

Research Article

Bearing Fault Diagnosis Based on Deep Belief Network and Multisensor Information Fusion

Jie Tao,^{1,2} Yilun Liu,^{1,3} and Dalian Yang^{1,4}

¹School of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China

²Key Laboratory of Knowledge Processing and Networked Manufacturing, Hunan University of Science and Technology, Xiangtan 411201, China

³Light Alloy Research Institute, Central South University, Changsha 410083, China

⁴Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, China

Correspondence should be addressed to Jie Tao; caroltaojie@126.com

Received 20 April 2016; Revised 3 August 2016; Accepted 11 August 2016

Academic Editor: Ganging Song

Copyright © 2016 Jie Tao 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.

In the rolling bearing fault diagnosis, the vibration signal of single sensor is usually nonstationary and noisy, which contains very little useful information, and impacts the accuracy of fault diagnosis. In order to solve the problem, this paper presents a novel fault diagnosis method using multivibration signals and deep belief network (DBN). By utilizing the DBN's learning ability, the proposed method can adaptively fuse multifeature data and identify various bearing faults. Firstly, multiple vibration signals are acquainted from various fault bearings. Secondly, some time-domain characteristics are extracted from original signals of each individual sensor. Finally, the features data of all sensors are put into the DBN and generate an appropriate classifier to complete fault diagnosis. In order to demonstrate the effectiveness of multivibration signals, experiments are carried out on the individual sensor with the same conditions and procedure. At the same time, the method is compared with SVM, KNN, and BPNN methods. The results show that the DBN-based method is able to not only adaptively fuse multisensor data, but also obtain higher identification accuracy than other methods.

1. Introduction

Bearing is one of the critical components which has a broad range of application in mechanical equipment. Due to the overload, fatigue, wear, corrosion, and other reasons, bearing is easily damaged in the process of machine operation. As a matter of fact, more than 50% of rotating machine malfunctions are related to bearing faults [1, 2]. Actually, a rolling bearing fault may lead to equipment intense shaking, apparatus shutdown, stopping producing, and even casualties. In general, the early weak fault of bearing is complicated and hard to detect [3, 4]. Therefore, bearing state monitoring and analysis is very important, in which it can discover early weak fault of the bearing and control the fault damage situation in time.

Recently, fault detection and diagnosis of bearing has been attracting considerable attention. Among all the kinds

of bearing fault diagnosis methods, vibration signal analysis is one of the most principal and useful tools [2]. In vibration-based bearing fault diagnosis, there are two kinds of approaches that have been proven effective to fault diagnosis: signal processing and pattern recognition [1, 3]. Conventional signal processing techniques such as fast Fourier transform (FFT), wavelet transforms (WT), and empirical mode decomposition (EMD) have been applied to bearing fault diagnosis and achieved some effectiveness [5, 6]. For pattern recognition approaches, artificial intelligence and machine learning are extensively used and studied, for example, fuzzy logic, support vector machine (SVM), and artificial neural network (NN) [7, 8]. However, most research only focused on single vibration analysis in bearing fault diagnosis. In fact, when using a single sensor vibration, the fault characteristics are very weak and useful information is limited. So, it requires intricate signal processing and feature

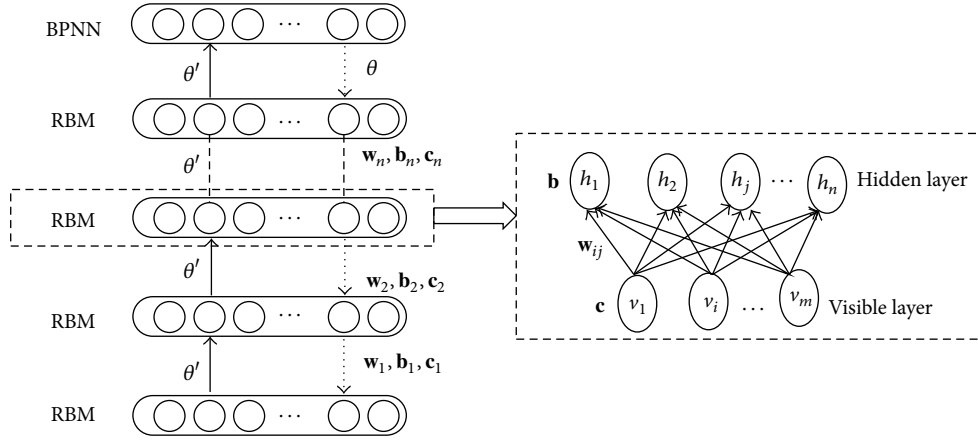


FIGURE 1: Basic structure of DBN.

extraction. Sometimes the accuracy of fault diagnosis is not stable.

To improve the diagnosis accuracy of bearing, some researches put forward the multisignals. At present, fusion of multisource signals mainly focused on three aspects: data level, feature level, and decision level. Among them, the data level fusion primarily mixed the diagnosis objects such as temperature, pressure, and vibration signals [9, 10]. This needs various kinds of sensors and instruments in the process of data gathering. The monitoring cost is expensive and the manipulation is complicated. However, in the convergence of feature level with the same kinds of signals, it needed complex signal analysis and weighted calculation [11, 12]. These methods had some shortcomings such as poor real-time property and weak generalization ability. In the decision level, the intelligent approaches introduced by the literature are, for example, expert systems, decision tree, and SVM [13, 14]. However, these methods all belong to the shallow learning method; the learning ability is lower.

Recently, deep learning became popular in artificial intelligence and machine learning [15]. As a key framework of deep learning, deep belief network (DBN) is primarily constituted by stacked restricted Boltzmann machines (RBM) which is a generative stochastic neural network that can learn probability distribution over abundant data [16]. In 2006, Hinton and colleagues utilized contrastive divergence to advance the RBM training process that greatly improved the learning efficiency of the DBN. The essence of DBN is the capability to automatically extract features through a successive learning process; it can mine the features from different aspects of the data in lower levels as input for the next layer [17, 18]. In addition, DBN accomplishes the learning process with an unsupervised pretraining and supervised fine-tuning. So, DBN has more mapping capability and extensive adaptability by a hierarchical structure. Due to the great advantages of DBN, it has obtained good effect in areas such as natural language understanding, image processing, speech recognition, and document recognition [18–20].

Lately, DBN gets the preliminary application in the field of fault diagnosis. Shao et al. [21] developed particle swarm to optimize the structure of the DBN and applied it

to analyze the simulation signals and experimental signals of a rolling bearing, which obtained more accurate and robust results than other intelligent methods. Tamilselvan et al. [22] originally presented a novel multisensor diagnosis methodology which used the DBN in system health diagnosis such as aircraft engine and electric power transformer. Gan et al. [23] constructed a two-layer DBN of rolling-element bearing fault diagnosis, and experiments showed that DBN got highly reliable results compared to those obtained by SVM and BPNN; Lei et al. [24] proposed a method for multistage gear fault diagnosis with deep learning, which can adaptively extract available fault characteristics from the original data and acquire higher diagnostic accuracy than subsistent methods. Tran et al. [25] presented an approach to implement DBN and multi-information for fault diagnosis of reciprocating compressors.

This paper focuses on the early weak fault of rolling bearing and applies the DBN to integrate the time-domain features of multivibration. The remainder of this paper is organized as follows. In Section 2, the methodologies of deep belief network are introduced. In Section 3, the process of multivibration signal fusion is described. In Section 4, a bearing test rig is explained and experiments are conducted for the proposed method. In Section 5, implementation of classifier based on the DBN model is presented. The obtained results and their evaluation are described. Finally, conclusions and future work are given in Section 6.

2. Deep Belief Network

2.1. Deep Belief Network Architecture. DBN is a model based on probability of energy generation, which comprises multiple layers of restricted Boltzmann machines (RBM) and a backpropagation neural network (BPNN) [16]. Figure 1 is the fundamental structure of DBN; the multilayered architecture makes sure that DBN can be trained through bottom-up learning in a sequence of RBMs and top-down fine-tuning by BPNN [17].

Restricted Boltzmann machine, the key prototype of DBN, is structured by a layer of visible (or input) units and a layer of hidden (or output) units. As every unit is binary,

it is trained by the activation probabilities. The units in the same layer are not connected to each other but have directed symmetrical connections to the units in the next layer. In DBN, the hidden layer of the RBM becomes the visible layer of the next RBM, so they set up a successive hierarchy by stacked RBMs.

In RBM, the visible node is denoted by v_i and the hidden node is represented by h_j . The weights between v_i and h_j are directed and denoted by \mathbf{w}_{ij} . The visible and hidden nodes have their biases represented by vectors \mathbf{c} and \mathbf{b} , respectively. \mathbf{b}_i , \mathbf{c}_j , and \mathbf{w}_{ij} of all RBMs make up the parameter set θ in DBN. As the values of θ define a probability distribution over the joint states of the visible and hidden nodes by an energy function,

$$E(\mathbf{v}, \mathbf{h}) = -\sum_{i=1}^m v_i \mathbf{c}_i - \sum_{j=1}^n h_j \mathbf{b}_j - \sum_{i=1}^m \sum_{j=1}^n v_i h_j \mathbf{w}_{ij}. \quad (1)$$

The ultimate purpose of DBN training is to find the best θ , which can minimize the model energy error and make the model at an equilibrium state. So, the energy function is utilized to define the joint probability distribution between \mathbf{v} and \mathbf{h} as follows:

$$P(\mathbf{v}, \mathbf{h} | \theta) = \frac{1}{Z(\theta)} e^{-E(\mathbf{v}, \mathbf{h} | \theta)}, \quad (2)$$

$$Z(\theta) = \sum_{\mathbf{v}, \mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h} | \theta)}.$$

Since DBN has no intralayer connections, the conditional probability distributions of visible and hidden nodes can be calculated by

$$P(v_i = 1 | \mathbf{h}) = \frac{1}{1 + \exp(-\mathbf{b}_i - \sum_j h_j \mathbf{w}_{ij})}, \quad (3)$$

$$P(h_j = 1 | \mathbf{v}) = \frac{1}{1 + \exp(-\mathbf{c}_j - \sum_i v_i \mathbf{w}_{ij})}. \quad (4)$$

2.2. The DBN Training Process. Generally, the DBN training procedure includes two parts: pretraining and fine-tuning. The pretraining is an unsupervised learning procedure which used the unlabeled data to train the individual RBM. The fine-tuning is a supervised learning process which utilized the backpropagation algorithm to further adjust the parameters.

In the pretraining, each layer is trained by the RBM rules. Since the RBM model is with binary units, it can be learned by stochastic gradient descent on the negative log-likelihood probability of the training data. The functions are as follows:

$$\begin{aligned} \frac{\partial \ln p(\mathbf{v}; \theta)}{\partial \mathbf{w}_{ij}} &= \langle v_i h_j \rangle_d - \langle v_i h_j \rangle_m, \\ \frac{\partial \ln p(\mathbf{v}; \theta)}{\partial \mathbf{b}} &= \langle h_j \rangle_d - \langle h_j \rangle_m, \\ \frac{\partial \ln p(\mathbf{v}; \theta)}{\partial \mathbf{c}} &= \langle v_i \rangle_d - \langle v_i \rangle_m, \end{aligned} \quad (5)$$

where $\langle \cdot \rangle_d$ denotes an expectation of the data distribution and $\langle \cdot \rangle_m$ is an expectation of the distribution defined by the model.

With the RBM property, it is easy to compute an unbiased sample of $\langle \cdot \rangle_d$ to the data distribution. However, obtaining an unbiased sample of $\langle \cdot \rangle_m$ is quite difficult [23]. Actually, the RBM learning method closely approximates the gradient objective function called contrastive divergence (CD) [17], in which $\langle \cdot \rangle_m$ is substituted by k iterations of Gibbs sampling as expressed in (6), where an iteration of alternating Gibbs sampling includes updating all parallel visible nodes by using (3), subsequently updating all parallel hidden nodes by (4).

$$\begin{aligned} \frac{\partial \ln p(\mathbf{v}; \theta)}{\partial \mathbf{w}_{ij}} &\approx \langle v_i h_j \rangle^0 - \langle v_i h_j \rangle^k, \\ \frac{\partial \ln p(\mathbf{v}; \theta)}{\partial \mathbf{b}} &\approx \langle h_j \rangle^0 - \langle h_j \rangle^k, \\ \frac{\partial \ln p(\mathbf{v}; \theta)}{\partial \mathbf{c}} &\approx \langle v_i \rangle^0 - \langle v_i \rangle^k. \end{aligned} \quad (6)$$

Actually, one-step Gibbs sampling has been shown to perform surprisingly well [17]. Based on (6), the updated methods for all parameters are given by the following equation, where σ represents learning rate whose value is between 0 and 1:

$$\begin{aligned} \Delta \mathbf{w} &\leftarrow \sigma \left(\langle v_i h_j \rangle^0 - \langle v_i h_j \rangle^1 \right), \\ \Delta \mathbf{b} &\leftarrow \sigma \left(\langle h_j \rangle^0 - \langle h_j \rangle^1 \right), \\ \Delta \mathbf{c} &\leftarrow \sigma \left(\langle v_i \rangle^0 - \langle v_i \rangle^1 \right). \end{aligned} \quad (7)$$

In the training process, dataset is usually divided into minibatches with a small number of data vectors and the values of θ are updated after handling each minibatch. To stabilize the RBM learning procedure, a momentum (λ) is often utilized in updating the synaptic weights and biases. With momentum (λ), the θ update, at the current epoch, can be associated with the θ update in the preceding epoch and calculated as

$$\begin{aligned} \mathbf{w}_n &\leftarrow \lambda \mathbf{w}_{n-1} + \sigma \left(\langle v_i h_j \rangle^0 - \langle v_i h_j \rangle^1 \right), \\ \mathbf{b}_n &\leftarrow \lambda \mathbf{b}_{n-1} + \sigma \left(\langle h_j \rangle^0 - \langle h_j \rangle^1 \right), \\ \mathbf{c}_n &\leftarrow \lambda \mathbf{c}_{n-1} + \sigma \left(\langle v_i \rangle^0 - \langle v_i \rangle^1 \right). \end{aligned} \quad (8)$$

After the bottom-up successive learning, the following step of the DBN training is top-down fine-tuning. Fine-tuning is a supervised learning process which used the backpropagation (BPNN) to further decrease the training error and advance the classification accuracy of the DBN. As the BPNN is supervised learning, fine-tuning uses labeled data for the DBN training. Unlike the unsupervised training in DBN that only deals one RBM at a time, the BPNN simultaneously trains all layers in DBN. The training error of BPNN is calculated with model outputs and the target label data. And the backpropagation learning is continued until the model output attains the maximum number of epochs.

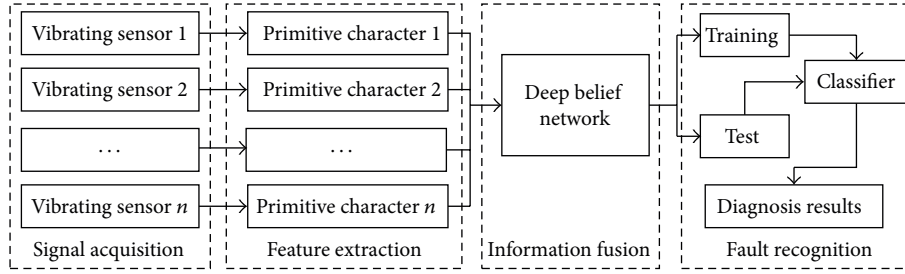


FIGURE 2: The flow diagram of multisensor information fusion.

3. Multisignal Fusion with DBN

Multisensor information fusion technology can obtain more accurate, rich fault features from vibration signals [12]. However, in the conventional information integration, signal processing needs to master a lot of signal processing technologies and to be combined with rich experience in engineering practice to extract fault features. Meanwhile, in the pattern recognition, traditional machine learning only contains single nonlinear transform structure; it cannot adaptively integrate the multi-information [20, 26].

In this paper, we apply the deep belief network (DBN) to adaptively fuse multivibrations. There are four main processes in the proposed bearing: multichannel signal acquisition, feature extraction, information fusion, and fault recognition.

As shown in Figure 2, firstly, the vibration signals are acquainted by each sensor. Secondly, some time-domain characteristics are extracted from original signal of every individual sensor. Thirdly, without any artificial selection, features data of all signal sensors are put into the DBN and generate appropriate DBN classifier. Finally, the integrated information is used to train or test the classifier, and then the classifier puts out the diagnosis results and completes fault diagnosis.

Since the DBN has a hierarchical structure which can extract the features from various aspects of the data by a layer-by-layer successive learning procedure [17], the multi-information fusion, based on deep belief network, can get rid of complex signal processing and complicated experience [24]. It takes the unsupervised learning with RBM and directly extracts feature from the multivibrations and then uses the best parameters to design DBN and completes the multi-information integration.

However, the structure of DBN is closely related to the number of hidden nodes and hidden layers; if the DBN structure is too simple, learning ability is so poor that it cannot effectively integrate the multi-information. Meanwhile, if the DBN structure is too complicated, it not only wastes running time but also produces problems such as overfitting, local extremism, and training failure [26]. Therefore, a method based on data reconstruction error is used to determine the structure of information fusion in DBN.

Figure 3 introduces the optimization process for the DBN structure. The reconstruct error is computed with the model outputs and the objective label data. At the beginning of the procedure, multichannel signal information is put into the

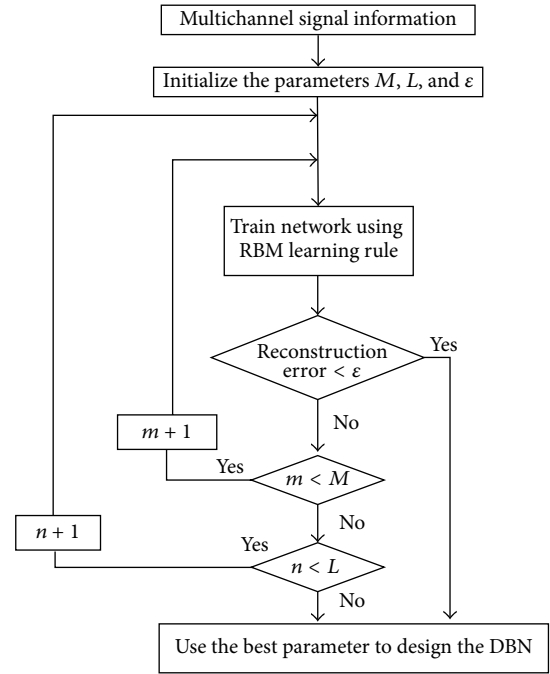


FIGURE 3: The flow chart of the optimization DBN in signal fusion.

DBN and the parameters of M , L , and ϵ are initialized, where the M , L , and ϵ are the max values of the hidden nodes, hidden layers, and reconstruction error, respectively. Then, DBN calculates the reconstruction error of training dataset by RBM learning rules. If the reconstruction error is less than ϵ , it finishes the optimization and puts out the parameters (θ) of DBN. Otherwise, it increases the number of hidden nodes or hidden layers. If the numbers overflow M or L , the procedure finds the best reconstruction error from history and builds the DBN for multi-information fusion.

Table 1 summarizes the procedure of bearing fault diagnosis using multi-information fusion with DBN. As shown in the table, the first step is gating the vibration signals from multichannels and collecting vibration data from each sensor. As the raw samples are nonlinear and unstable, it is necessary to extract some features from each sample. Then, the preprocessed vibration data are divided into training and testing datasets. The DBN structure is optimized by reconstruction error of training dataset and obtains the suitable DBN to accomplish the multi-information fusion.

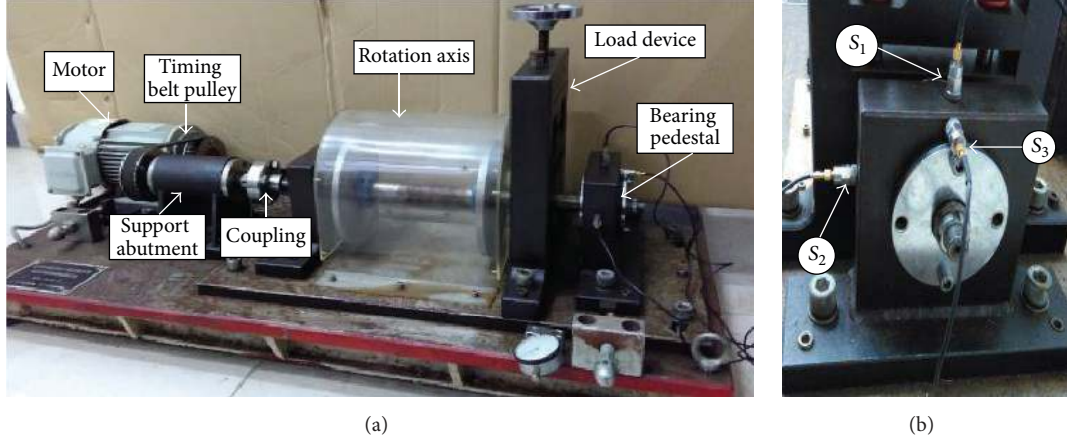


FIGURE 4: Rotating machinery fault simulation platform of QPZ-II: (a) experiment platform; (b) sensors locations.

TABLE 1: Procedure for bearing fault diagnosis using multi-information with DBN.

Step	Description
Step 1	Gather multichannel vibration signals
Step 2	Extract features of each channel sample
Step 3	Input all features of training samples and initial parameters of the DBN
Step 4	Optimize the DBN structure using reconstruction error of multi-information fusion
	(1) Each layer of the DBN is trained using RBM learning rule
	(2) Fine-tune the DBN using backpropagation learning
	(3) Calculate the reconstruction error using model outputs and the target label data
	(4) If the reconstruction error is smaller than ϵ output the DBN structure; otherwise, $m \leftarrow m + 1$ and return to step (1)
Step 5	(5) If the reconstruction error is not smaller than ϵ and $m = M$, $l \leftarrow l + 1$ and return to step (1) until l is more than n
	Develop the DBN using the structure with the best reconstruction error
Step 6	Perform diagnosis using the training DBN classifier model

4. Experimental Setup

In order to measure the validity of the suggested method, a bearing experimental platform is set up as shown in Figure 4. The bearing fault simulation platform was produced by Qian Peng Company with QPZ-II in China.

As shown in Figure 4, the experimental table is mainly constituted with motor, belt coupling, bearing pedestal, and so on. The bearing is installed in the pedestal, and three magnet acceleration sensors are installed in the pedestal, labeled by S_1 , S_2 , and S_3 , respectively. The position of S_1 is located on the vertical side of the bearing pedestal; S_2

and S_3 are, respectively, located on the lateral and front of the bearing pedestal. In the experiments, the variety of typical fault bearings can be installed and dismantled for multivibration collection.

The test bearings are produced by Harbin Bearing Manufacturing Company, China, with the bearing designation being NU205, which have 13 cylindrical rollers. The inner diameter is 25 mm, the outer diameter is 52 mm, and the thickness of the bearing is 15 mm. As shown in Figure 5, four experiments are carried out under each of the following bearing health conditions: the inner race fault, outer race fault, ball fault, and normal. All the faults are linear cutting with electrical discharge machining, and the cutting diameter is 0.5 mm; the cutting depth is 0.3 mm.

In the process of testing, a variable velocity motor directly drives a shaft. The belt on the right of the shaft brings along the coupling which runs with the same speed of motor. In the experiment, the sampling frequency is 10000 hz, the bearing speed is 1200 rpm, and the sampling time is 5 seconds.

According to the steps shown in Table 1, each experiment continuously acquainted 50000 signal points. Meanwhile, the bearing rotated 100 cycles. We select the signal points of a rotation cycle to construct a sample. So, 500 signal points constitute a data sample. There are four conditions defined for classification and 400 (100×4) training samples in dataset. Then, we randomly selected 200 samples constituting the test dataset. The dataset description is shown in Table 2. When the rolling bearing has local damage, it will cause the vibration signal mutation. The local damage position is different and the change of the vibration signal usually is not the same. Figure 6 is the amplitudes waveform of rolling bearing in different conditions.

It is seen from Figure 6 that the vibration signals waveforms are similar, and it is difficult to distinguish the various fault types of rolling bearings. So, some time-domain features are extracted from the original signals; the method is as follows:

- (1) X_i ($i = 1, 2, \dots, n$) is the discrete-time series of the i th sensor, and the vibration signals of bearing rotating a cycle are $X_i = [x_1, x_2, \dots, x_m]$, $m = (f \times 60) \div h$,

TABLE 2: Sample distribution of normal and different faults.

Rolling bearing condition	Fault diameter/mm	Fault depth/mm	Training samples	Test samples
Normal	—	—	100	
Inner race fault	0.5	0.3	100	200
Outer race fault	0.5	0.3	100	
Ball fault	0.5	0.3	100	



FIGURE 5: Pictures of the testing bearings: (a) inner race fault; (b) ball fault; (c) outer race fault; (d) normal.

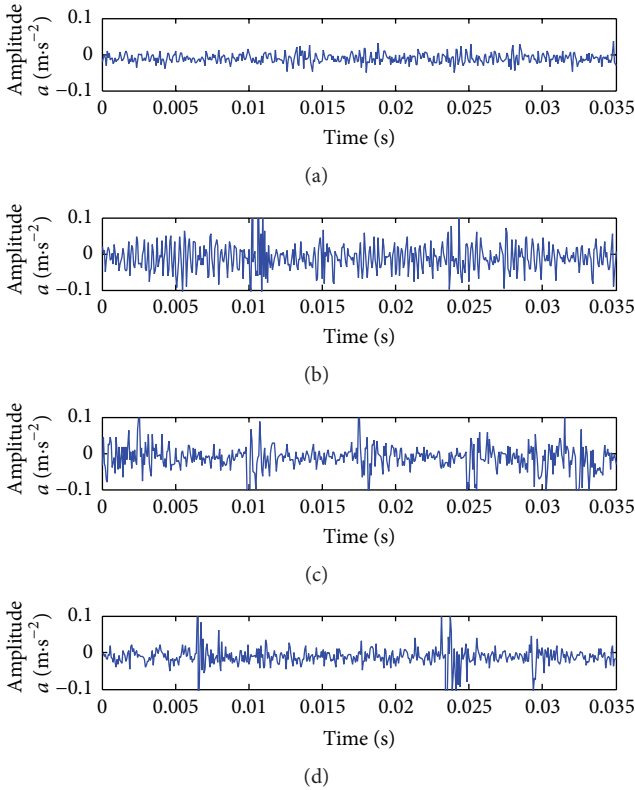


FIGURE 6: Instantaneous amplitudes waveform of rolling bearing in different conditions: (a) normal; (b) inner race fault; (c) outer race fault; (d) ball fault.

where f is the sampling frequency (Hz) and h is the rotational velocity (rpm).

- (2) According to Table 3, compute the time-domain statistics of the vibration signals X_i and get feature vector of the i th sensor as $A_i = [a_1, a_2, \dots, a_{14}]$.
- (3) Three sensors constitute the multisensor feature vector $T = [A_1, A_2, A_3]$.
- (4) Normalize the feature vector $T_i = [t_1, t_2, \dots, t_m]$, $m = n \times 14$.

5. Results and Discussions

5.1. Multisignals and Individual Signals. As the signals pre-processing is clearly explained in Section 4, 14 classical time-domain features are computed from the raw signals.

To illustrate the property of multivibrations fusion, the method based on each single sensor is also measured with the same conditions. These methods are represented as sensors 1, 2, and 3, corresponding to the three individual sensors. The input vectors of DBN in single sensor experiment have only 14 features extracted from vibration signals. The input vectors of DBN in multisensors have 42 features. The DBNs' structures are shown in Table 4.

The DBN structure of multisensors is 42-12-12-4. That is to say, the input layer contained 42 nodes and the output layer included 4 nodes, which depended on the dimensions of the input and output data. There are two hidden layers in the architecture; both hidden layers contained 12 hidden neurons, respectively. However, in sensors 1, 2, and 3 the input nodes are 14 and the hidden nodes are 8 in every hidden layer. The learning rate and momentum are used to adjust the model error and training efficiency. The learning rate in the experiment is selected as 0.01, and the momentum is 0.02 [20]. In the sequential training of every individual RBM, the pretraining of each RBM is accomplished with 20 iterations.

TABLE 3: Statistics features in time domain.

	Parameter	Definition
(1)	Average value \bar{x}	$\frac{1}{n} \sum x_i$
(2)	Maximum value x_{\max}	$\max(x_i)$
(3)	Minimum value x_{\min}	$\min(x_i)$
(4)	Peak-to-peak x_p	$\max(x_i) - \min(x_i)$
(5)	Mean absolute x_{abs}	$\frac{1}{n} \sum x_i $
(6)	Variance x_{var}	$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$
(7)	Standard deviation x_{std}	$\left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{1/2}$
(8)	Kurtosis x_{kur}	$\frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(n-1) x_{\text{std}}^4}$
(9)	Root mean square x_{rms}	$\left(\frac{1}{n} \sum_{i=1}^n x_i ^2 \right)^{1/2}$
(10)	Shape factor x_{sf}	$\frac{x_{\text{rms}}}{\bar{x}}$
(11)	Peak factor x_{pf}	$\frac{x_p}{x_{\text{rms}}}$
(12)	Skewness x_{ske}	$\frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(n-1) x_{\text{std}}^3}$
(13)	Impulse factor x_{if}	$\frac{x_p}{\bar{x}}$
(14)	Margin factor x_{mf}	$\frac{x_p}{x_{\text{rms}}}$

In the fine-tuning of model parameters, stochastic gradient descent (SGD) is used to further reduce the training error and improve the information fusion. However, in this research, the SGD takes the minibatch to globally adjust the parameters in DBN. Since there are 400 samples in the training dataset, the number of minibatch is 10 in the experiments.

We use the training dataset to train the DBN model and use the testing dataset to test the model identification accuracy. The classification process is repeated for 25 times and the classification results are averaged as shown in Figure 7.

The average accuracy of the training samples in multisensors is 97.5%, and the number of correct classification samples is 390. This is much higher than those using other methods, which are 91.5%, 85%, and 87.5% with 366, 340, and 350, respectively. The average accuracy of testing samples in multisensors fusion is 95.5%, and the number of correct classification samples is 191. This is much higher than those using other methods, which are 89%, 78.5%, and 75% with 178, 157, and 150, respectively.

Compared to those individual sensors, the training and testing accuracies of multisensors information fusion are obviously higher than other methods. In the three individual sensors, the classified accuracy of S_1 is better than S_2 and S_3 . It is indicated that the sensor put on the vertical location of the testing pedestal is most sensitive to the bearing faults. The results prove that it is more effective to integrate the signals from multisensors than to use the vibration from individual ones.

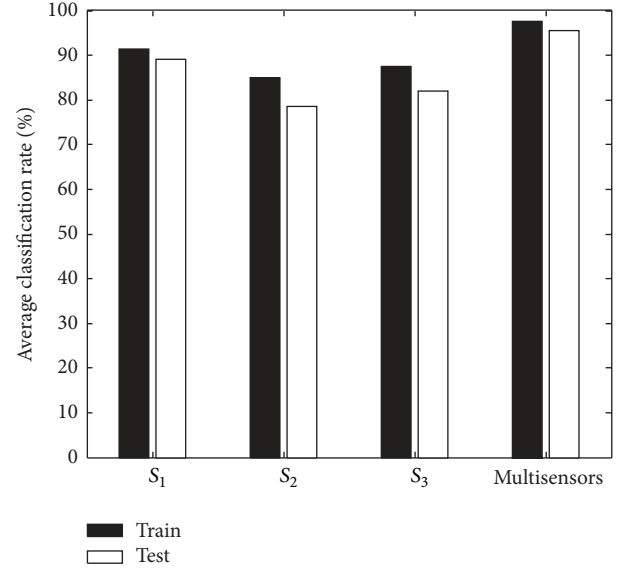


FIGURE 7: Classification rate of individual sensor and multisensors.

5.2. Deep Belief Network and Traditional Methods. To introduce the effectiveness of the suggested method, the comparative methods are also tested with the same data such as SVM, BPNN, and KNN. Without manual feature selection, the 42-dimension feature vector is used directly for rolling bearing fault diagnosis. The architecture of BPNN is 42-12-4; the decision function of KNN is Euclidean distance; the SVM used RBF kernel, which is the most popular and suitable kernel function applied. The penalty factor and the radius of the SVM kernel are 0.216 and 0.3027, each of which is determined through genetic algorithm. All the experiments are finished in the computer with Core i5 Intel central processing unit and 4G memory. In the DBN, SVM, BPNN, and KNN, the experiments are, respectively, repeated 15 times. The average classification accuracies are shown in Tables 5 and 6.

In the training experiments, DBN achieved 95.72% identification accuracy which is better than that of SVM (92.28%), KNN (90.06%), and BPNN (83.63%) for multivibration signals. For the testing experiments, the average accuracy of DBN is 93.17%; meanwhile SVM is 90.13%, KNN is 85.23%, and BPNN is 78.13%. In brief, the experiments results explain that the suggested methods have higher reliability and better accuracy than SVM, KNN, and BPNN in rolling bearing fault diagnosis.

The algorithms of DBN, SVM, KNN, and BPNN all trained the model by the same dataset and generated the classifier to carry out the bearing fault diagnosis. However, the stability and generalization ability are different in these methods. As shown in Tables 5 and 6, the training results of DBN approximately agree with the testing results in the 15 experiments. The testing accuracies in SVM and KNN are lower than training accuracies by 3–5%. The classification accuracy of BPNN is decreased obviously in test experiments.

In the experiments, the training datasets are selected from sequence data samples of all kinds of fault condition,

TABLE 4: The parameters of DBN in sensors 1–3 and multisensors.

Parameters	Sensor 1	Sensor 2	Sensor 3	Multisensors
Input	14	14	14	42
Output	4	4	4	4
Number of hidden layers	2	2	2	2
Number of neurons in a layer	8	8	8	12
Iteration	20	20	20	20
Minibatch size	10	10	10	10
Momentum	0.02	0.02	0.02	0.02
Learning rate	0.01	0.01	0.01	0.01

TABLE 5: Results of training datasets.

Method	Experiment results of each group (%)															Average (%)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
DBN	95.25	95	95.5	96	95	97.25	95	95.5	95.25	96	97	96.25	95.5	96	95.25	95.72
SVM	94.5	90	91	94	92.5	91	92.5	94	93.5	91	89.5	91	92	94.5	93.25	92.28
KNN	90	88	87.25	87	93.75	89.75	90	89	92	92	92.5	90.25	90	89	90.5	90.06
BPNN	84.5	82.5	80	80.5	82.5	84	86	80	82.5	85	86.5	84	83	88.5	85	83.63

TABLE 6: Results of testing datasets.

Method	Experiment results of each group (%)															Average (%)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
DBN	93.5	95	93.5	95.5	94.5	96.5	94.5	94.5	94	93.5	96	94.5	95	96	94.5	94.73
SVM	91	89	90	90.5	91	89.5	90	90	91	90	89	90	88.5	92	90.5	90.13
KNN	85	86	82.5	84.5	87.5	85.5	88	84.5	86	84.5	85	84	85.5	84.5	85.5	85.23
BPNN	82.5	80	78.5	77.5	75.5	75	74.5	78	75.5	78	76.5	78	80	80.5	82	78.13

and they orderly composite the training sample datasets. As the testing samples are randomly selected from the various states datasets, both sample category and sample order are random. SVM, to classify data identification, mainly depends on the kernel functions with the training set, which is closely related to the quantity and quality dataset. When using the KNN to classify the data identification, the results mainly are determined by the distance function, and once the distance function is selected, it will not be able to transform. So, the testing accuracies are much less than in training set of SVM and KNN. BPNN is a typically shallow learning model, which involved no more than one nonlinear feature transformation and has difficulty in representing complex functions with poor performance and generalization ability.

Compared with the traditional machine learning and signal processing technology, DBN has the merit to get rid of the dependence on signal processing technology. On the other hand, DBN can adaptively extract the fault feature without restrictive assumptions or complex parameter adjustment. Consequently, it is nothing strange that the DBN as a promising method has been effectively applied in multivibrations fusion.

6. Conclusions

Multiple sensors installed on various locations of bearing pedestal can supply abundant information for fault diagnosis

and detection. Based on this observation, a novel technique using deep belief network for the multivibrations fusion is put forward in this paper. Some conventional time-domain features are extracted from three accelerometer vibration sensors. Without manual feature selection, the features are used directly as the input vectors of the DBN. The obtained accuracy of multisensors is 97.5% which is about 10% higher than single sensor. At the same time, the mean accuracy of DBN, SVM, KNN, and BPNN is, respectively, 93.17%, 90.13%, 85.23%, and 78.13%. It suggests that DBN is more effective and stable for the identification of rolling bearing fault diagnosis than other methods. From the results, it can be realized that DBN is able to adaptively integrate available fault features from multisensors and it obtained higher identification accuracy than traditional methods.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This paper was supported by Nature Science Fund of China, under Grants nos. 61402167 and 51375500, and the National Science and Technology Support Project of China, under Grant no. 2015BAF32B01.

References

- [1] Y. Yang, D. Yu, and J. Cheng, "A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM," *Measurement*, vol. 40, no. 9-10, pp. 943-950, 2007.
- [2] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6077-6085, 2010.
- [3] I. El-Thalji and E. Jantunen, "A summary of fault modelling and predictive health monitoring of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 252-272, 2015.
- [4] G. Wang, Z. He, X. Chen, and Y. Lai, "Basic research on machinery fault diagnosis-what is the prescription," *Journal of Mechanical Engineering*, vol. 49, no. 1, pp. 63-72, 2013.
- [5] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: a review with applications," *Signal Processing*, vol. 96, pp. 1-15, 2014.
- [6] 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.
- [7] D. Yang, Y. Liu, S. Li, X. Li, and L. Ma, "Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm," *Mechanism and Machine Theory*, vol. 90, pp. 219-229, 2015.
- [8] A. Moosavian, H. Ahmadi, A. Tabatabaefar, and M. Khazaei, "Comparison of two classifiers; K-nearest neighbor and artificial neural network, for fault diagnosis on a main engine journal-bearing," *Shock and Vibration*, vol. 20, no. 2, pp. 263-272, 2013.
- [9] P. Shakyia, A. K. Darpe, and M. S. Kulkarni, "Bearing diagnosis using proximity probe and accelerometer," *Measurement*, vol. 80, pp. 190-200, 2016.
- [10] M. S. Safizadeh and S. K. Latifi, "Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell," *Information Fusion*, vol. 18, no. 1, pp. 1-8, 2014.
- [11] X. Li, D. Yang, D. Guo, and L. Jiang, "Fault diagnosis method based on multi-sensors installed on the base and KPCA," *Chinese Journal of Scientific Instrument*, vol. 32, no. 7, pp. 1551-1557, 2011.
- [12] M. Dong and D. He, "Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis," *European Journal of Operational Research*, vol. 178, no. 3, pp. 858-878, 2007.
- [13] T. P. Banerjee and S. Das, "Multi-sensor data fusion using support vector machine for motor fault detection," *Information Sciences*, vol. 217, no. 24, pp. 96-107, 2012.
- [14] L.-L. Jiang, Y.-L. Liu, X.-J. Li, and A.-H. Chen, "Gear fault diagnosis based on SVM and multi-sensor information fusion," *Journal of Central South University*, vol. 41, no. 6, pp. 2184-2188, 2010.
- [15] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85-117, 2015.
- [16] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504-507, 2006.
- [17] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527-1554, 2006.
- [18] D. Yu, L. Deng, I. Jang, P. Kudumakis, M. Sandler, and K. Kang, "Deep learning and its applications to signal and information processing," *IEEE Signal Processing Magazine*, vol. 28, no. 1, pp. 145-154, 2011.
- [19] R. Sarikaya, G. E. Hinton, and A. Deoras, "Application of deep belief networks for natural language understanding," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 4, pp. 778-784, 2014.
- [20] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" *Journal of Machine Learning Research*, vol. 11, pp. 625-660, 2010.
- [21] H. Shao, H. Jiang, X. Zhang, and M. Niu, "Rolling bearing fault diagnosis using an optimization deep belief network," *Measurement Science and Technology*, vol. 26, no. 11, Article ID 115002, 2015.
- [22] P. Tamilselvan, Y. Wang, and P. Wang, "Deep belief network based state classification for structural health diagnosis," *Reliability Engineering and System Safety*, vol. 115, no. 3, pp. 124-135, 2013.
- [23] M. Gan, C. Wang, and C. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 92-104, 2016.
- [24] Y. Lei, F. Jia, X. 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.
- [25] 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.
- [26] G.-Y. Pan, W. Chai, and J.-F. Qiao, "Calculation for depth of deep belief network," *Control and Decision*, vol. 30, no. 2, pp. 256-260, 2015.

