Hindawi
International Journal of Rotating Machinery
Volume 2017, Article ID 9602650, 10 pages
https://doi.org/10.1155/2017/9602650

*Research Article*

# Fault Feature Extraction and Diagnosis of Gearbox Based on EEMD and Deep Briefs Network

**Kai Chen,[1,2] Xin-Cong Zhou,[1,2] Jun-Qiang Fang,[1,2] Peng-fei Zheng,[1,2] and Jun Wang[1,2]**

[1] *Key Laboratory of High Performance Ship Technology (Wuhan University of Technology), Ministry of Education, Wuhan 430063, China*
[2] *Reliability Engineering Institute, School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 430063, China*

Correspondence should be addressed to Kai Chen; kaichen@whut.edu.cn

Received 30 December 2016; Revised 14 March 2017; Accepted 29 March 2017; Published 11 June 2017

Academic Editor: Tonghai Wu



A gear transmission system is a complex nonstationary and nonlinear time-varying coupling system. When faults occur on gear system, it is difficult to extract the fault feature. In this paper, a novel fault diagnosis method based on ensemble empirical mode decomposition (EEMD) and Deep Briefs Network (DBN) is proposed to treat the vibration signals measured from gearbox. The original data is decomposed into a set of intrinsic mode functions (IMFs) using EEMD, and then main IMFs were chosen for reconstructed signal to suppress abnormal interference from noise. The reconstructed signals were regarded as input of DBN to identify gearbox working states and fault types. To verify the effectiveness of the EEMD-DBN in detecting the faults, a series of gear fault simulate experiments at different states were carried out. Results showed that the proposed method which coupled EEMD and DBN can improve the accuracy of gear fault identification and it is capable of applying to fault diagnosis in practical application.

## 1. Introduction

Gearbox is an indispensable part in the modern industry, especially for lots of large equipment [1]. As a device of transmission power, it is often prone to failure because of its complex structure and the poor working conditions. Failure of the gearbox not only leads to shutting down or threatening personal safety but also causes considerable economic losses. Therefore, it is very important for engineers and researchers to monitor the gear conditions to prevent this kind of malfunction of the plants.

The internal structure of the gearbox system is interrelated and coupled inside [2]. During the operation process, a lot of factors need to be considered when evaluating the performance status of the equipment, such as vibration, noise temperature, the debris contaminants in the oil and grease, torque of the power input and output, and stress distribution on the tooth surface. There are many methods for gearbox fault diagnosis, including vibration analysis [3], noise analysis [4], and oil analysis. Being online, real-time, and nondamage, vibration signal analysis shows many advantages which make it widely used for gearbox fault diagnosis.

The most common signal processing methods include time-domain analysis and time-frequency analysis [5] such as amplitude spectrum analysis, order analysis [6], cepstrum analysis [7, 8], envelope spectrum analysis [6], Hilbert transform demodulation analysis [9], wavelet analysis, autocorrelation analysis [10], and empirical mode decomposition (EMD) [11]. Among the available vibration analysis methods, EMD is an effective signal analysis method which is suited for dealing with nonlinear and nonstationary signals [12]. It consists in local and fully data-driven separation of a signal in fast and slow oscillations. However, EMD experiences some problems, such as the presence of oscillations of very disparate amplitude in a mode or the presence of very similar oscillations in different modes, named as "mode mixing." To overcome shortcomings, the ensemble empirical mode decomposition (EEMD) was proposed [13]. EEMD performs the EMD over an ensemble of the signal plus Gaussian white noise. Considering that vibration signals of the gearbox are

nonlinear and nonstationary, the EEMD algorithm is suitable for analyzing and judging the gearbox signals [14].

There are several failure modes in gear breakdown; some researchers utilized intelligent pattern recognition techniques in fault detection and diagnosis. SVM was used in machine condition monitoring and fault diagnosis [15]. RBF neural network is applied to improve the fault recognition rate in the fault diagnosis of gearbox [16]. Artificial neural network expert system was proposed to detect and localize defects in rolling element bearings [17]. However, the conventional neural network is of slow convergence speed and the possibilities of finding optimal solution are small, which have limited the neural network's application in the practical projects. But since the Deep Learning was introduced by Hinton et al. in 2006 [18], a new method in machine learning techniques started to be applied in visual data classification [19], decoding analysis [20], language information retrieval [21], and classifying the faults [22]. Especially in 2016, AlphaGo, an outstanding product of the depth learning algorithm, defeated a famous chess player with the score of 4 : 1 [23]. Its excellent performance in the game revealed the power of the learning algorithm. Therefore, it is necessary to introduce such an advanced algorithm into the fault diagnosis of gearbox. In this paper, a new method coupled with EEMD and DBN will be presented to diagnose the gear fault.

## 2. Description of Fault Diagnosis Method

*2.1. EEMD Algorithm.* EMD is a time-frequency signal analysis method of nonlinear signals, which can decompose the data adaptively and obtain a series of IMFs [24]. These IMFs reflect the characteristics of the signal itself and are distributed from high frequency to low frequency. EMD algorithm is very suitable for analyzing the vibration signal of gearbox, because the vibration signal of gearbox is nonstationary.

However, in practical application, it is found that there is a problem of mode mixing in the EMD decomposition process [13]; that is, a single IMF contains different frequency components or the same frequency components are decomposed into different IMFs.

To alleviate the mode mixing problem occurring in EMD, an ensemble empirical mode decomposition (EEMD) is presented [25]. The essential of EEMD algorithm is to decompose the original signal which adds Gaussian noise by using the EMD method repeatedly, and the original vibration data is decomposed into a series of IMFs with different scales and continuous characteristics because of the characteristic of frequency uniform distribution of Gaussian white noise, which can suppress the appearance of modal mixing.

The EEMD algorithm can be summarized as follows:

(1) Adding a white noise $N(t)$ to the original signal $x(t)$ can get a new signal $X(t)$:

$$X(t) = x(t) + N(t). \tag{1}$$

(2) The signal $X(t)$ is decomposed by EMD to obtain a set of IMFs:

$$X(t) = \sum_{j=1}^{n} c_j(t) + r_n(t). \tag{2}$$

(3) Repeat steps (1) and (2) $m$ times; get $m$ groups of IMF by adding different amplitude of white noise $N_i(t)$, $i = 1 \sim m$:

$$\begin{aligned} X_i(t) &= x(t) + N_i(t) \\ X_i(t) &= \sum_{j=1}^{n} c_{ij}(t) + r_{in}(t). \end{aligned} \tag{3}$$

(4) Preserve the mean $c_j(t)$ $(i = 1, 2, \ldots, m)$ of each of the $m$ IMFs as the final IMFs:

$$c_j(t) = \frac{1}{m}\sum_{i=1}^{m} c_{ij}(t). \tag{4}$$

To verify the validity of EEMD, a simulated signal as shown in formula (5) is decomposed by using EMD and EEMD, respectively; the simulated signal and decomposition results are shown in Figure 1.

$$\begin{aligned} y_1 &= \sin(2\pi 30t) \\ y_2 &= \sin(2\pi 10t) \quad (t > 0.2) \\ y_3 &= \sin(2\pi 50t) \quad (t < 0.2) \\ y &= y_1 + y_2 + y_3. \end{aligned} \tag{5}$$

It can be seen from Figure 1 that EEMD can solve the mode mixing phenomenon which is better than EMD. Moreover, IMF2, IMF3, and IMF4 decomposed by EEMD are basically the same as the three components in the original simulated signal.

*2.2. Deep Belief Networks*

*2.2.1. Architecture of Restricted Boltzmann Machine.* Restricted Boltzmann machine (RBM), a generated random neural network, was proposed by Hinton and Sejnowski in 1986 [26]. There are two components in RBM: the visible layer ($v$) (input layer) in the bottom and the hidden layer ($h$) above. All visible units are connected to all hidden units, but there are no visible-visible or hidden-hidden connections. The weight between the visible layer and the hidden layer is denoted by $W$. The random variables of unit $(v, h)$ take values $\in (0, 1)$, and the joint probability distribution $P(v, h)$ under the model follows Boltzmann distribution [27]. The architecture of RBM is shown in Figure 2.

The learning process is as follows: when inputting the visible layer $v$, the hidden layer $h$ is obtained by $p(h \mid v)$, and vice versa. In other words, the visible layer is determined by $p(h \mid v)$ and the hidden layer. Then the parameter is adjusted until difference between the visible layer and derived visible layer is minimized. It means that the hidden layer obtained in the middle process is an alternative representation of the visible layer.

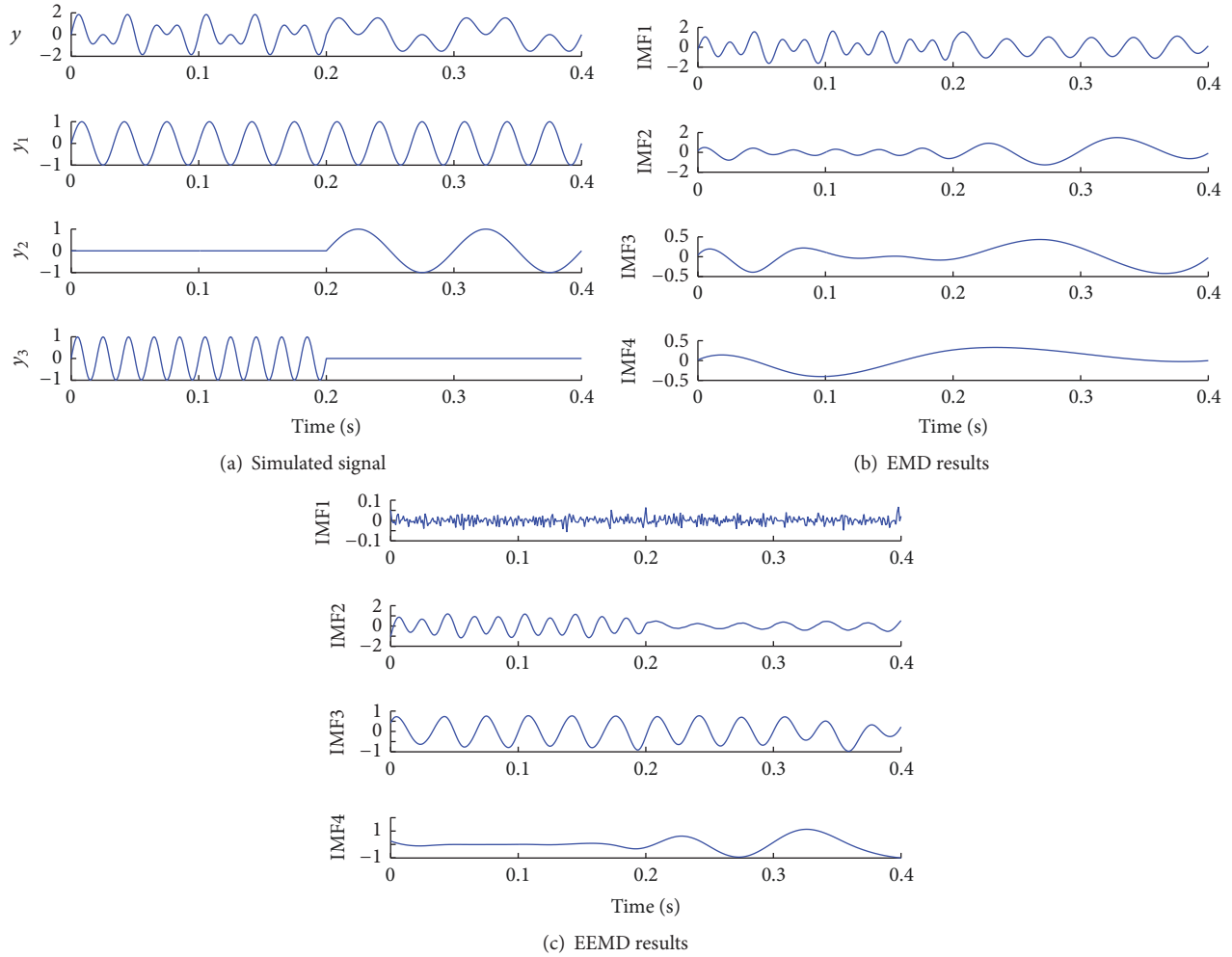



Figure 1: Comparison results of EMD and EEMD.

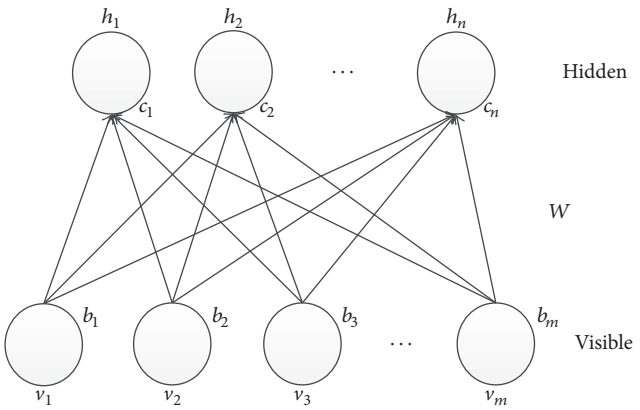



Figure 2: Architecture of restricted Boltzmann machine.

The derivation process of the RBM weight using contrastive divergence (CD) is described as follows [28]:

(1) Randomly choose a set of data in training sample as input $v^{(1)}$ (superscript denotes train time); determine learning rate $L$ and node number $m$ of the hidden layer.

(2) Randomly initialize $W$, $b$, and $c$.

(3) Update state of hidden variable according to formula (4):

$$P\left(h_j^{(k)} = 1 \mid v^{(k)}\right) = \frac{1}{1 + \exp\left(-\sum_i W_{ij} v_i^{(k)} - b_j\right)}, \quad (6)$$

where $j = 1, 2, 3, \ldots, m$ is the node number of the hidden layer.

(4) Reconstitute $v^{(k+1)}$ according to $h^{(k)}$ and formula (5):

$$P\left(v_i^{(k+1)} = 1 \mid h^{(k)}\right) = \frac{1}{1 + \exp\left(-\sum_j W_{ij} h_j^{(k)} - c_i\right)}, \quad (7)$$

where $i = 1, 2, 3, \ldots, n$ is the node number of the input layer.

(5) Then calculate $h^{(k+1)}$ according to $v^{(k+1)}$ and formula (4).

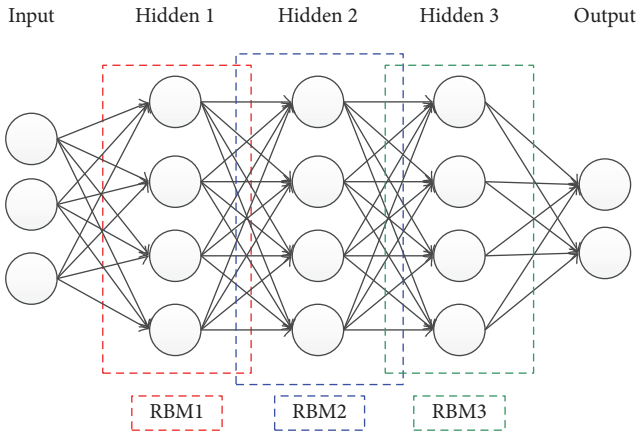



Figure 3: Deep belief network structure.

(6) Update weight and bias according to formula (6):

$$W = W + L\left(P\left(h_j^{(k)} = 1 \mid v^{(k)}\right)\left(v^{(k)}\right)^T - P\left(h_j^{(k+1)} = 1 \mid v^{(k+1)}\right)\left(v^{(k+1)}\right)^T\right)$$
$$c = c + L\left(v^{(k)} - v^{(k+1)}\right) \quad (8)$$
$$b = b + L\left(P\left(h_j^{(k)} = 1 \mid v^{(k)}\right) - P\left(h_j^{(k+1)} = 1 \mid v^{(k+1)}\right)\right).$$

(7) Iterate steps (4), (5), and (6) $K$ times; fulfill parameter update of RBM.

*2.2.2. Deep Belief Networks.* A DBN is formed by stacking a number of the RBMs layer by layer as shown in Figure 3. Each layer of the RBM is independent, and the layers are interconnected by weight. Deep belief network structure is shown in Figure 3.

A usable DBN network model can be obtained via pretraining and fine-tuning. Note that this learning procedure, so-called pretraining, is unsupervised. The training procedure is described as follows [29]:

(1) Each RBM layer is trained individually, insuring that the output of each layer contains input features as much as possible. The training process is as follows: the input data is mapped to the output data by weight; then the output data try to reconstruct the input data. The weight of the network is updated according to the difference after reconstruction. Repeat this process until the difference between input data and output data is very small; this is the RBM learning process presented in the previous section.

(2) In the previous step, the training only optimizes the mappings between the input layer and the output layer rather than the entire DBN. Therefore, it is necessary to set the last DBN layer as softmax classifier. The output feature vector of RBM is input feature vector of softmax classifier and it is trained under monitoring. All parameters of the network are fine-tuned according to the error of output and input to each RBM layer spreading from top to bottom. The process is illuminated as a flowchart in Figure 4.

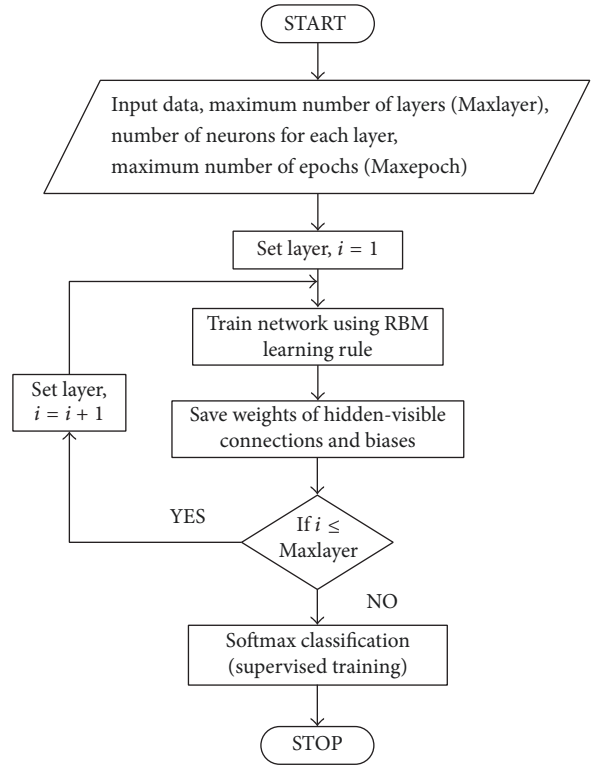



Figure 4: Flowchart of the DBN training process.

## 3. Experimental Setup and Fault Data

*3.1. Experimental Setup.* To evaluate the performance of the fault diagnosis method, experimental analysis of a gear transmission has been carried out in an experimental setup. As shown in Figure 5, the experimental setup is composed of electrical motor, timing belt pulley, coupling, gearbox, and magnetic powder brake. An electrical motor with frequency converter, whose speed is up to 1,500 r/min, is the drive. The number of teeth of the input gear is 55, while that of the output gear is 75; thus the speed ratio is 55/75. A magnetic powder brake is used as the external load. The sensors are the piezoelectric accelerometers (CA-YD-186).

The amplitude of the vibration signal is attenuated during transmission, and the amplitude attenuation of the high frequency component is much faster than the low frequency component. The location near the bearing, in which attenuation and distortion of the vibration signal were minimum, is the best sensor measuring point. In the experiments, the acceleration sensor is used to measure the vibration signal of the gearbox. And the sensor was arranged in the gearbox near each bearing seat on vertical direction, which is closer to the bearing than the horizontal direction in this experimental setup. The sensor measuring point was shown in Figure 5.

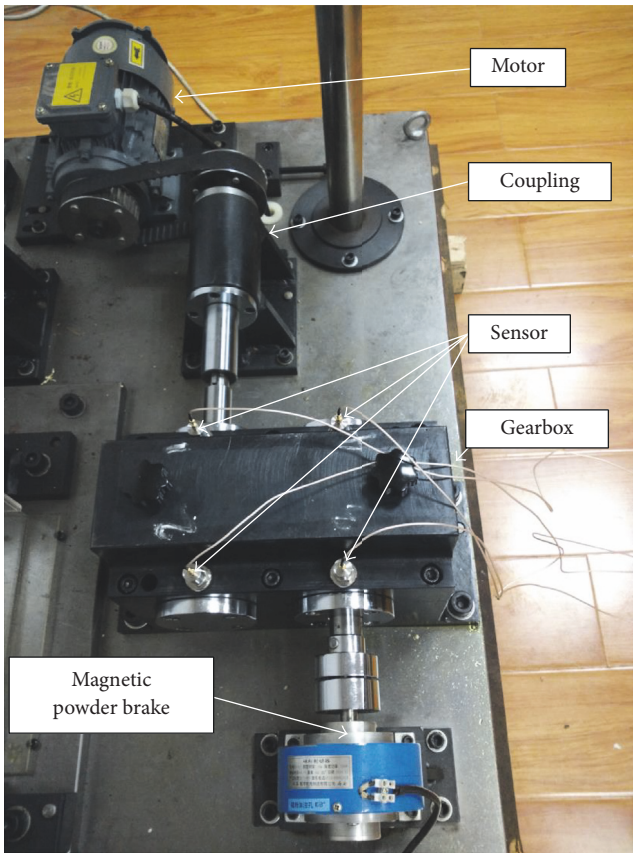



Figure 5: Experimental setup and sensor arranged.

*3.2. Fault Data Collection.* In order to collect fault data of gear fault type, a series of gears were processed to simulate common fault gears. As shown in Figure 6, three typical faults of gear were simulated, which are tooth breaking, root crack, and pitting corrosion.

Experiment was designed to collect the mass of data under various faults and work conditions, which involve four gear statuses (normal, break, wear, and pit), two types of motor speed (1500 r/min and 1200 r/min), and three kinds of load (0 A, 0.1 A, and 0.2 A). The sampling frequency is 10240 Hz. There are 131072 sampling points in each condition; then the data is divided into 64 parts, each of which is regarded as a sample with 2048 data points. Thus, 384 sets of data were generated in each gear status, and 64 sets of them were selected as test samples by random drawing, while the other 320 sets were chosen as training samples. In total, 1280 training samples and 256 test samples were extracted from the vibration data.

## 4. Application of Proposed Diagnosis Method

*4.1. EEMD-DBN Diagnosis Method.* The original vibration signals were decomposed into several IMFs components with EEMD method. The decomposed IMFs components represent the components of different frequencies in the original vibration signal. EEMD results in one sample of the different fault signals are indicated in Figure 7. There are 11 IMFs components in each EEMD result, and each decomposed fault signal is shown from high frequency to low frequency.

Table 1: Parameters of DBN.

| | |
|---|---|
| Network layer (including input layer) | 4 |
| Type of input data | Frequency domain |
| Length of the input data | 1024 |
| Node number of input layer | 1024 |
| Node number of RBM1 | 500 |
| Node number of RBM2 | 100 |
| Node number of output layer | 4 |
| Classifier of output layer | Softmax |
| Learning rate | 0.1 |
| Maximum iteration | 100 |

In real working condition, the collected vibration signal generally contains the noise background, so the pseudocomponent influences the EEMD decomposition. The frequency components contained in these pseudocomponents have the possibility of coinciding with the IMF characteristic frequency band and they should be eliminated. The process of gearbox fault diagnosis based on EEMD noise reduction and DBN is illuminated in Figure 8.

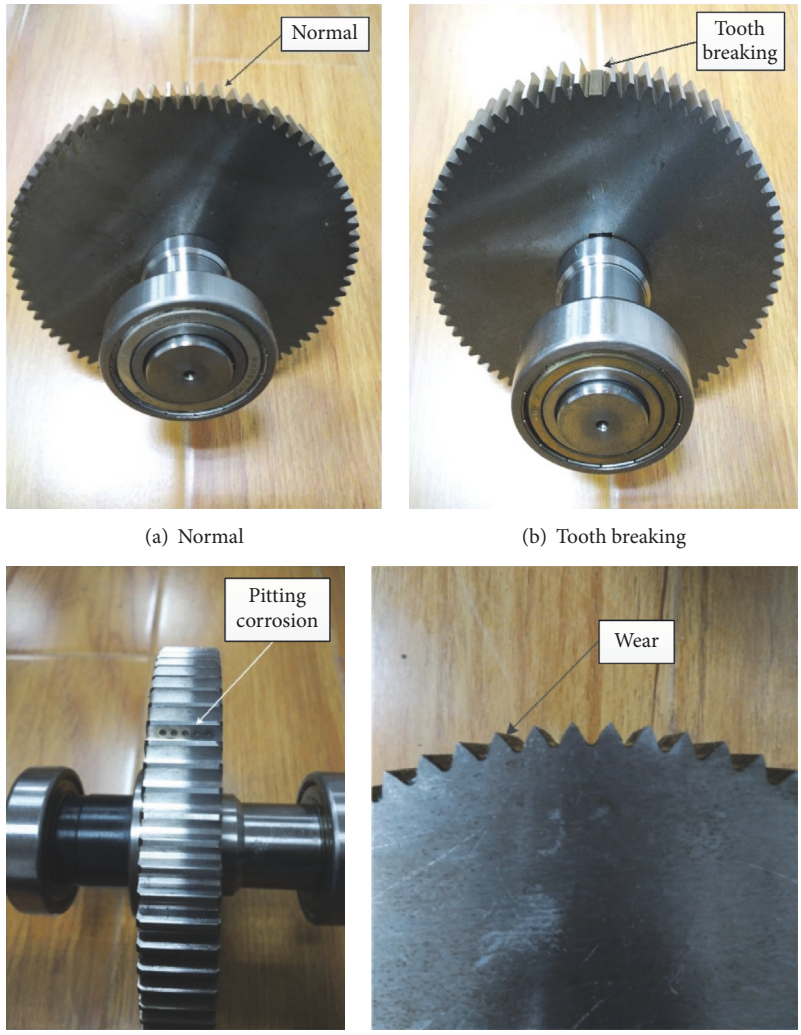
As the natural frequency of the gear is high, the noise in the signal is decomposed into the low frequency band in EEMD. However, the information contains the gearbox working status and fault characteristics which are decomposed into high frequency. Therefore, the high frequency parts of EEMD results are what is mainly studied. Front 4th-order IMFs decomposed by original data were chosen to reconstitute signals. After EEMD noise reduction, DBN is used to classify the four states of the gearbox.

*4.2. DBN Training Results.* Theoretically, the length of the input data should be 2048 points, but the type of input data is the frequency domain of reconstituted signal, which was symmetry, so the first half was taken. This choice will not affect the classification results. Moreover, it can reduce the training time. Therefore, the size of the input data is 1024. The parameters of DBN are shown in Table 1.

A comparative analysis was carried out between coupled EEMD with DBN and DBN without EEMD to verify the availability of the coupled method.

To increase the separation among the feature clusters in order to gain more diagnostic accuracy and reduce the feature dimensionality for effective computation, the principle component analysis (PCA) is successively applied to the obtained feature sets. The distribution of the features in the new feature space can be observable through the visualization.

The first three principal components of first layer and second RBM layer are extracted, and the visual graphic is shown in Figures 9 and 10. Figure 9 is the result of DBN




(a) Normal (b) Tooth breaking

(c) Pitting corrosion (d) Tooth wear

Figure 6: Simulation of gear fault type.

training without EEMD method, and Figure 10 is the result of DBN training with EEMD noise reduction.

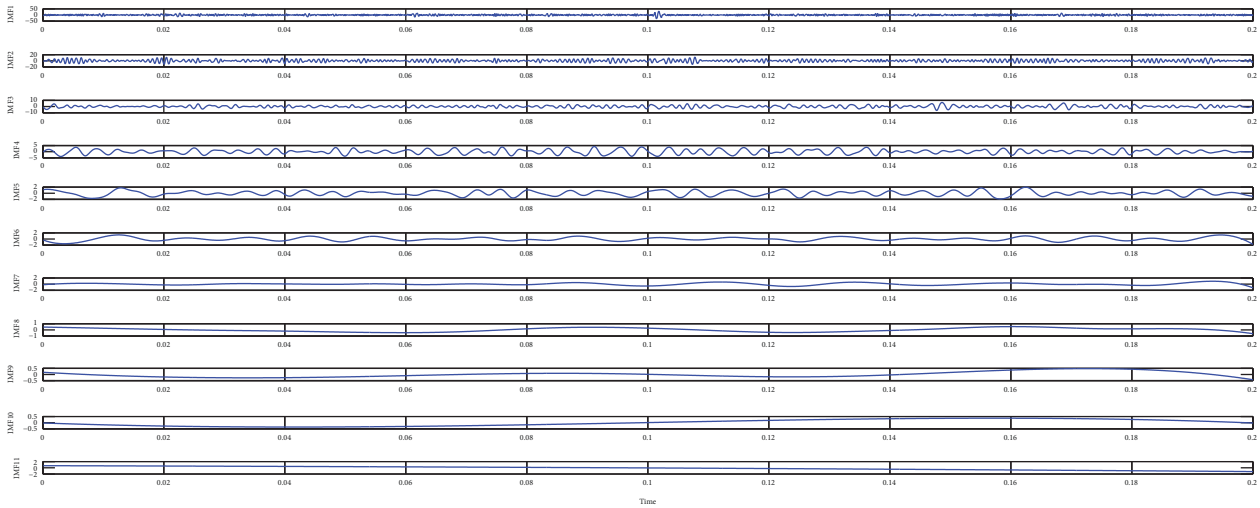
*4.3. DBN Test Results.* In order to estimate classification effect of the two training data samples on DBN, the data test was carried out. Test data was processed the same as the training data (if the training data is processed by EEMD, so is the test data; if the training data was unprocessed, so is the test data). The test of the DBN network was carried out 20 times, and the accuracy rate of each time is shown in Figure 11.

It can be observed from Figure 11 that the average recognition rate with EEMD processing is 99.25%, while the average recognition rate without processing is 96.25%. The recognition rate was enhanced almost 3% after EEMD processing. Moreover, recognition rate is more stable when the data is processed. It can be said that DBN methods effectively assist in improving the accuracy of gear fault diagnosis.

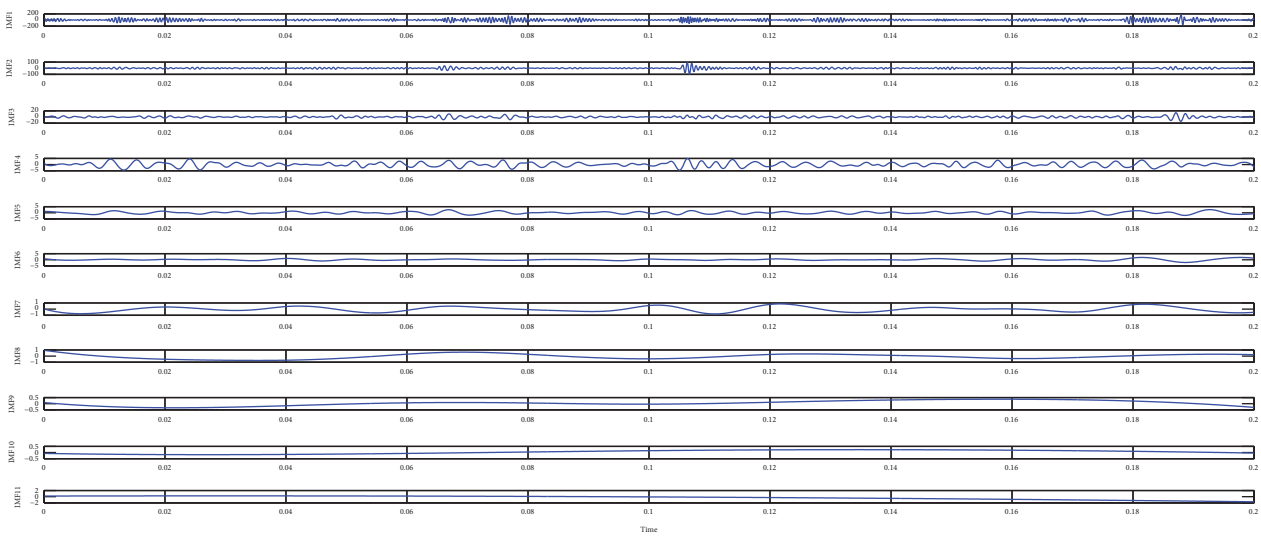
## 5. Conclusion

This paper has presented the new approach to diagnose gearbox fault using vibration signals. The original vibration signals were decomposed into several IMFs components with EEMD method. The decomposed IMFs components represent the different frequency components of the original vibration signal. To deal with the instabilities occurring in vibration signal, only main IMFs were chosen to reconstitute signals for eliminating noise effects in signal. The fault features were extracted by DBN training.

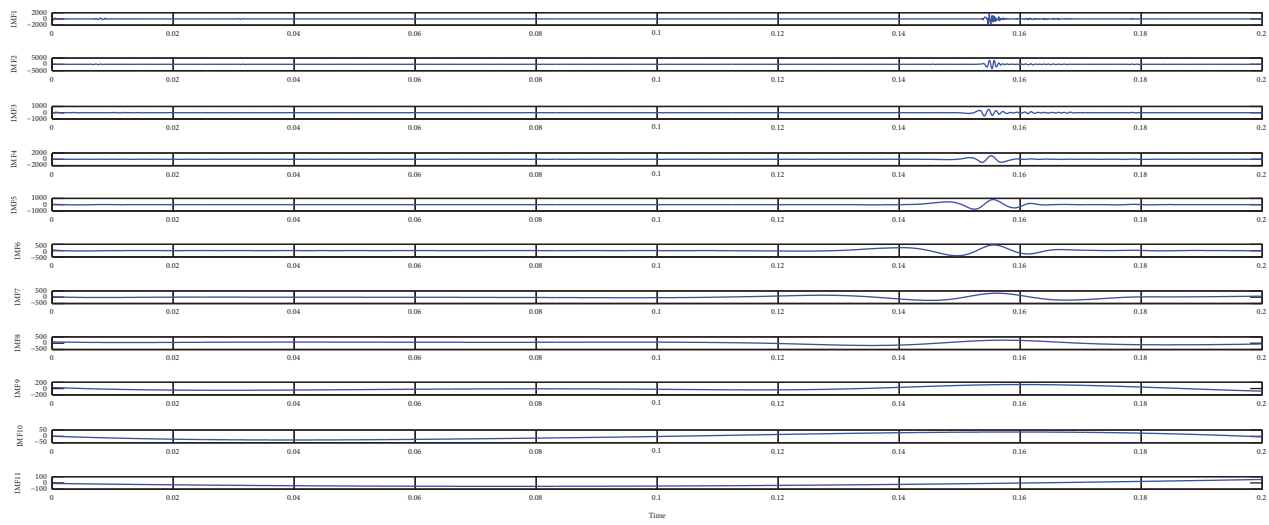
Finally, the features are inputted to DBN test, so that their performance is appraised. The results show that the proposed approach which coupled EEMD and DBN can improve the accuracy of gear fault identification and it is capable of applying to fault diagnosis in practical application.




(a) Normal

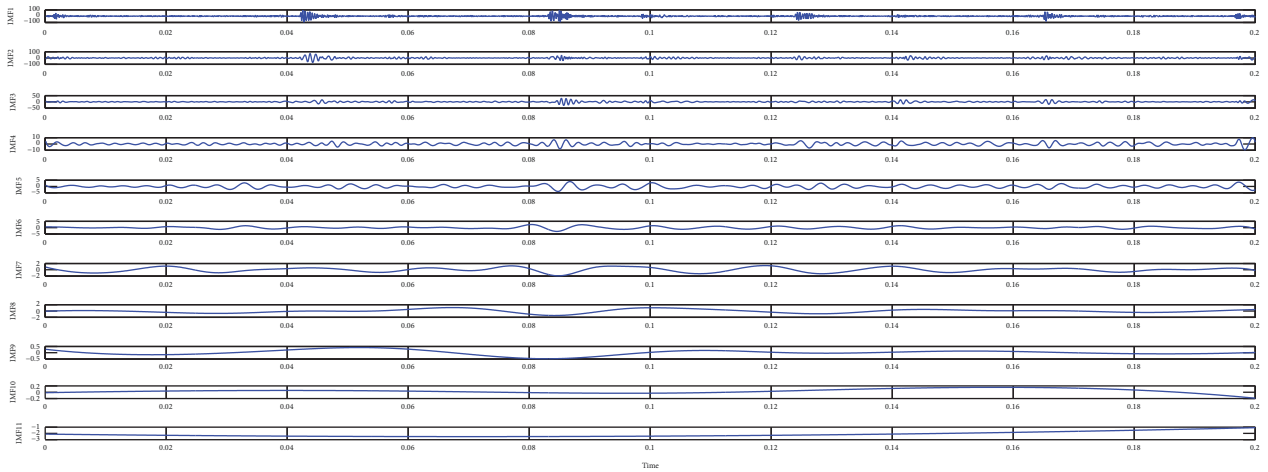



(b) Break




(c) Wear

Figure 7: Continued.




(d) Pit

Figure 7: EEMD results of fault signal.

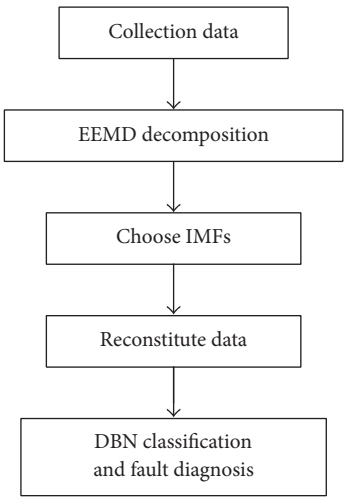
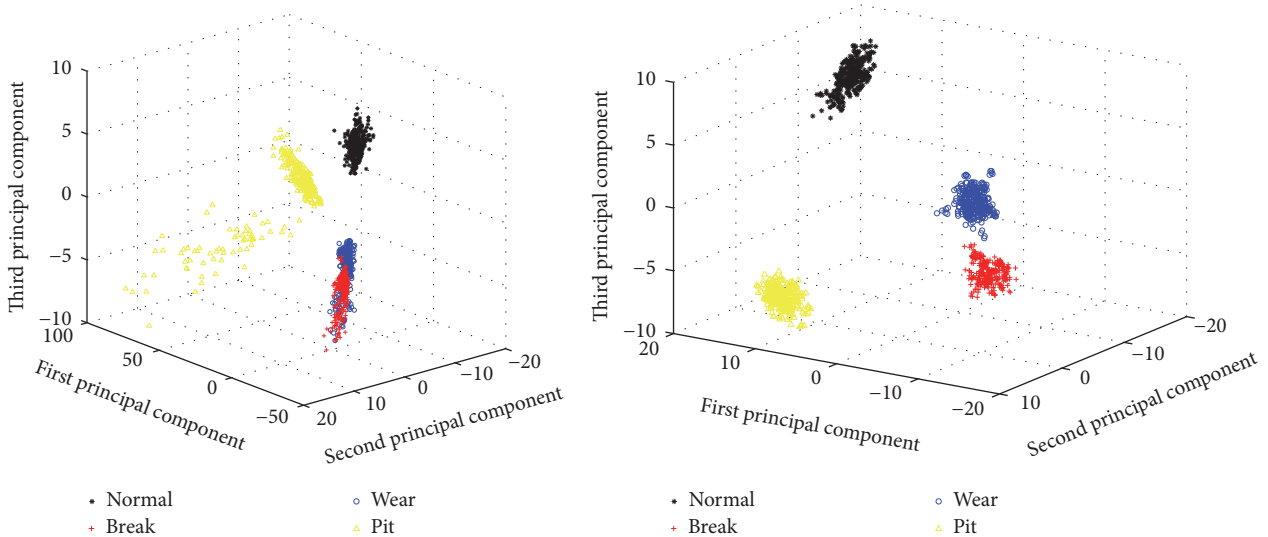



Figure 8: Procedure of fault diagnosis based on EEMD and DBN.




(a) First RBM layer

(b) Second RBM layer

Figure 9: Features distribution of DBN result.

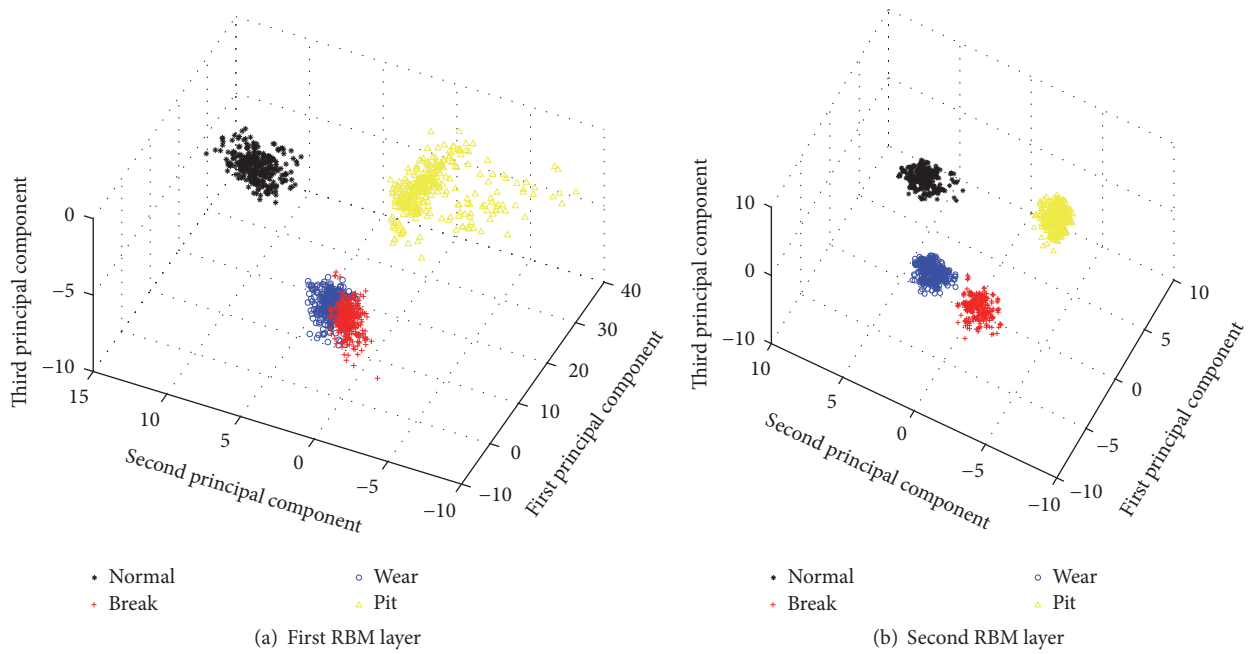



Figure 10: Features distribution of EEMD-DBN result.

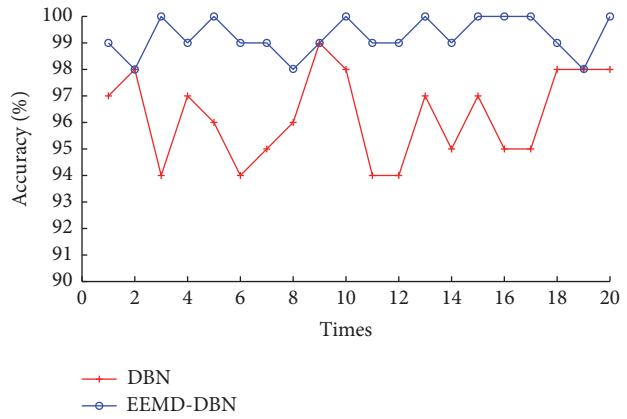



Figure 11: Accuracy of test result.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This project is sponsored by the grants from the Ministry of Transport Applied Basic Research Project (no. 2013329811360), the National Natural Science Foundation of China (no. 51139005 and no. 51509194), and the Fundamental Research Funds for the Central Universities (no. 2014-yb-018).

## References

[1] Z. Li, X. Yan, C. Yuan, J. Zhao, and Z. Peng, "Fault detection and diagnosis of a gearbox in marine propulsion systems using bispectrum analysis and artificial neural networks," *Journal of Marine Science and Application*, vol. 10, no. 1, pp. 17–24, 2011.

[2] Z. Li, X. Yan, Z. Tian, C. Yuan, Z. Peng, and L. Li, "Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis," *Measurement Journal of the International Measurement Confederation*, vol. 46, no. 1, pp. 259–271, 2013.

[3] B. Chen, Z. Zhang, C. Sun, B. Li, Y. Zi, and Z. He, "Fault feature extraction of gearbox by using overcomplete rational dilation discrete wavelet transform on signals measured from vibration sensors," *Mechanical Systems and Signal Processing*, vol. 33, pp. 275–298, 2012.

[4] Ş. Ulus and S. Erkaya, "An experimental study on gear diagnosis by using acoustic emission technique," *International Journal of Acoustics and Vibrations*, vol. 21, no. 1, pp. 103–111, 2016.

[5] Z. Li, Y. Jiang, C. Hu, and Z. Peng, "Recent progress on decoupling diagnosis of hybrid failures in gear transmission systems using vibration sensor signal: a review," *Measurement*, vol. 90, pp. 4–19, 2016.

[6] J. Cheng, Y. Yang, and D. Yu, "The envelope order spectrum based on generalized demodulation time–frequency analysis and its application to gear fault diagnosis," *Mechanical Systems & Signal Processing*, vol. 24, no. 2, pp. 508–521, 2010.

[7] H. Li and Zheng, "Gear fault diagnosis based on order cepstrum analysis," *journal of Vibration & Shock*, vol. 25, no. 5, pp. 65–68, 2006.

[8] L. Nacib, K.-M. Pekpe, and S. Sakhara, "Detecting gear tooth cracks using cepstral analysis in gearbox of helicopters," *International Journal of Advances in Engineering & Technology*, 2013.

[9] A. Djebala, N. Ouelaa, C. Benchaabane, and D. F. Laefer, "Application of the Wavelet Multi-resolution Analysis and Hilbert transform for the prediction of gear tooth defects," *Meccanica*, vol. 47, no. 7, pp. 1601–1612, 2012.

[10] M.-E. Morsy and G. Achtenova, "Value of autocorrelation analysis in vehicle gearbox fault diagnosis," *International Journal of Vehicle Noise & Vibration*, vol. 11, no. 2, pp. 181–193, 2015.

[11] Y. Li, Z. Han, J. Gao, and S. Ning, "Application of EMD in gear fault diagnosis based on autocorrelation analysis," *Journal of Chinese Agricultural Mechanization*, 2015.

[12] J. Cheng, D. Yu, and Y. Yang, "Fault diagnosis for rotor system based on EMD and fractal dimension," *China Mechanical Engineering*, 2005.

[13] N. E. Huang, M. C. Wu, S. R. Long et al., "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *The Royal Society of London. Proceedings. Series A. Mathematical, Physical and Engineering Sciences*, vol. 459, no. 2037, pp. 2317–2345, 2003.

[14] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4144–4147, Prague, Czech Republic, May 2011.

[15] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical Systems & Signal Processing*, vol. 21, no. 6, pp. 2560–2574, 2007.

[16] J. Leng, S. Jing, and Z. Wu, "Fault diagnosis for gearbox based on RBF neural network," *Journal of Mechanical Strength*, 2010.

[17] P. Jayaswal, S.-N. Verma, and A.-K. Wadhwani, "Development of EBP-Artificial neural network expert system for rolling element bearing fault diagnosis," *Journal of Vibration & Control*, vol. 17, no. 8, pp. 1131–1148, 2011.

[18] 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.

[19] Y. Liu, S. Zhou, and Q. Chen, "Discriminative deep belief networks for visual data classification," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2287–2296, 2011.

[20] Y. Hatakeyama, H. Kataoka, Y. Okuhara, and S. Yoshida, "Decoding analysis for fMRI based on Deep Brief Network," in *2014 World Automation Congress, WAC 2014*, pp. 268–272, usa, August 2014.

[21] J. Kim, J. Nam, and I. Gurevych, "Learning semantics with deep belief network for Cross-Language information retrieval," in *International Conference on Computational Linguistics*, pp. 579–588.

[22] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," *Reliability Engineering & System Safety*, vol. 115, no. 7, pp. 124–135, 2013.

[23] F.-Y. Wang, J. J. Zhang, X. Zheng et al., "Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond," *IEEE/CAA Journal of Automatica Sinica*, vol. 3, no. 2, pp. 113–120, 2016.

[24] Y. Lei, Z. He, and Y. Zi, "Application of the EEMD method to rotor fault diagnosis of rotating machinery," *Mechanical Systems & Signal Processing*, vol. 23, no. 4, pp. 1327–1338, 2009.

[25] X. Wang, C. Liu, F. Bi, X. Bi, and K. Shao, "Fault diagnosis of diesel engine based on adaptive wavelet packets and EEMD-fractal dimension," *Mechanical Systems & Signal Processing*, vol. 41, no. 1, pp. 581–597, 2013.

[26] G.-E. Hinton and T.-J. Sejnowski, "Learning and relearning in Boltzmann machines," pp. 282–317, 1986.

[27] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proceedings of the 24th International Conference on Machine learning (ICML '07)*, vol. 227, pp. 791–798, Corvallis, Oregon, June 2007.

[28] A. Fischer and C. Igel, *An Introduction to Restricted Boltzmann Machines*, Springer Berlin Heidelberg, 2012.

[29] G. Hinton, *Deep Belief Nets*, Springer US, 2011.