

# Rolling Element Bearing Fault Diagnosis based on Deep Belief Network and Principal Component Analysis

Guangxing Niu<sup>1</sup>, Bin Zhang<sup>2</sup>, Paul Ziehl<sup>3</sup>, Frank Ferrese<sup>4</sup>, Michael Golda<sup>5</sup>

<sup>1,2,3</sup> *University of South Carolina, Columbia, South Carolina, 29208, USA*

*gniue@mail.sc.edu*

*zhangbin@cec.sc.edu*

*ziehl@cec.sc.edu*

<sup>4,5</sup> *Naval Surface Warfare Center, Philadelphia Division (NSWCPD), Philadelphia, PA*

*frank.ferrese@navy.mil*

*eugene.m.golda1@navy.mil*

## ABSTRACT

Rolling element bearings are critical components in industrial rotating machines. Faults and failures of bearings can cause degradation of machine performance or even a catastrophe. Bearing fault diagnosis is therefore essential and significant to the safe and reliable operation of systems. For bearing condition monitoring, acoustic emission (AE) signals attract more and more attention due to its advantages on sensitivity over the extensively used vibration signal. In bearing fault diagnosis and prognosis, feature extraction is a critical and tough work, which always involves complex signal processing and computation. Moreover, features greatly rely on the characteristics, operating conditions, and type of data. With consideration of changes in operating conditions and increase of data complexity, traditional diagnosis approaches are insufficient in feature extraction and fault diagnosis. To address this problem, this paper proposes a Deep Belief Network (DBN) and Principal Component Analysis (PCA) based fault diagnosis approach using AE signal. This proposed approach combines the advantages of deep learning and statistical analysis, DBN automatically extracts features from AE signal, PCA is applied to dimensionality reduction. Different bearing fault modes are identified by least squares support vector machine (LS-SVM) using the extracted features. An experimental case is conducted with a tapered roller bearing to verify the proposed approach. Experimental results demonstrate that the proposed approach has excellent feature extraction ability and high fault classification accuracy.

## 1. INTRODUCTION

Rolling element bearings are key components in mechanical systems. They are subject to various stresses, transmissions and shocks, which may cause bearing fault, and eventually lead to system breakdown. The degradation of bearing condition will definitely affect the performance of the systems. To prevent unexpected system failure and reduce the maintenance cost, bearing fault diagnosis is desired to detect fault as early as possible.

A rolling element bearing is composed of rolling elements, an inner race, an outer race, and a cage. The faults can appear on any component, and these faults can be roughly divided as local faults and distributed faults (Zhang et al., 2008; Cerrada et al., 2018). Local faults are defined as a single localized fault, such as pitting, scratch, crack, hole, etc. Distributed faults are defined as irregularities of bearing structure, such as misalignment of shaft or races, eccentric races, off-size rolling elements, roughness, etc. These faults can be caused by many reasons, such as, overheating and load, improper installation, imperfect manufacturing. These distributed faults will cause excessive contact force and friction, which will finally lead to local faults. When a bearing operates under fault conditions, they cause certain characteristic signals in the form of sound, vibration, energy, or acoustic emission.

With the advancement of machine condition monitoring techniques, many different types of signals such as vibration, acoustic emission, ultrasound have been used for diagnosis (Rai & Upadhyay, 2016; Zhang et al., 2010; Li et al., 2018). Among these signals, acoustic emissions are the transient elastic waves, which are generated from a rapid release of localized stress energy caused by deformation or defect within or on the surface of a material (Al-Ghamd & Mba, 2006). Compared with most widely used vibration signals,

---

Guangxing Niu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

AE signals have many advantages: 1) insensitive to mechanical disturbances and noises caused by different operating conditions; 2) sensitive to fault size, which can offer earlier fault detection than vibration signals. These advantages make AE signals promising in bearing fault diagnosis.

Over the past decades, a lot of efforts have been presented for bearing fault diagnosis. These existing works are mainly divided into signal processing based approach and learning based approach (Cerrada et al., 2018; S. Guo, Yang, Gao, Zhang, & Zhang, 2018). For signal processing based approaches, feature extraction is needed to extract a fault indicator that is related to fault modes and fault state. Zhao et al. (S. Zhao, Liang, Xu, Wang, & Zhang, 2013) applied Empirical Mode Decomposition and Approximate Entropy based approach to detect different fault modes. Khanam et al. (Khanam, Tandon, & Dutt, 2014) employed Discrete Wavelet Transform (DWT) to decompose signals and estimate ball bearing faults. Hilbert transform (Bujoreanu, Monoranu, & OLARU N, 2014), matching pursuit method (Cui, Zhang, Zhang, & Lee, 2016), spectral analysis and statistical analysis (Gerber, Martin, & Mailhes, 2015) and envelop analysis (Sun, Guo, & Gao, 2015) are also very effective signal processing techniques in bearing monitoring.

In practical applications, however, fault characteristic signals are often corrupted by noises, which makes feature extraction and fault diagnosis difficult and challenging. For different monitoring signals and different fault modes, no generic feature extraction method is available. As a result, feature extraction is ad-hoc for different systems and this process is time-consuming, requires complex signal processing techniques, and needs extensive human involvement. All these limitations severely hinder the development of applications of signal processing based diagnosis approaches.

Learning-based approach, on the other hand, aims to learn potential signal patterns that are related to different fault modes or fault levels (Wang, Xiang, Zhong, & Zhou, 2018). Moreover, learning-based approaches, as a supervised learning process, require data including fault samples and their labels. The process involves feature extraction and feature classification. Feature extraction can be conducted in the time domain, the frequency-domain, and transform-domains, such as statistical parameters, signal energy of Intrinsic Mode Functions (IMFs) from Empirical Mode Decomposition (EMD), Discrete Wavelet Transform (DWT), Hilbert-Huang transform, etc. With feature extracted, a classifier is designed to classify features under different fault conditions for fault detection. Some widely used classifiers include neural networks, Support Vector Machines (SVM), and Bayesian estimation, among others.

Recently, with the successes of deep learning in image recognition and speech processing, many deep learning based fault diagnosis approaches were proposed for many applications

(Cococcioni, Lazzarini, & Volpi, 2013; Tang et al., 2018; Chen & Li, 2017). Guo et al. proposed a continuous wavelet transform scalogram (CWTS) and convolutional neural network (CNN) based approach for rotating machinery fault diagnosis (S. Guo, Yang, Gao, & Zhang, 2018). Qin et al. presented an optimized DBN and improved logistic sigmoid unit based fault diagnosis for planetary gearboxes of wind turbines (Qin, Wang, & Zou, 2019). These applications show that deep learning has great potentials in feature extraction and data mining.

These reported learning based approaches, although achieved good performance in some aspects, are insufficient in dealing with bearings under complex operating conditions and the continuously increase of volume and complexity of monitoring data. As a result, diagnosis accuracy can be affected. Another limitation of these proposed approaches is that signal transformation or feature extraction are also involved. Thus, the strong feature extraction and learning ability of deep learning cannot be fully exploited. Inspired by these limitations, a deep learning and statistics based approach is proposed in this paper. This proposed approach integrates DBN and PCA to extract features from raw data and, therefore, can avoid complex signal processing and corresponding feature extraction. LS-SVM is an effective pattern recognition approach, which is an extension and improvement of SVM. It has been widely used in fault diagnosis, and has shown excellent fault identification capability. Due to this feature, it is employed to process the extracted features for diagnosis in this paper. Experimental result shows that it can achieve high accuracy.

This paper is organized as follows. A brief introduction of DBN, PCA and LS-SVM is presented in Section 2. Section 3 describes the fault diagnosis steps of the proposed approach. Experiments are presented and results are analyzed and visualized in Section 4 to demonstrate the performance of the proposed approach. Finally, Section 5 provides concluding remarks and some future research directions.

## 2. PRINCIPLE OF RELATED APPROACHES

### 2.1. DBN principle

DBN can be regarded as a special neural network constructed from multiple Restricted Boltzmann Machines (RBMs) (G. Zhao et al., 2017). Fig. 1 shows the schematic representation of a two hidden layer DBN. DBN has a strong capability in capturing representative information from raw time series data. The output of the learning information is extracted features, which can be utilized as the input of supervised learning algorithms in classification or regression for fault diagnosis.

RBM is a special probabilistic model of Boltzmann machine,

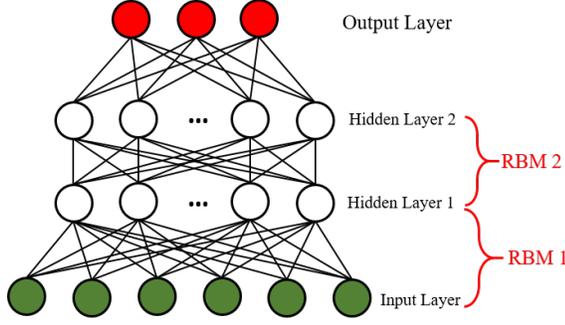


Figure 1. The structure of a 2 hidden layer DBN

which consists of a visible input vector  $v$  and a hidden vector  $h$ , connected by weighting factors, as shown in Fig. 2.

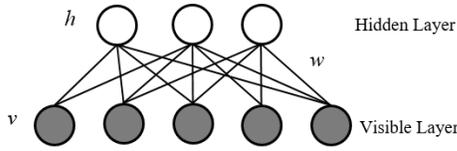


Figure 2. Structure of Restricted Boltzmann Machine

The joint configuration  $(v, h)$  can be given by the energy function (1).

$$E(v, h) = - \sum_{j=1}^m a_j v_j - \sum_{i=1}^n b_i h_i - \sum_{i=1}^n \sum_{j=1}^m v_j w_{ij} h_i \quad (1)$$

where  $v_j$  and  $a_j$  are the binary states and bias of the  $j$ -th element of the visible vector,  $h_i$  and  $b_i$  are the binary states and bias of the  $i$ -th element of the hidden vector,  $w_{ij}$  is the weight of the connection between the visible layer and the hidden layer. The joint distribution over the visible layer and hidden units is defined as

$$p(v, h) = \frac{1}{Z} e^{-E(v, h)} \quad (2)$$

where  $Z$  is a partition function, which can be described as

$$Z = \sum_{v, h} e^{-E(v, h)} \quad (3)$$

In DBN structure, the connections just exist between the visible layer and the hidden layer. The neurons in the same layer are independent with each other. The conditional probabilities of the hidden layer and the visible units are given as

$$p(h_i = 1|v) = \frac{1}{1 + e^{-b_i - \sum_{j=1}^m v_j w_{ij}}} \quad (4)$$

$$p(v_i = 1|h) = \frac{1}{1 + e^{-a_i - \sum_{j=1}^m h_j w_{ij}}} \quad (5)$$

The learning process of DBN can be divided into two stages: pre-training and fine-tuning (Hinton, Osindero, & Teh, 2006). In the pre-training process, the RBMs are trained layer by layer with an unsupervised manner. The forward pre-training process can be regarded as a construction and reconstruction process using Eq. (1). After all the RBMs in the DBN are pre-trained, the fine-tuning step will be applied to DBN using backpropagation algorithm (G. Zhao et al., 2018). In this fine-tuning process, the weights and biases of every layer are adjusted continuously until the error becomes smaller than predefined threshold. The trained DBN model is obtained after the fine-tuning step and can be used in describing the fault dynamics.

As mentioned earlier, the training procedure includes pre-training and fine-tuning. Pre-training stage aims to extract features based on its learning rules automatically. In pre-training, the stacked RBMs are trained layer by layer using greedy learning algorithm (Hinton et al., 2006). This is an unsupervised training process.

Given training input data and the initialization parameters, the first hidden layer can be trained greedily by (4) and (5). This process is a positive phase, in which the gradient of the log probability of the given training data can be described as:

$$\frac{\partial \log p(v^0)}{\partial \omega_{ij}} = \langle v_i^0 h_j^0 \rangle - \langle v_i^k h_j^k \rangle \quad (6)$$

The learning rule aims at maximizing the log probability of the data, which is equal to minimizing the divergence of the distribution defined by the model and the given training data. Based on the Contrastive Divergence (CD) algorithm (Hinton et al., 2006), in the training process, the parameters of DBN can be adjusted by

$$\omega_{ij}^{new} = \omega_{ij} + \gamma(\langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k) \quad (7)$$

$$a_{ij}^{new} = a_{ij} + \gamma(\langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k) \quad (8)$$

$$b_{ij}^{new} = b_{ij} + \gamma(\langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_k) \quad (9)$$

where  $\gamma \in [0, 1]$  denotes the learning rate, which can be used to adjust the learning speed.

The fine-tuning process is conducted to optimize the pre-trained network parameters and this supervised learning process will further adjust the structure to improve the classification accuracy. Conjugate gradient algorithm is applied to fine-tune the trained parameters using the labeled data. In this step, all parameters are updated at the same time until the fine-tuning threshold is reached. The trained DBN model can be got after these two steps are finished.

## 2.2. PCA principle

Principal component analysis is a traditional statistical analysis approach which can be used to get principal components.

This technology has been widely used in fault diagnosis for data dimension reduction, feature extraction and fusion, visualization of high-dimensionality, data regression, etc (Chang, Wang, Liu, & Wang, 2008; T. Guo, Zhou, Zhang, Chen, & Tai, 2019).

PCA can be assumed as a transformation that projects original data to a new space with lower dimension (Chang et al., 2008). Given original data vector  $x_i (i = 1, \dots, m)$ , the covariance matrix of the data vector can be calculated as:

$$Cov = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)(x_i - \mu)^T \quad (10)$$

where  $\mu$  is the mean value of the vector,  $\mu = \frac{1}{m} \sum_{i=1}^m x_i$ . Assume that the origin data vector is a  $n$  dimensional data, the eigenvalue of the covariance matrix can be described as:

$$\lambda_j u_j = Cov u_j \quad (j = 1, \dots, n) \quad (11)$$

where  $\lambda_j$  are the eigenvalues of the covariance matrix, which are sorted in descending order and  $u_j$  are the corresponding eigenvectors.

To get the first  $k$  eigenvectors ( $k < n$ ) that corresponding to the  $k$  largest eigenvalues, let

$$U_k = [u_1, u_2, \dots, u_k] \quad (12)$$

$$\Lambda_k = diag[\lambda_1, \lambda_2, \dots, \lambda_k] \quad (13)$$

The principal components of the original data can be computed as the orthogonal transformations of  $x_i$ :

$$P = U_k^T x_i \quad (14)$$

The obtained components are named as principal components. Dimensional reduction can be achieved by using the first several eigenvectors of the eigenvectors. In this paper, the distribution of the extracted features are assumed as Gaussian, this work mainly uses the characteristic of dimensional reduction of PCA.

### 2.3. LS-SVM principle

LS-SVM is an extension of SVM. It changes the inequality constraints in SVM to equality constraints, which transforms the quadratic problems in SVM into linear equations problems. Compared with traditional SVM, LS-SVM has higher operation efficiency and solution accuracy.

Given training dataset  $(x_i, y_i)$  with  $x_i$  being the input vector and  $y_i$  being its corresponding output label, we define  $\chi$  as the corresponding feature vector that can be used to map the input vector into a new feature space. Then a hyperplane can be described as:

$$w \cdot \chi(x_i) + b = 0 \quad (15)$$

where  $w$  denotes the weight of the orientation of the hyperplane,  $b$  is the bias. LS-SVM based identification problem can be regarded as the following optimization problem (Liu, Bo, & Luo, 2015):

$$\min J(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^N e_i^2 \quad (16)$$

$$\text{s.t. } y_i = w^T \chi(x_i) + b + e_i \quad i = 1, \dots, N \quad (17)$$

where  $J$  is the objective function,  $\gamma$  is the tradeoff coefficient. Sample  $x_i$  can be projected into high-dimensional space by nonlinear mapping  $\chi$ . To minimize the objective function, the first step is to define the corresponding Lagrange function:

$$L(W, b, e, \beta) = J(w, e) - \sum_{i=1}^N \beta_i \{w^T \chi_i + b + e_i - y_i\} \quad (18)$$

where  $\beta_i$  is a Lagrange multiplier. The conditions for optimality are given as:

$$\begin{aligned} \partial L / \partial w &= 0 \\ \partial L / \partial b &= 0 \\ \partial L / \partial e_i &= 0 \\ \partial L / \partial \beta_i &= 0 \end{aligned} \quad (19)$$

Eliminate  $w$  and  $e$ , a linear equation can be obtained as:

$$\begin{bmatrix} 0 & l^T \\ l & \Phi + I/\gamma \end{bmatrix} \begin{bmatrix} b \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (20)$$

where  $y = [y_1, y_2, \dots, y_N]^T$ ;  $l = [1, 1, \dots, 1]^T$ ;  $\beta = [\beta_1, \beta_2, \dots, \beta_N]^T$ ;  $\Phi_{i,j} = (x_i)^T (x_j)$ ,  $I$  is the unit matrix. Finally, LS-SVM classification decision-making model can be described as:

$$\hat{y}(x) = \sum_{i=1}^N \beta_i K(x, x_i) + b \quad (21)$$

where  $K(x, x_i) = e^{-(|x_i - x_j|^2 / \sigma^2)}$  is the Radial Basis Function (RBF) kernel function to be used in this paper, and  $\sigma$  is the kernel bandwidth.

Bearing fault diagnosis is an multi-classification problem in which the classification model is actually constructed by combining multiple two-class SVM classifiers. In this paper, one-against-one SVM is applied to identify different fault modes.

### 3. DBN-PCA-LSSVM BASED BEARING DIAGNOSIS

This research presents a bearing diagnosis using AE signals. Fig. 3 illustrates the implementation of the proposed diagnostic approach. The features of different fault modes are extracted by DBN from raw data automatically. PCA is then ap-

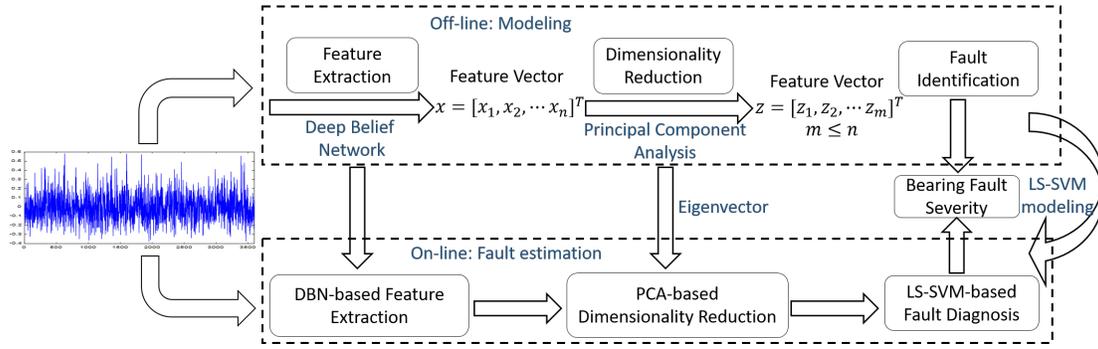


Figure 3. DBN-PCA-LSSVM based fault diagnosis

plied for dimensionality reduction, LS-SVM based fault classification and identification are finally implemented based on the extracted features for fault diagnosis.

The proposed approach includes offline modeling and online testing processes. In the modeling process, acoustic emission signals collected from bearings are pre-processed and fed into the initialized DBN structure, through which features can be extracted automatically layer by layer.

The detailed implementation diagnostic procedure steps are described as follows:

Step 1: Partition the time series data into segments based on the data sampling rate and roller bearing operation state, and divide the data into training data and testing data.

Step 2: Define the DBN structure, and train the DBN using the training data set. In this step, the number of hidden layers, the number of neurons in each hidden layer, the learning rate, and the initial weights should be defined. The training process is finished when the performance meets pre-defined requirements or iteration number reaches the threshold.

Step 3: PCA is applied on the extracted features ( $[x_1, x_2, \dots, x_n]$ ) from DBN to reduce dimensionality. Low-dimensional training features ( $[z_1, z_2, \dots, z_m]$ ,  $m < n$ ) can be obtained in this step. Besides, a projection matrix, which will be used in the testing process, is also obtained.

Step 4: Optimize LS-SVM by using training labels and the low-dimensional features obtained in Step 3 to identify different bearing fault modes.

Step 5: Test the trained model with the testing data set, trained DBN model, eigenvector of PCA, and trained LS-SVM. Analyze the performance on test data.

This proposed DBN-PCA-LSSVM based bearing diagnostic approach can extract features from raw data automatically, classify the fault mode, and estimate the fault severity. It avoids complex signal processing and human involvement in feature selection and extraction, which makes it more applicable, and easier to be extended to other applications.

## 4. EXPERIMENTAL RESULTS

### 4.1. Acoustic emission data preparation

The bearing used for the verification of the proposed approach is a tapered roller bearing: Timken LM501310 cup and LM501349 cone. The structure of the rolling bearing is described in Fig. 4 (Zhang et al., 2008) and the main geometric parameters are listed in Table 1. The data was collected from the AE sensors with the sampling rate of 50kHz, and the experimental test was conducted under different fault sizes with three different rotating speeds and three different loads, which makes nine different data sets for each fault size. The details of the collected AE data are described in Table 2.

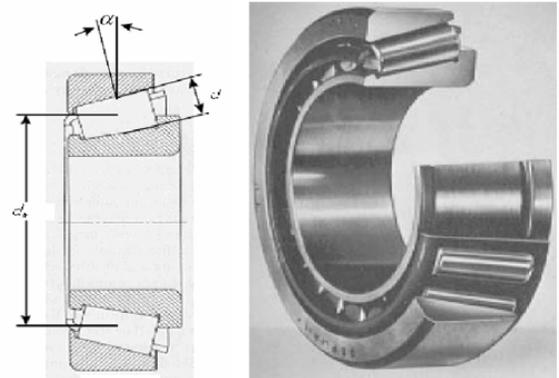


Figure 4. Tapered Roller Bearing Structure and Photo (Zhang et al., 2008)

Table 1. The geometric parameters of bearing

Parameter	Value	Parameter	Value
Pitch diameter	5.715cm	Roller diameter	0.784cm
Roller number	19	Contact angle	13.13°

For each test, the fault size is given by depth and width measured by microns to show fault severity. To utilize the data efficiently, the fault size is defined as the sum of width and

Table 2. AE testing data description

Test No.	Width	Depth	Speed(rpm)	Load(psi)
C0	0	0		
C1	35.33	2.46	800	200
C2	37.67	10.56		
C3	48.33	2.38	1200	400
C4	49.33	4.88		
C5	61.00	5.80	1600	600
C6	64.00	11.00		
C7	131.3	1.40		
D1	64	11	800	200
D2	64	11	1200	400
D3	64	11	1600	600

depth. For example, for C1 in Table 2 with fault width and depth of 35.33 and 2.46, respectively, the fault dimension is measured as  $35.33+2.46=37.79$ , as shown in Table 3. Based on this fault size definition, the fault mode is defined from F-1 (Health) to F-6 as shown in Table 3.

Table 3. Fault mode discretization

Dataset No.	Fault dimension	Fault mode
C0,G2	0	F-1(Health)
C1	37.79	F-2
C2,C3,C4	48.23~54.21	F-3
C5	66.8	F-4
C6,D1,D2,D3	75	F-5
C7	132.7	F-6

#### 4.2. Feature extraction and fault classification

Table 4. Training parameters

Parameter	Value
Iteration of each RBM	300
Learning rate	0.1
Fine-tuning iteration	1200
Momentum	0.9

For bearing fault diagnosis, the information in one cycle of rotation should be included in the input vector of DBN. For this reason, consider the lowest rotating speed, the input vector size is set as 3750. The DBN structure is given as 3750-1500-600-200. Other training parameters are given in Table 4. Each fault mode has 2250 samples in which 60% are used for training and the remaining 40% are used for testing. To

train the DBN with a fair way, each raw data set is separated into several segments, which are randomly selected to construct the training set and testing set.

To show the feature extraction performance, principle component analysis (PCA) is applied on the output of each layer. Only the first three main components are visualized in the 3-D space to make the performance of feature extraction clear. Fig. 5 is the visualization of raw AE data while Fig. 6 is the visualization of the extracted features by DBN.

Fig. 5 shows that the raw data of all the fault modes have severe degrees of overlapping. In other words, it will be very difficult to classify different fault modes and severity from raw data. With DBN learning outputs at different layers, the degrees of overlapping of the outputs from each DBN layer is decreasing. At the output layer, the extracted features are almost separated, as visualized in Fig. 6.

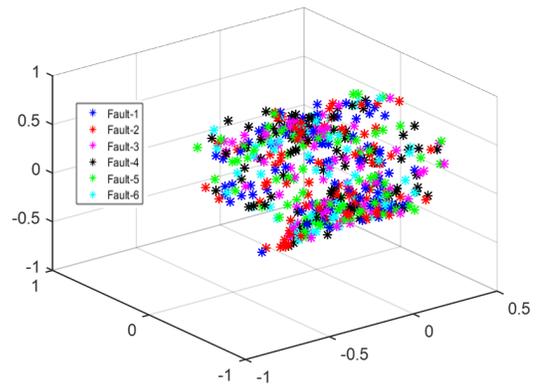


Figure 5. Raw data

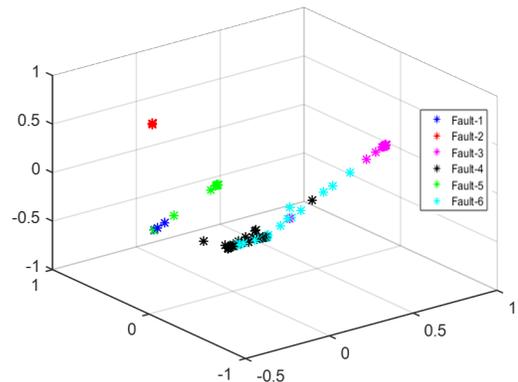


Figure 6. Visualization of extracted features

After the feature extraction step, features are obtained for training classifier. The training process is shown in Fig. 7. It is clear that the classification accuracy converges to a constant

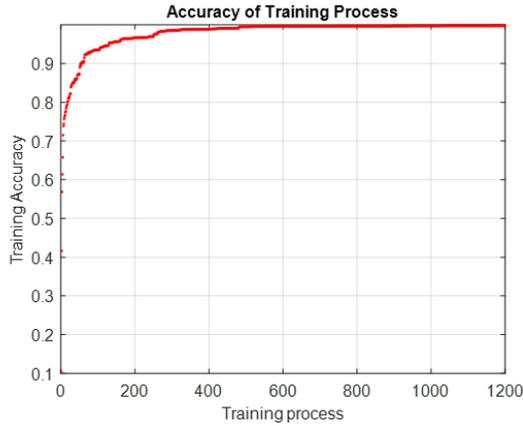


Figure 7. Training process

after about 500 iterations. Based on the extracted features, an accuracy of 93.46% is achieved. This result shows that DBN is an effective and powerful feature extraction approach for bearing fault diagnosis.

However, the accuracy and efficiency of fault diagnosis may get impacted due to the high dimensionality of the extracted features. PCA is applied to the extracted features to reduce the dimensionality of the feature vectors before performing fault identification. PCA can also cluster the main characteristics of different fault modes and increase the spatial distance among them. LS-SVM training process is implemented on the features after dimension reduction to obtain the optimal classification model. The dimension size for PCA is an important parameter that affects diagnosis results. Fig. 8 shows the diagnosis results obtained by the proposed DBN-PCA-LSSVM approach with dimension sizes of PCA changes from 10 to 200 with an increase step of 10. It is clear that the highest accuracy of 95.48% is achieved when the PCA dimension size is 50.

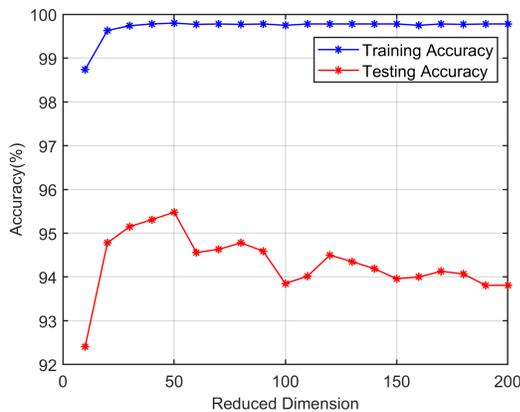


Figure 8. Diagnosis results obtained by DBN-PCA-LSSVM with different dimension sizes for PCA

Table 5 analyzed the misclassification results of each fault mode. The misclassification samples mainly appear on F-4 and F-6, 104 samples from F-4 are misclassified as F-6, and 75 samples from F-6 are misclassified as F-4. These results are also consistent with the results in Fig. 6, in which the degrees of overlapping of F-6 (cyan samples) and F-4 (black samples) are larger than others. Here is an example: One sample of F-6 is incorrectly classified as F-4. Based on the outputs of DBN, the probabilities of this sample belongs to F-4 and F-6 are 0.5909 and 0.4087, respectively. Both of the probabilities are not large enough for the classifier to make a strong decision. In this case, misclassification occurs. The potential causes and advanced approach to make correct classification will be further investigated in the future work.

## 5. CONCLUSIONS AND FUTURE WORK

This paper presents a DBN-PCA-LSSVM based rolling element bearing fault diagnosis approach. In the proposed approach, DBN is developed to extract feature automatically from raw sensor data, PCA is applied to reduce the dimensionality of the extracted features, and LS-SVM is performed to identify different fault modes. The rolling bearing fault diagnosis experiment is presented to validate the proposed approach.

The contributions of this paper are as follows: 1) By combining PCA with DBN, an integrated, accurate and intelligent fault diagnosis is proposed, in which DBN and PCA are used for bearing feature extraction, and LS-SVM is used for fault diagnosis. This integration takes full advantages of strong feature learning ability of DBN and statistic analysis of PCA. 2) Features with different dimension sizes are analyzed to find the optimal feature size for LS-SVM, which can achieve the highest classification accuracy.

With a DBN structure designed, the raw AE data are fed into the DBN and features are extracted automatically for fault diagnosis. PCA is applied to reduce the dimension of the extracted, then improve the accuracy and efficiency of fault classification. The experimental results show that this approach can achieve high diagnostic accuracy. From this perspective, the proposed DBN-PCA-LSSVM based approach provides a generic solution that can be applied to a variety of systems. Compared with the traditional signal processing and machine learning based approaches, the proposed method does not require complex signal processing techniques and human involvement.

The future research work will mainly focus on optimizing the structure and analyzing the misclassification causes of the proposed approach to enhance fault diagnosis accuracy and efficiency.

Table 5. Fault diagnosis results

Actual fault mode \ Diagnosis result	Diagnosis result						Accuracy(%)
	F-1	F-2	F-3	F-4	F-5	F-6	
F-1	894	0	0	0	6	0	99.3
F-2	0	900	0	0	0	0	100.0
F-3	0	0	891	0	0	9	99.0
F-4	0	0	0	796	0	104	88.4
F-5	25	0	0	0	875	0	97.2
F-6	0	0	25	75	0	800	88.9
Overall	—	—	—	—	—	—	95.5

### ACKNOWLEDGMENT

The work is partially supported by the NEEC project under contract N00174-17-1-0006 and ASPIRE grant program at the University of South Carolina.

### REFERENCES

- Al-Ghamd, A. M., & Mba, D. (2006). A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size. *Mechanical systems and signal processing*, 20(7), 1537–1571.
- Bujoreanu, C., Monoranu, R., & OLARU N, D. (2014). Study on the defects size of ball bearings elements using vibration analysis. *Applied Mechanics & Materials*(658).
- Cerrada, M., Sánchez, R.-V., Li, C., Pacheco, F., Cabrera, D., de Oliveira, J. V., & Vásquez, R. E. (2018). A review on data-driven fault severity assessment in rolling bearings. *Mechanical Systems and Signal Processing*, 99, 169–196.
- Chang, Y. W., Wang, Y. C., Liu, T., & Wang, Z. J. (2008). Fault diagnosis of a mine hoist using pca and svm techniques. *Journal of China University of Mining Technology*, 18(3), 327-331.
- Chen, Z., & Li, W. (2017). Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Transactions on Instrumentation and Measurement*, 66(7), 1693–1702.
- Cococcioni, M., Lazzerini, B., & Volpi, S. L. (2013). Robust diagnosis of rolling element bearings based on classification techniques. *IEEE Transactions on Industrial Informatics*, 9(4), 2256–2263.
- Cui, L., Zhang, Y., Zhang, F., Zhang, J., & Lee, S. (2016). Vibration response mechanism of faulty outer race rolling element bearings for quantitative analysis. *Journal of Sound and Vibration*, 364, 67–76.
- Gerber, T., Martin, N., & Mailhes, C. (2015). Time-frequency tracking of spectral structures estimated by a data-driven method. *IEEE Transactions on Industrial Electronics*, 62(10), 6616–6626.
- Guo, S., Yang, T., Gao, W., & Zhang, C. (2018). A novel fault diagnosis method for rotating machinery based on a convolutional neural network. *Sensors*, 18(5), 1429.
- Guo, S., Yang, T., Gao, W., Zhang, C., & Zhang, Y. (2018). An intelligent fault diagnosis method for bearings with variable rotating speed based on pythagorean spatial pyramid pooling cnn. *Sensors*, 18(11), 3857.
- Guo, T., Zhou, D., Zhang, J., Chen, M., & Tai, X. (2019). Fault detection based on robust characteristic dimensionality reduction. *Control Engineering Practice*, 84, 125–138.
- Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527–1554.
- Khanam, S., Tandon, N., & Dutt, J. (2014). Fault size estimation in the outer race of ball bearing using discrete wavelet transform of the vibration signal. *Procedia Technology*, 14, 12–19.
- Li, C., de Oliveira, J. L. V., Lozada, M. C., Cabrera, D., Sanchez, V., & Zurita, G. (2018). A systematic review of fuzzy formalisms for bearing fault diagnosis. *IEEE Transactions on Fuzzy Systems*.
- Liu, X., Bo, L., & Luo, H. (2015). Bearing faults diagnostics based on hybrid ls-svm and emd method. *Measurement*, 59, 145–166.
- Qin, Y., Wang, X., & Zou, J. (2019). The optimized deep belief networks with improved logistic sigmoid units and their application in fault diagnosis for planetary gearboxes of wind turbines. *IEEE Transactions on Industrial Electronics*, 66(5), 3814–3824.
- Rai, A., & Upadhyay, S. (2016). A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribology International*, 96, 289–306.
- Sun, S., Guo, Y., & Gao, Y. (2015). Experimental investigation on double impulses phenomenon of outer

race spalled rolling element bearings based on acoustic emission signals. In *Information and automation, 2015 IEEE international conference on* (pp. 1931–1935).

- Tang, S., Shen, C., Wang, D., Li, S., Huang, W., & Zhu, Z. (2018). Adaptive deep feature learning network with nesterov momentum and its application to rotating machinery fault diagnosis. *Neurocomputing*, 305, 1–14.
- Wang, S., Xiang, J., Zhong, Y., & Zhou, Y. (2018). Convolutional neural network-based hidden markov models for rolling element bearing fault identification. *Knowledge-Based Systems*, 144, 65–76.
- Zhang, B., Georgoulas, G., Orchard, M., Saxena, A., Brown, D., Vachtsevanos, G., & Liang, S. (2008). Rolling element bearing feature extraction and anomaly detection based on vibration monitoring. In *Control and automation, 2008 16th mediterranean conference on* (pp. 1792–1797).
- Zhang, B., Sconyers, C., Byington, C., Patrick, R., Orchard, M. E., & Vachtsevanos, G. (2010). A probabilistic fault detection approach: Application to bearing fault detection. *Ieee transactions on industrial electronics*, 58(5), 2011–2018.
- Zhao, G., Liu, X., Zhang, B., Liu, Y., Niu, G., & Hu, C. (2018). A novel approach for analog circuit fault diagnosis based on deep belief network. *Measurement*, 121, 170–178.
- Zhao, G., Liu, X., Zhang, B., Zhang, G., Niu, G., & Hu, C. (2017). Bearing health condition prediction using deep belief network. In *Proceedings of the annual conference of prognostics and health management society, orlando, fl, usa* (pp. 2–5).
- Zhao, S., Liang, L., Xu, G., Wang, J., & Zhang, W. (2013).

Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method. *Mechanical Systems and Signal Processing*, 40(1), 154–177.

## BIOGRAPHIES

**Guangxing Niu** received the B.A. degree from North University of China, Taiyuan, China, in 2014, and the M.S.E. degree from Harbin Institute of Technology, Harbin, China, in 2016, both in electrical engineering. He is currently working toward the Ph.D. degree in electrical engineering from the University of South Carolina, Columbia, SC, USA. His current research focuses on the data-driven based algorithms of diagnosis and prognosis for lithium ion batteries and industrial systems.

**Bin Zhang** received the B.E. and M.E. degrees from Nanjing University of Science and Technology, Nanjing, China, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2007. Before he joined the Department of Electrical Engineering, University of South Carolina, Columbia, SC, USA, he was with General Motors R&D, Detroit, MI; Impact Technologies, Rochester, NY; and Georgia Institute of Technology, Atlanta, GA, USA. He has published over 100 technical papers. His current research interests include prognostics and health management, intelligent systems and controls, and their applications to various engineering systems.

**Paul Ziehl** received the B.S. degree from California Polytechnic State University, San Luis Obispo and the M.S. and Ph.D. degrees from the University of Texas at Austin. Prior to joining the University of South Carolina, he was affiliated with Tulane University in New Orleans, Louisiana. He has published more than 170 technical papers. His current research interests include approaches for digital twin associated with advanced composites and structural prognostics for nuclear and other engineering systems.