IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 24, 2019, accepted August 21, 2019, date of publication August 27, 2019, date of current version September 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2937829

Fusion Recognition of Shearer Coal-Rock Cutting State Based on Improved RBF Neural Network and D-S Evidence Theory

LEI SI[1,2], ZHONG-BIN WANG[1,2], AND GAN JIANG[1]
[1]School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221116, China
[2]Jiangsu Engineering Technology Research Center on Intelligent Equipment for Fully Mining and Excavating, Xuzhou 221116, China

Corresponding author: Zhong-Bin Wang (wangzbpaper@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51605477 and Grand U1510117, and in part by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions.

ABSTRACT Accurate recognition of coal-rock cutting state is a prerequisite for intelligent operation of shearer, so as to achieve safe and efficient production in coal mines. This paper takes the sound signal, Y-axis and Z-axis vibration signals as analytic objects and proposes a fusion recognition method for shearer coal-rock cutting state via the combination of improved radical basis function neural network (RBFNN) and Dempster-Shafer (D-S) evidence theory. First of all, on the basis of original fruit fly optimization algorithm (FOA), the location updating mechanism of moth-flame optimization (MFO) is used to improve the convergence performance and exploration ability of FOA. Thus, a hybrid optimization algorithm of MFO-FOA is accordingly designed and some simulations are conducted to verify the effectiveness and superiority. Then, the optimal network parameters of RBFNN are found out by using proposed MFO-FOA to realize the excellent generalization ability and predictive performance. Moreover, the collected signals are decomposed by variational mode decomposition, and the envelope entropy and kurtosis are used to extract the features of first three intrinsic mode function components. The feature vectors obtained from three-type sensor data are utilized to construct the RBFNN classifiers. Besides, the D-S evidence theory with evidence correlation coefficient is introduced to fuse the preliminary identification results of three RBFNN classifiers. Finally, a self-designed experimental platform for shearer cutting coal-rock is built and some experiments are provided. The experimental results based on measured data demonstrate that the proposed method can effectively identify the coal-rock cutting state with higher accuracy.

INDEX TERMS Fusion recognition, improved fruit fly optimization algorithm, RBF neural network, shearer coal-rock cutting state.

I. INTRODUCTION

Coal is the most abundant and widely distributed fossil fuel on the earth. In China's proven fossil energy, coal accounts for about 94 percent. This phenomenon leads to coal-dominated energy structure of China will not change for a long time. But coal production is still a high-risk industry due to the harsh underground working environment and low reliability and automation level of coal mining equipment [1]. In 2018, there were 224 coal mine safety accidents in China. The mortality rate of one million tons was 3.1 times that of the United States and 6.6 times that of Australia. Therefore, in order to achieve safe and efficient production of coal mines, it is urgent to improve the intelligent level of mining equipment. As an important piece of coal mining equipment, shearer is responsible for cutting coal and transporting the coal to scraper conveyor. Accurately identifying the coal-rock cutting state of shearer is a prerequisite for achieving automatic mining and is becoming a research hotspot in the field of coal resource [2].

In recent years, in order to improve the intelligent operation level of shearer, domestic and foreign scholars mainly focused on two aspects of coal-rock interface recognition and memory cutting, and put forward many effective methods. In [3], radar technology was used to identify the coal-rock interface, which had been industrially tested at the Quecreek

The associate editor coordinating the review of this article and approving it for publication was Xianye Ben.

coal mine in Pennsylvania, USA. In [4], an opt-tactile sensor was newly developed to detect different types of material layers where a shearer machine can operate at the longwall face of underground coal mines. Ref. [5] proposed a low-activity spectrometric gamma-ray logging technique as a sensitive tool for the delineation of coal-rock interfaces. In [6], Li *et al.* proposed a coal-rock interface recognition method based on ultrasonic phased array by using the difference of acoustic impedance between coal and rock and the phased array technology. The vibration signals and infrared thermal images of cutting picks were deeply analyzed to realize effective dynamic identification of coal and rock in shearer cutting process [7], [8]. Moreover, many other methods based on acoustic waves [9], ground penetrating radar [10], natural gamma-ray [11] and Terahertz time domain spectrum [12] had been proposed to identify the coal-rock interface and good results were obtained in the laboratory. Through the above researches, it can be found that only two cutting states can be identified by the coal-rock interface recognition methods: coal or rock. When the geological condition of coal seam changes dramatically, there are some defects such as low identification accuracy and poor applicability. The memory cutting is the most widely used automatic control method for shearer. In [13], an improved approach through integration of improved genetic algorithm and fuzzy logic control method was proposed to reduce the enlargement of coal floor deformation in shearer memory cutting process. Refs. [14], [15] developed a hidden Markov model memory cutting method for shearer by employing data correlation of adjacent coal seams, and the adjustment frequency and accuracy can be improved. Although the memory cutting method can improve the automatic control level of shearer to a certain extent, the application effect is not ideal when the coal seam breaks suddenly.

Under this background, some scholars try to identify the coal-rock cutting state of shearer to provide the basis for its intelligent control. In [16], the study was to investigate the effects of previously ignored rock parameters along with engineering rock properties on specific cutting energy, and the results can be used to preliminarily identify the working state of shearer. In [17], the development of a new rippability classification system for coal measure rock based on specific energy was presented. Ref. [18] structured the theoretical model of cutting load and current, and proposed a novel analysis method for shearer cutting load characteristic based on particle filter to identify the cutting state of shearer. In [19]–[21], the vibration signals, sound signals and temperature of cutting area were applied in the recognition of shearer coal-rock cutting state by using some intelligent classifiers. Due to the complicated structure and rugged working conditions of shearer, the single-type sensor data are not reliable enough, and can not accurately reflect the coal-rock cutting state of shearer. Therefore, this paper presents a multi-sensor information fusion method for coal-rock cutting state identification through the integrated use of vibration signals and sound signals.

In the field of state recognition, there are many intelligent recognition methods, which have been widely used in fault diagnosis [22]–[24] and image classification [25], [26]. Among these methods, radical basis function neural network (RBFNN), put forward by Moody and Darken in 1988, has the advantages of simple network structure, strong approximation ability, fast learning speed, being not easy to fall into local minimum problems and good robustness [27], [28]. In order to enhance the learning performance of RBFNN, many researchers have adopted the RBFNN structure along with some nature-inspired meta-heuristic algorithms such as particle swarm optimization (PSO) [29], ant colony optimization (ACO) [30] and genetic algorithm (GA) [31], to implement the learning of the network. However, these techniques always exist with some drawbacks. Fruit fly optimization algorithm (FOA), as a novel swarm intelligence search algorithm, has been widely used in many fields [32], [33], because of its simple structure, strong global optimization ability, easy to understand and learn. Like other optimization algorithms, FOA may also have problems such as premature maturity and poor exploration ability. Hybridizing is a reasonable way to take strengths and avoid weakness. Considering that moth-flame optimization (MFO) algorithm has a unique population renewal mechanism [34], this study intends to design a mix algorithm of PSO with MFO, named MFO-FOA, for training RBFNN and finding out the optimal network parameters to make suitable performance verification and comparison.

However, every single technique always exists with some drawbacks. The single-type sensor data are not reliable enough in actual working condition, and will reduce the recognition effect of shearer coal-rock cutting state based on the proposed RBFNN model-based classifier. The D–S evidence theory, proposed by Dempster and perfected by Shafer [35], has been widely used for information fusion, such as intelligent obstacle sensing and recognizing [40], fault detection [41], target recognition [42] and so on. Therefore, the D–S evidence theory is introduced to achieve the decision-level fusion of the recognition results based on single-type signal source. The major contributions of the proposed recognition scheme can be summarized as follows:

(1) we propose a new optimization algorithm based on the combination of MFO and FOA to find out the optimal network parameters of RBFNN. The effectiveness and superiority of MFO-FOA is verified through the simulation comparison with other popular meta-heuristic algorithms.

(2) A new recognition technique based RBFNN optimized by using MFO-FOA is presented to achieve the coal-rock cutting state recognition based on three-type sensing data collected from one sound sensor and two vibration sensors. The D–S evidence theory is used to fuse the recognition results of three RBFNN classifiers and the accurate identification results can be accordingly obtained.

(3) Some experimental comparative study is performed to prove the effectiveness and superiority of proposed method.

The remainder of this work is organized as follows. In Section 2, the basic algorithms of RBF neural network

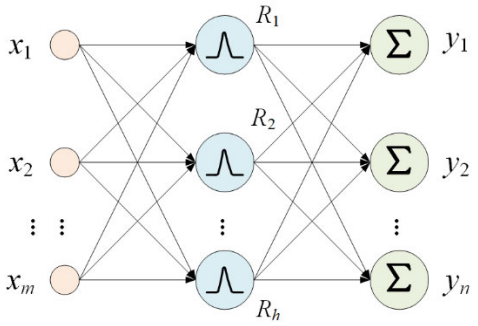


FIGURE 1. Structure of RBF neural network.

and D-S evidence theory are briefly introduced. In Section 3, the proposed hybrid optimization algorithm are proposed and the parameters optimization procedure of RBFNN using MFO-FOA are presented in detail. In Section 4, the proposed fusion recognition system is constructed. In Section 5, some experiments are conducted to verify the performance of the proposed method and the comparisons with other methods are performed to illustrate its excellent property in terms of recognition accuracy. Finally, the conclusions of this work are summarized in Section 6.

II. BASIC THEORY

A. RBF NEURAL NETWORK

RBF neural network is a three-layer feedforward neural network, which is generally composed of input layer, hidden layer and output layer, as shown in Fig. 1.

The commonly used radial basis function in RBF neural networks is Gaussian function, which can be expressed as:

$$R\left(x_p - c_i\right) = \exp\left(-\frac{1}{2\sigma^2}\left\|x_p - c_i\right\|^2\right) \tag{1}$$

where $\|x_p - c_i\|^2$ is the Euclidean norm, $x_p = (x_1^p, x_2^p, \ldots, x_m^p)$ is the p-th input sample, c_i is the Gaussian function center of the i-th node, σ is the variance, that is, the width of Gaussian function.

The output of RBF neural network can be calculated as:

$$y_j = \sum_{i=1}^{h} \omega_{ij} R\left(x_p - c_i\right) + \theta_j \tag{2}$$

where ω_{ij} is the connection weight between the hidden layer and the output layer, h is the number of nodes in the hidden layer, and θ_j is the threshold value of the j-th output node, which is the actual output value of the j-th output node in the output layer.

B. D-S EVIDENCE THEORY

In D-S evidence theory, if the elements of a finite complete set $\Theta = \{\theta_1, \theta_2, \cdots, \theta_N, \}$ are mutually exclusive, the set can be called a discernment framework. Defining function m: $2^\Theta \rightarrow [0, 1]$, $m(\emptyset) = 0$, $\sum_{A\subset\Theta} m(A) = 1$, $m(A)$ is termed as basic probability assignment (BPA).

The belief degree of proposition A can be measured by two concepts: $Bel(A)$ and $Pl(A)$. $Bel(A)$ represents the total belief level and can be defined as follows:

$$Bel(A) = \sum_{B\subseteq A} m(B) \quad (\forall A \subseteq \Theta) \tag{3}$$

Thus $Bel(A) = m(A)$ and the belief function satisfies the following conditions:

$$Bel(\emptyset) = 0, \quad Bel(\Theta) = 1 \tag{4}$$

$Pl(A)$ represents the plausibility belief level and can be defined as follows:

$$Pl(A) = \sum_{A\cap B\neq\emptyset} m(B) \quad (\forall A \subseteq \Theta, B \subseteq \Theta) \tag{5}$$

In D-S evidence theory, Dempster rule is used to combine information from multiple independent sources. Let m_1 and m_2 be the BPAs, the corresponding focal elements are $A_1, A_2, \ldots, A_k$ and $B_1, B_2, \ldots, B_k$, m represents the new evidence after the combination of m_1 and m_2. The Dempster combination rule can be described as follows:

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \frac{1}{1-k} \sum_{A_i\cap B_j=A} m_1(A_i) m_2(B_j) \end{cases} \tag{6}$$

where k reflects the conflict coefficient between evidences and can be represented as follows:

$$k = \sum_{A_i\cap B_j=\emptyset} m_1(A_i) m_2(B_j) \tag{7}$$

C. THE EVIDENCE CORRELATION COEFFICIENT

Obviously, the combination rule of D-S evidence theory has unavoidable disadvantages. That is, once there is a certain conflict between evidences, the fusion results would be counter-intuitive or not consistent with some evidences. In order to enhance the fusion effect and improve the applicability, the evidence correlation coefficient is introduced to measure the conflict coefficient between evidences and modify or preprocess the original evidences with a weighted mean combination model.

For two evidences m_i and m_j in discernment framework Θ, the correlation coefficient (denoted as r_{BPA}) and correlation degree (denoted as c) can be calculated by:

$$\begin{cases} r_{BPA}(m_i, m_j) = c(m_i, m_j) / \sqrt{c(m_i, m_i) \cdot c(m_j, m_j)} \\ c\left(m_i, m_j\right) = \sum_{ii=1}^{2^N} \sum_{jj=1}^{2^N} m_i(A_{ii}) m_j(A_{jj}) \frac{\left|A_{ii} \cap A_{jj}\right|}{\left|A_{ii} \cup A_{jj}\right|} \end{cases} \tag{8}$$

If the correlation coefficient between a certain evidence and other evidences is larger, the more support the evidence can receive and the higher the credibility of the evidence is, so the weight assigned to it should be greater. The credibility of evidence is taken as the standard to measure the weight assigned to each evidence, and the basic probability assignments of evidence are processed by using weighted mean method. Then, the combination rule of Dempster is used to realize the fusion of each modified evidence.

Through the above ideas, the correlation coefficient between m_i and m_j can be taken as the support degree of evidence, namely $Sup(m_i, m_j) = r_{BPA}(m_i, m_j)$, which is

abbreviated as S_{ij} later. Thus, the support degree matrix of evidence can be obtained.

$$SM = \begin{bmatrix} 1 & S_{12} & \cdots & S_{1j} & \cdots & S_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{i1} & S_{i2} & \cdots & S_{ij} & \cdots & S_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{n1} & S_{n2} & \cdots & S_{nj} & \cdots & 1 \end{bmatrix} \tag{9}$$

Obviously, the smaller the conflict between the two evidences is, the larger the correlation coefficient and mutual support degree will be. The total support of other evidences for evidence m_i is:

$$Sup\,(m_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} Sup\,(m_i, m_j) = \sum_{\substack{j=1 \\ j \neq i}}^{n} S_{ij} \tag{10}$$

The support degree after normalization can be used as the credibility $Crd(m_i)$ of evidence m_i, which can be calculated as:

$$\begin{cases} Crd\,(m_i) = Sup\,(m_i) / \sum_{i=1}^{n} Sup\,(m_i) \\ \sum_{i=1}^{n} Crd\,(m_i) = 1 \end{cases} \tag{11}$$

The credibility $Crd(m_i)$ can be used as the weight assigned to this evidence and new evidences can be obtained by averaging all evidences according to the weights, which can realize the fusion of highly conflicting evidences. The specific fusion process is shown in Fig. 2.

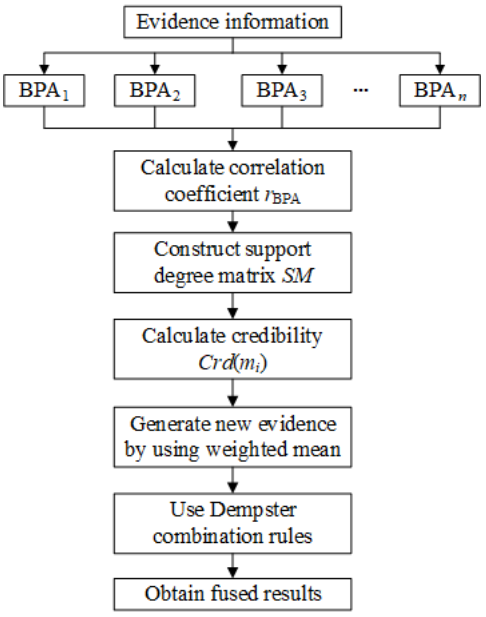


FIGURE 2. Fusion process of correlation coefficient-based D-S evidence theory.

III. THE IMPROVED RBF NEURAL NETWORK BASED ON A HYBRID OPTIMIZATION ALGORITHM

A. FRUIT FLY OPTIMIZATION ALGORITHM

Fruit fly optimization algorithms (FOA) was proposed by Pan [36] in 2012. Similar to other meta-heuristic intelligent algorithms, FOA derives from the natural fly population foraging process. In the process of foraging, drosophila flies constantly exchange food information in order to obtain the most effective path. The optimization process of FOA is shown in Fig. 3. The detailed implementation steps can be referred to some references [32].

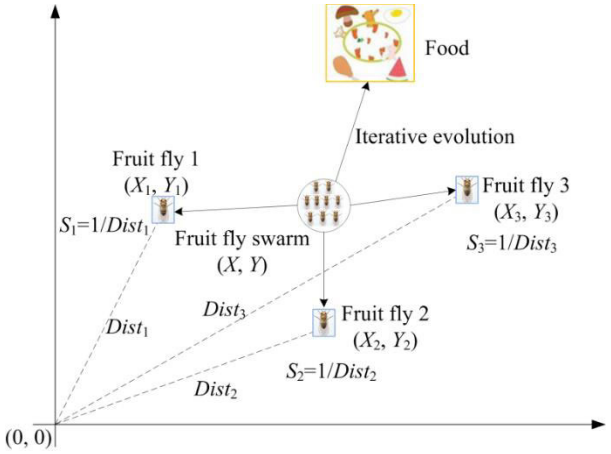


FIGURE 3. Schematic diagram of FOA optimization process.

B. MOTH-FLAME OPTIMIZATION ALGORITHM

The moth-flame optimization (MFO) algorithm, proposed by Mirjalili in 2015 [34], is a novel swarm intelligence optimization algorithm. The main inspiration of this algorithm is the navigation method of moths in nature called transverse orientation. In the d-dimensional search space, there exists a population M consisting of n moths. Each moth has a unique flame corresponding to it. The matrix F consisting of all flames has the same dimension as the moth population M. The moth is the actual search subject moving in the search space, and the flame is the best location for the moth to search so far. Therefore, if a better solution is found, each moth will search nearby and update it. Through this mechanism, the moth will never miss its optimal solution.

According to the helical flight path of moths, the location updating mechanism of each moth relative to the flame can be expressed as

$$M_i = D_i \cdot e^{bt} \cdot \cos(2\pi t) + F_j \tag{12}$$

where D_i denotes the distance between the i-th moth and the j-th flame, $D_i = |F_j - M_i|$. b is the shape constant of logarithmic helix. t is the path coefficient and can be set as a random number in $[r, 1]$, where r decreases linearly in the interval $[-2, -1]$ with the increase of iteration times. 'This can enable the moth to approach the corresponding flame more accurately with the iteration process.

Fig. 4 illustrates the location updating model of a math around its corresponding flame. The smaller the path coefficient t is, the closer the moth is to the flame and the faster it updates around the flame. It can be seen from the spiral

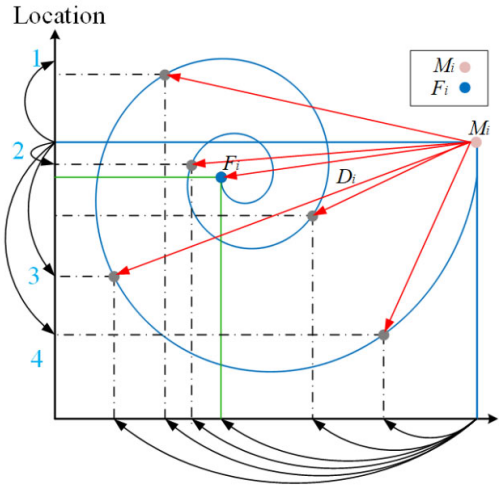


FIGURE 4. Location updating model of MFO.

equation that the moth can not only fly in the space between moth and flame, but also search around the flame, so that the global search ability and the local development ability of the algorithm are comprehensively considered.

In order to ensure that the local development ability of MFO does not decrease with the iteration, (13) is used to reduce the number of flames adaptively.

$$flame.no = round(N - l \times \frac{N-1}{T}) \tag{13}$$

where l is the current number of iterations, N is the maximum number of flames, T is the maximum number of iterations. In the initial step of iteration, there are N frames. At the end of the iteration, the moths use only the best flames to update their positions.

C. A HYBRID OPTIMIZATION ALGORITHM BASED ON MFO AND FOA

In FOA, the location updating of each individual fruit fly only depends on the location of the current fruit fly population, which makes the algorithm have better global search ability, but weakens the local search ability and easily falls into the local optimum. MFO algorithm has a special location updating mechanism, so that the search space is not limited to the space between the moth and the flame, but the whole space around the flame, expanding the search space of the algorithm. Based on the merits and demerits of FOA and MFO, this paper improves the FOA by using the location updating mechanism of MFO, and forms a new swarm intelligence optimization algorithm, termed as MFO-FOA. The specific improvement ideas can be described as follows. In the process of updating the location of flies, the random flight style of flies is changed into a spiral flight path. The center of the spiral flight path is the current location of the fly population, which increases the flight distance of flies and expands the search space, making the algorithm not easy to fall into local optimum. At the same time, the shape constant of the spiral equation can be reduced gradually with the increase of iteration times, thus reducing the search space of the fly population and ensuring the fast convergence of the algorithm in the later iteration.

The new location updating formula of fruit fly is as follows:

$$\begin{cases} dis\_X_i = \left|X_{i\_last} - X\_axis\right| \\ dis\_Y_i = \left|Y_{i\_last} - Y\_axis\right| \end{cases} \tag{14}$$

$$\begin{cases} X_i = dis\_X_i \cdot e^{bt} \cdot \cos(2\pi t) + X\_axis \\ Y_i = dis\_Y_i \cdot e^{bt} \cdot \cos(2\pi t) + Y\_axis \end{cases} \tag{15}$$

where dis_X_i and dis_Y_i are the distances between the individual location (X_{i_last} and Y_{i_last}) of the last generation and the current swarm location (X_axis and Y_axis), respectively. $b = 1 - n/(T+1)$, $t = (-2 - n/T) \times \text{rand} + 1$.

The specific implementation steps of MFO-FOA are as follows:

Step 1: Set the population size N, the maximum iteration number T, the location range of fly population LR and the random flight distance range FR. The initial fruit fly swarm location (X_axis, Y_axis) is randomly generated by using (16) The location of the first generation of individual fly is updated by using (17).

$$\begin{cases} X\_axis = rand\ (LR) \\ Y\_axis = rand(LR) \end{cases} \tag{16}$$

$$\begin{cases} X_i = X\_axis + rand\ (FR) \\ Y_i = Y\_axis + rand(LR) \end{cases} \tag{17}$$

Step 2: The distance between the current location and the original point is calculated, and the smell concentration judgment value (S_i) can be obtained.

$$\begin{cases} Dist_i = \sqrt{X_i^2 + Y_i^2} \\ S_i = 1/Dist_i \end{cases} \tag{18}$$

Step 3: Then, the smell concentration ($Smell_i$) of the individual fruit fly location is calculated by inputting the smell concentration judgment value (S_i) into the $Smell_i$ judgment function (also called the fitness function). The best smell concentration can be obtained by ranking all $Smell_i$, and the corresponding individual location $bestIndex$ is determined and found out.

$$\begin{cases} Smell_i = fit\ (S_i) \\ [bestSmell\ bestIndex] = best(Smell_i) \end{cases} \tag{19}$$

Step 4: The current $bestSmell$ is compared with the global optimal concentration $Smellbest$. If $bestSmell$ is better than $Smellbest$, then the global optimal concentration and the corresponding position should be updated.

$$\begin{cases} Smellbest = bestSmell \\ X\_axis = X\ (bestIndex) \\ Y\_axis = Y(bestIndex) \end{cases} \tag{20}$$

Step 5: $n = n + 1$. The individual position of fruit fly is updated by using (14) and (15).

TABLE 1. Seven used functions for comparison.

Function Name	Function Expression	Value Range	Optimum Solutions
Sphere function f_1	$f(x)=\sum_{i=1}^{5}x_i^2$	[-5, 5]	f(0,0,0,0,0) =0
Schwefel's 2.22 function f_2	$f(x)=\sum_{i=1}^{2}\lvert x_i\rvert+\prod_{i=1}^{2}\lvert x_i\rvert$	[-5, 5]	f(0,0)=0
Schwefel's 1.2 function f_3	$f(x)=\sum_{i=1}^{5}\left(\sum_{j=1}^{i}\lvert x_j\rvert\right)^2$	[-10, 10]	f(0,0,0,0,0)=0
Matyas function f_4	$f(x)=0.26(x_1^2+x_2^2)-0.48x_1x_2$	[-10, 10]	f(0,0)=0
Griewank function f_5	$f(x)=\frac{x_1^2+x_2^2}{200}-\cos x_1\cos\left(\frac{x_2}{\sqrt{2}}\right)+1$	[-10, 10]	f(0,0)=0
CrownedCross function f_6	$f(x)=0.0001\times\left[\left\lvert\sin x_1\sin x_2\exp\left(\left\lvert 100-\frac{\sqrt{x_1^2+x_2^2}}{\pi}\right\rvert\right)\right\rvert+1\right]^{0.1}$	[-10, 10]	f(0,0)=0.0001
Three-Hump Camel function f_7	$f(x)=2x_1^2-1.05x_1^4+\frac{x_1^6}{6}+x_1x_2+x_2^2$	[-5, 5]	f(0,0)=0

Step 6: If the termination condition is satisfied, the iteration is stopped. Otherwise, steps 2∼5 are repeated until the maximum iteration number T is reached.

In order to verify the effectiveness and superiority of MFO-FOA algorithm, seven widely used testing functions are used in this paper and their curve graphs usually have some peaks, valleys and channels. Therefore, the optimization performance of proposed algorithm and the ability to avoid falling into local optimum can be undoubtedly proved. The function expressions and theoretical optimum solutions are listed in Table 1.

In this simulation, other four popular intelligent algorithms, including FOA, MFO, bat algorithm (BA) and PSO, are compared with proposed algorithm to display the distinctiveness and superiority. To verify the performance of the improvement measures for FOA, an improved FOA (IFOA) presented in [37] is also used in the simulation. The main parameters are set as follows: population size $N = 30$, the maximum iteration number $T = 200$. Other parameters are set as default values. To ensure the accuracy, each algorithm is performed on the individual function for 20 times. The optimal value obtained after 20 times of simulation and the mean optimal value over the 20 replications are used to measure the performance of six algorithms. The optimization results over seven functions are shown in Table 2 and the mean optimal fitness curves are illustrated in Fig. 5.

As can be seen from Table 2, the optimal value precision grade of MFO-FOA is obviously higher than other five algorithms. For example, for Schwefel's 2.22 function, the precision grade of MFO-FOA can reach to E-38 and the approximate precision grade is E-26 of MFO. The precision grade of other four algorithms is only E-3. The comparison results indicate that the proposed algorithm has better search performance and exploration ability in the later iteration period.

The mean optimal fitness curves illustrated in Fig. 5 also show the rapid convergence and superior ability to overstep the local extremum for most functions. Especially for CrownedCross function, the optimal values based on MFO-FOA and MFO can reach to 1.0000E-04 and 1.0036E-04, which are basically equal to the theoretical optimal solution. Fig. 5(f) also indicates MFO-FOA has faster convergence speed than MFO algorithm. In addition, the location of the optimal solution of other algorithms is obviously far from the theoretical optimal point. All results prove the faster search speed and better convergence accuracy of MFO-FOA. It adopts new location updating strategy to make fruit fly well-distributed and benefit the global optimization, which can overcome effectively the premature problem of FOA, BA and PSO algorithms.

D. PARAMETER OPTIMIZATION OF RBF NEURAL NETWORK BASED MFO-FOA

In theory, the parameters of RBF neural network such as the center and width of the basis function, the connection weights between hidden layer and output layer, and the threshold values of the output nodes need to be optimized by MFO-FOA algorithm. However, too many parameters will lead to too high dimension of the algorithm and reduce the speed and accuracy of optimization. The MATLAB toolbox provides the creation function *newrb* of RBF neural network model, as shown in (21). The *newrb* function can automatically increase the number of hidden layer neurons until the network meets the target mean square error or the number of neurons reaches the maximum.

$$[net, tr]=newrb(P_{net}, T_{net}, goal, spread, MN, DF) \quad (21)$$

where P_{net} is the input vector; T_{net} is the output vector; *goal* is the target mean square error; *spread* is the expansion coefficient of radial basis function, i.e. the width of the basis function; *MN* is the maximum number of neurons; *DF* is the interval of increasing the number of neurons.

It can be seen that using *newrb* function to create RBF neural network only needs to select the appropriate width of basis function, so that RBF neurons can respond in the interval covered by the input vector. The larger the width of basis function is, the smoother the fitting process of Gaussian function will be. However, a large number of

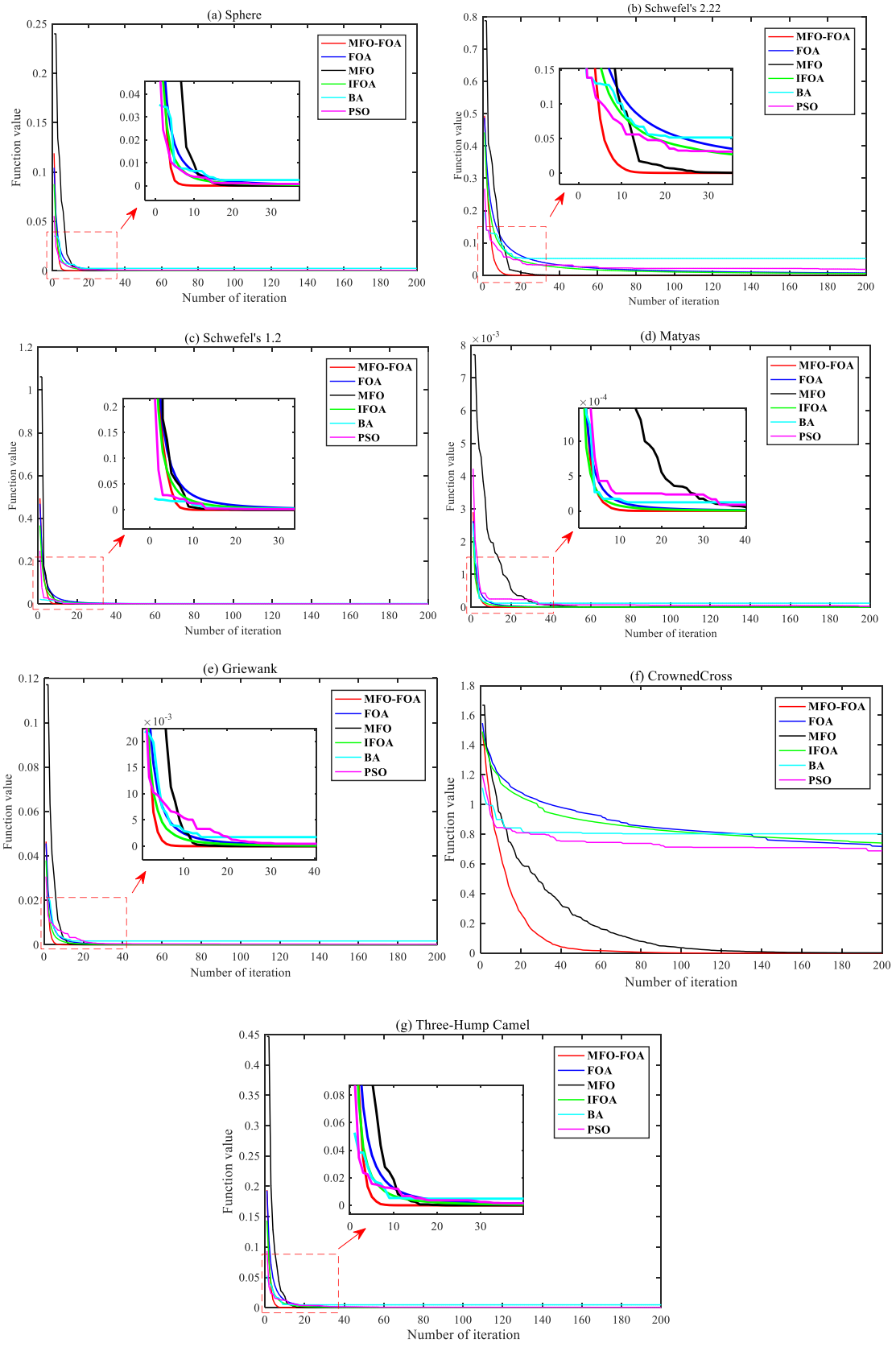


FIGURE 5. The mean optimal fitness curves of seven functions based on different optimization algorithms.

TABLE 2. The simulation results of different algorithms.

Function	Algorithm	Optimal Value	Mean Optimal Value
f_1	MFO-FOA	**1.8956E-73**	**6.3996E-68**
	FOA	2.0635E-05	2.1219E-5
	MFO	3.8603E-50	5.5551E-46
	IFOA	1.1288E-5	1.1702E-5
	BA	1.2023E-4	0.0025
	PSO	5.2855E-7	4.2980E-4
f_2	MFO-FOA	**6.4739E-38**	**9.8824E-35**
	FOA	0.0065	0.0065
	MFO	4.1273E-26	7.2192E-25
	IFOA	0.0047	0.0049
	BA	0.0032	0.0515
	PSO	0.0018	0.0180
f_3	MFO-FOA	**5.3941E-73**	**1.2478E-67**
	FOA	1.0074E-04	1.0306E-04
	MFO	1.5159E-45	6.5620E-38
	IFOA	5.6214E-05	5.8838E-05
	BA	4.7300E-05	0.0024
	PSO	2.1311E-05	2.9857E-04
f_4	MFO-FOA	**0**	**0**
	FOA	7.8491E-06	7.9653E-06
	MFO	**0**	**0**
	IFOA	4.2954E-06	4.4933E-06
	BA	1.4238E-05	5.2982E-04
	PSO	2.0198E-07	8.3099E-05
f_5	MFO-FOA	**1.6733E-55**	**9.5496E-46**
	FOA	4.2027E-07	4.2837E-07
	MFO	9.8563E-34	6.8696E-09
	IFOA	2.3083E-07	2.3961E-07
	BA	2.4572E-06	1.0239E-04
	PSO	1.6656E-07	3.7268E-05
f_6	MFO-FOA	**1.0000E-04**	**1.0000E-04**
	FOA	0.6987	0.7182
	MFO	1.0036E-04	0.0022
	IFOA	0.6580	0.7412
	BA	0.4270	0.8031
	PSO	0.5480	0.6882
f_7	MFO-FOA	**1.9833E-75**	**1.2446E-68**
	FOA	4.0744E-05	4.1987E-05
	MFO	2.1545E-52	1.0399E-45
	IFOA	2.2592E-05	2.3672E-05
	BA	1.4189E-05	0.0048
	PSO	5.6281E-06	3.5118E-04

neurons are needed to adapt to the rapid change of the function, and the sharpness of the Gaussian function will decrease, and the approximation error will increase. On the contrary, the smaller the width of basis function is, the smaller the approximation error is. However, the fitting process is not smooth and the performance of the network becomes worse, leading to the phenomenon of over-adaptation.

Therefore, MFO-FOA is used to optimize the basis function width of RBF neural network. The mean square error (MSE) between the expected output and actual output of the neural network is taken as the fitness function, so that the optimized RBF neural network achieves the best performance. The flowchart can be shown in Fig. 6.

IV. THE PROPOSED FUSION RECOGNITION SYSTEM FOR SHEARER CUTTING STATE

In this paper, based on the multi-sensor information fusion technology, three sensing signals are synthetically used to identify the coal-rock cutting state of shearer, including the cutting sound signal, Y-axis and Z-axis vibration signals of rocker arm. Firstly, the three sensing signals are preprocessed by variational mode decomposition (VMD) and the feature vectors are reasonably extracted. Then, the improved RBF neural network is used to identify the coal-rock cutting state and three independent recognition results can be obtained. Finally, the recognition results are fused by using D-S evidence theory to achieve the fusion recognition of coal-rock cutting state. The specific fusion recognition model is shown in Fig. 7.

A. SIGNAL PREPROCESSING

In the process of shearer cutting coal-rock, the components of sound signals and vibration signals are very complex, which have strong non-linearity, instability and impact. Traditional signal processing methods cannot deal with them effectively. In order to accurately extract the characteristic information contained in the signals, variational mode decomposition (VMD) is introduced to preprocess the three kinds of signals in this paper. Compared with empirical mode decomposition (EMD) and ensemble empirical mode decomposition (EEMD), VMD has better performance in noise-resistance, suppression of modal aliasing and endpoint effects. However, in the process of VMD, for the number of modes K, the manual trial method is generally applied, which has greater subjective influence on the decomposition effect. In view of this, EEMD is firstly used to decompose the signals and then the decomposition orders of VMD can be reasonably determined according to the result of EEMD.

B. FEATURE EXTRACTION

After VMD decomposition, a series of intrinsic mode function (IMF) components can be obtained. If the features of all components are extracted to establish the training samples for RBFNN, it will inevitably cause dimensional disaster

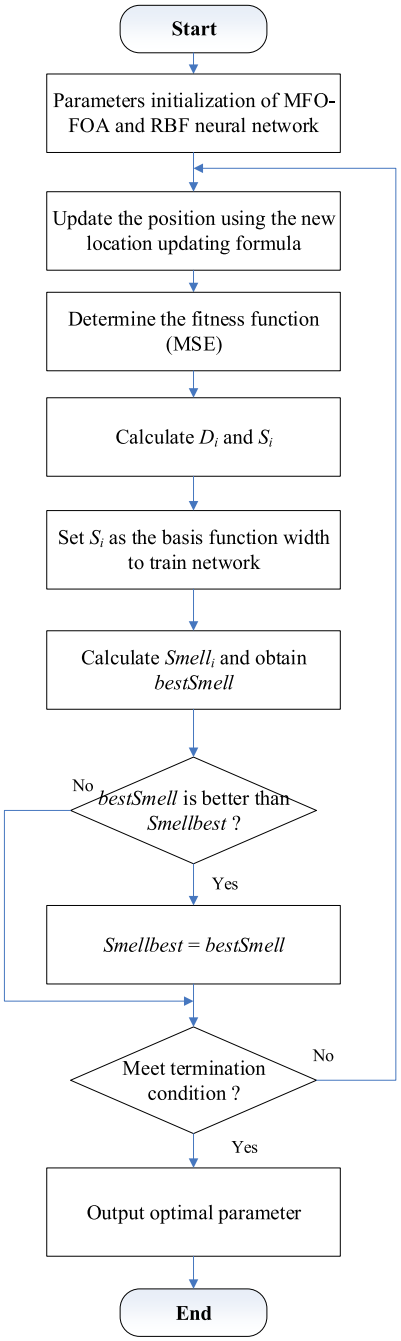


FIGURE 6. Flowchart of parameters optimization of RBFNN by using MFO-FOA.

and decrease the network performance. In addition, many studies have shown that the first several IMF components have already contained most information of the original signal [38], [39]. Therefore, this paper selects the first three IMF components to extract the key feature information. Through many comparative analyses, it is found that the envelope entropy and kurtosis of each IMF component are obviously discrepant in different coal-rock cutting states, which can be used as the features to characterize the cutting states.

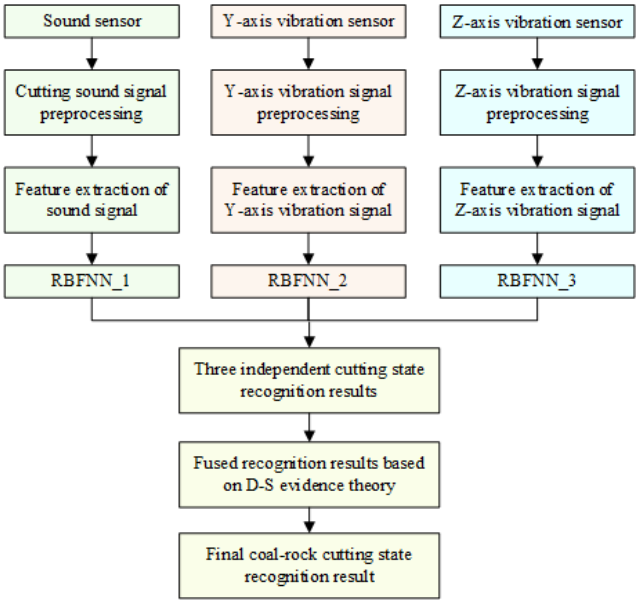


FIGURE 7. Recognition model of coal-rock cutting state based on multi-sensor information fusion.

Envelope entropy can effectively reflect the sparse characteristics of the signal. If the IMF component possesses abundant feature information, its sparsity is strong and the corresponding envelope entropy is small. For a zero-mean signal x with length N, the envelope entropy E_p can be calculated as follows:

$$\begin{cases} E_p = -\sum_{i=1}^{N} p_i lg p_i \\ p_i = a(i)/\sum_{i=1}^{N} a(i) \end{cases} \tag{22}$$

where a is the envelope of demodulation signal obtained by Hilbert transform of signal x.

The higher-order moments of signals can better reflect the weak changes. Kurtosis, belonging to the fourth-order moment of signal, is a dimensionless numerical statistic that can reflect the distribution characteristics of signal and the steepness of signal probability density function. It is very sensitive to the instantaneous characteristics of signal. the kurtosis K_u can be calculated as follows:

$$K_u = \frac{E(x-u)^4}{\sigma^4} \tag{23}$$

where μ and σ are the mean and standard deviation of signal x, respectively.

Through the above method, six eigenvalues with sensitive information are extracted and used to describe each sensing signal.

C. COAL-ROCK CUTTING STATE RECOGNITION BASED ON SINGLE-TYPE SIGNAL SOURCE AND ITS FUSION PROCESS

The improved RBF neural network is trained and tested by using the extracted samples and three network models can be constructed, marked as RBFNN_1, RBFNN_2 and RBFNN_3. The coal-rock cutting state recognition results

based on one sound sensor and two vibration sensors can be obtained. In D-S evidence theory, the three recognition results can be converted into the BPAs m_1, m_2 and m_3 of three evidences. The output vectors of RBF neural network are the judgment values of different coal-rock cutting states. In this paper, five cutting states are set up, including no-load, cutting the coal seam of hardness $f = 2$, cutting the coal seam of hardness $f = 3$, cutting coal seam with gangue and cutting rock, which can be marked as F_1, F_2, F_3, F_4, and F_5. Then the discernment framework $\Theta = \{F_1, F_2, F_3, F_4, F_5\}$. The basic probability assignment $m_i(F_j)$ of the i-th RBF neural network regarding the j-th coal-rock cutting state can be calculated as follows:

$$\begin{cases} m_i(F_j) \\ = |y_{ij}| \times R_i / \sum_{j=1}^{5} |y_{ij}| \\ m_i(\Theta) \\ = 1 - R_i \end{cases} \quad i = 1, 2, 3;\ j = 1, 2, \cdots, 5 \tag{24}$$

where y_{ij} is the actual output value of the j-th output layer neuron node of the i-th RBF neural network, R_i is the recognition accuracy of the i-th RBF neural network. $m_i(\Theta)$ is the uncertainty of the i-th RBF neural network.

Then, the correlation coefficients between each evidence are calculated by using (8), and the support matrix can be obtained. The credibility $Crd(m_i)$ is calculated according to (11). According to the credibility, the corresponding weight is allocated to each evidence, and the BPA of each evidence is processed by weighted mean method, which integrates multiple evidences into a new evidence body. The new evidence is fused by using Dempster combination rule, and the final recognition result of coal-rock cutting state can be accurately judged according to the fusion results.

V. EXPERIMENTAL VALIDATION

A. EXPERIMENTAL PLATFORM CONSTRUCTION AND SIGNAL ACQUISITION

In order to acquire more real sample data, a self-designed experimental platform for shearer cutting coal-rock is built as Fig. 8. The platform can be divided into two parts: one is the coal-rick cutting system, including shearer, scraper conveyor, hydraulic support, coal-rock specimen and fixed rack, and another part is the signal processing system, including one sound sensor, two vibration sensors (Y-axis and Z-axis), signal conditioner, data acquisition card and industrial computer. In fact, a vibration signal generally contains three directions. In the experiment, a coordinate system is established, in which the direction parallel to the hauling direction of shearer is X-axis, the direction parallel to the drum axis is Y-axis, and the direction perpendicular to the drum axis and hauling direction is Z-axis. In the process of shearer cutting coal-rock, the X-axis vibration signal is greatly affected by the hauling speed. The information contained in this vibration signal is more complex and cannot accurately characterize the coal-rock cutting state. Therefore,

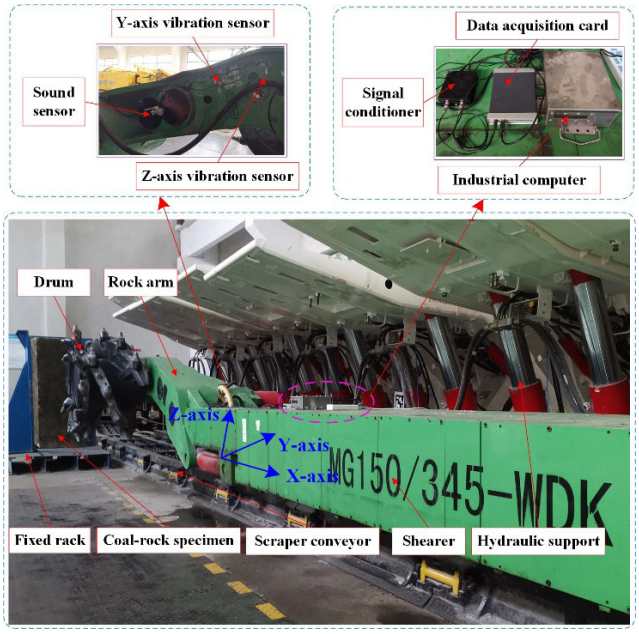


FIGURE 8. Experimental platform of shearer cutting coal-rock.

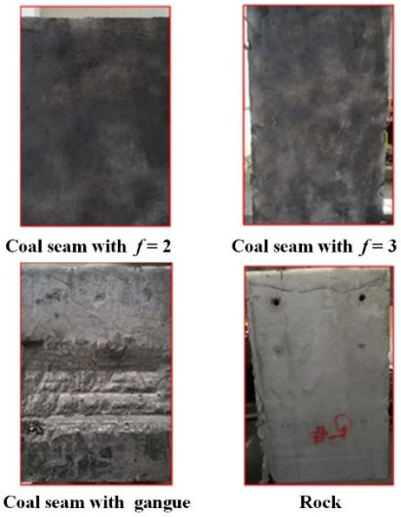


FIGURE 9. Four typical coal-rock specimens.

two vibration signals, Y-axis and Z-axis, are selected in this paper. In order to protect the sensor from being damaged during the experiment, three sensors are installed on the inside of the rocker arm (side of the hydraulic support).

Due to poor conditions of coal mining working face, it is difficult to directly acquire and transport the natural coal-rock with large-scale structure and regular form. Therefore, using different mixing ratios of pulverized coal, cement and sand, this paper developed four typical coal-rock specimens with different properties, including coal seam with hardness $f = 2$, coal seam with hardness $f = 3$, coal seam with gangue, and rock, as shown in Fig. 9. Thus, totally five coal-rock cutting states can be obtained, namely F_1, F_2, F_3, F_4, and F_5. The corresponding class labels are listed in Table 3.

TABLE 3. The coal-rock cutting states of shearer.

Cutting State	Coal-rock Specimen	Class Label
F_1	None	1
F_2	Coal seam with hardness $f = 2$	2
F_3	Coal seam with hardness $f = 3$	3
F_4	Coal seam with gangue	4
F_5	Rock	5

During the experiment, the hauling speed of shearer was set to 2 m/min, the acquisition frequency of sound signal was 44.1 kHz, and the acquisition frequency of vibration signal was 10 KHz. The data length of each sample was 0.5 s, and the time-domain waveforms under different cutting states collected by three sensors are shown in Fig. 10.

B. COAL-ROCK CUTTING STATE RECOGNITION BASED ON IMPROVED RBF NEURAL NETWORK WITH SINGLE SIGNAL SOURCE

In the experiment, 60 sets of sample data are selected from three sensors for each cutting state, and 300 sets of data samples for cutting sound signal, Y-axis and Z-axis vibration signals are obtained respectively. After feature extraction, three feature matrices of 300×6 can be acquired. 200 groups of eigenvectors (40 groups for each cutting state) are randomly selected from the three matrices as training samples and the remaining 100 vectors are used as testing samples as testing samples.

To verify the effectiveness of FOA-FOA and the proposed recognition method (marked as MFO-FOA-RBFNN), the effects of FOA, MFO and MFO-FOA optimizing the network parameters of RBFNN are compared. The relevant parameters are set as follows: the size of the fruit fly population and the moth population is 30, the maximum number of iterations is 100, the number of input layer neurons is 6, and the number of output layer neurons is 5. The expected outputs corresponding to five cutting state are [1 0 0 0 0], [0 1 0 0 0], [0 0 1 0 0], [0 0 0 1 0], and [0 0 0 0 1], respectively. RBF neural network is created by using *newrb* function, and the basic function width is optimized by FOA, MFO and MFO-FOA respectively. The iterative curves of the optimal fitness values with the cutting sound signal as the training sample are shown in Fig. 11.

It can be observed from Fig. 11 that MFO-FOA algorithm has the best effect on the optimization of basis function width,

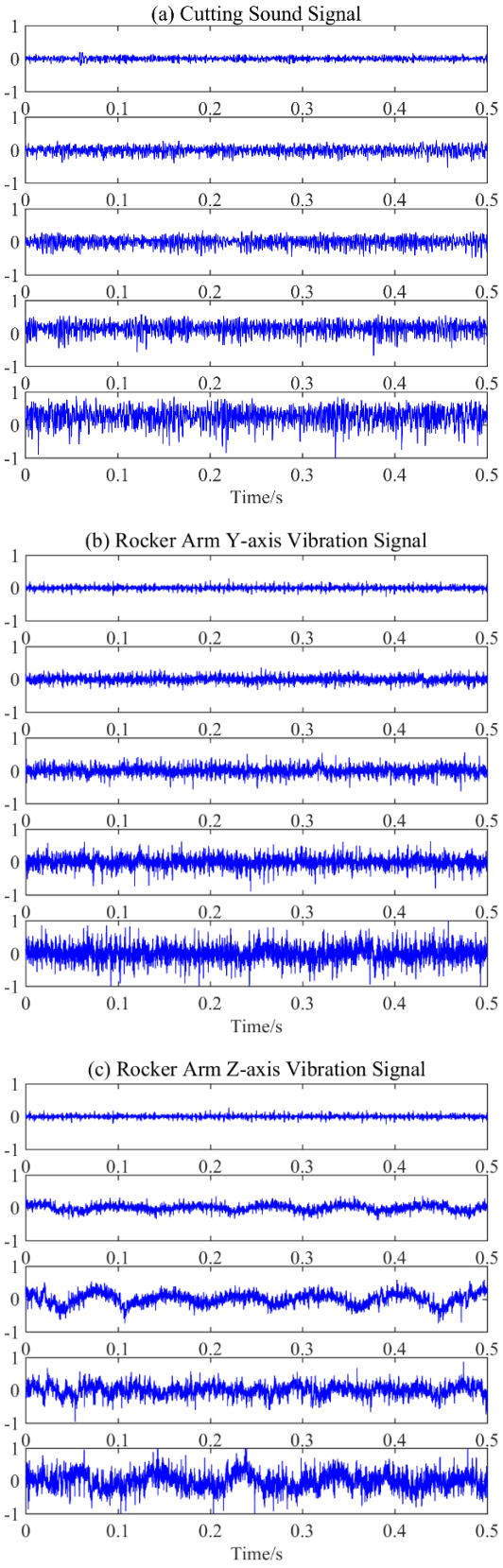


FIGURE 10. Time domain waveforms collected through three sensors.

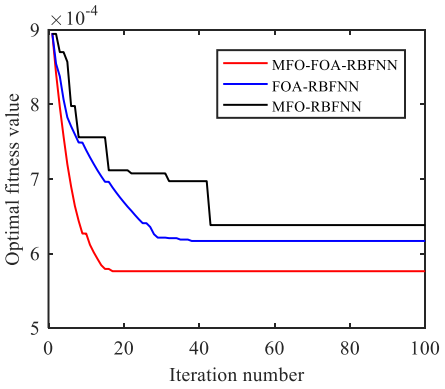


FIGURE 11. The iterative curves of the optimal fitness values with the cutting sound signal as the training sample.

TABLE 4. Optimization results of different algorithms.

Algorithms	MFO-FOA	FOA	MFO
Basic function width	0.0141	2.3889	2.3781
Optimal fitness value	5.7629×10^{-4}	6.1687×10^{-4}	6.3818×10^{-4}

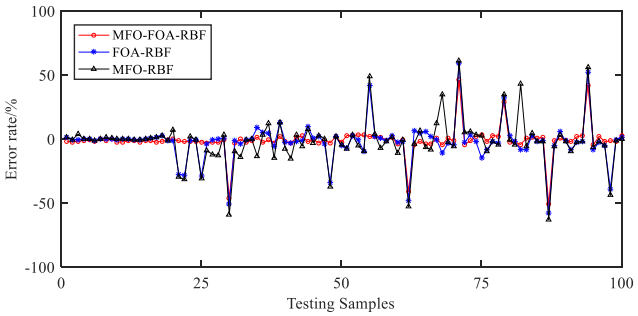


FIGURE 12. Error rates between the expected output and the actual output of three networks.

and can converge to the optimal value faster than FOA and MFO, with higher convergence accuracy. The specific optimization results are listed in Table 4. The number of hidden layer neurons in the optimized RBF neural network is 50, 52 and 52, respectively.

Then, the 100 testing samples of cutting sound signal are used to verify the recognition effect of three RBF neural networks. The error rates between the expected output value and the actual output value are plotted as Fig. 12, and the cutting state recognition results are shown in Fig. 13.

By observing Fig. 12, the error curve of MFO-FOA-RBFNN manifests smaller fluctuations than that of FOA-RBFNN and MFO-RBFNN. The mean error of proposed method is only 4.22%, which is obviously lower than FOA-RBFNN (7.34%) and MFO-RBFNN (10.21%). The results demonstrate that the hybrid algorithm of MFO-FOA

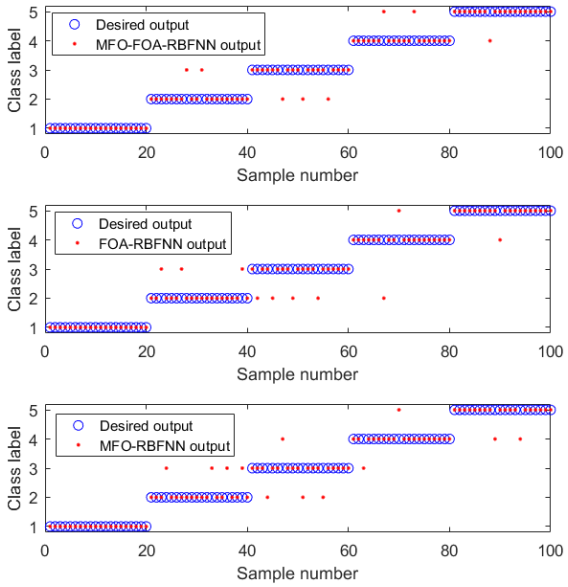


FIGURE 13. Coal-rock cutting state recognition results based on cutting sound signal.

performs better search ability and optimization performance than MFO and FOA. In addition, seen from Fig. 13, eight samples are misclassified by using proposed method and the recognition accuracy in the testing phase can reach to 92 percent, which is higher than FOA-RBFNN (90%) and MFO-RBFNN (88%). It can also be found from Fig. 13 that the misclassified samples mostly occur among the states of cutting coal-rock specimens with $f = 2$ and $f = 3$. By using the proposed MOF-FOA algorithm, a better basis function width can be acquired, so that the RBFNN can still maintain better recognition performance when distinguishing two kinds of coal-rock with similar properties. The comparison results indicate the proposed recognition method possesses good practicality and effectiveness.

For other two vibration signals of Y-axis and Z-axis, the simulation parameters are the same as the sound signal. Thus, the iterative curves of the optimal fitness values by using three signals can be illustrated as Fig. 14. The cutting state recognition results of MFO-FOA-RBFNN based on two vibration signals are shown in Fig. 15 and Fig. 16. It can be observed from Fig. 14 that three curves all tend to converge at about 20 iterations, proving fast search speed and superior convergence performance of proposed MFO-FOA. When using Y-axis and Z-axis vibration signals to identify the coal-rock cutting state, ten and eight samples are misclassified, and the recognition accuracies are 90% and 92% respectively. Obviously, the recognition accuracy of coal-rock cutting state based on single signal source is still a little low, which cannot meet the actual requirements.

Finally, the comparative analysis of RBFNN with other popular kernel functions, such as Rational Quadratic kernel and Multiquadric kernel, is provided. The simulation results are listed in Table 5. Seen from Table 5, the recognition results

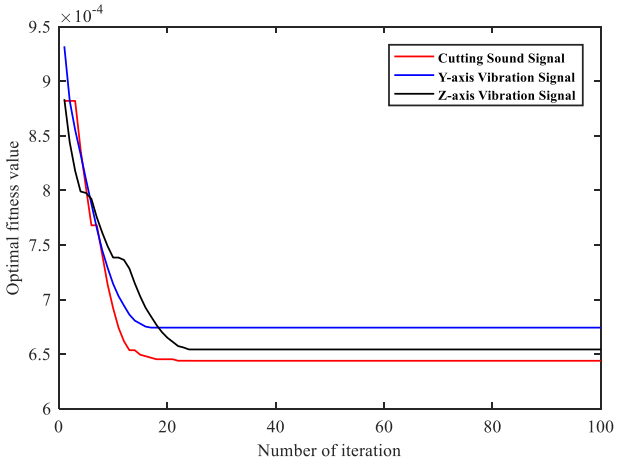


FIGURE 14. The iterative curves of the optimal fitness values based on proposed method.

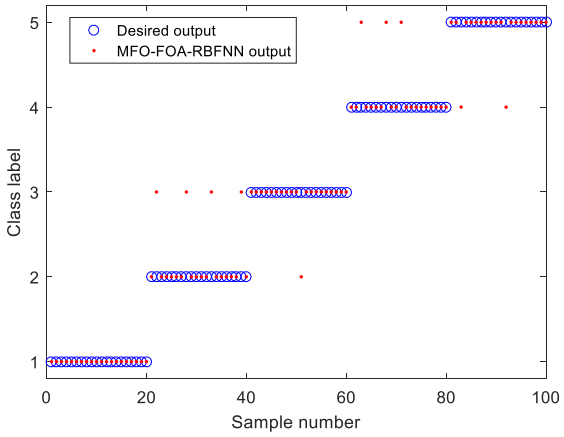


FIGURE 15. Coal-rock cutting state recognition results based on Y-axis vibration signal.

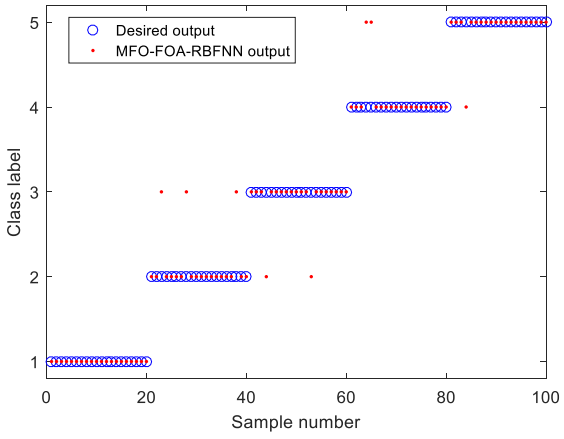


FIGURE 16. Coal-rock cutting state recognition results based on Z-axis vibration signal.

of RBFNN with other kernel functions are basically consistent with Gaussian kernel function, which can be concluded that three kernels have similar mapping performance and can be used as an alternative when using the Gaussian becomes too computationally intensive.

TABLE 5. Recognition accuracy of RBFNN with different kernel functions.

Kernel function	Gaussian	Rational Quadratic	Multiquadric
Sound signal	92%	91%	92%
Y-axis vibration signal	90%	90%	91%
Z-axis vibration signal	92%	92%	91%

TABLE 6. RBFNN output results based on single signal source.

Cutting State	F_1	F_2	F_3	F_4	F_5
Sound signal	-9.9269×10^{-4}	0.9272	-0.0626	0.0084	0.0105
Y-axis vibration signal	0.1093	-0.0320	0.7292	-0.0230	-0.0013
Z-axis vibration signal	-0.0919	0.8516	0.0126	-0.0039	0.0116

TABLE 7. The BPA of each evidence.

Evidence	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_5)$	$m(\Theta)$
1	9.2810 ×10^{-4}	0.8669	0.0585	0.0079	0.0098	0.0560
2	0.1143	0.0335	0.7628	0.0241	0.0014	0.0880
3	0.0900	0.8344	0.0123	0.0038	0.0114	0.0720

C. FUSION RECOGNITION

In order to improve the recognition accuracy, the D-S evidence theory is used to fuse the independent identification results based on three signal sources. To describe the fusion process in detail, a group of data sample with conflicting recognition results is selected as an example, shown in Table 6. Obviously, the network recognition result based on Y-axis vibration signal is inconsistent with the recognition results of other two networks based on Z-axis vibration signal and sound signal, and the results are conflicting. By using the D-S evidence theory with evidence correlation coefficient, the decision-level fusion process of the output results is as follows.

(1) Calculate the BPA of each evidence. The output results of RBFNNs are taken as three evidences respectively, and the BPA of each evidence is calculated by using (24), as shown in Table 7.

TABLE 8. The credibility of each evidence.

Evidence	m_1	m_2	m_3
Credibility(Crd)	0.4570	0.1007	0.4423

TABLE 9. The BPAs of new evidence and fused results.

Evidence	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_5)$	$m(\Theta)$
BPA	0.0517	0.7686	0.1090	0.0077	0.0097	0.0533
Fused results	0.0115	0.9485	0.0331	0.0012	0.0016	0.0040

(2) Calculate the credibility of each evidence. Firstly, the correlation coefficient between each evidence should be computed to create the support degree matrix.

$$\begin{aligned} S_{12} &= S_{21} = r_{BPA}(m_1, m_2) = 0.1435, \\ S_{13} &= S_{31} = r_{BPA}(m_1, m_3) = 0.9932, \\ S_{23} &= S_{32} = r_{BPA}(m_2, m_3) = 0.1070, \\ SM &= \begin{bmatrix} 1.0000 & 0.1435 & 0.9932 \\ 0.1435 & 1.0000 & 0.1070 \\ 0.9932 & 0.1070 & 1.0000 \end{bmatrix}. \end{aligned}$$

Then the credibility can be obtained by using (11) and the results are listed in Table 8.

(3) According to the credibility of each evidence, the corresponding weight is assigned to each evidence, and the BPAs are processed by using the weighted mean method. Then, three evidences are merged into a new evidence and the obtained new evidence is fused according to Dempster's combination rule. The results are shown in Table 9.

From Table 8, the fused BPA value of $m(F_2)$ can reach to 0.9485, indicating that the current coal-rock cutting state belongs to the F_2 category, that is, shearer is cutting the coal-rock specimen with hardness $f = 2$, which is consistent with the actual working condition.

Together with the fusion results, the final recognition results of 100 testing samples can be determined correctly, which can be shown as Fig. 17. From this figure, it can be observed that only two samples are misclassified and the recognition accuracy can increase to 98%, which is significantly higher than the one with respect to the single-type signal features. These encouraging results clearly indicate that the well-designed multi-sensor information fusion identification method can significantly improve the prediction accuracy of the shearer cutting coal-rock system.

D. COMPARISON WITH OTHER METHODS

In order to analyze the difference in terms of recognition accuracy between the proposed method and other methods in our research team, it is compared with the

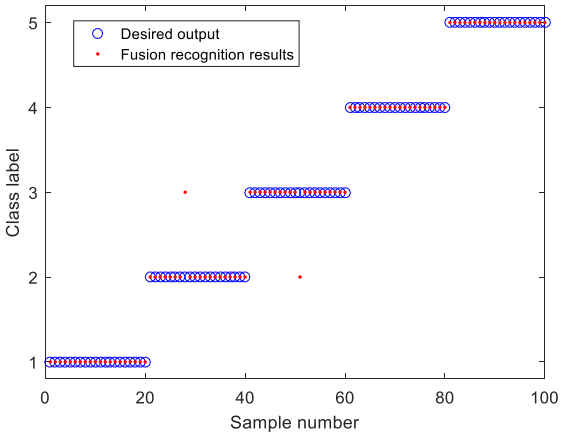


FIGURE 17. The fusion recognition results of shearer coal-rock cutting state based on three-type sensor data.

TABLE 10. Performance comparisons with other methods.

Method Name	Class 1	Class 2	Class 3	Class 4	Class 5	*Accuracy*
Ref. [19]	20	18	19	17	19	93%
Ref. [20]	20	17	18	18	19	92%
Ref. [21]	20	18	17	19	17	91%
Proposed method	20	19	19	20	20	98%

methods in [19]–[21]. In [19], the vibration signals are taken as analytic objects and are decomposed by local mean decomposition. The time-frequency features are extracted and the shearer cutting states are classified by the fuzzy C-means clustering algorithm. In [20], the cutting sound signals are taken as analytic objects and are decomposed by the improved EEMD. The probabilistic neural network is used as the classifier after obtaining the features. In [21], the temperature of cutting area is used as the analytic object and the cutting state is identified by using SVM. Experimental scheme is the same as above simulations. The examined results of four methods are summarized in Table 10.

It can be observed from Table 10 that the shearer coal-rock cutting state recognition accuracy of proposed method is obviously higher than that of other three methods in [19]–[21]. Although superior signal processing and feature extraction algorithms are used in both [19] and [20], there is also a certain amount of misjudgment, mainly due to not enough characteristic information in single-type sensor data. In addition, because of the hysteresis of temperature transfer, the temperature in cutting area may not reflect different coal-rock cutting states in time. This phenomenon makes ref. [21] perform the worst recognition effect.

In summary, the fusion recognition method based on RBFNN and D-S evidence theory can achieve better classification results and outperforms the competing approaches. For conflicting information or error information caused by sensor failure, the proposed fusion system can solve the fault tolerance and respond appropriately.

VI. CONCLUSION

In this study, the sound signals and vibration signals both are taken as analytic objects, and a novel method for coal-rock cutting state recognition of shearer based on improved RBFNN and D-S evidence theory is proposed. From theoretical analyses and experimental results, the major contributions of this study are summarized as follows:

(i) Based on the merits and demerits of FOA and MFO, FOA is improved by using the location updating mechanism of MFO and a hybrid swarm intelligence optimization algorithm of MFO-FOA is proposed in this study. The simulation results verify the effectiveness and superiority of MFO-FOA algorithm.

(ii) The sound signals and vibration signals are decomposed by VMD, and the envelope entropy and kurtosis are used to achieve feature extraction. The optimal network parameters of RBFNN are found out by using proposed MFO-FOA and the coal-rock cutting state is accordingly recognized based on the feature information of single-type signal source.

(iii) Based on the idea of integrated decision-making, a fusion model is constructed based on D-S evidence theory to compensate for the inconsistency of single recognition results. The corresponding experiments show that the recognition accuracy with respect to the fusion model can reach to 98%, which is significantly higher than the one with respect to the single-type signal features.

Under the same experimental samples, the performance of the proposed method is compared with other methods in [19]–[21]. It can conclude that the proposed method has certain improvements in terms of the pattern recognition and the higher recognition accuracy verifies the excellent performance of the proposed method.

REFERENCES

[1] G. Wang, G. R. Zhao, and H. W. Ren, "Analysis on key technologies of intelligent coal mine and intelligent mining," *J. China Coal Soc.*, vol. 44, no. 1, pp. 34–41, Jan. 2019.

[2] G. F. Wang and Y. B. Du, "German industry 4.0 and intelligent manufacturing development of Chinese coal machine equipment," *Coal Sci. Technol.*, vol. 47, no. 3, pp. 1–9, Mar. 2019.

[3] I. Y. Bausov, G. L. Stolarczyk, L. G. Stolarczyk, and S. D. S. Koppenjan, "Look-ahead radar and horizon sensing for coal cutting drums," presented at the 4th Int. Workshop, Adv. Ground Penetrating Radar, Jun. 2007, pp. 208–211. [Online]. Available: https://ieeexplore.ieee.org/document/4278876

[4] R. Sahoo and A. M. Mazid, "Application of opto-tactile sensor in shearer machine design to recognise rock surfaces in underground coal mining," presented at the IEEE Int. Conf. Ind. Technol., Feb. 2009, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4939645

[5] J. Asfahani and M. Borsaru, "Low-activity spectrometric gamma-ray logging technique for delineation of coal/rock interfaces in dry blast holes," *Appl. Radiat. Isot.*, vol. 65, no. 6, pp. 748–755, Jun. 2007.

[6] L. Li and C. P. Ouyang, "Research on coal-rock interface recognition based on ultrasonic phased array," *J. China Univ. Mining Technol.*, vol. 46, no. 3, pp. 485–492, May 2017.

[7] Q. Zhang, H. J. Wang, T. Guo, and Y. F. Liu, "Study on coal-rock interface recognition of coal shearer based on cutting infrared thermal image of picks," *Coal Sci. Technol.*, vol. 45, no. 5, pp. 22–27, May 2017.

[8] Q. Zhang, H. J. Z. H. Liu, Wang, Y. Tian, C. H. Huang, "Study on coal and rock identification based on vibration and temperature features of picks," *Coal Sci. Technol.*, vol. 46, no. 3, pp. 1–9, Mar. 2018.

[9] G. Zhang, Z. Wang, L. Zhao, Y. Qi, and J. Wang, "Coal-rock recognition in top coal caving using bimodal deep learning and Hilbert-Huang transform," *Shock Vib.*, vol. 2017, Jul. 2017, Art. no. 3809525.

[10] J.-L. Xie and J.-L. Xu, "Ground penetrating radar-based experimental simulation and signal interpretation on roadway roof separation detection," *Arabian J. Geosci.*, vol. 8, no. 3, pp. 1273–1280, Mar. 2015.

[11] N. Zhang and C. Liu, "Radiation characteristics of natural gamma-ray from coal and gangue for recognition in top coal caving," *Sci. Rep.*, vol. 8, no. 1, p. 190, 2018.

[12] X. Wang, S. G. Miao, and E. J. Ding, "Study of dielectric property of coal and rock medium in Terahertz domain," *J. China Univ. Mining Technol.*, vol. 45, no. 4, pp. 1–8, Apr. 2016.

[13] C. Tan, R. X. Xu, Z. Wang, L. Si, and X. Liu, "An improved genetic fuzzy logic control method to reduce the enlargement of coal floor deformation in shearer memory cutting process," *Comput. Intell. Neurosci.*, vol. 2016, Jan. 2016, Art. no. 3973627.

[14] W. Li, C. Luo, H. Yang, and Q. Fan, "Memory cutting of adjacent coal seams based on a hidden Markov model," *Arabian J. Geosci.*, vol. 7, no. 12, pp. 5051–5060, 2014.

[15] Q.-G. Fan, W. Li, Y.-Q. Wang, M.-B. Fan, and X.-F. Yang, "An algorithm of shearer memory cutting based on grey-Markovian model," *J. Central South Univ.*, vol. 42, no. 10, pp. 3054–3058, Oct. 2011.

[16] B. Tiryaki and A. C. Dikmen, "Effects of rock properties on specific cutting energy in linear cutting of sandstones by picks," *Rock Mech. Rock Eng.*, vol. 39, no. 2, pp. 89–120, Apr. 2006.

[17] H. Basarir, C. Karpuz, and L. Tutluoglu, "Specific energy based rippability classification system for coal measure rock," *J. Terramechanics*, vol. 45, nos. 1–2, pp. 51–62, Feb./Apr. 2008.

[18] Z.-P. Xu and Z.-B. Wang, "Characteristic analysis of shearer cutting load based on particle filter," *J. China Coal Soc.*, vol. 36, no. 4, pp. 696–700, Apr. 2011.

[19] L. Si, Z. B. Wang, C. Tan, and X. H. Liu, "Vibration-based signal analysis for shearer cutting status recognition based on local mean decomposition and fuzzy C-means clustering," *Appl. Sci.*, vol. 7, no. 2, p. 164, Feb. 2017.

[20] J. Xu, Z. B. Wang, C. Tan, L. Si, and X. Liu, "A cutting pattern recognition method for shearers based on improved ensemble empirical mode decomposition and a probabilistic neural network," *Sensors*, vol. 15, no. 11, pp. 27721–27737, Oct. 2015.

[21] L. Si, Z. B. Wang, Y. W. Liu, and C. Tan, "Online identification of shearer cutting state using infrared thermal images of cutting unit," *Appl. Sci.*, vol. 8, no. 10, p. 1772, Sep. 2018.

[22] K. Li, R. Zhang, F. Li, L. Su, H. Wang, and P. Chen, "A new rotation machinery fault diagnosis method based on deep structure and sparse least squares support vector machine," *IEEE Access*, vol. 7, pp. 26571–26580, 2019.

[23] J.-H. Zhong, J. Zhang, J. Liang, and H. Wang, "Multi-fault rapid diagnosis for wind turbine gearbox using sparse Bayesian extreme learning machine," *IEEE Access*, vol. 7, pp. 773–781, 2018.

[24] S. Liu, Y. Sun, and L. Zhang, "A novel fault diagnosis method based on noise-assisted MEMD and functional neural fuzzy network for rolling element bearings," *IEEE Access*, vol. 6, pp. 27048–27068, 2018.

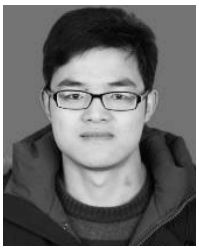
[25] G. Foody, "Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes," *Int. J. Remote Sens.*, vol. 25, no. 15, pp. 3091–3104, 2004.

[26] F. Yang, "Image classification algorithm based on deep neural network and multi-layer feature learning," *Basic Clin. Pharmacol.*, vol. 124, no. 3, p. 287, Apr. 2019.

[27] Z.-Y. Chen and R. J. Kuo, "Combining SOM and evolutionary computation algorithms for RBF neural network training," *J. Intell. Manuf.*, vol. 30, no. 3, pp. 1137–1154, Mar. 2019.

[28] S. N. Qasem, S. M. Shamsuddin, and A. M. Zain, "Multi-objective hybrid evolutionary algorithms for radial basis function neural network design," *Knowl.-Based Syst.*, vol. 27, pp. 475–497, Mar. 2012.

[29] H. Hamdi, C. Ben Regaya, and A. Zaafouri, "Real-time study of a photovoltaic system with boost converter using the PSO-RBF neural network algorithms in a MyRio controller," *Sol. Energy*, vol. 183, pp. 1–16, May 2019.
[30] J. B. Li and X. Liu, "Melt index prediction by RBF neural network optimized with an adaptive new ant colony optimization algorithm," *J. Appl. Polym. Sci.*, vol. 119, no. 5, pp. 3093–3100, Mar. 2011.
[31] M. Awad, "Forecasting of chaotic time series using RBF neural networks optimized by genetic algorithms," *Int. Arab Inf. Technol.*, vol. 14, no. 6, pp. 826–834, Nov. 2017.
[32] R. Dong, S. Liu, and G. Liang, "Research on control parameters for voltage source inverter output controllers of micro-grids based on the fruit fly optimization algorithm," *Appl. Sci.*, vol. 9, no. 7, p. 1327, Apr. 2019.
[33] S.-X. Lv, Y.-R. Zeng, and L. Wang, "An effective fruit fly optimization algorithm with hybrid information exchange and its applications," *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 10, pp. 1623–1648, Oct. 2018.
[34] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowl.-Based Syst.*, vol. 89, pp. 228–249, Nov. 2015.
[35] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976, pp. 10–40.
[36] W.-T. Pan, "A new fruit fly optimization algorithm: Taking the financial distress model as an example," *Knowl.-Based Syst.*, vol. 26, pp. 69–74, Feb. 2012.
[37] J. Xu, Z. B. Wang, C. Tan, L. Si, L. Zhang, and X. H. Liu, "Adaptive wavelet threshold denoising method for machinery sound based on improved fruit fly optimization algorithm," *Appl. Sci.*, vol. 6, no. 7, p. 199, Jul. 2016.
[38] X. Chen, Y. Yang, Z. Cui, and J. Shen, "Vibration fault diagnosis of wind turbines based on variational mode decomposition and energy entropy," *Energy*, vol. 174, pp. 1100–1109, May 2019.
[39] Z. Wang, J. Wang, and W. Du, "Research on fault diagnosis of gearbox with improved variational mode decomposition," *Sensors*, vol. 18, no. 10, p. 3510, Oct. 2018.
[40] Y. Zhao, F. Ding, J. Li, L. Guo, and W. Qi, "The intelligent obstacle sensing and recognizing method based on D–S evidence theory for UGV," *Future Gener. Comput. Syst.*, vol. 97, pp. 21–29, Aug. 2019.
[41] J. Lu, H. Zhang, and X. Tang, "A novel method for intelligent single fault detection of bearings using SAE and improved D–S evidence theory," *Entropy*, vol. 21, no. 7, p. 687, Jul. 2019.
[42] P. Xiuqin, C. Yongcun, L. Yong, L. Xiali, and Z. Yue, "Research on the algorithm of target recognition based on two-leveled RBF neural network and D-S evidence theory," in *Proc. Int. Symp. Intell. Signal Process. Commun.*, Tottori, Japan, Dec. 2006, pp. 911–914.



LEI SI was born in Xuzhou, China, in 1987. He received the B.S. degree in mechanical engineering and automation and the Ph.D. degree in mechanical engineering from the China University of Mining and Technology, in 2010 and 2015, respectively, where he is currently an Associate Professor. He has authored or coauthored about 15 journal articles (SCI) and holds about seven patents. His research interests include intelligent control and fault diagnosis of coal mining equipment.

ZHONG-BIN WANG was born in Suzhou, China, in 1972. He received the B.S. degree in mechanical engineering and automation from the China University of Mining and Technology, in 1995, and the Ph.D. degree in mechanical engineering from the Nanjing University of Aeronautics and Astronautics, in 2002. He is currently the Dean of the School of Mechatronic Engineering, China University of Mining and Technology. He was selected as the training object of the "Blue Project" in Jiangsu Province, in 2010, as the training object of the "333 High-level Talents Training Project" in Jiangsu Province (the third level), in 2011, and as the training object of the "Six Talents Peaks" in Jiangsu Province, in 2012. He has authored or coauthored 40 SCI articles and holds more than 50 authorized invention patents. He has and trained more than 20 doctoral and master's graduate students. His research interest includes remote monitoring and intelligent control of mechanical and electrical equipment in coal mining face.

GAN JIANG was born in Suzhou, China, in 1994. He received the B.S. degree in mechanical engineering and automation from the China University of Mining and Technology, in 2016, where he is currently pursuing the M.S. degree. His research interest includes the working state recognition of coal mining machine.