Fault Diagnosis Method of Rolling Bearings Based on Supercomplete Dictionary Learning

line 1: Dou An
line 2: *Faculty of Electronic and Information Engineering*
line 3: *Xi'an Jiaotong University*
line 4: Xian 710049, China
line 5: douan2017@xjtu.edu.cn

line 1: Chunlin Hu
line 2: *Faculty of Electronic and Information Engineering*
line 3: *Xi'an Jiaotong University*
line 4: Xian 710049, China
line 5: hucl0918@stu.xjtu.edu.cn

2023 35th Chinese Control and Decision Conference (CCDC) | 979-8-3503-3472-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/CCDC58219.2023.10327208

***Abstract*—As one of the most important technologies to guarantee the safe operation of industrial equipment, fault diagnosis technology has gained wide attention. Rolling bearings are indispensable and wearing parts for large machinery equipment. It's greatly significant to find out the fault type, fault severity, and fault location in time for maintaining the normal operation of a mechanical system. Based on the advantages of the supercomplete dictionary learning model, this paper proposes a new feature extraction method, which is combined with the Softmax classifier for rolling bearing fault diagnosis. We use a measured rolling bearing data set to prove the effect of our method. Then we design contrast experiments to compare our method with traditional methods. The experiment results show that our method can accurately diagnose multifarious bearing faults, and the supercomplete dictionary model can extract the characteristics of vibration signals well, which is superior to traditional research efforts.**

Keywords—fault diagnosis; rolling bearing; supercomplete dictionary learning model; machine learning; Softmax classifier

I. Introduction

In mechanical equipment, the bearings, which are responsible for supporting the rotating shaft and the parts on the shaft, are very important in the normal operation of mechanical equipment. Although the rolling bearing has the advantages of convenient maintenance and reliable quality, it is still possible to break down in production which may cause a bad impact on production and life.

Recently, the mainstream fault diagnosis includes two processes: extracting feature and classifying vibration signals. For the feature extraction method, the commonly used methods include fast Fourier transform [1], wavelet transforms [2], and empirical mode decomposition [3]. For the mode classification method, the commonly used methods include an expert system, vector machine supporting [4], BP neural network, and deep learning. Now the updating of fault diagnosis methods mainly lies in the different combinations and replacement of feature extraction methods and classifiers to achieve better diagnosis results.

The traditional fault diagnosis methods mainly rely on human experience judgment and developing signal processing technology to extract useful feature representation from various vibratory signals. Although the methods mentioned above already have satisfactory performance, the feature extraction representation method is mainly designed by human. It mainly depend on the existing knowledge about signal processing and diagnostic specialty, so it cannot be well generalized to other diagnostic fields. A fault diagnosis method based on deep reinforcement learning has been proposed, which can learn the advanced features of vibration signals and has significant diagnostic performance. Dictionary learning is more prominent in fault diagnosis and representation learning because of its simple and efficient characteristics.

However, in real-world applications, the input samples captured under different operating conditions vary widely within the class, and the classification performance of these methods is often affected. Background interference and noise erosion also have a significant influence on amplitude of signals, which may cause a obvious intra-class variation. Under the same fault mode, the vibration signals with different loads show great intra-class variation, which will obviously aggravate the decline of classification performance. Therefore, a new fault diagnosis method is urgently needed to solve these problems. Based on the advantages of the supercomplete dictionary learning model, a new feature extraction method is proposed in this paper that is combined with the Softmax classifier for rolling bearing fault diagnosis. The main works in this paper is as follows:

- We designed the whole process of experimental fault diagnosis by using the pretreatment of data sets and supercomplete dictionary learning.
- We established the fault diagnosis model by combining the Softmax classifier with the supercomplete dictionary learning algorithm.
- We conducted the experiment based on MATLAB to prove the accuracy and effect of the new method.

II. Basic Principle Of Rolling Bearing Fault Diagnosis

The general process of rolling bearing fault diagnosis is shown in Fig 1.

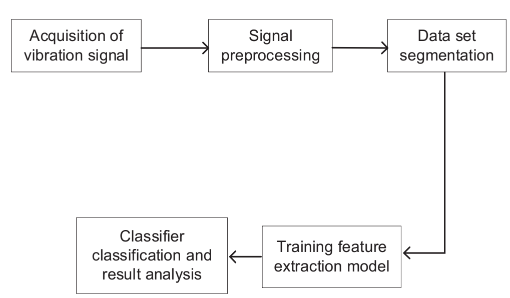

Fig 1 General process of rolling bearing fault diagnosis

The main process of bearing fault diagnosis is as follows: acquiring vibration signals, signal preprocessing, data set segmentation, training of the supercomplete dictionary learning model, computational feature characterization, training of Softmax classifier, and analysis of results.

III. Research On Rolling Bearing Fault Diagnosis

This paper assumes that the training set is $\left\{x^i\right\}_{i=1}^{M}$ and the corresponding set is formulated as $\left\{y^i\right\}_{i=1}^{M}$, where $x^i \in \Re^{N\times 1}$, $y^i \in \left\{1,2,\cdots,K\right\}$. For each input x^i, the model calculates the probability $p(y^i = k \mid x^i)$ for all labels of k, where $k = 1,2,\cdots,K$. Therefore, Softmax regression will output a vector and the vector will yield the estimated probability of K input samples belonging to each label. Specifically, we assume the form of $h_\theta(x^i)$ is

$$h_\theta(x^i) = \begin{bmatrix} p(y^i = 1 \mid x^i;\theta) \\ p(y^i = 2 \mid x^i;\theta) \\ \vdots \\ p(y^i = K \mid x^i;\theta) \end{bmatrix} = \frac{1}{\sum_{k=1}^{K} e^{\theta_K^T x^i}} \begin{bmatrix} e^{\theta_1^T x^i} \\ e^{\theta_2^T x^i} \\ \vdots \\ e^{\theta_K^T x^i} \end{bmatrix} \quad (1)$$

where θ denotes the parameter of the Softmax regression model which can be formulated as $\theta = \left[\theta_1,\theta_2,\cdots,\theta_K\right]^T$. $\sum_{k=1}^{K} e^{\theta_K^T x^i}$ denotes the distribution normalization which ensures that the sum of all elements is 1.

Based on the assumption of the model, this paper trains the model by minimizing the loss function $J(\theta)$.

$$J(\theta) = -\frac{1}{M}\left[\sum_{i=1}^{M}\sum_{k=1}^{K} 1\left\{y^i = k\right\} \log \frac{e^{\theta_K^T x^i}}{\sum_{l=1}^{N} e^{\theta_l^T x^i}}\right] + \frac{\lambda}{2}\sum_{k=1}^{K}\sum_{l=1}^{N}\theta_{kl}^2 \quad (2)$$

Where $\left\{y^i = k\right\}$ indicates that the function returns 1 if the condition is true and 0 otherwise. λ denotes the weight attenuation term.

The ultimate purpose of the weight attenuation term is to prevent overfitting and underfitting. It makes some parameters of the Softmax regression close to zero but allowing others to keep relatively large values. With the weight attenuation term, $J(\theta)$ is strictly convex and can ensure that the maximum regression model has a unique solution in theory for any $\lambda > 0$.

The linear weighted output values are converted into probability distributions through the Softmax operation. The principle of the Softmax classifier can be simply expressed as:

$$Soft\max(p(y^i = k \mid x^i) = \frac{\exp(p(y^i = k \mid x^i)}{\sum_{i=1}^{5}\exp(p(y^i = k \mid x^i)} \quad (3)$$

Where y^i indicates that the sample y corresponds to category I and k indicate the label.

The Softmax layer structure is shown in Fig 2. And the main diagnosis process is shown in Fig 3.

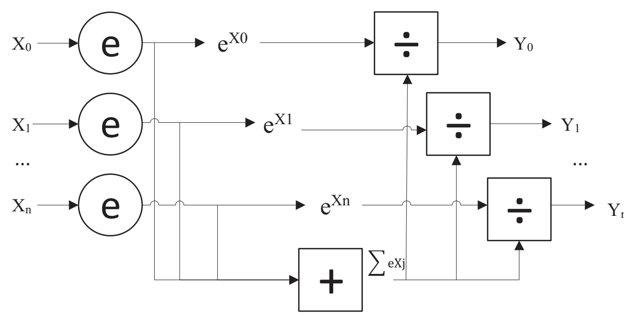

Fig 2 Softmax layer structure

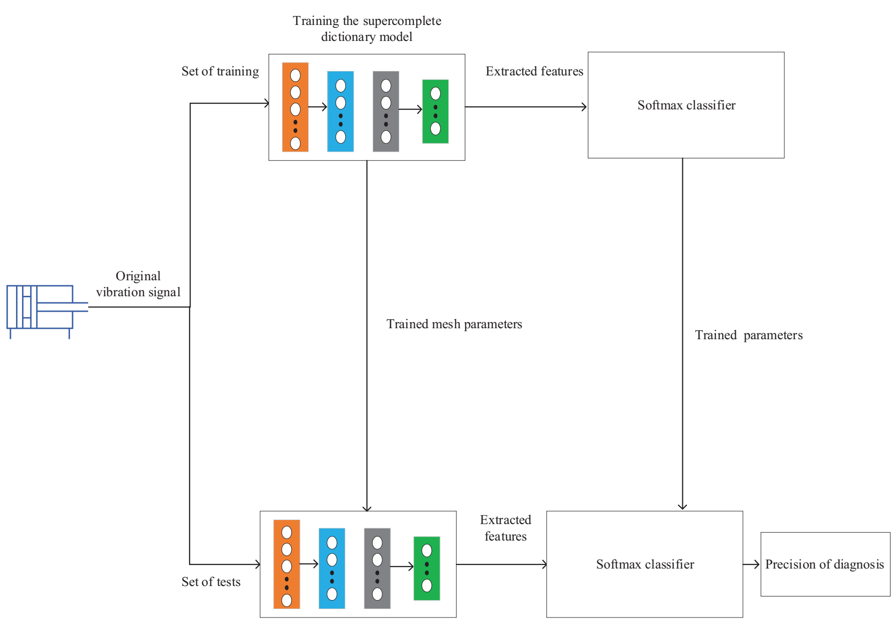

Fig 3 Main diagnosis process

Table 1 Gearbox bearing data set

health states	operating condition	number of samples	health label
normal state	O_1/O_2	300/300	1
ball fault	O_1/O_2	300/300	2
inner ring fault	O_1/O_2	300/300	3
outer ring fault	O_1/O_2	300/300	4
inner and outer ring compound fault	O_1/O_2	300/300	5

Table 2 initial experimental parameters table

initial experimental parameters	value
input size	100
output size	100
regularization parameter	0.3
Regular type	1
number of iterations	100
ratio of training	0.1
rate of overlap	0.95

IV. Experiment

A. Description of experimental data set

As mentioned above, in the era of mechanical big data, fault detection and diagnosis have made remarkable achievements in predicting environment conditions and remaining service life. The representation of the original sensory data is very important to the success of fault detection. Traditional methods are a lot of work because they often rely on handmade features and expertise heavily. To vertify the performance of the supercomplete dictionary model diagnosis method, a real case study, namely gearbox rolling bearing fault diagnosis, was conducted by constructing a data set. According to the vibration data selected in the paper under the minimum damage radius of 7 miles, we programmed it in MATLAB. Through preliminary research, this paper explores in detail the different health states that may appear in the working process of gearbox rolling bearings, and then builds a gearbox rolling bearing data set with five labels, as shown in Table 1.

The data set gear operating states are divided into five types, which are normal state, ball fault state, inner ring fault state, outer ring fault state, and inner and outer ring compound fault state. Under each working condition, there are 300 vibration signal samples for each health state and each sample has 2400 sampling points. The sampling frequency is to 1024Hz. O_1 and O_2 respectively represent different operating conditions: O_1 indicates the operating condition of the bearing is rotation frequency 20 Hz (1200 rpm), no load (0 Nm). And O_2 indicates the operating condition of the bearing is rotation frequency 30 Hz (1800 rpm), load 2 V (7.32 Nm).

The main parameters include input size and output size, regularization parameter, regular type, number of iterations, ratio of training, and rate of overlap between fragments. The initial parameter values are shown in Table 2.

B. Analysis of experimental results

In this experiment, 5 kinds of bearing fault combination states were selected in the data set, with a total of 3000 data points. In order to avoid the influence of randomness, this paper conducts 10 repeated experiments independently and obtains the average classification accuracy. The experimental results are shown in Figure 4. It can be seen that the accuracy rate reaches a high level under the default experimental conditions.

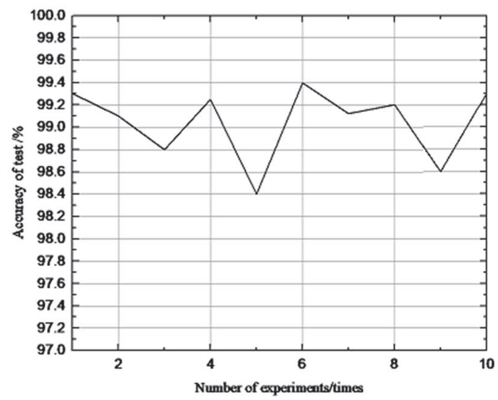

Fig 4 Diagnostic accuracy change line graph

Considering 8 different network structures, the diagnostic performance of the supercomplete dictionary model is evaluated. Except for different input and output sizes, the optimal trade-off parameter for each network structure were evaluated. Five repeated experiments were conducted independently for each kind of network framework, and the bar chart as shown in Fig 5 was drawn.

From Fig 5, the supercomplete dictionary model in this paper has achieved high test accuracy on the whole, all above 95%. When the input and output sizes were set to 50-50 and 60-75, although the iteration time was shortened, the diagnostic accuracy could not reach a very high level obviously. When the input and output sizes are 50-100 and 75-75, the diagnostic accuracy is above 98% in a limited time, but there is still much space to improve. When the input and output sizes were 75-100, 75-150, 100-100, and 100-150, the diagnostic accuracy reached a high level and remained within the highest diagnostic accuracy range of this method. Therefore, to eliminate the influence from different network architectures on the learning results of the supercomplete dictionary method, the 100-100 network structure type is adopted in this paper after considering the influence of iteration time and diagnostic accuracy comprehensively. In the remaining experiments, all experimental parameters share the same structure, so the number of elements for the input and output sizes is both 100 and 100.

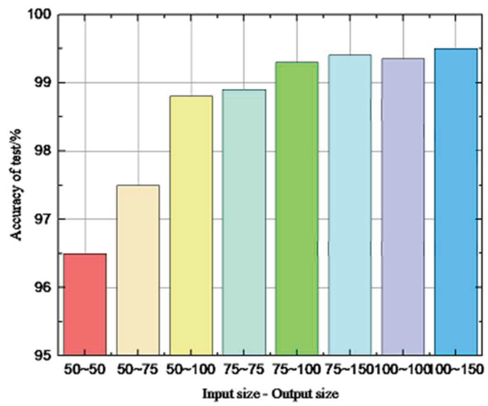

Fig 5 Influence of network framework on diagnostic accuracy

In this section, the influence degree of regularization parameters on diagnostic accuracy was studied, and then the optimal regularization parameter β was determined. As mentioned above, the network framework of this experiment adopts 100-100 type, and the regularization parameter β is [0.1:0.1:1], that is, 10 groups of data are taken every 0.1 intervals in this interval. In order to eliminate the experimental accidental error, each set of data was independently conducted three times, a total of 30 experiments. Finally, the influence curve of regularization parameter β on diagnostic accuracy was drawn, as shown in Fig 6.

From Fig 6, the regularization parameter β has a high diagnostic accuracy within the value range of [0.1, 0.8], reaching more than 96%. When β is 0.9 and 1.0, although iteration time is greatly reduced, the diagnostic accuracy is lower than in other experimental conditions. When β is 0.1 and 0.2, the iteration time is twice as long as other parameter conditions, and the diagnostic accuracy is not the highest. When β is in the range [0.3, 0.7], the iteration time is similar, and the conclusion can be drawn by observing the parameter curve. The diagnostic accuracy was high and the iteration time was reasonable and acceptable. Therefore, the regularization parameter β will be set to 0.3 in other experiments in this paper.

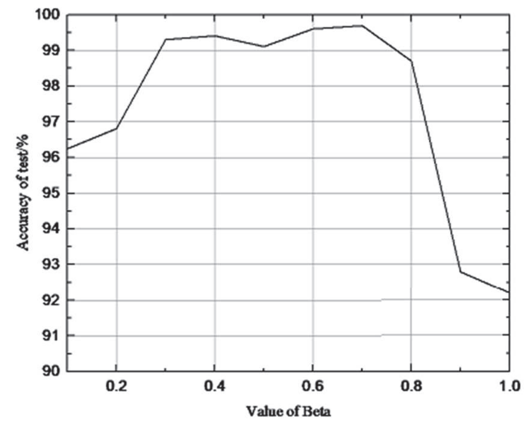

Fig 6 Influence of regularization parameter β on diagnostic accuracy

Then, to prove the validity and superiority of supercomplete dictionary learning model method, the accuracy of the method proposed in this paper is compared with that of several traditional methods on the selected data sets. The artificial extraction features are mainly based on the features of the time domain and frequency domain. The data set utilized in this experiment is the original vibration signals of rolling bearings with 5 categories of health statuses described in Table 1, and the comparison results are shown in Table 3. The results indicate that even when the proportion of training samples is 10%, the optimal recognition rate of the supercomplete dictionary model algorithm is 99.26±0.15%. Therefore, the method in this paper is better to all the usual manual diagnostic methods.

Table 3 Comparison with the method based on artificial feature extraction

Methods	Precision of diagnosis
Supercomplete Dictionary model + Softmax Classifier	99.26±0.15%
Supercomplete dictionary model + Extreme Learning machine classifier	96.53±0.87%
Artificial extraction feature + Softmax classifier	90.25±1.17%
Artificial extraction feature + Extreme Learning machine classifier	91.50±1.00%

As mentioned above, 300×5×2 data points in total were selected for the database sample in this paper. Where 2 represents O1 and O2 operating conditions, and 5 represents five different health states. 300 data points were selected for each working condition. As the training ratio was selected as 10%, the final result was a matrix of 2700×100. Among them, 2700 denotes the number of samples and 100 denotes the dimension. This paper uses the tsne dimension reduction technology of MATLAB and the drawing function to visualize the experimental results. The experiment results of comparison between the proposed method and the traditional methods are visualized in Fig 7 and Fig 8. From Fig 7, when the training ratio is 10%, the supercomplete dictionary model learning method adopted in this paper achieves high diagnostic accuracy and good aggregation. As can be seen from Fig 8, the traditional manual diagnosis method also contains 2700 data signals. When the feature similarity is high, it will be difficult to distinguish which category it belongs to, and the diagnostic accuracy is much lower.

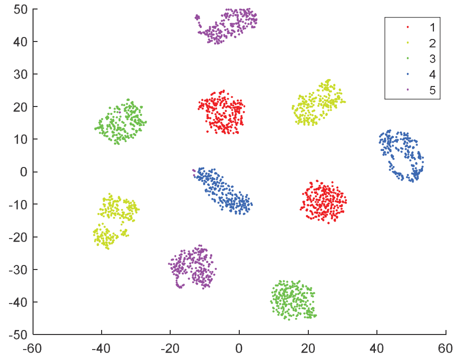

Fig 7 Visualization of feature recognition results of supercomplete dictionary learning method

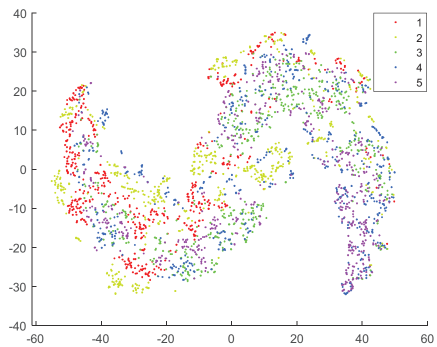

Fig 8 Visualization of feature recognition results of traditional manual diagnosis methods

V. CONCLUSION

This paper investigates and analyzes the historical background and development of machine learning in the field of fault diagnosis. Then the common fault categories of rolling bearings were explored and a gearbox rolling bearing data set containing five health states was constructed. Next, a new fault diagnosis method based on the supercomplete dictionary learning model is proposed by establishing and optimizing the loss function. Finally, the effect of the method is proved by a large number of experimental simulations. Compared with the traditional manual diagnosis method, the method in this paper has more advantages and prospects. The diagnostic time is greatly shortened and the diagnostic accuracy is greatly improved which proves that the method proposed in this paper has broad research prospects.

REFERENCES

[1] Rai V K, Mohanty A R. Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform[J]. Mechanical Systems and Signal Processing, 2007, 21(6): 2607-2615.

[2] Lou X, Loparo K A. Bearing fault diagnosis based on wavelet transform and fuzzy inference[J]. Mechanical systems and signal processing, 2004, 18(5): 1077-1095.

[3] Yu Y, Junsheng C. A roller bearing fault diagnosis method based on EMD energy entropy and ANN[J]. Journal of sound and vibration, 2006, 294(1): 269-277.

[4] Konar P, Chattopadhyay P. Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs)[J]. Applied Soft Computing, 2011, 11(6): 4203-4211.

[5] Ryllias K C, Antoniadis I A. A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments[J]. Engineering Applications of Artificial Intelligence, 2012, 25(2): 326-344.

[6] R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng. Self-taught learning in NIPS Workshop on learning when test and training inputs have different distributions[J]. Journal of sound and vibration, 2006, 76(8): 345-351.

[7] Y. Lei, F. Jia, J. Lin, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3137-3147.

[8] Z. Wei, J. Gao, X. Zhong, Z. Jiang, and B. Ma. Incipient fault diagnosis of rolling element bearing based on wavelet packet transform and energy operator[J]. WSEAS Trans. Syst, vol. 2011, 31(7): 81–90.

[9] Z. Gao, C. Cecati and S. X. Ding. A Survey of Fault Diagnosis and Fault-Tolerant Techniques —Part I: Fault Diagnosis with Model-Based and Signal-Based Approaches[J] IEEE Transactions on Industrial Electronics, 2015, 62(6): 3757-3767.

[10] Z. Gao, C. Cecati and S. X. Ding. A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part II: Fault Diagnosis With Knowledge-Based and Hybrid/Active Approaches[J]. IEEE Transactions on Industrial Electronics, 2015, 62(6): 3768-3774.

[11] L. Wen, X. Li, L. Gao and Y. Zhang. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998.

[12] T. Li, Z. Zhao, C. Sun, R. Yan and X. Chen. Multireceptive Field Graph Convolutional Networks for Machine Fault Diagnosis[J]. IEEE Transactions on Industrial Electronics, 2021, 68(12): 12739-12749.