

## Research Article

# A Multimodel Decision Fusion Method Based on DCNN-IDST for Fault Diagnosis of Rolling Bearing

Weixiao Xu , Luyang Jing , Jiwen Tan, and Lianchen Dou

*College of Mechanical Engineering, Qingdao University of Technology, Qingdao 266520, China*

Correspondence should be addressed to Luyang Jing; [jingluyang@qut.edu.cn](mailto:jingluyang@qut.edu.cn)

Received 27 August 2020; Accepted 27 November 2020; Published 15 December 2020

Academic Editor: Zhixiong Li

Copyright © 2020 Weixiao Xu 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.

Each pattern recognition method has its advantages and disadvantages to diagnose the state of rotating machinery. There are many fault types of rolling bearings with apparent uncertainty. The optimal fusion level is usually challenging to be selected for a specific fault diagnosis task, and extensive human labour and prior knowledge are also highly required during these selections. To solve the above problems, a multimodel decision fusion method based on Deep Convolutional Neural Network and Improved Dempster-Shafer Evidence Theory (DCNN-IDST) is proposed for the inspection of rolling bearing. To solve the defect of the original evidence theory method in the fusion of high-conflict evidence, the fuzzy consistency matrix is introduced. By calculating the factor weight, the reliability and rationality of D-S evidence theory are improved. The DCNN model can learn features from the original data and carry out adaptive feature extraction for multiple sensor information. The features extracted by DCNN adaptively are input into multiple network models for decision fusion. The new method of DCNN-IDST multimodel decision fusion is applied to detect the damage of rolling bearings. To evaluate the effectiveness of the proposed method, both the BP neural network and RBF neural network are used to set up a multigroup comparison test. The result demonstrates that the proposed method can detect the fault of the rolling bearing effectively and achieve the highest diagnosis accuracy among all the tested methods in the experiment.

## 1. Introduction

Mechanical fault diagnosis technology is a newly developing subject to monitor, diagnose, and predict the condition and the malfunction of continuously operating mechanical equipment to guarantee its safe operation. It can effectively improve the safety and the stability of rotating machinery equipment [1]. The rolling bearing plays an essential role in the running process of mechanical equipment, where its status will directly affect the running efficiency and life of the mechanical system. According to statistics, CNC machining by rolling bearings caused by mechanical failure accounted for about 30% of the total fault. It can be seen that timely diagnosis of the status information and fault diagnosis of rolling bearing is vital, which can reduce maintenance costs and ensure the regular operation of equipment.

There are many fault types of rolling bearing with apparent uncertainty. It is impossible to predict and diagnose

the bearing state comprehensively by collecting a single sensor data. Therefore, data from different sources are combined to achieve the purpose of effective diagnosis of complex systems [2, 3]. Dempster-Shafer evidence theory (DST) has the advantage of expressing “uncertain” and “unknown” without knowing prior probability, which has been widely used in information fusion technology [4, 5]. Niu et al. [6] collected the current signal and the vibration signal of electrical machinery, integrating voting decisions with Bayes’ detective method, raising the multiagent decision-making layer blending approach, to incorporate the information from the results of multisensor signals’ malfunction diagnosis. Li et al. using D-S evidence theory to process decision-making layer’s information integrate thus to diagnose the malfunction of the diesel engine [7]. Zhang et al. adopt the SVDD approach to improve D-S evidence theory, setting up a two-stage fusion model to process experimental verification in the malfunction diagnosis for

chisel engine and finally achieve hierarchical and multilevel malfunction diagnosis [8]. Cao et al. improved D-S evidence theory, set up the integrating relations between each evidence, and increased the recognition rate of malfunction diagnosis according to the problem of highly conflict evidence which appeared in malfunction diagnosis for large-scale equipment [9]. Liu et al. adopted the method of combining quantum medical neural network with D-S evidence theory to achieve the high precision diagnosis for mechanical characteristics of circuit breakers [10]. Jiang and Lin et al. attained the goal of improving D-S evidence theory by changing the combination rules in the evidence theory, which also had specific application effects [11, 12].

In the last few years, deep learning has played an essential role in the field of artificial intelligence. Deep convolutional neural network (DCNN) has strong ability of data mining and information integration [13–19]. It has been widely applied to the state monitoring and diagnosis research of rotating machinery. Ga et al. classified the faults of rolling bearings by analyzing vibration signals with wavelet transform and extracting features and feeding them into the deep learning model [17]. Jing et al. used the deep convolution model to adaptively extract the original fault signal characteristics and identified the fault of the motor [19]. Zhao et al. combined the depth convolution model with the LSTM model to monitor and estimate tool wear and achieved good fault diagnosis results [20].

As an emerging machine learning method, deep learning has strong capabilities of feature extraction and function mapping. Deep learning method can meet the analysis requirements of diverse, nonlinear, and high-dimensional health monitoring data, which can be applied in the life prediction of rotating machinery equipment. Babu GS et al. [21] first used the Convolution Neural Network (CNN) to predict the residual life of bearings and verified the effectiveness of the method. Li [22] combined the convolutional neural network with the short-time memory neural network and used CNN's features of convolution and weight sharing to extract deep features and input them into the LSTM network, effectively realizing the prediction of the remaining service life of the rolling bearing. Zhao et al. [20] combined the depth convolution model with the LSTM model to monitor and estimate tool wear, which achieved good fault diagnosis results. Shi et al. [23], based on complementing the advantages of CNN and LSTM, proposed the space-time series prediction method based on ConvLSTM. Luo et al. [24] proposed a ConvLSTM-AE framework which better encodes the change of appearance and motion for regular events. Qiao et al. used the TD-ConvLSTM time series model to analyze the space-time characteristics of multisensor data on different time scales [25].

There are many fault types of rolling bearings with apparent uncertainty. It is often challenging to select the optimal feature or fusion level in state monitoring. Therefore, combining the advantages of multiple intelligent identification methods, this paper proposes a multimodel decision fusion method based on DCNN-IDST for fault diagnosis of rolling bearing. By this method, the fault

information of rotating machinery can be diagnosed comprehensively.

## 2. Theoretical Background

*2.1. Dempster–Shafer Evidence Theory.* Dempster–Shafer theory has the advantage of dealing with uncertainty, and its basic formula is shown as follows.

- (1) Suppose there are proposition  $A$  and  $U$  as the framework:  $2^U \rightarrow [0, 1]$ .

There is a function  $m$  that satisfies the following conditions:

$$U = \{x_1, x_2, x_3 \dots\}, \quad (1)$$

$$m(\phi) = 0, \quad (2)$$

where  $m(A)$  is the Basic Probability Assignment, which represents the exact trust function for  $A$ .

The belief function can be defined as follows:

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B). \quad (3)$$

The plausibility functions are set to

$$\text{Pls}(A) = 1 - \text{Bel}(A), \quad (4)$$

where  $\text{Pls}(A)$  and  $\text{Bel}(A)$  are the upper and lower limit function of  $A$ . The relationship between them is shown as follows:  $\text{Pls}(A) \geq \text{Bel}(A)$ ,  $A \subseteq \Omega$ .

- (2) The combination rule of Dempster–Shafer evidence theory [26–29].

Suppose that the basic probability assignment of frame  $U$  is  $m_1$  and  $m_2$ ; if  $A \in U$  and  $m(A) > 0$ , the focal target elements are  $A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_n$ .

The combination rule of Dempster–Shafer evidence theory is defined as follows:

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i)m_2(B_j)}{1 - \sum_{A_i \cap B_j = \phi} m_1(A_i)m_2(B_j)}. \quad (5)$$

- (3) The new improved method of Dempster–Shafer evidence theory (IDST).

In this paper, the improved DST method proposed in the literature [30] is adopted to diagnose the state of rotating machinery. The specific steps are as follows.

Assuming  $m_1, m_2, \dots, m_n$  as the Basic Probability Assignment as the framework of  $U$ , the framework is  $\Theta = \{A, B, C, \dots\}$ .

The IDST method is defined as follows.

- (1) The original basic probability assignment is redistributed, and a new fuzzy consistent matrix is established.

Add the fuzzy matrix in rows:

$$R = (r_{ij})M \times N, \quad (6)$$

$$r_i = \sum_{k=1}^n r_{ik}, \quad (i = 1, 2, \dots, n). \quad (7)$$

On this basis, the consistent fuzzy matrix is obtained by transformation according to the following formula:

$$r_{ij} = \frac{(r_{ik} - r_{jk})}{2n + 0.5}. \quad (8)$$

- (2) Figure out the factor weight coefficient.

Suppose there is factor  $A_i$  and target of  $O_k$ , and the weight  $s_i^k$  is as follows:

$$s_i^k = \frac{1}{n} - \frac{1}{2\alpha} + \frac{\sum_{j=1}^n r_{ij}}{n\alpha}, \quad i = 1, 2, \dots, n, \quad (9)$$

where  $\alpha \geq (n-1)/2$ , to improve the resolution of the sorting result, and the values are  $\alpha = (n-1)/2$ .

- (3) Recalculate the basic probability allocation.

The adjusted value of basic probability distribution is defined as follows:

$$mi(A_K)' = s_i^k mi(A_K), \quad (10)$$

where  $k = 1, 2, 3, \dots$

Since the sum of the changed basic probability distribution values is not 1, the supplementary definition is as follows, to constitute the basic probability distribution function:

$$mi(\Theta)' = 1 - \sum_{k=1}^{d_i} mi(A_K)'. \quad (11)$$

- (4) Get the average evidence:

$$\bar{m}(A) = \frac{1}{n} \sum_{i=1}^n m_i(A), \quad \forall A \subseteq \Theta. \quad (12)$$

- (5) The improved evidence theory formula, as shown below:

$$m * (A) = \sum_{A_i \cap B_j = A} m_1(A_i) * m_2(B_j) + \bar{m}(A), \quad (13)$$

$$m(\theta) = 1 - \sum m(A), \quad (14)$$

where  $\bar{m}(A)$  is the moderate support degree after weighted the evidence.

**2.2. Basic Theory of Multiple Network Models.** Deep convolutional neural network (DCNN) was initially proposed by Hubel and Wiesel [31] based on the structural design of the visual nervous system. It has a strong ability of data mining and information fusion and can effectively realize the local connection, weight sharing, space pooling, and other functions [16, 18, 32–37]. It contains many convolutional layers, pooling layers, and fully connected layers with vital data mining and information integration capabilities. It can effectively realize the local connection, weight sharing, space pooling, and other functions. The classical network of convolutional neural networks (CNN) is shown in Figure 1. The convolutional layer is the core component of deep convolutional network, which is composed of multiple sets of two-dimensional filters. When the data enters the convolutional layer, it will conduct convolution operation with the weight of the two-dimensional filter. The result after procedure is the output of the convolutional layer. The input data is processed by the down-sampling algorithm. The commonly used down-sampling algorithms include maximum pooling, average pooling, and nonuniform pooling, among which the maximum pooling method is the most widely used. The full connection layer is the last layer where the data has been processed by the previous convolution layer and the pooling layer, which has been extracted and converted into high-level information features. The full connection layer classifies the high-level information features to obtain the final recognition results. Soft-max is the generalization of the logistic classifier, which mainly solves the problem of multiclassification.

Backpropagation neural network (BPNN) can realize the nonlinear mapping between input and output, with one input layer, one output layer, and multiple hidden layers, each layer having various neurons. The output layer transfer function uses the linear function. With the minimum mean square error as the training objective, the BP algorithm is used as the learning algorithm of the network. BP neural network can learn the network through the gradient descent method and adjust the weight of the system through the backpropagation error. Through the above practices, the overall error of the network can be minimized.

The structure of Radial Basis Function neural network (RBFNN) is simple, with self-learning adaptive ability, fast convergence speed, function approximation ability, and

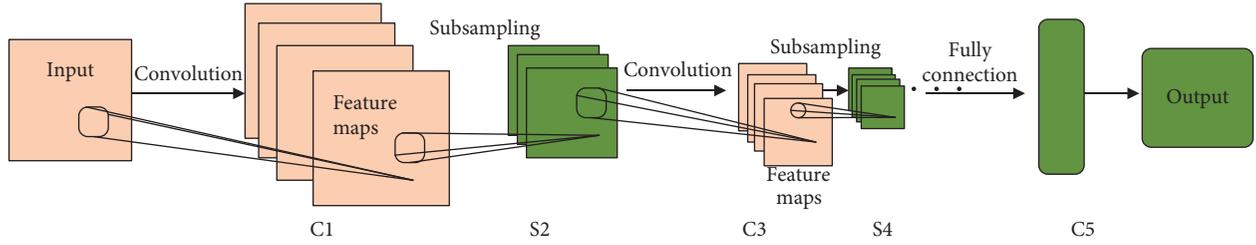


FIGURE 1: The classical network of convolutional neural networks (CNN).

superior advantages, thus having pervasive applications [38–40]. Its network structure is mainly composed of the input layer, hidden layer (using radial basis function as activation function), and output layer.

### 2.3. A Multimodel Decision Fusion Method Based on DCNN-IDST

**2.3.1. The Proposed Method.** In this paper, the fault information of rolling bearings is collected by multiple sensors, and the extracted characteristic values are input into various network models (BPNN, RBFNN, and DCNN) by using the adaptive feature extraction feature of deep convolutional neural network. The output results are normalized as the basic probability distribution of evidence theory. The original D-S synthesis formula is improved by consistent fuzzy matrix, and the final fusion result is obtained by decision level fusion.

The flowchart based on the DCNN-IDST method is shown in Figure 2. (1) Vibration signals of rolling bearings at three different speeds are collected by two vibration sensors. (2) Data preprocessing. (3) The vibration signals of three rates of the rolling bearing collected are fused at the data level and then input into the DCNN model. (4) The eigenvalues of DCNN adaptive output were input into three network models, respectively, for training. (5) The output results of the three network models are normalized and then input into IDST for decision fusion.

**2.3.2. Set Up Comparative Test Methods.** The flowcharts of comparative test methods are shown in Figure 3.

This research method verifies the feasibility of DCNN-IDST-based multimodel decision fusion method in rotating machinery equipment state monitoring by setting up multiple groups of comparative tests. The PCA method is used to compare the classification effect of artificial selection feature and DCNN adaptive feature extraction. The feature quantity extracted by artificial feature and the feature quantity removed by DCNN adaptively were input into BPNN and RBFNN, respectively, to verify the recognition effect of the model. The diagnosis effect of the new method based on IDST and D-S synthesis formula is proved.

## 3. Experiment and Discussion

**3.1. Experimental Setup.** Four states of rolling bearing (normal, inner ring fault, outer ring fault, and rolling body fault) are selected for test verification. The rolling bearing

fault diagnosis test is shown in Figure 4. Figure 5 shows the types of bearing fault. The sampling frequency is 20 kHz. The motor speed is set as 600 rpm, 900 rpm, and 1200 rpm. Each 2048 data point is used as a set of fault data samples. There are 300 samples for each health condition under each identical operating speed. The data of rolling bearings used in the test are shown in Table 1.

**3.2. Data Processing.** The original vibration signal of the rolling bearing at three rotational speeds (600 rpm, 900 rpm, and 1200 rpm) is preprocessed to improve the signal monitoring ratio. The preprocessed data are fused at the data level. Fifteen time-domain characteristics of vibration signals at three speeds of the rolling bearing are extracted manually. There are ten dimensionless time-domain indexes including mean value, root mean square value, root amplitude value, absolute mean value, variance, maximum value, and minimum value. There are five dimensionless time-domain indexes, including waveform index, peak value, pulse index, margin index, and kurtosis index. There are three frequency-domain eigenvalues, including mean square frequency, barycenter frequency, and frequency variance. These features will be input to BPNN and RBFNN for pattern recognition. Due to the limitation of space, only some time-domain indexes are listed in Table 2.

At the same time, the time-domain signal of the original rolling bearing is directly input into the DCNN model. Twenty feature values are adaptively extracted and input into BPNN and RBFNN for pattern recognition.

### 3.3. Model Design

**3.3.1. Backpropagation Neural Network.** The number of nodes in the input layer of Backpropagation Neural Network (BPNN) is set as fifteen (fifteen characteristic quantities). The number of nodes in the output layer is set as four (four fault types). The number of nodes in the hidden layer can be determined as fifteen after several tests. The learning rate is set at 0.05, and the expected error is 0.002. The desired output can be set as follows: normal [1,0,0,0], inner ring pitting [0,1,0,0], outer ring pitting [0,0,1,0], and rolling element pitting [0,0,0,1]. The output of the BP network is normalized and directly used as the basic probability distribution function of DST. Figure 6 shows the output of BP-adaptive network model.

**3.3.2. Radial Basis Function Neural Network.** For the RBF network model structure, the same training set and test set as

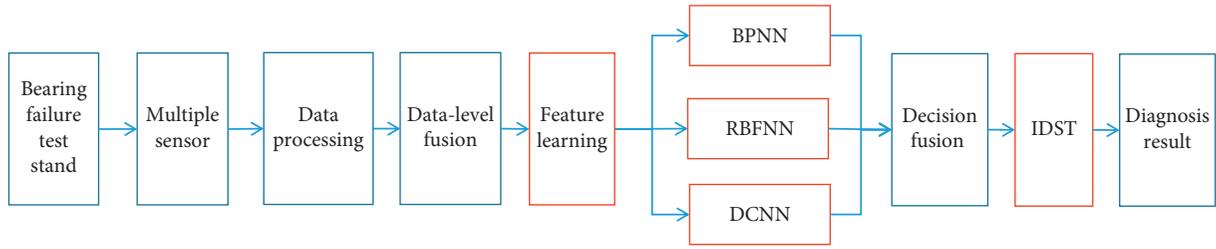


FIGURE 2: The flowchart based on the DCNN-IDST method.

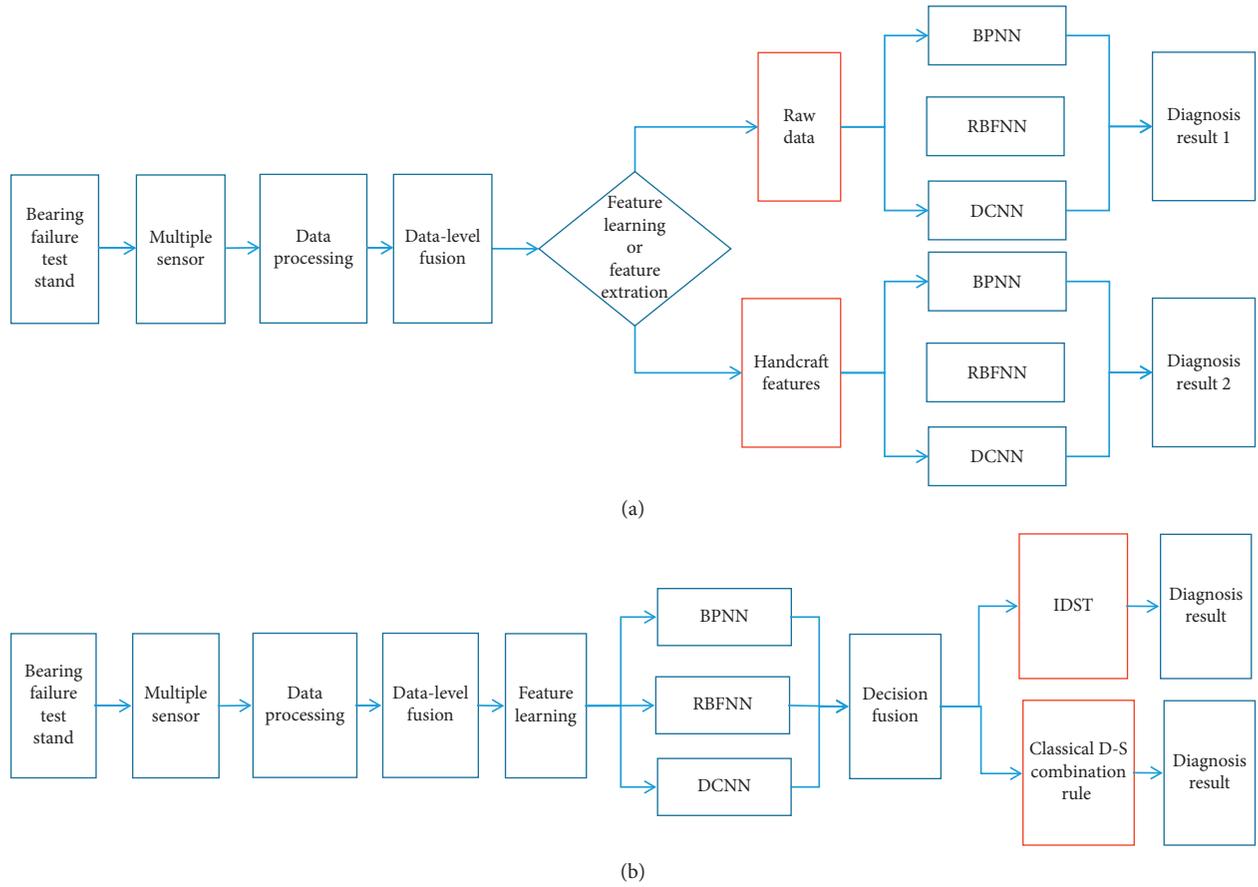


FIGURE 3: Flowcharts of comparative methods. (a) Advantages of artificial feature extraction and DCNN adaptive feature extraction. (b) The diagnosis effect of the new approach based on IDSET and D-S synthesis formula is verified.

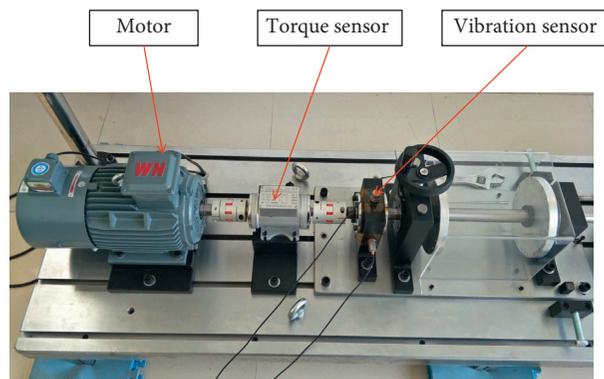


FIGURE 4: Rolling bearing fault diagnosis test.

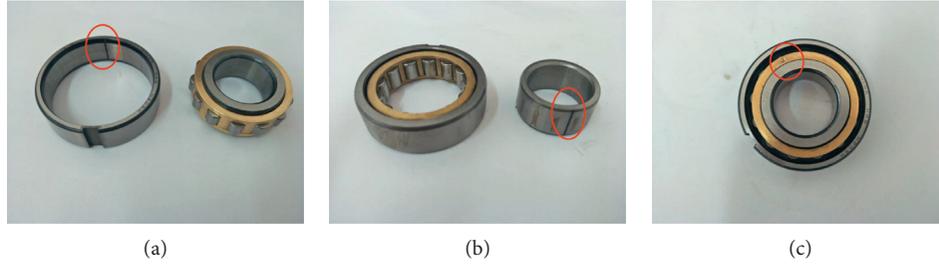


FIGURE 5: The types of bearing fault: (a) outer ring fault, (b) inner ring fault, and (c) rolling body fault.

BP neural network are used to train and test the system. The output result is normalized as the basic probability assignment function of evidence theory.

**3.3.3. Deep Convolutional Neural Network.** In this paper, the parameters of the selected model of DCNN are chosen by the method of ergodic reference. It is necessary to adjust the length of input data, the number of convolution kernel, size of convolution kernel, etc. Parameters of the selected model of DCNN are shown in Table 3. The main structure of DCNN is shown in Table 4.

## 4. Experimental Results

**4.1. Principal Component Analysis (PCA).** To verify the adaptive feature extraction capability of DCNN, principal component analysis (PCA) was used to analyze the manually selected features and the adaptive feature extraction capability of DCNN. The output graph of principal component analysis (PCA) method is shown in Figure 7.

As can be seen from Figure 7, the classification effect of features extracted by DCNN is better than the result of manual selection. It can be seen that DCNN has the advantage of adaptive feature extraction, which can lay a foundation for subsequent equipment fault diagnosis.

### 4.2. Experimental Results

**4.2.1. The Comparison Test Results of Several Network Models.** The fifteen time-domain, frequency-domain, and twenty DCNN adaptive features extracted manually were input into BPNN, RBFNN, and DCNN, respectively, for pattern recognition. 80% of these data are selected for training and 20% for testing. The comparison test results of several network models are shown in Table 5. Figure 8 shows the testing accuracy of several network models.

As can be seen from Table 5 and Figure 8, DCNN adaptive feature extraction was used to input the extracted feature into each network model, and the recognition accuracy was higher than that of the manually removed part. It can effectively improve the diagnostic accuracy of the model, which provides a new diagnostic idea for the state

TABLE 1: The data of rolling bearings used in the test.

Pattern label	Bearing state	Input speed (rpm)	Load
1	Normal	600, 900, 1200	0
2	Inner ring fault	600, 900, 1200	0
3	Outer ring fault	600, 900, 1200	0
4	Rolling body fault	600, 900, 1200	0

monitoring of rotating machinery and the decision fusion of multiple models.

**4.2.2. Decision Fusion Results.** In this paper, the generalization ability of the network model is used to construct the BPA of D-S evidence theory. The output results of BPNN, RBFNN, and DCNN are normalized and used as the BPA of D-S evidence theory. The output results of the three networks are shown in Tables 6–8, where  $m(A)$ ,  $m(B)$ ,  $m(C)$ , and  $m(D)$  represented the BPA of normal, inner ring fault, outer ring fault, and rolling body fault.  $m(\theta)$  is the uncertainty and  $\varepsilon$  is the reliability (FL = feature learning).

As shown in Table 6, the output of each network is uncertain. After the fusion of the three network models, the output results of the network are all inner ring faults. With the addition of credible evidence to the fusion process, the BPA of B is continuously increasing. Based on the proposed method in this paper, the final diagnosis accuracy rate reaches 94.82%, which is higher than that of any single network model.

To further compare the advantages of DCNN-IDST multimodel decision fusion method, two groups of comparative experiments were set up. (1) Three network models were fused under D-S synthesis formula. (2) Three network models were fused under the IDSET method. The  $\varepsilon$  between the two fusion rules is shown in Table 9. Figure 9 offers the fusing by D-S synthesis formula, and Figure 10 shows the fusing by the IDSET method.

By comparing Tables 7–9, it can be seen that the output results of the three networks were normalized and then fused again, which significantly improved the recognition accuracy. With the increasing of supporting evidence, the reliability of the IDST method for decision fusion has a steady rise and the pattern recognition rate is also greatly improved.

TABLE 2: Some of the dimension time-domain features.

Number	Mean value	Root mean square value	Root amplitude value	Crooked degree	Peak value	Variance	Maximum value	Feng-Feng value
1	0.008435	0.219334	0.136023	0.017431	0.017919	0.048107	1.440957	1.872766
2	-0.00601	0.198404	0.129573	0.009372	0.007169	0.039364	0.912906	1.281596
3	0.029603	0.245241	0.1616	0.018736	0.016446	0.060143	1.17327	1.594832
4	0.040591	0.25593	0.167282	0.021668	0.025214	0.0655	1.600798	2.093033
5	-0.0324	0.192693	0.126626	0.00949	0.007472	0.037131	0.977287	1.370114
6	-0.03391	0.167655	0.11473	0.004182	0.003009	0.028108	0.685338	1.04883
7	-0.00774	0.19325	0.126156	0.008213	0.006361	0.037346	0.827895	1.202846
8	-0.00372	0.219194	0.136029	0.020222	0.020906	0.048046	1.332877	1.701796

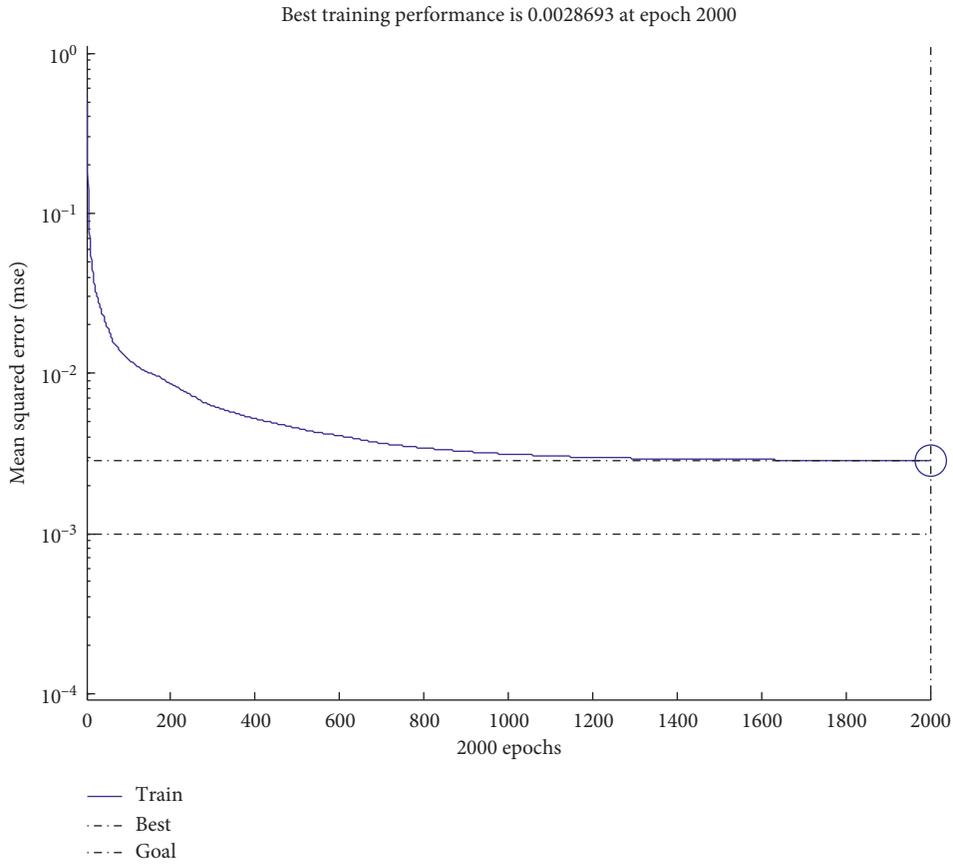


FIGURE 6: The output of the BP-adaptive network model.

TABLE 3: Parameters of the selected model of DCNN.

Number	Input data	Number of the convolution kernels	Size of the convolution kernel	Accuracy (%)
1	2048	16	10, 15	77.5
2	2048	32	10, 15	73.33
3	2048	64	10, 15	80.33
4	2048	128	10, 15	84.83
5	2048	256	10, 15	76.66

TABLE 4: Main structure of DCNN.

Layer	Type	Parameters
1	Convolution layer	Size 128, number10
2	Pooling layer	$S=2$
3	Convolution layer	Size 128, number15
4	Pooling layer	$S=2$
5	The connection layer	ReLU, Nodes10
6	Soft-max	4 outputs

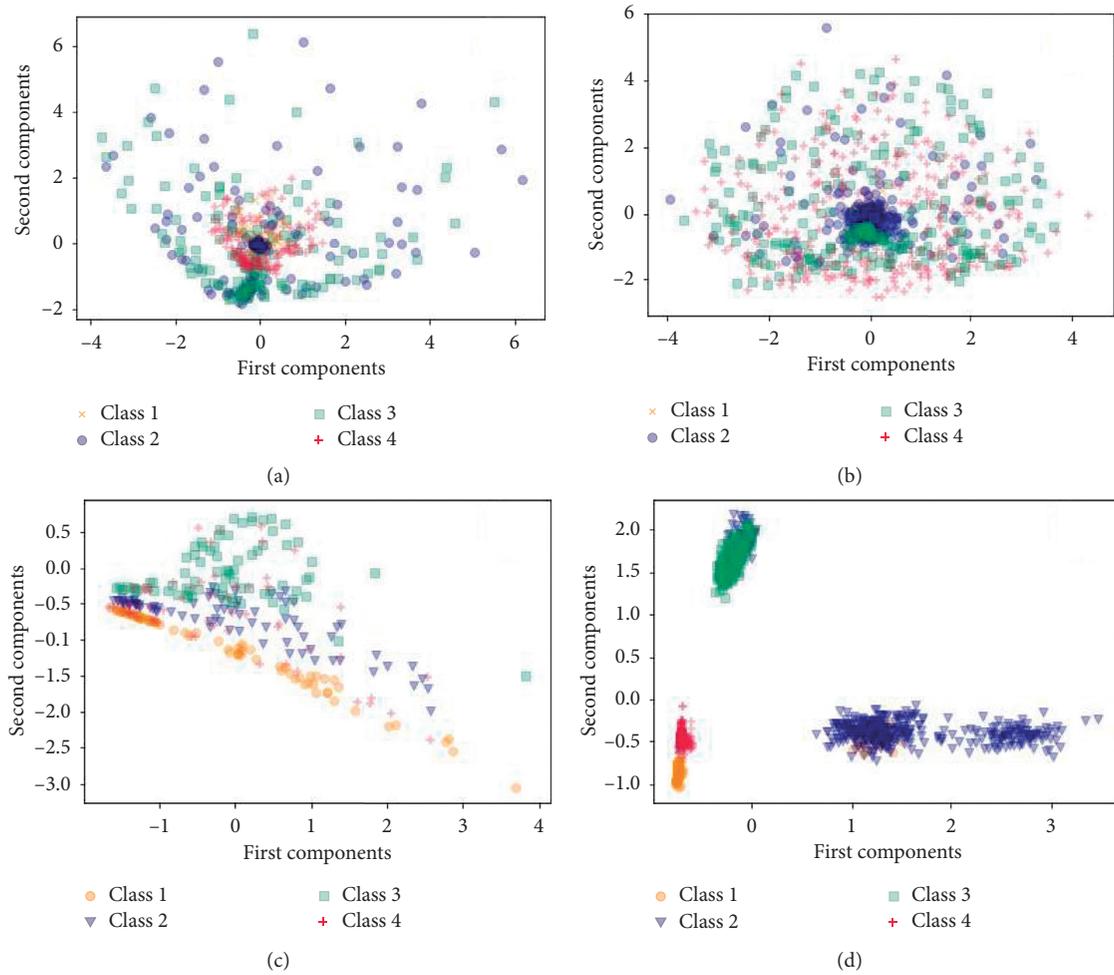


FIGURE 7: The output graph of principal component analysis (PCA) method. (a) Input data at a single rotation speed; (b) input data after data level fusion; (c) input data extracted by manual features; (d) adaptive feature learned through DCNN.

TABLE 5: The comparison test results of several network models.

Number	Model	Feature learning from raw data (%)	Time-domain features (%)	Manual feature extraction frequency-domain features (%)	Time-frequency-domain features (%)
Vibrating sensor1	BPNN	89.17	86.25	87.55	88.78
	RBFNN	87.08	65.83	70.35	73.33
	DCNN	83	72.50	78.33	80.62
Vibrating sensor2	BPNN	89.58	80.42	81.25	85.00
	RBFNN	91.42	77.92	85.66	88.66
	DCNN	90.92	80.51	81.55	83.33

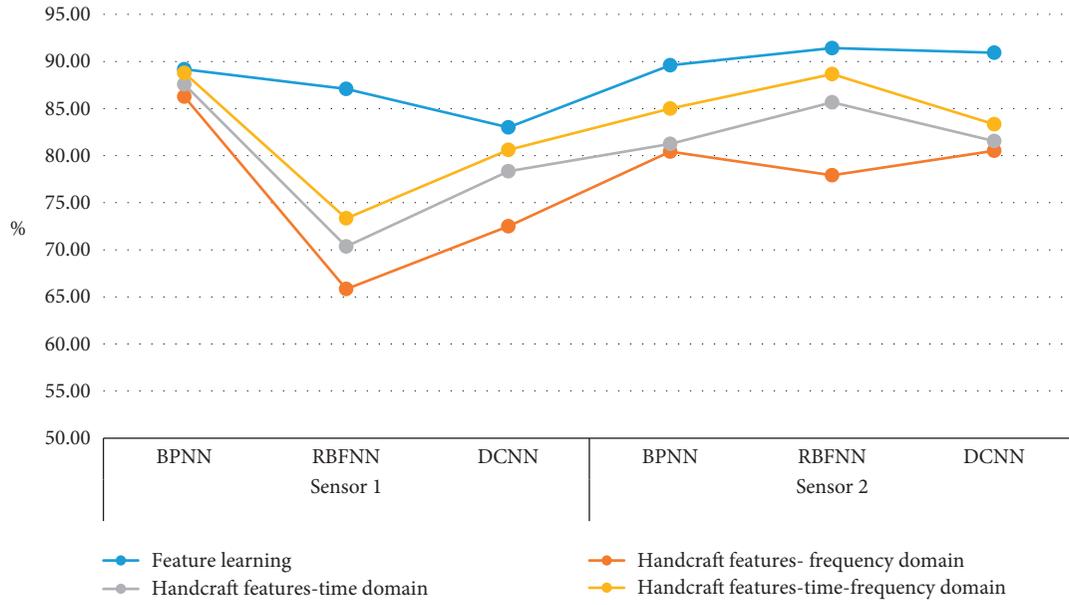


FIGURE 8: Testing the accuracy of several network models.

TABLE 6: The actual outputs of the three network models are compared with the expected results.

Models	Actual outputs			Expect results	Diagnosis results	
BPNN-FL1	0.000000	0.960499	0.000000	0.000045	0,1,0,0	Inner ring fault
	0.000000	0.986047	0.000002	0.000483	0,1,0,0	Inner ring fault
	0.000026	0.414734	0.408441	0.068637	0,1,0,0	Uncertainty
	0.000000	0.982463	0.004014	0.030370	0,1,0,0	Inner ring fault
	0.000000	0.925475	0.000000	0.000172	0,1,0,0	Inner ring fault
	0.000000	0.725717	0.000043	0.003560	0,1,0,0	Inner ring fault
RBFNN-FL1	-0.349173	0.283555	0.650220	0.415404	0,1,0,0	Uncertainty
	-0.048801	1.123021	0.128835	-0.203042	0,1,0,0	Inner ring fault
	0.221095	0.833929	0.189633	-0.244645	0,1,0,0	Inner ring fault
	-0.441047	0.937030	0.526606	-0.022593	0,1,0,0	Inner ring fault
	-0.041985	1.174574	-0.290296	0.157707	0,1,0,0	Inner ring fault
	0.312547	1.065893	-0.560675	0.182236	0,1,0,0	Inner ring fault
DCNN-FL1	0.000067	0.552400	0.446572	0.000961	0,1,0,0	Uncertainty
	0.000000	0.989853	0.000146	0.000001	0,1,0,0	Inner ring fault
	0.000000	0.940449	0.059551	0.000000	0,1,0,0	Inner ring fault
	0.000000	0.981931	0.018045	0.000025	0,1,0,0	Inner ring fault
	0.000000	0.976631	0.000069	0.000000	0,1,0,0	Inner ring fault
	0.000000	0.973404	0.026596	0.000001	0,1,0,0	Inner ring fault

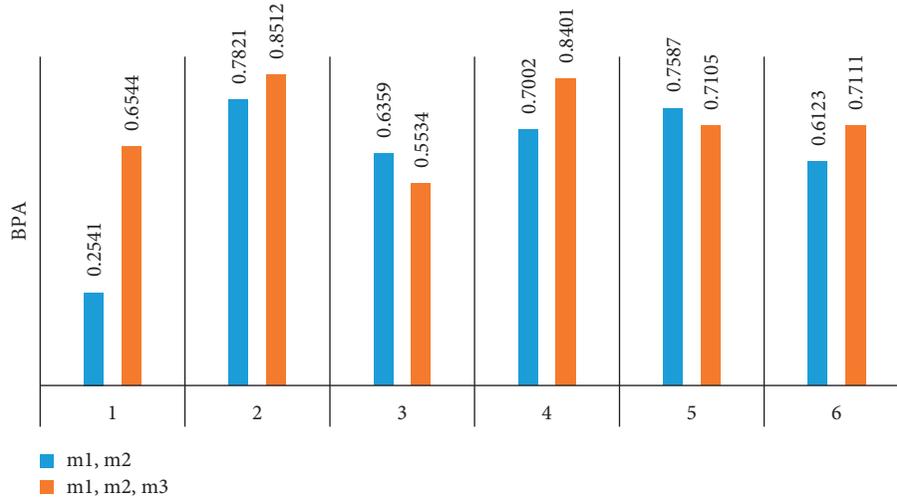


FIGURE 9: Fusing by D-S synthesis formula.

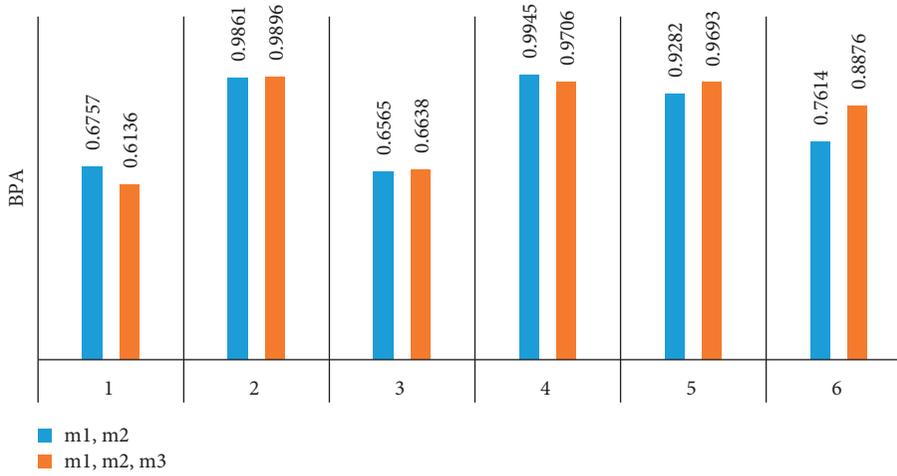


FIGURE 10: Fusing by the IDSET method.

TABLE 7: The result of the convergence of BPNN and RBGNN.

Number	$\epsilon$	$m(A)$	$m(B)$	$m(C)$	$m(D)$	$m(\theta)$	Results
1	0.675797	0.000000	0.819113	0.132410	0.101339	0.127043	Inner ring fault
2	0.986144	0.000800	0.999710	0.001722	0.000003	0.000193	Inner ring fault
3	0.656522	0.059662	0.610154	0.276491	0.009481	0.144536	Inner ring fault
4	0.994518	0.000000	0.987882	0.004749	0.010135	0.000030	Inner ring fault
5	0.928233	0.005859	0.992027	0.000000	0.010629	0.005345	Inner ring fault
6	0.761354	0.055722	0.904836	0.000004	0.049402	0.065069	Inner ring fault

TABLE 8: The result of the convergence of three networks.

Number	$\epsilon$	$m(A)$	$m(B)$	$m(C)$	$m(D)$	$m(\theta)$	Results
1	0.613672	0.000010	0.657967	0.145877	0.015425	0.188641	Inner ring fault
2	0.989621	0.000004	0.999838	0.000010	0.000000	0.000108	Inner ring fault
3	0.663839	0.008114	0.784689	0.062165	0.001289	0.137731	Inner ring fault
4	0.970560	0.000000	0.998597	0.000416	0.000148	0.000880	Inner ring fault
5	0.969325	0.000088	0.998571	0.000001	0.000161	0.000956	Inner ring fault
6	0.887605	0.002948	0.980156	0.001408	0.002614	0.013401	Inner ring fault

TABLE 9: The  $\varepsilon$  between the two fusion rules.

$k$ Or $\varepsilon$		1	2	3	4	5	6
Comparison test 1	m1, m2	0.2541	0.7821	0.6359	0.7002	0.7587	0.6123
	m1, m2, m3	0.6544	0.8512	0.5534	0.8401	0.7105	0.7111
Comparison test 2	m1, m2	0.6757	0.9861	0.6565	0.9945	0.9282	0.7614
	m1, m2, m3	0.6136	0.9896	0.6638	0.9706	0.9693	0.8876

The test results show that the DCNN-IDST multimodel decision fusion can achieve better diagnosis results and perfect diagnosis system.

## 5. Conclusions

This paper gives a multimodel decision fusion method based on DCNN-IDST for the damage detection of rolling bearing. Combining the advantages of DCNN and D-S evidence theory, the DCNN model presents a convincing adaptive feature extraction ability with a better effect than the manual feature extraction. The D-S evidence theory method is improved by fuzzy consistency matrix, and multimodel integration decision-making system of rotating mechanical malfunction diagnosis is established based on experimental research. The basic probability assignment after decision fusion is generally higher than the initial recognition result of a single network model. With the increase of supporting evidence, the BPA and reliability after fusion recognition increase further. The work demonstrates that the proposed method combines the advantages of every single network and overcomes the defects of a single network with uncertainty effectively, with the best recognition accuracy among all the tested methods.

## Data Availability

The basic data for the study were obtained from the laboratory.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The research was supported by the National Key Research and Development Program of China (Grant no. 2018YFB1701302), Key Research and Development Program of Shandong Province (Grant no. 2018GGX103016), and Shandong university science and technology plan project (Grant no. J15LB10).

## References

- [1] G. Wang, Z. J. He, X. F. Chen, and Y. N. Lai, "Basic research on machinery fault diagnosis-what is the prescription," *Journal of Mechanical Engineering*, vol. 49, no. 01, pp. 63–72, 2013.
- [2] F. Alam, R. Mehmood, I. Katib, N. N. Albogami, and A. Albeshri, "Data fusion and IoT for smart ubiquitous environments: a survey," *Institute of Electrical and Electronics Engineers Access*, vol. 5, pp. 9533–9554, 2017.
- [3] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: a review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013.
- [4] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," in *Classic Works of the Dempster-Shafer Theory of Belief Functions*, pp. 57–72, Springer, Berlin/Heidelberg, Germany, 2008.
- [5] G. Shafer, *A Mathematical Theory of Evidence*, vol. 42, Princeton Univ. Press, Princeton, NJ, USA, 1976.
- [6] G. Niu, T. Han, B.-S. Yang, and A. C. C. Tan, "Multi-agent decision fusion for motor fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 3, pp. 1285–1299, 2007.
- [7] H. K. Li, X. J. Ma, and Z. Wang, "Diesel engine fault diagnosis based on multi-symptom information fusion," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 35, no. 1, pp. 121–124, 2004.
- [8] Y. Zhang and J. Cui, "Crankshaft bearing fault diagnosis based on SVDD and D-S theory," *Journal of Chongqing University of Technology (Natural Science)*, vol. 11, 13 pages, 2019.
- [9] J. F. Cao and W. Cao, "Fault diagnosis of large manufacturing equipment based on improved evidence fusion theory," *Journal of Vibration, Measurement & Diagnosis*, vol. 32, no. 4, pp. 532–538, 2012.
- [10] J. H. Liu, M. P. Zhou, and C. C. Hou, "Mechanical characteristics fault diagnosis of circuit breaker based on quantum genetic neural network and D-S evidence theory," *High Voltage Apparatus*, vol. 54, no. 1, pp. 230–235, 2018.
- [11] W. Jiang, B. Wei, C. Xie, and D. Zhou, "An evidential sensor fusion method in fault diagnosis," *Advances in Mechanical Engineering*, vol. 8, 2016.
- [12] L. Yun, Y. Li, X. Yin, and D. Zheng, "Multisensor fault diagnosis modeling based on the evidence theory," *Institute of Electrical and Electronics Engineers Transactions on Reliability*, vol. 67, pp. 513–521, 2018.
- [13] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [14] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [15] L. Jing, M. Zhao, P. Li, and X. Xu, "A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox," *Measurement*, vol. 111, pp. 1–10, 2017.
- [16] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331–345, 2016.
- [17] P. Tamilselvan, Y. Wang, and P. Wang, "Deep belief network based state classification for structural health diagnosis," in *Proceedings of the Institute of Electrical and Electronics Engineers Aerospace Conference*, pp. 1–11, Big Sky, USA, March 2012.
- [18] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *Institute of Electrical and Electronics Engineers*

- Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016.
- [19] L. Jing, T. Wang, M. Zhao, and P. Wang, “An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox,” *Sensors*, vol. 17, no. 2, pp. 414–429, 2017.
- [20] R. Zhao, R. Yan, J. Wang, and K. Mao, “Learning to monitor machine health with convolutional bi-directional LSTM networks,” *Sensors*, vol. 17, no. 2, pp. 273–296, 2017.
- [21] G. S. Babu, P. Zhao, and X. L. Li, “Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life,” in *Proceedings of the International Conference on Database Systems for Advanced Applications*, pp. 214–228, Dallas, TX, USA, April 2016.
- [22] S. P. Li, *Research on Remaining Useful Life Prediction Method of a Rolling Bearing Combining CNN and LSTM*, Harbin University of Science and Technology, Harbin, Heilongjiang, China, 2019.
- [23] X. Shi, Z. Chen, and H. Wang, “Convolutional LSTM network: a machine learning approach for precipitation now-casting,” in *Proceedings of the 29th Annual Conference on Neural Information Processing Systems*, vol. 24, Montreal, QC, Canada, December 2015.
- [24] W. Luo, W. Liu, and S. Gao, “Remembering history with convolutional LSTM for anomaly detection,” in *Proceedings of the Institute of Electrical and Electronics Engineers International Conference on Multimedia and Expo*, vol. 25, Hong Kong, China, July 2017.
- [25] H. Qiao, T. Wang, P. Wang, S. Qiao, and L. Zhang, “A time-distributed spatiotemporal feature learning method for machine health monitoring with multi-sensor time series,” *Sensors*, vol. 18, no. 9, p. 2932, 2018.
- [26] X. Deng and W. Jiang, “Dependence assessment in human reliability analysis using an evidential network approach extended by belief rules and uncertainty measures,” *Annals of Nuclear Energy*, vol. 117, pp. 183–193, 2018.
- [27] L. Yin, X. Deng, and Y. Deng, “The negation of a basic probability assignment,” *Institute of Electrical and Electronics Engineers Transactions on Fuzzy Systems*, vol. 27, no. 1, pp. 135–143, 2019.
- [28] F. T. Wang, X. J. Ma, and H. Zhu, “Research on Fault Diagnosis method based on Dempster-Shafer evidential theory,” *Journal of Dalian University of Technology*, vol. 4, pp. 470–474, 2003.
- [29] Q. Sun, X. Q. Ye, and W. K. Gu, “A new combination rules of evidence theory,” *Journal of Electronic*, vol. 8, no. 8, pp. 117–119, 2000.
- [30] W. X. Xu, J. W. Tan, and H. Zhan, “Research and application of the improved DST new method based on fuzzy consistent matrix and the weighted average,” *Advanced Materials Research*, vol. 8, 2014.
- [31] D. H. Hubel and T. N. Wiesel, “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex,” *Journal of Physiology*, vol. 160, no. 1, 1962.
- [32] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [33] J. Schmidhuber, “Deep learning in neural networks: an overview,” *Neural Netw*, vol. 61, pp. 85–117, 2014.
- [34] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, “Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data,” *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 303–315, 2016.
- [35] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, “A sparse auto-encoder-based deep neural network approach for induction motor faults classification,” *Measurement*, vol. 89, pp. 171–178, 2016.
- [36] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, “Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks,” *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.
- [37] X. Guo, L. Chen, and C. Shen, “Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis,” *Measurement*, vol. 93, pp. 490–502, 2016.
- [38] W. X. Xu, J. W. Tan, and H. Zhan, “Improved D-S evidence theory in ball screw of CNC machine intelligent fault diagnosis applications,” *Manufacturing Technology and Machine Tools*, no. 9, pp. 163–166, 2019.
- [39] J. Li, L. W. Tian, and Y. Q. Wang, “Short-term load prediction of power system based on GA-RBF neural network,” *Journal of Shanghai University of Electric Power*, vol. 35, no. 3, pp. 205–210, 2019.
- [40] Y. Liu and H. X. Chen, “Fault diagnosis of gearbox based on RBF-PF and particle swarm optimization wavelet neural network,” *Netural Computing and Applications*, vol. 31, pp. 4463–4478, 2019.