

Research Article

Based on Soft Competition ART Neural Network Ensemble and Its Application to the Fault Diagnosis of Bearing

**Dan Yang,^{1,2} Hailin Mu,^{1,2} Zengbing Xu,^{1,2,3} Zhigang Wang,^{1,2}
Cancan Yi,^{1,2} and Changming Liu^{1,2,4}**

¹Key Laboratory of Metallurgical Equipment and Control Technology, Wuhan University of Science and Technology, Ministry of Education, Wuhan 430081, China

²Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

³The State Key Laboratory of Digital Manufacturing Equipment & Technology, Huazhong University of Science and Technology, Wuhan 430074, China

⁴The State Key Laboratory of Refractories and Metallurgy, Wuhan University of Science and Technology, Wuhan 430081, China

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

Received 27 March 2017; Accepted 15 May 2017; Published 3 July 2017

Academic Editor: Marek Lefik

Copyright © 2017 Dan Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a novel method for fault diagnosis based on an improved adaptive resonance theory (ART) neural network and ensemble technique. The method consists of three stages. Firstly, the improved ART neural network is comprised of the soft competition technique based on fuzzy competitive learning (FCL) and ART based on Yu's norm, the neural nodes in the competition layer are trained according to the degree of membership between the mode node and the input, and then fault samples are classified in turn. Secondly, with the distance evaluation technique, the optimal features are obtained from the statistical characteristics of original signals and wavelet coefficients. Finally, the optimal features are input into the neural network ensemble (NNE) based on voting method to identify the different fault categories. The proposed method is applied to the fault diagnosis of rolling element bearings, and testing results show that the neural network ensemble can reliably classify different fault categories and the degree of faults, which has a better classification performance compared with the single neural network.

1. Introduction

Presently, the fault diagnosis is increasingly intelligent because of wide applications capability of artificial neural networks (ANNs). However, some ANNs are unable to detect new fault category in the diagnosis process. In other words, the trained ANN can only identify the faults which are learned at the stage of training. If a new fault category occurs, the training data sample of the data set used to train networks need be added, and these networks are required to retrain and learn the knowledge using the complete data sets. This can result in a time-consuming and costly process [1]. And, in the real world, nobody knows what will happen next time, so it is impossible to obtain the complete training data set which contain all data samples of all fault categories.

These characteristics limit the application of these neural networks in fault diagnosis field. So we want an intelligent system which is able to adapt well to unexpected changes in the environment, the system should be able to deal with the so-called stability-plasticity dilemma. That is, the system should be designed to be stable enough to protect useful historical data from the corruption and have some degree of plasticity to learn new data [2, 3]. As a solution to this problem, the adaptive resonance theory (ART) networks were developed and have been applied to the field of pattern recognition and fault diagnosis [4]. Because of the superiority of ART network, some ART models such as ART2 and fuzzy ART have been developed and applied to the field of fault diagnosis [5–8]. Furthermore, some classification methods with the improvement of ART, such as the combination of

ART and Yu's norm, have been applied well to the field of fault diagnosis, which can deal with the uncertainty problem through the fuzzy formalism [9].

Although the classification method with the combination of ART and Yu's norm has a better performance in the field of fault diagnosis, there still exist two defects [9]. One is that the neurons are incompletely used. Some neurons in competition layer are too far away from samples to win. This is why "dead neurons" are created. The other problem is that only one neuron in competition layer can win with winner-takes-all mechanism. This mechanism can reduce the convergent speed of network and result in waste of more storage information between input neurons and competition ones.

This study proposes a novel competitive network—an adaptive resonance theory based on soft competition (SYART), whose topological structure is the same as Yu's norm based ART. In SYART, soft competition mechanism is introduced in competition layer instead of winner-takes-all. When the similarity between the input sample and the neuron is larger than the preset threshold, the neuron wins, and the network is adjusted according to the membership degrees between the winning neurons and the input sample. So there is more than one winning neuron in the network. This mechanism makes full use of competition neurons and makes the network model optimization.

However, because of the nonlinearity of fault signals and the diversification of fault categories [10, 11], it is difficult to maintain a high accuracy with the single SYART. In order to solve this problem, a novel method for fault diagnosis based on SYART and NNE is proposed. The extracted features from the statistical characteristics of raw signals and wavelet coefficients are selected with the distance evaluation technique, then different features are used to train each SYART as the input of the network, and the final result of the fault diagnosis is decided by the NNE with majority voting [12–14].

The proposed fault diagnosis method is applied to the fault diagnosis of rolling element bearing. The testing results show the effectiveness of SYART. And a better effect has been obtained using NNE. The rest of the paper is organized as follows. Section 2 reviews the adaptive resonance theory and Yu's norm ART. An improved method for fault diagnosis—Yu's norm ART based on soft competition is proposed in Section 3. Section 4 describes the neural network ensemble based on majority voting. Finally, the process and the conclusions about the proposed clustering method in the field of fault diagnosis are given in Section 5.

2. Review of ART and Yu's Norm Based ART

Adaptive resonance theory (ART) was designed by Carpenter and Grossberg [5]. And in order to meet the different application needs some ART network models are developed, the popular one of which is the ART1. As an online learning system, the ART1 network model can cluster inputs by using the unsupervised learning [6]. It can self-organize stable recognition codes in real time in response to arbitrary sequences of input patterns, and it uses the vigilance test

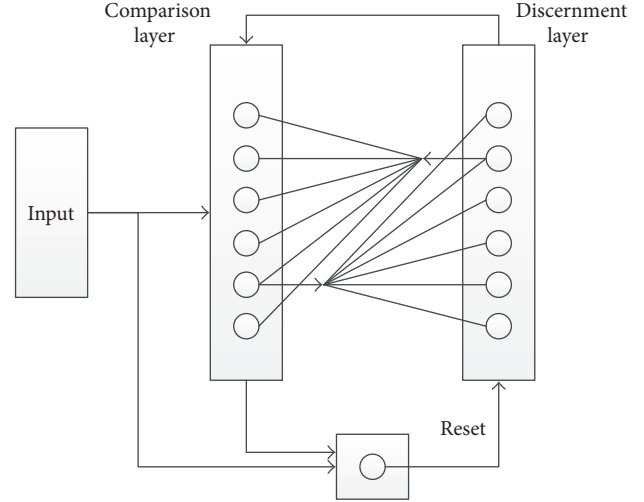


FIGURE 1: The architecture of Yu's norm ART.

to resolve the stability-plasticity dilemma. In the learning process, when a new sample is input into the ART network, it can be categorized by comparing with the weight vectors of existing cluster nodes. If the sample and the existing categories can be matched and the match degree is not less than the vigilance value, the sample will be classified into the matched category and the weight vector of the corresponding category will be updated. Otherwise, a new category will be created with retaining the existing memory.

With the development of clustering methods using similarity measure based on fuzzy set theory, a similarity measure based on Yu's norm was set up and the corresponding clustering method was proposed by Luukka in 2007 [15, 16]. Yu's norm is composed of Yu's T -norm and Yu's S -norm, which can be written as follows:

$$\begin{aligned} T(x, y) &= \max[0, (1 + \tau)(x + y - 1) - \tau xy], \\ S_n(x, y) &= \min[1, x + y + \tau xy], \end{aligned} \quad (1)$$

in which $\tau > -1$, and the similarity measure can be derived as follows:

$$S(x, y) = T(S_n(\bar{x}, y), S_n(x, \bar{y})). \quad (2)$$

Considering the complementary characteristics of ART and Yu's norm in clustering, a clustering method combining ART with Yu's norm for fault diagnosis was presented by Xu et al. in 2016 [9]. The architecture of Yu's norm ART is shown in Figure 1. Its architecture is similar to ART's, excluding the adaptive filter. The composition of Yu's norm ART is input layer which stores the input samples, comparison layer which receives bottom-up input from input layer and top-down input from discernment layer, and discernment layer which contains the active category and stores category nodes. For each input sample, the categorization with the clustering method of Yu's norm ART is performed by category choice, vigilance test, and learning.

(1) *Category Choice.* In this stage, one winner category node is selected from all existing category nodes with hard

competition mechanism. The choice function is a similarity measure between the j th category and the i th input sample, which can be written as

$$S \langle Y_i, V_j \rangle = \max \left(0, (1 + \tau) \left(Sn(\bar{Y}_i, V_j) + Sn(Y_i, \bar{V}_j) - 1 - \tau Sn(\bar{Y}_i, V_j) Sn(Y_i, \bar{V}_j) \right) \right), \quad (3)$$

where Y_i is the i th input sample and V_j is the weight vector of the j th category node. Having the biggest similarity degree, the winner category J is selected as follows:

$$J = \operatorname{argmax} \left(S(Y_i, V_j) \right). \quad (4)$$

(2) *Vigilance Test*. As a vigilance parameter to test the similarity degree to which Y_i is a subset of category V_j , ρ is introduced to determine whether the i th input sample matches the selected winner category J . If the similarity degree meets the vigilance criterion

$$S \langle Y_i, V_j \rangle \geq \rho, \quad (5)$$

where $\rho \in [0, 1]$, which means that the input sample Y_i is sufficiently similar to the winner category V_j , then the sample Y_i will be classified into the J th category, and the learning is performed at the same time. Otherwise, a new category node will be generated in the discernment to contain the input sample; and the corresponding weight vector of the category is given by the following formula:

$$V_{n+1} = Y_i, \quad (6)$$

where n is total number of current category nodes.

(3) *Learning*. When the vigilance criterion is satisfied by the winner category, the corresponding weight vector of the category will be updated by the following formula:

$$V'_J = \frac{(n \times V_J + Y_i)}{(n + 1)}, \quad (7)$$

where V'_J is the updated weight vector of category J , V_J is the original weight vector of category J , and n is the number of samples which belong to the category J .

3. ART Based on Soft Competition and Yu's Norm

In the competition layer of Yu's norm ART, the only one winning neuron is selected with hard competition mechanism, which makes neurons in competition layer used insufficiently. In order to solve the problem, a Yu's norm ART network based on soft competition (SYART) is proposed. Based on the membership degree formula of FCM, fuzzy competition learning (FCL) is introduced as soft competition mechanism, which has the advantages of flexible adjustment of soft competition and noise suppression [17]. Under soft competition mechanism, when the similarity between the input sample and the neuron is larger than the preset threshold, the

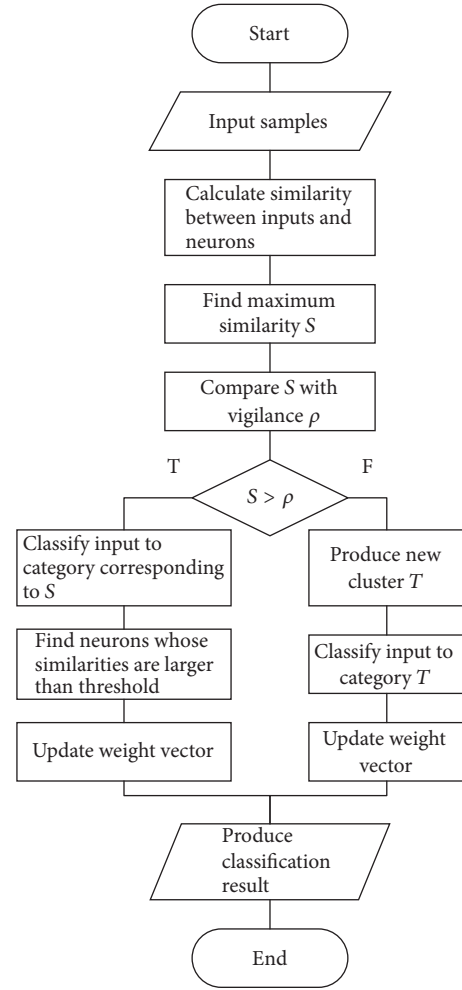


FIGURE 2: Flow chart of Yu's norm ART based on soft competition.

neuron wins; then the network is adjusted according to the membership degrees between the winning neurons and the input sample. So there is more than one winning neuron in the network.

The detailed fault diagnosis flow chart of SYART is shown in Figure 2. Just as Yu's norm ART, the similarity between the input sample and the neuron is calculated with formula (3), then get maximum similarity S , and compare S with vigilance parameter ρ , and if $S < \rho$, a new category node will be generated to contain the input sample. Otherwise, the input sample will be classified into the category corresponding to the maximum similarity S . And, under soft competition mechanism, the corresponding weight vectors of winning neurons whose similarities with the input sample are larger than the preset threshold are updated by the following formula:

$$V' = \frac{(n \times V + hi^m \times Y_i)}{(n + hi^m)}, \quad (8)$$

where V' is the updated weight vector of winning neuron, V is the original weight vector of winning neuron, n is the number of samples which belong to the winning neuron, and hi is

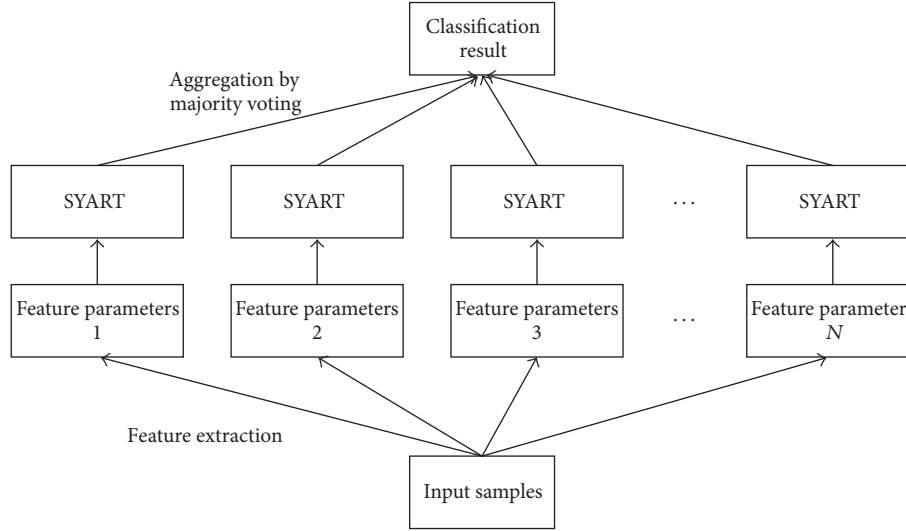


FIGURE 3: The proposed architecture of the neural network ensemble.

the membership degree between the winning neuron and the input sample, which is calculated by the following formula:

$$hi = \left(\frac{\|Y_i - V_j\|^2}{\sum_{k=1}^c \|Y_i - V_k\|^2} \right)^{-1/(m-1)}, \quad (9)$$

where c is the number of winning neurons and V_j is the corresponding weight vector of current neuron. As the fuzzy exponent, m makes the level of soft competition adjustable. When $m \rightarrow 1$, hi is 0 or 1, which makes soft competition degenerate into hard competition. When $m \rightarrow \infty$, the value of hi is $1/c$, which is the most fuzzy and meaningless. And because of making lower membership noise points have less contribution to the network, the larger m is able to restrain noise.

4. Neural Network Ensemble Based on Majority Voting

Recently, instead of using a single NN, combined NN for increasing classification accuracy is an active research area, which has been widely used in the field of data mining, pattern recognition, fault diagnosis, and so forth. Generally, neural network ensemble (NNE) can be considered as two-step processes, first training component networks and then combining them. As is shown in Figure 3, during the first step, a number of classification results are produced by using the same algorithm with different feature parameters. In the next step, all classification results are combined with ensemble technique to get the final classification result.

As a simple and intuitive ensemble technique, majority voting is most frequently used, which can be described as follows. Define the result of the m th clustering H_m as $h_{m,i} \in \{0, 1\}$ ($m = 1, 2, 3, \dots, M$, $i = 1, 2, 3, \dots, C$), where M is the number of based clustering results and C is the number of clusters. If the chosen m th clustering is cluster i , then $h_{m,i} = 1$;

otherwise $h_{m,i} = 0$. The result of majority voting as ensemble decision is cluster k if

$$\sum_{m=1}^M h_{m,k} = \max_i \sum_{m=1}^M h_{m,i}. \quad (10)$$

However, clustering results can be represented by number labels (1, 2, 3, etc.) or alphabet labels (A, B, C, etc.). Furthermore, the same label may represent different clusters, and the same cluster may be labeled differently under the same type of labeling scheme, which can lead to a bad performance in clustering with the majority voting as ensemble technique in this paper. In order to get rid of the problem, the clusters labels need to be unified; that is, the same label must represent the same cluster. Then a method of label unification based on maximum matching is proposed, which can be described as follows. Select two results of the partitions, standards of relabeling $C_r = \{1, 1, 2, 2, 2, 3, 3\}$ and relabeled partition $C_t = \{2, 3, 2, 2, 3, 1, 3\}$, and then the contingency matrix can be found as follows:

		C_r			
		1	2	3	
1		0	0	1	(11)
C_t	2	1	2	0	
	3	1	1	1	

Define a_{ij} as the value of contingency matrix, which means the number of objects with label i in the partition C_t and label j in the partition C_r . Cluster i and cluster j have more probability to be the same cluster with larger a_{ij} . Considering the row values may be all the same, which makes it difficult to get a larger a_{ij} , the contingency matrix shown in (11) can be updated by the following formula:

$$b_{i,j} = \frac{a_{i,j}^2}{\sum_{m=1}^k a_{im} \sum_{m=1}^k a_{mj}}, \quad (12)$$

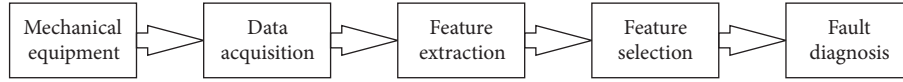


FIGURE 4: Architecture of fault diagnosis system.

where a_{ij} is the value in (11) and b_{ij} is the new value shown as follows:

$$C_t = \begin{matrix} & C_r \\ & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{6} & \frac{4}{9} & 0 \\ \frac{1}{6} & \frac{1}{9} & \frac{1}{6} \end{bmatrix} \end{matrix} \quad (13)$$

Find the larger b_{ij} of each row in the contingency matrix shown in (11), and then we can relabel C_t by the relabeling: 1 to 3, 2 to 2, and 3 to 1. The partition $C_t = \{2, 3, 2, 2, 3, 1, 3\}$ becomes $C'_t = \{2, 1, 2, 2, 1, 3, 1\}$, whose label unification is completed.

5. Diagnosis System Using Soft Competition Yu's Norm ART

5.1. Structure of Diagnosis System. The fault diagnosis system is shown in Figure 4. The system mainly includes four stages: data acquisition, feature extraction, feature selection, and fault diagnosis. In this paper, the data set of the rolling element bearings are used as fault signals to ensure the credibility of diagnosis results [18]. Then time-domain and frequency-domain features are extracted with modified distance discriminant technique from the fault signals after wavelet transform, which is more effective than FFT. And then the proposed soft competition Yu's norm ART is trained and used to diagnose the faults of bearings. Finally, the diagnosis results obtained with different sets of features are combined by the NNE with majority voting to identify the final result of the fault diagnosis.

5.2. Data Acquisition. As is shown in Figure 5, the experiment rig consists of a three-phase induction motor, a load motor, and a torque sensor [18]. The ball bearings whose types are 6205-2RS JEM SKF are installed in a motor driven mechanical system. The three-phase induction motor is connected to a load motor to get the desired torque load levels. An accelerometer is mounted on the driven end of the motor housing.

Faults are introduced into the drive-end bearing of the motor by the electrodischarge machining. Different defect sizes (0.007 and 0.021 inches) are introduced to simulate different fault severities of bearings, and each defect size is introduced into the outer race and inner race to simulate the different fault categories of bearings. Thus, six different fault data sets can be obtained from the experimental system: (1) in normal condition; (2) with ball fault; (3) with two different

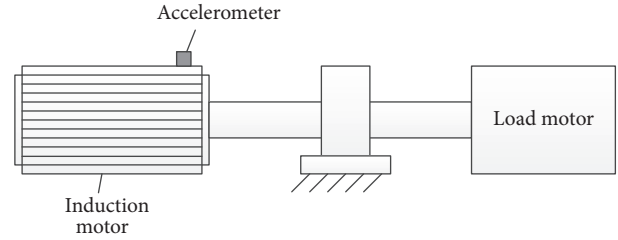


FIGURE 5: The schematic diagram of the experimental setup.

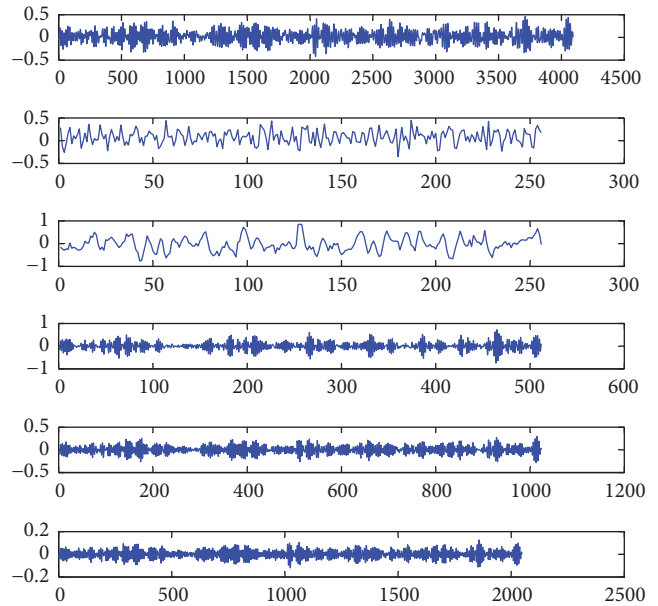


FIGURE 6: Wavelet decomposition of fault signal.

severity levels of inner race faults; (4) with two different severity levels of outer race faults.

5.3. Features Extraction. As the important parameter to describe the change of the vibration signal, feature parameters are usually used to depict the fault-related information about the bearings. In order to acquire more accurate information, wavelet decomposition with Haar wavelet is used to analyze vibration signals, which are decomposed into four levels. The decomposition results of one sample signal can be shown in Figure 6. And then feature parameters are extracted to describe the bearing fault from each layer of wavelet signals.

In this paper, six machines are used for fault diagnosis to accomplish the ensemble of clustering results; as the input of the NN, six sets of feature parameters which can be shown in Table 1 are extracted from the same vibration signal.

TABLE 1: Six sets of feature parameters.

NN	NN1	NN2	NN3	NN4	NN5	NN6
	Mean square	Mean square	Root mean square	Root mean square	Root mean square	Root mean square
Feature parameter	Root mean square	Root mean square	Kurtosis	Kurtosis	Peak	Peak factor
	Kurtosis	Margin	Variance	Waveform index	Peak factor	Impulse factor
	Variance	Energy entropy	Peak	Energy entropy	Energy entropy	Margin

5.4. *Features Selection.* As shown in Table 1, a total of 20 feature parameters are used as the input of each machine, which means the reduction of diagnosis accuracy and the increase of computation time because of the redundancy or irrelevance of some features. In order to maintain a better diagnosis performance, the selection of sensitive features which provide characteristic information for the diagnosis system needs to be done, and irrelevant or redundant features must be removed. Here, a new parameter evaluation technique based on distance discriminant technique [19] is proposed, which includes the similarity formula (3). The feature selection process can be described as follows.

Step 1. Extract all 20 feature parameters, and the parameter set can be described as follows:

$$\left\{ S_{n,f,p}, n = 1, 2, \dots, 100; f = 1, 2, \dots, 6; p = 1, 2, \dots, 20 \right\}, \quad (14)$$

where $S_{n,f,p}$ is the value of the feature parameter p of the sample n in the f type of fault.

Step 2. For each feature parameter p , calculate the within-class distance in each mode:

$$d_{p,f} = \frac{1}{100 \times 99} \sum_{n_i, n_j=1}^{100} S(S_{n_i, f, p}, S_{n_j, f, p}), \quad (15)$$

$$f = 1, 2, \dots, 6; n_i, n_j = 1, 2, \dots, 100, n_i \neq n_j,$$

where $S(S_{n_i, f, p}, S_{n_j, f, p})$ is the similarity which is calculated with formula (3) between the feature parameters $S_{n_i, f, p}$ and $S_{n_j, f, p}$. Then calculate the average within-class distance of the six types of fault:

$$d_p^{\text{in}} = \frac{1}{6} \sum_{f=1}^6 d_{p,f}. \quad (16)$$

Step 3. Calculate the average of the feature parameter p in the same type of fault as the center of within-class:

$$g_{f,p} = \frac{1}{100} \sum_{n=1}^{100} S_{n,f,p}. \quad (17)$$

Then calculate the average within-class distance:

$$d_p^{\text{out}} = \frac{1}{6 \times 5} \sum_{f_i, f_j=1}^6 S(g_{f_i, p}, g_{f_j, p}), \quad (18)$$

$$f_i, f_j = 1, 2, \dots, 6, f_i \neq f_j,$$

where $S(g_{f_i, p}, g_{f_j, p})$ is the similarity which is calculated with formula (3) of the center of within-classes $g_{f_i, p}$ and $g_{f_j, p}$.

Step 4. Calculate the ratio of the average within-class distance d_p^{out} to the average within-class distance d_p^{in} :

$$e_p = \frac{d_p^{\text{out}}}{d_p^{\text{in}}}. \quad (19)$$

Step 5. The sensitivity coefficient of each feature parameter can be defined and calculated as follows:

$$\lambda_p = \frac{e_p}{E_p}, \quad (20)$$

where E_p is the maximum of e_p .

The sensitivity coefficients of each feature parameter can be obtained, respectively, with this algorithm, and the feature parameter whose sensitivity coefficient is larger than the threshold value β is the sensitive feature parameter. The result of features selection of a model can be shown in Figure 7.

5.5. *Fault Diagnosis.* Some data samples of bearings are utilized to evaluate the performance of the proposed method in the phase of fault diagnosis, and the data samples contain six different fault conditions, which are labeled, respectively, by Arabic numerals (1, 2, 3, 4, 5, and 6). Each fault condition contains 100 data samples, and each data sample contains 4096 sample points. The detailed description is shown in Table 2.

The sensitive feature parameters are chosen as the input of the fault diagnosis model. The model is characterized by completing classification with the completion network training at the same time. When the first data sample is put into the blank network model, the first cluster node is built as one fault category by the sample and the node weight is the sample itself. When the next input sample enters the model, the similarity between the second input sample and the first cluster node of the model is calculated and then compared with the vigilance parameter; if the similarity is larger than the vigilance parameter, the input sample is classified into the first cluster, and the corresponding weight vector of the cluster node is updated by (8); otherwise, the second cluster node is produced. When the third input sample enters the model, it is compared with all the produced cluster nodes. If the biggest similarity degree meets the vigilance criterion, the sample is classified into the winner cluster, or else a new cluster node is produced. And the winners are the cluster

TABLE 2: Statistics of each fault condition of bearing.

Condition	Normal	Ball fault	Inner race fault	Outer race fault	Inner race fault	Outer race fault
Defect diameter	0	0.007	0.007	0.007	0.021	0.021
Label	1	2	3	4	5	6
Number of data samples	100	100	100	100	100	100

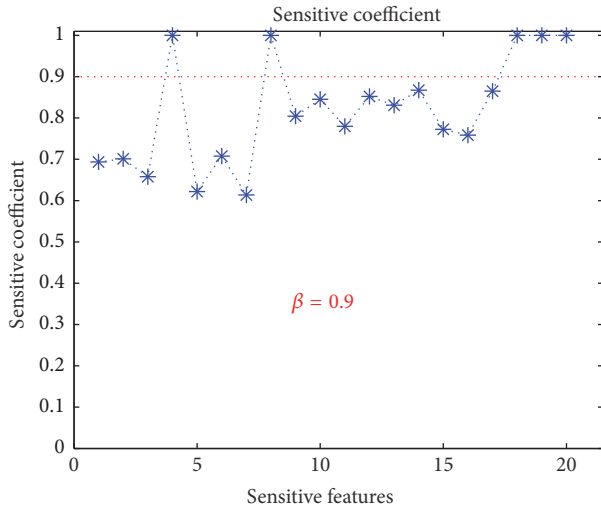


FIGURE 7: The result of features selection.

nodes whose similarity degrees are larger than the threshold value α . According to the above-mentioned reasoning, a trained classifier is obtained.

In the process of diagnosis, the threshold value β is set according to the ascending order to complete the feature selection and the condition of termination is the highest diagnostic accuracy, and then threshold value is 0.9. And each parameter left in the model is set in the same way without changing the other parameters, then the vigilance parameter ρ is 0.835, the parameter τ is 0.4, the fuzzy exponent m is 6, and threshold value α is 0.4. The diagnostic results of the six models can be shown in Figures 8, 9, 10, 11, 12, and 13, respectively.

For convenience of computation, the classification accuracy can be obtained by the following formula:

$$DA = \frac{C}{(T - N)} \times 100\%, \quad (21)$$

where DA is classification accuracy, C is the sample number of correct classification, T is the number of total samples, and N is the number of generated cluster nodes [18]. The classification accuracy of each model can be shown as Table 3.

As in Table 3, the classification accuracy of model 1 reaches 84.17%, which means that to a certain extent the fault diagnosis method by Yu's norm ART based on soft competition can not only distinguish different faults but also distinguish different fault degrees under the same fault type. However, the classification accuracy of model 3 is lower than 74%, which shows that different feature parameters lead to different diagnostic accuracy; in other words, it is difficult to maintain a high accuracy with the single model.

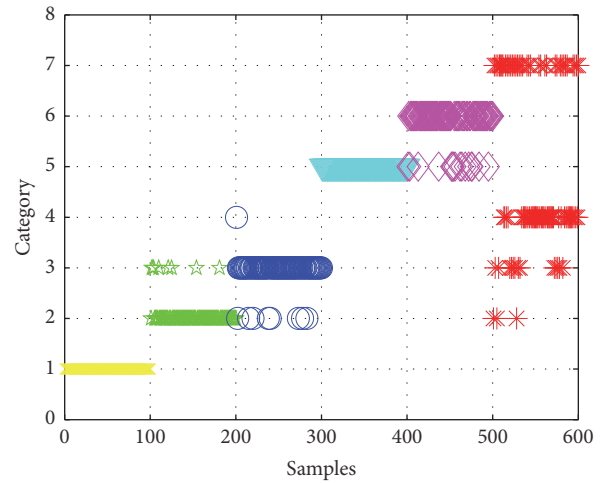


FIGURE 8: The diagnostic result of model 1.

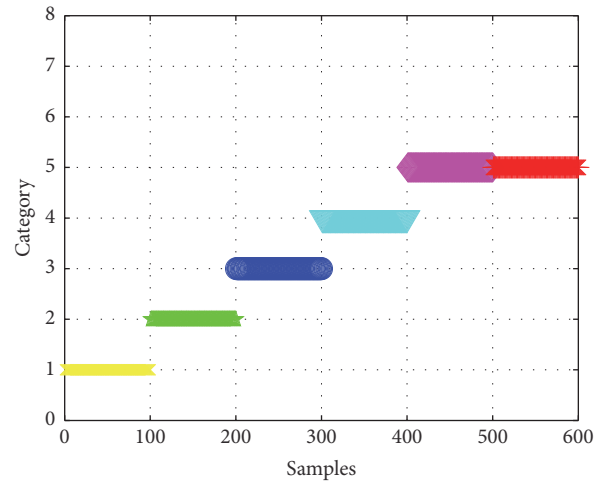


FIGURE 9: The diagnostic result of model 2.

After the diagnosis of each model, the final diagnostic result can be obtained by the neural network ensemble. Here, three models, four models, five models, and six models are chosen to ensemble, and the results are shown in Figures 14, 15, 16, and 17.

The classification accuracy of each type of ensemble can be shown in Table 4. As is shown Table 4, the classification accuracy of the ensemble of six models reaches 96.33%, and even if they are only three models, the classification accuracy of their ensemble is higher than 92%, which means that the ensemble can improve the performance of fault diagnosis of models significantly. Besides, with the increase of the number

TABLE 3: The classification accuracy of each model.

Model	Error points	Classification accuracy
Model 1	95	84.17%
Model 2	100	83.33%
Model 3	158	73.67%
Model 4	133	77.83%
Model 5	141	76.5%
Model 6	127	78.83%

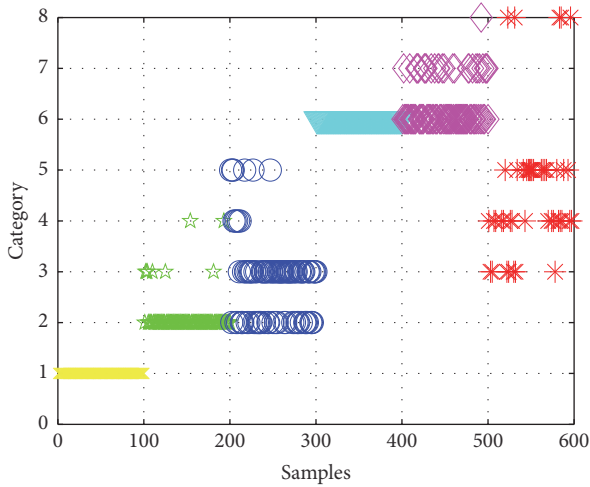


FIGURE 10: The diagnostic result of model 3.

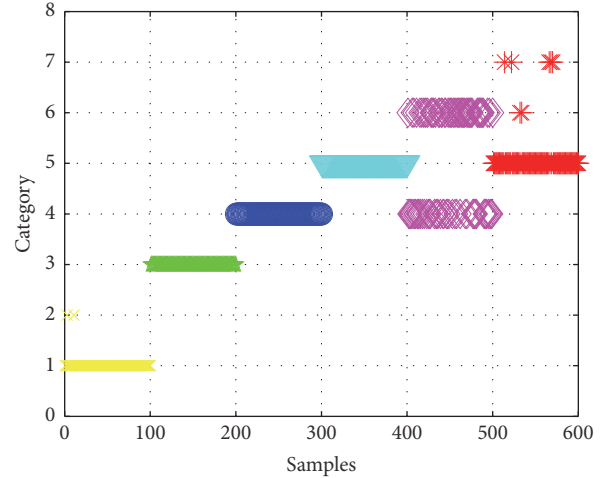


FIGURE 12: The diagnostic result of model 5.

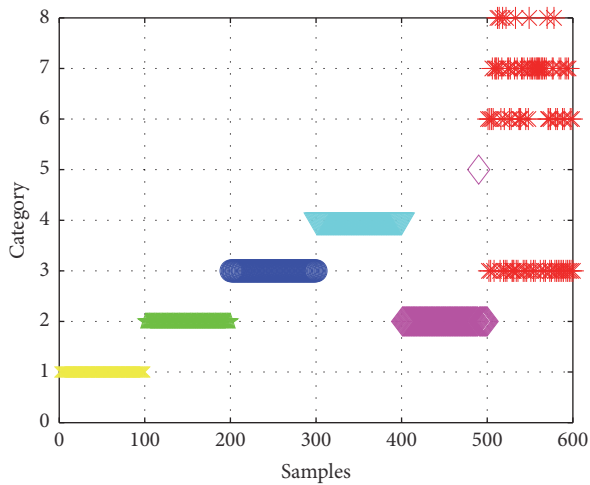


FIGURE 11: The diagnostic result of model 4.

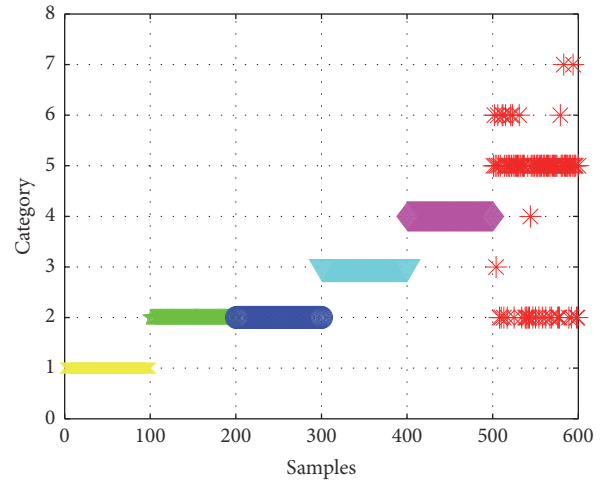


FIGURE 13: The diagnostic result of model 6.

of models involved in ensemble, the classification accuracy of fault diagnosis method is improved steadily.

5.6. Performance Test of Ensemble-SYART Clustering Method Based on Bootstrap Method. As we all know, the performance of ART clustering method can be affected by the initial conditions, which is the same as the ensemble-SYART clustering method. To study the stability and generalization of

the proposed method, the bootstrap method [20] is used to compute the estimated mean, standard deviation, and confidence interval for the classification accuracy, which is useful for estimating a parameter when the underlying distribution function of parameter is unknown. 15 bootstrap samples are generated by disrupting data order for statistical analysis of diagnosis accuracy. After calculation, the estimated statistical mean of diagnosis result is 91.52%, the standard deviation is 5.56%, and the 95% confidence interval achieves [88.44%,

TABLE 4: The classification accuracy of ensemble.

Type of ensemble	Error points	Classification accuracy
Three models	45	92.5%
Four models	30	95%
Five models	24	96%
Six models	22	96.33%

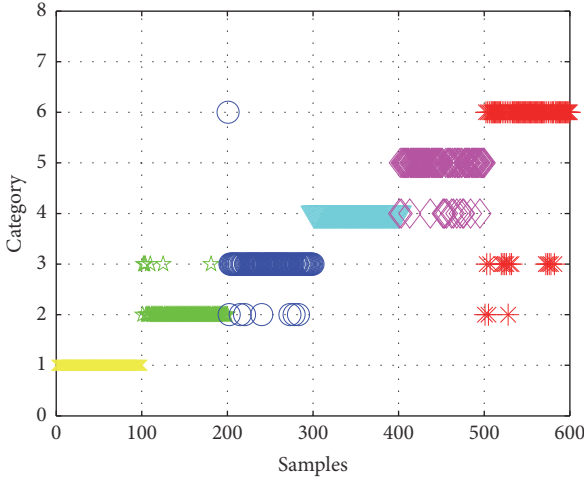


FIGURE 14: The ensemble result of three models.

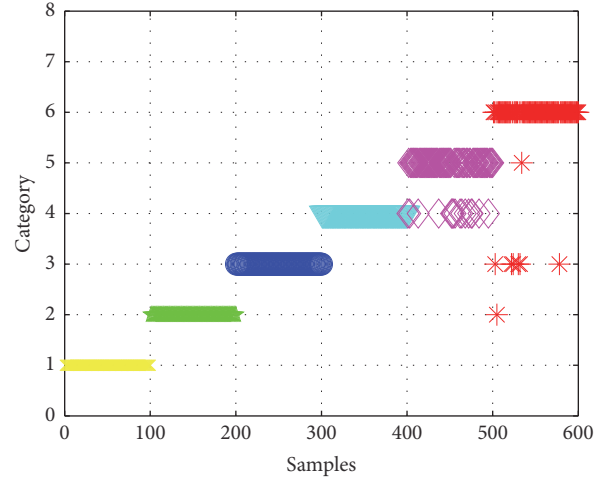


FIGURE 16: The ensemble result of five models.

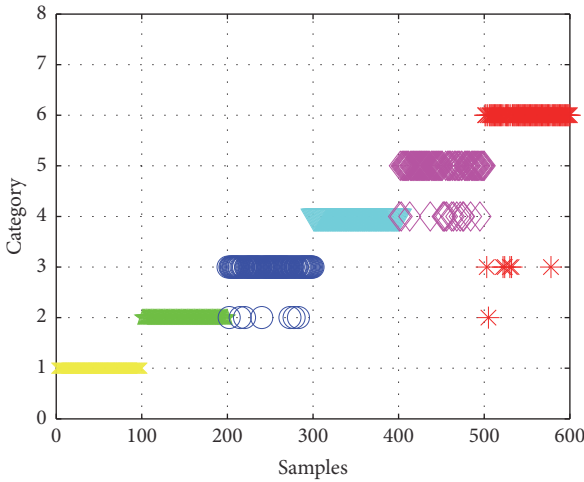


FIGURE 15: The ensemble result of four models.

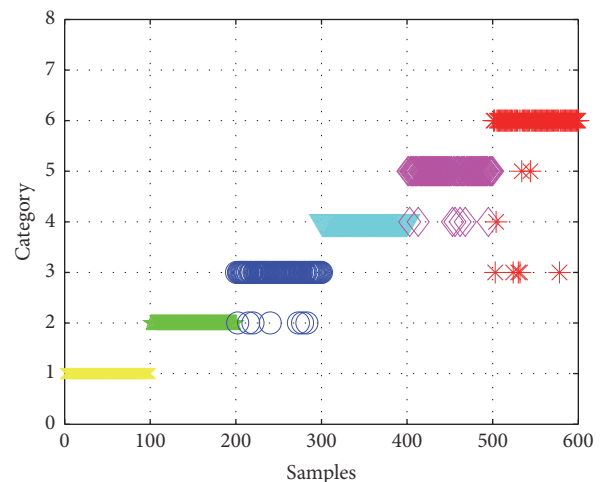


FIGURE 17: The ensemble result of six models.

94.60%]. These all can indicate that the performance of the proposed ensemble-SYART clustering method is stable and generalized.

5.7. Classification Performance Comparison with Other Methods. In order to validate the superiority of the proposed ensemble-SYART clustering method, the classification result produced by ensemble-SYART classifier is compared with that produced by other conventional unsupervised neural networks, such as the ART and fuzzy *c*-mean network [21]. The same data samples are utilized to evaluate these methods.

The classification results of the ensemble-SYART clustering method versus other classification methods with the same sensitive feature parameters are shown in Table 5.

From the table it can be drawn that the classification accuracy of the ensemble-SYART is the highest and reaches 96.33% while the classification accuracy of fuzzy *c*-mean is less than 70%. For the ART and Yu's norm ART, the classification accuracy is 73.33% and 79.17%, which means that Yu's norm ART can make more use of the neuron information and can diagnose the data samples in the fuzzy region more accurately with soft competition mechanism. These indicate

TABLE 5: Comparison of classification with different neural networks.

Neural networks	Classification accuracy
ART	73.33%
Yu's norm ART	79.17%
Fuzzy c -mean	<70%
SYART	84.17%
Ensemble-SYART	96.33%

that the proposed ensemble-SYART clustering method has superior classification performance comparatively.

6. Conclusions

In this paper, a novel method for fault diagnosis based on an improved Yu's norm ART and neural network ensemble (NNE) is presented to diagnose the faults of rolling element bearings, whose soft competition mechanism is involved to improve the performance of fault diagnosis.

The sensitive feature parameters which are selected with improved distance discriminant techniques to overcome the redundancy and irrelevance of some features are chosen to input the fault diagnosis model to diagnose the fault categories of bearings. The experiment result showed that the proposed ensemble-SYART clustering method can diagnose the faults of bearings successfully, and its diagnosis accuracy is higher than single SYART and some other conventional unsupervised neural networks. And because the initial conditions affect the performance of the proposed clustering method, the bootstrap method is utilized to analyze the diagnosis results. The statistical analysis result shows that the proposed clustering method is stable and generalized. All these indicate that the proposed method has better diagnosis ability and performance and further demonstrate that the proposed clustering method holds a good promise in the field of fault diagnosis of mechanical system.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work is supported financially by the National Natural Science Foundation of China (Grant no. 51405353 and no. 51505346) and the Natural Science Foundation of Hubei province under Grants no. 2015CFB306 and the research project of Hubei Education under Grants no. Q20151103 and the Open Fund of the Key Laboratory for Metallurgical Equipment and Control in Wuhan University of Science and Technology, Ministry of Education (2014B01), and the Open Research Foundation of State Key Lab. of Digital Manufacturing Equipment & Technology in Huazhong University of Science & Technology (Grant no. DMETKF2017010).

References

- [1] E. Llobet, E. L. Hines, J. W. Gardner, P. N. Bartlett, and T. T. Mottram, "Fuzzy ARTMAP based electronic nose data analysis," *Sensors and Actuators, B: Chemical*, vol. 61, no. 1, pp. 183–190, 1999.
- [2] S. Grossberg, "Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world," *Neural Networks*, vol. 37, pp. 1–47, 2013.
- [3] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, "Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 698–713, 1992.
- [4] J. W. Gardner, E. L. Hines, and C. Pang, "Detection of vapours and odours from a multisensor array using pattern recognition: Self-organising adaptive resonance techniques," *Measurement and Control*, vol. 29, no. 6, pp. 172–177, 1996.
- [5] G. A. Carpenter and S. Grossberg, "ART 2: self-organization of stable category recognition codes for analog input patterns," *Applied Optics*, vol. 26, no. 23, pp. 4919–4930, 1987.
- [6] G. A. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Computer Vision, Graphics and Image Processing*, vol. 37, no. 1, pp. 54–115, 1987.
- [7] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, vol. 4, no. 6, pp. 759–771, 1991.
- [8] I. Dagher, M. Georgiopoulos, G. L. Heileman, and G. Bebis, "An ordering algorithm for pattern presentation in fuzzy ARTMAP that tends to improve generalization performance," *IEEE Transactions on Neural Networks*, vol. 10, no. 4, pp. 768–778, 1999.
- [9] Z. Xu, Y. Li, Z. Wang, and J. Xuan, "A novel clustering method combining ART with Yu's norm for fault diagnosis of bearings," *Shock and Vibration*, vol. 2016, Article ID 5468716, 2016.
- [10] C. Yi, Y. Lv, Z. Dang, H. Xiao, and X. Yu, "Quaternion singular spectrum analysis using convex optimization and its application to fault diagnosis of rolling bearing," *Measurement*, vol. 103, pp. 321–332, 2017.
- [11] C. Yi, Y. Lv, M. Ge, H. Xiao, and X. Yu, "Tensor singular spectrum decomposition algorithm based on permutation entropy for rolling bearing fault diagnosis," *Entropy*, vol. 19, no. 4, p. 139, 2017.
- [12] D. Chen, Y. Yang, H. Wang, and A. Mahmood, "Convergence analysis of semi-supervised clustering ensemble," in *Proceedings of the 3rd International Conference on Information Science and Technology (ICIST '13)*, pp. 783–788, March 2013.
- [13] A. P. Topchy, M. H. C. Law, A. K. Jain, and A. L. Fred, "Analysis of consensus partition in cluster ensemble," in *Proceedings of the Fourth International Conference on Data Mining (ICDM '04)*, pp. 225–232, November 2004.
- [14] Z. Xu, Y. Li, Z. Wang, and J. Xuan, "A selective fuzzy ARTMAP ensemble and its application to the fault diagnosis of rolling element bearing," *Neurocomputing*, vol. 182, pp. 25–35, 2016.
- [15] Y. D. Yu, "Triangular norms and TNF-sigma-algebras," *Fuzzy Sets and Systems. An International Journal in Information Science and Engineering*, vol. 16, no. 3, pp. 251–264, 1985.
- [16] P. Luukka, "Similarity classifier using similarity measure derived from Yu's norms in classification of medical data sets," *Computers in Biology and Medicine*, vol. 37, no. 8, pp. 1133–1140, 2007.

- [17] C. C. Jou, "Fuzzy clustering using fuzzy competitive learning networks," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '92)*, pp. 714–719, Baltimore, MD, USA.
- [18] "The case western reserve university bearing data center, bearing data center fault test data [EB/OL]," <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.
- [19] B. S. Yang, T. Han, and J. L. An, "ART-KOHONEN neural network for fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 645–657, 2004.
- [20] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability, Chapman and Hall, New York, NY, USA, 1993.
- [21] B. Zhang, S. Qin, W. Wang, D. Wang, and L. Xue, "Data stream clustering based on Fuzzy C-Mean algorithm and entropy theory," *Signal Processing*, vol. 126, pp. 111–116, 2016.



Hindawi

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

