[[1]](#footnote-1)

**~~Differential Geometric~~ Multi-modal Feature~~s~~ Fusion based Deep Learning Method for Frequency-type Fault Diagnosis**

Funa Zhou1, Po Hu1\*, Ju H. Park2\**,* Shuai Yang1 , Chenglin Wen3

1. School of Computer and Information Engineering, Henan University, Kaifeng, China

E-mail:[zhoufn2002@163.com](mailto:yansuisong8086@163.com)

1. Department of Electrical Engineering, Yeungnam University, 280 Daehak-Ro, Kyongsan  
   38541, Republic of Korea
2. School of Automatic, Hangzhou Dianzi University, Hangzhou, 310018, PR China.

**Abstract:** Rotating mechanical equipment usually suffer from frequency-type fault which cannot be well diagnosed in the time domain since frequency feature is not well extracted. On the other hand, diagnosis in frequency can only results in offline fault classification not online fault diagnosis. To solve this problem, multi-modal feature fusion based deep learning method for frequency-type fault diagnosis is proposed in this paper. Multi-modal feature corresponding to the original data, slope data and curvature data are firstly extracted by 3 separate deep neural networks, respectively. Then multi-modal feature fusion is developed to get a new combined feature which can characterize the potential frequency feature involved in time domain data. Lastly, use the fused new feature as the input of the softmax classifier to get online diagnosis result of frequency-type fault. Simulation experiment and case study of rolling bearing shows the efficiency of the method proposed in this paper.

**Key Words:** Fault diagnosis; Deep learning; Multi-modal feature; DNN; Feature fusion

1. **Introduction**

At present, mechanical equipment is more and more inclined to large-scale and complicated. Since various parts of the mechanical equipment are closely connected, failure of any part may cause breakdown of the entire mechanical equipment or even led to large industrial accidents, accurate and timely fault diagnosis is a challenging issue [1-8]. As a core part of rotating machinery, bearings are used in a wide range of industrial field. In the whole life of bearing, different degree of wear phenomenon is inevitable, timely monitoring of the bearing failure is significant in the sense of avoiding serious losses or even risk of the operator life. In this regard, research on bearing fault diagnosis has attracted widely attention from experts both in academic and application [2-6, 8-14]. Therefore, bearing fault diagnosis in the actual industrial production process is a promising research field.

In general, the methods of fault diagnosis can be categorized into 3 classes: physical model based method, knowledge based method, and data-driven based method [7]. Since the precise physical model is unavailable, physical model based methods are difficult to adapt to increasingly complex mechanical equipment fault diagnosis. Quantity of prior knowledge is difficult to process which would invalidate the knowledge based fault diagnosis methods. The data-driven techniques are widely applied in industry for process monitoring and fault diagnosis [9–14]. Principal component analysis (PCA), support vector machine (SVM), and artificial neural network (ANN) are the popular data-driven techniques for fault diagnosis [15–22]. Data-driven methods include statistical feature extraction and Neural network method [8-11]. The method based on statistical feature extraction can only detect faults but not diagnose faults. In recent years, due to the rise of Neural network method, Neural network based methods are widely used in the fault diagnosis of complex systems.

Neural network based methods are the most advanced data-driven method for fault diagnosis. Wang et al. used WPT to extract non-stationary characteristics of the bearing’s vibration signal as the pre-extracted feature of ANN [16]. The method uses the nonlinear learning classification ability and self-organizing ability of ANN to classify and diagnose bearing faults. However, due to the non-linearity, instability, high latitude and the large amount of noise pollution, the failure characteristics of mechanical equipment make it impossible to accurately diagnose mechanical equipment [12-14]. Some scholars put forward the use of signal processing feature extraction methods combined with machine learning methods for mechanical equipment failure diagnosis. Widodo and Yang extract frequency-domain features as SVM data sources to detect mechanical problems [13-17]. When the number of samples is limited and the signal is non-stationary, Yu et al. propose a method of rolling bearing fault diagnosis using a combination of SVM and EMD methods [10]. Hu et al. extracted the energy of each node of Wavelet Packet Transform (WPT) from the vibration signal as the characteristic parameter of bearing fault diagnosis, which greatly preserved the time-frequency characteristic of the characteristic information. Fault diagnosis based on combination of WPT and SVM, can improve the accuracy of fault diagnosis [18-21]. Based on the non-stationary characteristics of vibration signals of rolling bearings, Wang et al. used WPT to de-noise the collected signals and extract the energy characteristics of wavelet bands of each frequency band as the input characteristics of Artificial Neural Network (ANN) Learning classification ability and self-organization ability to bearing fault classification and diagnosis [22-26]. Yang and Tang etc. Proposed a method that combined the expert system with the Back Propagation neural network (BPNN), which fully utilized the advantages of the expert system and ANN and successfully detected the bearing failure [27]; Jiang et al. Propose a method using a combination of high-order cumulants with BPNN. This method uses high-order statistics as the eigenvector to improve the accuracy of BPNN in bearing fault diagnosis [28, 29]. However, in these studies, SVM and BPNN have many shortcomings as a shallow learning method: the essence of SVM is a dichotomizer, which has low learning efficiency in multi-classification and large sample problems. In addition, expert experience is required to choose suitable kernel function and scale parameter. SVM method cannot be real-time monitoring and diagnosis, which make it unavailable to meet the performance requirement of real-time diagnosis. ANN also has some congenital defects: (1) As a shallow neural network, the convergence of ANN is easy to fall into local optimum, thus it cannot well characterize the signal characteristic; (2) Low efficiency of ANN learning and inaccuracy of feature extraction is another shortcoming of ANN. In summary, as a shallow learning method, SVM and ANN have not been able to effectively extract features involved in high-dimensional and non-steady data [30]. Deep learning is a promising feature extraction tool and has aroused widespread concern of scholars [21, 28-30]. Compared with shallow learning, deep learning can perform well feature extraction and in-depth study on nonlinear big data [31, 32].

As one of the most popular and hot machine learning methods in the world today, deep learning has revolutionized artificial intelligence. Deep neural network (DNN) adopts unsupervised layer-by-layer greedy training algorithm and global parameter tuning based on BP algorithm, which can not only avoid local optimal problems, but also solve the restriction of the number of labels and samples. Hinton and Salakhutdinov first proposed the theory of deep learning in 2006. DNN can obtain a more abstract high-level representation of the input data by combining lower-level features which is a nonlinear transformations of its previous layer. So that DNN can discover the distribution of data without artificial feature selection, and can learn complex function [33]. Due to its excellent feature extraction capabilities, deep learning (DL) quickly attracted the attention of fault diagnosis researchers [34-37]. Lu et al. used the good feature extraction ability of deep neural network to successfully diagnose the bearing fault with the innovation to diagnose the unknown type fault timely and effectively [34]; Jia et al. used deep neural network to detect rolling shaft bearing health status [35].But the above mentioned DNN fault diagnosis methods didn’t take the data’s multi-modal feature into consideration.

Gan et al. proposed a fault classification method based on hierarchical neural network [36]. By constructing a two-layer neural network, the method not only could locate the position of bearing fault, but also effectively mine the fault size of the bearing in the same position. Li et al. addresses a multi-modal deep support vector classification (MDSVC) approach in [37] to propose a separation-fusion based deep learning method for gearboxes fault diagnosis [37]. The main innovation of this method is to fuse the multi-modal feature extracted in time domain, frequency domain and wavelet domain respectively. But it is not a real-time fault diagnosis method since Fourier transform is used to get the frequency data for the reason that Fourier transform is a full-time domain to frequency domain transformation. At present, existed studies about bearing fault diagnosis have not completely solved the problem of frequency-type faults with high diagnosis accuracy and real-time performance. The method considering the multi-modality of features, but it cannot guarantee real-time performance.

There are many zero-crossing values in the bearing data since they are mostly vibration signal. So bearing is common to suffer frequency fault whose fault size may be very larger in frequency value but very small in amplitude value. This character of frequency makes DNN available to distinguish frequency fault type based simply on the amplitude value of the observation data for the reason that multi-modal feature involves in vibration signal. On the other hand, differential geometry feature of the vibration signal, such as slope and curvature, can reveal the frequency fault feature in time domain. So research on real-time and accurate fault diagnosis method for bearing fault diagnosis is required to develop an efficient multi-modal feature fusion mechanism in time domain.

In order to resolve the limitation stated above, this paper developed an online differential geometric feature fusion based DNN (DGFFDNN) fault diagnosis method to improve the accuracy of frequency-type fault diagnosis. State of the art of this research is to design a fusion mechanism by combining multi-modal differential geometric feature. By this means, frequency failure can be well detected online which is significant to real-time frequency fault diagnosis. The main innovation is to solve the problem of inaccurate time-domain detection when frequency faults have similar amplitudes and frequent zero crossings by fusing the geometric features of the raw data. The original, slope and curvature data are respectively constructed into an Automatic Encoder (AE) model to fully excavate the implied frequency information of the data in time domain. Then, the obtained multi-modal features are fused and the feature dimension is increased to make the classification model training more accurate. First calculate the slope and curvature of the raw data that can represent the frequency characteristics of the value. The second step is to construct three DNN networks for the extraction of different multi-modal. The third step is to fuse the obtained features. The final step is to use the well trained network to diagnose the new samples in real time.

The remainder of this paper is organized as follows: Section 2 is the review of deep learning. In Section 3, a new DGFFDNN fault online diagnosis method is originally developed; In Section 4, the validity of the proposed fault diagnosis method is obtained through experiments and simulation analysis; Sections 5 is the conclusion and future work of this paper. (文章剩余部分组织结构)

1. **Review of deep learning theory**

Deep learning is an unsupervised learning method. By multi-layer nonlinear transformation, low-level features are combined to form more abstract high-level representation, making the learning system cannot find the distributed representation of the data without relying on artificial feature selection. Deep learning adopts unsupervised learning to pre-train DNN layer by layer, which helps DNN effectively excavate the fault features in mechanical signals. Then, DNN is fine-tuned by supervised learning to optimize the expression of DNN's fault features and make them have monitoring and diagnostic capabilities. In this paper, DNNs are pre-trained by stacking Auto Encoders (AE).

Auto-Encoder is a three-layer unsupervised neural network that is the input layer, multiple hidden layer and output layer. Among them, the input layer and the hidden layer constitute the coding network, while the hidden layer and the output layer constitute the decoding network. As shown in Fig.1, the input and output of the AE are the same. AE converts input data in a high-dimensional space into coded vectors in a low-dimensional space through a coding network, and coded vectors in a low-dimensional space are reconstructed into raw input data by a decoding network. Since the input signal can be reconstructed at the output layer, the encoded vector becomes a representation of the input data.



**Fig. 1.** The model of Auto-encoder

Given an unlabeled datasetconsisting ofobservation features, each observation variable hassamples. The encoder network encodes the sampleto the hidden activate valuewith an activation function. The encoder process is described as follows:

 (1)

where is the encoder function, Sigmoid functionis usually taken as the activation function in the encoder process. *W* is the weight matrix of the network between input layer and the hidden layer, *b* is the bias vector generated by the encoder network, is the connection parameter between the input layer and the hidden layer. The Sigmoid function can be depicted via Eq.(2)

 (2)

Similarly, for the decoder network, the feature matrixobtained from encoder network is used to reconstructthrough the decoder network such that the reconstructedis equal to the input. The decoder process is described as follows:

 (3)

where is the decoder function, is the activation function of the decoder process, represents the weight matrix between the hidden layer and the output layer of the network , *d* is the bias vector generated by the decoder process.

The essence of training AE is to train and optimize the network parametersand. In order to make the output as close as possible to the input, we need to optimize the network training parameters. Describe the proximity between input and output by minimizing the reconstruction error.The reconstruction error can be shown in Equation 4.

 (4)

Where  denotes predictive value, denotes actual value.

In training process, gradient descent method is used for AE parameter optimization, the concrete update form of network parameters can be formulated as formula (5) - (6).

 (5)

 (6)

Where  is the learning rate,  and  can be calculated by BP algorithm.

DNN can be simply seen as a multi-hidden layer neural network in which multiple AE layers are stacked. The bottom-up unsupervised learning method is used to extract features layer-by-layer and fine-tune the entire network with supervised learning methods. So that the DNN can extract the most essential feature potentially involved in the raw input data. The structure of DNN can be shown in Fig. 2.



**Fig. 2.** The structure of DNN

First, the DNN network is pre-trained by the unsupervised layer-by-layer greedy training algorithm. Given an untagged input data set as the input to the coding network to practice the first automatic encoder, the coded vector is obtained as the feature extracted on the first layer. Training parameter is obtained by setting x as the output of.is then used as the input to the second automatic encoder  and trained on the network parameter  of .  is the hidden layer of , which can be seen as a feature representation of . This process is repeated to obtain the feature of the  automatic encoder  and the corresponding network training parameter.

Second, add a classifier to the top of the DNN network. The DNN pre-training process is accomplished through unsupervised training layer by layer, and feature representation on each layer are extracted. However, up to now, DNN does not have the classification function, in order to achieve the output classification function, an efficient classifier is required on the top layer. In this paper, we use the Softmax classifier as the output layer of DNN. Use feature of the last hidden layer and the labeled data to train the Softmax classifier. Gradient descent is used to update the network parameter of deep learning. The fine-tuning process uses the labeled data to improve the performance of DNN.

1. **Differential geometric feature fusion based DNN fault diagnosis method**

There are many frequency-type faults in the actual system. For the processing of such faults, the main method is fourier transformed to obtain the frequency domain information and then processed. Although it is easily detected in the frequency domain, it is difficult to detect in the time domain. To ensure real-time fault detection, need to be able to effectively detect such faults in the time domain. This section first analyzes frequency-type faults and then describes in detail of the methods proposed in this paper.

* 1. **Frequency-type fault analysis**

Since the bearing data is a periodic vibration signal, there will be a large amount of zero-crossing data. These zero-crossing data are equal in amplitude, which makes it difficult to discriminate the health of the bearing based only on the amplitude data. Fig. 3.illustrates fault 1 and fault 2 as an example, where blue is data for fault 1, and red is data for fault 2. In a and b points are the zero-crossing points of fault 2 and fault 1, and their amplitudes are 0. At this time, these two points cannot be distinguished based on the amplitude information. But if considering their slope data, we can clearly distinguish these two fault types. So the method based on differential geometric feature fusion proposed in this paper is essentially to increase the frequency domain information of the data in order to maximize the utilization of useful information. Therefore, this paper first calculates the slope and curvature values of the raw data as characteristics of the frequency domain information.



**Fig.3.** Slope feature describe

It can be seen from Fig. 4.The difference between the fault data 1 and the fault data 2 in the time domain is small, so that the diagnosis effect based on the time domain amplitude is greatly reduced, and the difference between the two is more obvious in the frequency domain. This makes the diagnostic effect in the frequency domain greatly improved. However, this diagnostic effect is of no practical significance, because in the health monitoring of actual industrial systems, people hope that the real-time performance of the system is high, which can minimize the security risks. The method based on differential geometric feature fusion proposed in this paper is very practical. This method can obtain higher bearing fault diagnosis accuracy without using Fourier transform and time domain to frequency domain transform. This also provides an innovative way of real-time monitoring of equipment health.



**Fig.4.** Comparison between time domain and frequency domain

* 1. **DGFFDNN-based fault online detection method**

The DGFFDNN fault online diagnosis method is presented in this section. This section is divided into three parts to introduce the algorithm step by step: DNN-based multi-modal feature extraction, multi-modal feature fusion, fault online diagnosis. The complete fault diagnosis algorithm is as follows:

* + 1. **Multi-modal feature extraction**

The first step of DGFFDNN proposed in this paper is to extract data feature. This article uses a stacking Auto Encoders to extract data features. The feature extraction algorithm is as follows:

***Step* 1:** obtain data that characterize the differential geometric features of raw data. Summarized in algorithm 1.

|  |
| --- |
| **Algorithm 1** Get data that characterizes the differential geometric features of the raw data. |
| **Require:**  : Raw vibration signal data with N fault data sets  : Data preprocessing function, with parameters |
| 1:**function**  2: for n=1:N  calculate the slope of  calculate the curvature of  5:  6:  7: **return** |

***Step* 2:** Training DNN network model.

Respectively reorganize  into a matrix of n rows and m columns:.Construct observation, each observation has  samples. This can make more efficient use of deep learning feature mining capabilities. Respectively construct three DNN network with N hidden layers, and initialize DNN training parameters.

 (7)

Among them,**is the weight matrix of raw data, similarity,  is the weight matrix of the slope data, only the original data will be used as an example for description later, and will not be described again. ** is bias vector is the number of DNN hidden layer neurons. Network configuration saved in.  represent raw data set. The number of input neurons of DNN is expressed by using (8).

 (8)

The parameters of DNN are initialized by (9) - (10)

 (9)

 (10)

Applying unsupervised layer-by-layer feature extraction to datasets by training DNN for parameters

 (11)

Layer-by-layer feature extraction as shown in Fig. 2. get the feature of the top .

Similarly build DNN network model of slope and curvature data. Repeat the above AE forward encoding process to extract the characteristic values of the original data slope data and curvature data as follows:

 (12)

Whereis the features of raw data;is the features of slope data;is the features of curvature data;

Calculate the reconstruction error of the original data, slope data, and curvature data separately according to Equation 13.

 (13)

Then gradient descent method is used for AE parameter optimization, the concrete update form of network parameters can be formulated as formula (5) - (6). When the reconstruction error reaches a minimum, it means that the AE-trained parameters can be a good representation of the characteristics of the data., and  respectively represents the feature values of AE extracted from raw data, slope data and curvature data. Then add a classifier to the top of the DNN network.

* + 1. **Feature fusion**

In this paper, the model of DNN which is established by the original, slope and curvature data respectively can fully extract the fault components of the frequency class contained in the fault signal. In order to make the final diagnosis accuracy higher, we must make full use of the trained network to extract the frequency and time domain features of the data. We establish three DNN model to extract the fault feature respectively. The feature fusion process network structure shown in Fig.5.



**Fig. 5.** Feature fusion network structure

Feature fusion is to recombine the high-level features extracted by these three automatic encoders. The combination of features is as Fig. 5.



**Fig.5.** New feature vector

The features extracted from the above model are recombined to obtain a new feature vector f, as shown in Equation 14.

 (14)

Then use the trained Softmax classifier to classify the recombined features. The softmax classifier will give a detailed formula description in the next online diagnostic section.

* + 1. **Online diagnosis**

On-line diagnosis is the use of off-line learned model parameters to identify faults in real-time data collection. In order to accurately use the DNN model to extract the essential characteristics of the health status of mechanical equipment from the input samples, the following steps are required: First, the data of the collected vibration signals should be preprocessed. Secondly, the preprocessed data is used as the DNN model input unsupervised layer pre-training to extract the characteristics of mechanical equipment health status. Finally, based on the finite number of samples of the sample, the entire network is fine-tuned using the BP algorithm to update the entire network parameters. In this way we can carry out an effective diagnosis of the health status of mechanical equipment. Mechanical equipment fault diagnosis is usually divided into two processes: Offline learning and Online diagnosis. The purpose of offline learning is to train a DNN model using the collected historical data. First, the data is preprocessed on historical data, and then the forward training process of DNN is performed. Finally, the DNN network parameters are adjusted in reverse using labels, and the well trained network parameters are saved. On-line diagnosis is to use the well trained parameters and online data to perform corresponding calculations, and then determine whether the data is fault data. If there is a fault, an early warning is issued. If not, the data collection and diagnosis are continued. Usually, DNN for mechanical equipment fault diagnosis steps shown in Fig. 6.



**Fig. 6.** DNN based fault diagnosis

Detailed steps for online diagnosis are as follows:

***Step* 1:** Add a classifier to the top of the three DNN network.

The DNN pre-training process is completed through unsupervised training layer by layer, and the layers of feature information are extracted. Here need to add a DNN classifier on the top. In this paper, we use the Softmax classifier as the output layer of DNN. We use new feature  and the labeled data to train the Softmax classifier. The probability of each type can be calculated by the following hypothetical function:

 (14)

 (15)

Where, is the new feature vector,  is the model parameter of Softmax. The model parameters can also be optimized by the same means to minimize the cost function. The cost function of Softmax classifier can be defined as follows:

 (16)

Where is indicate function.

A softmax classifier is added at the top of DNN, and a limited set of labels  are used for the reverse adjustment of DNN.

 (17)

 (18)

, can be calculated by (14) - (15), is the number of samples. indicates the output of , is the learning rate of the reverse trimming process.

***Step* 2:** Fine-tune the parameters.

For the accuracy of feature extraction and output layer classification, the entire DNN training parameters are supervised finely with a finite number of sample tags and a BP algorithm, and the fine tuning process is completed by minimizing the reconstruction error. The update of training parameters through (19) - (20).

 (19)

 (20)

where  represents the actual output value,  is a parameter set generated from the whole network training, , BP algorithm is used to update the network parameter, and  is the learning rate in the process of deep learning. The fine-tuning process uses the labeled data to improve the performance of DNN. Fault Diagnosis Flowchart Based on GFFDNN is shown in Fig. 7.



**Fig.7.** Fault Diagnosis Flowchart Based on GFFDNN

1. **Experiment and analysis**

Rolling bearings play a crucial role in rotating machinery, bearing health will directly affect the reliability and stability of the entire system. In this paper, the rolling bearing is used as an experimental platform to verify the effectiveness of DGFFDNN. The proposed method is compared with the neural network method without feature fusion. This paper validates the algorithm using simulation data and real rolling bearing data.

**4.1 Simulation Study**

This article mainly solves the problem that frequency class faults are difficult to detect effectively in the time domain. When such fault signals appear, it is difficult for traditional learning methods to effectively identify different types of faults. This paper validates the effectiveness of the proposed algorithm by simulating multiple sets of different types of test data. Among them, there are data with different amplitudes of the same frequency, data with different amplitudes of the same amplitude, and data with different amplitudes and different frequencies

**4.1.1 Description of simulation experimental data**

Simulation data generation method is shown as Table 1.and the data Waveform is shown in Fig. 8, 9, and 10, respectively.

**Tab.1.** Simulation data

|  |  |  |  |
| --- | --- | --- | --- |
| Different experimental scenes | Sampling interval | Signal 1 | Signal 2 |
| different amplitudes with same frequency | 0.1 |  |  |
| different frequency with same amplitude | 0.1 |  |  |
| different amplitudes with different frequencies | 0.1 |  |  |

**4.1.2 Analysis of simulation experiment result**

In order to verify the effectiveness of the algorithm, different types of simulation data are also verified in this paper. The situation when the signal frequency has same frequency with different amplitudes is shown in Fig.8. On the left is the waveform of the signal. On the right is the diagnostic effect of the proposed algorithm. Red and blue are different fault signals in the waveform diagram. In the diagnostic effect chart, red represents the predicted category, and blue represents the real category.



**Fig.8.** Same frequency with different amplitudes

The situation when the signal same amplitudes with different frequency is shown in Fig.9. It can be seen that frequency-type faults are difficult to detect in the time domain. Compared to other types of fault results, detailed data can be found in Tab 2.



**Fig.9.** Same amplitudes with different frequency

The situation when the signal different amplitudes with different frequency is shown in Fig.10. From the above simulation results, it can be seen that when the amplitude and frequency of the signal are not the same, the diagnostic effect is the best. In addition to the simulation data, this paper also uses the data collected in the test platform built by ourselves to verify.



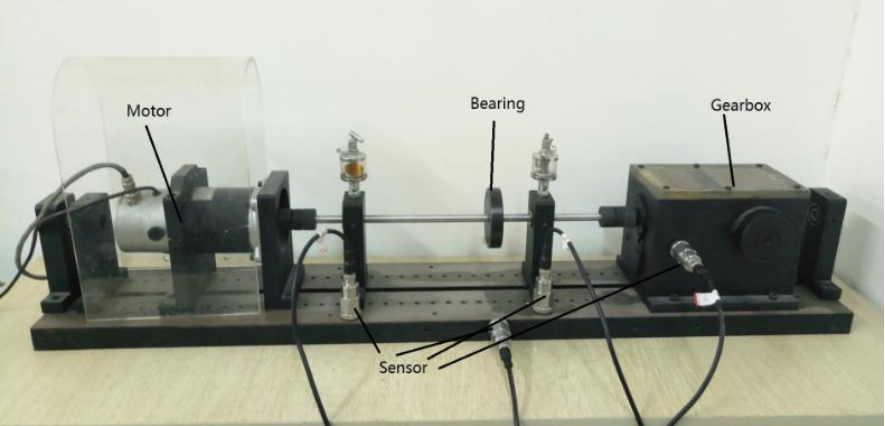
**Fig.10.** Different amplitudes with different frequency

**4.2 Case study**

In order to further verify the validity of the algorithm, this paper uses our own platform and the public test dataset of Western Reserve University. In our own test platform, we have carried out research on different fault diameter with same fault type, and different fault types with same fault diameter.

**4.2.1 Description of experimental platform**

The experimental data set is the bearing data that collected by Data-driven research team of Henan University built a bearing fault diagnosis test platform. In the bearing fault diagnosis, the vibration signal is the easiest to collect. And the vibration signal is more sensitive to bearing failure. Therefore, vibration signals are widely used in bearing fault diagnosis. Acceleration sensor is used in the experiment to collect the vibration signal of motor drive end as experimental data of bearing fault diagnosis. In this experiment, the acceleration sensor is used to collect the motor-driven vibration signal of 0hp, the sampling frequency is 48kHz. There are four types of fault diameter :(1)0.007, (2)0.014, (3)0.021 (4) normal condition. The experimental platform shown in Fig.11.



**Fig. 11.** Experiment platform of rolling bearing

**4.2.2 Case study result analysis**

Our proposed deep learning based on differential geometric features is applied to bearing fault diagnosis. There are 4,500 samples under each data type and 4 different data types to characterize the frequency fault types of rotating mechanical systems. In order to reduce the influence of randomness, Experiment repeated 10 times. DNN training uses a stochastic gradient descent method, and the maximum number of iterations of DNN in each layer is 1000, 800, 1000 times respectively. In this paper, DNN's pre-training initialization parameters are shown in Table 2.

**Tab.2.** DNN model parameters

|  |  |  |  |
| --- | --- | --- | --- |
| Training parameter |  |  |  |
| Hidden layers | 6 | 4 | 5 |
| Number of neurons | 500/400/200/100/50/10 | 500/100/50/20/10 | 500/200/100/50/20/10 |
| Max number of epochs | 1000 | 800 | 1000 |
| Learning rate | 0.01 | 0.02 | 0.01 |

This article selects three sets of fault data and a set of normal data. The three sets of fault data are internal fault types, and the load is 0. The difference is that the fault diameter is 0.007, 0.014, and 0.021. The four data types were chosen because these data types have similar amplitude and frequent zero crossings. It makes difficult to effectively distinguish fault types based on time domain signals. The time domain signal is shown in Fig.12.



**Fig. 12.** Same fault type with different fault diameter

The diagnostic results of same fault type with different fault diameter are shown in Fig. 13. It can be seen that there is a good diagnostic accuracy for bearing data.



**Fig.13.** Diagnosis accuracy of same fault type with different fault diameter



**Fig. 14.** Same fault diameter with different fault type



**Fig.15.** Diagnosis accuracy of same fault diameter with different fault type

In order to further verify the effectiveness of the algorithm, we have also carried out research the same fault diameter with different fault type on our own test platform. The time domain signal is shown in Fig.14. The diagnostic results of same fault diameter with different fault type are shown in Fig. 15. It can be seen that there also a good diagnostic accuracy for bearing data.

**4.2.3 Benchmark data set testing**

In addition, we also use the data from the Case Western Reserve University Bearing Center to verify this algorithm. The diagnostic results are shown in Fig. 16. It can be seen that the proposed algorithm is effective.



**Fig.16.** Detection accuracy of Western Reserve University Bearing Data Center data

Compared with the results without feature fusion, the proposed method is validated. It can be seen from Table 3 that the diagnostic accuracy of the proposed method in frequency class and other faults is higher than that of ordinary neural networks and traditional deep neural networks.

**Tab.3.** Comparison of the accuracy of fault diagnosis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data | GFFDNN | DNN | GFFBP | BP |
| Case Western Reserve University Bearing Center | 98.06 | 89.52 | 87.73 | 73.56 |
| Same fault type with different fault diameter | 98.54 | 90.14 | 88.16 | 80.13 |
| Same fault diameter with different fault type | 97.63 | 89.53 | 86.42 | 70.84 |
| Same frequency with different amplitudes | 94.34 | 93.01 | 92.69 | 91.04 |
| Same amplitudes with different frequency | 95.06 | 90.23 | 86.27 | 84.31 |
| Different amplitudes with different frequency | 98.40 | 95.24 | 94.36 | 92.86 |

1. **Conclusion and future work**

In this paper, aiming at the problem that frequency-type ~~class~~ fault cannot be accurately ~~effectively~~ diagnosed in the time domain, a new DGFFDNN ~~fault online diagnosis~~ method is developed to get an online fault diagnosis algorithm with high diagnosis accuracy. ~~in detail.~~ The main innovation of this method is to characterize the potential frequency feature involved in time domain data by exploring the dynamic trend feature characterizing ability of differential geometry feature. Frequency-type fault whose fault size is larger in frequency domain while very small in time domain can be well diagnosed by the proposed method. The main idea is to extract multi-modal feature corresponding to the original data, slope data and curvature data are firstly extracted by 3 separate deep neural networks, respectively; Then the ~~obtained data~~ extracted multi-modal features are merged to augment ~~increase~~ the feature dimension, so that the fault classifier ~~diagnosis model~~ can be trained more accurately. In addition, this method enables diagnosis ~~detection~~ of frequency-type faults in the time domain, which will facilitate the online ~~real-time detection~~ ~~diagnosis of the~~ health diagnosis of the mechanical equipment. The case study of rolling bearing test platform validates the effectiveness of the method.

~~Large-scale mechanical equipment is getting more and more complicated, and the requirements for fault diagnosis are getting higher and higher.~~ How to effectively implement health monitoring of devices through self-learning algorithms without tagged data is a very important research direction in the future.

**Reference**

[1] D. H. Zhou, Y. Liu and X. He, "Review on fault classification techniques for closed-loop systems," Acta Automatica Sinica, vol.39, no. 11, pp. 1933-1943,2013.

[2] F. N. Zhou, J. H. Park and Y. J. Liu, "Differential feature based hierarchical PCA fault detection method for dynamic fault," Neurocomputing ,vol.202, pp. 27-35, 2016.

[3] C. Li, RV. Sanchez, G. Zurita, M. Cerrada, D. Cabrera and R. E. Vásquez , "Multimodal deep support vector classification with homologous features and its application to gearbox fault classification," Neurocomputing, vol.168, pp.119-127, 2015.

[4] X. He, Z. Wang, Y. Liu and D. H. Zhou, "Least-Squares Fault Detection and Classification for Networked Sensing Systems Using A Direct State Estimation Approach," in IEEE Transactions on Industrial Informatics, vol. 9, no. 3, pp. 1670-1679, Aug. 2013.

[5] D. Zhao, D. Shen and Y. Q. Wang, "Fault classification and compensation for two-dimensional discrete time systems with sensor faults and time-varying delays, " Int. J. Robust Nonlinear Control, doi: [10.1002/rnc.3742](http://dx.doi.org/10.1002/rnc.3742), 2017.

[6] H. Li and D.Y. Xiao, "Surver on data driven fault classification methods", Control and Decision, vol. 26, no. 1, pp. 1-9+16, 2011.

[7] R. M. An and Y. Gao, "Spacecraft fault classification based on hierarchical neural network," Spacecraft environment engineering, vol. 30, no. 2, pp. 203-208, 2013.

[8] F. N. Zhou , C. L. Wen, Y. B. Leng and Z. G. Chen, "A data-driven fault propagation analysis method", Journal of Chemical Industry and Engineering(China), vol. 61, no. 8, pp. 1993-2001,2010.

[9] H. Q. Ji, X. He and D. H. Zhou, "On the use of reconstruction-based contribution for fault classification, "Journal of Process Control, vol. 40, pp. 24-34, 2016.

[10] D. J. Yu, M. F. Chen, J. S. Cheng and Y. Yang, "A fault classification approach for rotor systems based on empirical mode decomposition method and support vector machines, " Proceedings of the Chinese society for electrical engineering, vol. 26,no. 16, pp. 162-167, 2006.

[11] M. Gan, C. Wang and C. A. Zhu, "Construction of hierarchical classification network based on deep learning and its application in the fault pattern recognition of rolling element bearings" Mechanical Systems and Signal Processing, vol.72-73, pp. 92–104, 2016.

[12] G. F. Bin, J. J. Gao, [X. J. Li](http://www.sciencedirect.com/science/article/pii/S0888327011003207) and [B. S. Dhillon](http://www.sciencedirect.com/science/article/pii/S0888327011003207), "Early fault classification of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network," Mechanical Systems and Signal Processing, vol. 27, pp. 696–711,2012.

[13] A. Widodo and BS. Yang, "Support vector machine in machine condition monitoring and fault classification," Mechanical Systems and Signal Processing, vol. 21,no. 6, pp.2560-2574, 2007.

[14] D. Zhao, Z. P. Lin and Y. Q. Wang, "Integrated state/disturbance observers for two-dimensional linear systems," in IET Control Theory & Applications, vol. 9, no. 9, pp. 1373-1383, 2015.

[15] Q. Hu, Z. J. He, Z. S. Zhang and Y. Y. Zi , "Fault classification of rotating machinery based on improved wavelet package transform and SVMs ensemble ," Mechanical Systems and Signal Processing ,vol. 21, no.2, pp. 688–705, 2007.

[16] L. Y. Wang, W. G. Zhao, Y. Liu, "Rolling Bearing Fault Classification Based on Wavelet Packet- Neural Network Characteristic Entropy," Advanced Materials Research, Vols. 108-111, pp. 1075-1079, 2010.

[17] Y. Yang and W. Tang, "Study of remote bearing fault classification based on BP Neural Network combination," 2011 Seventh International Conference on Natural Computation, pp. 618-621,Shanghai, 2011.

[18] L. Jiang, Q. Li, J. Cui and J. Xi, "Rolling bearing fault classification based on higher-order cumulants and BP neural network," The 27th Chinese Control and Decision Conference (2015 CCDC), pp. 2664-2667, Qingdao, 2015.

[19] T. Kuremoto, S. Kimura, K. Kobayashi and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," Neurocomputing, vol. 137, pp. 47–56, 2014.

[20] S. M. Zhang, F. L. Wang, S. Tan and S. Wang, "A fully automatic onine mode identiflcation method for multi-mode processes," Acta Automatica Sinica, vol. 42, no.1, pp.60-80, 2016.

[21] Jürgen Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85–117, 2015.

[22] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," Science, vol.313, pp. 504-507, 2006.

[23] H. Ze, A. Senior and M. Schuster, "Statistical parametric speech synthesis using deep neural networks," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 7962-7966, Vancouver, BC, 2013.

[24] B. Song, S. Tan and H. B. Shi, "Key principal components with recursive local outlier factor for multimode chemical process monitoring," Journal of Process Control, vol.47,pp.136–149, 2016.

[25] L. P. Zhao, C. H. Zhao and F. R. Gao, "Inter-batch-evolution-traced process monitoring based on inter-batch mode partition for multiphase batch processes," Chemometrics and Intelligent Laboratory Systems, vol.138, pp.178–192, 2014.

[26] Y. W. Zhang, C. Wang and R. Q. Lu, "Modeling and monitoring of multimode process based on subspace separation," Chemical Engineering Research and Design, vol.91, no.5, pp.831–842, 2013.

[27] F. N. Zhou, C. L. Wen, T. H. Tang and Z. G. Chen, "DCA based multiple faults classification method," Acta Automatica Sinica, vol.35, no.7, pp. 971-982, 2009.

[28] P. Tamilselvan and P. F.Wang, "Failure classification using deep belief learning based health state classification," Reliability Engineering and System Safety, vol. 115, pp. 124–135, 2013.

[29] R. B. Huang, C. Liu, G. Q. Li, and J. L. Zhou, "Adaptive Deep Supervised Autoencoder Based Image Reconstruction for Face Recognition," Mathematical Problems in Engineering, vol. 2016, Article ID 6795352, 14 pages, 2016.

[30] H. M. Liu, L. F. Li, and J. Ma, "Rolling Bearing Fault Classification Based on STFT-Deep Learning and Sound Signals," Shock and Vibration, vol. 2016, Article ID 6127479, 12 pages, 2016.

[31] P. L. Wang , C. J. Xia, "Fault detection and self-learning identification based on PCA-PDBNs," Chinese Journal of Scientific Instrument, vol.36, no. 5, pp.1147-1154, 2015.

[32] R. Pang, Z. B. Yu,  W. Y. Xiong and H. Li , "Faults recognition of high -speed train bogie based on deep learning ," Journal of Railway Science and Engineering, vol.12, no.6, pp. 1283-1288, 2015.

[33] C. Lu，Z.Y. Wang, W. L. Qin and J. Ma, "Fault classification of rotary machinery components using a stacked denoising autoencoder-based health state identification," Signal Processing, vol.130, pp.377-388, 2017.

[34] F. Jia, Y. G. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent classification of rotating machinery with massive data," Mechanical Systems and Signal Processing, vol.72-73, pp. 303–315, 2016.

[35] Bearing data Centre, Case Western Reserve University, Available:

<http://csegroups.case.edu/bearingdatacenter/home>

1. This research was supported in part by the Natural Science Fund of China (Grant No. U1604158).

   \*Corresponding author E-mail: [hupo4210@163.com](mailto:hupo4210@163.com), [jessie@ynu.ac.kr](mailto:jessie@ynu.ac.kr) [↑](#footnote-ref-1)