

Received November 22, 2019, accepted January 14, 2020, date of publication January 21, 2020, date of current version January 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968519

A Bearing Fault Diagnosis Method Based on Feature Selection Feedback Network and Improved D-S Evidence Fusion

XIANGHONG TANG^{1,2,3}, XIN GU¹, JIACHEN WANG¹, QIANG HE¹,
FAN ZHANG¹, AND JIANGUANG LU^{1,2,3}

¹Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, Guiyang 550025, China

²State Key Laboratory of Public Big Data, Guizhou University, Guiyang 550025, China

³School of Mechanical Engineering, Guizhou University, Guiyang 550025, China

Corresponding author: Xin Gu (gxgx521521@126.com)

This work was supported in part by the Science and Technology Major Project of Guizhou Province under Grant [2013]6019, in part by the Project of Guizhou High-Level Study Abroad Talents Innovation and Entrepreneurship under Grant 2018.0002, in part by the Project of China Scholarship Council under Grant 201806675013, in part by the Open Fund of Guizhou Provincial Public Big Data Key Laboratory under Grant 2017BDFJ019, in part by the Guizhou University Foundation for the Introduction of Talent under Grant (2016) No.13, in part by the Science and Technology Platform, and in part by the Talent Team Plan of Guizhou Province under Grant [2017]5789-24.

ABSTRACT Bearings running state affects the normal operation of mechanical equipment. It is of great theoretical and practical value to carry out bearing fault diagnosis. In bearing fault diagnosis research, the extraction and selection of fault features can help improving the accuracy of bearing fault diagnosis. However, these researches suffer from the following weaknesses. (1) High dimension of the selected features. (2) Uncertainty of single sensor for data sampling. Therefore, in this paper, a feature selection feedback network (FSFN) is proposed to overcome the first weakness. At the same time, we proposed an improved Dempster-Shafer (IDS) evidence theory fusion method based on the kappa coefficient to deal with the second weakness. Extensive evaluations of the proposed method on the CUT-2 experimental platform dataset showed that FSFN can not only reduce the dimension of the final selected feature without decreasing the diagnostic accuracy but also shorten the time of feature selection. Moreover, compared with the existing DS evidence theory fusion method, IDS can achieve higher average fusion precision and improve the accuracy and reliability of bearing fault diagnosis.

INDEX TERMS Bearing fault diagnosis, feature selection, feedback network, D-S evidence theory.

I. INTRODUCTION

Due to improvement of digitalization and network level of manufacturing enterprises, a large amount of data of mechanical equipment has been accumulated [1], [2]. Bearings are usually one of the most common components of mechanical equipments, and their operating conditions have a considerable impact on mechanical equipments [3], [4]. Therefore, it is of great significance in theory and practice to carry out bearing fault diagnosis. The effect of bearing fault diagnosis is largely determined by fault features. The feature selection of bearing fault diagnosis can shorten the model training time and simplify fault diagnosis process by eliminating irrelevant features and reducing feature dimensions. In addition, information of multi-sensors is beneficial to the reliability

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Wang¹.

and accuracy of diagnosis results from the single sensor, and ensures the authenticity of bearing fault diagnosis results.

Feature selection can effectively improve the diagnostic accuracy and reduce the feature dimensions in bearing fault diagnosis. Hui *et al.* [5] proposed an improved WFS technique before integration with a support vector machine (SVM) model classifier as a complete fault diagnosis system for a rolling element bearing case study. Liu *et al.* [6] used an SVM as the fault decision maker, adopted wrapper-type feature subset evaluation criteria, and used the evolutionary Monte Carlo method to search for optimal feature subsets to realize bearing fault diagnosis. Yu *et al.* [7] proposed a novel feature extraction procedure incorporating an improved feature dimensionality reduction method, and the experimental results show that the proposed fault diagnosis model can serve as an effective and adaptive bearing fault diagnosis system. Cui *et al.* [8] proposed a rolling bearing

fault diagnosis method based on local and correlation analysis and carried out feature selection by calculating intrinsic mode function (IMF) correlations. Shi [9] proposed a new correlation coefficient of simplified neutrosophic sets (SNSs) to diagnose bearing fault types and extract the feature vector corresponding to the energy feature value of each frequency band for fault feature extraction. Zheng *et al.* [10] proposed a new rolling bearing fault diagnosis method based on multi-scale fuzzy entropy (MFE), Laplacian score (LS) and variable predictive model-based class discrimination (VPMCD). However, there are several similar problems in the feature selection work of the above researchers, that is, the dimension of the feature selection is high.

With growth of the types and number of sensors, data from huge amounts of sensors help people to be aware of the condition of equipments. The technique of combining information from multiple sources to form a unified understanding is called multi-sensor information fusion [11]. Currently, multi-sensor information fusion technology has been applied in many fields [11], [12], such as sensor networks, robotics, video and image processing, intelligent system design, target tracking, and automotive automation.

Multi-sensor information fusion is divided into data fusion, feature fusion and decision fusion [13], in which decision fusion is the highest level. First, a single decision is made based on the data of each sensor; then, multiple single decision results are combined to achieve the final fusion results. Decision fusion is a fusion based on specific goals, and the result of the fusion determines the accuracy of the decision. Compared to other fusion level methods, decision fusion discards a large amount of data but uses the fewest resources. Therefore, it has advantages in terms of strong anti-interference ability and low cost. Common decision fusion methods include Bayesian inference, expert systems, Dempster-Shafer evidence theory, and fuzzy theory. Coussement *et al.* [14] proposed a Bayesian decision support framework that formalizes subjective human expert opinions with more objective organizational information. Ren *et al.* [15] proposed a Bayesian inference system based on a Gaussian process for realizing intelligent surface measurements of multi-sensor instruments. Chen *et al.* [16] proposed a hybrid expert system to simulate the decision-making process of clinical sleep staging through symbol fusion. Rikalovic and Cosic [17] proposed an expert system for industrial location factor analysis based on a fuzzy inference system, which solved the problem of nonlinear optimization through available knowledge. Gong *et al.* [18] proposed a fault diagnosis method for the primary cooling system of nuclear power plants based on D-S evidence theory. Chen and Feng [19] proposed a new method for addressing conflicting evidence in the DS evidence theory and applied an improved method to pulsed gas tungsten arc welding, which combines the information obtained by arc, sound and vision sensors in the process of GTAW. Liu *et al.* [20] proposed an intelligent fault diagnosis method based on D-S evidence theory that comprehensively analyzes vibration and

temperature signals to diagnose bearing faults. Xie *et al.* [21] proposed a stochastic decision-making intuitionistic fuzzy method using applied foreground theory and gray relational analysis, which verified the effectiveness and feasibility of an example. Sadrykia *et al.* [22] proposed a new multi-criteria decision-making (MCDM) method based on a geospatial information system, which was applied to predict the extent of building damage before a potential earthquake occurs.

Although DS evidence theory has certain advantages in addressing multi-sensor information, with the gradual deepening of DS evidence research and the expansion of its application range, DS evidence theory has also revealed certain disadvantages. For example, measuring the weight of evidence by the distance function is not based on the precision of the data; thus, it is not sufficiently accurate for weight distribution. Additionally, the Delphi method [23] is used to correct the sensor weight dependence expert experience; thus, it is too subjective in terms of weight distribution.

Based on the above problems, a bearing fault feature selection feedback network and a bearing fault information fusion method based on the kappa coefficient are proposed. The main contributions of this paper are as follows:

(1) This paper proposes a feature selection network with a feedback effect. The network takes the feature selection time and the feature selection dimension as feedback conditions. Then, comparing the number of feature selections with the FF-FC-MIC (Feature-to-Feature and Feature-to-Category-Maximum Information Coefficient) feature selection method, it is verified that the network can select a smaller number of features in a shorter time and achieve better diagnosis results; with the diagnostic accuracy as the evaluation index, the network and the current feature selection method are used to verify the diagnostic accuracy of different feature selection algorithms on SVM and KNN. The experimental results show that the performance of the network is better than the current feature selection method.

(2) Aiming at the uncertainty of single sensor data diagnosis results in bearing fault diagnosis, an improved DS evidence fusion algorithm based on kappa coefficient is proposed. The algorithm improves the traditional DS evidence theory by data consistency check. By comparing with the improved DS evidence theory based on the distance function, the experimental results show that the performance of the proposed algorithm is better than the current improved DS evidence theory method.

The remainder of this paper is structured as follows. Section 2 introduces the basic theory and algorithm applied in this paper. Section 3 provides the details of the proposed method. Section 4 shows the experimental results. A conclusion is presented at the end of this paper.

II. RELATED BASIC THEORY AND ALGORITHM

A. FEATURE WEIGHT BASED ON A DECISION TREE

The decision tree (DT) method is a supervised machine learning method that is often used for classification and regression.

Because of its high calculation speed and high precision, it is often used in state-of-the-art research [24], [25]. The establishment of a decision tree model is to select a feature with the best distinguishing ability based on a certain criterion, divide the nodes of the tree, and repeat the feature selection to divide new nodes until the maximum depth of the tree is reached. This completes the establishment of the decision tree model. The construction process of the model is to divide and deepen the tree nodes by selecting features that exist in the feature selection process, which allows the model to contain most of the features that are strongly distinguishable [26]. According to the different feature selection criteria in the decision tree model, common decision tree algorithms are ID3, C4.5, and CART (classification and regression tree).

The ID3 algorithm uses the information gain to select features, and features with large information gains are preferred. The C4.5 algorithm uses the information gain rate selection feature to reduce the information gain and select features with more eigenvalues. Therefore, ID3 and C4.5 are decision tree algorithms based on the entropy model of information theory. To simplify the model and not completely lose the advantages of the entropy model, Breiman proposed the CART algorithm [27], [28].

The CART algorithm uses the Gini index to measure the purity of the sample data. The purity of a dataset D is defined as:

$$Gini(D) = \sum_{k=1}^{|y|} \sum_{k' \neq k} p_k p_{k'} = 1 - \sum_{k=1}^{|y|} p_k^2 \quad (1)$$

The Gini index of a is defined as:

$$Gini_{index}(D,a) = \sum_{v=1}^{|y|} \frac{|D^v|}{|D|} Gini(D^v) \quad (2)$$

The ability to distinguish features increases as the purity decreases.

B. DEMPSTER-SHAFFER EVIDENCE THEORY

D-S evidence theory consists of four important concepts: the recognition framework, basic probability assignment, uncertainty representation, and synthetic rules.

Recognition framework: In D-S evidence theory, the recognition framework is represented by a finite nonempty set:

$$\Theta = \{\theta_1, \theta_2, \dots, \theta_n\} \quad (3)$$

where θ_i ($i = 1, 2, \dots, n$) represents the i th hypothesis and reflects the i th possible recognition result and n represents the number of hypotheses.

Based on the recognition framework, the power set 2^Θ can be expressed as

$$2^\Theta = \{\emptyset, \{\theta_1\}, \{\theta_2\}, \dots, \{\theta_n\}, \{\theta_1, \theta_2\}, \dots, \{\theta_1, \theta_n\}, \dots, \{\theta_1, \theta_2, \dots, \theta_n\}\}$$

where $\theta_i \in \Theta, \theta \subseteq 2^\Theta$.

Basic probability assignment (BPA) [29]: The basic probability assignment is also known as the mass function,

which is used to describe the degree of support for hypotheses. The mass function can be expressed as $m : 2^\Theta \rightarrow [0, 1]$, which satisfies the following two equations:

$$\begin{aligned} m(\emptyset) &= 0 \\ \sum_{A \in \Theta} m(A) &= 1 \end{aligned} \quad (4)$$

where A is a proposition in 2^Θ , which contains one or more hypotheses; $m(A)$ indicates the basic support of evidence to proposition A .

Uncertain representation [20]: According to the mass function, the belief function (Bel) and the plausibility function (Pl) can be derived.

$$\begin{cases} Bel(\emptyset) = 0 \\ Bel(A) = \sum_{B \subseteq A} m(B), \quad A \in 2^\Theta, A \neq \emptyset \end{cases} \quad (5)$$

$$\begin{cases} Pl(\emptyset) = 0 \\ Pl(A) = \sum_{A \cap B \neq \emptyset} m(B), \quad A \in 2^\Theta, A \neq \emptyset \end{cases} \quad (6)$$

where $Bel(A)$ is the sum of the probabilities of all subsets of A , and $Pl(A)$ is the sum of the probabilities of B without intersection with A . $Bel(A)$ and $Pl(A)$ can be used as the lower and upper bounds of A , respectively. $[Bel(A), Pl(A)]$ is called the confidence interval of A , which is used to indicate the uncertainty of A [30], as shown in Figure 1.

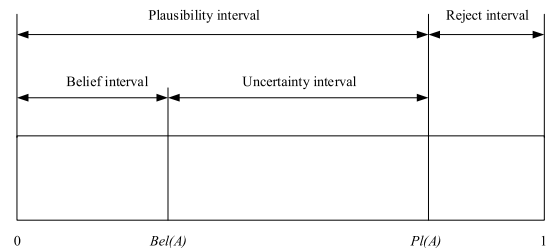


FIGURE 1. Certain representation of proposition A.

Synthetic Rule: Assume that there are n evidence bodies m_1, m_2, \dots, m_n under the system framework $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$; then the D-S evidence theory synthetic rule of m_1, m_2 is defined as:

$$m_1 \oplus m_2 \dots \oplus m_n = \begin{cases} \frac{1}{1-K} \sum_{A_1 \cap A_2 \dots \cap A_n = A} m_1(A_1) m_2(A_2) \dots m_n(A_n), & \text{if } A \neq \emptyset \\ \text{if } A = \emptyset \end{cases} \quad (7)$$

where k is the collision factor, which reflects the degree of collision between the evidence bodies.

$$\begin{aligned} K &= \sum_{A_1 \cap A_2 \dots \cap A_n = \emptyset} m_1(A_1) m_2(A_2) \dots m_n(A_n) \\ &= 1 - \sum_{A_1 \cap A_2 \dots \cap A_n \neq \emptyset} m_1(A_1) m_2(A_2) \dots m_n(A_n) \end{aligned} \quad (8)$$

C. KAPPA TEST

The kappa test is a method used to assess consistency in statistics and measures the consistency of two judgments (e.g., the consistency of the test results of two diagnostic test methods for the same sample or subject, where two medical workers performed two observations of the same group of patients to make a diagnosis) by the kappa coefficient. Its value is between -1 and 1 , and it is generally between 0 and 1 in practical applications.

The calculation of the kappa coefficient is based on the confusion matrix obtained by two judgments. Assume that there is a confusion matrix:

$$D_{ij} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (9)$$

The kappa coefficient is calculated as follows:

$$k = \frac{p_o - p_e}{1 - p_e} \quad (10)$$

where p_o is the sum of the observations in the diagonal unit of the confusion matrix:

$$p_o = \frac{\sum_{i=1}^n a_{ii}}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}} \quad (11)$$

p_e is the sum of the expected values in the diagonal unit as follows:

$$p_e = \frac{c_1 \times d_1 + c_2 \times d_2 + \dots + c_n \times d_n}{n \times n} \quad (12)$$

where c_i represents the number of real samples of the i th category:

$$c_i = \sum_{j=1}^n a_{ij} \quad (13)$$

d_i denotes the number of samples predicted by the i th category:

$$d_i = \sum_{i=1}^n a_{ij} \quad (14)$$

III. PROPOSED METHOD

In this section, we first improves the FF-FC-MIC feature selection algorithm in the literature [31] and proposes a new feature selection method, FSN (feature selection feedback network), for the feature dimension and feature selection time feedback function, which can perform online fault diagnosis feature selection. Then, a multi-sensor information fusion method is proposed based on the kappa coefficient to improve the accuracy and reliability of bearing fault diagnosis. The method improves the traditional D-S evidence theory from two aspects: evidence body weight and sensor weight. The kappa coefficient is used as the similarity measure between each evidence body to correct the weight of the evidence body, and the sensor weight is corrected based on the diagnostic accuracy of different sensor data.

A. FEATURE SELECTION FEEDBACK NETWORK(FSN)

Although the FF-FC-MIC feature selection algorithm achieves reduced complexity, it has a long running time and cannot satisfy the requirements of online diagnosis. It is found through experiments that the FF-FC-MIC feature selection algorithm presents a phenomenon in which the dimension of the feature selection is higher than that of the traditional feature selection algorithm. Further analysis shows that the correlation calculation time is long and accounts for the vast majority of the algorithm’s running time. Therefore, this paper proposes a feature selection network algorithm that can feedback the feature selection dimension and feature selection time to realize the intelligent features and real-time selection mode in bearing fault diagnosis.

Figure 2 shows the feature selection feedback network framework. First, we set the feature quantity threshold and feature selection time threshold. Then, we use the FF-FC-MIC algorithm proposed by [31] to select the feature for which the selected feature dimension and feature time are compared with the set two thresholds. When the feedback condition is satisfied, the weight of each feature in the feature set is calculated by CART, and features with a weight of zero are eliminated. Next, the FF-FC-MIC algorithm is used in the remaining features to complete the feature selection. Thus, the selected feature dimension is reduced, and the feature selection time is shortened. The entire process of FSN feature selection is shown in Figure 3.

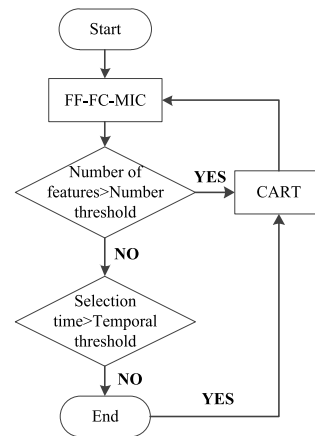


FIGURE 2. Feature network framework.

B. DS EVIDENCE THEORY FUSION ALGORITHM BASED ON THE KAPPA COEFFICIENT

The kappa coefficient of different sensor data is obtained by the consistency test, which is used as the weight of the evidence. Because the fault diagnosis models based on different sensors have different decision diagnosis results for different information sources of the same device state, kappa coefficient can be used to measure the similarity of the two judgments. Therefore, the weight of the different evidence is determined by testing the consistency of the data, and a similarity judgment can be made more scientifically and accurately.

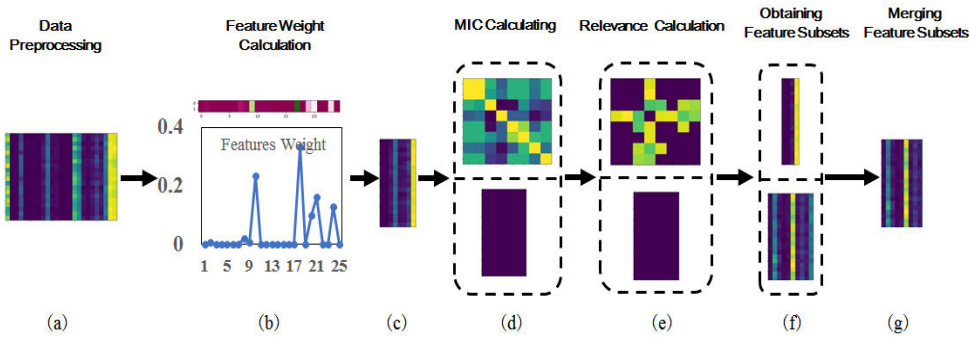


FIGURE 3. Flowchart of the proposed method. (a) Data preprocessing: Features are obtained through the time domain and frequency domain, forming a feature set (FS) matrix with a size of $n \times m$. (b) Feature weight calculation: The weight of each feature is calculated by CART, and features whose weight is not 0 are selected. (c) MIC (maximum information coefficient) calculation: The relevance between features is calculated to obtain the FF-MIC matrix with a size of $l \times l$ by MIC, and the relevance between features and fault categories is calculated to obtain the FC-MIC matrix with a size of $l \times p$ by MIC. (d) Relevance calculation: An $FF_threshold$ is set to distinguish strong irrelevance features from FS, and an $FF_threshold$ is set to distinguish strong relevance features from FS. (e) Obtaining feature subsets: According to the thresholds, Subset1 and Subset2 are obtained. (f) Merging feature subsets: An intersection operation is applied to merge Subset1 and Subset2, obtaining a final F-Subset.

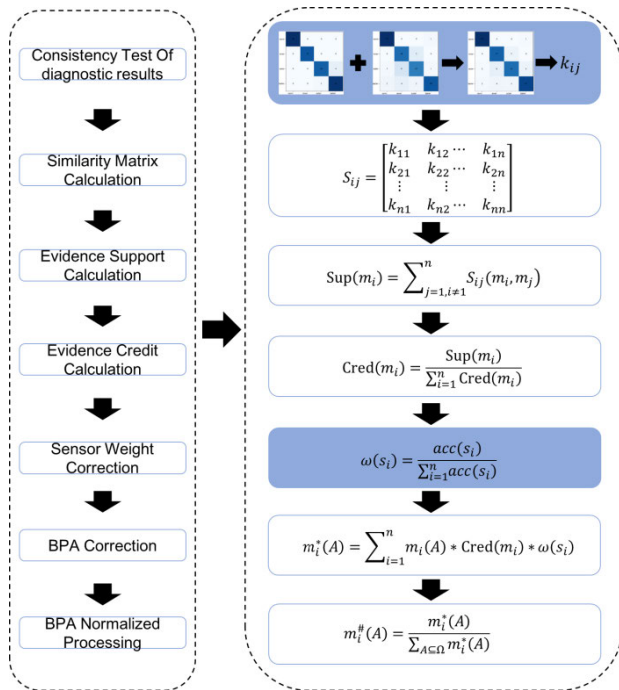


FIGURE 4. The process of IDS fusion.

The IDS fusion method flowchart is shown in Figure 4. First, the consistency of the two evidence bodies m_i and m_j is tested to obtain the kappa coefficient k_{ij} .

Taking k_{ij} as the measure of similarity between the evidence m_i and m_j , the similarity matrix of the evidence is obtained as follows:

$$S_{ij} = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & k_{22} & \dots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix} \quad (15)$$

Then, based on the similarity matrix, the $Sup(m_i)$ of the evidence m_i is calculated:

$$Sup(m_i) = \sum_{j=1, i \neq j}^n S_{ij}(m_i, m_j) \quad (16)$$

where $Sup(m_i)$ denotes the support between the evidence m_i and the other evidence bodies.

Moreover, based on the support, the $Cred(m_i)$ of the evidence m_i is calculated:

$$Cred(m_i) = \frac{Sup(m_i)}{\sum_{i=1}^n Cred(m_i)} \quad (17)$$

where $Cred(m_i) \in [0, 1]$ and $\sum_{i=1}^n Cred(m_i) = 1$.

The kappa coefficient of the two evidence bodies is obtained through a consistency test; the closer it is to 1, the higher the degree of consistency between the two evidence bodies. Specifically, the shorter the distance between the two evidence bodies is, the higher the support between each body; in addition, the higher the trust degree is, the greater the weight of the final evidence.

Next, this paper uses the weighted average of the diagnostic accuracy corresponding to each sensor data item to correct the sensor weight. Compared with the commonly used Delphi method, correcting the sensor weight by the weighted average does not depend on expert experience, and the expression result is more objective and accurate. The sensor weight $\omega(s_i)$ is calculated as follows:

$$\omega(s_i) = \frac{acc(s_i)}{\sum_{i=1}^n acc(s_i)} \quad (18)$$

where $acc(s_i)$ represents the diagnostic accuracy and $\sum_{i=1}^n acc(s_i)$ represents the sum of the diagnostic accuracy of the sensor n .

Then, we use the kappa coefficient to modify the weights of evidence and the weighted average of the diagnostic accuracy

to modify the common corrected BPA:

$$m_i^*(A) = \sum_{i=1}^n m_i(A) * Cred(m_i) * \omega(s_i) \quad (19)$$

Finally, the corrected BPA is normalized:

$$m_i^\#(A) = \frac{m_i^*(A)}{\sum_{A \subseteq \Omega} m_i^*(A)} \quad (20)$$

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL SETUP AND DATASET

To validate the effectiveness of the proposed feature selection method, vibration signals were collected, and experiments were conducted on the CUT-2 platform. As shown in Figure 5, the test rig was composed of an oscilloscope, a data acquisition system, a speed governor, the CUT-2 platform and a computer. The test bearing was placed at the right end of the test bench. The sampling frequency was 2 kHz, the motor speed is 2,000, 2,500 and 3,000 rpm, and the bearing model is a deep groove ball bearing 6,900 zz. The data acquisition of the bearing vibration signal is completed without load. To ensure the integrity of the bearing vibration signal, the sensors are installed in the X, Y and Z directions, as shown in Figure 6. In addition, an electric discharge machine was utilized to set three different types of faults: outer race faults,

ball faults, and inner race faults, as shown in Figure 6. The fault diameters were 0.2 and 0.3 mm.

The vibration signal construction verification dataset was collected by the CUT-2 bearing experimental platform. The CUT-2 dataset contains four bearing states: normal, outer ring fault, inner ring fault, and ball fault. Each class contains 200 samples, each consisting of 1,024 consecutive points. The dataset of the same structure was constructed at 2,000, 2,500, 3,000 rpm, and the training and test processes were performed on the dataset at the same speed. Table 1 shows the details of the datasets. A training set with 560 samples is generated, along with a test set of 240 samples.

TABLE 1. CUT-2 bearing data set at 2,000, 2,500, 3,000 rpm.

Conditions of the Bearing	Fault Size (mm)	Number of Samples	Class
Normal condition		200	0
Inner race fault	0.2	100	1
	0.3	100	
Outer race fault	0.2	100	2
	0.3	100	
Baller fault	0.2	100	3
	0.3	100	

B. ANALYSIS OF EXPERIMENTAL RESULTS OF THE FSFN ALGORITHM ON THE CUT-2 DATASET

Table 2 shows the feature weights calculated by the CARTA algorithm at 2,000, 2,500, and 3,000 rpm.

At 2,000 rpm, the FSFN feature selection algorithm calculates the feature weights through the CART algorithm. A total of 13 features were selected under the X-direction sensor dataset, and the feature numbers were 0, 1, 3, 8, 13, 14, 15, 17, 20, 21, 22, 23, and 24. A total of 14 features were selected under the Y-direction sensor dataset, and the feature numbers were 0, 1, 3, 4, 6, 7, 10, 14, 15, 17, 18, 20, 23, and 24. A total of 6 features were selected under the Z-direction sensor dataset, and the feature numbers were 1, 5, 10, 13, 16, and 21.

At 2,500 rpm, 9 features were selected under the X-direction sensor dataset, and the feature numbers were 5, 6, 10, 13, 15, 17, 20, 21, and 22. A total of 8 features were selected under the Y-direction sensor dataset, and the feature numbers were 0, 1, 16, 17, 19, 20, 22, and 24. A total of 7 features were selected under the Z-direction sensor dataset, and the feature numbers were 1, 3, 14, 15, 20, 21, and 24.

At 3,000 rpm, 14 features were selected under the X-direction sensor dataset, and the feature numbers were 0, 1, 2, 3, 10, 11, 13, 15, 16, 18, 19, 21, 22, and 24. A total of 8 features were selected under the Y-direction sensor dataset, and the feature numbers were 2, 5, 6, 7, 15, 18, 20, and 23. A total of 6 features were selected under the Z-direction sensor data set, and the feature numbers were 1, 3, 7, 13, 20, and 21.

The effectiveness of the FSFN feature selection algorithm was verified on the CUT-2 dataset. As shown in Table 3

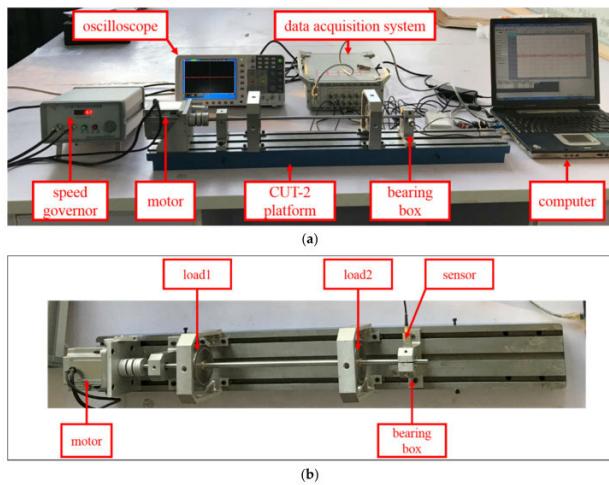


FIGURE 5. Experimental test rig: (a) CUT-2 data acquisition system; (b) CUT-2 platform.

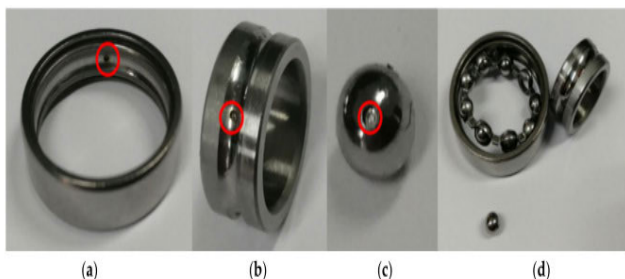


FIGURE 6. Locations of bearing faults: (a) outer race fault; (b) ball fault; (c) inner race fault; and (d) combination of parts.

TABLE 2. The feature weight of the CUT-2 dataset.

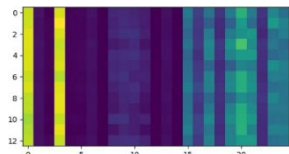
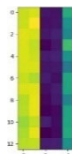
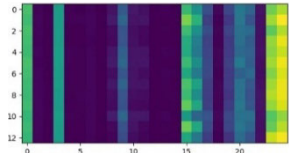
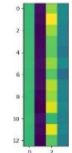
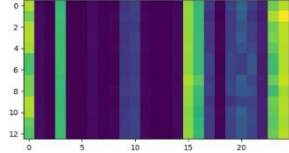
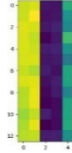
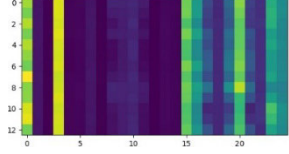
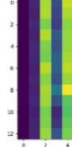
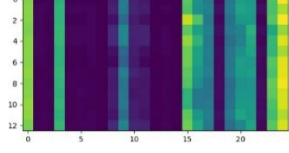
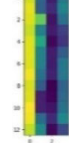
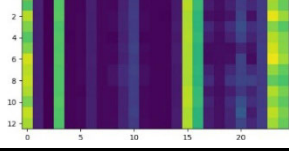
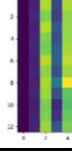
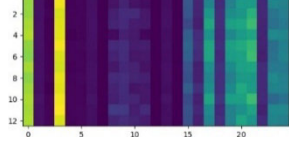
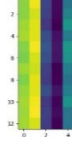
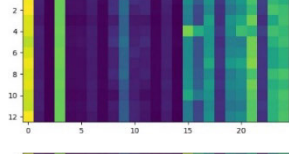
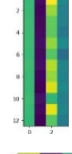
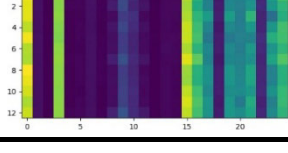
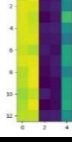
Feature	2000 rpm			2500 rpm			3000 rpm		
	X	Y	Z	X	Y	Z	X	Y	Z
f0	0.2503	0.01659	0	0	0.11152	0	0.00165	0	0
f1	0.01211	0.05767	0.0444	0	0.02574	0.05117	0.02593	0	0.13194
f2	0	0	0	0	0	0	0.00764	0.02175	0
f3	0.00251	0.00206	0	0	0	0.0017	0.18315	0	0.10014
f4	0	0.0067	0	0	0	0	0	0	0
f5	0	0	0.11344	0.17982	0	0	0	0.00999	0
f6	0	0.00256	0	0.00165	0	0	0	0.02767	0
f7	0	0.21675	0	0	0	0	0	0.01322	0.09257
f8	0.00125	0	0	0	0	0	0	0	0
f9	0	0	0	0	0	0	0	0	0
f10	0	0.00228	0.00988	0.20033	0	0	0.01208	0	0
f11	0	0	0	0	0	0	0.03014	0	0
f12	0	0	0	0	0	0	0	0	0
f13	0.16666	0	0.33198	0.00654	0	0	0.0033	0	0.23444
f14	0.36348	0.00722	0	0	0	0.33716	0	0	0
f15	0.12559	0.27977	0	0.24485	0	0.01553	0.32894	0.00577	0
f16	0	0	0.04261	0	0.00985	0	0.05845	0	0
f17	0.00333	0.02488	0	0.02571	0.00984	0	0	0	0
f18	0	0.00889	0	0	0	0	0.00167	0.33801	0
f19	0	0	0	0	0.01602	0	0.00495	0	0
f20	0.02348	0.3141	0	0.01152	0.31616	0.14489	0	0.29829	0.10145
f21	0.00942	0	0.4577	0.10699	0	0.11239	0.01995	0	0.33362
f22	0.03248	0	0	0.22259	0.21231	0	0.26089	0	0
f23	0.00287	0.04871	0	0	0	0	0	0.28532	0
f24	0.00652	0.01182	0	0	0.29857	0.33716	0.06126	0	0

(The figures are the visualization results for the corresponding speed samples. The horizontal axis in the figure represents the number of features and the vertical axis represents the first 13 samples. For the sake of beauty, the first 13 samples are chosen), the FSFN feature selection algorithm selects the number of features in the CUT-2 datasets. From the three-sensor datasets for the three rotational speeds, the FSFN feature selection algorithm performs the feature selection with fewer than six of the 25 features.

Additionally, the number of features of the two types of algorithms is compared, as shown in Figure 7. Compared with the FF-FC-MIC feature selection algorithm, the algorithm proposed in this paper selects fewer features, thus the feature selection time is also shortened.

For further analysis, the diagnostic accuracies of the different feature selection algorithms are verified on the SVM and KNN methods, as shown in Table 4. The diagnostic accuracy of the classification model and the average

TABLE 3. CUT-2 dataset feature selected by the proposed.

Motor Speed (rpm)	Sensor Position	All Features	Selected Features	Selected Number
2000	X			0, 3, 8, 10, 20
	Y			3,10,15,20
	Z			1,5,10,13,15
2500	X			5, 10, 15, 17, 20
	Y			0,16,19,20
	Z			1,14,20,21,24
3000	X			0, 3, 15, 16, 19
	Y			2,6,16,20
	Z			1, 7, 13, 20, 21

diagnostic accuracy of the FSFN feature selection algorithm are all above 95%, which are higher than the diagnostic accuracy of the FF-FC-MIC and SFS mRMR feature selection

algorithms. However, at 2,500 rpm, compared to the mRMR feature selection algorithm, the average diagnostic accuracy of the FSFN feature selection algorithm decreases by 1%,

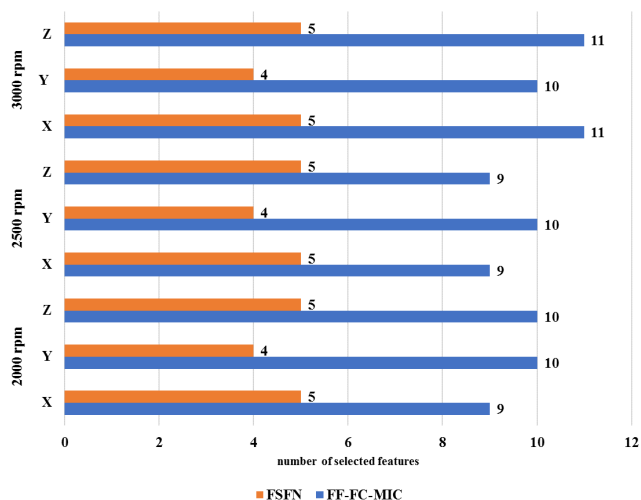


FIGURE 7. Comparison of feature numbers selected by two types of algorithms.

TABLE 4. Comparison of classification accuracy of SVM and KNN at motor speeds of 2,000, 2,500 and 3,000 rpm using different feature selection methods.

Motor Speed(rpm)	Feature Selection Methods	Classification Models		Average Accuracy
		SVM	KNN	
2000	FF-FC-MIC	0.9225	0.9125	0.9175
	SFS	0.9425	0.9363	0.9394
	mRMR	0.9833	0.9708	0.9775
	FSFN	0.9917	0.9833	0.9875
2500	FF-FC-MIC	0.9625	0.9313	0.9469
	SFS	0.9688	0.9525	0.9607
	mRMR	0.9833	0.9799	0.9892
	FSFN	0.9833	0.9625	0.9729
3000	FF-FC-MIC	0.9938	0.9875	0.9907
	SFS	0.9863	0.9763	0.9813
	mRMR	0.9958	0.9833	0.9896
	FSFN	0.9958	0.9958	0.9958

which is possible because of the instability of the vibration signal at 2,500 rpm. In other words, the FSFN algorithm shows high precision.

C. EVALUATION OF THE IMPROVED D-S(IDS) EVIDENCE FUSION ALGORITHM AND ITS PERFORMANCE IN BEARING FAULT DIAGNOSIS

The proposed improved DS evidence theory algorithm (IDS) is verified on the CUT-2 dataset in this paper. For the CUT-2 dataset, the fusion verification is performed for the X, Y directions, the Y,Z directions, the X,Z directions and the X, Y, Z directions.

First, IDS fusion is performed based on the diagnostic accuracy corresponding to the feature selected by the feature network proposed in this paper. To eliminate the influence of random factors, ten experiments were repeated, and the average value was taken as the average fusion precision. The average fusion precision of IDS is compared with the average accuracy of single sensor diagnosis and traditional evidence theory. The results are shown in Figure 8 (a)-(d).

Based on the analysis of Figure 8 (a)-(d), the average fusion accuracy is approximately 1% higher than the diagnostic accuracy of a single sensor, regardless of the combination of two-sensors or three-sensor data diagnostic accuracy. Specifically, the traditional DS evidence theory and IDS algorithm perform the decision fusion. The reason for the small improvement in fusion accuracy is that before the decision fusion, the corresponding datasets of each sensor selected the feature with the highest diagnostic accuracy by the feature network under the current conditions. The diagnostic accuracy of single-sensor data is close to 100% at different speeds. Therefore, when the decision fusion is performed again, the average fusion precision is close to the maximum value, and the effect of the decision fusion cannot be shown.

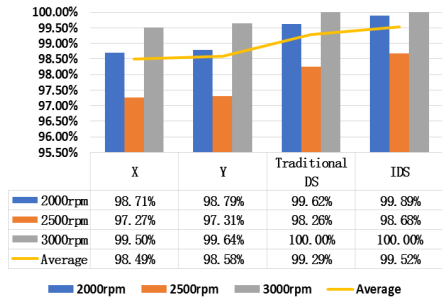
To highlight the advantages in this study with respect to the improvement of D-S evidence theory and the accuracy and effectiveness of the experiment, the control variables are used to select the same features for all sensor-constructed datasets for training and verification. Therefore, based on features in the time domain and frequency domain, this study selected 0-1-2-3-4, i.e., a total of five features for verification.

In addition, we compare the proposed IDS algorithm with other algorithms (traditional DS evidence theory, DS evidence theory based on the Euclidean distance correction of the weight of evidence (Euclidean distance DS), D-S evidence theory based on the Minkowski distance correction of the weight of evidence (Minkowski distance DS), and DS evidence theory based on the Mahalanobis distance correction of the weight of evidence (Mahalanobis distance DS)) in fusion precision. Ten verifications were performed for each sensor dataset and fusion method, and the average was taken as the final result.

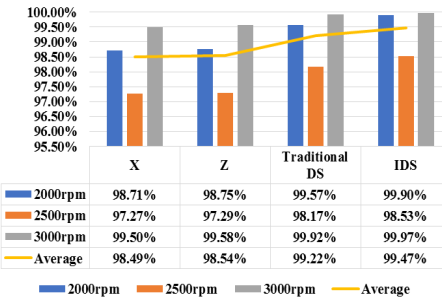
1) CONFUSION MATRIX VISUALIZATION

As shown in Table 5, the confusion matrix at each speed on the CUT-2 dataset is visualized, and the kappa coefficient is obtained by consistency testing of the two-sensor data at the same speed.

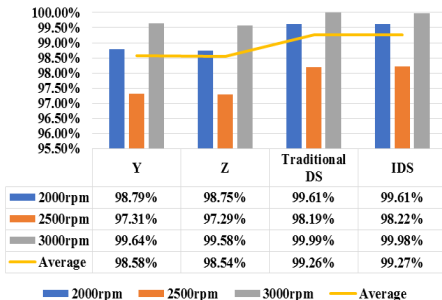
In this study, the weight of evidence is corrected by performing a consistency test on the diagnostic results of two sensors at the same speed. First, the confusion matrix corresponding to the diagnosis result of each sensor dataset is calculated based on different sensor data diagnoses. Second, the elements on the corresponding positions of different confusion matrices are added to a summation confusion matrix, which reflects the predicted and true values under the two sensors. Different prediction conditions represent the diagnostic accuracy of each sensor dataset. Through the consistency



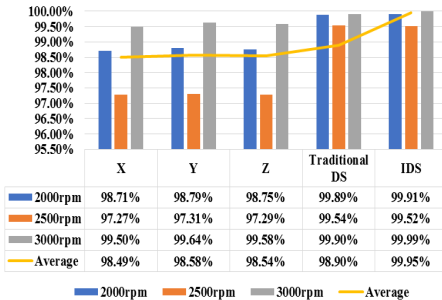
(a)



(b)



(c)



(d)

FIGURE 8. (a) Comparison of diagnostic accuracy and fusion method results of the X- and Y-direction sensors after the network characterization. (b) Comparison of diagnostic accuracy and fusion method results of X- and Z-direction sensors after the network characterization. (c) Comparison of diagnostic accuracy and fusion method results of Y- and Z-direction sensors after the network characterization. (d) Comparison of diagnostic accuracy and fusion method results of X-, Y- and Z-direction sensors after the network characterization.

testing of the confusion matrix obtained from different sensor data, the degree of difference in the diagnostic consistency between the two sensors can be obtained. The degree of this

difference is expressed by the kappa coefficient, which is closer to 1, indicating that the difference between the two diagnoses is smaller and that the two evidence bodies are closer to each other. Thus, the calculation based on the kappa coefficient can reflect the distance between different evidence bodies and thus correct the weight of the evidence.

The kappa coefficient represents the similarity of the two-sensor data diagnostic results. Therefore, regardless of whether the sensor is in the X, Y, or Z directions, the similarity between the diagnostic results is relatively high.

2) ANALYSIS OF EXPERIMENTAL RESULTS ON CUT-2 DATASET

This study validated the applicability and stability of IDS on the CUT-2 dataset. Because the CUT-2 experimental platforms contain sensors in three positions, the IDS is verified by a combination of four sensors, i.e., sensor combinations in the X and Y directions, X and Z directions, Y and Z directions, and X, Y, and Z directions. The SVM model is used to diagnose data in the X, Y, and Z directions, and the average accuracy is used as the corrected sensor weight, as shown in Table 6.

Figure 9 shows that IDS improves the diagnostic accuracy by an average of 3.55% compared to the single X and Y sensors. Compared with the other DS evidence theory, the fusion accuracy increases by at least an average of approximately 0.75% and by at most an average of approximately 1.36%. Overall, although IDS achieves a lesser improvement in terms of fusion accuracy compared to the other improved DS evidence theory, it demonstrates a certain effectiveness and adaptability in the decision fusion of sensor data in the X and Y directions.

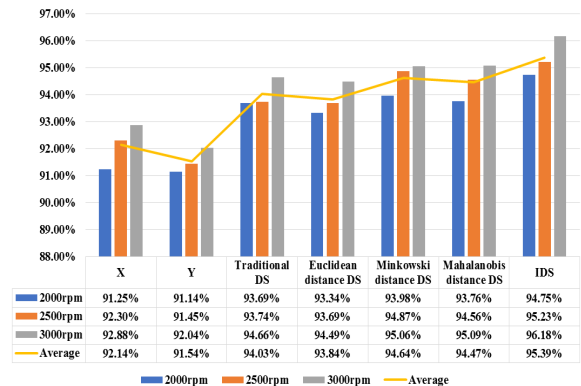


FIGURE 9. Comparison of diagnostic accuracy and fusion method results for X, Y direction sensors.

Figure 10 shows that IDS improves the diagnostic accuracy by an average of 2.65% compared to the single sensors in the X and Z directions. Compared with the other improved DS evidence theory, the fusion accuracy increases by at least an average of approximately 0.53% and by at most an average of approximately 1.18%. At 2,500 rpm, the fusion performances of IDS and Mahalanobis DS are almost identical. Therefore, the fusion precision of IDS is less than that of the other

TABLE 5. Confusion matrix visualization of CUT-2 dataset diagnosis results.

Motor Speed (rpm)	Confusion Matrix Of Sensor Data Diagnosis Results In Different Directions				Summation Confusion Matrix	Kappa Coefficient																																																																										
	X	Y	Z																																																																													
2000	<table border="1"> <tr><td>norm</td><td>62</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>49</td><td>2</td><td>0</td></tr> <tr><td>outer</td><td>1</td><td>2</td><td>50</td><td>7</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>3</td><td>63</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	62	0	1	0	inner	0	49	2	0	outer	1	2	50	7	baller	0	0	3	63		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>51</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>0</td><td>0</td><td>59</td><td>1</td></tr> <tr><td>baller</td><td>1</td><td>0</td><td>0</td><td>65</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	0	51	0	0	outer	0	0	59	1	baller	1	0	0	65		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>125</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>100</td><td>2</td><td>0</td></tr> <tr><td>outer</td><td>1</td><td>2</td><td>109</td><td>8</td></tr> <tr><td>baller</td><td>1</td><td>0</td><td>3</td><td>128</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	125	0	1	0	inner	0	100	2	0	outer	1	2	109	8	baller	1	0	3	128		norm	inner	outer	baller		0.9498
	norm	62	0	1	0																																																																											
	inner	0	49	2	0																																																																											
	outer	1	2	50	7																																																																											
	baller	0	0	3	63																																																																											
		norm	inner	outer	baller																																																																											
norm	63	0	0	0																																																																												
inner	0	51	0	0																																																																												
outer	0	0	59	1																																																																												
baller	1	0	0	65																																																																												
	norm	inner	outer	baller																																																																												
norm	125	0	1	0																																																																												
inner	0	100	2	0																																																																												
outer	1	2	109	8																																																																												
baller	1	0	3	128																																																																												
	norm	inner	outer	baller																																																																												
	<table border="1"> <tr><td>norm</td><td>62</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>49</td><td>2</td><td>0</td></tr> <tr><td>outer</td><td>1</td><td>2</td><td>50</td><td>7</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>3</td><td>63</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	62	0	1	0	inner	0	49	2	0	outer	1	2	50	7	baller	0	0	3	63		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>1</td><td>50</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>2</td><td>0</td><td>58</td><td>0</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>0</td><td>66</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	1	50	0	0	outer	2	0	58	0	baller	0	0	0	66		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>125</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>1</td><td>99</td><td>2</td><td>0</td></tr> <tr><td>outer</td><td>3</td><td>2</td><td>108</td><td>7</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>3</td><td>129</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	125	0	1	0	inner	1	99	2	0	outer	3	2	108	7	baller	0	0	3	129		norm	inner	outer	baller		
norm	62	0	1	0																																																																												
inner	0	49	2	0																																																																												
outer	1	2	50	7																																																																												
baller	0	0	3	63																																																																												
	norm	inner	outer	baller																																																																												
norm	63	0	0	0																																																																												
inner	1	50	0	0																																																																												
outer	2	0	58	0																																																																												
baller	0	0	0	66																																																																												
	norm	inner	outer	baller																																																																												
norm	125	0	1	0																																																																												
inner	1	99	2	0																																																																												
outer	3	2	108	7																																																																												
baller	0	0	3	129																																																																												
	norm	inner	outer	baller																																																																												
	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>51</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>0</td><td>0</td><td>59</td><td>1</td></tr> <tr><td>baller</td><td>1</td><td>0</td><td>0</td><td>65</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	0	51	0	0	outer	0	0	59	1	baller	1	0	0	65		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>1</td><td>50</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>2</td><td>0</td><td>58</td><td>0</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>0</td><td>66</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	1	50	0	0	outer	2	0	58	0	baller	0	0	0	66		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>126</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>1</td><td>101</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>2</td><td>0</td><td>117</td><td>1</td></tr> <tr><td>baller</td><td>1</td><td>0</td><td>0</td><td>131</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	126	0	0	0	inner	1	101	0	0	outer	2	0	117	1	baller	1	0	0	131		norm	inner	outer	baller		
norm	63	0	0	0																																																																												
inner	0	51	0	0																																																																												
outer	0	0	59	1																																																																												
baller	1	0	0	65																																																																												
	norm	inner	outer	baller																																																																												
norm	63	0	0	0																																																																												
inner	1	50	0	0																																																																												
outer	2	0	58	0																																																																												
baller	0	0	0	66																																																																												
	norm	inner	outer	baller																																																																												
norm	126	0	0	0																																																																												
inner	1	101	0	0																																																																												
outer	2	0	117	1																																																																												
baller	1	0	0	131																																																																												
	norm	inner	outer	baller																																																																												
	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>51</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>0</td><td>0</td><td>54</td><td>6</td></tr> <tr><td>baller</td><td>0</td><td>1</td><td>11</td><td>54</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	0	51	0	0	outer	0	0	54	6	baller	0	1	11	54		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>49</td><td>0</td><td>2</td></tr> <tr><td>outer</td><td>0</td><td>1</td><td>57</td><td>2</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>7</td><td>59</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	0	49	0	2	outer	0	1	57	2	baller	0	0	7	59		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>126</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>100</td><td>0</td><td>2</td></tr> <tr><td>outer</td><td>0</td><td>1</td><td>111</td><td>8</td></tr> <tr><td>baller</td><td>0</td><td>1</td><td>18</td><td>113</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	126	0	0	0	inner	0	100	0	2	outer	0	1	111	8	baller	0	1	18	113		norm	inner	outer	baller		
norm	63	0	0	0																																																																												
inner	0	51	0	0																																																																												
outer	0	0	54	6																																																																												
baller	0	1	11	54																																																																												
	norm	inner	outer	baller																																																																												
norm	63	0	0	0																																																																												
inner	0	49	0	2																																																																												
outer	0	1	57	2																																																																												
baller	0	0	7	59																																																																												
	norm	inner	outer	baller																																																																												
norm	126	0	0	0																																																																												
inner	0	100	0	2																																																																												
outer	0	1	111	8																																																																												
baller	0	1	18	113																																																																												
	norm	inner	outer	baller																																																																												
	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>51</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>0</td><td>0</td><td>54</td><td>6</td></tr> <tr><td>baller</td><td>0</td><td>1</td><td>11</td><td>54</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	0	51	0	0	outer	0	0	54	6	baller	0	1	11	54		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>62</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>6</td><td>45</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>2</td><td>0</td><td>58</td><td>0</td></tr> <tr><td>baller</td><td>0</td><td>1</td><td>0</td><td>65</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	62	0	1	0	inner	6	45	0	0	outer	2	0	58	0	baller	0	1	0	65		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>125</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>6</td><td>96</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>2</td><td>0</td><td>112</td><td>6</td></tr> <tr><td>baller</td><td>0</td><td>2</td><td>11</td><td>119</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	125	0	1	0	inner	6	96	0	0	outer	2	0	112	6	baller	0	2	11	119		norm	inner	outer	baller		
norm	63	0	0	0																																																																												
inner	0	51	0	0																																																																												
outer	0	0	54	6																																																																												
baller	0	1	11	54																																																																												
	norm	inner	outer	baller																																																																												
norm	62	0	1	0																																																																												
inner	6	45	0	0																																																																												
outer	2	0	58	0																																																																												
baller	0	1	0	65																																																																												
	norm	inner	outer	baller																																																																												
norm	125	0	1	0																																																																												
inner	6	96	0	0																																																																												
outer	2	0	112	6																																																																												
baller	0	2	11	119																																																																												
	norm	inner	outer	baller																																																																												
	<table border="1"> <tr><td>norm</td><td>63</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>inner</td><td>0</td><td>49</td><td>0</td><td>2</td></tr> <tr><td>outer</td><td>0</td><td>1</td><td>57</td><td>2</td></tr> <tr><td>baller</td><td>0</td><td>0</td><td>7</td><td>59</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	63	0	0	0	inner	0	49	0	2	outer	0	1	57	2	baller	0	0	7	59		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>62</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>6</td><td>45</td><td>0</td><td>0</td></tr> <tr><td>outer</td><td>2</td><td>0</td><td>58</td><td>0</td></tr> <tr><td>baller</td><td>0</td><td>1</td><td>0</td><td>65</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	62	0	1	0	inner	6	45	0	0	outer	2	0	58	0	baller	0	1	0	65		norm	inner	outer	baller	<table border="1"> <tr><td>norm</td><td>125</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>inner</td><td>6</td><td>94</td><td>0</td><td>2</td></tr> <tr><td>outer</td><td>2</td><td>1</td><td>115</td><td>2</td></tr> <tr><td>baller</td><td>0</td><td>1</td><td>7</td><td>124</td></tr> <tr><td></td><td>norm</td><td>inner</td><td>outer</td><td>baller</td></tr> </table>	norm	125	0	1	0	inner	6	94	0	2	outer	2	1	115	2	baller	0	1	7	124		norm	inner	outer	baller		
norm	63	0	0	0																																																																												
inner	0	49	0	2																																																																												
outer	0	1	57	2																																																																												
baller	0	0	7	59																																																																												
	norm	inner	outer	baller																																																																												
norm	62	0	1	0																																																																												
inner	6	45	0	0																																																																												
outer	2	0	58	0																																																																												
baller	0	1	0	65																																																																												
	norm	inner	outer	baller																																																																												
norm	125	0	1	0																																																																												
inner	6	94	0	2																																																																												
outer	2	1	115	2																																																																												
baller	0	1	7	124																																																																												
	norm	inner	outer	baller																																																																												

TABLE 5. (Continued.) Confusion matrix visualization of CUT-2 dataset diagnosis results.

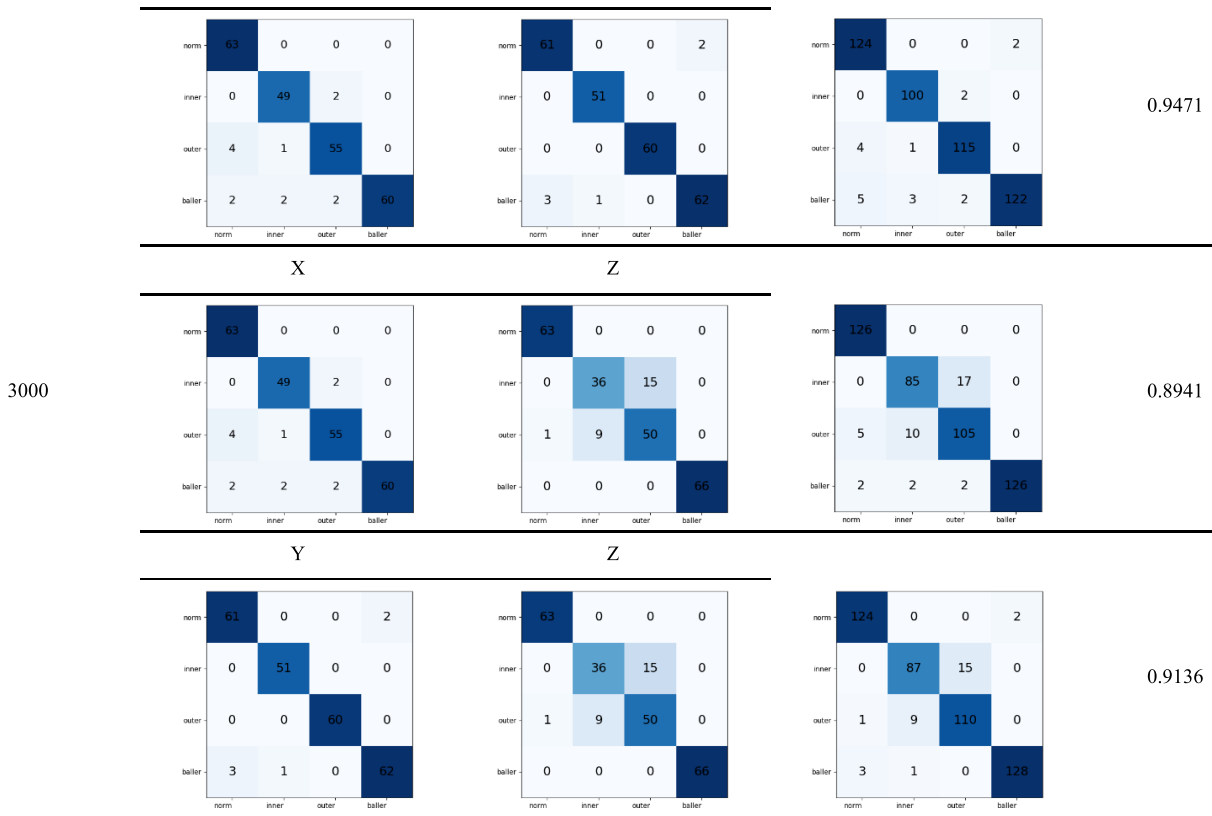


TABLE 6. Corrected sensor weights in X, Y, and Z directions.

Motor Speed (rpm)	Sensor combination	Sensor weight
2000	X,Y	X:0.59 Y:0.41
	X,Z	X:0.55 Z:0.45
	Y,Z	Y:0.56 Z:0.44
2500	X,Y	X:0.58 Y:0.42
	X,Z	X:0.53 Z:0.47
	Y,Z	Y:0.52 Z:0.48
3000	X,Y	X:0.60 Y:0.40
	X,Z	X:0.52 Z:0.48
	Y,Z	Y:0.54 Z:0.46

improved DS evidence theory, and the decision fusion of sensor data for the X and Z directions is not as good as the decision fusion of sensor data for the X and Y directions.

It can be seen from Figure 11 that IDS improves the diagnostic accuracy by an average of 2.84% in the sensor data fusion for the Y and Z directions. Compared with the other improved DS evidence theory, the average fusion accuracy

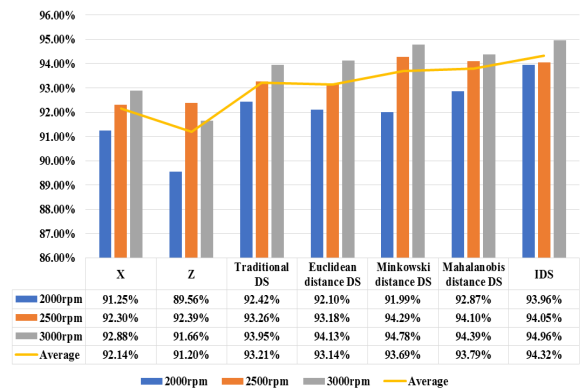


FIGURE 10. Comparison of diagnostic accuracy and fusion method results for X- and Z-direction sensors.

increases by at least an average of approximately 0.42% and by at most an average of approximately 0.95%. Therefore, compared with the other improved DS evidence theory, IDS achieves a lower fusion precision; however, it demonstrates a certain effectiveness and adaptability in the decision fusion of the sensor data of the Y and Z directions.

The improved DS evidence theory is fused on the diagnostic results of the sensor data in the X, Y, and Z directions, as shown in Figure 12. Compared to other fusion methods, IDS still achieves the highest average fusion accuracy.

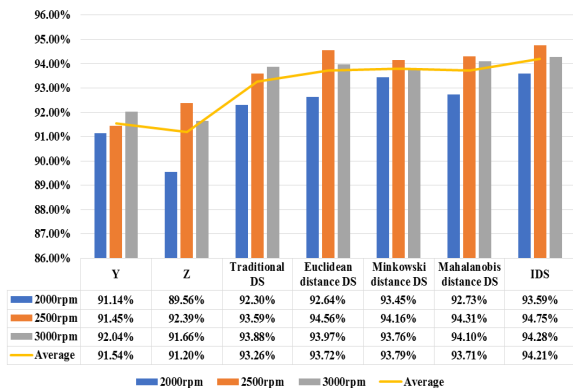


FIGURE 11. Comparison of diagnostic accuracy and fusion method results for Y- and Z- direction sensors.

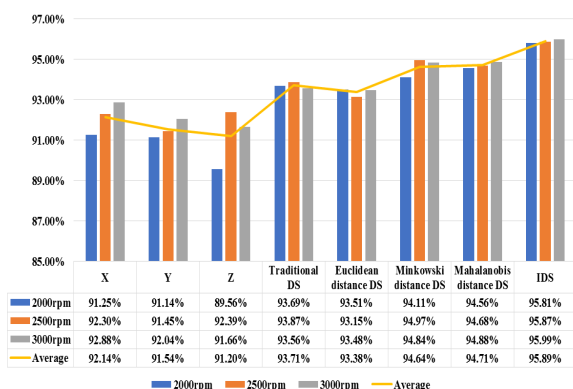


FIGURE 12. Comparison of diagnostic accuracy and fusion method results for X-, Y-, and Z- direction sensors.

Specifically, IDS is equally applicable and effective for the fusion of three-sensor data.

As shown in Figure 9-12, the average accuracy of the three-direction sensor information fusion reached 95.89%, which was higher than the two-direction sensor information fusion because when the number of sensors increases, the decision on the same sample also increases. IDS weighs the similarity measure between multiple evidence bodies by the kappa coefficient and corrects multiple weights of evidence. Therefore, a higher average fusion accuracy is achieved by integrating multiple sensor information fusions. In addition, as shown in Figure 12, the IDS fusion accuracy improves by an average of approximately 4.26% compared to the X-, Y-, and Z-direction sensors.

Based on Figure 9 – Figure 12, the following conclusions can be drawn: First, the IDS proposed in this paper can fuse multi-sensor information effectively; then, compared with DS evidence theory based on a corrected distance for evidence bodies, IDS is superior in terms of average fusion precision; finally, IDS demonstrates its validity and applicability on the CUT-2 dataset. Therefore, IDS can integrate multi-sensor information effectively, which provides a new idea for the improvement of evidence theory.

V. CONCLUSION

This paper proposes a bearing fault diagnosis method based on a feature selection feedback network and multi-sensor information fusion. The results can be summarized as follows:

1) This study proposes a feature network algorithm with a feedback effect. To realize the online selection of bearing fault diagnosis features, this study constructs a feature selection network by setting the feedback conditions of the feature quantity and feature selection time. The feature selection feedback network can achieve the same diagnostic accuracy as FF-FC-MIC; however, the feature selection dimension is lower.

2) To address the uncertainty and inaccuracy of single-sensor data-diagnosis results, this paper proposes an improved DS evidence theory based on the kappa coefficient. By comparing with traditional DS evidence theory and three DS evidence theories based on distance functions to correct the weight of the evidence body, this study proposes that the improved DS evidence theory can achieve higher average diagnostic accuracy.

Given the increase in the number and types of sensors in manufacturing, the amount of data has also increased exponentially. Therefore, the quick and effective online analysis and processing of massive data have become important research trends. Improving the fault diagnosis method in terms of time, complexity and other indicators while also realizing online fault detection, the online management and scheduling of diagnostic resources, and the online update of the diagnostic model are the focus of future research in online fault diagnosis.

REFERENCES

- [1] O. Abdeljaber, S. Sassi, O. Avci, S. Kiranyaz, A. A. Ibrahim, and M. Gabbouj, "Fault detection and severity identification of ball bearings by online condition monitoring," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 8136–8147, Oct. 2019.
- [2] W. Zheng, M. Tynes, H. Gorelick, Y. Mao, L. Cheng, and Y. Hou, "FlowCon: Elastic flow configuration for containerized deep learning applications," in *Proc. 48th Int. Conf. Parallel Process. (ICPP)*, Aug. 2019, pp. 87:1–87:10.
- [3] L. Ciabattoni, F. Ferracuti, A. Freddi, and A. Monteriu, "Statistical spectral analysis for fault diagnosis of rotating machines," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4301–4310, May 2018.
- [4] V. Tra, J. Kim, S. A. Khan, and J.-M. Kim, "Bearing fault diagnosis under variable speed using convolutional neural networks and the stochastic diagonal Levenberg-Marquardt algorithm," *Sensors*, vol. 17, no. 12, p. 2834, Dec. 2017.
- [5] K. H. Hui, C. S. Ooi, M. H. Lim, M. S. Leong, and S. M. Al-Obaidi, "An improved wrapper-based feature selection method for machinery fault diagnosis," *PLoS ONE*, vol. 12, no. 12, Dec. 2017, Art. no. e0189143.
- [6] X. P. Liu, H. Q. Zheng, and T. Y. Zhu, "Feature selection in machine fault diagnosis based on evolutionary Monte Carlo method," *J. Vib. Shock*, vol. 30, no. 10, pp. 98–101, 2011.
- [7] X. Yu, F. Dong, E. Ding, S. Wu, and C. Fan, "Rolling bearing fault diagnosis using modified LFDA and EMD with sensitive feature selection," *IEEE Access*, vol. 6, pp. 3715–3730, 2018.
- [8] B.-Z. Cui, H.-X. Pan, and Z.-B. Wang, "Fault diagnosis of roller bearings base on the local wave and approximate entropy," *J. North Univ. China (Natural Sci. Ed.)*, vol. 33, no. 5, pp. 552–558, 2012.
- [9] L. Shi, "Correlation coefficient of simplified neutrosophic sets for bearing fault diagnosis," *Shock Vib.*, vol. 2016, no. 2, pp. 1–11, 2016.

- [10] J. Zheng, J. Cheng, Y. Yang, and S. Luo, "A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination," *Mechanism Mach. Theory*, vol. 78, pp. 187–200, Aug. 2014.
- [11] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Inf. Fusion*, vol. 14, no. 1, pp. 28–44, 2013.
- [12] M. L. Fung, M. Z. Q. Chen, and Y. H. Chen, "Sensor fusion: A review of methods and applications," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 3853–3860.
- [13] Z. Duan, T. Wu, S. Guo, T. Shao, R. Malekian, and Z. Li, "Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: A review," *Int. J. Adv. Manuf. Technol.*, vol. 96, nos. 1–4, pp. 803–819, Apr. 2018.
- [14] K. Coussement, D. Benoit, and M. Antioco, "A Bayesian approach for incorporating expert opinions into decision support systems: A case study of online consumer-satisfaction detection," *Decis. Support Syst.*, vol. 79, pp. 24–32, Nov. 2015.
- [15] M. Ren, C. Cheung, and G. Xiao, "Gaussian process based Bayesian inference system for intelligent surface measurement," *Sensors*, vol. 18, no. 11, p. 4069, Nov. 2018.
- [16] C. Chen, A. Ugon, C. Sun, W. Chen, C. Philippe, and A. Pinna, "Towards a hybrid expert system based on sleep event's threshold dependencies for automated personalized sleep staging by combining symbolic fusion and differential evolution algorithm," *IEEE Access*, vol. 7, pp. 1775–1792, 2019.
- [17] A. Rikalovic and I. Cosic, "A fuzzy expert system for industrial location factor analysis," *Acta Polytechnica Hungarica*, vol. 12, no. 2, pp. 34–51, 2015.
- [18] Y. Gong, X. Su, H. Qian, and N. Yang, "Research on fault diagnosis methods for the reactor coolant system of nuclear power plant based on D-S evidence theory," *Ann. Nucl. Energy*, vol. 112, pp. 395–399, Feb. 2018.
- [19] B. Chen and J. Feng, "Multisensor information fusion of pulsed GTAW based on improved D-S evidence theory," *Int. J. Adv. Manuf. Technol.*, vol. 71, nos. 1–4, pp. 91–99, Mar. 2014.
- [20] J. Liu, A. Chen, and N. Zhao, "An intelligent fault diagnosis method for bogie bearings of metro vehicles based on weighted improved D-S evidence theory," *Energies*, vol. 11, no. 1, p. 232, Jan. 2018.
- [21] N. Xie, Z. Li, and G. Zhang, "An intuitionistic fuzzy soft set method for stochastic decision-making applying prospect theory and grey relational analysis," *J. Intell. Fuzzy Syst.*, vol. 33, no. 1, pp. 15–25, Jun. 2017.
- [22] M. Sadrykia, M. R. Delavar, and M. Zare, "A GIS-based decision making model using fuzzy sets and theory of evidence for seismic vulnerability assessment under uncertainty (case study: Tabriz)," *J. Intell. Fuzzy Syst.*, vol. 33, no. 3, pp. 1969–1981, Aug. 2017.
- [23] G. Rowe and G. Wright, "The Delphi technique as a forecasting tool: Issues and analysis," *Int. J. Forecasting*, vol. 15, no. 4, pp. 353–375, Oct. 1999.
- [24] A. Mosavi, P. Ozturk, and K.-W. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, p. 1536, Oct. 2018.
- [25] J. Dou, A. P. Yunus, D. Tien Bui, A. Merghadi, M. Sahana, Z. Zhu, C.-W. Chen, K. Khosravi, Y. Yang, and B. T. Pham, "Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic island, Japan," *Sci. Total Environ.*, vol. 662, pp. 332–346, Apr. 2019.
- [26] A. N. Richter and T. M. Khoshgoftaar, "A review of statistical and machine learning methods for modeling cancer risk using structured clinical data," *Artif. Intell. Med.*, vol. 90, pp. 1–14, Aug. 2018.
- [27] S. Hamze-Ziabari and T. Bakhshpoori, "Improving the prediction of ground motion parameters based on an efficient bagging ensemble model of M5' and CART algorithms," *Appl. Soft Comput.*, vol. 68, pp. 147–161, Jul. 2018.
- [28] M. Seera, C. P. Lim, and S. C. Tan, "A hybrid FAM-CART model for online data classification," *Comput. Intell.*, vol. 34, no. 2, pp. 562–581, May 2018.
- [29] F. Ye, J. Chen, Y. Li, "Improvement of DS evidence theory for multi-sensor conflicting information," *Symmetry*, vol. 9, no. 5, p. 69, May 2017.
- [30] Y. Li, J. Chen, F. Ye, and D. Liu, "The improvement of DS evidence theory and its application in IR/MMW target recognition," *J. Sensors*, vol. 2016, no. 6, pp. 1–15, 2016.
- [31] X. Tang, J. Wang, J. Lu, G. Liu, and J. Chen, "Improving bearing fault diagnosis using maximum information coefficient based feature selection," *Appl. Sci.*, vol. 8, no. 11, p. 2143, Nov. 2018.



XIANGHONG TANG received the Ph.D. degree in computer software and theory from the Huazhong University of Science and Technology, Wuhan, China, in 2010.

From 2016 to 2017, he had an academic visit as a Visiting Scholar with the University of Illinois at Chicago, USA. He is currently a Professor and also the Deputy Dean of the State Key Laboratory of Public Big Data, Guizhou University, Guiyang, China. His current research interests include intelligent manufacturing, fault diagnosis, and data mining.

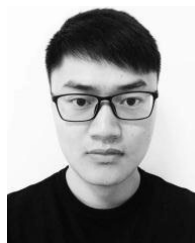


XIN GU received the B.S. degree in mechanical engineering from the Guizhou Institute of Technology, Guiyang, China, in 2018. He is currently pursuing the M.S. degree with the Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, China.

His research interests include fault diagnosis, deep learning, and data mining.



JIACHEN WANG received the B.S. degree in mechanical engineering from Northwest University, in 2014, and the M.S. degree from the Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, China, in 2019. During the M.S. degree, he mainly studied fault diagnosis of rolling bearing.



QIANG HE received the B.S. degree in mechanical engineering from the University of Science and Technology Liaoning, in 2018. He is currently pursuing the M.S. degree with the Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, China.

His research interests include machine learning, deep learning, and data processing.



FAN ZHANG received the B.S. degree in mechanical engineering from the Hunan University of Science and Technology, in 2017. He is currently pursuing the M.S. degree with the Key Laboratory of Advanced Manufacturing Technology, Ministry of Education, Guizhou University, China.

His research interests include machine learning, deep learning, compressed sensing, and data mining.



JIANGUANG LU received the Ph.D. degree in computer software and theory from the Chengdu Institute of Computer Application, Chinese Academy of Science, China, in 2016. He is currently an Associate Professor with Guizhou University, China. His research interests include machine learning, data mining, and intelligent manufacturing.

• • •