

# Multi-view feature fusion for rolling bearing fault diagnosis using random forest and autoencoder

Sun Wenqing Deng Aidong Deng Minqiang Zhu Jing Zhai Yimeng Cheng Qiang Liu Yang

(National Engineering Research Center of Turbo-Generator Vibration, Southeast University, Nanjing 210096, China)

(School of Energy and Environment, Southeast University, Nanjing 210096, China)

**Abstract:** To improve the accuracy and robustness of rolling bearing fault diagnosis under complex conditions, a novel method based on multi-view feature fusion is proposed. Firstly, multi-view features from perspectives of the time domain, frequency domain and time-frequency domain are extracted through the Fourier transform, Hilbert transform and empirical mode decomposition (EMD). Then, the random forest model (RF) is applied to select features which are highly correlated with the bearing operating state. Subsequently, the selected features are fused via the autoencoder (AE) to further reduce the redundancy. Finally, the effectiveness of the fused features is evaluated by the support vector machine (SVM). The experimental results indicate that the proposed method based on the multi-view feature fusion can effectively reflect the difference in the state of the rolling bearing, and improve the accuracy of fault diagnosis.

**Key words:** multi-view features; feature fusion; fault diagnosis; rolling bearing; machine learning

**DOI:** 10.3969/j.issn.1003-7985.2019.03.005

The rolling bearing is one of key parts in a wind turbine drive train. Due to the harsh operating environment of the wind turbine, rolling bearing failures occur frequently. According to statistics, 30% of the rotating machinery failures are caused by rolling bearings<sup>[1]</sup> and about 80% of the wind turbine gearbox failures are caused by bearing failure<sup>[2]</sup>. Therefore, bearing fault diagnosis is essential for efficient and reliable operation of the wind turbines.

Traditionally, the fault diagnosis of wind turbine rolling bearing is based on the spectrum analysis of vibration signals<sup>[3]</sup>. The key technology of it is to extract the fault characteristic frequency from noisy signals. Methods of spectrum analysis include the Fourier transform<sup>[4]</sup>, Hilbert transform along with some joint time-frequency analysis methods such as empirical mode decomposition (EMD)<sup>[5]</sup>

and variational mode decomposition (VMD). The traditional methods only study the bearing vibration signal from a certain perspective, and the features are manually extracted depending on much prior knowledge about signal processing techniques and diagnostic expertise<sup>[6]</sup>, which cannot meet the requirements of real-time and portability of fault diagnosis in the era of big data.

In order to comprehensively analyze the difference between faults, it is necessary to examine the vibration signal of the bearing from multiple perspectives, so as to grasp the overall state of the bearing. In this paper, both the spectrum analysis of the vibration signal and the time-frequency analysis are performed. Then, the features are extracted from the time domain, frequency domain (frequency spectrum and envelope spectrum) and time-frequency domain (EMD). Although multi-view features are highly complementary, these features tend to be redundant, which is not conducive to fault diagnosis. Therefore, feature selection and feature fusion before classification are necessary.

Feature selection is the process of selecting some of the most effective features of the original features to reduce the dimension of the feature set, and it is an important means to improve the performance of the learning algorithm<sup>[7]</sup>. The way to select features is to sort them based on certain evaluation criteria, and then determine the number of features selected according to needs. Commonly used evaluation criteria are feature missing values, feature variance, Pearson correlation coefficients, etc. In recent years, many studies have used the random forest (RF) model for feature selection<sup>[8-9]</sup>. Random forest derives the importance of features based on the performance of the training data, by calculating the importance of each feature in each tree, then taking a weighted average to achieve the final feature importance assessment.

Feature fusion is to reduce feature redundancy after feature selection. The methods can be divided into linear fusion and nonlinear fusion. Common linear fusion methods are principal components analysis (PCA) and linear discriminant analysis (LDA). Locally linear embedding (LLE) and autoencoder are typical nonlinear fusion methods. In the case of autoencoder (AE), compression and decompression of features are implemented in an unsupervised manner via a neural network<sup>[10]</sup>. Fea-

**Received** 2019-04-16, **Revised** 2019-08-15.

**Biographies:** Sun Wenqing (1994—), male, graduate; Deng Aidong (corresponding author), male, doctor, professor, dnh@seu.edu.cn.

**Foundation item:** The National Natural Science Foundation of China (No. 51875100).

**Citation:** Sun Wenqing, Deng Aidong, Deng Minqiang, et al. Multi-view feature fusion for rolling bearing fault diagnosis using random forest and autoencoder[J]. Journal of Southeast University (English Edition), 2019, 35(3): 302 – 309. DOI: 10.3969/j.issn.1003-7985.2019.03.005.

ture selection and fusion make it easy to train the model and visualization. The deeper meaning is to transform raw features into a new and concise representation<sup>[11]</sup>.

The multi-view feature set can make full use of the information of the original signal to reflect the difference among states. The extraction of multi-view features and the reduction of redundancy by feature fusion are two essential components in this model.

In this paper, a new strategy for rolling bearing fault diagnosis using multi-view features is proposed. The scheme consists of two parts including feature extraction and feature fusion. Specifically, features are extracted from the time domain, frequency domain and time frequency domain. Then, the random forest model (RF) and autoencoder are employed for feature selection and feature fusion, respectively. Finally, the support vector

machine (SVM)<sup>[12]</sup> is introduced to evaluate the features.

The main contributions are summarized as follows:

- 1) The technology of feature extraction from multiple perspectives is proposed in this work, which takes the advantages of signal processing in feature extraction.
- 2) Aiming at the redundancy of multi-view features, a novel feature fusion strategy based on the random forest and the autoencoder is developed.
- 3) The validity and superiority of the proposed method in practical applications are verified through experimental analyses.

## 1 Signal Processing and Feature Extraction

The flow chart of the proposed method is shown in Fig. 1. The main part of the method includes feature extraction and feature fusion.

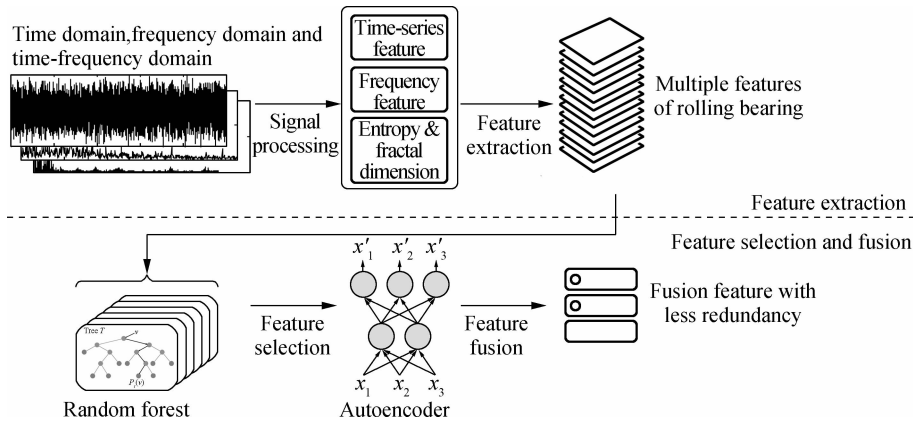


Fig. 1 The framework of the proposed method

In this section, the method of feature extraction of the vibration signal from the time domain, frequency domain and time frequency domain is introduced.

### 1.1 Time domain features

The original vibration signal contains the information about the normal vibration component, fault vibration component and environmental noise. In order to reduce dependence on prior knowledge, 43 statistical features are extracted from the timing waveform. Some features are as follows:

Mean

$$f_1 = \frac{1}{N} \sum_{n=1}^N x(n) \quad (1)$$

Maximum amplitude

$$f_2 = \frac{1}{N} \sum_{n=1}^N |x(n)| \quad (2)$$

Maximum peak

$$f_3 = \frac{1}{2} (\max(x(n)) - \min(x(n))) \quad (3)$$

Standard deviation

$$f_4 = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (x(n) - p_1)^2} \quad (4)$$

Square root amplitude

$$f_5 = \left( \frac{1}{N} \sum_{n=1}^N \sqrt{|x(n)|} \right)^2 \quad (5)$$

Skewness

$$f_6 = \frac{1}{N-1} \sum_{n=1}^N \left( \frac{x(n) - p_1}{p_4} \right) \quad (6)$$

Kurtosis

$$f_7 = \frac{1}{N-1} \sum_{n=1}^N \left( \frac{x(n) - p_1}{p_4} \right)^4 \quad (7)$$

Clearance factor

$$f_8 = \frac{f_3}{f_5} \quad (8)$$

### 1.2 Frequency domain features

The data points in the time domain must first be transferred to the frequency domain by the Fourier transform and Hilbert transform, and then the signal can be ana-

lyzed in the frequency domain. The frequency domain features include all the basic 43 statistical features, and other features are constructed for the frequency, shown as follows:

Average frequency

$$f_1 = \frac{1}{N_{\text{fft}}} \sum_{j=1}^{N_m} X(j) \quad (9)$$

Center frequency

$$f_2 = \frac{\sum_{j=1}^{N_m} [f(j)X(j)]}{\sum_{j=1}^{N_m} X(j)} \quad (10)$$

Frequency root mean square

$$f_3 = \sqrt{\frac{\sum_{j=1}^{N_m} [f(j)^2 X(j)]}{\sum_{j=1}^{N_m} X(j)}} \quad (11)$$

Frequency standard deviation

$$f_4 = \sqrt{\frac{\sum_{j=1}^{N_m} \{[f(j) - f_{\text{FC}}]^2 X(j)\}}{\sum_{j=1}^{N_m} X(j)}} \quad (12)$$

### 1.3 Time frequency domain features

EMD decomposes signals into a finite number of signals defined as intrinsic mode functions (IMFs) from high frequency to low frequency of unequal bandwidth adaptively. The internal volatility of the signal is reflected in the extracted IMFs which include real physical information of the signal. Except the basic 43 statistical features, entropy features and fractal dimension are also extracted from selected IMFs.

1) Entropy feature. Entropy usually reflects the degree of chaos in the signal, which is a function that characterizes the uncertainty of the signal. Assuming that the IMF sequence is  $X = \{x_1, x_2, \dots, x_n\}$ , the probability of each data point occurring is  $p_i = P(x(i))$ .

$$\sum_{i=1}^{N_i} p_i = 1 \quad (13)$$

Then, the information entropy of the IMF signal can be expressed as

$$F = - \sum_{i=1}^{N_i} p_i \ln p_i \quad (14)$$

2) Fractal dimension. Box dimension is the degree of irregularity and complexity of characterizing fractal sets at different scales. The fractal dimension can describe the structural features of signals. Specific calculation methods can refer to Ref. [13].

## 2 Feature Selection and Fusion

### 2.1 Feature selection

In this step, a random forest is applied to the training

data, then the decision of all trees in the forest is aggregated for all the data based on the most voting for the classification<sup>[14]</sup>. The feature importance assessment using RF is to calculate contributes of each features to each tree in the random forest, then take the average and finally compare the contribution between different features<sup>[15]</sup>.

Metrics of contribution include Gini index (or  $G$  for short) and out-of-bag data (OOB). In the evaluation method, the calculation of the Gini index is

$$G_m = 1 - \sum_{k=1}^{|K|} P_{mk}^2 \quad (15)$$

where  $K$  is the number of categories;  $P_{mk}$  represents the proportion of category  $k$  in node  $m$ . The importance of the feature  $X_i$  at node  $m$  is the change in the Gini index before and after branching at node  $m$ :

$$V_{\text{im}}^{\text{gini}} = G_m - G_l - G_r \quad (16)$$

where  $V_{\text{im}}^{\text{gini}}$  is the importance of the feature  $X_i$  at node  $m$ ;  $G_l$  and  $G_r$  represent the Gini index of two new nodes after branching, respectively. If the node where feature  $X_i$  appears in decision tree  $i$  is in set  $m$ , then the importance of  $X_i$  in the  $i$ -th tree is

$$V_{ij}^{\text{gini}} = \sum_{m \in M} V_{jm}^{\text{gini}} \quad (17)$$

Assume that the forest has a total of  $n$  trees, then

$$V_j^{\text{gini}} = \sum_{i=1}^n V_{ij}^{\text{gini}} \quad (18)$$

All the obtained importance scores are normalized to obtain the score of the importance of the feature  $X_i$ :

$$V_j = \frac{V_j}{\sum_{i=1}^c V_i} \quad (19)$$

### 2.2 Feature fusion

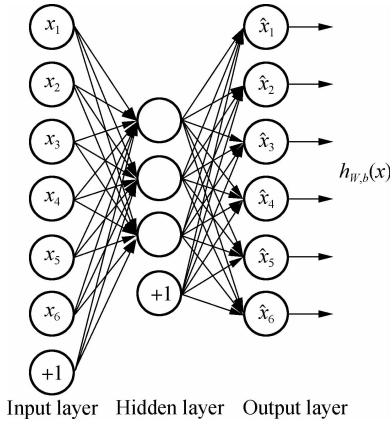
Standardization before data input can effectively improve the convergence speed and training effect, since the autoencoder is a neural network model. All the feature vectors were standardized by removing the mean and scaling to unit variance. The standard score of a sample  $x$  is calculated as

$$z = \frac{x - u}{s} \quad (20)$$

where  $u$  is the mean of the training samples; and  $s$  is the standard deviation of the training samples.

The autoencoder consists of an encoder and a decoder. To learn more abstract features, feature fusion can be performed by operating on the trained encoder. The basic architecture of an autoencoder is shown in Fig. 2.

The first layer is the input layer, the middle layer is the



**Fig. 2** Architecture of an autoencoder

hidden layer and the last layer is the output layer. The autoencoder compares the output of the feedforward network with the input and feeds the loss back into a recursive structure, which is powerful and easy to implement in an unsupervised manner<sup>[16]</sup>. The autoencoder network output layer has the following relationship with the input layer:  $\hat{x}_i \approx x_i$ . The objective function is

$$J(W, b) = \frac{1}{m} \sum_{i=1}^m (\hat{x}_i - x_i)^2 \quad (21)$$

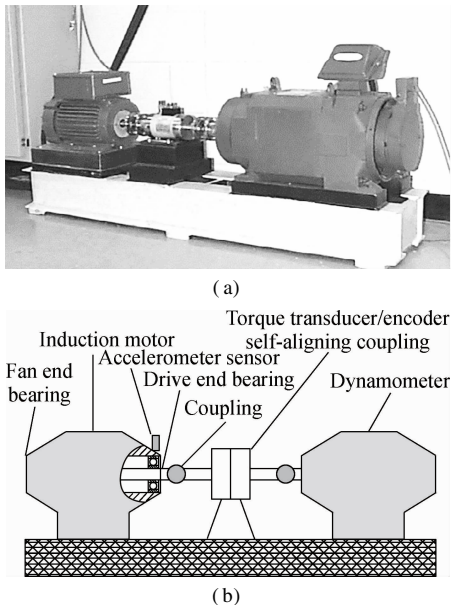
where  $W$  is the weight of hidden layer nodes;  $b$  is the bias of hidden layer neurons;  $m$  is the number of hidden layer neurons.

The trained encoder can conduct a nonlinear transformation to reduce the dimensionality of high-dimensional features by back propagation and gradient descent.

### 3 Experiments

#### 3.1 Datasets

Bearing vibration data is obtained from Case Western Reserve University<sup>[17]</sup>. The test stand is shown in Fig. 3.



**Fig. 3** Schematic diagram of the experiment platform. (a) Test rig; (b) Schematic

The test stand consists of a 2-hp motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown). The test bearings support the motor shaft. Single point faults were introduced to the test bearings using electro-discharge machining with fault diameters of 2.13, 4.26 and 6.39 mm. An accelerometer was attached to the motor housing at the drive end of the motor.

Vibration data was collected at a sampling frequency of 12 kHz and each sample contains 4 000 points. In order to more clearly demonstrate the gain of the proposed method for classification accuracy, noise is added to the bearing vibration signal used, with a SNR of  $-2$ . The details about the bearing datasets are shown in Tab. 1.

**Tab. 1** Descriptions of bearing datasets

Number of samples	Fault diameter/mm	Defect	Class label
60	0	Normal	0
30	2.13	Inner race	1
30	2.13	Balls fault	2
30	2.13	Outer race@ 3	3
30	2.13	Outer race@ 6	4
30	2.13	Outer race@ 12	5
30	4.26	Inner race	6
30	4.26	Balls fault	7
30	4.26	Outer race@ 6	8

The dataset contains 300 signal samples covering nine different conditions, i. e., normal condition, ball fault, inner race fault, and outer race fault. When the model is trained, the entire sample set is randomly divided into the training set and the test set. The training set contains 200 samples, and the test set includes 100 samples.

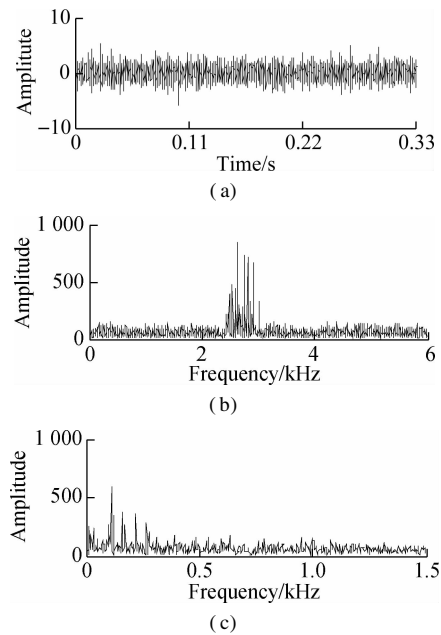
#### 3.2 Signal transformation and decomposition

The feature extraction for the vibration signal is from three aspects: time domain, frequency domain and time-frequency domain. The time domain waveform, frequency spectrum and envelope spectrum of the signal with the outer race fault are given in Fig. 4.

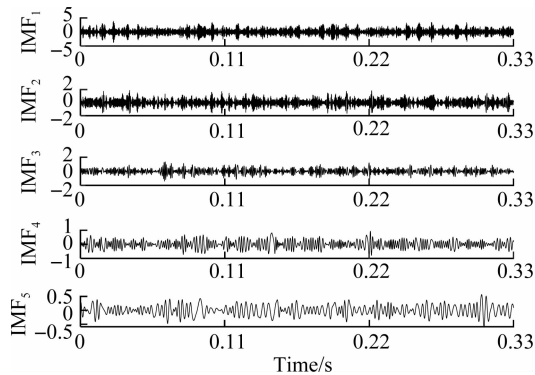
The waveforms of the first five IMFs containing the main feature information are given in Fig. 5. Time-frequency domain features are extracted from the first five IMF components obtained after EMD decomposition.

#### 3.3 Feature extraction

The time domain features are statistical features. The frequency domain features include statistical features and four frequency related features based on the frequency spectrum and envelope spectrum, and the time-frequency domain features include statistical features, three entropy features, and a box-dimensional feature. The dimensions of time domain features, frequency domain features and time-frequency features are 43, 94 and 235, respectively, as shown in Tab. 2. Therefore, features with a dimension of 372 are extracted for each sample.



**Fig. 4** Analysis of measured vibration signal. (a) Time domain waveform; (b) Frequency spectrum; (c) Envelope spectrum



**Fig. 5** The first five IMFs yielded by EMD

Tab. 2 Feature set		
Time domain	Frequency domain	Time-frequency domain
43 statistic features	86 statistic features	215 statistic features
	8 frequency features	15 entropy feature
		5 fractal dimension
43 features	94 features	235 features

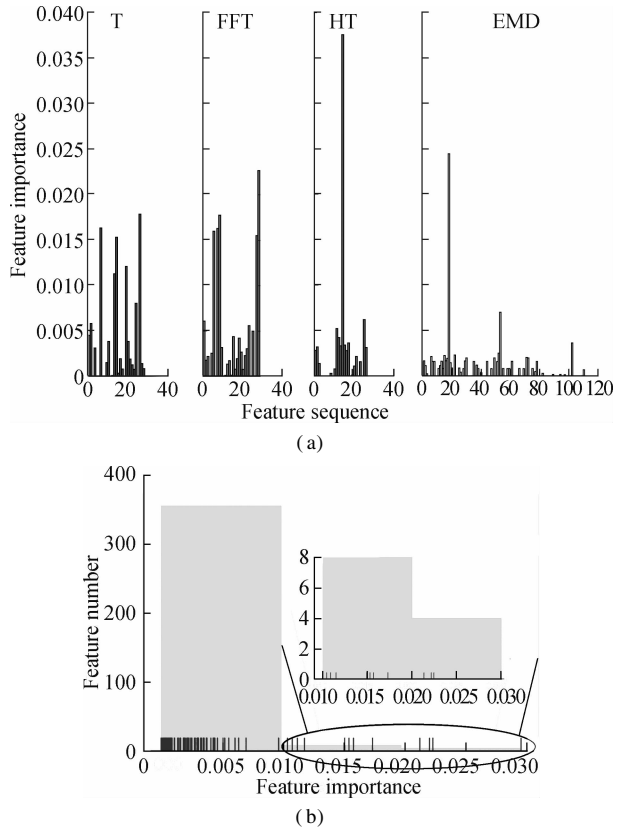
A multi-view feature set can comprehensively reveal the conditions of the bearing from multiple perspectives. However, among them there are many features that do not change with fault conditions, which make fault recognition more difficult. Usually, a good classification result is not guaranteed if all the features are directly fed to the classifier without feature fusion.

3.4 Feature selection and fusion

When using the random forest model (RF) for feature selection, the features and labels of the training set are fed into the model for training. The trained RF model can give the value of each feature based on its performance on the decision tree branches. The specific parameters of RF are shown in Tab. 3.

Tab. 3 Parameters of random forest and autoencoder		
Fusion model	Hyperparameter	Optimization results
Random forest	$n\_estimators$	100
	$learning\_rate$	0.1
	$min\_leaf$	1
	Loss	Deviance
	$max\_depth$	3
	$min\_split$	2
Autoencoder	Number of hidden layers	1
	Node per layer	120-30-120
	Regularizer	$L1(10^{-4})$
	Loss	Binary_crossentropy
	Epoch	250
	Batch_size	32

Fig. 6 shows the importance of each feature from four different feature sets including time domain features (T), the frequency domain consisting of frequency spectrum (FFT) and envelope spectrum (HT), and the time frequency domain from the IMFs (EMD). The histogram of feature importance is shown in Fig. 6(b). For the convenience of observation, the first picture has features removed that are less than 0.0001 in importance.

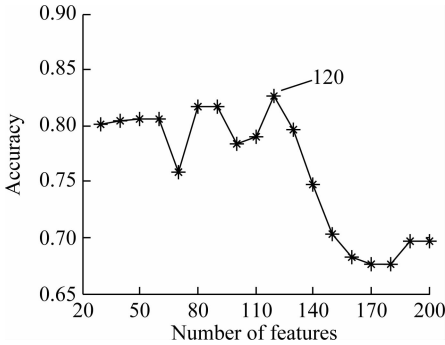


**Fig. 6** The importance of the features given by the random forest. (a) The importance of each feature; (b) The histogram of feature importance

As can be seen from Fig. 6(a), the features extracted from the original waveform and frequency spectrum are generally more important than those from the envelope spectrum and IMFs. To more clearly reveal the difference between states, it is necessary to take the important fea-

tures and discard useless ones based on feature importance.

It is evident in Fig. 6(b) that only a few features are of high importance. However, the optimization of the number of selected features still needs to be executed since a small number of important features may fail to guarantee a good reconstruction of fault characteristics. The optimization of feature selection is shown in Fig. 7.



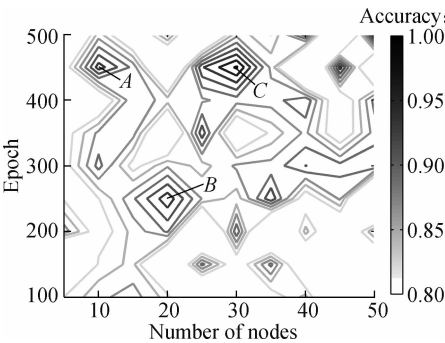
**Fig. 7** Relationship between the number of selected features and the accuracy of SVM classification

To find the optimal number of the selected features, the model is evaluated by setting the validation set, where the training set is further divided into a small training set and a validation set. The validation size is set to be 0.33. It can be found in Fig. 7 that the classification accuracy is not the best when the 20 most important features are selected. Therefore, in order to improve the accuracy of the classifier, it is necessary to select some features that are not very important. In this experiment, the number of selected features is 120.

The random forest (RF) can eliminate the irrelevant features which fusion methods cannot remove completely and the autoencoder can further reduce the redundancy of the selected features.

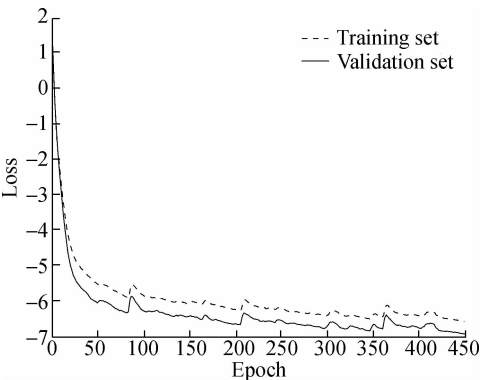
The autoencoder model used in this experiment has a hidden layer and uses L1 regularization to prevent model overfitting. Specific parameters are shown in Tab. 3.

As an unsupervised machine learning model, the autoencoder has two important parameters to optimize, the number of hidden layer nodes and the number of epochs. The results optimized by grid search are shown in Fig. 8.



**Fig. 8** Grid research optimization results

As shown in Fig. 8, there are three combinations that allow the SVM to obtain good classification results, marked as points A, B and C. The number of hidden layer nodes is 10, 20, and 30, and the epoch is 450, 250, and 450, respectively. To maximize the accuracy, point C is selected. Therefore, the number of hidden layer nodes and the number of epoches are determined as 30 and 450. The errors of autoencoder on the training set and the validation set are shown in Fig. 9.



**Fig. 9** Autoencoder errors

3.5 SVM classification results

To prove the superiority of the proposed feature extraction and fusion method, the SVM classification results of the proposed method and other methods are listed in Tab. 4.

From the perspective of error generation, the error mainly comes from the outer race fault classification. The differences of these four outer race faults are merely the angle or depth of the crack. The time domain features cannot fully reveal the difference between the faults. After introducing the features of the frequency domain and the time-frequency domain, the classification error is reduced.

According to the results in Tab. 4, compared with the SVM performance of the original features of the time domain and all the original features, the improvement of the classification accuracy is not obvious. The increase in the features in the frequency domain and the time-frequency domain will introduce a large number of irrelevant features, which cause the difference between faults to be obscured. Therefore, when random forest is used to remove a large number of less important features, the accuracy of the classifier has been greatly improved. The accuracy rate is increased from 89% to 97%. Further-more, when using the autoencoder to reduce the redundancy of the feature set, the classification accuracy is improved by 2% and achieves 99%.

To further reveal the superiority of feature dimension reduction using the RF and autoencoder, comparisons are made between three different feature fusion methods, including PCA, kernel PCA and locally linear embedding (LLE) with the proposed methods. The results are shown

Tab. 4 SVM classification accuracy of different feature sets

Feature set	N	Normal	Balls fault		Inner-race fault		Outer-race fault				Average accuracy/%
		Fault 0	Fault 2	Fault 7	Fault 1	Fault 6	Fault 3	Fault 4	Fault 5	Fault 8	
The original features of time domain	43	89.47	100	100	78.57	84.62	60.00	100	100	58.33	87.02
All the original features	372	100	100	100	81.25	100	42.86	100	80.00	52.94	89.10
RF	120	100	100	100	100	100	85.71	100	80.00	81.82	96.71
PCA	40	100	100	76.92	61.90	100	100	100	50.00	77.78	87.59
KPCA	40	17.95	100	100	65.00	100	100	100	23.08	90.00	91.47
LLE	10	100	100	85.50	100	100	100	100	75.00	88.59	96.13
RF + AE	40	100	100	100	100	100	100	100	100	90.00	99.10

in Tab. 4, where  $N$  represents the dimension of the fused features.

KPCA selects a better performing polynomial kernel function instead of the Gaussian kernel. As for the PCA and KPCA, parameter  $n\_components$  is set to be 40. Limited by the algorithm, LLE integrates the original features into 10 dimensions and  $n\_neighbors$  is set to be 100.

It can be observed that the features of PCA fusion are the worst in SVM classification, and they even fail to achieve the accuracy before fusion. In contrast, the KPCA shows the difference clearly and the accuracy increases by 2.3% compared to that of the original features. LLE demonstrates a great improvement in accuracy, mainly due to its good nonlinear mapping ability. Among these models, the proposed RF + AE model has the highest accuracy, which further illustrates the robustness of the methods for extracting effective information and reducing feature redundancy.

4 Conclusions

1) Multi-view features can fully grasp the fault state of the bearing. After feature selection and fusion, features from multiple views can clearly reveal the state difference between normal and fault conditions. Experiments show that the fault feature set can be constructed well when the features of the vibration signals are extracted from the time domain, frequency domain and the time-frequency domain.

2) Combined with the feature selection and fusion method of the random forest and autoencoder in this paper, the accuracy of bearing fault classification can be effectively improved. The classification accuracy reaches 99.10%, which exceeds the accuracy of the feature set from the single perspective and outperforms other feature fusion methods.

3) In future studies, more features will be added to achieve better classification results and the performance of the fused features can be enhanced by using a deeper autoencoder. In addition, the proposed method can be applied to the fault diagnosis of gearboxes and life prediction of rotating machinery.

References

[1] Hossain M, Abu-Siada A, Mueeen S. Methods for advanced wind turbine condition monitoring and early diagnosis: A literature review[J]. *Energies*, 2018, **11**(5): 1309. DOI: 10.3390/en11051309.

[2] Yang W, Tavner P J, Crabtree C J, et al. Cost-effective condition monitoring for wind turbines[J]. *IEEE Transactions on Industrial Electronics*, 2010, **57**(1): 263 – 271. DOI: 10.1109/tie.2009.2032202.

[3] Kay S M, Marple S L. Spectrum analysis: A modern perspective[J]. *Proceedings of the IEEE*, 1981, **69**(11): 1380 – 1419. DOI: 10.1109/proc.1981.12184.

[4] Boashash B. *Time-frequency signal analysis and processing: A comprehensive reference* [M]. Oxford: Elsevier Science, 2016: 8 – 9.

[5] Chen Y, Zhou C, Yuan J, et al. Application of empirical mode decomposition in random noise attenuation of seismic data [J]. *Journal of Seismic Exploration*, 2014, **23**: 481 – 495.

[6] Jia F, Lei Y G, Lin J, et al. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data [J]. *Mechanical Systems and Signal Processing*, 2016, **72/73**: 303 – 315. DOI: 10.1016/j.ymssp.2015.10.025.

[7] Bishop C. *Pattern recognition and machine learning*[M]. New York: Springer-Verlag, 2006: 559 – 561.

[8] Cao W H, Xu J P, Liu Z T. Speaker-independent speech emotion recognition based on random forest feature selection algorithm [C]// *Proceedings of the 36th Chinese Control Conference (CCC)*. Dalian, China, 2017. DOI: 10.23919/chicc.2017.8029112.

[9] Kursa M B, Rudnicki W R. Feature selection with the BorutaPackage[J]. *Journal of Statistical Software*, 2010, **36**(11): 1 – 13. DOI: 10.18637/jss.v036.i11.

[10] Liou C Y, Cheng W C, Liou J W, et al. Autoencoder for words[J]. *Neurocomputing*, 2014, **139**: 84 – 96. DOI: 10.1016/j.neucom.2013.09.055.

[11] Zhao X P, Wu J X, Zhang Y H, et al. Fault diagnosis of motor in frequency domain signal by stacked de-noising auto-encoder [J]. *Computers, Materials & Continua*, 2018, **57**(2): 223 – 242. DOI: 10.32604/cmc.2018.02490.

[12] Zhang X Y, Liang Y T, Zhou J Z, et al. A novel bearing fault diagnosis model integrated permutation entropy, en-

semble empirical mode decomposition and optimized SVM [J]. *Measurement*, 2015, **69**: 164 – 179. DOI: 10.1016/j.measurement.2015.03.017.

[13] Karperien A. *Defining microglial morphology: Form, function, and fractal dimension* [M]. Queensland, Australia: Charles Sturt University, 2004: 99 – 102.

[14] Forouzannezhad P, Abbaspour A, Cabrerizo M, et al. Early diagnosis of mild cognitive impairment using random forest feature selection[C]//2018 *IEEE Biomedical Circuits and Systems Conference (BioCAS)*. Cleveland, OH, USA: IEEE, 2018. DOI: 10.1109/biocas.2018.8584773.

[15] Li B Q, Cai Y D, Feng K Y, et al. Prediction of protein cleavage site with feature selection by random forest[J]. *PLoS One*, 2012, **7**(9): e45854. DOI: 10.1371/journal.pone.0045854.

[16] Chen Z Y, Li W H. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network [J]. *IEEE Transactions on Instrumentation and Measurement*, 2017, **66**(7): 1693 – 1702. DOI: 10.1109/tim.2017.2669947.

[17] Zheng J D, Cheng J S, Yang Y. Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis[J]. *Mechanical Systems and Signal Processing*, 2013, **40**(1): 136 – 153. DOI: 10.1016/j.ymssp.2013.04.005.

基于随机森林和自编码的滚动轴承多视角特征融合

孙文卿 邓艾东 邓敏强 朱 静 翟怡萌 程 强 刘 洋

(东南大学火电机组振动国家工程研究中心, 南京 210096)

(东南大学能源与环境学院, 南京 210096)

**摘要:**为了提高复杂工况下滚动轴承故障诊断的准确性和鲁棒性,提出了基于振动信号多视角特征深度融合的故障诊断方法. 首先,通过傅里叶变换、希尔伯特变换以及经验模态分解(EMD)提取滚动轴承振动信号的时域、频域及时频域特征;然后,利用随机森林模型(RF)选择多视角特征集中与轴承运行状态相关性较大的特征,将选出特征通过自编码器(AE)进行深度融合以进一步降低特征的冗余性;最后,将融合后的多视角特征作为支持向量机(SVM)的输入,以对其有效性进行评价. 实验结果表明,提出的多视角特征融合的故障诊断方法能有效地反映轴承运行状态之间的差异,提高了滚动轴承故障诊断的准确率.

**关键词:**多视角特征;特征融合;故障诊断;滚动轴承;机器学习

**中图分类号:**TH133.33