Research Article

Reliable Fault Diagnosis of Rolling Bearing Based on Ensemble Modified Deep Metric Learning

Zengbing Xu [ID] **,[1,2,3] Xiaojuan Li,[1,2] Jinxia Wang,[4] and Zhigang Wang[1,2]**

[1]*Key Laboratory of Metallurgical Equipment and Control Technology (Wuhan University of Science and Technology), Ministry of Education, Wuhan 430081, China*
[2]*Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering (Wuhan University of Science and Technology), Wuhan 430081, China*
[3]*The State Key Lab. of Digital Manufacturing Equipment & Technology, Huazhong University of Science and Technology, Wuhan 430074, China*
[4]*Rootcloud Technology Co., Ltd., Changsha 410199, China*

Correspondence should be addressed to Zengbing Xu; xuzengbing@163.com

A novel ensemble Yu's norm-based deep metric learning (DMLYu) is proposed to diagnose the fault of rolling bearing in this paper, which can diagnose the fault classes through the information fusion method that combines the different diagnosis results produced by several Yu's norm-based deep metric learning models with different scale signals. The suggested method is composed of three steps: firstly the vibration signal is decomposed into multiple IMF components by the EEMD method, then these IMF components are input into the DMLYu models which is called the modified deep metric learning model based on Yu's norm-based similarity measure, respectively, to extract the feature parameters to diagnose the fault of rolling bearings from the different scales, and finally the final diagnosis decision is made by fusion strategy based on Bayesian belief method (BBM). At last, through a multifaceted diagnosis test of rolling bearing on different datasets, the effectiveness of the proposed ensemble DMLYu based on BBM is verified, and the superiority of the proposed diagnosis method is validated by comparing its diagnosis accuracy and generalization with DMLYu based on voting method and the individual DMLYu model.

1. Introduction

The rolling bearing is the key component of rotating machinery; its failure can lead to the equipment shutdown and the enterprise economic losses. In order to ensure the normal operation of the rolling bearing, it is necessary to monitor and diagnose the fault conditions of the bearing [1–4]. Currently, the mechanical fault diagnosis has entered the era of mechanical big data; it is of great significance to study how to effectively use the big data to diagnose the bearing faults [5, 6].

Presently deep learning has been widely applied to the field of fault diagnosis of mechanical equipment because of the strong ability of automatically learning discriminative feature parameters from the mechanical big data through the multiple-layer nonlinear transformation. Different kinds of deep learning models, such as deep belief network (DBN), the convolution neural network (CNN), and deep autoencoder model, have been developed to diagnose the fault categories of different mechanical equipment [7–13]. Although these deep learning models can obtain higher diagnosis accuracy than the shallow neural network models, their diagnosis accuracy and generalization ability need to be improved, and their diagnosis mechanism is unexplainable in the process of diagnosis [14].

Deep metric learning (DML) which can map original feature parameters to discriminative feature space by maximizing interclass variation and minimizing intraclass variation is also suggested to be applied to the field of pattern recognition [15–18]. These DML models can use the distance

metric criterion to classify the data samples with explainable classification mechanism, but in the field of fault diagnosis some mechanical signals are too difficult to be diagnosed because of the complexity of signal transmission path and insensitiveness of the fault parameter features to fault categories; in particular, some data samples in the boundary region of different fault categories can be misclassified by the DML based on distance metric criterion [18]. Afterwards a DML based on Yu's norm similarity measure (DMLYu) is proposed to diagnose the fault of rolling bearings, which can automatically extract feature parameters through the fuzzy formalism and the multiple-layer nonlinear transformation, and it can recognize the faulty data samples in the boundary region of different fault categories with higher accuracy [19].

Although the individual DMLYu can take advantages of the end-to-end learning ability of deep learning mechanism to obtain discriminative feature parameters from the original vibration signals, like other deep learning models it can not reveal thoroughly the nonstationary dynamic fault characteristics concealed in the time domain vibration signals. To solve this problem, different kinds of time frequency analysis methods, such as short-time Fourier transform (STFT), wavelet packet transform (WPT), and empirical mode decomposition (EMD), have been combined with deep learning for fault diagnosis [20, 21].

Furthermore, the diagnosis accuracy and reliability of the DMLYu model are degraded because of the small faulty data samples and overfitting. Ensemble learning based on the decision fusion strategy has been verified to overcome these limitations of individual deep learning model and achieve higher accuracy and reliability because of the complementary classification behaviors among different classifiers [22]. When multiple deep learning models combined with the different scale components of original signal are applied to the same fault diagnosis problems, its final diagnosis performance is superior to that of the individual deep learning model. Therefore, some ensemble deep learning models have been developed to apply to the field of fault diagnosis with higher accuracy and reliability [23–25].

In view of the above principles, a novel ensemble DMLYu model based on the Bayesian belief method and ensemble empirical mode decomposition (EEMD) is proposed to diagnose the faults of rolling bearings, and the contribution and innovation of the proposed fault diagnosis method are written as follows:

(1) The deep metric learning based on Yu's norm can classify the data samples in the boundary region of different fault categories with higher accuracy and explainable classification mechanism because of similarity measure based on the fuzzy rule of Yu's norm

(2) The individual DMLYu model can extract more insensitive fault feature parameters from the IMF component which can describe the fault related information from the different scale view

(3) The ensemble DMLYu model based on the Bayesian belief method can diagnose the fault of bearings with

the highest accuracy and generalization ability because of the fusion mechanism of several DMLYu models with different IMF components

The diagnosis schematic flowchart is shown in Figure 1. Firstly, the vibration signal is gradually truncated through the sliding window and divided into N data sample segments, and then the data sample segment is decomposed into multiple IMF components. Secondly, each IMF component is fed into the DMLYu to diagnose the fault of bearings, respectively. Finally, the final diagnosis result can be obtained by the Bayesian belief fusion technique. The remainder of the paper is organized as follows. The proposed ensemble Yu's norm-based deep metric learning model which synthesizes the EEMD algorithm and the basic theory of DMLYu model with the Bayesian belief method is described in Section 2; the fault diagnosis experiment of rolling bearings is conducted in Section 3. At last the conclusions are drawn in Section 4.

2. Ensemble Yu's Norm-Based Deep Metric Learning Model

Owing to the fact that ensemble deep metric learning model inherits the advantages of both the deep metric learning models and the ensemble learning, the ensemble deep metric learning has better generalization performance and higher diagnosis accuracy. Accordingly, in order to improve the diagnosis accuracy and robustness of individual DMLYu model, referring to Figure 1 the different scale components of vibration signal which is decomposed by the EEMD method to describe the fault related information from the different viewpoint are input into multiple DMLYu models to form the proposed ensemble DMLYu model based on the Bayesian belief fusion method.

2.1. EEMD Method. EEMD which is the improved version of the EMD is mainly to decompose the original signal into multiple intrinsic mode functions (IMFs) and solve the problem of mode mixing in the EMD method by a noise-assisted analysis method [26]. It produces a collection of series by adding the white noise with statistical property of uniform distribution in frequency range to the primitive signal firstly and then processing the newly acquired series with EMD method. The algorithm of EEMD is described as follows:

(1) Given a time series $x(t)$, a new time series $x_i(t)$ can be generated by adding a white noise with the given amplitude $w_i(t)$ to the primitive data series $x(t)$, namely,

$$x_i(t) = x(t) + w_i(t), \tag{1}$$

where $w_i(t)$ is the added white noise of the ith trial.

(2) Use EMD to decompose the time series $x_i(t)$, then the jth IMF components $c_{ij}(t)$ and one residual component $r_i(t)$ are obtained as the following formula:
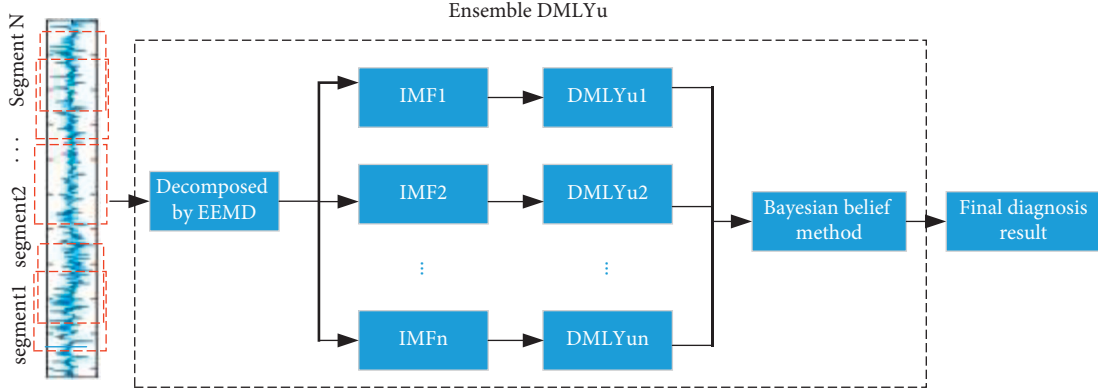
FIGURE 1: The flowchart of the proposed fault diagnosis method.

$$x_i(t) = \sum_{j=1}^{n} c_{ij}(t) + r_i(t), \tag{2}$$

where $c_{ij}(t)$ is the jth IMF component of n IMF components of the ith trial.

(3) Repeat step (1) and step (2) with the given M trials; calculate the ensemble mean of all trials. The corresponding formulae are depicted in the following equations:

$$c_j(t) = \frac{1}{M} \sum_{i=1}^{M} c_{ij}(t),$$
$$\tag{3}$$
$$r(t) = \frac{1}{M} \sum_{i=1}^{N} r_i(t).$$

Finally, the decomposition result of the original time series $x(t)$ by the EEMD can be written as the following equation:

$$x(t) = \sum_{j} c_j(t) + r(t), \tag{4}$$

where $r(t)$ is the final residual component of EEMD.

In addition, it should be noted that amplitude of the added white noise and the number M of trials are two key parameters in this algorithm.

2.2. The Basic Theory of the Deep Metric Learning Based on Yu's Norm

2.2.1. Deep Metric Learning. Deep metric learning (DML) is developed on the basis of the deep learning which can learn hierarchical feature representation directly from original data sample, which use the deep neural network structure to integrate feature learning and metric learning into a joint learning framework [27]. The network architecture of DML consists of input layer, multiple hidden layers, and output layer, which can be seen in Figure 2. It can also compute the feature representation $h^{(N)}$ of a data sample x by passing it to multiple-layer nonlinear transformations and map the original feature parameters to discriminative feature space

by maximizing interclass variation and minimizing intraclass variation [15, 16].

2.2.2. Deep Metric Learning Based on Yu's Norm. The traditional DML based on Euclidean distance can misdiagnose the data samples in the overlapping region of the different fault classes because of the nonlinearity of the classification boundary line. Especially when the faulty data samples are fuzzy and insensitive to the fault classes, the misdiagnosis ratio is even higher. But the metric learning based on Yu's norm depends on similarity between different data samples rather than the distance, which classify the data samples by the similarity measure effectively. Correspondingly the DML based on Yu's norm-based similarity (DMLYu) is proposed.

Assuming that there are $N + 1$ layers in the deep network and $p^{(n)}$ units in the nth layer, where $n \in [1, 2, \ldots, N]$, the output of x at the nth layer is computed as

$$f^{(n)}(x) = h^{(n)} = \varphi\left(W^{(n)} h^{(n-1)} + b^{(n)}\right), \tag{5}$$

where $W^{(n)} \in R^{p^{(n)} \times p^{(n-1)}}$ and $b^{(n)} \in R^{p^{(n)}}$ are the weight and bias of the parameters in the nth layer. φ is the nonlinear activation function of each layer, which is set as a tanh function here. The nonlinear mapping $f^{(n)}$ is a function parameterized by $\{W^{(i)}\}_{i=1}^{n}$ and $\{b^{(i)}\}_{i=1}^{n}$. For each pair of input samples x_i and x_j, their corresponding representation at the nth layer of the deep network can be written as $f^{(n)}(x_i)$ and $f^{(n)}(x_j)$. Correspondingly the Euclidean distance of the data sample points x_i and x_j in the deep metric network space is substituted by the similarity based on Yu's norm which is written as follows:

$$d_{f^{(n)}}(X_i, X_j) = S < f^{(n)}(X_i), f^{(n)}(X_j) >, \tag{6}$$

where $S\langle X_i, X_j \rangle = \max(0, (1 + \lambda)(Sn(\overline{X}_1, X_j) + Sn(X_i, \overline{X}_j) - 1) - \lambda Sn(\overline{X}_1, X_j) Sn(X_i, \overline{X}_J))$, $Sn(X_i, X_j) = \min[1, X_i + X_j + \lambda X_i X_j]$, $\overline{X}_1 = 1 - X_i$, and λ is a constant.

Based on the graph embedding framework, the Marginal Fisher Analysis (MFA), which is a supervised descendent dimension algorithm that measures the similarity between every data sample and its neighbor samples, is conducted on the output of all the training samples at the top layer of deep
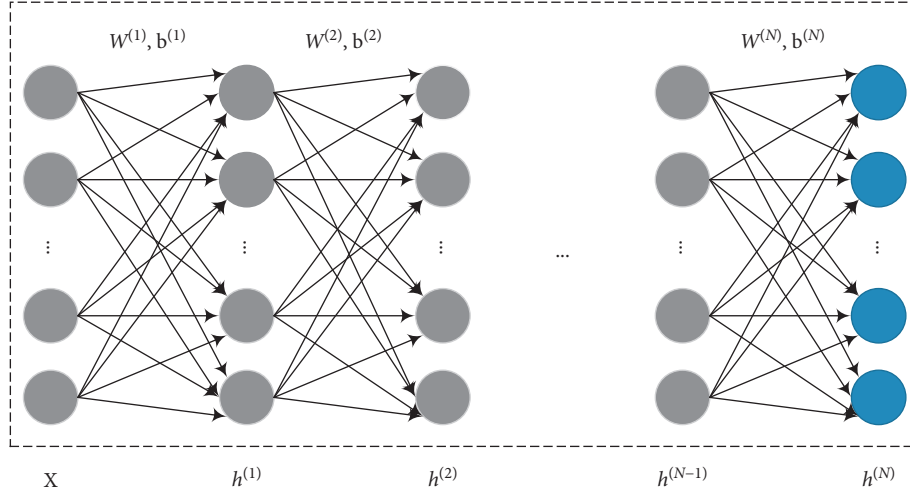
FIGURE 2: The network architecture. X is the input data sample of the network, $(h^{(1)} \ldots h^{(N-1)})$ are the output of the hidden layers, and $h^{(N)}$ is the output of the top layer. $W^{(N)}$ and $b^{(N)}$ are the parameters of the network to be learned, $1 \le n \le N$.

neural network; a strongly supervised deep metric learning model is constructed and formulated as follows:

$$\min_{f^{(N)}} J = S_c^{(N)} - \alpha S_c^{(N)} + \gamma \sum_{n=1}^{n} \left(\left\| W^{(n)} \right\|_F^2 + \left\| b^{(n)} \right\|_2^2 \right), \quad (7)$$

where α is the free parameter which balances the importance between intraclass compactness and the interclass separability; the larger α is, greater the interclass scatter is; γ is the adjustable regularization parameter, $\gamma > 0$; $\|Z\|_F$ denotes the Frobenius norm of the matrix Z; $S_c^{(n)}$ and $S_b^{(n)}$ are the intraclass compactness and interclass separability, respectively, and their formula can be written as follows:

$$S_c^{(n)} = \frac{1}{Mk_1} P_{ij} d_{f^{(n)}} \left(X_i, X_j \right),$$

$$\tag{8}$$

$$S_b^{(n)} = \frac{1}{Mk_2} Q_{ij} d_{f^{(n)}} \left(X_i, X_j \right),$$

where M is the number of data samples in the training set; P and Q are adjacency matrixes. If X_j is one of the k_1-intraclass nearest neighbors of X_i, then P_{ij} is set to 1, otherwise 0; if X_j is one of the k_2-interclass nearest neighbors of X_i, Q_{ij} is set to 1, otherwise 0.

The subgradient descent method is utilized to optimize the parameters $\{W^{(n)}, b^{(n)}\}$ in equation (7). The gradient of the objective function J with respect to the parameters $W^{(n)}$ and $b^{(n)}$ is computed as follows:

$$\frac{\partial J}{\partial W^{(n)}} = \frac{1}{Mk1} P_{ij} \left(L_{ij}^{(n)} h_i^{(n-1)T} + L_{ji}^{(n)} h_j^{(n-1)T} \right) - \frac{\alpha}{Mk2} Q_{ij} \left(L_{ij}^{(n)} h_i^{(n-1)T} + L_{ji}^{(n)} h_j^{(n-1)T} \right) + 2\gamma W^{(n)},$$

$$\tag{9}$$

$$\frac{\partial J}{\partial b^{(n)}} = \frac{1}{Mk1} P_{ij} \left(L_{ij}^{(n)} + L_{ji}^{(n)} \right) - \frac{\alpha}{Mk2} \sum_{i=1}^{M} \sum_{j=1}^{M} Q_{ij} \left(L_{ij}^{(n)} + L_{ji}^{(n)} \right) + 2\gamma b^{(n)},$$

where $h_i^{(0)} = x_i$ and $h_j^{(0)} = x_j$ are the original input data samples of the network; for all other layers $n = 1, 2, \ldots, N - 1$, the updated equations are written as follows:

$$L_{ij}^{(N)} = \left(h_i^{(N)} - h_j^{(N)} \right) \odot \varphi' \left(Z_i^{(N)} \right),$$
$$L_{ji}^{(N)} = \left(h_j^{(N)} - h_i^{(N)} \right) \odot \varphi' \left(Z_j^{(N)} \right),$$
$$L_{ij}^{(n)} = \left(W^{(n+1)T} L_{ij}^{(n+1)} \right) \odot \varphi' \left(Z_i^{(n)} \right), \quad (10)$$
$$L_{ji}^{(n)} = \left(W^{(n+1)T} L_{ji}^{(n+1)} \right) \odot \varphi' \left(Z_j^{(n)} \right),$$

where the operation $\odot$ denotes the elementwise multiplication and $W^{(n)}$ and $b^{(n)}$ can be updated by the following gradient descent algorithm until convergence:

$$W^{(n)} = W^{(n)} - \tau \frac{\partial J}{\partial W^{(n)}},$$

$$\tag{11}$$

$$b^{(n)} = b^{(n)} - \tau \frac{\partial J}{\partial W^{(n)}},$$

where τ is learning rate.

In addition, the backpropagation neural network (BPNN) is introduced into the top output layer of the DMLYu and is used to further fine-tune the parameters of the network [26, 27] and diagnose the data samples in this paper.

2.3. Decision Fusion Based on Bayesian Belief Method.

Based on the assumption of mutual independency of classifiers and the diagnosis error of each classifier, the Bayesian belief method (BBM) can obtain the final diagnosis result by fusion of the belief measure of each classifier which is computed by the confusion matrix of each classifier [28]. Assume that there are M known fault classes and K classifiers in the same diagnostic task; the classifier e_k can be depicted as a function:

$$e_k(x) = j, \tag{12}$$

where $k = 1, 2, \ldots, K$, $j \in \{1, 2, \ldots, M, M + 1\}$, $M + 1$ is unknown fault class label, $e_k(x)$ signifies that the sample x is assigned to class j by the classifier e_k, and its two-dimensional confusion matrix can be calculated by the following equation:

$$CM_k = \begin{pmatrix} n_{11}^k & n_{12}^k & \cdots & n_{1M}^k & n_{1(M+1)}^k \\ n_{21}^k & n_{22}^k & \cdots & n_{2M}^k & n_{2(M+1)}^k \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ n_{M1}^k & n_{M2}^k & \cdots & n_{MM}^k & n_{M(M+1)}^k \end{pmatrix}, \tag{13}$$

which is obtained by executing $e_k(x)$ on the test dataset after $e_k(x)$ is trained. Each row i represents class c_i, each column j represents $e_k(x) = j$. The matrix unit n_{ij}^k represents that the input samples from class c_i are assigned to class c_j by the classifier $e_k(x)$. The number of samples in class c_i is $n_{i.}^k = \sum_{j=1}^{M+1} n_{ij}^k$, where $i = 1, 2, \ldots, M$, and the number of samples labeled j by $e_k(x)$ is $n_{.j}^k = \sum_{i=1}^{M} n_{ij}^k$, where $j = 1, 2, \ldots, M + 1$. A belief measure of classifier e_k can be calculated by the following belief function:

$$b_k(x \in c_i \mid e_k(x)) = P(x \in c_i \mid e_k(x) = j) = \frac{n_{ij}^k}{n_{.j}^k}, \tag{14}$$

where $i = 1, 2, \ldots M$, $j = 1, 2, \ldots, M + 1$. But this belief function is only suitable when the number of samples in each class is the same. When the number of data samples in each class is different, the diagnosis accuracy decreases because of the imprecise beliefs. So an improved belief function is calculated as follows:

$$ib_k(x \in c_i \mid e_k(x)) = P(x \in c_i \mid e_k(x) = j) = \frac{n_{ij}^k / n_{i.}^k}{\sum_{i=1}^{M}(n_{ij}^k / n_{i.}^k)}. \tag{15}$$

When K classifiers $e_1, e_2, \ldots, e_k$ are utilized, their corresponding belief measures $ib_1, ib_2, \ldots, ib_k$ can be computed by (15). Fusion strategies of all K classifiers can result in the final belief measure of the ensemble classifiers, which is the average algorithm. And its formula is depicted as follows:

$$b(i) = b(x \in c_i \mid e_1(x), e_2(x), \ldots, e_k(x), EN)$$
$$= P(x \in c_i \mid e_1(x), e_2(x), \ldots, e_k(x), EN) = \frac{1}{M+1} \sum_{k=1}^{K} b_k(x \in c_i \mid e_1(x), EN), \tag{16}$$

where $i = 1, 2, \ldots M + 1$, EN denotes the common classification environment. Thus, the sample x is classified into a class $c_j (j = 1, 2, \ldots, M + 1)$ according to belief of making the final decision $B(j) = \max_{i=1}^{M+1} b(i)$.

2.4. General Diagnosis Procedure of the Ensemble DMLYu Model.

In order to obtain the higher diagnosis accuracy and stronger generalization, the ensemble DMLYu is proposed to diagnose the fault of rolling bearings. The corresponding general diagnosis procedure is summarized as follows:

Step 1: collecting the data samples of different fault classes of rolling bearings by the sliding window from the vibration data.

Step 2: selecting the training data samples and decomposing each data sample into n IMF components by the EEMD method.

Step 3: inputting n IMF components into different N DMLYu models, respectively, and obtaining N initial diagnosis results. Then the final diagnosis decision can be obtained by the fusion strategy based on the BBM.

Step 4: using the ensemble DMLYu model to diagnose the testing data sample.

3. Fault Diagnosis of Rolling Bearings

3.1. Acquisition of Vibration Data. In order to verify the validity of the proposed ensemble DMLYu model and ensure the credibility of diagnosis results, the vibration signal used is obtained from the dataset of the rolling element bearings [29]. Figure 3 shows the photo and the schematic diagram of experiment rig. A three-phase induction motor is connected to a dynamometer and a torque sensor by a self-aligning coupling. The rolling element bearings are installed in a motor driven mechanical system. The dynamometer is used to control the desired torque load levels. An accelerometer is mounted at the 12 o'clock position at the driven end of the motor housing. The vibration data are acquired with the 12 K/s sample rate. The test bearing type is 6205-2RSJEMSKF, deep groove ball bearing.

To simulate the different fault categories and severities of bearings, the single point defects are introduced by the electrodischarge machining. Four different defect diameters (0 mm, 0.18 mm, 0.36 mm, and 0.54 mm) are introduced into the inner race, the ball, and outer race, respectively, and the defect depths are all 0.28 mm. Each bearing is tested under the 0 hp loads and 1800 rpm. The dataset contains 10 fault categories. The number of data samples for each fault class is 500, in which the number of training samples is 350 and the test samples is 150 and the total number of training data samples and test samples is 3500 and 1500, respectively. Additionally each data sample has 512 sample points. The detailed data statistics is described in Table 1.

3.2. Diagnosis Analysis of Rolling Bearing

3.2.1. Decomposition of Vibration Signal by EEMD Method. To obtain the different scale components which are input into these multiple DMLYu models, respectively, the training samples and test samples are all decomposed into the IMF components by the EEMD method in which the ratio of the standard deviation of the added white noise is set as 0.1 and the trial M is set as 100 here. Figure 4 shows the different IMF components which are derived from the slight fault signal of inner race. From the figure it can be seen that the original vibration is decomposed into 8 IMF components and one residual signal, and these IMF components can describe their respective dynamic characteristics from the different scale and viewpoints.

3.2.2. Diagnosis Performance Comparison and Discussion. To verify the effectiveness and superiority of the proposed ensemble DMLYu model based on BBM, the individual DMLYu model and the ensemble DMLYu models based on voting method are all used to diagnose the fault of bearings.

In addition, in all the diagnosis tests the parameters of the ensemble DMLYu model are described as follows: firstly, the number of DMLYu models for ensemble is set as 8. Secondly, the DMLYu model is comprised of 1 input layer

and 2 hidden layers, the corresponding node number is 512-100-100, respectively, and the second hidden layer is connected with the BPNN classifier, and the node number of the output layer in BPNN is set as 10 which represents the number of fault classes. At last, according to the empirical experience for the hyperparameters of DMLYu, α and λ are set as 4.0 and 0.2, respectively, the maximum number of iterations T is set as 10, the regularization parameter γ is set as 0.5, the initial learning rate τ is set as 0.2, and its corresponding learning rate decline factor is set as 0.95.

(1) Performance Comparison with the Individual DMLYu and Other Ensemble DMLYu Models. In this section, comparison of the diagnosis performance of the proposed ensemble DMLYu based on BBM with the individual DMLYu model and ensemble DMLYu model based on voting method is conducted. The original vibration signal and each of its IMF components which are decomposed by EEMD method are all used as the input of the individual DMLYu model, respectively, to diagnose the fault of rolling bearings. Figures 5(a)–5(i) show the corresponding diagnosis results, in which the X axis is the data sample sequence and Y axis is the fault class name. From the figure, it can be seen that the diagnosis accuracy of the individual DMLYu model with the original vibration signal is 95.4% and some data samples are misclassified into other fault classes, and it is lower than the diagnosis accuracy produced by the individual DMLYu model with the 1st IMF component, but it is higher than the diagnosis accuracy produced by the individual DMLYu model with the other IMF components. These demonstrated that the fault information concealed in time domain cannot reveal the nonstationary dynamic characteristics of the vibration signal thoroughly; the 1st IMF component can contain more fault related information than other IMF components.

Meanwhile, the figure also shows that the highest diagnosis accuracy is 98% which is obtained by the individual DMLYu model with the 1st IMF component; the lowest diagnosis accuracy is 25.67% which is produced by the individual DMLYu model with the 8th IMF component. But the diagnosis accuracy does not decrease with the increase of order number of the IMF component. For example, the diagnosis accuracy of the individual DMLYu model with the 6th IMF component is higher than that of the individual DMLYu model with the 5th IMF component. These are mainly because that each IMF component depicts the different local fault related information from the perspective of different time scales, the contribution of the fault information to the diagnosis capability of DMLYu model is different, and the corresponding diagnosis accuracy is also different.

Figure 6 shows the diagnosis result produced by the ensemble DMLYu based on BBM with the first n IMF components. In the figure, the bar on the vertical measurement represents the diagnosis accuracy of the individual DMLYu with the different IMF component; the curve plot represents the diagnosis accuracy produced by the ensemble DMLYu model based on BBM with the first n IMF components against the number of IMF component. From the

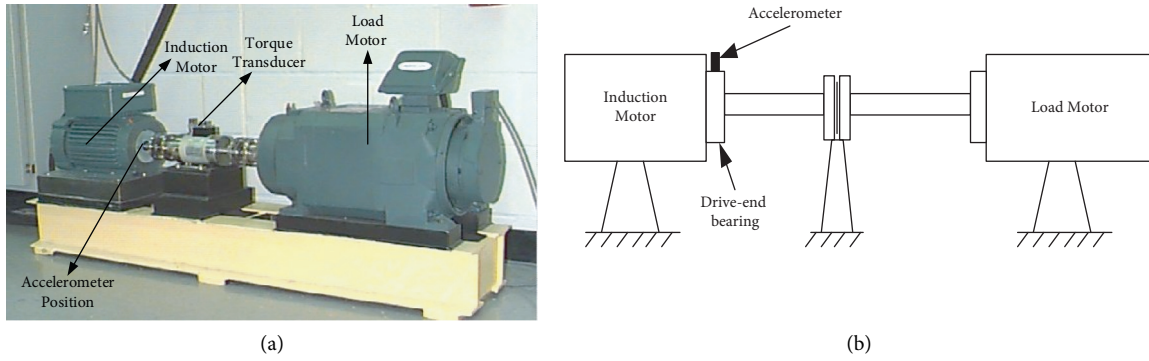(a)   (b)

FIGURE 3: The photo and the schematic diagram of experiment rig. (a) The experimental setup; (b) the schematic diagram of the experimental setup.

TABLE 1: The faulty data statistics of rolling bearing.

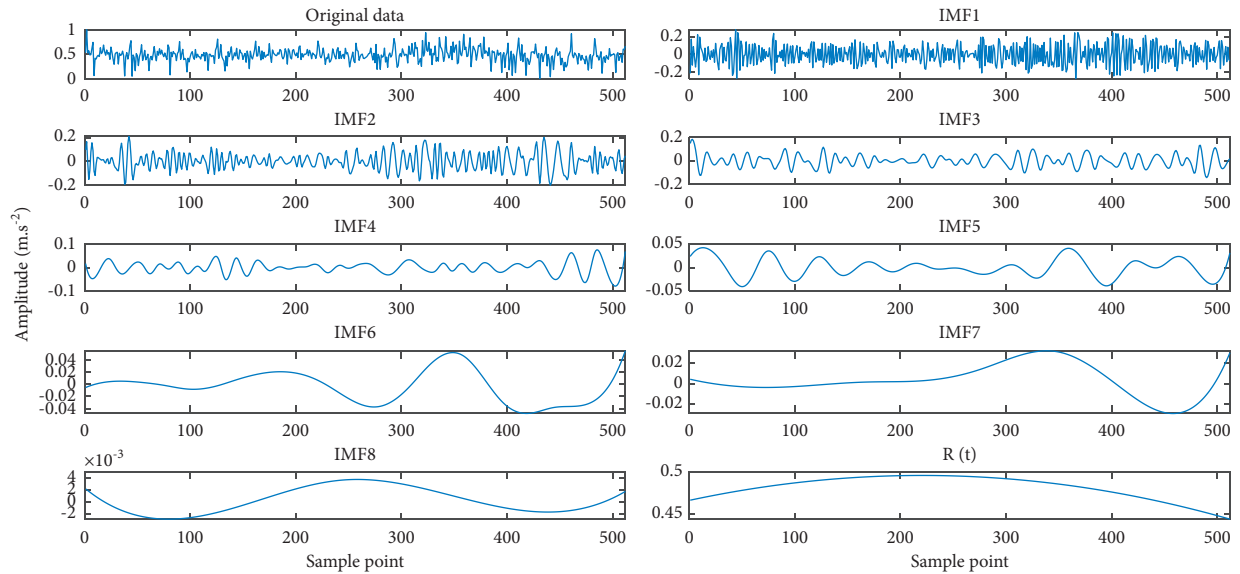| Fault class label	Fault condition	Fault size (mm)	Training samples	Test samples
1	Normal	0	350	150
2	Light inner race	0.18	350	150
3	Moderate inner race	0.36	350	150
4	Severe inner race	0.54	350	150
5	Light ball	0.18	350	150
6	Moderate ball	0.36	350	150
7	Severe ball	0.54	350	150
8	Light outer race	0.18	350	150
9	Moderate outer race	0.36	350	150
10	Severe outer race	0.54	350	150

FIGURE 4: The decomposition of slight fault signal of inner race.

figure it can be seen that the diagnosis accuracy of ensemble DMLYu model increases with the increase of the number of DMLYu models used for ensemble. When the number of the DMLYu models with the IMF components used for ensemble is 8 the diagnosis accuracy of ensemble DMLYu model based on BBM is up to 100%. And from Figures 5 and 6, it can be seen that the diagnosis accuracy of ensemble model with different number of DMLYu models is higher than that of the individual DMLYu model with any IMF component and original vibration signal because of the complementary classification behaviors of all individual DMLYu models used for ensemble. All these can demonstrate that ensemble DMLYu model based on BBM is superior to the individual DMLYu model obviously.
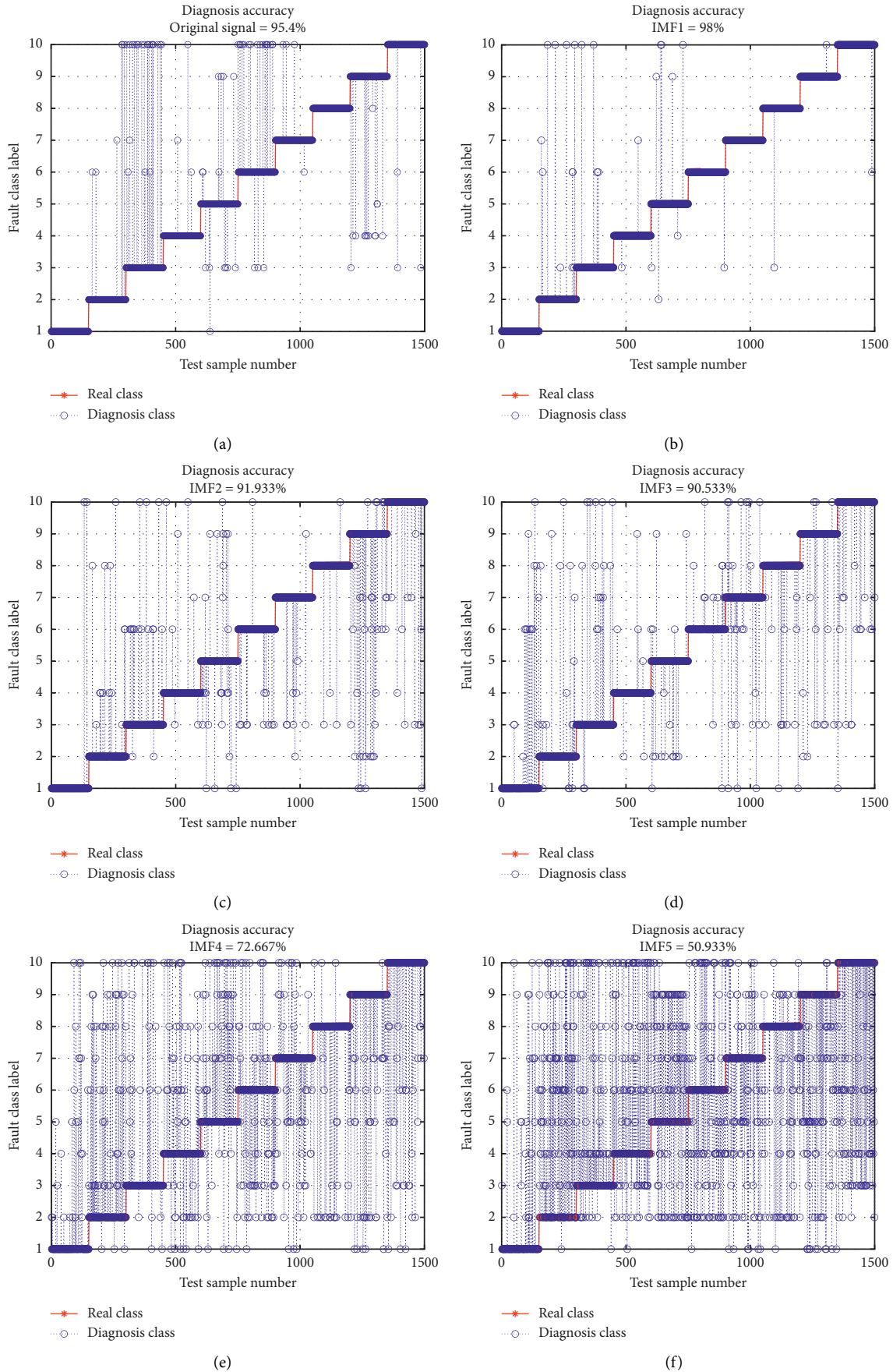
Diagnosis accuracy
Original signal = 95.4%

(a)

Diagnosis accuracy
IMF1 = 98%

(b)

Diagnosis accuracy
IMF2 = 91.933%

(c)

Diagnosis accuracy
IMF3 = 90.533%

(d)

Diagnosis accuracy
IMF4 = 72.667%

(e)

Diagnosis accuracy
IMF5 = 50.933%

(f)

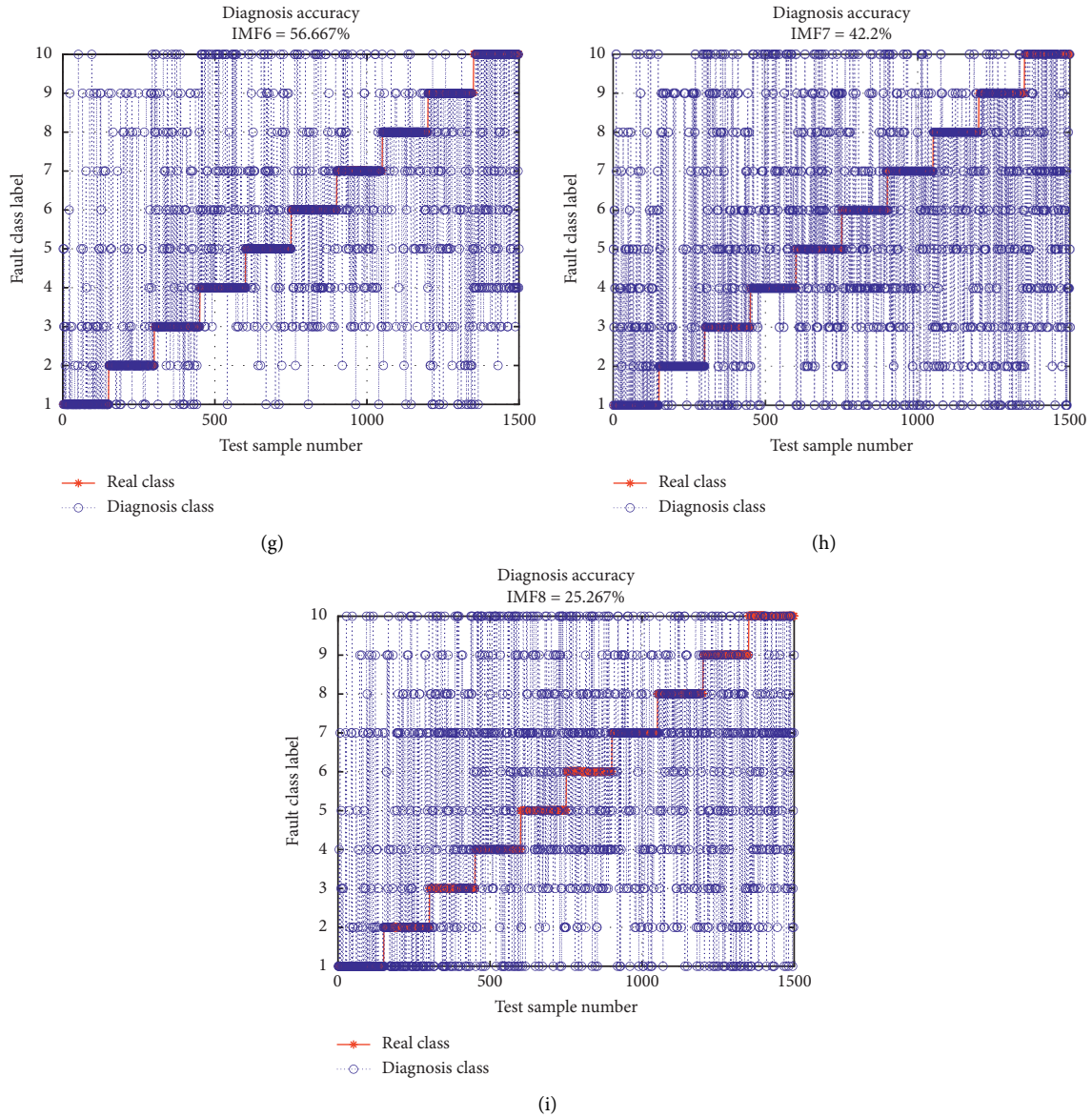Figure 5: Continued.

(g)

(h)

(i)

FIGURE 5: Diagnosis accuracy produced by individual DMLYu with original signal and IMF component. Diagnosis accuracy produced by individual DMLYu (a) with original signal, (b) with IMF1 component, (c) with with IMF2 component, (d) with IMF3 component, (e) with IMF4 component, (f) with IMF5 component, (g) with IMF6 component, (h) with IMF7 component, and (i) with IMF8 component.

At the same time, the diagnosis accuracy produced by the ensemble DMLYu model based on voting method with the first n IMF components against the number of IMF component can be seen from the curve plot in Figure 7. From the figure, it can be found that when the number of the DMLYu models used for ensemble is not less than 2 the diagnosis accuracy of the ensemble DMLYu model is higher than that of the individual DMLYu model with the vibration signal. And the ensemble DMLYu model with the first 3 IMF components can obtain the highest diagnosis accuracy which is up to 98.27%, but the diagnosis accuracy of ensemble DMLYu model is high and low with the number of DMLYu model used for ensemble increasing; this can demonstrate that the diagnosis accuracy of ensemble DMLYu model does not increase with the increase of the number of DMLYu models used for ensemble. In addition, from Figures 6 and 7, it can be seen that the diagnosis accuracies of ensemble DMLYu model based on BBM with first n IMF components are all higher than that of the ensemble DMLYu model based on voting method when the number of DMLYu models used for ensemble is the same. And the highest diagnosis accuracy of ensemble DMLYu model based on BBM with 8 IMF components is also higher than that of ensemble DMLYu model based on voting method with 3 IMF components which is higher than that of the individual DMLYu model with the original vibration signal. These can indicate that the diagnosis performance of these two kinds of ensemble DMLYu model based on different fusion strategy is better than that of the individual DMLYu model, and the
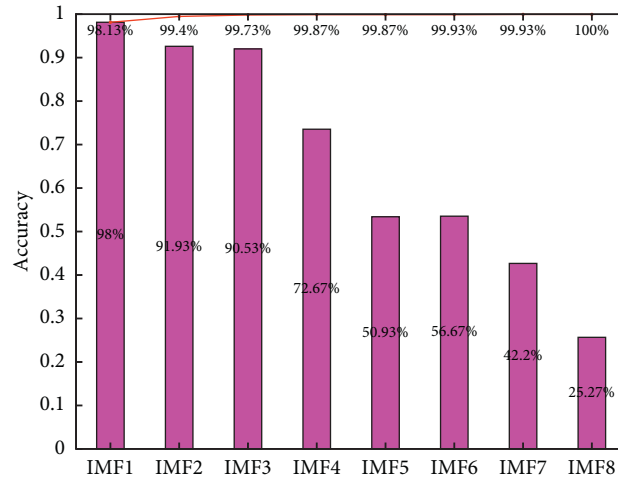
FIGURE 6: Diagnosis accuracy produced by ensemble DMLYu based on BBM with first *n* IMF components.
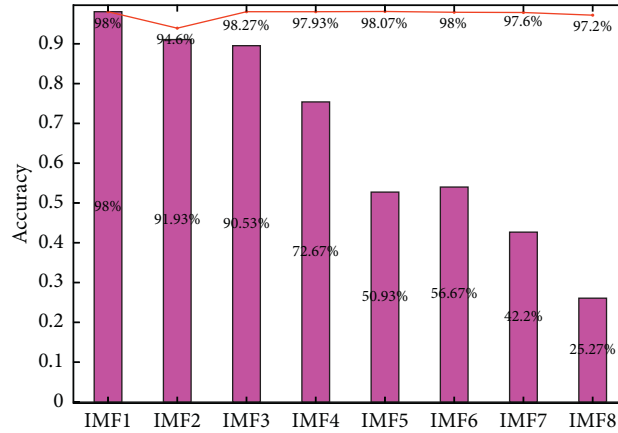

FIGURE 7: Diagnosis accuracy produced by ensemble DMLYu based on voting method with first *n* IMF components.

performance of the ensemble DMLYu model based on BBM is the best.

(2) Generalization Analysis of the Ensemble DMLYu. To study the stability and generalization of the proposed ensemble DMLYu model based on BBM, three bearing datasets under three working conditions (0 hp load and 1797 rpm, 1 hp load and 1772 rpm, and 2 hp load and 1750 rpm) are utilized, and for convenience these datasets are referred to as dataset 1, dataset 2, and dataset 3. Each dataset contains 10 fault classes, in each dataset the number of data samples for each fault class is 500, among them the number of training samples is 350, and the number of the test samples is 150. And each data sample has 512 sample points. In each dataset five diagnosis tests are conducted by the ensemble DMLYu model based on BBM, ensemble DMLYu model based on voting method, and individual DMLYu model, respectively. The average diagnosis accuracy produced by these three models is shown in Table 2.

Table 2 can display that the average diagnosis accuracy produced by ensemble DMLYu based on BBM with three datasets is 99.95%, 99.96%, and 99.96%, respectively, the

maximum difference between these three average diagnosis accuracies is 0.01%, and the minimum difference between them is 0. The average diagnosis accuracy produced by ensemble DMLYu based on voting method with three datasets is 98.24%, 98.69%, and 98.01%, the maximum difference between these three average diagnosis accuracies is 0.68%, and the minimum difference between them is 0.23%. The average diagnosis accuracy produced by individual DMLYu with three datasets is 93.57%, 93.28%, and 92.68%, the maximum difference between these three average diagnosis accuracies is 0.89%, and the minimum difference between them is 0.11%. Obviously the diagnosis accuracy of ensemble DMLYu based on BBM is the highest among these three methods on the same dataset; the diagnosis accuracy of ensemble DMLYu based on voting method is higher than that of individual DMLYu. The maximum difference and the minimum difference of the average diagnosis accuracy of DMLYu based on BBM with three datasets are all lower than those of DMLYu based on voting method and individual DMLYu model. These can indicate that diagnosis stability and generalization of the ensemble DMLYu based on BBM are stronger than those of

TABLE 2: Performance comparison between individual DMLYu and ensemble DMLYu.

Diagnosis model	Average diagnosis accuracy (%)		
	Dataset 1	Dataset 2	Dataset 3
DMLYu based on BBM	99.95	99.96	99.96
DMLYu based on voting method	98.24	98.69	98.01
DMLYu	93.57	93.28	92.68

the ensemble DMLYu based on voting method and individual DMLYu model and the diagnosis accuracy produced by the ensemble DMLYu based on BBM is the highest. All these can demonstrate that the ensemble DMLYu based on BBM can diagnose the fault of rolling bearings effectively with strong generalization and high accuracy.

4. Conclusion

In this paper, a novel ensemble DMLYu model based on BBM and EEMD is proposed and applied to the fault diagnosis of the rolling bearings. In order to solve the misdiagnosis problem of data sample in the overlapping region at the boundary of different fault classes and improve the diagnosis accuracy and robustness of deep metric learning model, by the EEMD method the original vibration data of rolling bearing is decomposed into multiple IMF components which are input into the deep metric learning model based on Yu's norm, respectively, then the initial fault results are obtained, respectively, and at last the final diagnosis decision is made by the BBM fusion strategy.

Through a multifaceted comparison of three methods on different experimental datasets, the effectiveness and generalization of the proposed ensemble DMLYu model based on BBM were verified by comparison with the ensemble DMLYu model based on voting method and individual DMLYu model. The diagnosis results have demonstrated that the proposed ensemble method was more effective and robust than other ensemble DMLYu models based on voting method and individual DMLYu model for fault diagnosis of rolling bearings under different working conditions and verified that the proposed ensemble method can diagnose the fault of rolling bearings with high accuracy and reliability. All these can show that the proposed ensemble DMLYu model will have a prosperous application prospect in the field of fault diagnosis. But in the process of fault diagnosis it is found that the number of DMLYu models used for ensemble can affect the diagnosis accuracy, so the selectiveness of the DMLYu models used for ensemble needs to be studied further in future.

Data Availability

The data used to support this study are available at the website http://csegroups.case.edu/bearingdatacenter/pages/download-data-file.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] A. Glowacz, "Ventilation diagnosis of angle grinder using thermal imaging," *Sensors*, vol. 21, no. 8, p. 2853, 2021.

[2] A. Glowacz, R. Tadeusiewicz, S. Legutko et al., "Fault diagnosis of angle grinders and electric impact drills using acoustic signals," *Applied Acoustics*, vol. 179, no. 5, Article ID 108070.

[3] X. Zhang, Z. Min, Z. Xiang, and J. Mao, "Research on diagnosis algorithm of mechanical equipment brake friction fault based on MCNN-SVM," *Measurement*, vol. 186, Article ID 110065, 2021.

[4] D. H. Pandya, S. H. Upadhyay, and S. P. Harsha, "Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN," *Expert Systems with Applications*, vol. 40, no. 10, pp. 4137–4145, 2013.

[5] Y. Lei, F. Jia, W. Zhou, and J. Lin, "A deep learning-based method for machinery health monitoring with big data," *Journal of Mechanical Engineering*, vol. 51, no. 21, pp. 49–56, 2015.

[6] X. Su, Y.-y. Wu, H.-J. Pei, X. Y. Lan, and J. S. Gao, "Progress in the application of big data technology in process industry," *Chemical Industry Progress*, vol. 35, no. 6, pp. 1652–1659, 2016.

[7] H. Zhao, H. Liu, Y. Jin, X. Dang, and W. Deng, "Feature extraction for data-driven remaining useful life prediction of rolling bearings," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[8] W. Deng, S. Shang, X. Cai et al., "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," *Knowledge-Based Systems*, vol. 224, Article ID 107080, 2021.

[9] W. Deng, J. Xu, X.-Z. Gao, and H. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–10, 2020.

[10] K. Anil, V. Govind, C. P. Gandhi, Z. Yuqing, and G. Adam, "Novel convolutional neural network (NCNN) for the diagnosis of bearing defects in rotary machinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, Article ID 3510710, 2021.

[11] G. Niu, X. Wang, M. Golda, S. Mastro, and B. Zhang, "An optimized adaptive PReLU-DBN for rolling element bearing fault diagnosis," *Neurocomputing*, vol. 445, pp. 26–34, 2021.

[12] T. Jin, C. Yan, C. Chen, Z. Yang, H. Tian, and S. Wang, "Light neural network with fewer parameters based on CNN for fault diagnosis of rotating machinery," *Measurement*, vol. 181, Article ID 109639, 2021.

[13] X. Kong, G. Mao, Q. Wang, H. Ma, and W. Yang, "A multi-ensemble method based on deep auto-encoders for fault diagnosis of rolling bearings," *Measurement*, vol. 151, no. 151, Article ID 107132, 2020.

[14] H. Shi, Y. Yang, X. Zhu et al., "Embedding deep metric for person Re-identification: a study against large variations," in *Proceedings of the European Conference on Computer Vision*, Spinger, Amsterdam, The Netherlands, October 2016.

[15] B. Wu, Z. Chen, J. Wang, and H. Wu, "Exponential discriminative metric embedding in deep learning," *Neurocomputing*, vol. 290, 2018.

[16] W. Yang, Y. Yan, and S. Chen, "Adaptive deep metric embeddings for person re-identification under occlusions," *Neurocomputing*, vol. 340, pp. 125–132, 2019.

[17] J. Hu, J. Lu, and Y. P. Tan, "Discriminative deep metric learning for face verification in the wild," in *Proceedings of the Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, IEEE, Columbus, OH, USA, June 2014.

[18] X. Li, Z. Xu, and W. Xiong, "Bearing fault diagnosis method based on deep metric learning," *Shock and Vibration*, vol. 39, no. 15, pp. 25–31, 2020.

[19] Z. Xu, X. Li, H. Lin, Z. Wang, and T. Peng, "fault diagnosis of rolling bearing based on modified deep metric learning method," *Shock and Vibration*, vol. 2021, Article ID 6635008, 11 pages, 2021.

[20] M. Zhao, M. Kang, B. Tang, and M. Pecht, "Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4290–4300, 2018.

[21] K. Chen, X.-C. Zhou, J.-Q. Fang, P.-f. Zheng, and J. Wang, "Fault feature extraction and diagnosis of gearbox based on EEMD and deep briefs network," *International Journal of Rotating Machinery*, vol. 2017, Article ID 9602650, 10 pages, 2017.

[22] C. F. Tsai and J. W. Wu, "Using neural network ensembles for bankruptcy prediction and credit scoring," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2639–2649, 2008.

[23] H. Shao, H. Jiang, Y. Lin, and X. Li, "A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders," *Mechanical Systems and Signal Processing*, vol. 102, pp. 278–297, 2018.

[24] Y. Liu, X. Yan, C.-a. Zhang, and W. Liu, "An ensemble convolutional neural networks for bearing fault diagnosis using multi-sensor data," *Sensors*, vol. 19, no. 23, p. 5300, 2019.

[25] Y. Han, B. Tang, and L. Deng, "Multi-level wavelet packet fusion in dynamic ensemble convolutional neural network for fault diagnosis," *Measurement*, vol. 127, pp. 246–255, 2018.

[26] J. Yachao, L. Guolong, D. Xin, and H. Kun, "A novel denoising method for vibration signal of hob spindle based on EEMD and grey theory," *Measurement*, vol. 169, Article ID 108490, 2021.

[27] J. Lu, J. Hu, and J. Zhou, "Deep metric learning for visual understanding: an overview of recent advances," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 76–84, 2017.

[28] L. Lam and C. Y. Suen, "Optimal combinations of pattern classifiers," *Pattern Recognition Letters*, vol. 16, no. 9, pp. 945–954, 1995.

[29] The Case Western Reserve University Bearing Data Center, "Bearing data center fault test data," 1998, http://csegroups. case.edu/bearingdatacenter/pages/download-data-file.