

# Attention mechanism based multi-scale feature extraction of bearing fault diagnosis

LEI Xue<sup>1</sup>, LU Ningyun<sup>1,2,\*</sup>, CHEN Chuang<sup>1,3</sup>, HU Tianzhen<sup>1</sup>, and JIANG Bin<sup>1,2</sup>

1. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; 2. State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; 3. College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211816, China

**Abstract:** Effective bearing fault diagnosis is vital for the safe and reliable operation of rotating machinery. In practical applications, bearings often work at various rotational speeds as well as load conditions. Yet, the bearing fault diagnosis under multiple conditions is a new subject, which needs to be further explored. Therefore, a multi-scale deep belief network (DBN) method integrated with attention mechanism is proposed for the purpose of extracting the multi-scale core features from vibration signals, containing four primary steps: preprocessing of multi-scale data, feature extraction, feature fusion, and fault classification. The key novelties include multi-scale feature extraction using multi-scale DBN algorithm, and feature fusion using attention mechanism. The benchmark dataset from University of Ottawa is applied to validate the effectiveness as well as advantages of this method. Furthermore, the aforementioned method is compared with four classical fault diagnosis methods reported in the literature, and the comparison results show that our proposed method has higher diagnostic accuracy and better robustness.

**Keywords:** bearing fault diagnosis, multiple conditions, attention mechanism, multi-scale data, deep belief network (DBN).

**DOI:** [10.23919/JSEE.2023.000129](https://doi.org/10.23919/JSEE.2023.000129)

## 1. Introduction

Bearing is one of the most critical components in rotating machines. Nearly 40% to 50% of all rotating machinery failures are in connection with bearing faults [1]. Accurate diagnosis and prognosis of rotating machinery is vital for maintaining the safe and reliable operation of rotating machinery. The fault diagnosis of bearings under constant operating conditions has been extensively studied [2]. However, in practice, bearings often run under

multiple conditions along with different rotational speeds and external loads. Bearings running under multiple conditions are prone to various failures. Therefore, there is an urgent need for reliable multi-working condition diagnosis methods, however, there is little in-depth research on bearing fault diagnosis under multi-working conditions [2].

With the massive data collected in many systems, data-driven fault diagnosis methods are faced with the challenges of big data processing and complex working conditions. Different from the shallow neural network methods, deep learning methods have remarkable ability to process a vast amount of data, which are attracting more and more attention.

In recent years, great progress has been made in fault diagnosis using deep learning methods [3–11]. For instance, some interesting results based on vibration signal processing using deep learning methods have been reported in [12–21]. Deep belief network (DBN) [22] as a widely used deep learning technique [23–25], can unify feature extraction, feature reduction as well as pattern recognition. In addition, it is able to extract representative information and sensitive features automatically. In the context of massive data, DBN is a proper algorithm for bearing fault diagnosis.

As discussed previously, bearings often run under multiple conditions, with their signals being unstationary. If the deep learning algorithm is used for blind training of vibration signals collected from different conditions, the obtained features may have serious divergences even for the same fault type [2]. Such divergence leads to the fact that those bearing fault diagnosis methods developed for constant running conditions are unsuitable for multi-condition cases. Thus, it is the primary requirement for successful data-driven bearing diagnosis methods under various conditions to obtain comprehensive intrinsic characteristics. Many researchers have been committed to fault diagnosis using multi-scale method, and have achieved some positive results [26–28]. Obviously, vibration sig-

Manuscript received August 16, 2021.

\*Corresponding author.

This work was supported by the National Natural Science Foundation of China (62020106003; 61873122; 62303217), Aero Engine Corporation of China Industry-university-research Cooperation Project (HFZL2020CXY011), and the Research Fund of State Key Laboratory of Mechanics and Control of Mechanical Structures (Nanjing University of Aeronautics and Astronautics) (MCMS-I-0121G03).

nals at different scales can provide diversified information. Using a multi-scale resampling technique, different features of vibration signals at different scales can be obtained. Nevertheless, the features at different scales often have different contributions to bearing fault diagnosis. It is important to eliminate redundant features to enhance the performance of fault diagnosis. The attention mechanism will be adopted for such a purpose.

The attention mechanism [27–29] has been successfully applied to different research tasks. The main advantage lies in that it can highlight the pivotal features. Inspired by this idea, numerous researchers have begun to apply the attention mechanism to fault diagnosis to obtain favorable performance. In [29], an attention mechanism was introduced into fault diagnosis in an efficient way.

In this regard, in this paper, a multi-scale DBN model based on attention mechanism is proposed, which can be used for fault diagnosis of bearings under multiple conditions. Firstly, the vibration signals are processed to obtain different scales. Then, multi-scale features are extracted by DBN. Next, the combined scales with the highest diagnostic reliability are selected. On the basis of the principle of attention mechanism, the contributions of different scale features are calculated to assign the scale weights. Based on the weighted features, multi-scale features are fused for fault diagnosis.

The rest of this paper is described below. The basic theory of the proposed method is introduced in Section 2.

Section 3 describes the multi-scale DBN model integrated with attention mechanism. In Section 4, a benchmark dataset for bearing fault diagnosis under multiple conditions is used to validate the effectiveness of the proposed method. Moreover, some comparative results with other fault diagnosis methods are provided. Finally, Section 5 summarizes this paper.

## 2. Basic theory

### 2.1 Multi-scale data processing

Bearings often work at different rotational speeds and load conditions. In this paper, the bearing fault under multiple conditions is mainly considered, in which the vibration signals of bearing failure will be more complicated. The multi-scale vibration signals can be obtained by down-sampling frequency.

As is known to all, the micro-scale data can reflect the detailed characteristics, while the large-scale data can describe the global characteristics. In general, the macro-scale characteristics are able to provide more comprehensive information for fault diagnosis.

Set  $X = \{x_1, x_2, \dots, x_n, \dots, x_N\}$  as the input of multi-scale data input layer, where  $N$  denotes the initial length of the vibration signals data,  $x_n$  is the  $n$ th value of the sampling point of the original vibration signal. Thus, the corresponding data length is  $\frac{N}{2^{s-1}}$ . As illustrated in Fig. 1, the data form with different scale numbers ( $s = 1, 2, 3$ ).

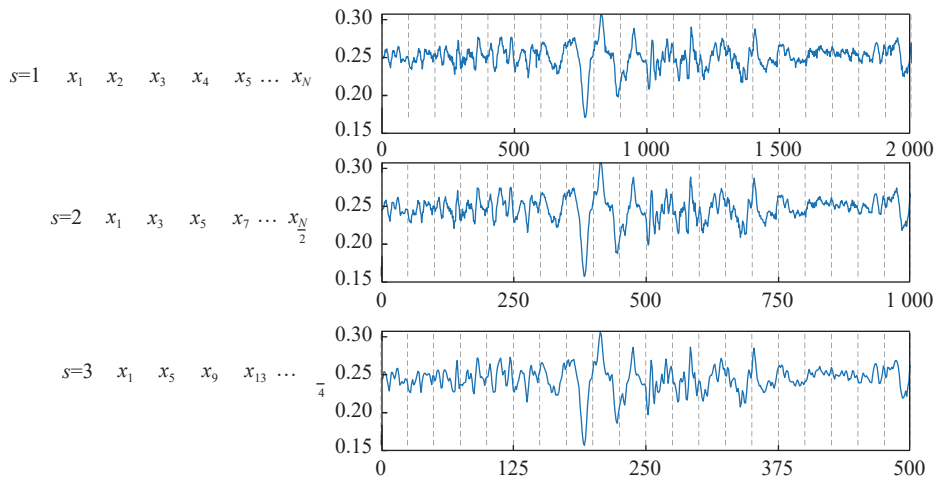


Fig. 1 Structure of multi-scale DBN method based on attention mechanism

Due to the limitations described in Section 1, the final scale number is not arbitrarily given. The larger the number of the scales, the longer the training time. To ensure the timeliness of fault diagnosis, it is necessary to pre-determine the number of scales, which will be discussed in detail in Section 4.

### 2.2 DBN

DBN generally consists of multiple restricted Boltzmann

machine (RBM) [22]. Each RBM is composed of a visual layer and a hidden layer. Fig. 2 shows the structure of DBN. The feature extraction for fault diagnosis is realized through the iteration of RBM and fine-tuning using the backpropagation (BP) neural network. In order to solve the training problems of DBN, connection weights of networks are optimized through layer by layer greedy learning.

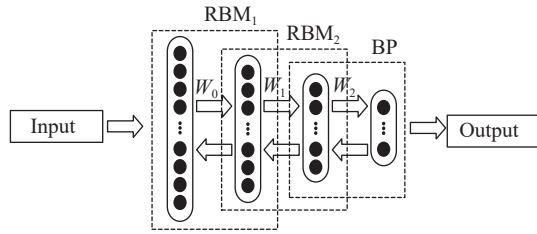


Fig. 2 DBN structure

2.3 Attention mechanism

To enhance the fault diagnosis performance, the attention mechanism is applied to DBN, concentrating on the characteristics that have a great impact on fault diagnosis. Fig. 3 shows the structure of attention mechanism. On the basis of attention mechanism, this paper adopts the strategy of “feature recalibration” [30]. According to the contribution of each channel to the fault diagnosis, the attention mechanism assigns different weights to the final outputs of the hidden layer of each channel.

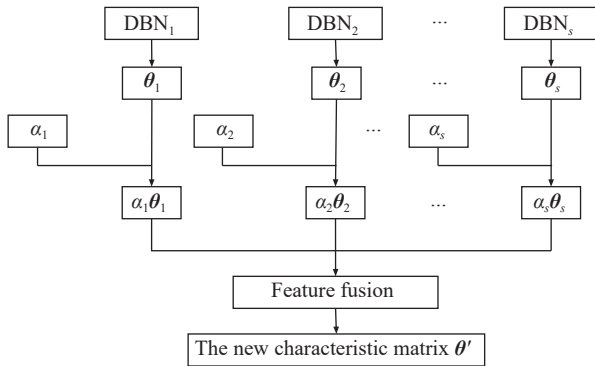


Fig. 3 Feature fusion process of attention mechanism

Different scale features extracted from each channel exert different influence on the fault diagnosis results. The key is how to assign weights to different scale features.

Set  $\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_s$  to be the output matrix of hid-

den layer for the  $i$ th scale ( $i = 1, 2, \dots, s$ ), and there are  $k$  hidden layer output nodes. Then,  $\theta_i$  can be given by

$$\begin{cases} \theta_i = [\theta_i(1), \theta_i(2), \dots, \theta_i(k)] \\ \vdots \\ \theta_i = [\theta_i(1), \theta_i(2), \dots, \theta_i(k)] \end{cases} \quad (1)$$

On the other hand, the average accuracy rate of fault diagnosis in each scale can be regarded as confidence level (noted as  $e_1, e_2, \dots, e_i$  ( $i = 1, 2, \dots, s$ )). Thus, the weight of each channel can be determined by

$$\begin{cases} \alpha_i = \frac{e_i}{e_1 + e_2 + \dots + e_s}, \quad i = 1, 2, \dots, s \\ \alpha_1 + \alpha_2 + \dots + \alpha_s = 1 \end{cases} \quad (2)$$

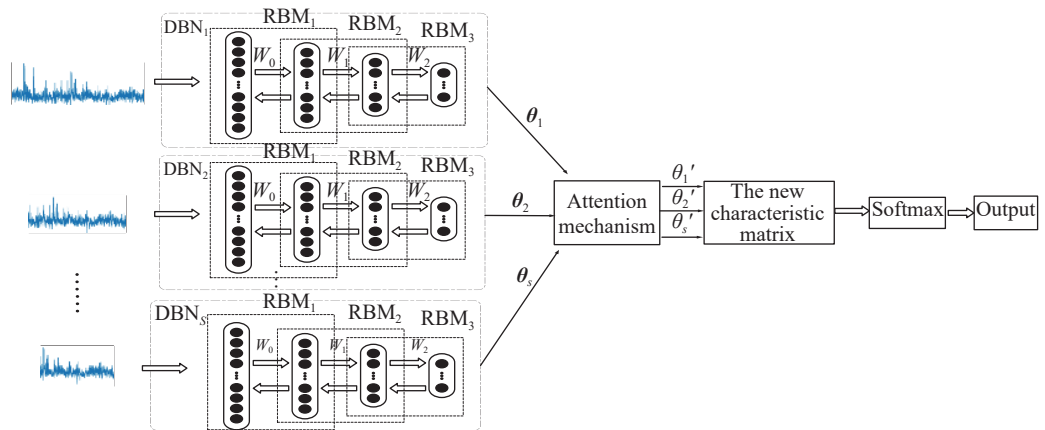
Ultimately, the new characteristic matrix  $\theta'$  can be updated by

$$\begin{cases} \theta' = [\alpha_1\theta_1(1) + \alpha_2\theta_2(1) + \dots + \alpha_s\theta_s(1), \\ \alpha_1\theta_1(2) + \alpha_2\theta_2(2) + \dots + \alpha_s\theta_s(2), \dots, \\ \alpha_1\theta_1(k) + \alpha_2\theta_2(k) + \dots + \alpha_s\theta_s(k)] \\ \theta' = \sum_{i=1}^s \alpha_i\theta_i \end{cases} \quad (3)$$

3. Multi-scale DBN model integrated with attention mechanism

The diagnosis flow of the proposed multi-scale DBN method integrated with attention mechanism is shown in Fig. 4, which consists of four modules:

- (i) The multi-scale data input layer, which is used for receiving the vibration signals in different scales;
- (ii) The multi-scale feature extraction layer, which aims to extract the features of each scale;
- (iii) The attention mechanism layer, which is applied to obtain multi-scale features using the attention mechanism;
- (iv) The diagnosis results output layer, which utilizes the fused multi-scale features to achieve fault classification.



— The multi-scale data input layer — DBN extraction feature — Attention mechanism layer — Classification layer —

Fig. 4 Diagnosis flow of multi-scale DBN method based on attention mechanism

The multi-scale vibration data are input to the DBN, and the process of training the data of each scale is called a channel. Different from single-scale DBN training, after multi-scale DBN iteration, the hidden layer output characteristics of each channel are included in the attention mechanism. The details of attention mechanism are discussed in Subsection 2.3.

As shown in Fig. 3, every scale has a channel, and DBN is used for feature extraction. The weights are assigned according to aforementioned principles, and a new characteristic matrix  $\theta'$  is obtained.

With the obtained fused multi-scale features, a “softmax” activation function, which can generate the conditional probability of each type of output, is employed to realize the multi-class fault classification.

Specifically, as shown in Fig. 5, the fault diagnosis process based on DBN with attention mechanism is described as follows. Firstly, the multi-scale vibration signals are obtained using the multi-scale method. Secondly, the multi-scale features are obtained through multiple RBMs with different channels. Thirdly, the attention mechanism is used to recalibrate the features of different channels. Subsequently, the fault diagnosis is realized by using the “softmax” activation function in the classification layer. It should be pointed out that the proposed multi-scale DBN model with attention mechanism can handle multiple time-varying conditions. Since the multi-scale method can weaken random noise to some certain extent, the influence of noise on fault diagnosis is also considered.

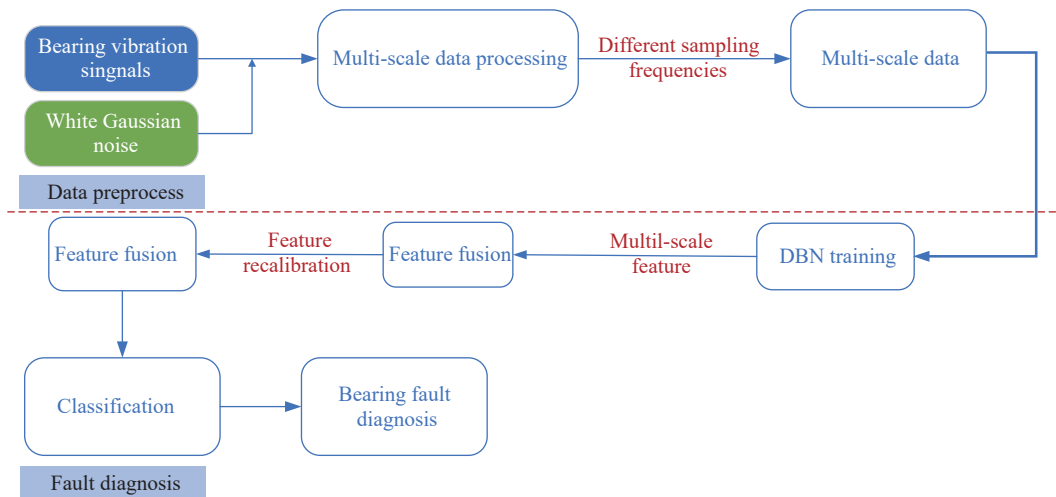


Fig. 5 Fault diagnosis process based on DBN with attention mechanism

## 4. Experimental validation

### 4.1 Data introduction

In this section, the bearing vibration data under multiple conditions provided by University of Ottawa is applied to verify the performance effectiveness of the proposed method.

Fig. 6 demonstrates the experimental platforms. Table 1 introduces the dataset, and four variable speed conditions are used for fault diagnosis, including the increasing speed (marked as IS), decreasing speed (marked as DS), increasing and then decreasing speed conditions (marked as IDS), as well as decreasing and then increasing speed (marked as DIS). In addition, in order to highlight the superiority of this method, the results of diagnosis are compared with four existing fault diagnosis methods. In practice, considering that the vibra-

tion signals are polluted by noise, the Gaussian noise is added to the bearing vibration data so as to verify the robustness of this method.

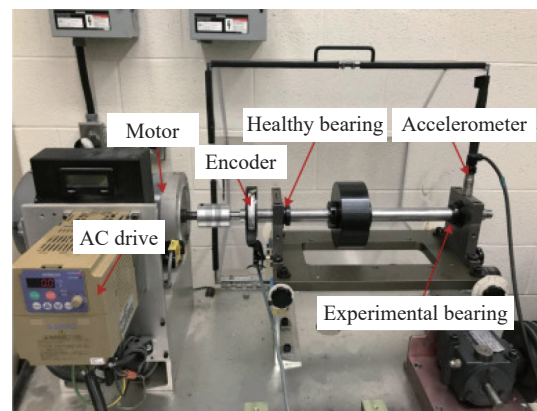


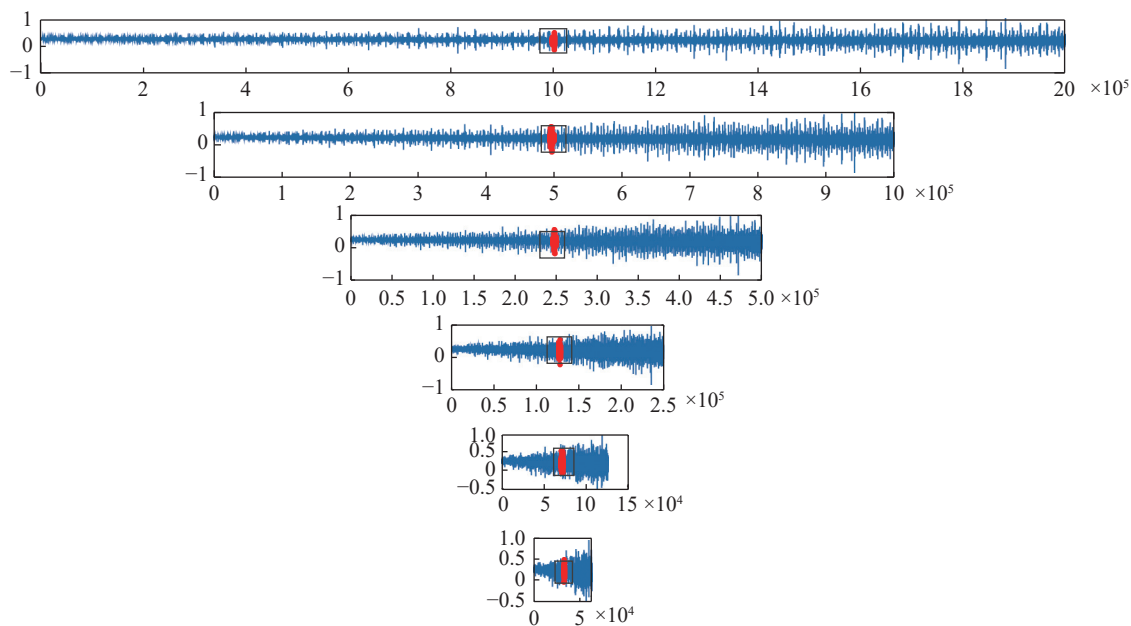
Fig. 6 Experimental set-up [31]

**Table 1 Introduction of bearing dataset**

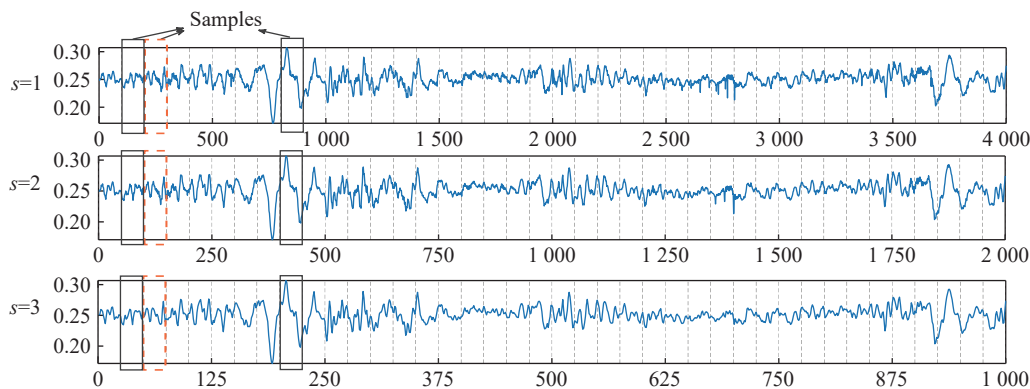
Speed varying conditions	Bearing health condition	Label	Training dataset	Testing dataset
IS	Healthy	1	10 500	4 500
	Inner race fault	2	10 500	4 500
	Outer race fault	3	10 500	4 500
DS	Healthy	4	10 500	4 500
	Inner race fault	5	10 500	4 500
	Outer race fault	6	10 500	4 500
	IDS	Healthy	7	10 500
Inner race fault		8	10 500	4 500
Outer race fault		9	10 500	4 500
DIS	Healthy	10	10 500	4 500
	Inner race fault	11	10 500	4 500
	Outer race fault	12	10 500	4 500

### 4.2 Multi-scale data processing

As mentioned earlier, multi-scale data can provide more comprehensive information for fault diagnosis. The multi-scale data are obtained by multi-scale processing of vibration signals. Fig. 7 illustrates the overall situation of vibration signals after processing six scales. According to Fig. 7, the vibration data obtained synchronously have different data lengths due to the reduction of the sampling frequency. Moreover, the vibration signals obtained are different with different sampling frequencies. In this figure, the area around 1 000 000 sampling points of the raw data is marked in red. As the scale increases, it can be seen that the position of the point is changing. The bearing signals from one to three scales are used as the signals, while Fig. 8 merely shows the multi-scale process of the first 4 000 samples.



**Fig. 7 Bearing vibration signals at six scales**



**Fig. 8 Multi-scale process to the first 4000 samples**



### 4.3 Scale selection and fault diagnosis results

Information contained in different scales has different contributions to fault diagnosis. There are some invalid scales affecting the timeliness of diagnosis. Therefore, it is necessary to select a reasonable scale for fault diagnosis. In order to select the appropriate scale, fault diagnosis should be carried out on a single-scale, and the scale with good diagnosis should be preferred. Fault diagnosis is conducted on the vibration data of different scales. The accuracy, stability and time of fault diagnosis are taken as the measurement standard. To eliminate random factors, 20 trials are collected for each scale. The results of fault diagnosis with scales from 1 to 5 under single time-varying condition (with increasing speed) is shown in Table 2. It can be seen that the accuracy and stability of fault diagnosis have been improved at scale 3.

**Table 2** Fault diagnosis results at different scales

Number of scales	Average accuracy/%	Average standard deviation	Average diagnostic time/s
1	96.26	2.041	311.594
2	96.48	0.749	538.051
3	98.73	0.153	767.237
4	98.48	0.166	1006.977
5	98.55	0.166	1226.475

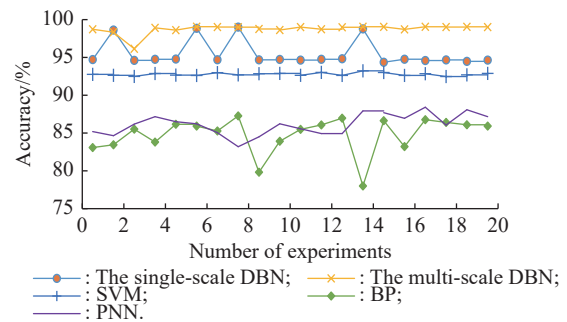
Compared with the first two scales, the accuracy and stability of diagnosis have not changed significantly with the increase of scales in later stage, but the time required for diagnosis is increasing. Thus, the number of scales is set to 3.

### 4.4 Comparative experimental results

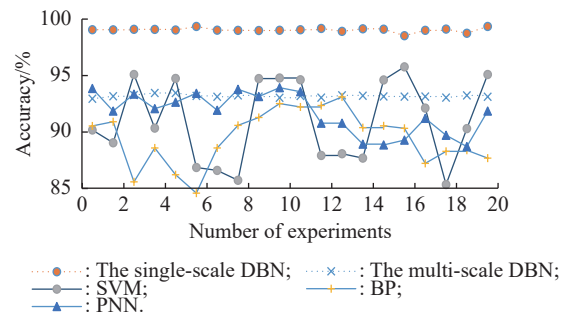
To highlight the advantages of the proposed method, four other methods (single-scale DBN algorithm, support vector machines (SVM) algorithm, back propagation neural network (BPNN) algorithm and probabilistic neural network (PNN) algorithm for fault diagnosis based on the same datasets) are used for comparisons. The multiple conditions including single time-varying condition and mixed time-varying conditions are taken into consideration. Additionally, the influence of noise on fault diagnosis is considered.

#### 4.4.1 Results of Experiment 1

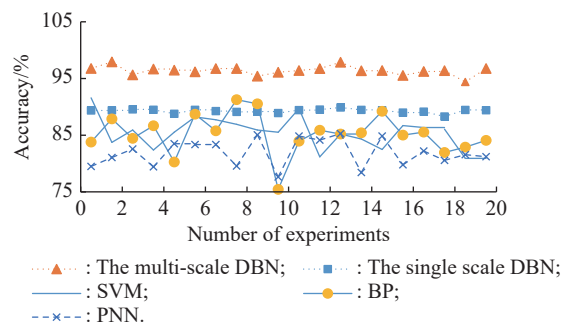
The bearing vibration data under single time-varying conditions (including the IS condition, DS condition, IDS condition and DIS condition) are used for fault diagnosis. Figs. 9–11 list the diagnostic accuracy of 20 experiments of several algorithms under different conditions, respectively. According to Figs. 9–11, the average accuracies as well as standard deviations of 20 fault diagnoses under three conditions are listed in Table 3, respectively.



**Fig. 9** Diagnostic accuracy of two fault diagnosis experiments under increasing speed conditions



**Fig. 10** Diagnostic accuracy of fault diagnosis experiments under decreasing speed conditions



**Fig. 11** Diagnostic accuracy of fault diagnosis experiments under increasing then decreasing speed conditions

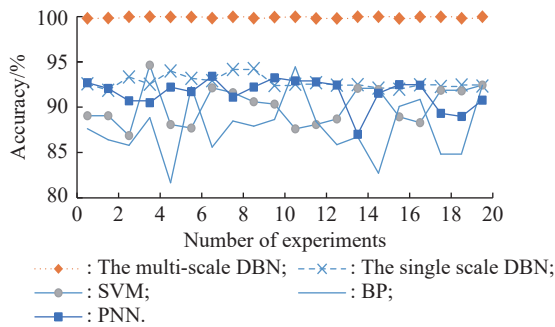
**Table 3** Fault diagnosis results

Method	Speed varying condition							
	Increasing speed		Decreasing speed		Increasing then decreasing speed		Combo time-varying rotational speed conditions	
	Average accuracy/%	Standard deviation	Average accuracy/%	Standard deviation	Average accuracy/%	Standard deviation	Average accuracy/%	Standard deviation
Multi-scale DBN	98.73	0.153	99.041	0.178	99.886	0.083	96.593	0.515
Single-scale DBN	96.48	0.749	93.131	0.190	92.739	0.693	89.177	0.527
SVM	92.77	0.190	90.986	3.673	90.082	2.118	85.368	2.861
BPNN	84.79	2.341	89.499	2.413	87.717	3.231	85.202	3.588
PNN	86.13	1.390	91.686	1.817	91.517	1.611	81.900	2.366

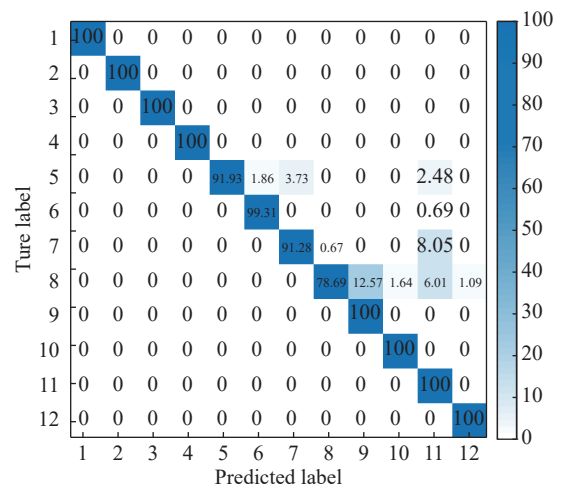
As shown in the figures, compared with the classical four methods, the proposed method has better diagnostic effect and stability. It can be seen from Table 3 that in 20 experiments under different conditions, the average accuracy of this method is higher, and the stability of diagnosis is better in terms of standard deviation.

#### 4.4.2 Results of Experiment 2

Bearing vibration data under combo time-varying conditions (including IS, DS, IDS and DIS conditions) are used for fault diagnosis. The average accuracy and standard deviation of 20 fault diagnoses are listed in Table 3, and the diagnostic accuracy of 20 experiments of several algorithms is shown in Fig. 12.

**Fig. 12** Diagnostic accuracy of fault diagnosis experiments under mixed time-varying rotational speed conditions

As can be seen from Table 3, the average accuracy of the proposed method is higher in 20 experiments under combo time-varying conditions, compared with the other four methods. Moreover, the proposed method has a better performance in diagnostic effects, and the diagnostic results are also stable, as shown in Fig. 12. In addition, it can be observed that the accuracy rate of mixed working conditions is less than that of single time-varying rotational speed condition in Fig. 13.

**Fig. 13** Confusion matrix for fault diagnosis

#### 4.4.3 Results of Experiment 3

The strong noise in the actual environment can easily interfere with the bearing vibration signal. However, the data collected from the experimental platform may be less disturbed by strong noise. The multi-scale processing method is used to filter out high-frequency interference and random noise partly. To verify the robustness of the proposed method, white Gaussian noise is added to the bearing vibration data (under the condition of increasing and then decreasing the speed). The signal to noise ratio (SNR) is expressed as

$$\text{SNR} = 10 \lg \frac{\text{signalpower}}{\text{noisepower}}. \quad (4)$$

The vibration signal under time-varying conditions mixed noise of  $-4$  dB to  $10$  dB is used to verify the robustness of the proposed method. Fig. 14 shows the original signal and the vibration signal of  $10$  dB mixed noise. The single-scale DBN algorithm, the SVM algorithm and the PNN algorithms are utilized to diagnose the faults with the same datasets. The results of fault diagno-

sis is shown in Fig. 15. It can be seen from Fig. 15 that the multi-scale method based on attention mechanism has a higher diagnostic accuracy than the single-scale method, the SVM method and the PNN method. Especially, when SNR=4, the obtained accuracy rate is greater than 98%.

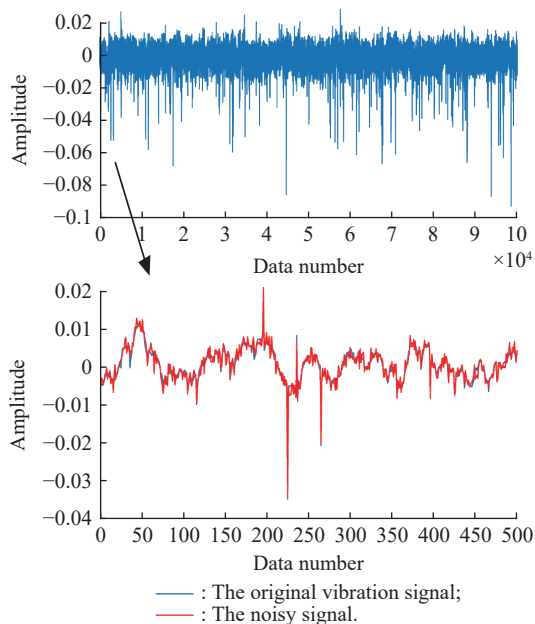


Fig. 14 Original signal and vibration signal of 10 dB mixed noise

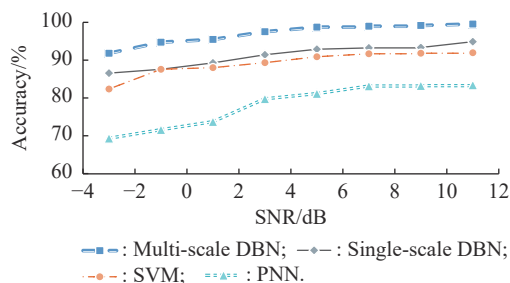


Fig. 15 Fault diagnosis results under noise conditions

## 5. Conclusions

In this study, to develop the bearing fault diagnosis method under multiple conditions, a multi-scale DBN method based on attention mechanism is proposed. The multi-scale vibration signal is obtained by multi-scale method, and multi-scale features are obtained by the multiple RBMs in different channels. The features are recalibrated by using the attention mechanism. At the classification level, the “softmax” activation function is used to realize fault diagnosis. Since the multi-scale method can weaken random noise to a certain extent, the influence of noise on fault diagnosis is also considered.

Experimental results demonstrate that the multi-scale DBN method based on attention mechanism has higher diagnostic accuracy and stability in fault diagnosis under

mixed time-varying conditions, compared with single-scale of DBN, SVM and BPNN algorithm. In addition, this proposed method has better diagnostic recognition rates and stronger robustness under different noise conditions.

## References

- [1] NANDI S, TOLIYAT H A, LI X D. Condition monitoring and fault diagnosis of electrical motors—a review. *IEEE Trans. on Energy Conversion*, 2005, 20(4): 719–729.
- [2] MOSHREFZADEH A. Condition monitoring and intelligent diagnosis of rolling element bearings under constant/variable load and speed conditions. *Mechanical Systems and Signal Processing*, 2021, 149: 107153.
- [3] WANG B, LEI Y G, LI N, et al. Multi-scale convolutional attention network for predicting remaining useful life of machinery. *IEEE Trans. on Industrial Electronics*, 2020, 68(8): 7496–7504.
- [4] WANG C S, LUN Y, CHEN Y H, et al. A data-driven aero-engine degradation prognostic strategy. *IEEE Trans. on Cybernetics*, 2021, 51(3): 1531–1541.
- [5] LEI Y G, YANG B, JIANG X W, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mechanical Systems and Signal Processing*, 2020, 138: 106587.
- [6] LI X, JIA X D, ZHANG W, et al. Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation. *Neurocomputing*, 2020, 383(28): 235–247.
- [7] CHEN Y H, PENG G L, ZHU Z Y, et al. A novel deep learning method based on attention mechanism for bearing remaining useful life prediction. *Applied Soft Computing*, 2019, 86: 105919.
- [8] PATHIRAGE C S N, LI J, LI L, et al. Application of deep autoencoder model for structural condition monitoring. *Journal of Systems Engineering and Electronics*, 2018, 29(4): 873–880.
- [9] CHEN H T, JIANG B, LU N Y. Probability-relevant incipient fault detection and diagnosis methodology with applications to electric drive systems. *IEEE Trans. on Control Systems Technology*, 2018, 27(6): 2766–2773.
- [10] LI Y, ZIO E, LU N Y, et al. Joint distribution-based test selection for fault detection and isolation under multiple faults condition. *IEEE Trans. on Instrumentation and Measurement*, 2020, 70: 1–13.
- [11] CHEN C, LU N Y, JIANG B, et al. Condition-based maintenance optimization for continuously monitored degrading systems under imperfect maintenance actions. *Journal of Systems Engineering and Electronics*, 2020, 31(4): 841–851.
- [12] JAVED K, GOURIVEAU R, ZERHOUNI N, et al. Enabling health monitoring approach based on vibration data for accurate prognostics. *IEEE Trans. on Industrial Electronics*, 2014, 62(1): 647–656.
- [13] NIU G X, WANG X, GOLDA M, et al. An optimized adaptive PReLU-DBN for rolling element bearing fault diagnosis. *Neurocomputing*, 2021, 445: 26–34.
- [14] LI X Q, JIANG H K, WANG R X, et al. Rolling bearing fault diagnosis using optimal ensemble deep transfer network. *Knowledge-Based Systems*, 2020, 213: 106695.
- [15] NASIRI A, TAHERI-GARAVAND A, OMID M, et al. Intelligent fault diagnosis of cooling radiator based on deep learning analysis of infrared thermal images. *Applied Thermal Engineering*, 2019, 163: 114410.
- [16] AZAMFAR M, LI X, LEE J. Intelligent ball screw fault diagnosis using a deep domain adaptation methodology. *Mechanism and Machine Theory*, 2020, 151: 103932.
- [17] YANG Z B, ZHANG J P, ZHAO Z B, et al. Interpreting net-



- work knowledge with attention mechanism for bearing fault diagnosis. *Applied Soft Computing*, 2020, 97: 106829.
- [18] JIAO J Y, ZHAO M, LIN J, et al. A comprehensive review on convolutional neural network in machine fault diagnosis. *Neurocomputing*, 2020, 417: 36–63.
- [19] LI C, ZHANG S H, QIN Y, et al. A systematic review of deep transfer learning for machinery fault diagnosis. *Neurocomputing*, 2020, 407: 121–135.
- [20] WANG Y R, SUN G D, JIN Q. Imbalanced sample fault diagnosis of rotating machinery using conditional variational auto-encoder generative adversarial network. *Applied Soft Computing*, 2020, 92: 106333.
- [21] IRFAN M, SAAD N, IBRAHIM R, et al. An intelligent fault diagnosis of induction motors in an arbitrary noisy environment. *Journal of Nondestructive Evaluation*, 2016, 35(1): 1–13.
- [22] HINTONG E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313(5786): 504–507.
- [23] ZHONG T, QU J F, FANG X Y, et al. The intermittent fault diagnosis of analog circuits based on EEMD-DBN. *Neurocomputing*, 2021, 436: 74–91.
- [24] LIU K, WU J K, LIU H B, et al. Reliability analysis of thermal error model based on DBN and Monte Carlo method. *Mechanical Systems and Signal Processing*, 2021, 146: 107020.
- [25] PAN Y B, HONG R G, CHEN J, et al. A hybrid DBN-SOM-PF-based prognostic approach of remaining useful life for wind turbine gearbox. *Renewable Energy*, 2020, 152: 138–154.
- [26] HU Q, SI X S, ZHANG Q H, et al. A rotating machinery fault diagnosis method based on multi-scale dimensionless indicators and random forests. *Mechanical Systems and Signal Processing*, 2020, 139: 106609.1–106609.22.
- [27] XU Z F, LI C, YANG Y. Fault diagnosis of rolling bearings using an improved multi-scale convolutional neural network with feature attention mechanism. *ISA Transactions*, 2021, 110: 379–393.
- [28] JIANG J R, LEE J E, ZENG Y M. Time series multiple channel convolutional neural network with attention-based long short-term memory for predicting bearing remaining useful life. *Sensors*, 2019, 20(1): 166.
- [29] LI X, ZHANG W, DING Q. Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism. *Signal Processing*, 2019, 161: 136–154.
- [30] JIE H, LI S, GANG S, et al. Squeeze-and-excitation networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023.
- [31] HUANG H, BADDOUR N, LIANG M. Bearing fault diagnosis under unknown time-varying rotational speed conditions via multiple time-frequency curve extraction. *Journal of Sound & Vibration*, 2018, 414: 43–60.

## Biographies



**LEI Xue** was born in 1991. She received her M.S. degree in control theory and control engineering from Lanzhou University of Technology, Lanzhou, China, in 2018. She is currently pursuing her Ph.D. degree in control theory and control engineering with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics. Her current research interests

include data-driven fault diagnosis, fault prognosis, and rotating machinery fault and health management.

E-mail: leixue@nuaa.edu.cn



**LU Ningyun** was born in 1978. She received her Ph.D. degree from Northeastern University, Shenyang, China, in 2004. From 2002 to 2005, she worked as a research associate and post-doctoral fellow in Hong Kong University of Science and Technology. She is currently a full professor at the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics,

Nanjing, China. Her research interests include data-driven fault prognosis and diagnosis and their applications to various industrial processes.

E-mail: luningyun@nuaa.edu.cn



**CHEN Chuang** was born in 1992. He received his M.Eng. degree in control engineering from Nanjing Tech University, Nanjing, China, in 2018, and Ph.D. degree in control theory and control engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2022. From March 2021 to April 2022, he became a China Scholarship Council (CSC)-funded joint

Ph.D. student at York University, Canada. He is currently a lecturer in College of Electrical Engineering and Control Science at Nanjing Tech University, Nanjing, China. His current research interests include data-driven fault prognosis and health management.

E-mail: chenchuang@nuaa.edu.cn



**HU Tianzhen** was born in 1996. She received her M.S. degree from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2020. She is currently pursuing her Ph.D. degree in control science and engineering at Nanjing University of Aeronautics and Astronautics. Her research interests include data-driven fault diagnosis and fault propagation analysis.

E-mail: hutianzhen@nuaa.edu.cn



**JIANG Bin** was born in 1966. He received his Ph.D. degree in automatic control from Northeastern University, Shenyang, China, in 1995. He had ever been a post-doctoral fellow, a research fellow, and a visiting professor in Singapore, France, USA and Canada, respectively. He is currently a chair professor of Cheung Kong Scholar Program in Ministry of Education, and the vice president of

Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include fault diagnosis and fault tolerant control and their applications in aircrafts, satellites and high-speed trains.

E-mail: binjiang@nuaa.edu.cn