2603. A hydraulic fault diagnosis method based on sliding-window spectrum feature and deep belief network

Xinqing Wang[1], Jie Huang[2], Guoting Ren[3], Dong Wang[4]
College of Field Engineering, People's Liberation Army University of Science and Technology, Nanjing, China
[2]Corresponding author
E-mail: [1]*wwwxxxqqq@126.com*, [2]*huangjie051501@126.com*, [3]*871625417@qq.com*, [4]*dyhkxydfbb@126.com*

Check for updates

Abstract. The vibration signal of hydraulic system contains abundant state information, so vibration testing technology is an effective way to realize the fault diagnosis of hydraulic system. However, the mapping relation between signal characteristic and system state is complex and the expression meaning of characteristic is obscure, which brings a great challenge to the hydraulic fault diagnosis. The DBN, a newly proposed deep learning model, has an advantage of autonomously learning and reasoning. And it is good at studying the concealed representation of data and highlighting the feature expression. So, it is contributive to deal with the problems of large capacity data like high dimension, redundancy, and nonlinear etc. Therefore, DBN is chosen as the fault diagnosis method in this paper. Meanwhile, given that the difficulty in feature extraction of hydraulic vibration signal and the important influence of input feature vector to the diagnosing of DBN, a fast and effectively feature extraction method based on sliding-window spectrum feature (SWSF) is proposed. It is effective in remaining the integrity of feature, avoiding the risking of relative shifting of characteristic spectrum, and decreasing the dimensions of feature vector. The experimental results demonstrate that the combination of SWSF and DBN is a fast and effective approach to realize the fault diagnosis of hydraulic system.

Keywords: hydraulic system, fault diagnosis, feature extraction, sliding-window spectrum feature, deep belief network.

1. Introduction

Hydraulic system plays an important role in modern industry equipment, so it is of great significance to accurately detect and diagnose fault of hydraulic equipment in order to ensure its safety running. The structure of hydraulic system and its sealed working characteristic make the hydraulic fault present the following features [1, 2], such as invisibility, susceptible to random factors, the complex mapping relation between signal characteristic and system state, etc. Therefore, it is difficult to judge the fault state through the signal directly, so selecting an appropriate diagnosis method is critical. Deep belief network (DBN), a newly proposed deep neural network by Hinton [3], has a strong ability of autonomously learning and reasoning. It is able to focus on studying the hiding representation of data and highlighting the expression of characteristic data, so it is contributive to deal with the problems of large capacity data like high dimension, redundancy, and nonlinear etc. DBN can obtain higher level characteristics from the low-level features by means of unsupervised and greedy training, and is widely used in fields of machine learning and pattern recognition [4]. In recent years, the applying of DBN to fault diagnosis problems has drawn more and more attention of scholars. For instance, Tamilselvan [5] uses DBN for fault diagnosis of aircraft engine based on the health state classification. Shao [6] employs an optimization DBN in rolling bearing fault diagnosis through several time-domain features. But Tamilselvan adopts the raw data directly, which needs a high quality data. And Shao selects several kinds of features factitiously, which excessively depends on experimenter's experience.

The essence of fault intelligent recognition is the recognizing of specific characteristics describing the corresponding fault by machine learning method, so feature extraction is critical to realize fault intelligent recognition. When DBN is applied to fault diagnosis, the input feature vectors can be varied, whose availability and dimension have a great influence on the diagnosis accuracy and efficiency of DBN [7]. With the development in vibration testing field, most of the test equipment has the trait of large capacity, high frequency and speediness. Consequently, the original measured signal is usually in great amount and high-dimensional, if the original measured signal is directly taken as the input sample of DBN, it will lead to a complicated processing and a long running time, which is adverse to the real-time and rapid diagnosis. So, a certain preprocessing for the original data is essential to improve the diagnosis accuracy and efficiency of DBN.

Generally speaking, the means of obtaining fault feature can be divided into two kinds. One is directly extracting the parameters of original data in time domain, frequency domain or energy domain [8, 9], such as average, kurtosis, etc. And these parameters are used as input features of DBN. Another is applying the approach of decomposition [10], such as wavelet decomposition [2, 11], empirical mode decomposition [12, 13], singular value decomposition [14], etc. Then the further disposing results of decomposed components are taken as input features. As for the first means, it mainly depends on experience, which is susceptible to artificial subjective factors. For another decomposition means, an appropriate decomposition method should be chosen firstly, and then a series of onerous processes are carried on, which have not made full use of the powerful function of such a deep learning model like DBN. Therefore, aiming at the fast and effectively generating of input feature vector, a feature extraction method based on sliding-window spectrum feature (SWSF) is put forward in this paper. This method starts with the spectrum function of original signal, gets the input feature vector of DBN through the approach of sliding window, thereby avoiding the manual process of feature extraction and selection, enhancing the intelligence of recognition process, and giving full play to the strong ability of autonomous learning and reasoning of DBN. Through experimental verification, it indicates that the combination of SWSF method and DBN is a quick and effective approach to realize the fault diagnosis of hydraulic system.

2. The DBN principle

2.1. The DBN structure

A DBN is a feedforward neural network stacked with several Restricted Boltzmann Machines (RBMs), which consists of a visible layer, a number of hidden layers, and an output layer [15]. As shown in Fig. 1, a DBN with four layers is taken as the example to introduce the principle of DBN. This DBN includes two unsupervised RBMs and one supervised BP classifier. In the DBN, the visible layer used to accept input data and the 1st hidden layer compose the RBM1, the 1st and 2nd hidden layers compose the RBM2, and the 2nd hidden layer and the output layer compose the BP classifier.
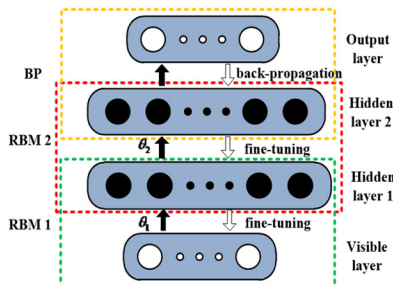

Fig. 1. Structure diagram of a DBN with two hidden layers

The DBN training process consists of two parts, including unsupervised training and supervised training. In the unsupervised training part, the entire network parameters are initialized firstly, and then the RBMs are trained through greedy algorithm. After the training of last RBM is completed, the output of its hidden layer is taken as the input of visible layer of next RBM, and by thus layer-by-layer training, the parameters of each RBM are acquired. In the supervised training part, the label layer of data is brought into training process. The training error is distributed to each RBM by BP classifier through back-propagation algorithm, thereby realizing the updating of the parameters.

2.2. The training process of RBM

RBM is a special kind of Markov random field, as well as an energy-based stochastic neural network [16], its energy function of visible layer v and hidden layer h can be defined as:

$$E(v, h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i h_j w_{ij}, \tag{1}$$

where v_i and h_j are the binary state of visible unit i and hidden unit j, w_{ij} is the weight between them, a_i and b_j are the biases of corresponding units.

Each RBM has one visible layer and one hidden layer, for the two layers, all the visible units and hidden units are connected to each other symmetrically and bidirectionally, and there is no connection between any two units within the same layer. Any of the unit state can be either 1 or 0. The joint probability distribution of visible layer v and hidden layer h, and the conditional probabilities of visible unit v_i and hidden unit h_j can be given as:

$$P(v, h) \propto \exp(h^T W v + b^T v + a^T h), \tag{2}$$

$$P(v_i = 1|h) = \frac{1}{1 + \exp(-a_i - \sum_j h_j w_{ij})}, \tag{3}$$

$$P(h_j = 1|v) = \frac{1}{1 + \exp(-b_j - \sum_i v_i w_{ij})}, \tag{4}$$

where W is the weight matrix between visible unit and hidden unit, respectively a and b are the biases matrix of corresponding units.

Researches from experts like Hinton have proved that the parameters $\theta = \{W, a, b\}$ making $p(v, h)$ maximized can be obtained by Contrastive Divergence (CD) algorithm. According to CD algorithm, the update rule of the RBM parameters $\theta = \{W, a, b\}$ is:

$$\Delta W = \varepsilon_W (\langle vh^T \rangle_{data} - \langle vh^T \rangle_{recon}), \tag{5}$$
$$\Delta a = \varepsilon_a (\langle v \rangle_{data} - \langle v \rangle_{recon}), \tag{6}$$
$$\Delta b = \varepsilon_b (\langle h \rangle_{data} - \langle h \rangle_{recon}), \tag{7}$$
$$W \leftarrow W + m\Delta W, \tag{8}$$
$$a \leftarrow a + m\Delta a, \tag{9}$$
$$b \leftarrow b + m\Delta b, \tag{10}$$

where ε_W, ε_a, ε_b are learning rate of weight, biases of hidden layer and visible layer. $\langle \cdot \rangle_{data}$ refers to the expectation decided by the probability distribution of training data, and is easy to calculate its unbiased estimation. $\langle \cdot \rangle_{recon}$ refers to the expectation decided by the probability distribution of reconstructed data. m is the inertiafactor, which associates the anteroposterior update parameters, and is used to keep the stability of parameter updating.

2.3. The advantages of DBN

A large number of experiments prove that, compared with the traditional BP neural network, DBN doesn't need many labeled training samples, and is no longer limited to the effect of random initialization. RBM is trained layer-by-layer through unsupervised and greedy training, and the training result is taken as the initial value of probability model of supervised learning. This approach can speed up the convergence speed of network, solve the problem of tending to fall into local optimal of BP neural network, improve the learning ability, and provide a technical support for the efficient and deep learning.

3. The principle of SWSF method

3.1. Problems of applying characteristic spectrum

The spectrum curve of vibration signal contains the system's state information, when the state varies, the spectrum would also change. Spectrum curve is actually a multimodal function, which usually consists of multiple local peaks [17]. As shown in Fig. 2, the spectrum curve of a hydraulic signal taken from the article's experiment set, contains nearly 10 significant local peaks. Actually, each local peak is corresponding to a vibration mode. Therefore, in view of the signal's generation mechanism, a local peak should be a complete feature. In one local peak, pick out the spectrum of maximum amplitude, and define its frequency as the characteristic spectrum of this local peak, such as the frequency of 95.21 Hz in Fig. 3. Obviously, a characteristic spectrum is a vibration mode, and a local peak can be considered as a frequency band decided by the certain characteristic spectrum, and the frequency bands vary in width.

In the fault diagnosis based on vibration testing technology, the spectrum is often taken as the input features of classifier. In reference [18], the whole Hilbert envelope spectrum of bearing fault signals are used as the feature vectors of DBN, which may cause the following defects. For example, the frequency band containing characteristic spectrum would be dismembered, which destroys the completeness of feature. And make the dimension of input vector very high, which would add the burden of classifier and increase the running time. Besides, for the same type of test sample, there is the risk of relative shifting of characteristic spectrum. For example, Fig. 3 shows the spectrogram of two samples of the same type, judging from the significant frequencies of the two figures, the frequencies of 383.3 Hz and 385.7 Hz should be the representation of the same fault feature. However, the relative shifting of characteristic spectrum appears as a result of test environment or man-made factors, which would have a negative impact on the subsequent classification process.
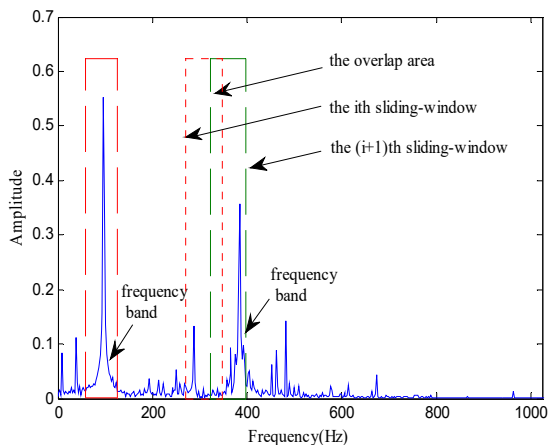

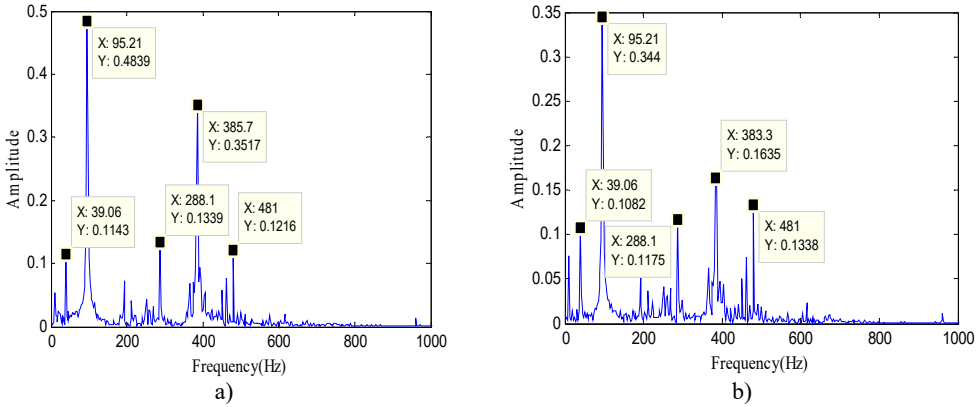
Fig. 2. The spectrum curve of a hydraulic signal

Fig. 3. The relative shifting of characteristic spectrum in different samples

Obviously, the division of spectrum curve should be an effective method for above defects. Krakovsky [19] uses the moving-window discrete Fourier transform as a multichannel filter. Ye [20] do the modal decomposition by segmenting the spectrum curve. And Ma [21] uses sliding window FFT algorithm to analysis low frequency oscillation. These papers mainly utilize sliding window for modal analysis or frequency filter, and they split the spectrum function directly into several fragment. But in this paper, the sliding method is used as a feature extraction approach and takes the overlap rate into account. Ideally, by dividing each frequency band corresponding to the characteristic spectrum should be well separated, and these frequency bands would be used as feature vectors. In this case, the integrity of characteristic frequency band is kept, the relative shifting of characteristic spectrum avoided and the dimensions of feature vectors reduced. If adopting the method of equidistant division, it would be difficult for the complete characteristic frequency band to avoid being broken down. If using the method of artificial division, the workload would be increased greatly.

3.2. Specific steps of SWSF method

Therefore, in this article, the way of sliding for division is adopted and a feature extraction method based on sliding-window spectrum feature (SWSF) is proposed. The specific steps are as follows.

Step 1, perform spectrum transform for vibration signal $x(t)$. In this article, calculate the Hilbert envelope spectrum of $x(t)$, and take its unilateral spectrum sequence $F(\omega)$. The subsequent dealing objects are pointing at $F(\omega)$.

Step 2, determine the width W of sliding-window and the overlap rate O of adjacent sliding-windows, and ensure that the product of W and O is an integer (see Fig. 2).

Step 3, the abscissae of starting and ending of the first sliding-window can be given as:

$$start(1) = 1, \tag{11}$$
$$end(1) = W. \tag{12}$$

Sum up spectrum sequence of $F(\omega)$ in the first sliding-window, and get the first sliding spectrum feature $H(1)$:

$$H(1) = \sum_{w=1}^{W} F(\omega). \tag{13}$$

Step 4, conduct the ith sliding. Judge whether the ending abscissa of the last sliding is not

greater than the sequence length of $F(\omega)$. If not, the loop terminates. Otherwise, continue sliding. At the moment, the abscissae of starting and ending of the ith sliding-window can be given as:

$$start(i) = start(i-1) + (1-O)W, \qquad (14)$$
$$end(i) = end(i-1) + (1-O)W. \qquad (15)$$

Sum up spectrum sequence of $F(\omega)$ in the ith sliding-window, and get the ith sliding spectrum feature $H(i)$:

$$H(i) = \sum_{w=start(i)}^{end(i)} F(\omega). \qquad (16)$$

Step 5, obtain the new feature sequence $H(v)$ of sliding spectrum.

By the feature extraction process SWSF for vibration signal, the feature sequence $H(v)$ is obtained, which is similar to the histogram sequence in statistical method (hereinafter referred to as sliding spectrum).

The Hilbert envelope spectrum is adopted in the SWSF method, its process can be given as:

1) Perform the Hilbert transform for signal $x(t)$:

$$H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(t)}{t-\tau} d\tau. \qquad (17)$$

2) Calculate the envelope signal $e(t)$:

$$e(t) = \sqrt{x^2(t) + H^2[x(t)]}. \qquad (18)$$

3) Carry on Fourier transform for the envelope signal $x(t)$, select its unilateral spectrum sequence $E(f)$ as the envelope spectrum of signal $e(t)$:

$$E(f) = \frac{1}{\pi} \int_{-\infty}^{+\infty} e(t)dt. \qquad (19)$$

3.3. An example of applying SWSF

The feature extraction process SWSF is performed for a hydraulic signal under the sliding width of 2 and 4, and overlap rate of 0 respectively. According to the result shown in Fig. 4, the dimensions of sliding spectrum reduce significantly, which are 512 and 256, compared with the 1024 dimensions of the signal's Hilbert envelope spectrum.

4. Experiment validation of the proposed diagnosis solution for hydraulic system

4.1. Experiment introduction

The vibration signals used in this paper are measured on the synthetical hydraulic fault experimental platform as shown in Fig. 5, which is entirely made of steel. And these signals are respectively tested under the five states of normal, blockage, leakage, cavitation and impulsion. The measuring points are distributed on the hydraulic cylinder, the sampling frequency is 5000 Hz and the number of sampling points is 2048. Fig. 6 illustrates the time-domain graphs of signals in normal and other four failure states. The figure indicates that the tested vibration signals

present different variation rule with different working states, but it is not enough to distinguish these states just depending on the time domain analysis.
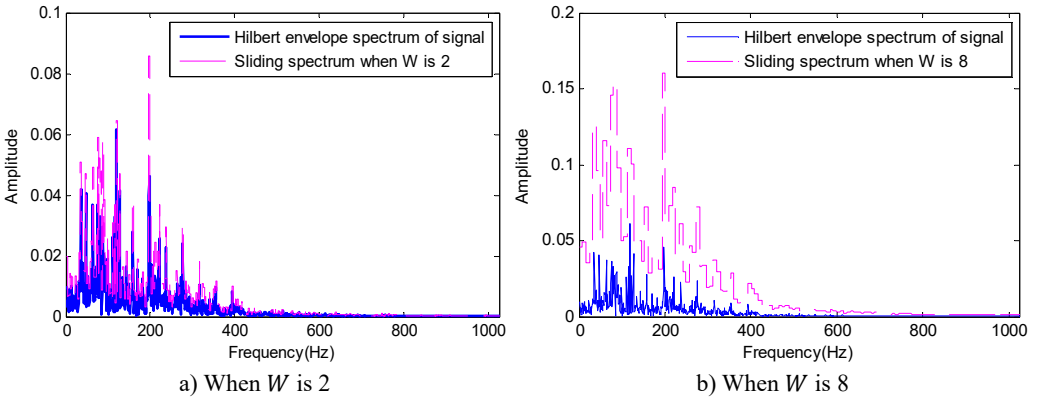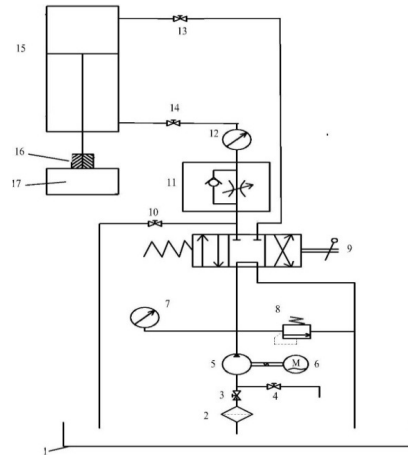

a) When W is 2

b) When W is 8

Fig. 4. Sliding spectrums in two cases

a)

b)

Fig. 5. The synthetical hydraulic fault experimental platform and its schematic diagram. 1 – fuel tank, 2 – suction filter, 3 – control valve of oil-absorbing blockage, 4 – control valve of cavitation, 5 – hydraulic pump, 6 – electromotor, 7 – piezometer, 8 – relief valve, 9 – hand-directional valve, 10 – control valve of leakage, 11 – one-way throttle valve, 12 – flowmeter, 13 – control valve of oil inlet blockage, 14 – control valve of oil outlet blockage, 15 – hydraulic cylinder, 16 – clamping sleeve, 17 – load

4.2. Analysis of experiment results

To validate the proposed hydraulic fault diagnosis method combining SWSF and DBN, the data tested from the synthetical hydraulic fault experimental platform is used as the experiment sample set. The training set is consist of 300 samples in normal state, 300 in blockage, 300 in leakage, 300 in cavitation and 180 in impulsion, totaling 1380 samples. And the corresponding testing set is made up of 200 samples in normal state, 200 in blockage, 200 in leakage, 200 in cavitation and 100 in impulsion, totaling 900 samples. Before diagnosis of the experiment sample set, the main structure parameters of DBN need to be determined, and then situations in different sliding width and overlap rate. During the discussing process, the evaluation criteria mostly depend on testing accuracy, while allowing for the elapsed time flexibly.
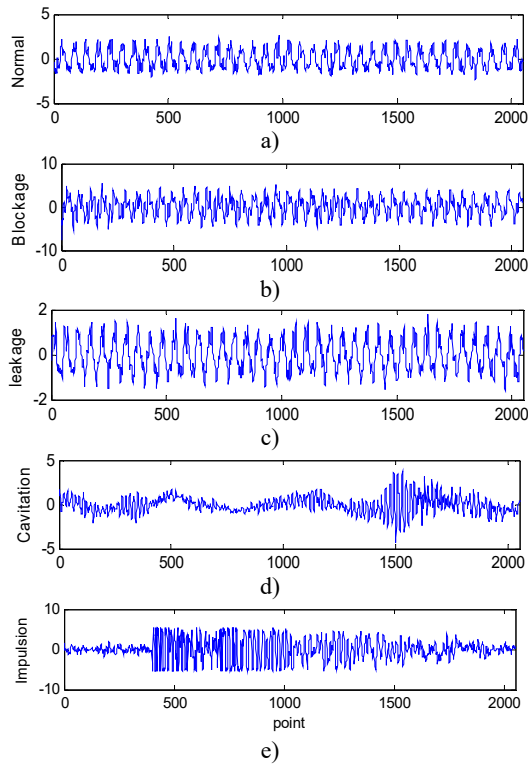
Fig. 6. Time-domain plots of hydraulic vibration signals in five different states

4.2.1. Discussion on the main structure parameters of DBN

The structure of adopted DBN is M-n-n-5, where M is the dimension of input data, n is the number of units in the hidden layer which usually vary from 10 to 100, and the 5 represents 5 types of states. Obviously, n is an important parameter to for DBN, besides, training times of DBN has a great effect on the classification accuracy and consuming time. Thus, the n and training times are discussed in the following three cases, when the W and O are 2 and 0, 4 and 0.5, 16 and 0.25 respectively.

1) Discussion on the number of hidden units.

When discussing about the value of the number n of hidden units, define the training times as 50 firstly. Then repeating classification testing 10 times, get the averages of these testing results shown in Table 1. By comparison，it is found that the classification accuracy is more better when the number of hidden units is 100, therefore, 100 is decided as the value of n finally.

Table 1. Classification results of the DBN with different hidden units

Classification accuracy (%)	Number of hidden units					
	20	40	60	80	100	20
W/O	2/0	93.11	94.56	97.56	97.78	98.00
	4/0.5	92.56	96.89	96.00	97.33	97.33
	16/0. 25	89.56	89.78	90.00	90.22	91.56

2) Discussion on training times.

According to the previous result, when discussing the training times, choose 100 hidden units and repeat classification testing 10 times. The averages of testing results are displayed in Fig. 7. Combined with the three cases, the curves of classification accuracy tend to be smooth after training times reaches 40. Because the consuming time of DBN's diagnosis is positive correlated

with the training times, so 40 is confirmed as the final value of training times.

Based on the above discussion results, finally, the network structure of DBN can be given as M-100-100-5, and the training times is 40.
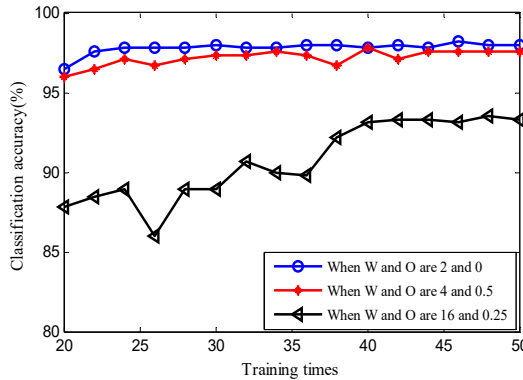

Fig. 7. Classification results with different training times

4.2.2. Cases in different sliding width and overlap rate

The original vibration signals measured in the synthetical hydraulic fault experimental platform are time-domain signals, their length is 2048. By the process of SWSF, the sliding spectrum is obtained, in the next, the sliding spectrum is taken as input sample of the DBN to perform training and classifying, and then what state each sample in is recognized, thereby the fault diagnosis of hydraulic system is realized. It is known that if different sliding width W and overlap rate O are chosen, different sliding spectrum will be got, and the difference mainly reflects in the constitution and dimension of feature, which has a great negative influence on classification accuracy and running time (During the whole running time of DBN, the time spent on the training process occupies a major portion, compared with which, the part time for classifying process is negligible). Then，if the most optimal solution of W and O can be sought out, which can ensure both a high classification accuracy and a less consuming time, thereby achieving the purpose of rapid and real-time diagnosis for hydraulic system. In this article, the choice of W is varied in the range of $\{1, 2, 4, 8, 16, 32, 64\}$, and the choice of O is $\{0, 0.125, 0.25, 0.5\}$. The detail diagnostic result is shown in Table 2 and Table 3.

Table 2 presents the varied classification result of DBN with different sliding width and overlap rate. There are some vacancies in the table because of the corresponding products of sliding width and overlap rate are not integers, which lead to the sliding process unable to be conducted. It can be seen from Table 2 that the classification accuracy can reach more than 90 % when the sliding width is not more than 16, and the accuracy is above 97 % when the sliding width is 2 or 4 From the horizontal analysis of Table 2, it is found that when the sliding width is greater than 4, the classification accuracy declines gradually with the increase of sliding width. On the contrary, when the sliding width is less than 4, the classification accuracy ascends gradually with the decrease of sliding width.

Table 2. Classification results of the DBN with different sliding width and overlap rate

Classification accuracy (%)		W						
		1	2	4	8	16	32	64
O	0	96.89	98.00	98.44	93.11	89.56	76.89	56.66
	0.125	–	–	–	93.33	91.78	80.67	57.55
	0.25	–	–	97.11	95.11	92.22	84.67	60.65
	0.5	–	97.11	98.00	94.00	90.00	80.00	67.11

Thus, when the sliding width is 4, the difference of features is most obvious and the classification accuracy reaches the highest. The reasons of above result can be described in the following two aspects. On the one hand, when sliding width is small, the relative shifting of characteristic spectrum and some non- characteristic spectrum have an interference in classifying. On the other hand, when sliding width is relatively large, too much spectrum is mixed together, which confuses the distinction between different characteristic spectrum and increases the difficulty of classification. From the vertical analysis of Table 2, it is also found that the overlap rate has little influence on the classification accuracy while the sliding width is small. But once that the sliding width is larger than 4, the classification accuracy when the overlap rate is not 0 are higher than that is 0, and the larger the sliding width is, the more obvious the situation is. It is because the overlap rate is used to dispose the border problem. Each characteristic spectrum does not exist in isolation, this characteristic spectrum and part of its adjacent spectrums belong to one characteristic frequency band decided by one vibration mode. So, in some degree increasing the overlap rate can make the sliding process smoother, which contributes to retaining the integrity of characteristic frequency band and avoiding it being dismembered, thereby improving the classification accuracy to some extent.

Table 3. Dimension of sliding spectrum and diagnostic consuming time
of DBN when sliding width and overlap rate are different

Time(s) / dimension		W						
		1	2	4	8	16	32	64
O	0	27.88/1024	18.54/512	13.63/256	11.21/128	9.27/64	8.45/32	8.11/16
	0.125	–	–	–	11.39/146	9.53/170	8.73/36	8.39/18
	0.25	–	–	15.13/341	12.07/170	10.09/85	9.02/42	8.77/21
	0.5	–	27.33/1023	18.13/511	13.83/255	11.12/127	9.45/63	8.74/31

Table 3 describes the changes of the dimension of sliding spectrum and the diagnostic consuming time of DBN. According to Table 3, when the overlap rate remains unchanged, the larger the sliding width is, the less the dimension of sliding spectrum is and the shorter the consuming time is. When the sliding width remains unchanged, the higher the overlap rate is, the higher the dimension of sliding spectrum is and the longer the consuming time is. In a word, with a larger sliding width and a lower overlap rate, the dimension of sliding spectrum will be less and the consuming time shorter. Combining Table 2 and 3, it can be discovered that the difference between the longest consuming time and the shortest one is nearly three times. And when the sliding width is 4, the classification accuracies are about 98 %, and the computation time are 13.63 s,15.13 s and 18.13 s respectively, with a difference of 4.5 s between the minimum and the maximum. So, choose a larger sliding width and a smaller overlap rate in order to decrease the consuming time when the required classification accuracy is guaranteed.

In conclusion, the proposed method combining SWSF and DBN is fast and effective for the intelligent fault diagnosis of this hydraulic experimental sample set. Meanwhile it is very crucial to choose an appropriate sliding width and overlap rate for the certain sample set with taking the classification accuracy and time cost into consideration. And the optimal solution for this hydraulic experimental sample set is that sliding width and overlap rate are 4 and 0 respectively.

4.2.3. Change process of features in DBN

In order to study the change process of the feature in each layer of DBN, the input data and output data of two hidden layers are disposed by the principal component analysis (PCA) method. For an alternative visualization, the 3D plots of the first three principal components are shown in Fig. 8, when sliding width is 3 and overlap rata is 0. Fig. 8(a) indicates a poor aggregation with the mixing of a great part features of the 5 states. After the first hidden layer, as shown in Fig. 8(b), there are significant differences between features in each state, except for that between leakage and impulsion, besides a slight mixing is exist between normal and blockage. After the

second hidden layer (Fig. 8(c)), in addition to the mixing of a few features between leakage and impulsion, the aggregation extent and the difference between each state improve greatly. The above result suggests that through the feature extracting in hidden layers of DBN layer by layer, the distribution of feature is greatly optimized, the expression ability improved and the discrimination enhanced, thereby performing the following state classifying effectively.
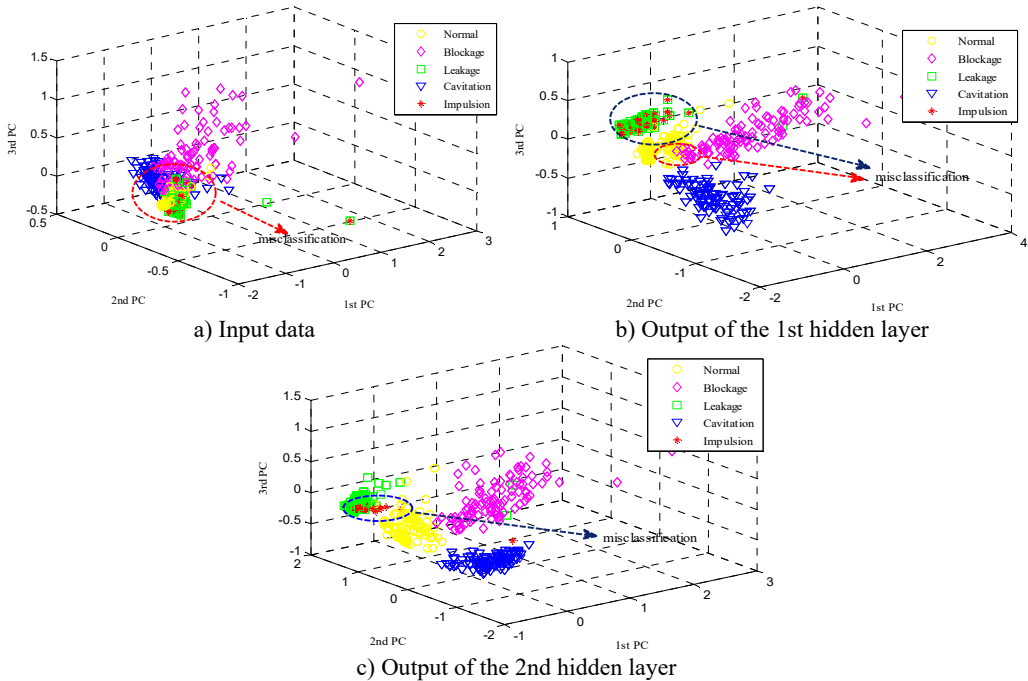

a) Input data

b) Output of the 1st hidden layer

c) Output of the 2nd hidden layer

Fig. 8. The first three principal components of testing samples

4.2.4. Contrastive analysis of different classifiers

Numerous studies have demonstrated that the SVM [22] is outstanding in shallow network model among kinds of classification algorithms, such as AdaBoost [23], BP neural network, etc. Therefore, to examine the classification performance of DBN, a representative of deep network model, the contrastive analysis is conducted among DBN, SVM and SVM optimized by PSO. Take the sliding spectrum when sliding width is 3 and overlapping rate is 0 as input features (the data dimension is 256), and get the test results as shown in Table 4. When the classifier is SVM, the parameters using default values, a classification accuracy of 80.67 % is acquired and the consuming time is 1.22 s. When the parameters c and g of SVM are optimized by the PSO algorithm, the classification accuracy raise to 94.67 %, but the occupation time is as long as 830 s. According to the comparison of Table 4, when adopting DBN as classifier, the classification accuracy is obviously the highest and the occupation time is relatively short. It is further evidence that, when feature dimension is high and the mapping relation between feature and state is complex, DBN has an advantage over shallow network models liking SVM.

Table 4. Comparison of different classifiers

Classifier	Classification accuracy (%)	Consuming time (s)
DBN	98.44	13.63
SVM	80.67	1.22
SVM+PSO	94.67	830

5. Conclusions

Aiming at the intractable problem of fault diagnosis for hydraulic system, a novel diagnosis solution combining a new feature extraction method based on sliding-window spectrum feature and DBN is presented. Through experimental verification, it is demonstrated that the proposed diagnosis solution is able to realize the intelligent fault diagnosis of hydraulic system quickly and efficiently.

1) DBN is a deep learning model adopting the greedy learning algorithm layer by layer, which has a powerful learning ability. In the aspects of processing high-dimensional, nonlinear and large volume data, DBN outperform shallow network models such as SVM, which can better represent the complex mapping relation between measured signal and equipment, and is useful to realize the fault diagnosis of hydraulic system.

2) The proposed feature extraction method based on sliding-window spectrum feature is contributive to the rapidly and efficiently generating of input feature vectors of DBN. To some extent, by means of this method, the dismembering of frequency band containing characteristic spectrum can be avoided, the integrity of feature kept, the risking of relative shifting of characteristic spectrum reduced, and the dimensions of feature vectors decreased.

3) For different diagnostic objects, when the SWSF method is adopted, it is needed to choose a suitable sliding solution accordingly. Taking into account both diagnosis accuracy and time cost, as far as possible a larger sliding width and a smaller overlap rate should be selected for a better solution.

Acknowledgements

References

[1] **Shen M. X., Zhao L. Z., Zhou B., Zhou C. M.** Hydraulic pump fault diagnosis of RBF network based on PSO optimization. Chinese Hydraulics and Pneumatics, Vol. 5, Issue 1, 2016, p. 87-92.

[2] **Zhao X. X., Zhang S. S., Zhou C. L.** Experimental study of hydraulic cylinder leakage and fault feature extraction based on wavelet packet analysis. Computers and Fluids, Vol. 10, Issue 6, 2015, p. 33-40.

[3] **Hinton G. E., Osindero S., Teh Y. W.** A fast learning algorithm for deep belief nets. Neural computation, Vol. 18, Issue 7, 2006, p. 1527-1554.

[4] **Chen Z. Y., Zeng X. Q., Li W. H., et al.** Machine fault classification using deep belief network. IEEE International Instrumentation and Measurement Technology Conference Proceedings, 2016, p. 831-836.

[5] **Tamilselvan P., Wang P.** Failure diagnosis using deep belief learning based health state classification. Reliability Engineering and System Safety, Vol. 115, 2013, p. 124-135.

[6] **Shao H. D., Jiang H., Zhang X., et al.** Rolling bearing fault diagnosis using an optimization deep belief network. Measurement Science and Technology, Vol. 26, Issue 11, 2015, p. 101-118.

[7] **Hinton G. E.** A practical guide to training restricted Boltzmann machines. Momentum, Vol. 9, Issue 1, 2010, p. 1-21.

[8] **Feng Z. P., Liang M., Chu F. L.** Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples. Mechanical Systems and Signal Processing, Vol. 38, Issue 1, 2013, p. 401-421.

[9] **Tang H., Wu Y., Hua G., etc.** Internal leakage fault diagnosis of hydraulic cylinder using PCA and BP network. Journal of Central South University of Science and Technology, Vol. 42, Issue 12, 2011, p. 3709-3714.

[10] **Van T. T., Althobiani F., Ball A.** An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks. Expert Systems with Applications, Vol. 41, Issue 9, 2014, p. 4113-4122.

[11] **Goharrizi A. Y., Sepehri N.** A Wavelet-based approach to internal seal damage diagnosis in hydraulic actuators. IEEE Transactions on Industrial Electronics, Vol. 57, Issue 5, 2010, p. 1755-1763.

[12] **Lei Y. G., Lin J., He Z. J., Zuo M. J.** A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mechanical Systems and Signal Processing, Vol. 35, Issue 2, 2013, p. 108-126.

[13] **Zhao X. M., Patel T. H., Zuo M. J.** Multivariate EMD and full spectrum based condition monitoring for rotating machinery. Mechanical Systems and Signal Processing, Vol. 27, Issue 5, 2012, p. 712-728.

[14] **Jiang H., Chen J., Dong G.** Study on Hankel matrix-based SVD and its application in rolling element bearing fault diagnosis. Mechanical Systems and Signal Processing, Vol. 52, Issue 3, 2015, p. 338-359.

[15] **Hinton G. E.** Where do features come from. Cognitive Science, Vol. 38, Issue 6, 2014, p. 1078-1101.

[16] **Fisher A., Igel C.** Training restricted Boltzmann machines: an introduction. Pattern Recognition, Vol. 47, Issue 1, 2014, p. 25-39.

[17] **Ye Q. W., Wang T. Q.** Optimization modal analysis with PSO based on spectrum segmentation. Chinese Journal of Scientific Instrument, Vol. 30, Issue 8, 2009, p. 1584-1589.

[18] **Wang X. X., Li Y. F., Rui T., et al.** Bearing fault diagnosis method based on Hilbert envelope spectrum and deep belief network. Journal of Vibroengineering, Vol. 17, Issue 3, 2015, p. 1295-1308.

[19] **Krakovsky V. Y.** Moving-window discrete Fourier transform. Journal of Real-Time Image Processing, Vol. 1, Issue 2, 2006, p. 153-161.

[20] **Ye Q. W., Wang T. Q., Ye J. Y.** Simulation study of modal parameter estimation algorithm with pso based on spectrum segmentation. Journal of System Simulation, Vol. 21, Issue 19, 2009, p. 6225-6228.

[21] **Ma J. W., Yang F.** Low frequency oscillation mode recognition based on the sliding window FFT algorithm. Electric Switchgear, Vol. 6, Issue 1, 2014, p. 69-72.

[22] **Gu B., Sheng V. S.** A Robust Regularization Path Algorithm for nu-Support Vector Classification. IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, Issue 5, 2017, p. 1241-1248.

[23] **Wen X. Z., Shao L., Xue Y., et al.** A rapid learning algorithm for vehicle classification. Information Sciences, Vol. 39, Issue 5, 2015, p. 395-406.

Xinqing Wang received Ph.D. degree in mechanical engineering from Tianjin University, Tianjin City, China, in 1998. Now he works as a Professor at PLA University of Science and Technology, China. His current research interests include mechanical electronic engineering, fault diagnosis and signal processing.
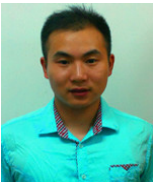
Jie Huang received his Master degree in mechanical engineering from PLA University of Science and Technology, Nanjing City, China, in 2015. Now he is a Ph.D. candidate. His current research interests include deep learning and mechanical system fault diagnosis.

Guoting Ren received his Bachelor degree in Tsinghua University, Beijing in 2015. Now he is a Master candidate. His current research interests include artificial neural networks and multi-line laser radar application.

Dong Wang received Ph.D. degree in vehicle engineering from PLA University of Science and Technology, Nanjing City, China, in 2013. Now he works as a Lecturer at PLA University of Science and Technology, China. His current research interests include fault diagnosis and signal processing.