An Intelligent Bearing Fault Diagnosis Method Based on SF-SVM

Bao'an Qiu, Pan Sun* and Lili Li
The 14th Research Institute of China Electronics Technology, Nanjing, Jiangsu, China
*corresponding author’s e-mail: 562251022@qq.com

Abstract. Rolling bearing, as a key component of rotating machinery, its health status directly determines the stability and reliability of the whole machine. The research on its intelligent diagnosis method has important engineering value and academic significance. However, due to actual engineering conditions, the types of bearing failures and the amount of data are limited. Aiming at the difficulty of extracting and selecting bearing vibration features under limited sample constraints, this paper proposes an intelligent fault diagnosis method of SF-SVM. On the basis of the short-time Fourier change, the L2 regularized sparse filter is used to extract the unsupervised feature of the bearing vibration time-frequency map. After obtaining the typical features of the bearing, the support vector machine is used for diagnosis.

1. Introduction
With the rapid development of perception technology and information technology, people's demand for digitization, informatization and intelligence of rotating machinery is higher and higher. Rolling bearing, as a key component of rotating equipment. Its health state during operation directly determines the stability and reliability of equipment[1,2]. According to statistics, the proportion of rotating machinery faults caused by rolling bearing failure is 45% ~ 55%[3].

The research on the fault diagnosis method of rolling bearing can provide guidance for the predictive maintenance of bearing and overall equipment, prolong the service life of bearing and reduce the operation and maintenance cost of bearing, which has important engineering value and academic significance.

During the operation of the bearing, the signals obtained by the sensor in real time are high-dimensional and large quantities, but limited by the actual situation, the type and number of faults are limited. Traditional diagnosis methods rely too much on expert experience for feature extraction and feature selection, and deep learning methods need a lot of data for training and learning. Therefore, aiming at the difficulty of bearing vibration feature extraction and selection under small samples, this paper proposes a fault diagnosis method of SF-SVM.

2. Research status at home and abroad
Bearing fault diagnosis is to extract features from the collected signals to judge the health status of bearings. Traditional fault diagnosis methods include KNN [4-6], support vector machine SVM [7,8], neural network [9,10], etc. W. Qin et al. [7] used IMF information entropy as the feature input of SVM to realize the fault diagnosis of rolling bearing. Tang G [11] solved the problem of effectively extracting fault features and selecting high-precision classifier, and proposed a rolling bearing fault diagnosis method based on improved fast spectral correlation and optimized random forest.

With its strong expression ability and hierarchical learning ability, deep learning algorithm has developed rapidly in the fields of speech recognition and target detection, and is also favored by...
researchers in the field of rotating machinery fault diagnosis. Aiming at the problem that traditional feature extraction depends on expert experience and prior knowledge, Jia f et al.[12] proposed a fault diagnosis method based on deep neural network (DNN). Olivier Janssens et al.[13] proposed a feature learning model based on convolutional neural network, which directly learns the typical features used for bearing state judgment from the original data. Qi y et al.[14] proposed a fault diagnosis method based on stacked sparse self encoder, which can extract more representative high-dimensional features from bearing vibration signals. Gan m et al.[15] extracted the fault characteristics of bearing signals with the help of wavelet packet transformation and used them as the input of deep confidence network to realize the fault diagnosis of rolling bearing.

To sum up, traditional methods need to use a priori knowledge and engineering experience for feature extraction, while deep learning methods can be fully trained on the premise of enough training samples. Therefore, it is necessary to study the bearing diagnosis method under the condition of limited samples.

3. Fault diagnosis method based on SF-SVM

The algorithm framework of fault diagnosis is shown in Fig.1, including signal preprocessing, feature extraction and fault diagnosis. Signal preprocessing is the preliminary processing of the collected signals, including signal truncation, short-time Fourier transform, division of training set and test set, etc. Feature extraction, that is, using the sparse filtering method to extract the features with sparsity within and between samples from the time-frequency map, so as to reduce the dimension of the signal. In fault identification, under the constraint of limited samples, soft interval support vector machine is selected to suppress over fitting by introducing relaxation variables into the loss function.

**Figure 1.** Algorithm framework

3.1. Signal preprocessing - short time Fourier transform

The standard Fourier transform has the problem that the frequency domain information and time domain information cannot be localized at the same time. The short-time Fourier transform divides the signal into many small time intervals. Analyzing each time interval with Fourier transform can master the local frequency of the signal and the time period information locked out of the frequency.

Given a signal \( x(t) \in L^2(\mathbb{R}) \), its STFT is defined as follows.
\[ STFT_t(\tau, \Omega) = \int x(\tau) g(t-\tau) e^{-j\Omega \tau} d\tau \] (1)

The meaning of this formula is to intercept the signal \( x(T) \) with the window function \( g(T) \) in the time domain, and Fourier transform the intercepted local signal. By continuously moving the center of the window function, the Fourier transform at different times can be obtained. The set of these Fourier transforms is \( STFT_t(\tau, \Omega) \).

3.2. Feature extraction

3.2.1. Sparse filtering principle. Sparse filtering is an unsupervised feature learning method proposed by Ngiam [16], which extracts the features of high dispersion and population sparsity by optimizing the L1 / L2 norm of the feature matrix. Suppose the sample data is \( X = [x_1, x_2, ..., x_i, ..., x_M] \), where \( x_i \in \mathbb{R}^N \), i.e. \( x \in \mathbb{R}^{NM} \), and its corresponding weight matrix is \( w \in \mathbb{R}^{NF \times N} \), then the feature matrix \( F \) corresponding to \( F = Wx \), where \( F \in \mathbb{R}^{N \times M} \). The feature matrix is expressed as \( F = [f_1, f_2, ..., f_i, ..., f_M] \), where \( f_i \) represents feature vector corresponding to sample data \( x_i \) and it can be calculated as follows.

\[ f_i^{(j)} = W^{(j)} \times x_i \] (2)

Then, the same features of all training samples are normalized by L2 norm, and the feature vector after L2 norm normalization of the \( j \)-th feature of all samples \( f^{(j)} \) is recorded as follows.

\[ f_i^{(j)} = \frac{f_i^{(j)}}{\| f_i^{(j)} \|_2} \] (3)

The feature matrix \( F \) is obtained after row normalization, and then its each column is normalized by L2 norm, that is, all transformed features of the \( i \)-th sample are normalized by column, and the transformed sample vector is recorded as \( \hat{f}_i \).

\[ \hat{f}_i = \frac{\hat{f}_i}{\| \hat{f}_i \|_2} \] (4)

Finally, the characteristic evaluation function, namely the cost function, is obtained.

\[ \min_{W} \sum_{i=1}^{M} \| f_i \|_1 = \sum_{i=1}^{M} \| \hat{f}_i \|_1 / \| \hat{f}_i \|_2 \] (5)

The loss function can be solved by finite memory quasi Newton method (L-BFGS).

3.2.2. Feature extraction based on sparse filtering. Sparse filtering is used for unsupervised feature extraction to obtain the typical features of bearings with dispersion and sparsity. In order to improve the feature extraction ability of sparse filter under sample limited constraints, L2 regularization is carried out, that is, L2 norm of feature mapping matrix is introduced into its objective function \( W \), as shown below.

\[ \min_{W} \sum_{i=1}^{M} \| \hat{f}_i \|_1 = \sum_{i=1}^{M} \| \hat{f}_i \|_1 / \| \hat{f}_i \|_2 + \lambda \| W \|_2 \] (6)

3.3. Fault diagnosis

In order to alleviate the over fitting phenomenon of support vector machine under limited samples and noise interference, soft interval support vector machine is selected to introduce relaxation variables for
each sample to characterize the degree that the sample does not meet the constraints. The corresponding objective function is as follows.

$$\min_{w, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{m} \xi_i$$

(7)

C > 0 is called the penalty parameter. While maximizing the interval, the number of samples that do not meet the constraints is required to be as small as possible.

4. Experimental verification

4.1. Experiment introduction

This experiment selects the public data set of the bearing data center of Case Western Reserve University [17]. As shown in the figure 2, the bearing fault simulation test-bed is composed of drive motor, frequency converter, load motor, vibration acceleration sensor, collector, etc. There are 4 test conditions in total, as shown in Table 1 below.

In this experiment, the fault of the bearing is simulated by machining defects of different scales at different parts of the bearing. The specific types are shown in Table 1. The fault of the bearing is expressed in the form of xx-x, where XX represents the fault location, if represents the inner ring fault, of represents the outer ring fault, Ba represents the rolling element fault, X represents the size of the fault scale, and the larger the number, the larger the fault scale.

![Figure 2. Experiment table](image)

Table 1. Bearing fault parameters.

| Fault type                  | Fault scale (mm) | Fault symbol representation |
|-----------------------------|------------------|-----------------------------|
| inner race/outer race/ball  | 1.78             | IF-1/OF-1/BA-1              |
| inner race/outer race/ball  | 3.56             | IF-2/OF-1/BA-2              |
| inner race/outer race/ball  | 5.15             | IF-3/OF-3/BA-3              |

4.2. Parameter selection

The method proposed in this paper needs to select several parameters, including the length of input samples, the number of extracted features, sparse filtering and regularization parameters \(\lambda\), Penalty coefficient C of support vector machine. Therefore, the parameter selection was studied, and each test was repeated 30 times to reduce the influence of random factors.

4.2.1. Sample length. In order to study the influence of sample length, that is, the input dimension of sparse filtering, on the diagnosis results, 0.5% samples are randomly selected for model training and regularization parameters \(\lambda=0.01\), penalty coefficient \(C=0.1\), the number of features extracted by sparse filtering is set to 100, and the diagnosis results are shown in Figure 4. When the sample length is within \([10005000]\), the diagnostic accuracy and stability are almost unchanged, indicating that the sample length has little impact on the diagnostic results. In this paper, the sample length is set to 1000.
4.2.2. **Feature number.** The influence of the number of features extracted by sparse filtering on the diagnosis results is shown in Fig.4 With the gradual increase of the number of features extracted, the diagnosis accuracy and stability are also gradually improved. When the number of features is greater than 30, the diagnosis tends to be stable, so this paper is set $\lambda = 30$.

![Graph showing the influence of feature number on accuracy](image)

**Figure 4.** The influence of feature number

4.2.3. **Parameter $\lambda$.** Sparse filter regularization parameters $\lambda$. The influence on the diagnosis results is shown in Fig.5, with the regularization parameters $\lambda$. The diagnostic accuracy and stability are also gradually improved. When $\lambda \geq 0.01$, the diagnosis tends to be stable, so it is set in this paper $\lambda = 0.01$.

![Graph showing the influence of parameter $\lambda$ on accuracy](image)

**Figure 5.** The influence of parameter $\lambda$
4.2.4. Parameter C. The influence of the penalty coefficient C on the diagnosis result is shown in Figure 6. When C is in the [0,0.1] range, the diagnosis accuracy and stability are almost unchanged with the change of C, indicating that the penalty coefficient C has little influence on the diagnosis result, so C = 0.1 is set in this paper.

4.2.5. Train set ratio. The influence of the proportion of training set on the diagnosis results is shown in Fig.7. When the proportion of training set is in the interval of [0.01,0.1], the diagnosis accuracy and stability gradually improve with the increase of the proportion of training set. When the proportion of training set is greater than 0.05, the diagnosis accuracy tends to be stable, indicating that the model is trained more fully with the increase of training sample data.
4.3. Comparison of results
In order to verify the fault diagnosis model in this paper, the diagnosis methods proposed in this paper are compared with the existing methods, as shown in Table 2, from the ratio of training samples, operating conditions of bearings, number of bearing States, recognition rate, etc.

Compared with the results of the existing literature, this method has better recognition performance under complex working conditions with less training samples.

Table 2. Comparison of diagnosis results

| Method            | Condition | Train set ratio | Number of bearing types | Accuracy |
|-------------------|-----------|----------------|-------------------------|----------|
| Wavelet+SVMS[18]  | 1         | 75%            | 10                      | 88.9%    |
| TR-LD[19]         | 4         | 10%            | 10                      | 92.6%    |
| SF+softmax[20]    | 1,2,3,4   | 10%            | 10                      | 99.7%    |
| Proposed Method   | 1,2,3,4   | 0.1%           | 10                      | 99.9%    |

5. Conclusion
Aiming at the difficulty of bearing vibration feature extraction and fault recognition under limited sample constraints, this paper proposes a fault diagnosis method of SF-SVM, uses L2 regularized sparse filter to automatically extract the representative features of bearing from time-frequency diagram, and realizes fault diagnosis by support vector machine. Compared with the existing literature, the superiority of this method is verified.

References
[1] Qi X X, Ji J W, Han X W 2011 Fault Diagnosis Methods of Rolling Bearing: A General Review Key Engineering Materials 480-1-986-92
[2] S. Tyagi, et al 2017 A DWT and SVM based method for rolling element bearing fault diagnosis and its comparison with Artificial Neural Networks Journal of Applied and Computational Mechanics 3(1)
[3] Rai A, Upadhyay S H 2016 A Review on Signal Processing Techniques Utilized in the Fault Diagnos of Rolling Element Bearings Tribology International 96 289–306
[4] Cai E., et al 2015 Fault diagnosis method based on LMD and KNN algorithms for rolling bearing Modern Electronics Technique
[5] Lu D, Qian N and Yang X 2018 Fault Diagnosis of Rolling Bearing Based on KNN-Naive Bayesian Algorithm Computer Measurement & Control
[6] Khelil J, Ramdani M, Khelil K, et al 2019 Bearing Faults Diagnosis Using Discrete Wavelets and Artificial Intelligence Approaches//2019 1st International Conference on
Sustainable Renewable Energy Systems and Applications (ICSRESA)

[7] Qin W, Zhang W and Lu C 2016 Rolling bearing fault diagnosis: A data-based method using EEMD, information entropy and one-versus-one SVM. //2016 12th World Congress on In-telligent Control and Automation (WCICA), Guilin 1016-20

[8] Huo Z, Zhang Y, Shu L and Gallimore M 2019 A New Bearing Fault Diagnosis Method Based on Fine-to-Coarse Multiscale Permutation Entropy, Laplacian Score and SVM. IEEE Access 7 17050-66

[9] Gunerkar R S, Jalan A K & Belgamwar S U 2019 Fault diagnosis of rolling element bearing based on artificial neural network Journal of Mechanical Science and Technology 33(2) 505–11

[10] Al-Raheem K F, et al 2008 Application of the Laplace-Wavelet Combined With ANN for Rolling Bearing Fault Diagnosis International journal of comadem 11(4) 18-24.

[11] Tang G; Pang B; Tian T, Zhou C 2018 Fault Diagnosis of Rolling Bearings Based on Im-proved Fast Spectral Correlation and Optimized Random Forest Applied Science 8(10) 1859

[12] Jia F, Lei Y, Lin J, et al 2016 Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data Mechanical Systems and Signal Processing 72-73 303-15

[13] Olivier Janssens, Viktor Slavkovikj, Bram Vervisch, Kurt Stockman, Mia Loccufier. 2016 Con-volutional Neural Network Based Fault Detection for Rotating Machinery Journal of Sound and Vibration 377 331-45

[14] Qi Y, Shen C, Liu J, et al 2017 An Automatic Feature Learning and Fault Diagnosis Method Based on Stacked Sparse Autoencoder//International Workshop of Advanced Manu-factoring & Automation. Springer, Singapore 7 367-75

[15] Gan M, Wang C, Zhu C 2016 Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings Mechanical Systems and Signal Processing 72(73) 92-104

[16] Ngiam J, Pang W K, Chen Z, et al 2011 Sparse filtering International Conference on Neural Information Processing Systems

[17] http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-univers-

[18] Dy W, Tao J, Li Y, et al 2014 Wavelet leaders multifractal features based fault diagnosis of rotating mechanism Mechanical Systems and Signal Processing 43(s 1–2) 57–75

[19] Jin X, Zhao M, Chow T W S, et al 2014 Motor Bearing Fault Diagnosis Using Trace Ratio Lin-ear Discriminant Analysis IEEE Transactions on Industrial Electronics 61(5) 2441-51

[20] Lei Y, Jia F, Lin J, et al 2016 An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data IEEE Transactions on Industrial Electronics 1-1