Application of stack marginalised sparse denoising auto-encoder in fault diagnosis of rolling bearing

Junling Zhang\(^1\), Zhigang Chen\(^1,2\)\(\textsuperscript{3}\), Xiaolei Du\(^1\), Xu Xu\(^1\), Miao Yu\(^1,2\)

\(^1\)Institute of Electrical and Mechanical and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, People’s Republic of China
\(^2\)The Construction Safety Monitoring Engineering Technology Research Center of Beijing, Beijing 100037, People’s Republic of China
\(^3\)E-mail: gangz22@163.com

Abstract: When a fracturing vehicle is working, it generally needs to bear high loads, media corrosion and erosion. For this special working environment, this study proposes a rolling bearing fault diagnosis method based on stack marginalised sparse denoising auto-encoder (SDAE). This method combines the sparse auto-encoder (SAE) and the denoising auto-encoder (DAE) and combines the characteristics of dimensionality reduction and robustness. The method adds marginalisation to optimise the SDAE. Finally, it uses a two-layer stacking method. The output results of the second marginalised SDAE are used as input to the softmax classifier for learning training and classification testing. This improved method (stack SDAE) improves the denoising ability, reduces the computational complexity, solves the problems of difficult parameter adjustment and slows training convergence. The experimental tests were carried out on the failure of pitting corrosion of the outer ring of the bearing, pitting failure of the inner ring, and cracking of the rolling element. The results show that the algorithm can effectively improve the accuracy of fault diagnosis of rolling bearings, and it has greatly improved than the algorithms of SAEs and DAE.

1 Introduction

The fracturing truck is the main equipment for increasing oil production in some oilfields in China. Its main role is to inject high-pressure, large-volume fracturing fluids into the wells, extract and increase oil and gas reservoir permeability. The main working part of the fracturing car is the power end and the hydraulic end. The power end provides power. It should pay attention to daily maintenance. The hydraulic end is the main execution part. Its rolling bearings are subjected to high-fatigue loads, media corrosion and erosion during operation, and are prone to failure. If a fault occurs, the work of oilfield exploitation cannot continue. This will not only affect the production quality of the product and the economic efficiency of the manufacturer but also pose a great threat to the life safety of the operator. Due to this complicated working condition, the collected signals mostly exhibit nonlinear and non-stationary characteristics; at the same time, they are inevitably affected by various noises and signal modulation interferences, making the early fault signals extremely easy to be affected. It is very important to study the fault diagnosis of rolling bearing if it can be judged early on by this weak information to determine whether a fault has occurred and accurately determined the type of fault [1].

The rolling bearing fault diagnosis mainly has two aspects. One is to extract the characteristic index that can reflect the running state of the rolling bearing through the signal processing method. The traditional feature extraction methods are not auto-adaptive and cannot reflect the essential characteristics of the signal [2]. Based on this, an adaptive time-frequency analysis method called Hilbert transform is proposed [3]. However, the extreme value of the endpoint of this method is indefinite, making the decomposition result inaccurate. On this basis, an improved empirical mode decomposition algorithm is proposed [4]. This algorithm can basically completely characterise the fault and reduce the amount of calculation, but it requires certain prior knowledge. With the continuous development of deep learning, a fault diagnosis method based on auto-encoder (AE) is proposed [5]. A single AE cannot extract the characteristics of the signal very well, so there are many improved AEs. The authors of [6] proposed sparse AE (SAE) for fault diagnosis of rolling bearings.

This type of diagnostic method can select some important features, reduce the dimension of signal processing, but it is less robust. In [7], denoising AE (DAE) was proposed. The robustness of this type of encoder is good, but the dimensionality reduction effect is not as good as the SAE. Therefore, in [7], sparse DAE (SDAE) was proposed for fault diagnosis. It well combines the robustness of DAE and the dimensionality reduction of the SAE. However, this algorithm requires layer-by-layer training. Therefore, the amount of calculation is large, the training speed is slow, and it takes time to adjust to the optimal parameters. Based on the above analysis, this study proposes stack marginalised sparse DAE (SMDAE). It combines the SAE and the DAE. Then it is marginalised to form marginalised SDAE (MSDAE), and two layers are stacked to form SMDAE so as to troubleshoot and distinguish the type of rolling bearing.

2 Stack sparse denoising auto-encoder

2.1 Auto-encoder

AE is an artificial neural network, which contains three layers: an input layer, a hidden layer, and an output layer [8]. AE, the simplest artificial neural network model, is an unsupervised learning algorithm [9]. A simple system model of AE is shown in Fig. 1.

AE is a system that tries to restore the original input signal. It consists of an encoder and a decoder. Essentially, it does some transformation on the input signal. In other words, AE uses another way to express the input signal. The final goal is that the output signal is basically consistent with the input signal. Therefore, the input and output of AE should have the same structure. The structure is shown in Fig. 2.

As can be seen from Fig. 2, the hidden layer plays a very important part in AE. If it can restore the original signal, it should basically carry all the features of the original signal. So the hidden layer is also called the feature extraction layer. The forward conduction process of AE is divided into coding and decoding. So the AE is the feature extraction. Set the input signal to be \(\{x_1, x_2, \ldots, x_n\}\) and the coding function to be \(f_{\phi}\). The formula for converting from the input layer to the hidden layer is
\[ h_{W, \theta}(x) = f_{\theta}(x_0) = s(Wx_0 + b). \]  
(1)

\( \theta \) is the parameter matrix of the coding network, \( \theta = \{W, b\} \); \( s \) is the activation function of sigmoid; \( W \) is the weight matrix from the input layer to the hidden layer; \( b \) is the coefficient of offset.

After getting the hidden layer, it reaches the output layer through the decoding process. In this process, it needs the decode function. Set the decode function to be \( g(\theta' \cdot \text{output layer}) + \text{input layer} \) to the hidden layer; \( \theta' \) is the parameter matrix of the decoding network, \( \theta' = \{W', b'\} \), \( W' \) is the weight matrix from the hidden layer to the output layer and \( b' \) is the coefficient of offset.

AE is to choose the optimal parameter matrix. The goal is to make the error between the output layer and the input layer as small as possible in order to reduce the error of input and output data. To improve the practicality of the algorithm, it is necessary to limit the learning ability of AE. On this basis, many types of AEs have been developed, such as DAE, SAE and so on.

2.2 Denoising auto-encoder

When the input signal has no noise, the output of AE is basically the same as the input. However, when the input error is slightly higher, AE usually performs very poorly. For this reason, Vincent et al. proposed a DAE algorithm. The principle of DAE is shown in Fig. 3.

The DAE first adds the noise to the original signal \( x \) and obtains \( \tilde{x} \). Then the noise-contaminated signal is put into the DAE. The characteristic expression \( y \) of \( x \) is obtained by the encoding function \( g \), and the reconstructed signal \( z \) is obtained by \( f \). The reconstruction error is characterised by the loss function \( L_H(x, z) \). The DAE's overall loss function is

\[ L_{DAE}(\theta) = \frac{1}{N} \sum_{x \in D} J(x, y) + \frac{\lambda}{2} \| W \|_1^2 + \| W' \|_1^2. \]  
(4)

where \( J(x, y) = (1/2) \| x - y \|_1^2 \) and \( \lambda \) is a weight constraint that prevents overfitting.

2.3 Sparse auto-encoder

The sparse model plays a more and more important role in machine learning and image processing [10]. It has the function of variable selection, which simplifies the model and retains the most important information in the data. It has effectively solved many problems in the modelling of high-dimensional data [11]. Owing to this characteristic of the sparse model, it has a good effect on the AE. Therefore, the SAE is proposed [12]. When \( n \) of the hidden layer is greater than \( n \) of the input, the sparseness of the hidden layer needs to be limited. The most neurons in the hidden layer are in a 'suppressed' state, and a small number of neurons that can basically express all the information of the input data are in an 'active' state.

To make the AE sparse, it has introduced penalty items. They are used to punish \( \hat{\rho}_j \) which is far away from \( \rho \) (\( \rho \) is the sparsity parameter, \( \hat{\rho}_j \) is the average activation value of the \( j \)th hidden node). The specific form of the penalty item is as follows:

\[ KL(\rho | | \hat{\rho}) = \sum_{j=1}^{N} \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j}. \]  
(5)

\[ s_2 \] represents the number of neurons in the hidden layer; this expression is also called relative entropy. From the expression, we can see that the relationship between relative entropy \( \rho \) and \( \hat{\rho}_j \) is large. If the difference between the two is zero, the relative entropy is also zero too. The larger the value of the difference, the bigger the relative entropy value, i.e. the greater the value of the penalty item. After the penalty item is added, the SAE objective function is expressed as

\[ J_{space}(W, b) = J(W, b) + \beta \sum_{j=1}^{N} KL(\rho | | \hat{\rho}_j). \]  
(6)

where \( \beta \) is the weight used to control sparse penalty terms.

2.4 Stack sparse denoising autoencoder

The SAE has sparse characteristics, while the DAE has robust characteristics. Based on the advantages of these two methods, an algorithm that has both sparsity and robustness is proposed. SDAE combining the SAE and the DAE is obtained. Its loss function is also a combination of the two

\[ L_{SDAE}(\theta) = \frac{1}{N} \sum_{x \in D} J(x, y) + \beta \sum_{j=1}^{N} KL(\hat{\rho}_j | | \rho) + \frac{\lambda}{2} \| W \|_1^2 + \| W' \|_1^2. \]  
(7)

where \( \beta \) is the weight of the sparseness penalty factor.

Although SDAE already includes the advantages of two algorithms (SAE and DAE), it cannot express the most essential characteristics of data. Such a single algorithm is unable to meet the needs, so the stack SDAE (SSDAE) is proposed. SSDAE is a multi-layer SDAE. The output of the previous layer is used as the input of the next layer.
introduce the idea of limitation [14]. When the encoder is formed. Here, the value of \( \lambda \) is the number of layers to be used. To restore the original image at the most and improve the image processing to get damaged copies. The four bearing states were sampled through the experimental equipment, it measures the original and the obtained results are up to standard. If the target is met, SMSDAE training and the other part for testing. The time-frequency analysis is performed on the collected signals and the obtained results are used as training data and 400 were used as test data. They are classified the collected signals, and then use one part for training and the other part for testing. The time-frequency analysis is performed on the collected signals and the obtained results are pre-processed. After preprocessing, the input range is [0, 1] and the sigmoid function is selected as the activation function.

3.2 Diagnosis process

Based on the above analysis and SMSDAE model, this study designed the diagnosis process combined with the fault diagnosis of rolling bearings. The process is shown in Fig. 6.

The Taylor expansion of \( J(x, y) = J(x, m_{d}(\hat{x})) \) at \( x \) in the loss function of SDA is

\[
J(x, m_{d}(\hat{x})) \approx J(x, m(\mu_{i})) + \frac{\partial J}{\partial x} \hat{x} - \mu_{i} \frac{\partial^{2} J}{\partial x^{2}} \hat{x}^{2} + \ldots.
\]

For details on the derivation process, see [3]. Also, then get the loss function of MSDAE as follows:

\[
J_{MSDAE}(\theta) = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} \left( J(x, m_{d}(\mu_{i})) + \frac{\partial J}{\partial x} \hat{x} - \mu_{i} \frac{\partial^{2} J}{\partial x^{2}} \hat{x}^{2} \right) + \beta \sum_{j=1}^{M} KL(\hat{\rho}_{j} | \rho).
\]

Drawing on the idea of SSDAE, we put the two MSDAEs together to form the SMSDAE with two hidden layers. Its loss function is

\[
J_{SMSDAE}(\theta) = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} \left( J(x, m_{d}(\mu_{i})) + \frac{\partial J}{\partial x} \hat{x} - \mu_{i} \frac{\partial^{2} J}{\partial x^{2}} \hat{x}^{2} \right) + \beta \sum_{j=1}^{M} KL(\hat{\rho}_{j} | \rho).
\]

Finally, the output results of the second MSDAE are input into the softmax classifier for classification [15]. This will complete the training [16]. The overall schematic of MSDAE is shown in Fig. 5.
categorised and labelled to give Table 1. The inner and outer ring faults were made by machining a hole with a diameter of 0.3 mm on the outer ring and the inner ring; the wire cutting technology processes a groove with a width of 0.3 mm and a depth of 0.3 mm in the inner and outer rings of the bearing.

Figs. 7 and 8 are the time domain and spectrum diagrams, which are obtained by analysing the measured original vibration signals. The sampling frequency of the data file is 16 kHz, the motor speed is 1350 r/min, and the sampling point is 5 × 2000. Its rolling bearing model is SKF NJ202EC, with nine rollers, an outer diameter of 35 mm, an inner diameter of 15 mm, a thickness of 11 mm, and a contact angle of 0°.

This experiment only shows the frequency spectrum within the frequency range of 2000 Hz. From the comparison of the four small graphs in Fig. 8, it can be found that when the bearing works normally (the first figure in Fig. 8), its frequency signal characteristics are prominent, and the frequency peaks near 700 Hz and some small peaks are at frequencies <500 Hz. When the inner ring fails (the second picture in Fig. 8), the frequency signal is also obvious. It can be clearly observed that the peak value is reached near the frequency of 500 Hz, and basically, there are no other prominent frequencies. When a rolling element failure or other failure occurs (Figs. 3 and 4 in Fig. 8), it can be observed that the frequency signal is not very prominent and there are many very small local peaks. This shows that noise has submerged useful information in the signal.

### Table 1 Results on a miner

| Fault type          | Normal | Outer ring failure | Inner ring failure | Rolling element failure |
|---------------------|--------|--------------------|--------------------|------------------------|
| collected data      | 2000   | 2000               | 2000               | 2000                   |
| training data       | 1600   | 1600               | 1600               | 1600                   |
| test data           | 400    | 400                | 400                | 400                    |

4.2 SMSDAE training and analysis

In this experiment, SMSDAE sets the following parameters: add white noise with $\sigma = 35$; in order to make each hidden layer neuron satisfy sparsity, the average activation value is 0; the stack layer number is two layers; the minimum square error is used as the final loss function.

Select frequency domain samples for SMSDAE training and analysing the diagnostic results. With the SMSDAE system, it can automatically extract the inherent characteristics of the bearing fault data and complete the corresponding training. Finally, the softmax classifier classifies the fault types. The first three principal components of these features were extracted and visualised using principal component analysis. The results are shown in Figs. 9–11. Figs. 9–11 show the main characteristics of the frequency domain signal extracted from SSDAE, MSDAE, and SMSDAE, respectively. It can be seen from the figure that the scatter pattern of each type of rolling bearing is not shown in Fig. 9. However, in Fig. 10, it is only obvious that it is not ideal. It shows that the feature extraction ability of SSDAE for frequency domain signal is not strong, and MSDAE is stronger than SSDAE in feature extraction ability, but it still does not reach a very clear separation purpose. By observing Fig. 11, we can see that the characteristics...
of the main component of the frequency domain signal extraction distribution are very regular, and the types of rolling bearing motion states can be well separated. It explains that the effectiveness of SMSDAE is well for feature extraction, and also shows it is very accurate to judge the work state of rolling bearings.

To prove the good effectiveness and high accuracy of the SMSDAE method, the analysis of MSDAE (single-layer analysis) and SSDAE (double-layer unimproved analysis) was also performed in this experiment. 15 tests were conducted. The test results are shown in Fig. 12 and Table 2.

It can be seen from Fig. 10 that the diagnostic accuracy of MSDAE and SSDAE is very high, basically above 95%. The average accuracy of MSDAE is as high as 98.39%, which is nearly 1% higher than that of SSDAE. In contrast, the accuracy of MSDAE is higher. Also, the standard deviation of MSDAE is 0.0034%, which is relatively stable compared to SSDAE. This shows that the improved results are better. The accuracy of SMSDAE is basically 100%, which is the highest. The standard deviation is almost 0. Experiments have proved the accuracy and feasibility of the method.

5 Conclusion
This study made some improvements to the shortcomings in fault diagnosis of rolling bearings based on SSDAE. It proposes SMSDAE. This method firstly analyses the time-frequency of signals and passes the frequency information through a two-layer MSDAE in order to get the purpose of fault diagnosis for rolling bearings. Experiments show that this method is very effective for the fault diagnosis of rolling bearings, and by comparing with several other different AEs, it is proved that the diagnostic accuracy is high.

6 Acknowledgments
This research was supported by the Projects of the National Natural Science Foundation of China (no. 51004005) and the Beijing Excellent Talent Training (no. 2013D005017000013) and the Ministry of Housing and Urban-Rural Development of the People's Republic of China (no. 2016-K4-081) and the Beijing Municipal Commission of Education Science and Technology Plan (SQKM 201710016014) and the Fundamental Research Funds for Beijing Universities (X18217).

7 References
[1] Patel, V.N., Tandon, N., Pandey, R.K.: ‘Vibrations generated by rolling element bearings having multiple local defects on races *’, *Procedia Technol.*, 2014, 14, pp. 312–319
[2] Humphrey, E.J., Bello, J.P., Lecun, Y.: ‘Feature learning and deep architectures: new directions for music informatics’, *J. Intell. Inf. Syst.*, 2013, 41, (3), pp. 461–481
[3] Ali, H.A., Sai Ho, L., Mohammed, H.I.: ‘Diagnosis system for Parkinson's disease using speech characteristics of patients and deep belief network’, *CAAI Trans. Intell. Technol.*, 2017, 2, (4), pp. 246–253
[4] Cao, Y.: ‘Intelligent recommendation system based on neural network and fuzzy logic’ (Chongqing University, Chongqing, 2006) (in Chinese)
[5] Yu, K., Jia, L., Chen, Y., et al.: ‘Deep learning, yesterday, today and tomorrow’, *J. Comput. Res. Dev.*, 2013, 50, (9), pp. 1799–1804 (in Chinese)
[6] Tang, F., Long, H.: ‘Sparse auto-coded deep neural network and its application in fault diagnosis of rolling bearings’, *J. Mech. Eng. Tech.*, 2018, 3, (in Chinese)
[7] Chen, R., Yang, X., Yang, L., et al.: ‘Diagnosis of rolling bearing damage degree based on stacked sparse and noise auto-encoding deep neural network’, *J. Vib. Shock*, 2017, 36, (21), pp. 125–131 (in Chinese)
[8] Rumelhart, D.E., Hinton, G.E., Williams, R.J.: ‘Learning representations by back-propagating errors’ in *Neurocomputing: foundations of research* (MIT Press, 1988), pp. 533–536
[9] Lecun, Y., Bottou, L., Orr, G.B., et al.: ‘Efficient BackProp’ in *Neural networks: tricks of the trade* (Springer, Berlin, Heidelberg, 1998), pp. 9–50
[10] Sun, W., Shao, S., Yan, R.: ‘Induction motor fault diagnosis based on sparse automatic coded deep neural network’, *J. Mech. Eng.*, 2016, 52, (9), pp. 65–71 (in Chinese)
[11] Minh, V., Kavukcuoglu, K., Silver, D., et al.: ‘Playing atari with deep reinforcement learning’, *Comput. Sci.*, 2013, arXiv preprint arXiv: 1312.5602
[12] Yin, H., Jiao, X., Chai, Y., et al.: ‘Scene classification based on single-layer SAE and SVM’, *Expert Syst. Appl.*, 2015, 42, (7), pp. 3368–3380
[13] Ma, H., Ma, S., Xu, Y., et al.: ‘Image denoising based on improved stacked sparse denoising auto-encoder’, *Comput. Eng. Applic.*, 2018, (4), pp. 199–204 (in Chinese)

[14] Yuan, K., Ling, Q., Yin, W.: ‘On the convergence of decentralized gradient descent’, *Mathematics*, 2016, 26, (3), pp. 1835–1854

[15] Cheriyadat, A.M.: ‘Unsupervised feature learning for aerial scene classification’, *IEEE Trans. Geosci. Remote Sens.*, 2013, 52, (1), pp. 439–451

[16] Chen, M., Weinberger, K., Sha, F., et al.: ‘Marginalized denoising auto-encoders for nonlinear representations’, Int. Conf. on Machine Learning, Beijing, China, June 2014, pp. 1476–1484