

Research Article

Multifeatures Fusion and Nonlinear Dimension Reduction for Intelligent Bearing Condition Monitoring

Liang Guo, Hongli Gao, Haifeng Huang, Xiang He, and ShiChao Li

School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Correspondence should be addressed to Hongli Gao; hongli_gao@home.swjtu.edu.cn

Received 17 October 2015; Revised 7 January 2016; Accepted 17 January 2016

Academic Editor: Peng Chen

Copyright © 2016 Liang Guo 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.

Condition-based maintenance is critical to reduce the costs of maintenance and improve the production efficiency. Data-driven method based on neural network (NN) is one of the most used models for mechanical components condition recognition. In this paper, we introduce a new bearing condition recognition method based on multifeatures extraction and deep neural network (DNN). First, the method calculates time domain, frequency domain, and time-frequency domain features to represent characteristic of vibration signals. Then the nonlinear dimension reduction algorithm based on deep learning is proposed to reduce the redundancy information. Finally, the top-layer classifier of deep neural network outputs the bearing condition. The proposed method is validated using experiment test-bed bearing vibration data. Meanwhile some comparative studies are performed; the results show the advantage of the proposed method in adaptive features selection and superior accuracy in bearing condition recognition.

1. Introduction

Bearing degradation is one of the most common fault sources in rotating machinery system. Unexpected bearing failure can lead to large costs of maintenance and loss of revenue. Traditionally, the maintenance activity is selected from preventive and corrective maintenance. However, preventive maintenance may typically involve the high maintenance costs, and corrective maintenance may reduce the productive efficiency [1, 2]. So condition-based maintenance based on condition monitoring is critical to assure safe and efficient operation of rotating machinery system [3].

Generally, among the various approaches of bearing condition monitoring, it can be classified into two categories: signal processing-based approaches and pattern recognition-based approaches. In signal processing-based approaches, some mathematical or statistical operations are performed on the measurement signal and then the bearing condition is judged through the prior knowledge of human beings. The kurtosis coefficient is a major indicator of bearing performance degradation detection [4]. The other statistical parameters such as the first-order moment (e.g., mean \bar{x}), the second-order moment (e.g., variance σ^2), and Crest Factor

are utilized in the study of bearing condition monitoring usually. The statistical analysis method is widely used for its simplicity and computation [5]. Spectrum analysis is also effective in tracking fault of bearing [6]. Spectral kurtosis in frequency domain is an effective tool to enhance the machinery fault features [7]. For bearing condition monitoring and fault diagnosis, envelope spectrum analysis is used for diagnostics bearing where faults signal has an amplitude modulating effect on the characteristic frequencies [8]. However, the spectral analysis method has no time scale information. In order to overcome that shortcoming, some new approaches based on time-frequency analysis are proposed for bearing faults detecting and localizing. Shi et al. [9] proposed a wavelet-based envelop spectrum method to detect the defect in rolling element bearings.

Meanwhile, pattern recognition-based condition monitoring is composed of two parts: feature extraction and pattern recognition. The signal processing-based feature extraction approach is a basic step in condition monitoring. Pattern recognition is another critical issue. Classification-based machine learning algorithms are commonly used in industry and academic fields for condition recognition [10]. Support

vector machines (SVMs) map the input data into high-dimension space with different kernel functions to efficiently perform a nonlinear classification. Abbasion et al. [11] present a new method based on wavelet analysis (WP) and support vector machine (SVM) to diagnosis of multifault of bearing. Artificial neural networks (ANN) are able to learn expert knowledge through a representative set of data. So ANN is commonly adopted in automated detection of machine conditions [12]. Bin et al. [13] extract the fault characteristics combined empirical mode decomposition and wavelet packet decomposition. Then BP neural network is taken to identify the model of bearing. In addition to BP neural network, fuzzy neural network [14], conditional random field [15], recurrent neural network [16], and radial basis function neural network [17] are also applied in bearing intelligent condition monitoring. Deep neural network (DNN) is a new kind of neural network architectures that attempts to abstract the high-level features from the raw signal through multilayer nonlinear transformation [18]. Hinton and Salakhutdinov [19] point out that deep learning can convert high-dimensional data to low-dimension nonlinearly by training a multilayer neural network. Since then deep neural network has widely been used in speech recognition [20], image recognition [21], natural language processing [22], and some other classification and recognition applications. But recently, the bearing condition monitoring based on deep neural network almost cannot be found in academic and industrial.

In this paper, we propose a new method based on multifusion and DNN nonlinear dimension reduction to recognize bearing condition. The time domain features, frequency domain features, and time-frequency features are fused for integrating the bearing condition features. Those features are input into the deep neural network (DNN) as the input vector. Then the DNN model will extract the high-level abstraction of input data and recognize the bearing condition. The remainder of this paper is organized as follows. The proposed method is presented in Section 2. Section 3 discusses the case study, where the proposed method is validated through real-world bearing vibration signal. Section 4 presents the conclusion and future work.

2. Condition Recognition Using the Proposed Method

The proposed method for bearing condition monitoring mainly includes three steps: feature extraction, feature dimension reduction, and condition recognition. The detailed process is as follows.

2.1. Feature Extraction. The presence of defect of machinery components can barely be determined from the raw acceleration signal. To get a better understanding of the raw vibration signal we tried to extract the time domain, frequency domain, and time-frequency domain features.

2.1.1. Time Domain Features. Time domain features of vibration signals have proved to be useful to present the machinery condition. From Table 1, we apply those classical time domain

TABLE 1: Time domain features extraction.

Name	Formula
RMS	$P_{t1} = \sqrt{\frac{1}{N} \sum_{n=1}^N x(n)^2}$
Skewness	$P_{t2} = \frac{1}{N} \sum_{n=1}^N \frac{(x(n) - \bar{x})^3}{\sigma^3}$
Kurtosis	$P_{t3} = \frac{1}{N} \sum_{n=1}^N \frac{(x(n) - \bar{x})^4}{\sigma^4} - 2$
Peak to Peak	$P_{t4} = x_{\max} - x_{\min}$
Crest Factor	$P_{t5} = \frac{ x(n) _{\max}}{\text{RMS}}$
Shape Factor	$P_{t6} = \frac{\text{RMS}}{(1/N) \sum_{n=1}^N x(n) }$
Impulse Factor	$P_{t7} = \frac{ x(n) _{\max}}{(1/N) \sum_{n=1}^N x(n) }$
Margin Factor	$P_{t8} = \frac{ x(n) _{\max}}{\text{RMS}}$

statistical features to represent the bearing condition. The Peak to Peak value calculates the difference between the maximum and minimum value in the signal. The Root Mean Square (RMS) measures the power content of vibration signal. Those are the first moment and second moment of probability distribution. Skewness, an asymmetry measure, is defined as the third moment. Kurtosis, a descriptor of signal shape, is based on a scaled calculation of the fourth moment of the data. The Crest Factor, Shape Factor, Impulse Factor, and Margin Factor are some dimensionless quantities that govern the bearing degeneration behavior.

2.1.2. Frequency Domain Features. The signal spectrum contains rich condition information. The spectral kurtosis, indicating the presence of short transients and their frequency locations, is obtained by calculating the kurtosis from each frequency band. For a number M of N -point realization at the frequency index m , the spectral kurtosis is given as [23]

$$G_{\text{SK}}(m) = \frac{\sum_{i=1}^M |X_N^i(m)|^4}{(\sum_{i=1}^M |X_N^i(m)|^2)^2} - 2, \quad (1)$$

where M is the realization of the process that contains N points. And X^i is the data from frequency transform. In practice, the categories of failures are not known. Similar to the rolling bearing, the inner, outer races and the cage faults have the chance to appear. So the frequency signatures are difficult to obtain due to the fact that the fault may concern all the components of the test bearing at the same time. We proposed a new condition indicator which is obtained

by calculating the correlation coefficient between the two vibration signals

$$CO_{XY} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}, \quad (2)$$

where x_i and y_i are two spectral kurtosis vectors of size N . x_i is the standard spectral kurtosis frequency distribution. y_i is the spectral kurtosis of current signal. And \bar{x} and \bar{y} are their corresponding means.

2.1.3. Time-Frequency Domain Features. Wavelet transfer (WT) decomposes signal into time-frequency space and has the excellent performance to present the signal characteristic. However, WT just subdivides signal into low-frequency bands. To get more detailed signal frequency representation, the wavelet package transform (WPT) is performed. Through the WPT, a signal can be divided into high-frequency and low-frequency bands with the binary tree form. At each decomposition levels, the signal is divided into two mutual orthogonal subspaces:

$$U_{j,k} = U_{j+1,2k} \oplus U_{j+1,2k+1}, \quad (3)$$

where j indicates the tree level and k is the node index in level j . The dividing is operating until decomposition level J ; then the process will produce 2^J subspaces which are mutual orthogonal subspaces. The WP function $W_{j,k}^n(t)$ is mathematically expressed as below [24]:

$$W_{j,k}^n(t) = 2^{j/2} W^n(2^j t - k), \quad (4)$$

where j and k are the scale and translation parameters, respectively. n is the oscillation parameter. So the WP coefficients $S_{j,k}^n$ are obtained by the inner product between the signal $x(t)$ and the WP functions $W_{j,k}^n$, as below:

$$S_{j,k}^n = \langle x, W_{j,k}^n \rangle = \int_{-\infty}^{\infty} x(t) W_{j,k}^n(t) dt. \quad (5)$$

2.2. Deep Neural Network. Feature vector selection and dimension reduction facilitate the pattern classification. Deep neural network (DNN) attempts to abstract the high-level data structure by using multiple nonlinear transformation. From Figure 1, the main processing step includes layered pretraining and fine-tuning. In this paper, the Autoencoder DNN structure is applied to process the multidimensional feature data.

2.2.1. Pretraining. During the pretraining stage, an Autoencoder network consists of three layers: the input layer, the hidden layer, and the output layer. Firstly, the input layer x is mapped into the hidden layer y with the equation

$$y = s(\mathbf{W}x + b), \quad (6)$$

where \mathbf{W} is the weight matrix, b is the bias value, and s is a transfer function; usually it is nonlinear such as the sigmoid

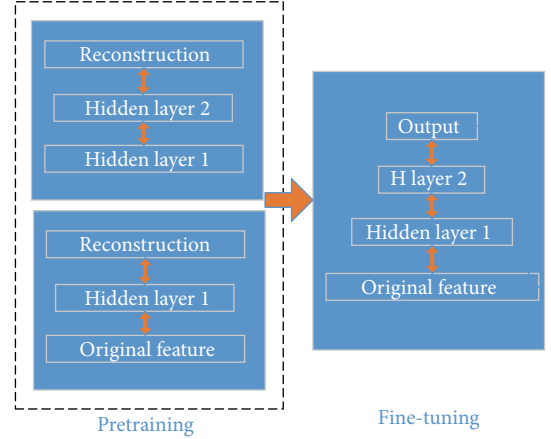


FIGURE 1: The framework for deep learning.

function. This processing is called encoder. Secondly, the decoder processing maps the hidden layer y into the output layer z with the equation:

$$z = s(\mathbf{W}'y + b'), \quad (7)$$

where the weight matrix \mathbf{W}' is constrained to the transpose of the encoder mapping: $\mathbf{W}' = \mathbf{W}^T$. The output value z is seen as the prediction of input x with the code value y . So the network is optimized by minimizing the error between the values z and x , such as $L(xz) = \|x - z\|^2$. So the loss function can be described as

$$L_H(x, z) = -\sum_{k=1}^d [x_k \log z_k + (1 - x_k) \log (1 - z_k)]. \quad (8)$$

The code layer (hidden layer) y has fewer nodes than the input and output layer. It can be seen as the compressed representation of the input layer. This is similar to the PCA data dimensionality reduction. But the Autoencoder is the nonlinear method.

2.2.2. Fine-Tuning. The multihidden layers of the Autoencoder are unsupervised pretraining. Then those hidden layers are combined with the Softmax classifier to construct the deep neural network. The fine-tuning trains the whole neural network with the supervised learning method to improve the performance. During the fine-tuning, all layers are stacked into a single model, so all of the model parameters can be optimized. Backpropagation algorithm for fine-tuning Autoencoder is used. And the detailed equations and derivation process can be found in [25].

2.3. Structure of the Proposed Method. To estimate the operation condition of rolling bearing, a new feature extraction and process scheme is proposed in this paper. The goal of the method is to extract the bearing condition's indicator features and fuse those features to the recognition model to indicate the bearing operation condition. Then the deep neural network reduces the features matrix dimension nonlinearly

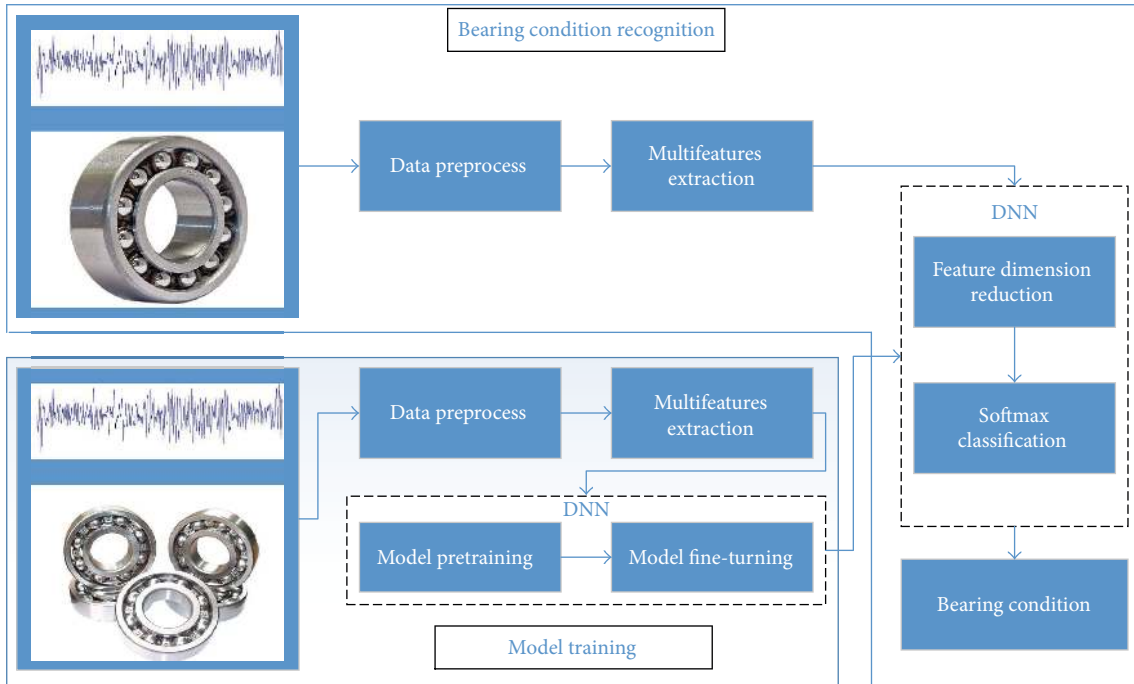


FIGURE 2: The framework for bearing condition estimation.

and recognizes the bearing condition. As shown in Figure 2, the proposed method is described as follows.

Step 1. Preprocess the raw signal with normalization.

Step 2. Extract features of the input signal with time domain, frequency domain, and time-frequency domain.

Step 3. Smooth the features by moving average filter, and then normalize the filtered features.

Step 4. Pretrain the Autoencoder neural network with the features of training and testing set by unsupervised learning.

Step 5. Fine-tune the stacked Autoencoder neural network with the label training set.

Step 6. Estimate the bearing operation condition under the trained model.

3. Experiments Analysis and Discussion

3.1. Description of the Platform PRONOSTIA. The experimentation platform is PRONOSTIA (Figure 3) which is designed to validate bearing performance degradation and condition monitoring assessment. In the present experimental test-bed a series of experiments about bearing degradation process is performed. All kinds of failure, inner race fault, outer race fault, ball fault, and so on, will have the chance to occur in the experiment.

The platform is mainly consisting of drive part and load part. In the driven part, the asynchronous motor (250 W) with a gearbox drives the shaft rotating. The testing bearing

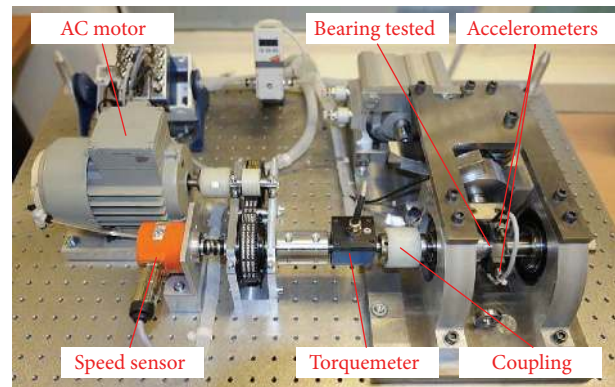


FIGURE 3: The experimental platform PRONOSTIA.

inner race is fixed through the shaft. The load part is produced by a force actuator (*pneumaticjack*), which exerts an exceeding force of the bearing's maximum dynamic load on the testing bearing to reduce the bearing's life cycle. Two accelerometers (*DYTRAN3035B*) are mounted horizontally and vertically on the housing of the test roller bearing to pick up the horizontal and the vertical accelerations as shown in Figure 3. The sampling frequency of the data acquisition is 25600 Hz and the vibration data provided by the two accelerometers are collected every 1 second [26, 27].

The bearing degradation data consists of three categories: the first operating condition (1800 rpm and 4000 N), the second operating condition (1650 rpm and 4200 N), and the third operating condition (1500 rpm and 5000 N). Empirically, the bearing full life cycle can be divided into four

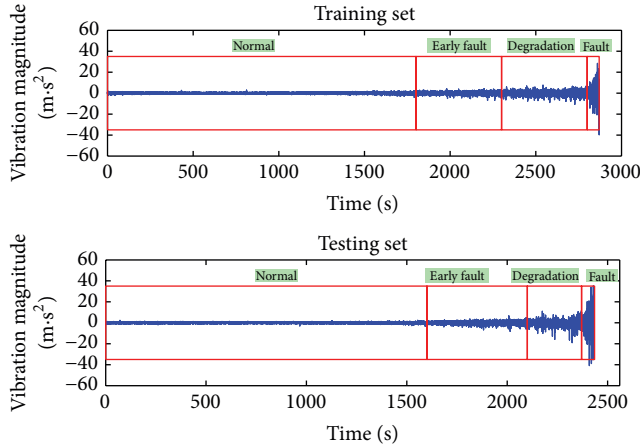


FIGURE 4: The run-to-fail vibration signal of bearing1.1 and bearing1.3.

failure levels: normal, early fault, degradation, and failure as shown in Figure 4. There are six run-to-fail datasets. And the number of different testing bearings are 2830, 871, 911, 797, 515, and 1637. So the total samples are 7534. Half of them (3767) are chosen randomly as the training samples. The other half are used as the testing samples.

3.2. Feature Extraction. The raw vibration signals are non-stationary and nontrendable which makes it difficult to obtain the condition estimation. According to Section 2, the time domain, frequency domain, and time-frequency domain feature extraction methods are applied to indicate the characteristics of bearing vibration signal.

As Table 1, the eight classical time domain features are shown in Figures 5(a)–5(h). From those figures, the features are almost trend but also with some impulse noise. Moving average filter is performed to smooth those features with an average size 15, determined empirically. The red curve is the filtered feature which is more smooth especially eliminating the influence of impulse. Normalization of features are conducted to transfer the features within the same scale.

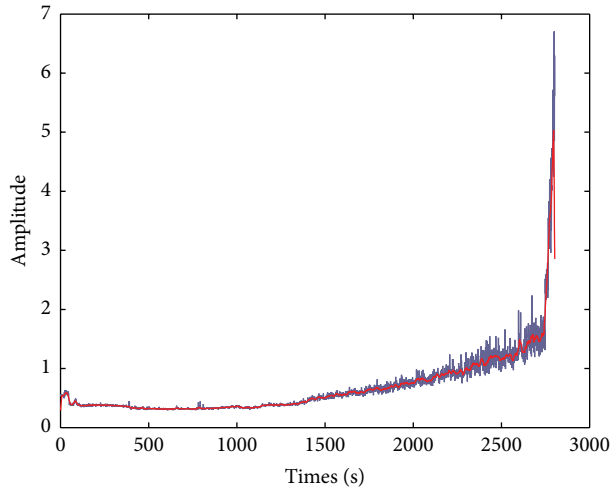
For the spectrum kurtosis (SK) in frequency domain, Welch's estimate of the SK is employed to indicate the locations in the frequency domain of the raw signal [28, 29]. Set window length $W_n = 2^8$, overlap length $O_n = W_n * 3/4$, and FFT length $F_n = 2 * W_n$. The length of raw signal is 2560; through (1), we can get the spectrum kurtosis G_{SK} . Select mean of the first five G_{SK} as the standard template of each vibration dataset and the correlation coefficient can be calculated by (2). Hence, Figure 5(i) is regarded as the bearing monitoring index from the frequency domain feature. As visual inspection, the correlation coefficient curves of spectrum kurtosis are of significant tendency and noiseless. Time-frequency domain features based on wavelet packet decomposition provide arbitrary time-frequency resolution of specified signal. From Figure 6(a), at the bearing initial operation stage, the frequency component is mainly focused on 400~600 Hz. But with the increase of running time, there are certain changes that come out for the concentrated

frequency band. At the end of bearing service, it moves down to 0~200 Hz as shown in Figure 6(b).

Based on these observations, we quantify the spread of energy in the time-frequency plane using the sum of absolute amplitude. Over time, the different time-frequency distributions have different curve trends. From Figure 6(c), we separate the six frequency bands 0~150 Hz, 151~300 Hz, 301~450 Hz, 451~600 Hz, and 601~800 Hz. For the severe oscillation of frequency band curve, the length of moving average filter is set to 21. Figure 6(d) is the filtered and normalized curves. Thus, eight time domain features, one frequency domain feature, and five time-frequency domain features for individual horizontal and vertical vibration signal are detected. They are smoothed by the average moving filter. And for improving the convergence speed of the regression algorithm, all the features are normalized.

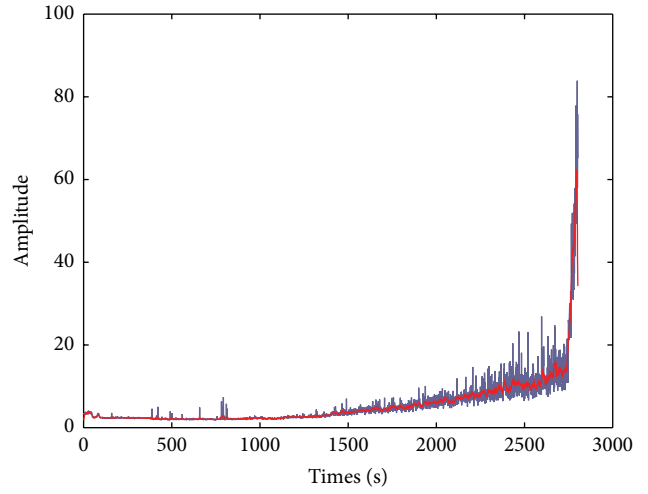
3.3. Results and Discussion. Through all above feature figures, not all of them have the same important degrees to identify the bearing condition. So before those features are fed into the classifier, we reduce their dimension through deep Autoencoder neural network. And it includes two stages: pretraining stage and fine-tuning stage. The pretraining is an unsupervised learning process. So in this case, we use the feature data from training and testing set to train the Autoencoder networks. Thus in the first operation condition, there are 7534 samples from the six run-to-fail sets with different operation condition. In our experiment, the neuron's number of input layer is 28, the first hidden layer is 20, the second hidden layer is 5, and the output layer is 4. During the first hidden layer training, the input layer's neurons are 28, the hidden layer's neurons are 20, and the output layer's neurons are 28. Through (6)~(8), the BP algorithm is used to train the Autoencoder network. When the training network arrives at the stop condition, the first hidden layer L_1 can be got. Then the first hidden layer L_1 is assigned to input layer for the second hidden layer training. The hidden layer's neurons are 5. Each layer is trained as a denoising Autoencoder by minimizing the reconstruction of its input. Once all layers are pretrained, the network goes through the second stage of training (fine-tuning). Fine-tuning is a supervised learning process where we want to minimize prediction error on a supervised task. A Softmax classifier layer is added to the top of the network. We then train the whole neural network to get the optimal parameters. And the transfer function of the last layer is tanh. The results are shown in Table 2. The degeneration stage works out the most accuracy rate with 98.2%. For comparison, PCA dimension reduction, SVM classification, and unpretraining method is applied to the dataset to test the classification rate.

3.3.1. PCA Dimension Reduction Method. Principle component analysis (PCA) is an effective linear data dimension reduction method [30] and is often used for bearing fault diagnosis and classification. PCA has ability to extract the most significant representation features. In the experiment, 28 features are extracted in every sample. PCA is applied to the training feature sets, and just five most distinct principle



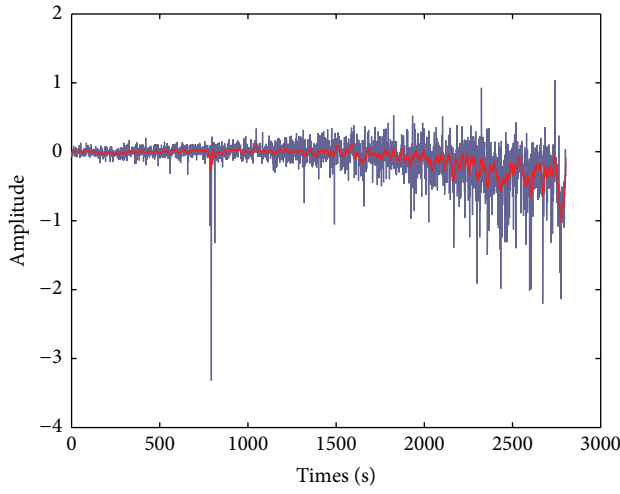
— Corresponding

(a) RMS



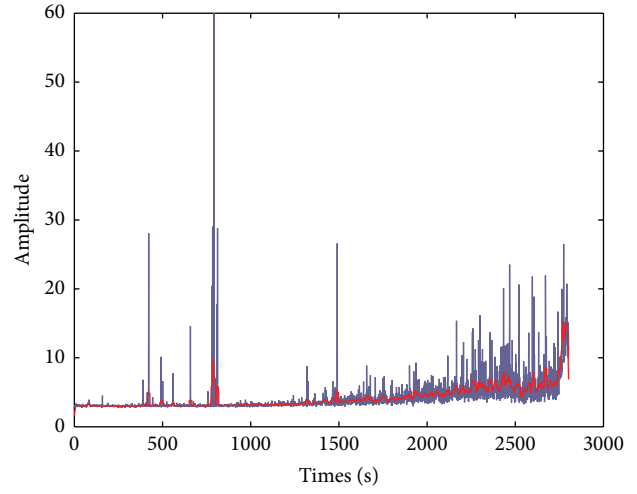
— Corresponding

(b) Peak to Peak



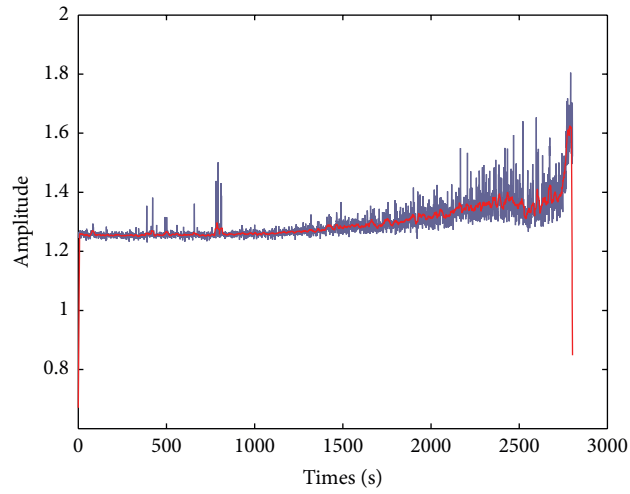
— Corresponding

(c) Skewness



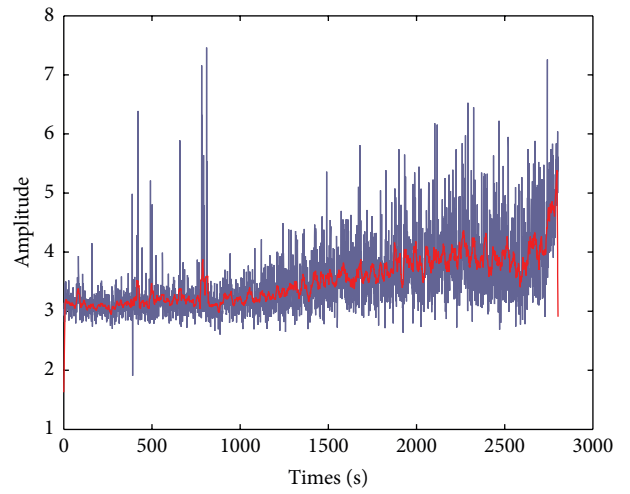
— Corresponding

(d) Kurtosis



— Corresponding

(e) Crest Factor



— Corresponding

(f) Shape Factor

FIGURE 5: Continued.

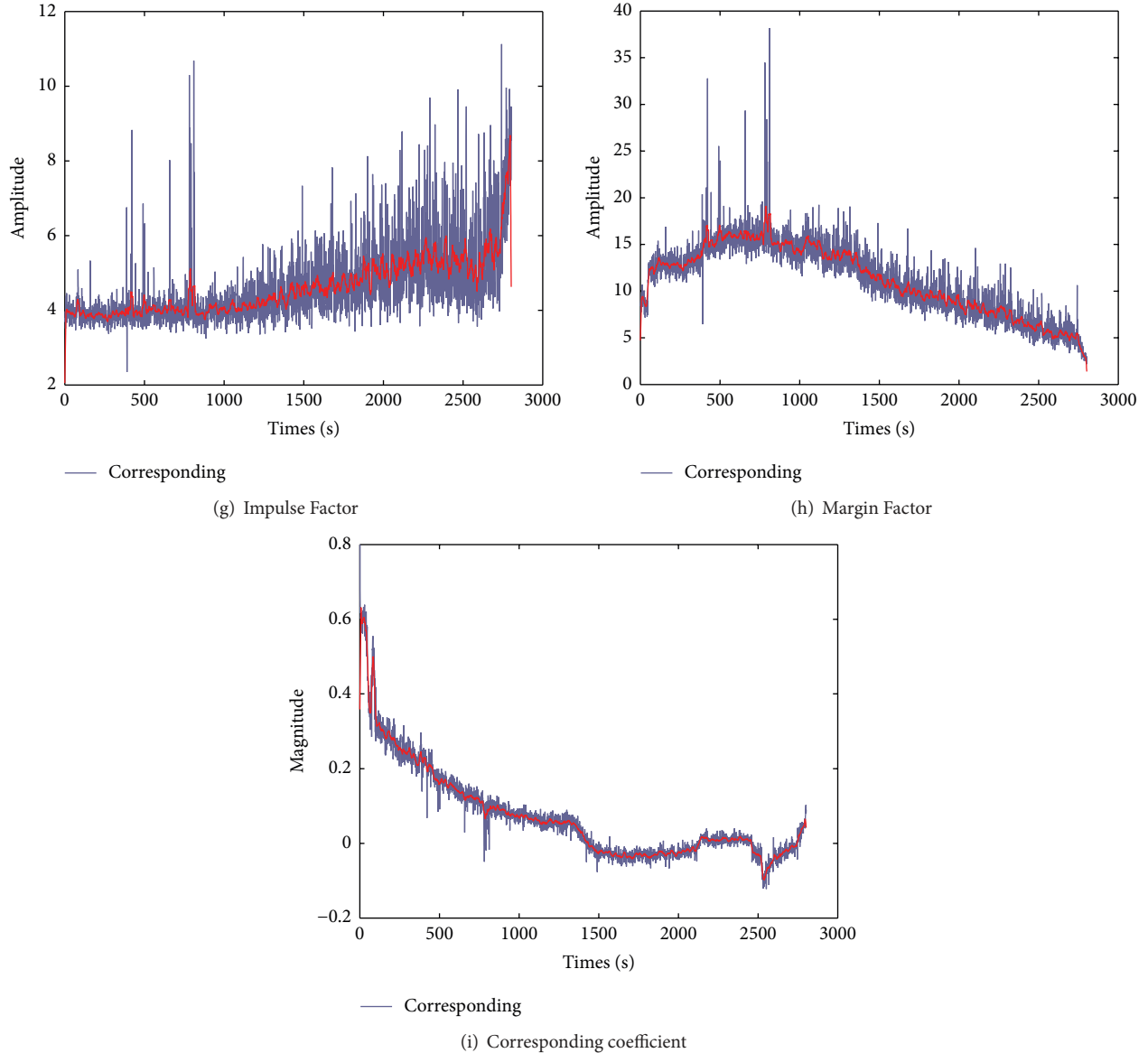


FIGURE 5: (a)~(h) are the time domain features and (i) is the frequency domain feature of bearing1_1.

components are accounted. Then the transformation matrix based on the training set is applied to the test signal. At last the compressed features are used for classification by BP neural network with three layers. The results can be found in Table 2.

3.3.2. SVM Classification Method. Support vector machine (SVM) is a classical supervised learning model with the rigorous statistical learning theory. The kernel of SVM model is the construction of kernel function implicitly mapping the input data into the high-dimensional spaces. And SVMs have been found to be remarkably effective in many machinery fault diagnoses [31]. In many bearing fault diagnosis applications [32], the RBF kernel obtains the significant classification accuracy rate. In the experiment, we mainly consider the RBF kernel. And the raw features input directly into SVM classification; the result is shown in Table 2.

3.3.3. The Influence of Unsupervised Prelearning. An important improvement of deep learning is the unsupervised learning which does not need the label data to train the neural network [33]. And the unsupervised learning can find the appropriate initial weights values which are helpful for optimizing the weights in nonlinear deep learning. For validating the significant of pretraining, the deep learning model is built without unsupervised prelearning; its result is shown in Table 2. The result indicates that the unsupervised prelearning can improve the recognition accurate rate of bearing condition.

3.3.4. The Influence of Different Features. To investigate the influence of different features, different feature combination is analyzed. First, the features are divided into three categories: traditional features (Feature 1), traditional features and

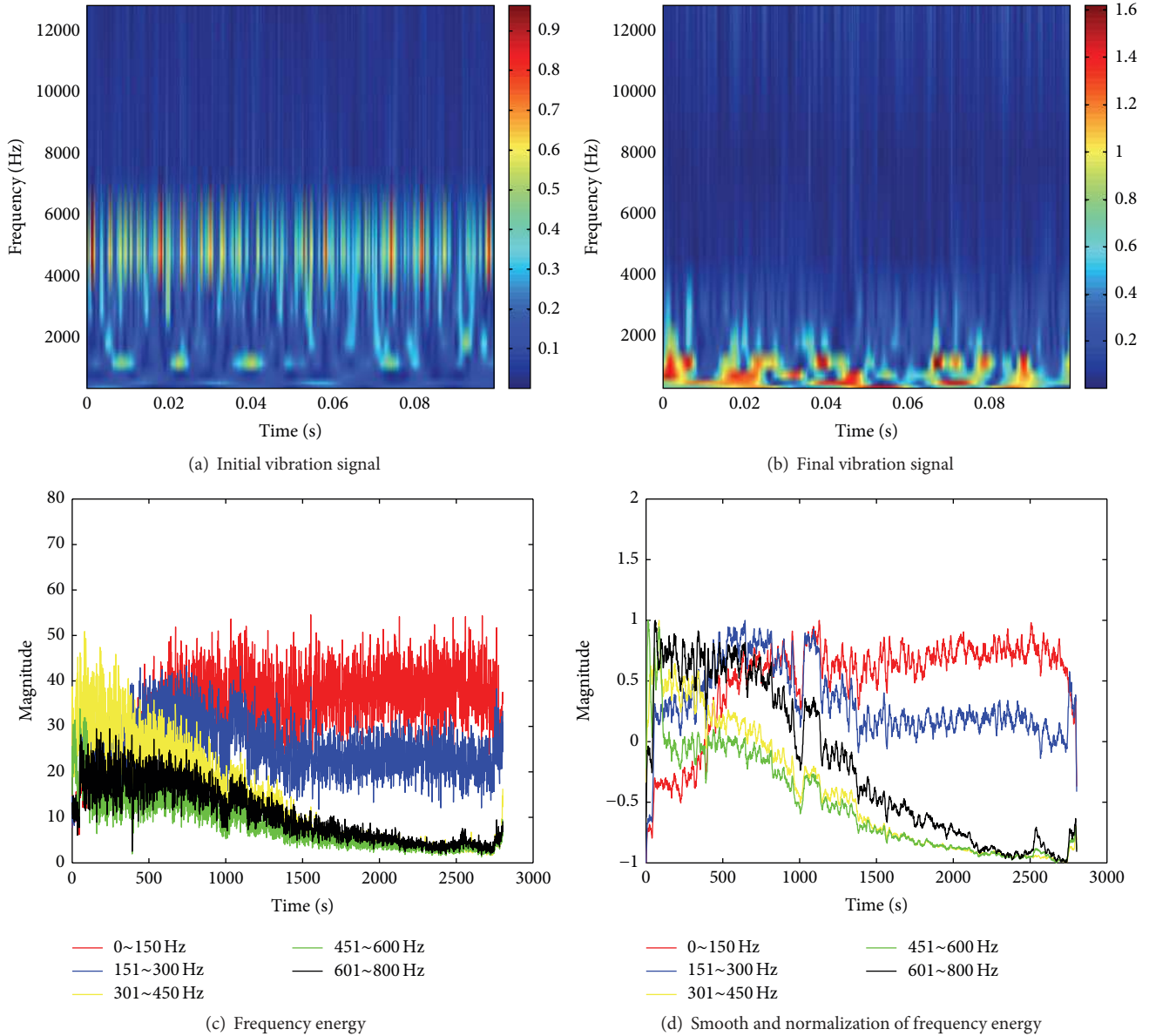


FIGURE 6: The time-frequency domain features of bearing1_1.

TABLE 2: The results of different algorithms.

	Proposed method				Proposed	PCA	SVM	Unpretraining
	Normal	Early fault	Degeneration	Fault				
Normal	0.965	0.035	0	0	0.965	0.915	0.940	0.915
Early	0.132	0.783	0.085	0	0.783	0.788	0.774	0.777
Degeneration	0	0.004	0.982	0.014	0.982	0.950	0.935	0.960
Fault	0	0	0.048	0.952	0.952	0.922	0.933	0.922
Success rate	0.965	0.783	0.982	0.952	0.935	0.898	0.909	0.897

spectrum kurtosis (Feature 2), traditional features, spectrum kurtosis, and WPT (Feature 3). Then those features' combination is reduced dimension and distinguished condition classification by deep learning model. The result can be found

in Table 3. As you can see, the proposed feature extraction method can improve the bearing condition recognition accuracy. And the spectrum kurtosis and WPT features are also helpful.

TABLE 3: The results of different features combination.

Features	Accuracy
Feature 1	0.892
Feature 2	0.913
Feature 3	0.935

4. Conclusion and Future Works

Condition-based maintenance based on condition monitoring is critical to reduce the cost of maintenance and improve production efficiency of industry. In this paper, a novel method based on multifeatures extraction, deep neural network features dimensional reduction, and condition recognition is proposed. Time, frequency, and time-frequency domain features are fused to represent the characteristic of bearing operation condition. Then the nonlinear dimension reduction method based on deep learning is proposed to highlight the hidden patterns and to compress the information. At last, a classification in which the different bearing operation stage is identified is added to the top level of deep neural network. This condition monitoring system is validated with real-world vibration monitoring data collected from bearing. A comparative study is performed between the proposed method and PCA dimension reduction method, SVM classification method, unpretraining method, and three different features' combination. Those results demonstrate the advantage of the proposed method in achieving more accurate condition estimation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (no. 51275426) and Innovation Funds of Southwest Jiaotong University.

References

- [1] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [2] P.-L. Zhang, B. Li, S.-S. Mi, Y.-T. Zhang, and D.-S. Liu, "Bearing fault detection using multi-scale fractal dimensions based on morphological covers," *Shock and Vibration*, vol. 19, no. 6, pp. 1373–1383, 2012.
- [3] H. Liao, E. A. Elsayed, and L.-Y. Chan, "Maintenance of continuously monitored degrading systems," *European Journal of Operational Research*, vol. 175, no. 2, pp. 821–835, 2006.
- [4] S. A. McInerney and Y. Dai, "Basic vibration signal processing for bearing fault detection," *IEEE Transactions on Education*, vol. 46, no. 1, pp. 149–156, 2003.
- [5] R. B. W. Heng and M. J. M. Nor, "Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition," *Applied Acoustics*, vol. 53, no. 1, pp. 211–226, 1998.
- [6] L. C. K. Reuben and D. Mba, "Bearing time-to-failure estimation using spectral analysis features," *Structural Health Monitoring*, vol. 13, no. 2, pp. 219–230, 2014.
- [7] C. Ruiz-Cárcel, E. Hernani-Ros, Y. Cao, and D. Mba, "Use of spectral kurtosis for improving signal to noise ratio of acoustic emission signal from defective bearings," *Journal of Failure Analysis and Prevention*, vol. 14, no. 3, pp. 363–371, 2014.
- [8] P. Borghesani, R. Ricci, S. Chatterton, and P. Pennacchi, "A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions," *Mechanical Systems and Signal Processing*, vol. 38, no. 1, pp. 23–35, 2013.
- [9] D. F. Shi, W. J. Wang, and L. S. Qu, "Defect detection for bearings using envelope spectra of wavelet transform," *Journal of Vibration and Acoustics—Transactions of the ASME*, vol. 126, no. 4, pp. 567–573, 2004.
- [10] 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.
- [11] S. Abbasion, A. Rafsanjani, A. Farshidianfar, and N. Irani, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine," *Mechanical Systems and Signal Processing*, vol. 21, no. 7, pp. 2933–2945, 2007.
- [12] B. Samanta and K. R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features," *Mechanical Systems and Signal Processing*, vol. 17, no. 2, pp. 317–328, 2003.
- [13] G. F. Bin, J. J. Gao, X. J. Li, and B. S. Dhillon, "Early fault diagnosis of rotating machinery based on wavelet packets—empirical mode decomposition feature extraction and neural network," *Mechanical Systems and Signal Processing*, vol. 27, no. 1, pp. 696–711, 2012.
- [14] C. Chen and G. Vachtsevanos, "Bearing condition prediction considering uncertainty: an interval type-2 fuzzy neural network approach," *Robotics and Computer-Integrated Manufacturing*, vol. 28, no. 4, pp. 509–516, 2012.
- [15] G. Wang, X. Feng, and C. Liu, "Bearing fault classification based on conditional random field," *Shock and Vibration*, vol. 20, no. 4, pp. 591–600, 2013.
- [16] S. Şeker, E. Ayaz, and E. Türkcan, "Elman's recurrent neural network applications to condition monitoring in nuclear power plant and rotating machinery," *Engineering Applications of Artificial Intelligence*, vol. 16, no. 7-8, pp. 647–656, 2003.
- [17] G. S. Vijay, S. P. Pai, N. S. Sriram, and R. B. K. N. Rao, "Radial basis function neural network based comparison of dimensionality reduction techniques for effective bearing diagnostics," *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, vol. 227, no. 6, pp. 640–653, 2013.
- [18] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–27, 2009.
- [19] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [20] L. Deng, G. Hinton, and B. Kingsbury, "New types of deep neural network learning for speech recognition and related applications: an overview," in *Proceedings of the 38th IEEE International Conference on Acoustics, Speech and Signal Processing*

- (ICASSP '13), pp. 8599–8603, IEEE, Vancouver, Canada, May 2013.
- [21] D. Ciregan, U. Meier, and J. Schmidhuber, “Multi-column deep neural networks for image classification,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 3642–3649, IEEE, Providence, RI, USA, June 2012.
- [22] Y. Goldberg and O. Levy, “word2vecexplained: deriving Mikolov et al.’s negative-sampling word-embedding method,” <http://arxiv.org/abs/1402.3722>.
- [23] J. J. G. de la Rosa and A. Moreno Muñoz, “Higher-order cumulants and spectral kurtosis for early detection of subterranean termites,” *Mechanical Systems and Signal Processing*, vol. 22, no. 2, pp. 279–294, 2008.
- [24] G. G. Yen and K.-C. Lin, “Wavelet packet feature extraction for vibration monitoring,” *IEEE Transactions on Industrial Electronics*, vol. 47, no. 3, pp. 650–667, 2000.
- [25] A. Ng, *Sparse Autoencoder*, vol. 72 of *CS294A Lecture Notes*, 2011.
- [26] P. Nectoux, R. Gouriveau, K. Medjaher et al., “Pronostia: an experimental platform for bearings accelerated degradation tests,” in *Proceedings of the Conference on Prognostics and Health Management*, pp. 1–8, Minneapolis, Minn, USA, September 2012.
- [27] E. Sutrisno, H. Oh, A. S. S. Vasan, and M. Pecht, “Estimation of remaining useful life of ball bearings using data driven methodologies,” in *Proceedings of the IEEE Conference on Prognostics and Health Management (PHM '12)*, pp. 1–7, Denver, Colo, USA, June 2012.
- [28] J. Antoni, “The spectral kurtosis: a useful tool for characterising non-stationary signals,” *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 282–307, 2006.
- [29] J. Antoni and R. B. Randall, “The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines,” *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308–331, 2006.
- [30] M. D. Farrell Jr. and R. M. Mersereau, “On the impact of PCA dimension reduction for hyperspectral detection of difficult targets,” *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 2, pp. 192–195, 2005.
- [31] Q. Hu, Z. He, Z. Zhang, and Y. Zi, “Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble,” *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 688–705, 2007.
- [32] J. Yang, Y. Zhang, and Y. Zhu, “Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension,” *Mechanical Systems and Signal Processing*, vol. 21, no. 5, pp. 2012–2024, 2007.
- [33] Q. V. Le, “Building high-level features using large scale unsupervised learning,” in *Proceedings of the 38th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '13)*, pp. 8595–8598, IEEE, Vancouver, Canada, May 2013.



Hindawi

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

