adopted, because transferring knowledge of a general object from classification to recognition task has been found to be successful in visually similar categories [20]. In the present study, bimodal transfer DNN (BT-DNN) is initially pretrained by simulating samples (coal and rock hitting the tail beam in the laboratory environment) and irrelevant samples (hand knocking the table), and solving the problem in which massive samples are necessary to train DNN. Then, the proposed model is pretrained by real training samples. Last, supervised fine-tuning is performed. With bimodal learning and transfer learning processes, the proposed method not only has high recognition accuracy but also has the advantages of low cost, extensive adaptability, insensitivity to the environment, and simplified technical difficulty.

Another important step in the recognition is extracting effective features from measured signals. The most commonly applied time-frequency analysis methods include Fourier transform (FT) [21–23], wavelet transform (WT) [24–27], and Hilbert-Huang transform (HHT) [28–31] methods. These methods have their own advantages and application areas. In this work, HHT is utilized to process signals. In addition, five statistics characteristics are computed, which include four characteristics based on intrinsic mode functions (IMFs) and one characteristic based on Hilbert marginal energy spectrum. However, a few extracted features may not provide several contributions for recognition. By contrast, combining different features as input yields an efficient combination of features.

The rest of this paper is organized as follows. Section 2 introduces the used methods. Section 3 describes the recognition system of coal-rock in top coal caving based on the proposed method, including signal acquisition, features extraction, and experiments. Section 4 summarizes this paper and proposes future work.

## 2. Methods

*2.1. Deep Belief Network.* A deep belief network (DBN) is a neural network that contains several layers of restricted Boltzmann machine (RBM), in which the input layer of the

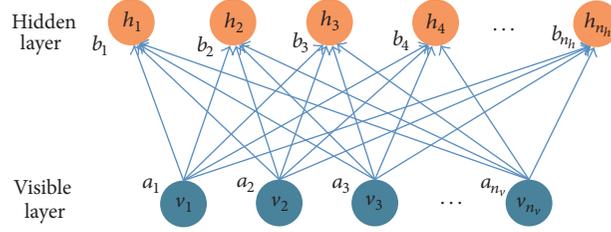


FIGURE 1: Schematic of RBM.

next RBM is the output layer of the previous RBM. RBM was proposed by Smolensky in 1986 [32]; it is a probabilistic graphical model that can be explained by stochastic neural network. RBM is a binary graph in which visible units are connected to hidden units; however, no connections exist between visible-visible units or hidden-hidden units. The visible layer represents observations, and the hidden layer learns features. RBM is applied to numerous machine learning methods due to its desirable properties. In particular after Hinton et al. proposed DBN based on RBM as a basic component [33], RBM has been successfully implemented in dimensionality reduction [34], feature learning [35], and classification [36]. The schematic of RBM is shown in Figure 1.

RBM is an energy based model, and its energy function is as follows:

$$E_{\theta}(v, h) = -\sum_{i=1}^{n_v} a_i v_i - \sum_{j=1}^{n_h} b_j v_j - \sum_{i=1}^{n_v} \sum_{j=1}^{n_h} h_j w_{i,j} v_i, \quad (1)$$

where  $n_v$  and  $n_h$  denote the number of neuron units in the visible and hidden layers, respectively.  $v_i$  and  $h_j$  denote the states of the  $i$ th neuron in the visible layer and the  $j$ th neuron in the hidden layer, respectively.  $a_i$  and  $b_j$  denote the bias of the  $i$ th neuron in the visible layer and the  $j$ th neuron in the hidden layer, respectively.  $w_{i,j}$  is the weight associated with the connection between units  $i$  and  $j$  in the hidden and visible layers, respectively.

Given the independence of the activations of hidden and visible units, the individual activation probabilities are as follows:

$$P(h_j = 1 | v) = \sigma \left( b_j + \sum_{i=1}^{n_v} w_{i,j} v_i \right), \quad (2)$$

$$P(v_i = 1 | h) = \sigma \left( a_i + \sum_{j=1}^{n_h} w_{i,j} h_j \right).$$

In this study, Gaussian-Bernoulli RBM was used, and its energy of joint configuration is

$$E(v, h) = \sum_{i=1}^{n_v} \frac{(v_i - b_i)^2}{2} - \sum_{i=1}^{n_v} \sum_{j=1}^{n_h} w_{i,j} v_i h_j - \sum_{j=1}^{n_h} a_j h_j. \quad (3)$$

Meanwhile, the conditional distribution is

$$p_{\theta}(v_i | h) = N \left( b_i + \sum_{j=1}^{n_h} w_{i,j} h_j, 1 \right), \quad (4)$$

Apart from the preceding differences, Gaussian-Bernoulli RBM is the same as binary RBM.

DBN was pretrained layer by layer by RBM, and back-propagation (BP) algorithm was adopted to fine-tune the entire network for the optimization of all network parameters.

**2.2. Bimodal Learning.** Multimodal learning was proposed by Ngiam et al. in 2011 [17] to learn features over multiple modalities (image and audio modalities). The authors proved that if multiple modalities were present at feature learning, then one modality can be better learned. Since the idea of multimodal learning has been proposed, many researchers have applied this idea and achieved good results [18, 19].

As shown in Figure 2, the proposed bimodal coal-rock recognition system in this study included bimodal DBNs, a joint representation layer, and an output layer. Each DBN had bimodal RBMs, and the architecture of each DBN was determined after testing 100 different architectures. The details on the determination of each DBN and joint representation layer architecture are provided in Section 3.3.1 and Section 3.3.2, respectively.

The training process of this model could be performed in four steps.

First, two DBNs were trained for acceleration and sound pressure, respectively.

Second, two DBNs and a joint representation layer were combined as a joint RBM, which was trained afterward.

Third, the two DBNs, joint RBM, and the output layer were combined to form a DNN. The DNN was subjected to greedily layer-wise training. The training process was still an unsupervised training, which is also called pretraining.

Fourth, this network was fine-tuned to strengthen the recognition capability; this process was a supervised training.

**2.3. Transfer Learning.** Deep networks usually need a large number of training samples to optimize their parameters. The limited labeled sample is the weakness of deep learning. To overcome this difficulty, transfer learning is employed to transfer knowledge from related data. In transfer learning, knowledge obtained from different but related works with sufficient samples is used. Moreover, this method has already achieved some success in the identification field [17, 34].

In this study, whether this transfer learning property of DNN could be generalized to coal-rock was explored. First, samples of simulating coal and rock hitting the tail beam were used to pretrain the DNN model. Second, samples of

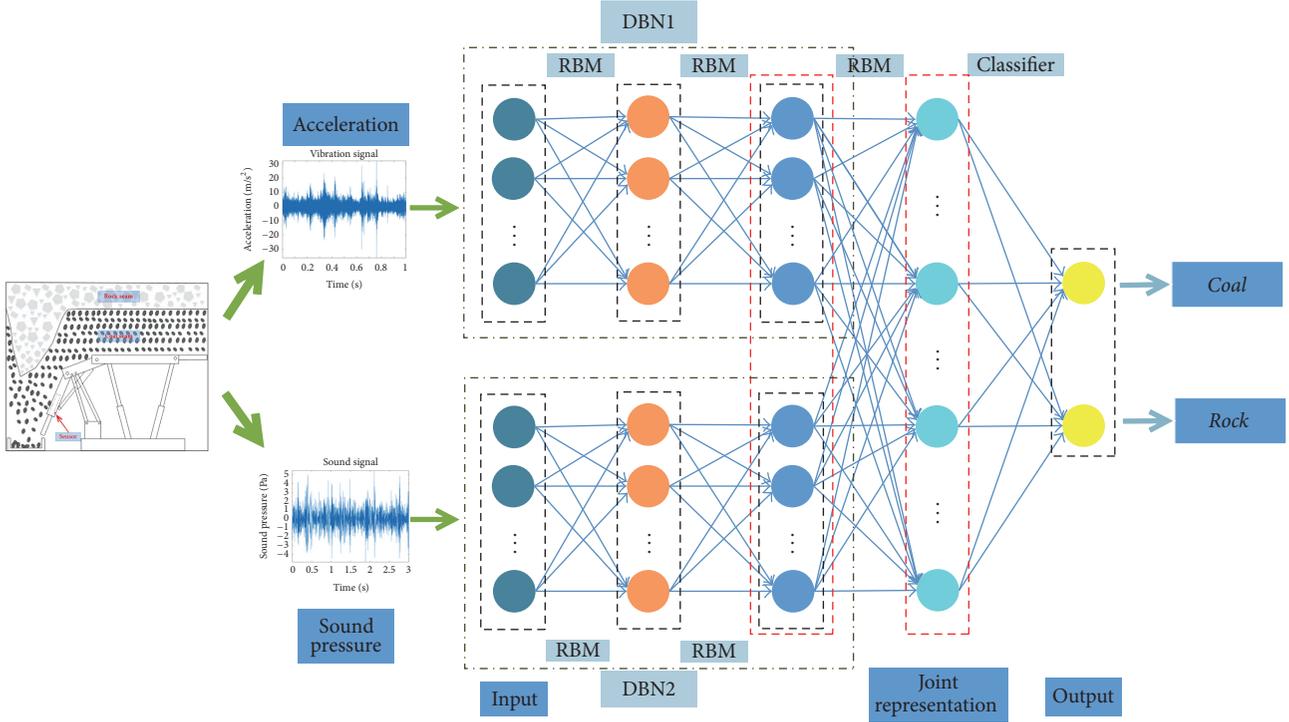


FIGURE 2: Architecture of bimodal learning.

hand knocking the table were used to pretrain the DNN model. Because these samples are also acceleration and sound pressure which belong to the same category with training samples, they played a pretraining effect and optimized the parameters. In the two preceding processes, the bimodal learning method is used. After the unsupervised training, the parameters of DNN are transferred to BT-DNN (Figure 3).

**2.4. Hilbert-Huang Transform.** At present, the most commonly applied time-frequency analysis methods are FT, WT, and Hilbert-Huang transform (HHT). HHT comprises two parts: ensemble empirical mode decomposition (EEMD) and Hilbert spectral analysis (HSA). EEMD was proposed in 2004 [37]. Different from EMD, EEMD solves the mode mixing problem by adding a certain amount of Gaussian white noise to the original signal each time before decomposing [38]. Moreover, EEMD has better performance in nonlinear and nonstationary signals than FT and WT, which are extensively applied in many industries. EEMD was selected in this study due to the strong background noise of the caving environment and the nonlinear and nonstationary measured signals. By using EEMD, several IMFs were obtained:

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t), \quad (5)$$

where  $x(t)$  is the original signal,  $c_i(t)$  is  $i$ th IMF, and  $r_n(t)$  is the trend item.

The obtained IMFs were transformed by Hilbert, and Hilbert energy spectra were obtained.

$$z_i(t) = c_i(t) + jH(c_i(t)) = a_i(t) e^{j\theta_i(t)}, \quad (6)$$

where  $H(c_i(t))$  is the Hilbert transform of  $c_i(t)$ ,  $a_i(t)$  is the amplitude function of  $z_i(t)$ , and  $\theta_i(t)$  is the phase function of  $z_i(t)$ .

Hilbert time-frequency spectra of the signal can be obtained as

$$H(\omega, t) = \text{Re} \sum_{i=1}^n a_i(t) e^{j \int \omega_i(t) dt}, \quad (7)$$

where  $\omega_i(t) = d\theta_i(t)/dt$ .

Finally, Hilbert marginal energy spectra were obtained.

$$E(\omega) = \int_0^T H^2(\omega, t) dt. \quad (8)$$

### 3. Recognition System for the Coal and Rock

Data acquisition, feature extraction, and state recognition combined the intelligent coal-rock recognition system, as shown in Figure 4. Data acquisition comprised the experimental system design, selected sensors, installed sensors, and signal acquisition. Feature extraction comprised signal processing, statistics feature computation, and selection of effective features combination. State recognition consisted of model building, parameter optimization, and pattern recognition.

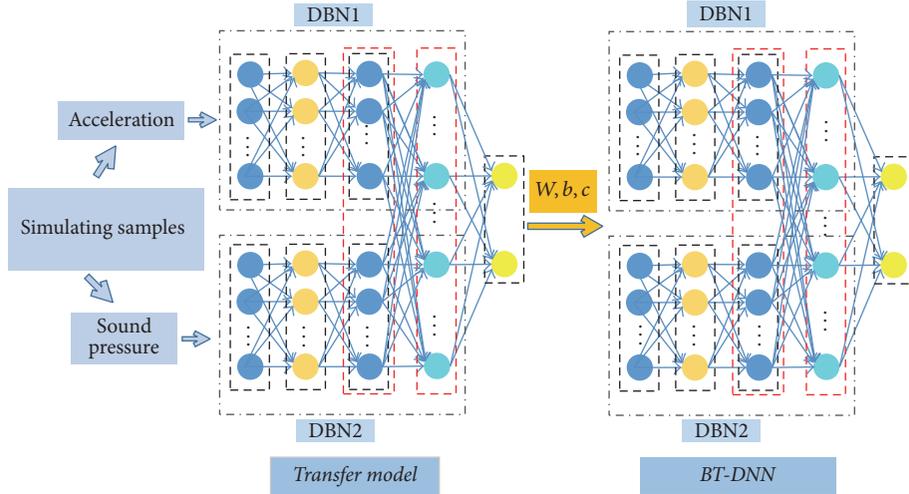


FIGURE 3: The architecture of transfer learning.

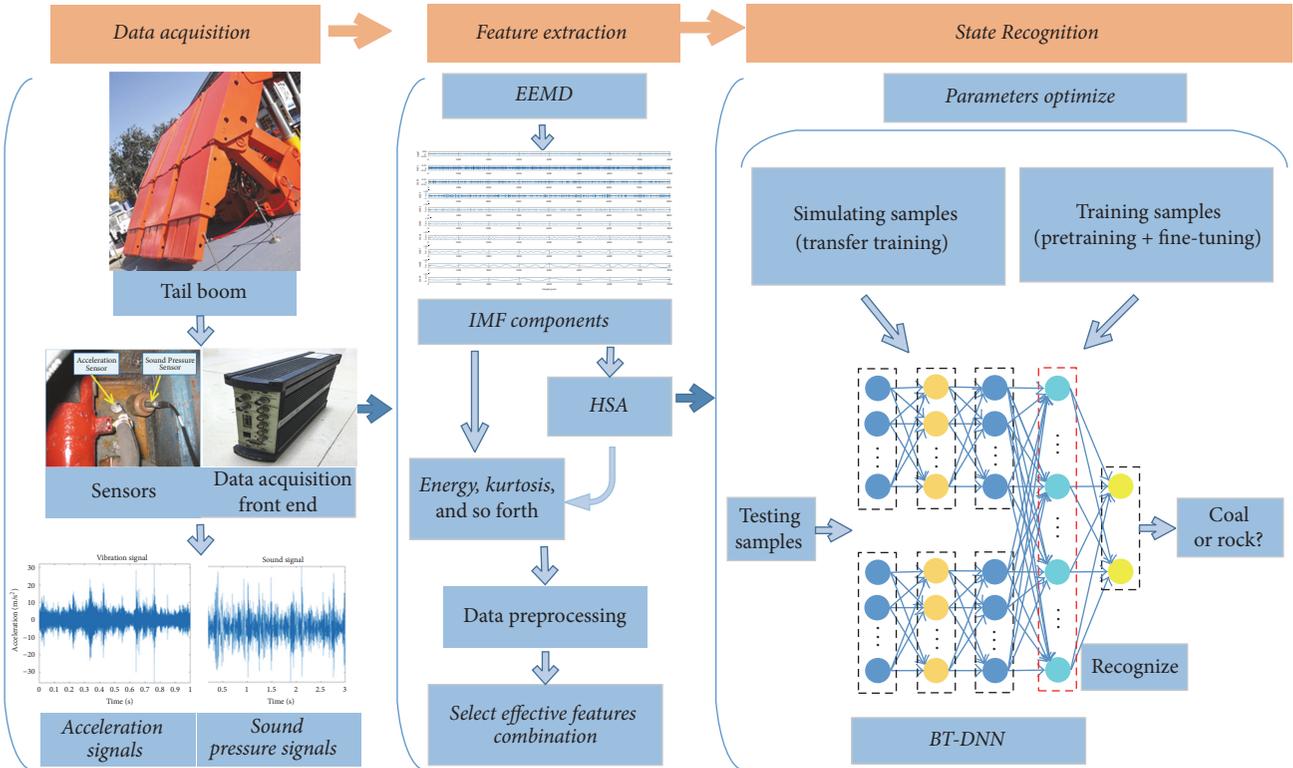


FIGURE 4: The proposed recognition system.

3.1. *Signal Acquisition.* Top coal caving working site and the self-designed experimental system are shown in Figures 5(a) and 5(b), respectively. The acceleration and sound pressure varied when coal and rock hit the tail beam supported by a hydraulic supporter. Thus, identifying coal-rock by detecting the tail beam acceleration and sound pressure is feasible. The sensors were installed on the back of the tail beam due to the following reasons. First, the main impact locations of coal and rock falling are the tail chute and beam. However, if the sensors are installed on the tail chute, then they will be easily

buried by falling coals and gangues. In addition, sensors may be damaged by the continuous spraying of water, which is used for dustproofing. Moreover, the running conveyor chute generates numerous noise that will add to the difficulty of data processing and analysis. Therefore, the ideal location of sensors is at the back of the tail beam. Two sensors were employed, acceleration and sound pressure sensors; their installation method was magnetic base. Sensors and specific installation location are shown in Figures 6(a) and 6(b), respectively. Their parameters are shown in Table 1.

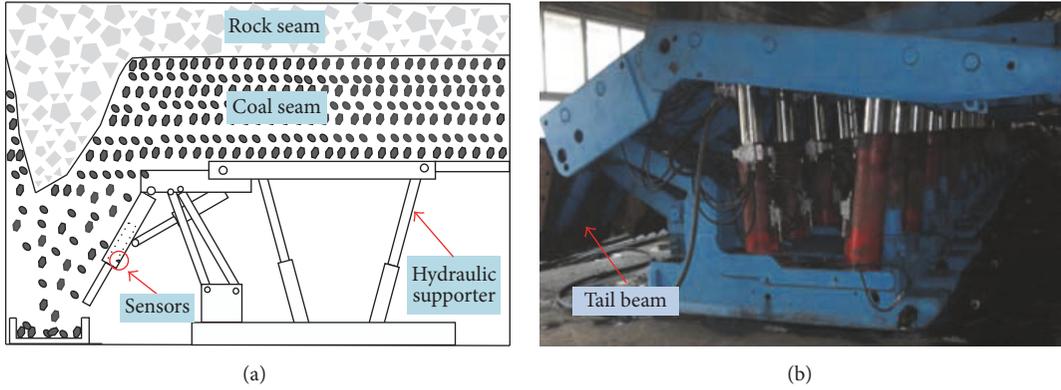


FIGURE 5: (a) Top coal caving working site; (b) self-designed experimental system.

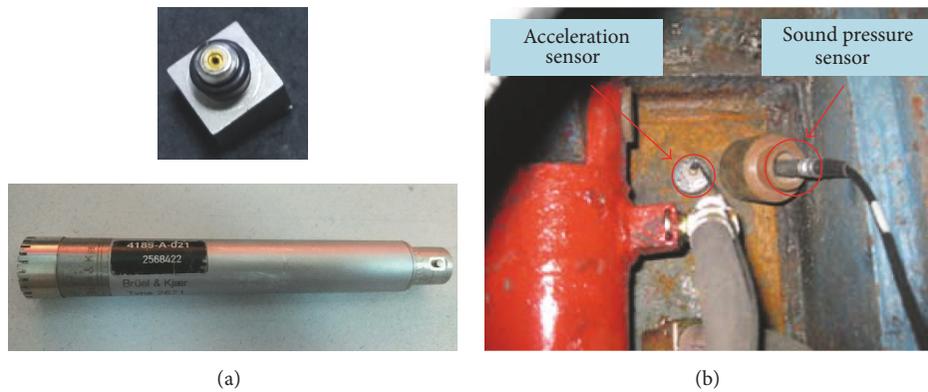


FIGURE 6: (a) Acceleration sensor and sound pressure sensor; (b) specific installation location.

TABLE 1: Parameters of acceleration sensors.

Parameters	Range	Inherent noise
Acceleration sensor	-71 g to 71 g	2 $\mu\text{g}\sqrt{\text{Hz}}$
Sound pressure sensor	16.5 to 134 dB (dynamic)	16.5 dB

Data acquisition system is composed of data acquisition front end, wireless router, and software system. The data collection process is as follows. The acceleration and sound pressure of the tail beam were detected by sensors, obtained by the data acquisition front end, and then transmitted to a notebook by a wireless router. The data were obtained in No. 1306 working platform in Xinglongzhuang Coal Mine. The thickness of coal is 7.34 m to 8.90 m. The caving method is one knife; in one caving, the distance of the caving step is 0.8 m. The frequency of the sampling is set at 65536 Hz; the time of each pattern sampling is 52 s. The sampling time of each sample is determined in Section 3.3.6. The signals are shown in Figure 7.

**3.2. Feature Extraction.** Feature extraction is important in pattern recognition, because selecting effective features significantly contributes to the improvement of recognition accuracy. By using EEMD, several IMFs were obtained, as shown in Figure 8.

The selection of statistics is significantly important for feature extraction. In this study, five statistics characteristics were computed. The four characteristics based on IMFs included skewness, kurtosis, energy, and variance, and one characteristic based on Hilbert marginal energy spectrum is the energy statistics of frequency division. The characteristics were calculated as follows.

(a) Energy based on IMFs is

$$e_i = \int_0^T |x_i(t)|^2 dt. \quad (9)$$

(b) Kurtosis based on IMFs is

$$\text{Kurtosis} = \frac{E(x - \mu)^4}{\sigma^4}. \quad (10)$$

(c) Skewness based on IMFs is

$$\text{skewness} = \frac{E(x - \mu)^3}{\sigma^3}. \quad (11)$$

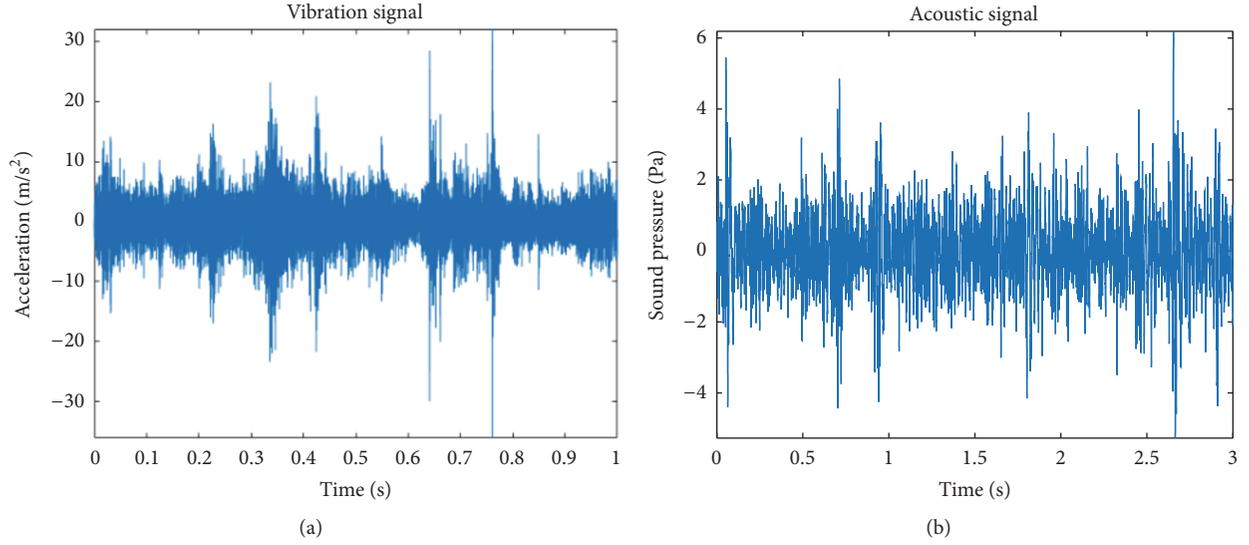


FIGURE 7: Measured signals: (a) acceleration; (b) sound pressure.

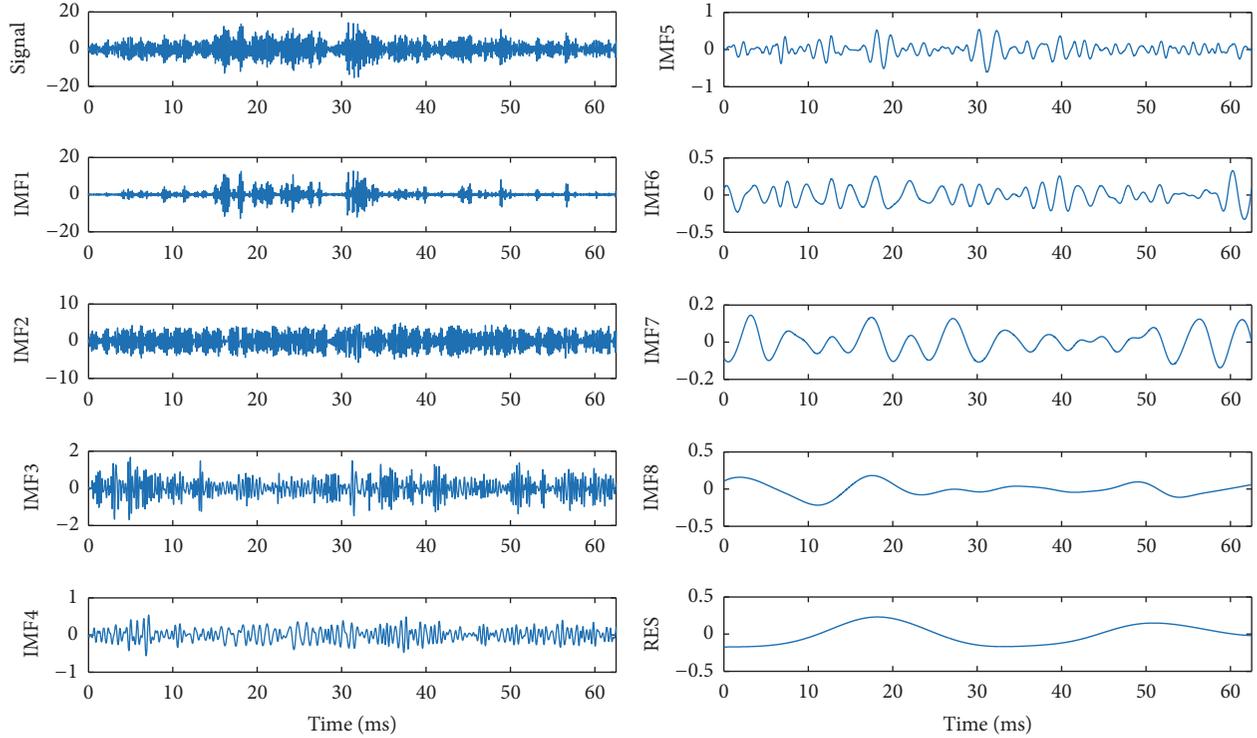


FIGURE 8: IMFs decomposed from measured signal.

(d) Variance based on IMFs is

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N |x_i - \mu|^2. \quad (12)$$

(e) Segmented energy ratio based on Hilbert marginal energy spectrum: the frequency was divided into 10 bands, and

the ratio of each frequency band to the total energy was calculated:

$$E_{Ni} = \frac{E_i}{\sum_{j=1}^{10} E_j}. \quad (13)$$

When using acceleration as the algorithm's only input, we chose  $i$  statistics from the 5 statistics as feature input;  $i$  varied from 1 to 5. After  $\sum_{i=1}^5 C_5^i = 31$  experiments, the most effective combination of statistics for acceleration input is

TABLE 2: Recognition rates with different combinations feature.

Number of acceleration statistics	Number of sound statistics					
	0	1	2	3	4	5
0		66.57%	79.27%	82.37%	81.73%	81.97%
1	69.63%	64.10%	68.00%	72.17%	75.90%	55.43%
2	78.17%	72.40%	74.77%	73.77%	72.33%	68.87%
3	88.90%	96.57%	80.93%	89.07%	82.30%	81.73%
4	90.73%	90.03%	97.93%	89.40%	92.43%	66.23%
5	83.30%	<b>99.13%</b>	88.03%	92.83%	92.13%	74.47%

determined. It is found that when the number of statistics is set to 4, corresponding to kurtosis, energy, variance, and energy statistics of frequency division, the recognition rate reached the maximum value (90.73%) as shown in Table 2. In the same way, we determined the feature when the sound pressure is the algorithm's only input. The corresponding combination of statistics is kurtosis, variance, and energy statistics of frequency division (82.37%).

When the acceleration and sound pressure are set as input together, we chose  $i$  statistics from 5 acceleration statistics and  $j$  statistics from 5 sound pressure statistics as feature input; the total number of combinations is  $\sum_{i=1}^5 C_5^i \times \sum_{j=1}^5 C_5^j = 961$ . After 961 experiments, it was found that when the number of acceleration statistics is set to 5 and the number of sound pressure statistics is set to 1, corresponding to 5 all statistics of acceleration and skewness of sound pressure, the recognition rate reached the maximum value (99.13%). And other comparison experiments use all 5 statistics of acceleration and skewness of sound pressure as feature input.

### 3.3. Experiments

**3.3.1. Comparison among Algorithms with Different DBN Structures.** The network structure, which includes the unit number of the visible and hidden layers, has a profound influence on network performance. A total of 100 combinations were tested for each DBN to seek the optimal structure. The first and second layers of acceleration and sound pressure DBNs both varied from 50 to 500 at intervals of 50. The result of acceleration DBN is shown in Figure 9. Good network performance was observed when the first hidden layer was set from 50 to 150, and the second hidden layer was set at more than 150. The structure of acceleration DBN was set to [42, 100, 250] to build the best recognition deep network. The structure of sound pressure DBN could also be determined in the same manner, with its optimal structure set to [8, 150, 100].

**3.3.2. Comparison among Algorithms with Different Numbers of Joint Representative Units.** The number of joint representative RBM units is considerably important for the proposed model performance, which belongs to the hyperparameter of the network. The number of joint representative units varied from 25 to 1000 to obtain a better performance, results shown in Figure 10.

It shows that the performance gradually and rapidly becomes better initially, then stabilizes, and finally decreases

because of the increasing number of joint representative units. This result could be attributed to having more units resulting in more parameters for learning more information regarding inputs. However, excessive parameters could not be learned, because the number of training samples was limited. If the number of units was too small, then learning the information would be difficult. The network performed best when the unit number was set to 150.

**3.3.3. Comparison of the Algorithms Using Unimodality and Bimodality.** One of the key ideas of this study is integrating acceleration and sound pressure for coal-rock recognition. Thus, algorithms with unimodality and bimodality were compared. Figure 11 is the results of comparison of algorithms using acceleration and sound pressure, respectively. It shows that the method using acceleration (90.73%) has better performance than using sound pressure (82.37%), but inferior to the combination of two modalities (99.13%).

Another issue is that the modality number may affect the performance of network, algorithms with different input methods were compared. One input method was treating acceleration and sound pressure as one vector input fed into the recognition network. The recognition result is shown as the green bar in Figure 12. The other input method was treating acceleration and sound pressure features separately as two modalities fed into the recognition network. The recognition result is shown as the yellow bar. Separately treating two modalities evidently helped the algorithm obtain a better performance (3.63%) than combining features together.

**3.3.4. Comparison among Algorithms with and without RBMs.** Pretraining plays an important role in the recognition of coal and rock. Figure 13 shows the comparison of algorithms with and without pretraining. The comparison shows that the model with unsupervised pretrained RBMs (99.13%) has 4.83% improvement compared to the model without pretraining (94.30%). This result is due to the BP algorithm that used gradient descent to gain the optima and pretraining, which allowed the network to be fine-tuned at a relatively good initial value rather than a few random initial points, thereby potentially avoiding the risk of falling into poor local optima. Therefore, pretraining is significantly necessary for model training.

**3.3.5. Comparison among Algorithms with and without Transfer Learning.** Transfer learning plays the same role with

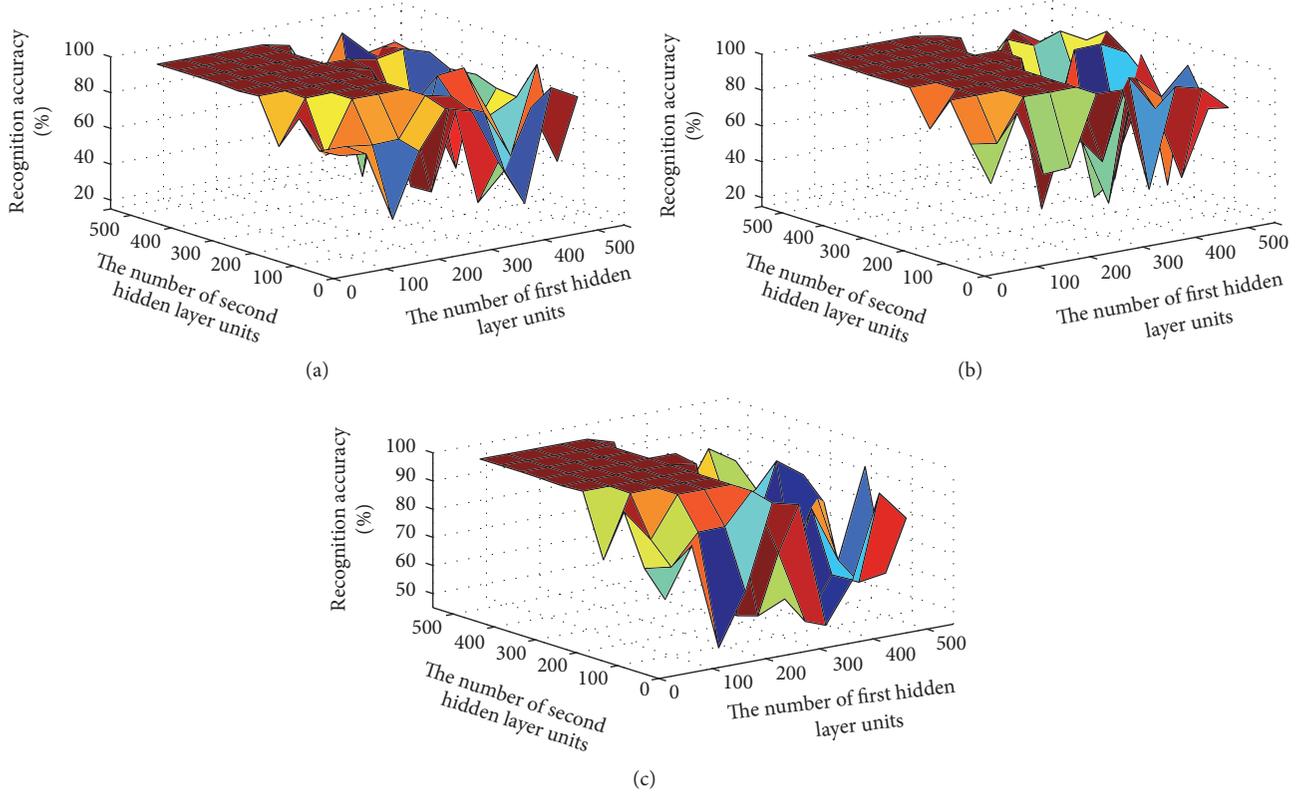


FIGURE 9: Recognition results for different DBN1 structures: (a) coal; (b) rock; (c) Ave.

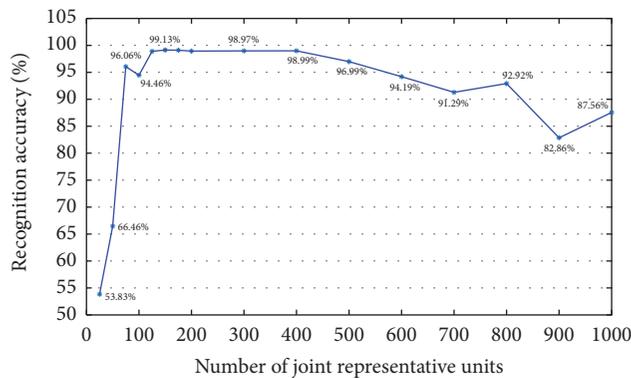


FIGURE 10: The recognition results with different numbers of joint units.

pretraining in the recognition of coal and rock (Table 3). In this experiment, algorithms with and without transfer learning were compared. Herein, the algorithm is pretrained first by the data obtained from the laboratory environment simulating coal and rock hitting the tail beam, then pretrained by the irrelevant data (hand knocking the table). The result shows that the model with transfer learning has 0.70% improvement compared with the model without transfer learning. This result is due to the transfer learning optimizing the initial value of formal training to enable the network to easily obtain the global optimum, which is similar to pretraining.

3.3.6. Comparison of Algorithms Using Different Lengths of Time. In this experiment, algorithms with different lengths of time per sample, from 7.8125 (each sample contains 512 sampling points) to 125 ms (each sample contains 8192 sampling points), were compared. Figure 14 shows that given the increasing length of sample time, recognition accuracy rapidly improves initially, then stabilizes, and finally decreases slightly. The algorithm obtained the best performance when the sample time length was set to 62.5 ms (each sample comprised 4096 sampling points). A longer sample time resulted in more included information, and performance gradually improved with the increasing length

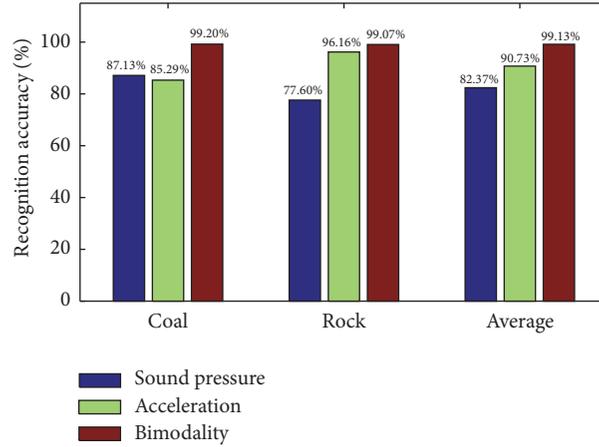


FIGURE 11: Comparison of algorithms with unimodality and bimodality.

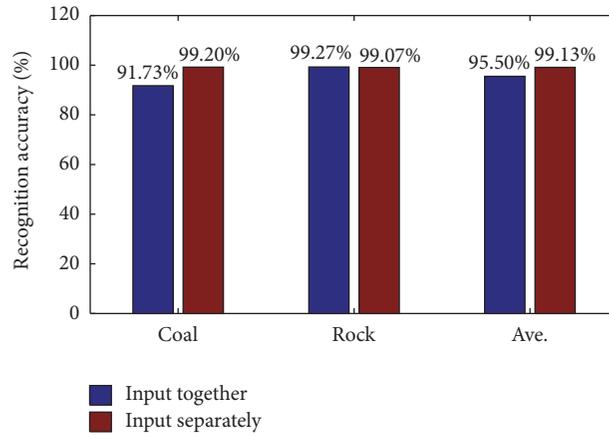


FIGURE 12: Comparison of algorithms with different input methods.

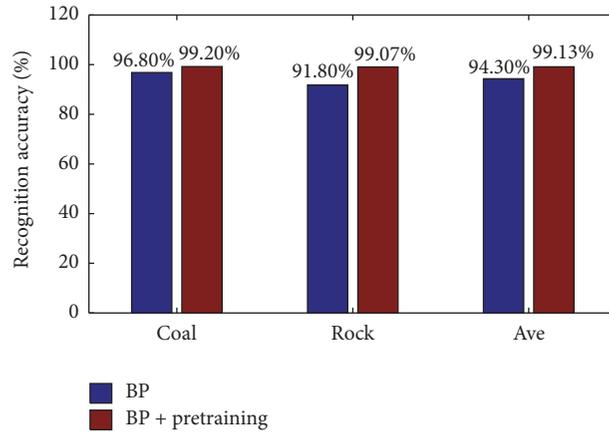


FIGURE 13: Comparison of algorithms with and without pretraining.

of time. However, when the time length of the sample was extremely long, a few samples may contain two states; and a longer time would increase the data processing time, thereby affecting the real-time performance.

*3.3.7. Comparison among Algorithms with Different Classifiers.* In this experiment, to prove the superiority of the proposed

method, other five algorithms were compared, including *k*-nearest neighbor (KNN), support vector machine (SVM), Naïve Bayes, decision tree, and DBN. As shown in Figure 15, decision tree is inferior to the other algorithms in this experiment. Naïve Bayes performs better than the other methods in rock recognition, except for DBN and the proposed method. SVM and KNN fairly perform, and KNN performs better

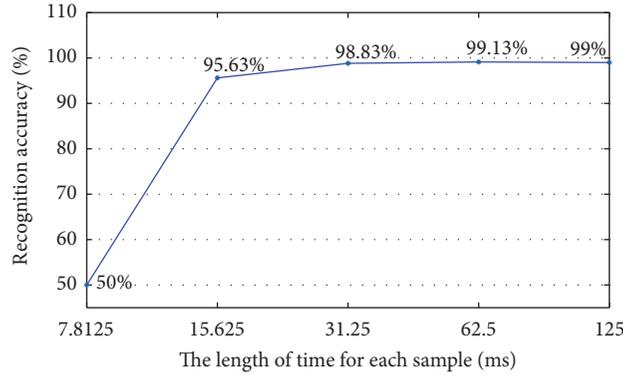


FIGURE 14: Comparison among algorithms with different frame rates.

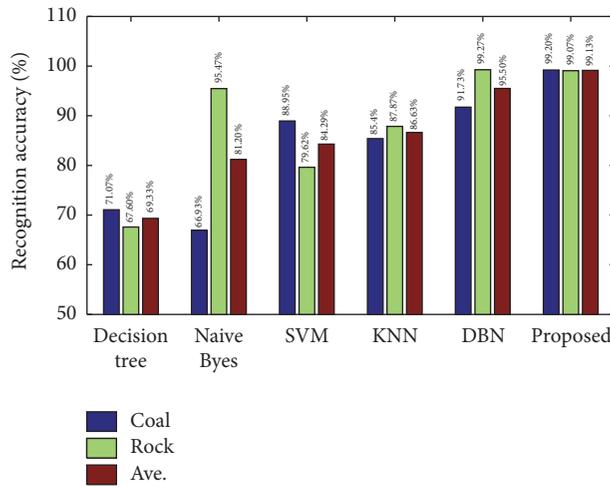


FIGURE 15: Proposed method versus other methods.

TABLE 3: Comparison among algorithms with and without transfer learning.

Method	Coal	Rock	Ave
Without transfer learning	96.96%	99.90%	98.43%
With transfer learning	99.20%	99.07%	99.13%

than SVM, especially in rock recognition. DBN performs well in rock recognition; however, in the coal recognition, DBN is inferior to the proposed method. Evidently, the proposed method performs best in this comparison experiment.

3.3.8. *Real-Time Comparison of Algorithms.* In this experiment, the real time of methods was compared, because fast and real-time identification of coal or rock during the process of top coal caving is very important to achieve the automated mining. The processing time measured using the computer system clock is estimated using Matlab R2017a on an Intel(R) Core (TM) i7-6700 CPU @ 3.40 GHz with 8 GB RAM running on Windows 10 operating system. The average data processing time and recognition time of each algorithm per sample are shown in Table 4. It can be seen

that data processing time accounts for the vast majority of the total time (more than 99.9%). Because the total time is less than 343 ms, in the actual application, it can be identified twice per second; all these methods can meet the real-time requirement.

#### 4. Conclusions

A novel method is proposed in this study based on bimodal deep learning and Hilbert-Huang transform for the identification of coal-rock in top coal caving. Four main innovations are present in this study. First, the bimodal learning method is adopted in enabling DNN to study the characteristics of coal and rock completely. Second, the transfer learning method is adopted to solve the problem wherein a large number of samples are necessary to train the DNN. Third, the extracted features of acceleration and sound pressure signals are combined to extract the most efficient features. Fourth, the most suitable installation location for the sensors is selected.

For future works, the authors plan to obtain substantial coal-rock data in different coal mines to improve the applicability of the proposed method, use more effective feature

TABLE 4: Comparison of the recognition time.

Method	Data processing/ms	Recognition/ms	Total/ms
Decision tree	342.02	$2.77 \times 10^{-3}$	342.02
Naïve Byes	342.02	$1.96 \times 10^{-3}$	342.02
SVM	342.02	$4.43 \times 10^{-2}$	342.06
KNN	342.02	$3.26 \times 10^{-1}$	342.34
DBN	342.02	$5.02 \times 10^{-3}$	342.03
Proposed	342.02	$1.93 \times 10^{-2}$	342.04

extraction methods, improve the real-time performance, and strengthen the stability and robustness. Moreover, producing an intelligent identification with high precision and speed, adaptability, and better practical products is the aim for future works.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Authors' Contributions

Zengcai Wang designed the experimental system; Guoxin Zhang and Lei Zhao formulated the proposed algorithm; Yazhou Qi and Jinshan Wang performed the experiments; Guoxin Zhang analyzed the data and wrote the paper.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 51174126).

### References

- [1] J. Qin, J. Zheng, X. Zhu, and D. Shi, "Establishment of a theoretical model of sensor for identification of coal and rock interface by natural  $\gamma$ -ray and underground trials," *Journal of China Coal Society*, vol. 21, pp. 513–516, 1996.
- [2] S. L. Bessinger and M. G. Nelson, "Remnant roof coal thickness measurement with passive gamma ray instruments in coal mines," *IEEE Transactions on Industry Applications*, vol. 29, no. 3, pp. 562–565, 1993.
- [3] Z. C. Wang, R. L. Wang, and J. H. Xu, "Research on coal seam thickness detection by natural Gamma ray in shearer horizon control," *Journal of China Coal Society*, vol. 27, pp. 425–429, 2002.
- [4] A. D. Strange, 2007, Robust thin layer coal thickness estimation using ground penetrating radar.
- [5] F. Ren, Z. Liu, and Z. Yang, "Harmonic response analysis on cutting part of shearer physical simulation system paper title," in *Proceedings of the 2010 IEEE 10th International Conference on Signal Processing, ICSP2010*, pp. 2509–2513, October 2010.
- [6] J. Sun and B. Su, "Coal-rock interface detection on the basis of image texture features," *International Journal of Mining Science and Technology*, vol. 23, no. 5, pp. 681–687, 2013.
- [7] W. Hou, "Identification of coal and gangue by feed-forward neural network based on data analysis," *International Journal of Coal Preparation and Utilization*, pp. 1–11, 2017.
- [8] G. Zhang, Z. Wang, and L. Zhao, "Recognition of rock-coal interface in top coal caving through tail beam vibrations by using stacked sparse autoencoders," *Journal of Vibroengineering*, vol. 18, no. 7, pp. 4261–4275, 2016.
- [9] L. Si, Z. Wang, X. Liu, C. Tan, and L. Zhang, "Cutting state diagnosis for shearer through the vibration of rocker transmission part with an improved probabilistic neural network," *Sensors (Switzerland)*, vol. 16, no. 4, 2016.
- [10] B. Wang, Z. Wang, and S. Zhu, "Coal-rock interface recognition based on time series analysis," in *Proceedings of the 2010 International Conference on Computer Application and System Modeling, ICCASM 2010*, pp. V8356–V8359, October 2010.
- [11] F. Ren, Z. Liu, and Z. Yang, "Dynamics analysis for cutting part of shearer physical simulation system," in *Proceedings of the IEEE International Conference on Information and Automation (ICIA '10)*, pp. 260–264, IEEE, Harbin, China, June 2010.
- [12] L. Si, Z. Wang, C. Tan, and X. Liu, "Vibration-based signal analysis for shearer cutting status recognition based on local mean decomposition and fuzzy C-means clustering," *Applied Sciences*, vol. 7, no. 2, p. 164, 2017.
- [13] J. Xu, Z. Wang, J. Wang, C. Tan, L. Zhang, and X. Liu, "Acoustic-based cutting pattern recognition for shearer through fuzzy C-means and a hybrid optimization algorithm," *Applied Sciences (Switzerland)*, vol. 6, no. 10, article no. 294, 2016.
- [14] Y.-w. Liang and S.-b. Xiong, "Forecast of coal-rock interface based on neural network and Dempster-Shafter theory," *Meitan Xuebao/Journal of the China Coal Society*, vol. 1, p. 17, 2003.
- [15] Y.-l. Zhang and S.-x. Zhang, "Analysis of coal and gangue acoustic signals based on Hilbert-Huang transformation," *Journal of China Coal Society*, vol. 1, p. 44, 2010.
- [16] J. R. Markham, P. R. Solomon, and P. E. Best, "An FT-IR based instrument for measuring spectral emittance of material at high temperature," *Review of Scientific Instruments*, vol. 61, no. 12, pp. 3700–3708, 1990.
- [17] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proceedings of the 28th International Conference on Machine Learning (ICML '11)*, pp. 689–696, Omnipress, Bellevue, Wash, USA, July 2011.
- [18] W. Zhang, Y. Zhang, L. Ma, J. Guan, and S. Gong, "Multimodal learning for facial expression recognition," *Pattern Recognition*, vol. 48, no. 10, pp. 3191–3202, 2015.
- [19] L. Zhao, Z. Wang, X. Wang, Y. Qi, Q. Liu, and G. Zhang, "Human fatigue expression recognition through image-based dynamic multi-information and bimodal deep learning," *Journal of Electronic Imaging*, vol. 25, no. 5, Article ID 053024, 2016.

- [20] B. Q. Huynh, H. Li, and M. L. Giger, "Digital mammographic tumor classification using transfer learning from deep convolutional neural networks," *Journal of Medical Imaging*, vol. 3, no. 3, Article ID 034501, 2016.
- [21] G. Prudhomme, L. Berthe, J. Bénier, O. Bozier, and P. Mercier, "Radiometric short-term fourier transform analysis of photonic doppler velocimetry recordings and detectivity limit," in *Proceedings of the Conference of the American Physical Society Topical Group on Shock Compression of Condensed Matter*, p. 160007, AIP Publishing, Tampa Bay, Fla, USA.
- [22] P. Neri, "Bladed wheels damage detection through non-harmonic fourier analysis improved algorithm," *Mechanical Systems and Signal Processing*, vol. 88, pp. 1–8, 2017.
- [23] W. Zhao, Z. Wang, J. Ma, and L. Li, "Fault diagnosis of a hydraulic pump based on the CEEMD-STFT time-frequency entropy method and multiclass SVM classifier," *Shock and Vibration*, vol. 2016, Article ID 2609856, 8 pages, 2016.
- [24] M. Zhang, G. Wang, and G. M. Evans, "Flow visualizations around a bubble detaching from a particle in turbulent flows," *Minerals Engineering*, vol. 92, pp. 176–188, 2016.
- [25] T. Putra, S. Abdullah, D. Schramm, M. Nuawi, and T. Bruckmann, "Reducing cyclic testing time for components of automotive suspension system utilising the wavelet transform and the fuzzy C-Means," *Mechanical Systems and Signal Processing*, vol. 90, pp. 1–14, 2017.
- [26] M. A. Brdys, M. T. Brdys, and S. M. Maciejewski, "Adaptive predictions of the euro/zloty currency exchange rate using state space wavelet networks and forecast combinations," *International Journal of Applied Mathematics and Computer Science*, vol. 26, no. 1, pp. 161–173, 2016.
- [27] D. Sun and Q. Ren, "Seismic damage analysis of concrete gravity dam based on wavelet transform," *Shock and Vibration*, vol. 2016, Article ID 6841836, 8 pages, 2016.
- [28] N. E. Huang and Z. Wu, "A review on Hilbert-Huang transform: method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, 2008.
- [29] M. Lozano, J. A. Fiz, and R. Jané, "Performance evaluation of the Hilbert-Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization," *Signal Processing*, vol. 120, pp. 99–116, 2016.
- [30] H. Xu, J. Liu, H. Hu, and Y. Zhang, "Wearable sensor-based human activity recognition method with multi-features extracted from Hilbert-Huang transform," *Sensors (Switzerland)*, vol. 16, no. 12, article no. 2048, 2016.
- [31] X. Yu, E. Ding, C. Chen, X. Liu, and L. Li, "A novel characteristic frequency bands extraction method for automatic bearing fault diagnosis based on Hilbert Huang transform," *Sensors (Switzerland)*, vol. 15, no. 11, pp. 27869–27893, 2015.
- [32] P. Smolensky, *Information Processing in Dynamical Systems: Foundations of Harmony Theory*, DTIC Document, 1986.
- [33] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [34] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *American Association for the Advancement of Science. Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [35] A. Coates, H. Lee, and A. Y. Ng, *An Analysis of Single-Layer Networks in Unsupervised Feature Learning*, vol. 1001 of 2, 2010.
- [36] H. Larochelle and Y. Bengio, "Classification using discriminative restricted boltzmann machines," in *Proceedings of the 25th International Conference on Machine Learning (ICML '08)*, pp. 536–543, July 2008.
- [37] Z. H. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2046, pp. 1597–1611, 2004.
- [38] R. Ditommaso, M. Mucciarelli, S. Parolai, and M. Picozzi, "Monitoring the structural dynamic response of a masonry tower: comparing classical and time-frequency analyses," *Bulletin of Earthquake Engineering*, vol. 10, no. 4, pp. 1221–1235, 2012.



**Hindawi**

Submit your manuscripts at  
<https://www.hindawi.com>

