A Bearing Fault Diagnosis Method Based on Dictionary Learning and Parameter-Optimized Support Vector Machine

Jing Yang*, Yan Hao, Ting Xu, Huichao Yan, Hongping Hu and Yanping Bai
Department of Mathematics, School of Science, North University of China, Tai Yuan, China

*Corresponding author

Abstract. Aiming at the problem of fault diagnosis, a novel method based on dictionary learning and parameter-optimized Support Vector Machine (SVM) was proposed in this paper and applied it to bearing fault diagnosis. Firstly, the collected bearing fault signals are transformed into gray images after data processing. Then, using dictionary learning, the gray images are denoised and output them as signal data. Finally, the SVM multi-classification model obtained by using Grid Search (GS) algorithm to optimize penalty parameter c and kernel function parameter g is used to classify and identify the fault type. This paper is based on data from Case Western Reserve University Bearing Center for experimental verification. The results show that the proposed model can continuously achieve the accuracy of 100% in the process of bearing fault diagnosis in different environments, which proves that the proposed method can accurately and effectively realize the fault diagnosis classification of bearings.

1. Introduction
Bearing is one of the most important parts in mechanical equipment [1-3]. In addition, the development of machinery and equipment will directly affect the development of multiple industries such as manufacturing, transportation, and aerospace. Therefore, it is of great significance to study bearing fault diagnosis methods.

Different solutions available in the literature to solve this problem have been reviewed [4-6]. For example, In [7], the authors use empirical mode decomposition and energy operator demodulation to extract the features of the fault signal, and then use the feature vector for classification and recognition. By processing the original signal data from multiple methods such as time domain, frequency domain and time-frequency, features are extracted from the encoder to obtain a good diagnosis result in [8]. Sun [9] uses compressed sensing to compress the data, and then uses SAE to extract fault features. This method can still have a good diagnosis effect when it is used in an industrial environment with a large amount of data. Sohaib [10] and others combine a hybrid feature pool with a deep neural network based on sparse stacked self-coding to propose a two-layer rolling bearing fault diagnosis model. In [11], Eren et al proposed a general-purpose rolling bearing intelligent fault diagnosis system based on a compact adaptive one-dimensional convolutional neural network, which provides a new idea to solve the problem of traditional rolling bearing intelligent fault diagnosis methods being limited by training data sets and test data sets. Although some achievements have been made in bearing fault diagnosis research, there is still a lot of research space. Such as, in some studies, high accuracy can only be obtained when using data in a specific environment for classification. And there are some algorithms that do not show good performance when using small samples. In addition,
during the experiment, the accuracy of the fault diagnosis of the bearing changes with the changes in the number of samples in the training set and the test set. This may be due to the instability of the fault diagnosis model. Therefore, we need a bearing fault diagnosis method that has stable performance and can adapt to complex environments.

In view of the above analysis, a bearing fault diagnosis method combining dictionary learning and SVM optimized by GS was proposed to realize the intelligent diagnosis of rolling bearing faults. First, the original signals are denoised through dictionary learning after data processing, and then the bearing fault categories are identified using the SVM classification model optimized by GS. Experiments show that the proposed method can not only achieve accuracy of 100% when diagnosing sample data collected in different environments, but also can maintain accuracy of 100% steadily in experiments on small samples or different training samples and test samples.

2. Theoretical Basis

2.1. Dictionary Learning

The concept of a dictionary was proposed by Mallet, Stephane G, and Zhang Zhifeng [12] for the field of signal processing in 1993. The main idea is to sparsely represent signals through a dictionary defined by prior knowledge. Dictionary learning was proposed by Michal Aharon, Michael Elad, and Alfred Bruckstein [13] in 2006. The difference from the previous dictionary is that dictionary learning no longer uses a dictionary defined based on prior knowledge, but dictionary and signal representation are alternately trained according to sample data to obtain the final dictionary. The dictionary learning can be divided into different types according to the different update method of dictionary. In the experiment, the update method of dictionary is K-SVD. The main idea of K-SVD algorithm is shown below.

Algorithm: Dictionary Learning(K-SVD)

Input: original sample, dictionary, sparse matrix

1. Initialization: Randomly take \( K \) columns vector from the original samples \( Y \in \mathbb{R}^{m \times n} \) or take \( K \) columns vectors \( \{d_1, d_2, \cdots, d_K\} \) from its left singular matrix as the initial dictionary. So, we get the dictionary \( D^{(0)} \in \mathbb{R}^{m \times K} \). order \( j = 0 \), repeat steps 2-3 below until the specified number of iteration steps is reached or converge to the specified error.

2. Sparse coding: use the dictionary \( D^{(j)} \) obtained in the previous step, and sparse encoding to get \( X^{(j)} \in \mathbb{R}^{K \times n} \).

3. Dictionary update: gradually update the dictionary \( D^{(j)} \), columns of dictionary \( d_k \in \{d_1, d_2, \cdots, d_K\} \).
   i. when updating \( d_k \), calculate the error matrix \( E_k \)
   \[
   E_k = Y - \sum_{j \neq k} d_j X^j_f
   \]
   ii. Take out the index set of the \( k \)-th row vector \( X^k_f \) of the sparse matrix that is not 0
   \[
   \omega_k = \{i \mid 1 \leq i \leq n, X^k_f(i) \neq 0\}, X^k_f = \{X^k_f(i) \mid 1 \leq i \leq n, X^k_f(i) \neq 0\}
   \]
   iii. Take out columns corresponding to \( \omega_k \) which is not 0 from \( E_k \) and get \( E'_k \)
   iv. Perform singular value decomposition for \( E'_k \), and take the first column of \( U \) to update the \( k \)-th column of the dictionary. Order \( X^k_f = \sum (1,1)^\top \)
   v. \( j = j + 1 \)
2.2. The Model of GS-SVM

SVM was first proposed by Vapnik [14]. SVM is mainly composed of two-classification support vector machine and multi-classification SVM. Among them, the multi-classification SVM is constructed by many two-classification SVM through certain rules. In the two-classification SVM, the C-SVC model is the most common model. And the mathematical expression is as follows:

(1) Suppose the known training set:

\[ T = \{(x_i, y_i) \} \in (X \times Y) \]

where \( x_i \in X = R^n \), \( y_i \in Y = \{1, -1\} \); \( x_i \) is feature vector.

(2) Select suitable radial basis kernel function \( K(x_i, y_j) \) from various kernel functions. Then introduce Lagrange multipliers \( \alpha \) and KKT conditions to construct the optimization problem:

\[
\min_{\alpha} \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i} \alpha_i \\
\text{s.t.} \sum_{i} y_i \alpha_i = 0, \ 0 \leq \alpha_i, i = 1, 2, \cdots, l
\]

Get the optimal solution: \( \alpha^* = (\alpha_1^*, \cdots, \alpha_l^*) \)

(3) \( b^* \) can be obtained through \( \alpha^* \):

\[
b^* = y_i - \sum_{i} y_i \alpha_i^* K(x_i, x_j)
\]

(4) Get the decision function:

\[
f(x) = \text{sgn}\left( \sum_{i=1}^{l} \alpha_i^* y_i K(x_i, x_j) + b^* \right)
\]

When dealing with multi-classification problems of rolling bearing faults, it is necessary to construct appropriate multi-class classifiers. The construction methods of multi-class classifier mainly include one-to-one method, one-to-many method and directed acyclic graph method. In the experiment, the multi-classifier of SVM is constructed by one-to-one method.

The GS algorithm is an exhaustive search method for specifying parameter values to select the best hyper parameters for the model. The best learning algorithm is obtained by using the cross-validation method to optimize the parameters of the estimation function. That is, the possible values of each parameter are arranged and combined, and all possible combinations are listed to generate a "grid". Then, each combination is used for SVM training, and cross-validation is used to evaluate the performance. The hyper parameter combination with the highest average score is used as the best choice, and the model object is returned.

Because the performance of SVM depends on its parameter selection, this paper chooses an intelligent optimization algorithm to optimize the parameters of SVM. Among many intelligent optimization algorithms, ACO [15], GA [16] and PSO [17] have many problems in performance such as large computational overhead and slow convergence when optimizing parameters. Therefore, in this paper, the grid search algorithm with fast convergence speed and high accuracy is used to optimize the parameters of the support vector machine to improve the accuracy of bearing fault diagnosis.

3. The Method Proposed In This Paper(K-SVD-GS-SVM)

Based on the above analysis and theoretical basis, a fault diagnosis method based on dictionary learning and parameter optimization support vector machine was proposed and applied to bearing fault diagnosis. The flow chart is shown in Figure 1. The specific steps are as follows.

Step 1: data pre-processing. Take 1000 data points as a sample to intercept 10 types of fault data without overlapping, and this paper intercept 100 samples for each type of data. Then, we need to convert each type of data sample matrix into a grayscale image.
Step 2: set the dictionary learning parameters and denoise the grayscale images of each type of sample through the dictionary learning algorithm, then output the grayscale images as the signal sample data. 
Step 3: build the SVM model and set the parameters of the GS optimization algorithm to optimize the parameters of the SVM to classify the fault data. 
Step 4: get the accuracy of fault diagnosis.

**Figure 1.** Algorithm flow chart of this paper

4. Validation of Proposed K-SVD-GS-SVM Model

4.1. Data Description

The experimental data in this article comes from the Bearing Data Center of Case Western Reserve University in the United States. The 6205-2RSSKF deep groove ball bearing that supports the motor drive shaft end is used as the test bearing. The data is collected in four states: 0HP, 1HP, 2HP, and 3HP. The sampling frequency is 12 kHz. The bearing states identified in the experiment mainly include normal states, inner ring faults, outer ring faults, and ball faults, and each fault state includes three types of faults with different degrees of damage: diameters of 0.007inch, 0.014inch, and 0.021inch. Therefore, we have a total of 10 types of fault states. In the experiment, 100 samples of each type of fault signal are intercepted without overlap, and each sample includes 1000 data points. It is important to note that the sample data set used in experiment 4.1 in this paper is shown in Table 1. In the experiment of 4.2, we selected data in Table 1 under 0 load and adjusted the ratio of training samples to test samples.
Table 1. Description of rolling bearing datasets

| Fault location | None | Ball | Inner race | Out race | Load |
|----------------|------|------|------------|----------|------|
| Category Labels | 1    | 2    | 3          | 4        | 5    |
| Diameter(inch)  | 0    | 0.007| 0.014      | 0.021    | 0.007|
| Dataset A       | Train 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 0 |
|                 | Test 30  | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 1 |
| Dataset B       | Train 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 2 |
|                 | Test 30  | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 3 |
| Dataset C       | Train 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 0 |
|                 | Test 30  | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 1 |
| Dataset D       | Train 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 2 |
|                 | Test 30  | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 3 |

4.2. Performance under Different Load

In the experiment, dictionary learning is first used to denoise the grayscale images after data preprocessing, and then the SVM model optimized by GS is used for classification and recognition. Denoising through dictionary learning can eliminate noise in the industrial environment, thereby get better features for classification. Figure 2 show the grayscale images converted from the original sample under 0 load. The grayscale images after dictionary learning are shown in Figure 2.

It can be seen from Figure 2 and Figure 3 that dictionary learning has particularly obvious effect on sample images. When dictionary learning is not performed, the overall sample image is in a fuzzy state, and the identification degree between each type of data sample is not high. After performing dictionary learning to denoise the sample images, the clarity of the entire sample has been significantly improved. At the same time, the degree of identification between various types of sample data also becomes obvious. Therefore, the application of dictionary learning to denoise sample data is of great significance to the final classification and recognition.

Figure 2. Grayscale image of raw data samples

According to the analysis in Table 2, the accuracy of SVM in bearing fault diagnosis is greatly affected by different environments and the diagnostic accuracy is lower than other methods. After using dictionary learning to denoise the data, the accuracy of A and C samples has improved to varying degrees compared to SVM, and the accuracy of B and D samples has decreased, which...
indicates that the sample classification results after dictionary learning denoising are not stable. After using GS to optimize the SVM parameters, the accuracy of bearing fault diagnosis has been significantly improved compared to the SVM model. In the experiments of sample sets C and D, the accuracy of GS-SVM improved by 13.43% and 14.14% compared with SVM. This shows that using GS to optimize SVM is of great significance to improve the accuracy for bearing fault diagnosis. It can be seen from Table 2 that the accuracy of the method proposed in this paper can reach 100% when identifying and classifying various types of sample data collected in different environments.

Table 2. Experimental results of different methods for bearing fault diagnosis.

| Sample set | SVM  | K-SVD-SVM | GS-SVM | K-SVD-GS-SVM |
|------------|------|-----------|--------|---------------|
| A          | 66%  | 68.33%    | 76.28% | 100%          |
| B          | 79%  | 74%       | 90.28% | 100%          |
| C          | 60%  | 66.33%    | 73.43% | 100%          |
| D          | 75%  | 71.33%    | 89.14% | 100%          |

4.3. Effect of the Different Ratio between Training Samples to Test Samples

In the course of the experiment, some methods may cause disparity in the accuracy of bearing fault diagnosis due to different ratios between training samples and test samples. In other words, the performance of these algorithms is not stable enough in experiments. Therefore, this paper tests under multiple sample ratios to verify the stability of the method. The experimental results are shown in Table 3.

It can be seen from Table 3 that in the bearing fault diagnosis, the accuracy of the SVM model is greatly affected by the ratio of the training samples to the test samples, and the diagnostic accuracy is lower than other methods. And the diagnostic accuracy rate differs by 9.2%. After the dictionary learning denoising, the accuracy rate is relatively stable when using SVM to test under different sample proportions, and the accuracy difference of diagnosis is 5%. Similarly, when using the GS-optimized SVM model to test under different sample proportions, the diagnostic accuracy differs by 3.91%. In summary, regardless of whether the SVM, K-SVD-SVM or GS-SVM model is tested at different sample proportions, it will be affected to a certain extent. But on the issue, the method proposed in this paper can consistently achieve the accuracy of 100% in many experiments.

Table 3. Effect of the ratio between training samples and test samples.

| Sample ratio | SVM | K-SVD-SVM | GS-SVM | K-SVD-GS-SVM |
|--------------|-----|-----------|--------|---------------|
| 9:1          | 70% | 66%       | 79.11% | 100%          |
| 8:2          | 66% | 71%       | 78.37% | 100%          |
| 7:3          | 66% | 68.33%    | 76.28% | 100%          |
| 6:4          | 64.5% | 69.75% | 77.50% | 100%          |
| 5:5          | 60.80% | 67% | 75.2% | 100%          |

5. Conclusion

This paper proposed a bearing fault diagnosis method based on dictionary learning and GS-optimized SVM. In the experiment, the signal sample data is converted into grayscale images for denoising. Secondly, SVM classifier with high generalization ability for small sample data and avoiding local minimum values is selected. The parameters of SVM are optimized by the GS optimization algorithm to solve the problem that performance of SVM is greatly affected by parameters. Through experimental verification, the proposed method not only achieves accuracy of 100% on the sample sets collected in different environments, but also shows stable and good performance for different ratios between train samples and test samples in the experimental process. Therefore, the method proposed in this paper can be used to bearing fault diagnosis in complex realistic and experimental environments, which provides new ideas for bearing fault diagnosis. In addition, the method presented in this paper is also applicable to small-scale data bearing fault diagnosis experiments. In summary,
the bearing fault diagnosis method proposed in this paper not only can achieve high diagnostic accuracy, but can be applied to bearing fault detection in actual industrial production.

Acknowledgments
The support of Science foundation of North University of China(Grant No.2017027), Shanxi Scholarship Council of China (Grant No.2016-088), Shanxi Natural Science Foundation (Grant No. 201701D221121, 201701D121012 and 201801D121026) and National Nature Science Foundation of China (Grant No. 61774137) in carrying out this research is gratefully acknowledged.

References

[1] Wu S D, Wu P H, Wu C W, et al. Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and Support Vector Machine[J]. Entropy, 2012, 14(12):1343-1356.
[2] Zhou W, Habetler T G, Harley R G. Bearing Fault Detection Via Stator Current Noise Cancellation and Statistical Control[J]. IEEE Transactions on Industrial Electronics, 2009, 55(12):4260-4269.
[3] Nandi S, Toliyat H, Li X. Condition monitoring and fault diagnosis of electrical motors-a review[J]. IEEE Trans Energy Convers, 2005, 20(4):719–729.
[4] Chen X., Wang S, Qiao B, et al. Basic research on machinery fault diagnostics: Past, present, and future trends[J]. Frontiers of Mechanical Engineering, 2018, 13(2).
[5] Cui L, Huang J, Zhang F. Quantitative and Localization Diagnosis of a Defective Ball Bearing Based on Vertical-Horizontal Synchronization Signal Analysis[J]. IEEE Transactions on Industrial Electronics, 2017, PP(99):11.
[6] Lei Y, Qiao Z, Xu X, et al. An underdamped stochastic resonance method with stable-state matching for incipient fault diagnosis of rolling element bearings[J]. Mechanical Systems and Signal Processing, 2017, 94:148-164.
[7] Wang Y, Liang M. An adaptive SK technique and its application for fault detection of rolling element bearings[J]. Mechanical Systems and Signal Processing, 2011, 25(5):1750-1764.
[8] Chen Z, Deng S, Chen X, et al. Deep neural networks-based rolling bearing fault diagnosis. Microelectronics Reliability, 2017, 75:327-333.
[9] Sun J, Yan C, Wen J. Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning. IEEE Transactions on Instrumentation and Measurement, 2018, 67(1):185-195.
[10] Sohaib M, Kim C H, Kim J M. A Hybrid Feature Model and Deep-Learning-Based Bearing Fault Diagnosis[J]. Sensors, 2017, 17(12): 2876-2891.
[11] Eren L, Ince T, Kiranyaz S. A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier[J]. Journal of Signal Processing Systems, 2019, 91(2):179-189.
[12] Stéphane G, Mallat, Member. et al. Matching Pursuits With Time-Frequency Dictionaries[J]. IEEE Trans. on Signal Processing, 1993, 41(12):3397 - 3415.
[13] Aharon M. An algorithm for designing over complete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54.
[14] Vapnik V N. The Nature of Statistical Learning Theory[M]. Springer, 2000.
[15] Zhang X L, Chen W, Wang B J, et al. Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization[J]. Neurocomputing, 2015, 167:260-279.
[16] Chen F, Tang B, Chen R. A novel fault diagnosis model for gearbox based on wavelet support vector machine with immune genetic algorithm[J]. Measurement, 2013, 46(1):220-232.
[17] Kennedy J, Eberhart R. Particle swarm optimization[C]// Proceedings of ICNN'95 - International Conference on Neural Networks. IEEE, 1995.