Research Article

A Novel Bearing Fault Diagnosis Method Based on Gaussian Restricted Boltzmann Machine

Xiao-hui He,[1] Dong Wang,[1,2] Yan-feng Li,[3] and Chun-hua Zhou[1]

[1]*Institute of Battlefield Engineering, PLA University of Science and Technology, Nanjing 210007, China*
[2]*Second Institute of Engineering Research and Design, Southern Theatre Command, Kunming 650222, China*
[3]*Equipment Support Department, Bengbu Automobile NCO Academy, Bengbu 233011, China*

Correspondence should be addressed to Dong Wang; dyhkxydfbb@163.com

To realize the fault diagnosis of bearing effectively, this paper presents a novel bearing fault diagnosis method based on Gaussian restricted Boltzmann machine (Gaussian RBM). Vibration signals are firstly resampled to the same equivalent speed. Subsequently, the envelope spectrums of the resampled data are used directly as the feature vectors to represent the fault types of bearing. Finally, in order to deal with the high-dimensional feature vectors based on envelope spectrum, a classifier model based on Gaussian RBM is applied. Gaussian RBM has the ability to provide a closed-form representation of the distribution underlying the training data, and it is very convenient for modeling high-dimensional real-valued data. Experiments on 10 different data sets verify the performance of the proposed method. The superiority of Gaussian RBM classifier is also confirmed by comparing with other classifiers, such as extreme learning machine, support vector machine, and deep belief network. The robustness of the proposed method is also studied in this paper. It can be concluded that the proposed method can realize the bearing fault diagnosis accurately and effectively.

1. Introduction

Bearing is the most important component of rotating machinery. The majority of problems in rotating machinery arise from the faulty bearing [1, 2]. When fault occurs in bearing, it may lead to fatal breakdown and serious damage. Therefore, the fault diagnosis of bearing is very important and can yield good results.

Vibration signals usually contain lots of information which can reflect the fault types of bearing. Therefore, vibration-based monitoring techniques have been widely used [3]. Feature extraction from vibration signals is a critical step in bearing fault diagnosis. There are many feature extraction methods, such as wavelet transform [4], empirical mode decomposition (EMD) [5], and morphological analysis [6]. In order to further improve the accuracy and efficiency of the diagnosis work, some new feature extraction methods have been proposed in recent years. Reference [7] combined wavelet leaders' multifractal features and wavelet package energy features together for the bearing fault diagnosis. Liu

et al. [8] extracted statistical features from several intrinsic mode functions (IMFs) both in time and in frequency domains. The aforementioned feature extraction methods have been successfully used in the bearing fault diagnosis. However, most of them are complex, and the computation of the features is really a burden. Excessive irrelevant features are often extracted from vibration signals. In order to further improve the diagnosis accuracy, many feature selection techniques have to be used [9]. This often makes the fault diagnosis much more complex.

Envelope spectrum analysis is a common and wildly used feature extraction method. When fault occurs in bearing, the envelope spectrums of vibration signals would change and the character frequency could be found. Envelope spectrums can reflect the fault types of bearing in frequency domain clearly. To make the feature extraction process more simple, envelope spectrums are used directly as the feature vectors to represent different fault types in this paper. The variety of speed would influence the envelope spectrum greatly. To solve this problem, the vibration signals are resampled to

the same equivalent speed [3]. Therefore, the feature vectors based on envelope spectrum could be more robust and not sensible to the change of speed.

Pattern recognition is another task of bearing fault diagnosis process. Artificial neural networks (ANN) [10] and support vector machine (SVM) [11] are the most widely used classifiers. However, these classifiers have some disadvantages such as local optimal solution, low convergence rate, and time-consuming, especially when dealing with high-dimensional data.

In the past decades, pattern recognition techniques have moved into a new platform of learning procedure called deep learning [12]. In deep learning, Gaussian restricted Boltzmann machine (Gaussian RBM) has the ability to provide a closed-form representation of the distribution underlying the training data [13]. Therefore, Gaussian RBM can be used for modeling high-dimensional real-valued data. Successful implementations of Gaussian RBM have been reported in digit recognition [14], image recognition [15], and so forth. However, Gaussian RBM has not been applied in the field of bearing fault diagnosis so far.

In this paper, we propose a novel bearing fault diagnosis method based on Gaussian RBM. In the proposed method, the vibration signals are resampled to the same equivalent speed. Then, the envelope spectrums of the resampled vibration signals are used directly as the feature vectors to represent the fault types of bearing. Finally, a Gaussian RBM classifier model is established to realize the fault diagnosis of bearing. The experiments on 10 different data sets are used to demonstrate the effectiveness of the proposed method. Gaussian RBM classifier is also compared with other classifiers, and the results indicate that Gaussian RBM classifier can have a better performance. The robustness of the proposed method is also studied in this paper.

In this paper, the feature extraction method based on envelope spectrum is described briefly in Section 2. In Section 3, Gaussian RBM is introduced, and a Gaussian RBM classifier model is established. Section 4 introduces the fault diagnosis methodology. In Section 5, the performance of the proposed method is studied. Finally, a conclusion is drawn in Section 6.

2. Feature Extraction

When fault occurs in bearing, the envelope spectrum of vibration signal contains a considerable amount of fault information. Statistical features extracted from envelope spectrum are often used for bearing fault diagnosis. However, the process of selecting the most sensitive features is a task which needs considerable expertise. In order to avoid too complex feature extraction process, the envelope spectrum of vibration signal is used directly to construct the feature vector in this paper. Because the feature vector is influenced greatly by the speed, the vibration signal should be resampled firstly. By resampling, we can get the same number of sample points at every revolution of bearing. Therefore, the influence of different speeds can be reduced. In the resampling process, an antialiasing (low-pass) FIR filter is applied to compensate
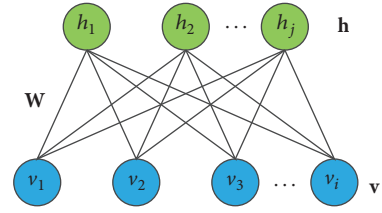

FIGURE 1: Architecture of Gaussian RBM.

for the signal delay [3]. In this paper, the feature extraction method can be described as below.

Step 1. Resample the vibration signals to the same equivalent speed.

Step 2. Get the envelope spectrum of the resampled vibration signals and use it directly as the feature vector to represent the fault type of bearing. The feature vector can be formulated as

$$\mathbf{T} = [f_0, f_1, \ldots, f_n],\tag{1}$$

where $\mathbf{T}$ is the feature vector and $f_0, f_1, \ldots, f_n$ are the values of the envelope spectrum lines.

3. Gaussian RBM Classifier Model

3.1. Architecture of Gaussian RBM. The architecture of Gaussian RBM is shown in Figure 1. Gaussian RBM is made up by a visible layer $\mathbf{v}$ and a hidden layer $\mathbf{h}$, and it is a bipartite undirected graphical model. For the visible layer $\mathbf{v}$ and the hidden layer $\mathbf{h}$, all visible units are connected to all hidden units, and there are no connections between any two units in the same layer [13]. The visible units of Gaussian RBM are linear units with independent Gaussian noise, whereas the hidden units are binary stochastic units [14, 15].

Gaussian RBM is an energy-based stochastic neural network [14, 16]. The joint probability distribution over visible layer $\mathbf{v}$ and hidden layer $\mathbf{h}$ is defined by an energy function. The energy function is given as

$$E(\mathbf{v}, \mathbf{h}; \theta) = \sum_i \frac{(v_i - b_i)^2}{2\sigma_i^2} - \mathbf{a}^T\mathbf{h} - \mathbf{v}^T\mathbf{W}\mathbf{h}.\tag{2}$$

The joint distribution is formulated as

$$p(\mathbf{v}, \mathbf{h}; \theta) = \frac{1}{Z(\theta)} \exp(-E(\mathbf{v}, \mathbf{h}; \theta)),$$
$$Z(\theta) = \int \sum_h \exp(-E(\mathbf{v}, \mathbf{h}; \theta))\, dv,\tag{3}$$

where $Z(\theta)$ is the normalizing factor.

Based on the joint distribution of Gaussian RBM, the probability that the model assigns to a visible vector is given as

$$p(\mathbf{v}; \theta) = \sum_h \frac{\exp(-E(\mathbf{v}, \mathbf{h}; \boldsymbol{\theta}))}{\int \sum_h \exp(-E(\mathbf{v}, \mathbf{h}; \boldsymbol{\theta}))\, dv}.\tag{4}$$

For Gaussian RBM, since there are no visible-visible or hidden-hidden connections, the conditional probability $p(h \mid \mathbf{v}; \theta)$ and $p(v \mid \mathbf{h}; \theta)$ can be expressed as

$$p\left(h_j = 1 \mid \mathbf{v}; \theta\right) = \frac{1}{1 + \exp\left(-\left(a_j + \sum_i w_{ij}\left(v_i / \sigma_i\right)\right)\right)}, \quad (5)$$

$$p\left(v_i = x \mid \mathbf{h}; \theta\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{\left(x - b_i - \sigma_i \sum_j h_j w_{ij}\right)^2}{2\sigma_i^2}\right), \quad (6)$$

where a_j and b_i are biases, σ_i is standard deviation of visible unit v_i, and x is real number.

In order to make the model implementation more simple, each component of the input data is normalized to zero mean and unit variance by a simple linear transformation [17].

3.2. Training Gaussian RBM.

Training a Gaussian RBM means adjusting its parameters such that the probability distribution of the model represents fits the training data as much as possible [13]. For Gaussian RBM, $p(\mathbf{v}; \theta)$ is the probability of the model simply on the input data. By maximizing $p(\mathbf{v}; \theta)$, the training of Gaussian RBM can be realized [18, 19]. Therefore, the gradient of the negative log probability of the visible layer $\mathbf{v}$ with respect to the parameters θ can be obtained from

$$\frac{\partial \log p(\mathbf{v}; \theta)}{\partial w_{ij}} = \left\langle v_i h_j \right\rangle_{\text{data}} - \left\langle v_i h_j \right\rangle_{\text{model}},$$

$$\frac{\partial \log p(\mathbf{v}; \theta)}{\partial a_j} = \left\langle h_j \right\rangle_{\text{data}} - \left\langle h_j \right\rangle_{\text{model}}, \quad (7)$$

$$\frac{\partial \log p(\mathbf{v}; \theta)}{\partial b_i} = \left\langle v_i \right\rangle_{\text{data}} - \left\langle v_i \right\rangle_{\text{model}},$$

where $\langle \cdot \rangle_{\text{data}}$ is an expectation with respect to the training data's distribution and $\langle \cdot \rangle_{\text{model}}$ denotes an expectation with respect to the distribution defined by Gaussian RBM.

According to (7) and stochastic gradient descent, the update rule of the parameters θ is given as

$$\Delta w_{ij} = \epsilon\left(\left\langle v_i h_j \right\rangle_{\text{data}} - \left\langle v_i h_j \right\rangle_{\text{model}}\right),$$

$$\Delta a_j = \epsilon\left(\left\langle h_j \right\rangle_{\text{data}} - \left\langle h_j \right\rangle_{\text{model}}\right), \quad (8)$$

$$\Delta b_i = \epsilon\left(\left\langle v_i \right\rangle_{\text{data}} - \left\langle v_i \right\rangle_{\text{model}}\right),$$

where ϵ is a learning rate.

Due to the special architecture of Gaussian RBM, it is very easy to get the unbiased sample of $\langle \cdot \rangle_{\text{data}}$. However, the unbiased sample of $\langle \cdot \rangle_{\text{model}}$ is intractable to computer. In order to solve this problem, Hinton developed a fast learning procedure based on contrastive divergence (CD) [13, 16]. The algorithm starts by setting the states of the visible units to a training vector. Then the binary states of the hidden units are all computed in parallel using (5). Once the binary states have been chosen for the hidden units, a "reconstruction" can be
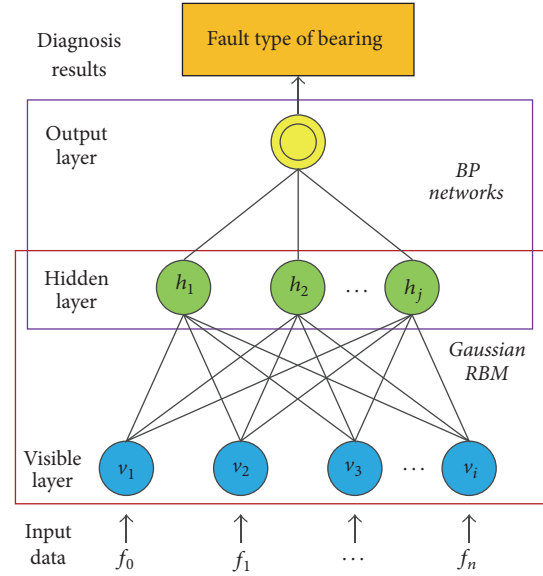

FIGURE 2: Architecture of Gaussian RBM classifier.

produced by setting each v_i to 1 with a probability given by (6). The change of the parameters θ is then given by

$$\Delta w_{ij} \approx \epsilon\left(\left\langle v_i h_j \right\rangle_{\text{data}} - \left\langle v_i h_j \right\rangle_{\text{recon}}\right),$$

$$\Delta a_j \approx \epsilon\left(\left\langle h_j \right\rangle_{\text{data}} - \left\langle h_j \right\rangle_{\text{recon}}\right), \quad (9)$$

$$\Delta b_i \approx \epsilon\left(\left\langle v_i \right\rangle_{\text{data}} - \left\langle v_i \right\rangle_{\text{recon}}\right).$$

In order to train Gaussian RBM more efficiently, the training set is divided into many small "minibatches" of several cases. After estimating the gradient on the minibatch, the parameters of the model are updated. To avoid having to change the learning rate when the size of a minibatch is changed, it is helpful to divide the total gradient computed on a minibatch by the size of the minibatch [19].

3.3. Classifier Model.

After Gaussian RBM's training process, the hidden layer could extract features from the visible layer. The extracted features (the output of the hidden layer) are better for classification than the input data. As shown in Figure 2, the Gaussian RBM classifier consists of three layers, namely, visible layer, hidden layer, and output layer. The visible layer and the hidden layer form a Gaussian RBM. The hidden layer and the output layer form a BP network.

The training of Gaussian RBM classifier contains two main steps, namely, pretraining and backpropagation training. In the pretraining process, Gaussian RBM is trained in an unsupervised manner. The pretraining process is described in Section 3.2. In the backpropagation training process, Gaussian RBM classifier is to be trained with the target values in a supervised manner.
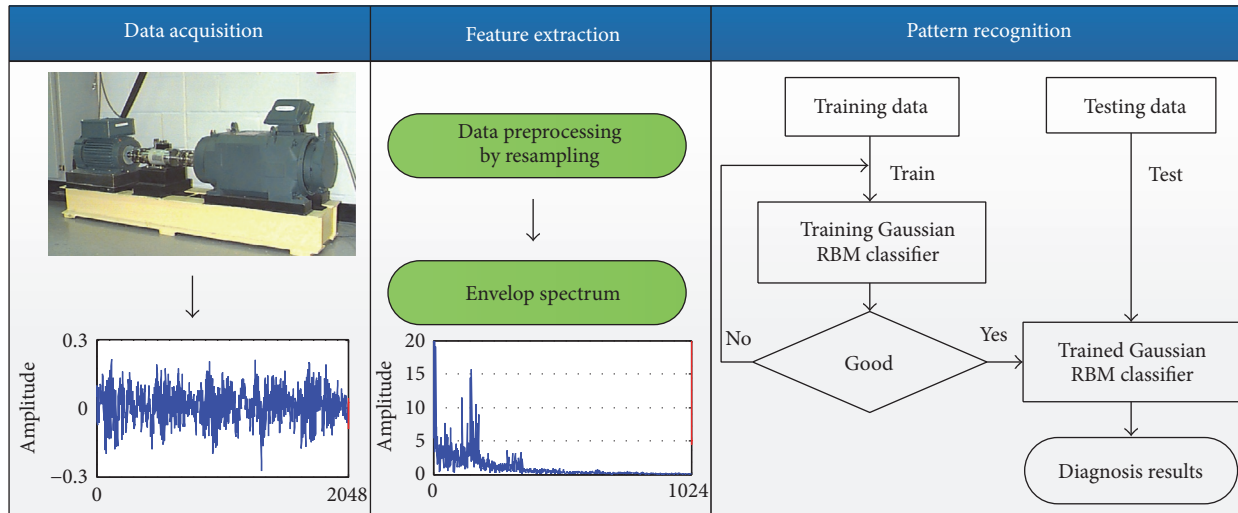
FIGURE 3: Structure diagram of the proposed bearing fault diagnosis method.

4. The Proposed Bearing Fault Diagnosis Method

The diagram of the proposed bearing fault diagnosis method is depicted in Figure 3. The whole procedure can be described as follows.

Step 1. Acquire vibration signals through bearing experiment system.

Step 2. Preprocess vibration signals by resampling and make them have the same equivalent speed.

Step 3. Get the envelope spectrums of the resampled vibration signals and use them directly as the feature vector to represent the fault type of bearing.

Step 4. Build a training data set and testing data set.

Step 5. Develop the Gaussian RBM classifier model and train it by inputting the training data set.

Step 6. Realize the bearing fault diagnosis by using the trained Gaussian RBM classifier model.

5. Experiments

5.1. Experiment System and Experiment Data. The experiment data used in this paper are obtained from the Case Western Reserve University Bearing Data Center [20]. The bearing experiment system consists of a 2 hp induction motor, a torque transducer, accelerometer, a dynamometer, and so on, as shown in Figure 4. The bearing used in the experiment is 6205-2RS JEM SKF deep groove ball bearing. Single-point faults with different fault diameters (7, 14, and 21 mil) are introduced to the driving end bearing using electrodischarge machining. The bearings with different fault diameters (7, 14, and 21 mil) are tested under four different
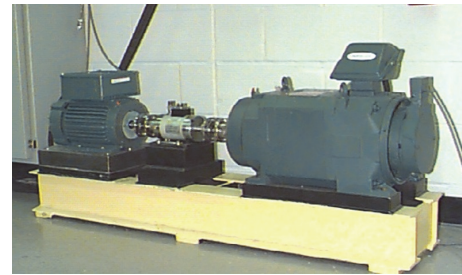

FIGURE 4: Bearing experiment system.

loads (0, 1, 2, and 3 hp). In the experiment, the speeds of the motor would change because of different loads. In order to acquire the vibration signals, an accelerometer is placed at the drive end of the motor. A data recorder is used to collect the vibration signals with the sample frequency 12 kHz. All the experiment data used in this paper are listed in Table 1.

5.2. Feature Extraction and Data Sets. Figure 5 displays the time domain signals of some fault types under four different loads. Each signal contains 2048 points. From these signals, it is not easy to identify the fault types.

All the data in Table 1 are resampled to the same equivalent speed 1772 r/min. Then, the resampled data are truncated into time-series with 2048 points, and the envelope spectrums of these time-series are used directly as the feature vectors to represent the fault types of bearing. Finally, all the feature vectors (samples) are separated into 10 different data sets, which include different fault types under four different loads (0 hp, 1 hp, 2 hp, and 3 hp). Each sample of the data sets contains 1024 points.

In many published references [7, 21, 22], the authors analyzed only a few fault types, and the training data and the testing data have the same speed and load. However, the
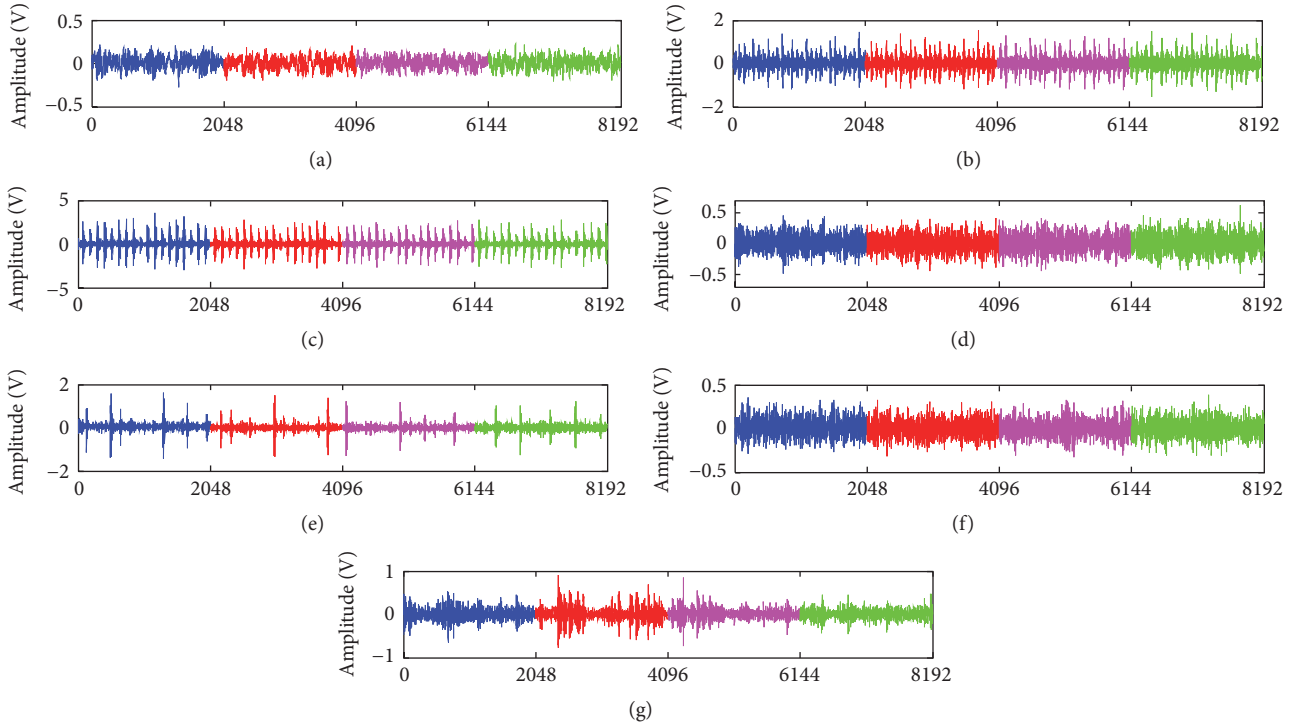
FIGURE 5: Time domain signals of some fault types under four different loads: (a) normal; (b) inner race fault (7 mil); (c) outer race fault (7 mil); (d) ball fault (7 mil); (e) inner race fault (14 mil); (f) outer race fault (14 mil); and (g) ball fault (14 mil). The colors blue, red, purple, and green represent the 0 hp, 1 hp, 2 hp, and 3 hp load, respectively.

TABLE 1: Experiment data description.

Fault diameter (mil)	Load (hp)	Speed (r/min)	Normal	Inner race	Outer race	Ball
0	0	1797	Normal_0	—	—	—
	1	1772	Normal_1	—	—	—
	2	1750	Normal_2	—	—	—
	3	1730	Normal_3	—	—	—
7	0	1797	—	IR007_0	OR007@6_0	B007_0
	1	1772	—	IR007_1	OR007@6_1	B007_1
	2	1750	—	IR007_2	OR007@6_2	B007_2
	3	1730	—	IR007_3	OR007@6_3	B007_3
14	0	1797	—	IR014_0	OR014@6_0	B014_0
	1	1772	—	IR014_1	OR014@6_1	B014_1
	2	1750	—	IR014_2	OR014@6_2	B014_2
	3	1730	—	IR014_3	OR014@6_3	B014_3
21	0	1797	—	IR021_0	OR014@6_0	B021_0
	1	1772	—	IR021_1	OR014@6_1	B021_1
	2	1750	—	IR021_2	OR014@6_2	B021_2
	3	1730	—	IR021_3	OR014@6_3	B021_3

running conditions of bearing are very complex in engineering application, and the fault diagnosis method should be robust and insensible to the variation of speed and load. For all the data sets in this paper, only the samples under 0 hp load are used for training, and the rest of samples under 0 hp, 1 hp, 2 hp, and 3 hp load are used for testing. The details of 10 different data sets are presented in Table 2.

The feature vectors of some fault types under four different loads are shown in Figure 6. The dimension of each feature vector is 1024. The feature vectors of the same fault type are close to each other, while that of different fault types can be clearly distinguished. It means that the feature vectors extracted by the proposed method can represent the fault types of bearing very well. Furthermore, the feature vectors
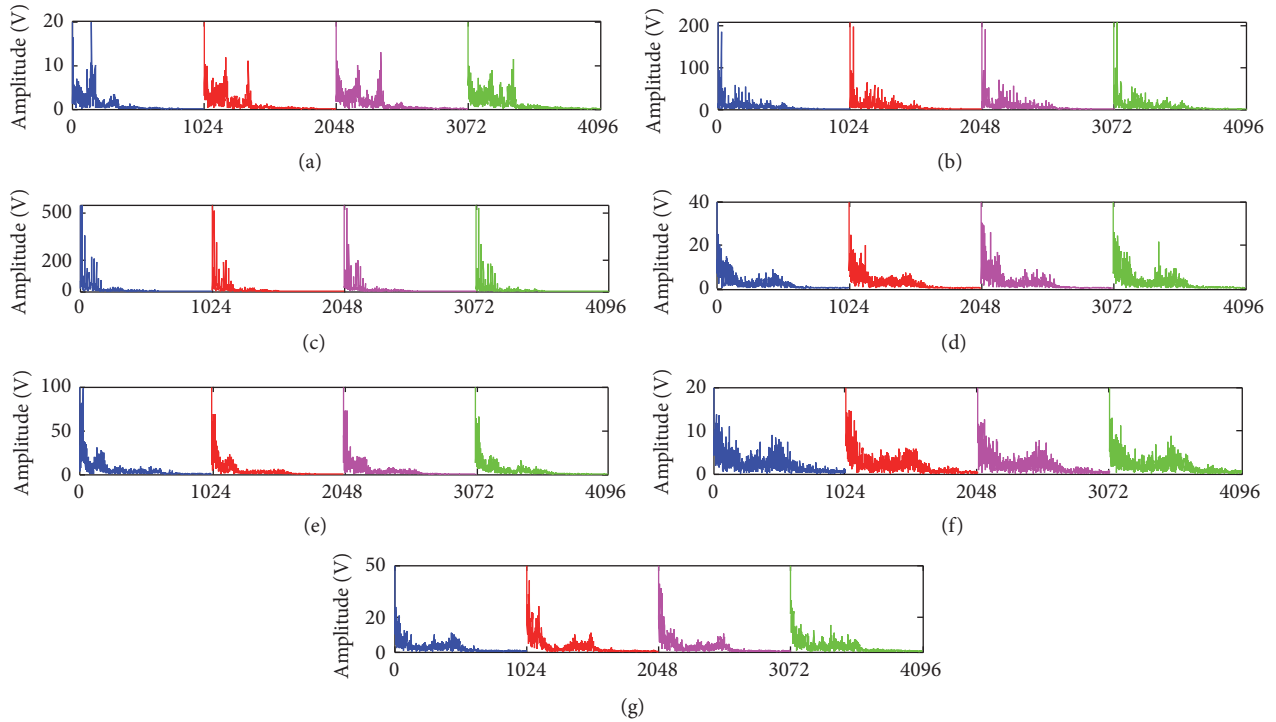
FIGURE 6: Feature vectors of some fault types under four different loads: (a) normal; (b) inner race fault (7 mil); (c) outer race fault (7 mil); (d) ball fault (7 mil); (e) inner race fault (14 mil); (f) outer race fault (14 mil); and (g) ball fault (14 mil). The colors blue, red, purple, and green represent the 0 hp, 1 hp, 2 hp, and 3 hp load, respectively.

TABLE 2: Details of 10 different data sets.

Data set	Training data	Testing data	Fault diameter (mil)	Fault type	Classification label
D070707	160	240	0 7 7 7	N I O B	1 2 3 4
D141414	160	240	0 14 14 14	N I O B	1 2 3 4
D212121	160	240	0 21 21 21	N I O B	1 2 3 4
DINN	160	240	0 7 14 21	N I I I	1 2 3 4
DOUT	160	240	0 7 14 21	N O O O	1 2 3 4
DBALL	160	240	0 7 14 21	N B B B	1 2 3 4
D071421	160	240	0 7 14 21	N I O B	1 2 3 4
D142107	160	240	0 14 21 7	N I O B	1 2 3 4
D210714	160	240	0 21 07 14	N I O B	1 2 3 4
DALL	400	600	0	N	1
			7 14 21	I	2 3 4
			7 14 21	O	5 6 7
			7 14 21	B	8 9 10

N—normal, B—ball fault, I—inner race fault, O—inner race fault.

are robust and not very sensible to the change of speed and load.

5.3. Results and Analysis.
Gaussian RBM classifier is utilized to realize the automatic recognition of the bearing fault types. For Gaussian RBM classifier, the visible layer and the output layer are constructed with neurons denoting the input data and the target classes, respectively, whereas 1000 neurons are used for the hidden layer. The minibatch size is set equal to the number of target classes. The learning rates for weights and biases are set to 0.001. The numbers of epochs for pretraining and backpropagation training are 50 and 100, respectively.

In order to show the superiority of Gaussian RBM classifier, other methods, such as extreme learning machine (ELM), support vector machine (SVM), and deep belief network (DBN), are also used for comparison in this paper. For ELM, the number of neurons in the hidden layer is set to 1000. For SVM, its type is one-against-all and the kernel function is selected as Gaussian function. For DBN, there are two hidden layers with 1000 and 100 neurons, respectively. The visible

TABLE 3: Classification results of the data sets.

Data set	ELM (%)	SVM (%)	DBN (%)	Gaussian RBM classifier (%)
D070707	95.92	95.88	99.33	100
D141414	91.53	97.44	98.79	100
D212121	97.15	98.69	99.58	100
DINN	100	94.50	87.25	100
DOUT	94.63	98.81	99.41	99.75
DBALL	76.18	82.75	89.7	95.33
D071421	85.12	89.69	92.2	97.99
D142107	92.99	93.56	96.01	100
D210714	95.49	99.19	99.28	100
DALL	71.43	89.52	90.75	93.25

layer of DBN is selected as the Gaussian neurons. In DBN, the minibatch size, the learning rates, and the numbers of epochs for pretraining and backpropagation training are the same as Gaussian RBM classifier. To account for the stochastic nature of machine learning, the classification process is repeated for 10 times and then the results are averaged. Table 3 summarizes the classification results of the above methods.

From Table 3, it can be seen that Gaussian RBM classifier performs better than ELM, SVM, and DBN for all the data sets. ELM is suitable for dealing with high-dimensional data, and its training process is very fast. However, the random choice of input weights and biases easily causes the so-called hidden layer output matrix not full column rank, and this lowers the effectiveness of ELM. For most of the data sets, ELM could achieve high classification accuracy. Because the complexity of DBALL and DALL is higher than the other data sets in Table 2, their classification accuracies achieved by ELM are only 76.18% and 71.43%, respectively. SVM performs better than ELM for most data sets except D070707 and DINN. However, the training process of SVM is quite time-consuming. DBN is a deep neural network; it can learn high complexity relationship between the input data and the target classes in the deep learning process. DBN performs better than ELM and SVM for most data sets. In this paper, the training processes of Gaussian RBM and DBN are the same, but the former has a simpler architecture. This makes Gaussian RBM classifier learn the relationship between the input data and the target classes much more efficiently. Therefore, Gaussian RBM classifier could outperform DBN for all the data sets.

5.4. Robustness of the Proposed Method. In engineering application, the running conditions of bearing are often very complex. The robustness is very important for the fault diagnosis method. DALL is the most complex data set in Table 2. For a clear understanding of the robustness of the proposed method, we use DALL to illustrate the robustness of the proposed method in this paper.

The training data of DALL are samples under 0 hp load, and they are replaced by the samples under 1 hp, 2 hp, and 3 hp load. Therefore, we can get four different data sets which are denoted as DALL_0, DALL_1, DALL_2, and DALL_3, respectively. For each data set, the classification process is
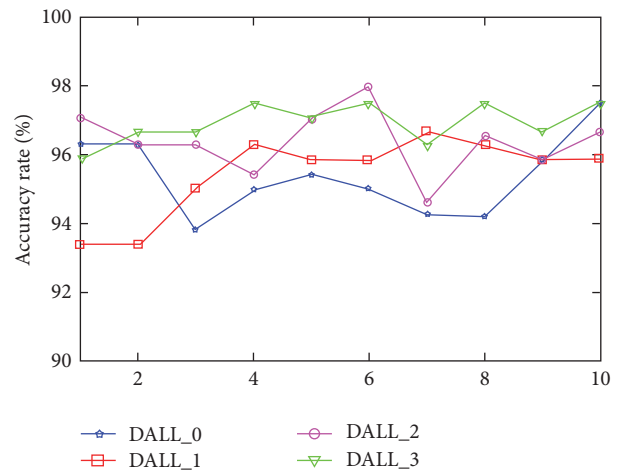

FIGURE 7: Classification results for DALL_0, DALL_1, DALL_2, and DALL_3.

repeated for 10 times, and the results are displayed in Figure 7. Although the training data are under different loads, the classification accuracies of four different data sets are still very high. The average classification accuracies for DALL_0, DALL_1, DALL_2, and DALL_3 are 93.25%, 95.15%, 96.08%, and 94.57%, respectively. In the experiment, the speeds of the motor would change under different loads. The results in Figure 7 demonstrate that the proposed method has a high robustness and is not sensible to the change of load and speed.

6. Conclusions

This paper has proposed a novel bearing fault diagnosis method based on Gaussian RBM. In the proposed method, the vibration signals are firstly resampled to the same equivalent speed. Then, the envelope spectrums of the resampled vibration signals are used directly as the feature vectors to represent the fault types of bearing. Finally, to deal with the high-dimensional feature vectors based on envelope spectrum, a Gaussian RBM classifier is applied to realize the bearing fault diagnosis. Experiment on ten different data sets verifies the performance of the proposed method. Gaussian RBM classifier is also compared with ELM and DBN, and the

result demonstrates that Gaussian RBM classifier has a better performance. The robustness of the proposed method is also verified in this paper. It can be concluded that the proposed method can realize the fault diagnosis of bearing accurately and effectively.

Competing Interests

The authors declare no conflict of interests.

Acknowledgments

References

[1] X. Li, A. Zheng, X. Zhang, C. Li, and L. Zhang, "Rolling element bearing fault detection using support vector machine with improved ant colony optimization," *Measurement*, vol. 46, no. 8, pp. 2726–2734, 2013.

[2] P. Zhang, Y. Du, T. G. Habetler, and B. Lu, "A survey of condition monitoring and protection methods for medium-voltage induction motors," *IEEE Transactions on Industry Applications*, vol. 47, no. 1, pp. 34–46, 2011.

[3] X. Wang, Y. Li, T. Rui, H. Zhu, and J. Fei, "Bearing fault diagnosis method based on Hilbert envelope spectrum and deep belief network," *Journal of Vibroengineering*, vol. 17, no. 3, pp. 1295–1308, 2015.

[4] Y. Pan, J. Chen, and X. Li, "Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means," *Mechanical Systems and Signal Processing*, vol. 24, no. 2, pp. 559–566, 2010.

[5] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.

[6] H. Li, H.-Q. Zheng, and L.-W. Tang, "Bearing multi-fault diagnosis based on improved morphological component analysis," *Journal of Vibration & Shock*, vol. 31, no. 12, pp. 135–140, 2012.

[7] W. Du, J. Tao, Y. Li, and C. Liu, "Wavelet leaders multifractal features based fault diagnosis of rotating mechanism," *Mechanical Systems and Signal Processing*, vol. 43, no. 1-2, pp. 57–75, 2014.

[8] Z.-W. Liu, H.-R. Cao, X.-F. Chen, Z. He, and Z. Shen, "Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings," *Neurocomputing*, vol. 99, no. 1, pp. 399–410, 2013.

[9] M. Saimurugan, K. I. Ramachandran, V. Sugumaran, and N. R. Sakthivel, "Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3819–3826, 2011.

[10] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech, "Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals," *Applied Acoustics*, vol. 89, no. 3, pp. 16–27, 2015.

[11] K. C. Gryllias and I. A. Antoniadis, "A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 2, pp. 326–344, 2012.

[12] T. Kuremoto, S. Kimura, K. Kobayashi, and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," *Neurocomputing*, vol. 137, pp. 47–56, 2014.

[13] A. Fischer and C. Igel, "Training restricted Boltzmann machines: an introduction," *Pattern Recognition*, vol. 47, no. 1, pp. 25–39, 2014.

[14] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–27, 2009.

[15] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[16] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," *Momentum*, vol. 9, no. 1, pp. 599–619, 2010.

[17] C. Shang, F. Yang, D. Huang, and W. Lyu, "Data-driven soft sensor development based on deep learning technique," *Journal of Process Control*, vol. 24, no. 3, pp. 223–233, 2014.

[18] N.-N. Ji, J.-S. Zhang, and C.-X. Zhang, "A sparse-response deep belief network based on rate distortion theory," *Pattern Recognition*, vol. 47, no. 9, pp. 3179–3191, 2014.

[19] V. T. Tran, F. AlThobiani, and A. Ball, "An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks," *Expert Systems with Applications*, vol. 41, no. 9, pp. 4113–4122, 2014.

[20] K. A. Loparo, "Bearing data set," Case Western Reserve University Bearing Data Center, http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website.

[21] Q. Hu, Z.-J. He, Z.-S. Zhang, and Y. Zi, "Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble," *Mechanical Systems & Signal Processing*, vol. 21, no. 2, pp. 688–705, 2007.

[22] C. Liu, G. Wang, Q. Xie, and Y. Zhang, "Vibration sensor-based bearing fault diagnosis using ellipsoid-ARTMAP and differential evolution algorithms," *Sensors (Basel, Switzerland)*, vol. 14, no. 6, pp. 10598–10618, 2014.