DISCUSSION ON FAULT DIAGNOSIS OF AND SOLUTION SEEKING FOR ROLLING BEARING BASED ON DEEP LEARNING

Qiming NIU
[1]*School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China;*
[2]*Department of Computer Teaching , Hebei University, Hebei, China.*
Email: niu_qimingcit@163.com

ABSTRACT *In order to realize the intelligent fault diagnosis of automobile gear box by image analysis, the signal that simulates the vibration signals in the process of actual impact fault and mixed fault were created first; secondly, the structure parameters of convolutional neural network in application are analysed, better network parameters are chosen; finally, to identify the fault gear box under variable speed by using convolutional neural network with optimized parameters, the performance of classification and recognition of time-frequency images of the convolutional neural network, deep belief networks, stacked auto-encoder were compared. The results show that the top of the three deep learning algorithms is stacked auto-encoder, followed by convolution neural network and deep belief network. It shows that the convolution neural network can be applied to the classification and identification of time-frequency images of vibration signals and the intelligent diagnosis of equipment under complex conditions.*

KEYWORDS*: pore pressure; fire; self-consolidating concrete; fiber reinforcement; spalling*

1 INTRODUCTION

Gear and rolling bearings are common parts of automotive transmission, which affected by the complex and variable operation conditions of vehicles and working under long time alternating load, are fault prone parts for automobile. Their fault vibration signals tend to exhibit strong non-linear and non-stationary feature (Guo et al., 2016). The time-frequency images obtained by time-frequency analysis represent the joint distribution information in both time domain and frequency domain. It directly reflects the relationship between the frequency components and time. The time frequency graph contains rich information of the equipment state, and the intelligent diagnosis of the gear box failure can be realized through the analysis of the image.

Convolution neural network is one of the algorithms in the field of deep learning. It has good classification performance for image recognition. Convolution neural network is applied to the processing of time-frequency images of vibration signals to achieve fault classification and recognition (Tarolli et al., 2017). First, the fault vibration signal of the gear box is collected, and the signal is transformed into time-frequency image by time-frequency transformation. Then the convolutional neural network is used to classify and identify time-frequency images, following are identifying the state of the transmission.

2 RESEARCH ON TIME FREQUENCY IMAGE RECOGNITION BASED ON DEEP LEARNING

The transmission gearbox with three axles and five speeds is chosen for the test. The transmission gear box has five forward gears and one reverse gear. In the experiment, the transmission is engaged with the fifth gear. The fault gear is the driven wheel in gear five. In order to simulate different types of mix fault simulation, different faults are respectively arranged in the inner ring of the bearing and fifth gear, the bearing is at the end of output shaft, the model is NPU311EN. The bearing fault is set in the inner ring of the bearing, including two fault states, the width of the fault is 0.2mm and 2mm and the depth is 1mm, plus the normal state of the bearing's inner ring, there are three states of bearing.

2.1 Fault diagnosis procedures

The concrete steps of a convolution neural network in the gearbox failure diagnose is shown in Figure 1.

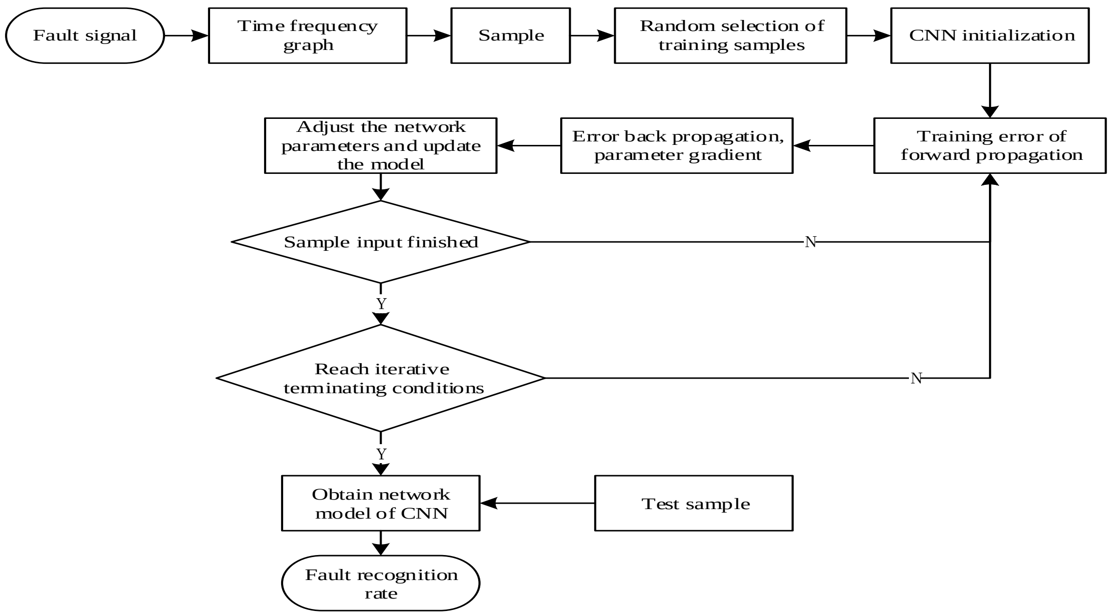

Figure 1. Process of fault diagnosis for gearbox

When the convolution neural network is trained, a large number of samples are needed. Therefore, for every vibration signal collected, each signal is divided into several small segments with a certain length, and then the time-frequency transformation of these small segments is made to get multiple time-frequency maps. The pixels of the time frequency graph are adjusted to the appropriate size to fit into the convolution neural network. A large number of samples are formed with the size-adjusted time frequency graph. A sample of training samples is randomly selected from the time frequency graph corresponding to each kind of signal, and the remaining one is used as the test sample (Tarolli et al., 2017). The training samples are input into the convolution neural network to train the network, and a trained network model will be obtained after the training process is finished. Input test samples to the trained neural network and use soft-max as classifier to get the classification results.

2.2 Time-frequency domain and signal domain analysis

In the actual experiment, 9 kinds of fault state and a normal state are set up, that is to say, 10 states need to be identified. In order to facilitate the analysis of simple time domain, select only the time domain signal of normal state and of three serious fault states for analysis. The four states are: normal, fifth gear teeth broken, fault in the inner ring of the bearing 0.2mm + fifth gear teeth broken, fault in the inner ring of the bearing 2mm + fifth gear teeth broken.

The time domain chart under the condition of the fifth broken gear teeth is compared with the normal state of the gear. In the normal state, the vibration signal has no obvious impact, the vibration amplitude is smaller, and the gear fault not only increases the vibration amplitude, also pushes the impact to a certain extent. In the case of the fifth broken gear teeth, when the fault is the inner bearing ring is 0.2mm wide, the impact phenomenon is more obvious, and the amplitude continue to increase; increase the inner bearing width to 2mm, the vibration amplitude is larger and the impact is more serious.

From the above analysis, we can see that if we only analyze the signal in time domain, we can only observe the magnitude of the vibration amplitude and whether there is any impact phenomenon. For fault diagnosis, this information is far from enough. The time frequency graph can clearly reflect the relationship between the frequency components of the signal and the time. It contains a large amount of information, which is more suitable for fault diagnosis.

In view of a wealth of information contained in the time-frequency map, and in the study of convolutional neural network recognition performance of hybrid transmission failure, the input of convolutional neural network must be two-dimensional shape, therefore, after the acquisition of the vibration signal of the transmission under various working conditions, first needs to transform time-domain signal into time-frequency (Tra et al., 2017). Considering that the signal after short time Fourier transform and continuous wavelet transform, time-frequency presented thereof will be affected by the choice of window function or wavelet function. Therefore, S transform is used in the time-frequency transform when selecting structure parameters on the convolutional neural network.

According to the parameters of the transmission system, we can calculate the rotating frequency, f_r, and fifth gear meshing frequency, f_z, as follows:

$$f_r = n \times \frac{26}{38} \times \frac{42}{22} = \frac{1000}{60} \times \frac{26}{38} \times \frac{42}{22} = 21.77(Hz) \quad (1)$$

$$f_z = f_r z = 21.77 \times 22 = 478.92(Hz) \quad (2)$$

When the gear fault occurs, the amplitude of the vibration increases with the degree of the gear fault. With the increase of the bearing inner ring fault, with the rotation of the bearing, the surface of the fault will impact with the roller, and cause the natural vibration of the bearing system in middle to high frequency, so there will be a relatively obvious high frequency band. When the fault width of the inner ring of the bearing is increased from 0.2mm to 2mm, the fault degree of the inner ring is deepened, and the amplitude of the vibration becomes larger. According to the difference of the type of fault and the degree of failure, the time frequency diagram of all kinds of faults is different. However, it is obviously difficult to distinguish all kinds of faults by experience, especially to distinguish between these 10 kinds of fault states. At this time, it is necessary to have a method to identify the common characteristics of the time frequency diagram of the same type of fault, and at the same time to distinguish the different types of faults. In view of the excellent performance of convolutional neural network in image recognition, such as face recognition and handwritten digit recognition, convolution neural network is selected to carry out the fault image recognition of transmission.

3 ANALYSIS OF EXPERIMENTAL RESULTS FOR FAULT IDENTIFICATION OF GEARBOX UNDER VARIABLE SPEED

In the course of automobile operation, the speed of the engine transferred to the gearbox is not constant in most cases. The speed of the input shaft of the gearbox varies with time, and the vibration is more complex than the input axis is at a stable speed. Convolution neural network is characterized by its unique structure and invariance to translation, scaling and torsion. Therefore, this chapter adopts the combination of time-frequency analysis and convolutional neural network to diagnose gearbox faults under variable speed.

3.1 Experiment arrangement

The experiment is carried out in transmission with three axles and multi speed; experimental equipments are consistent with these in the last chapter. The driven wheel in fifth the gear is set as fault. By way of cutting different degrees of gears, to simulate gear wear: mild and moderate wear and broken teeth. The bearing fault is located in the inner ring of the rolling bearing of the output shaft, which is a fault of 0.2mm in width. In the failure state and normal state of the combined gear and bearing, a total of eight states of the signal need to be identified, as shown in the following table 1.

Table 1. Fault description

Number of fault type	Fault type	Number of fault type	Fault type
1	Fifth gear normal	5	Inner ring 0.2mm+fifth gear normal
2	Fifth gear mild wear	6	Inner ring 0.2mm+ fifth gear mild wear
3	Fifth gear moderate wear	7	Inner ring 0.2mm+ fifth gear moderate wear
4	Fifth gear teeth broken	8	Inner ring 0.2mm+ fifth gear teeth broken

Signal analysis under rising speed

The variable speed is the speed of the input axis changes with time, including three kinds of conditions: acceleration, deceleration and acceleration and deceleration. The procedures of acceleration and deceleration are basically the same, so only the speed of the lift is chosen and analyzed in the paper. The acceleration means that the speed of the input shaft of the gearbox increases with time.

The sampling frequency is set to 12 kHz, and the vibration signals of each fault state and normal state are collected. In order to show the overall trend of signal with the change of time, the time domain signal with a fault length of 60s in each type is selected. At the rising speed, the speed of the input shaft gradually increases, and the amplitude of the vibration of the signal increases gradually. The degree of vibration can also be affected by the degree of failure, the greater the degree of failure, the stronger the vibration, and the greater the amplitude of the vibration. Since the input shaft in the experiment is set as rising overall, it cannot guarantee that every kind of fault state at the same time will have the same speed. A fixed time in the eight states time domain diagram, the amplitude of vibration fault degree may be lower than the amplitude of small fault (Sobie et al., 2018). The time domain signal can only reflect the amplitude of the vibration signal and the occurrence of the impact phenomenon, and the time domain signal cannot be relied upon on the identification of fault state.

The rotation frequency of each axis varies under variable speed, and the meshing frequency of the gear that participates in the transmission also changes, and the corresponding frequency component is an interval. The formula of the meshing frequency of constant mesh gear, f $_{m1,}$ and five output gears, $f_{m2,}$ are as follows:

$$f_{m1} = z_1 \times f_{n1} \quad (3)$$

$$f_{m2} = z_4 \times f_{n3} \quad (4)$$

Among them, z_1 is the number of driving gear tooth on the most meshing gear, z1=26; f_{n1} is the input shaft frequency; z_4 is the number of the driven gear, z_4 =22; f_{n3} is the output frequency.

When the speed of the input axis increases slowly from 0 to 1500rpm, the conversion of the input shaft, $f_{n1,}$ is gradually increased from 0 to 25Hz. The maximum value of the output frequencies of constant mesh gear and fifth gear, f_{m1max} and f_{m2max} are calculated as follows:

$$f_{m1max} = z_1 \times f_{n1max} = 26 \times 25 = 650(Hz) \quad (5)$$

$$f_{m2max} = z_4 \times f_{n3max} = 22 \times f_{n1max} \times \frac{26}{38} \times \frac{42}{22} = 718.42(Hz) \quad (6)$$

The meshing frequency range of a constant meshing gear is 0~650Hz, fifth gear meshing frequency range is 0~718.42Hz. If the two in the transmission gear have a second vibration response, the range of meshing frequency range of a constant meshing gear and fifth gear are respectively into 0~1300Hz and 0~1436.84Hz.

3.2 Fault diagnosis under the condition of rising speed

When identifying the time-frequency images of different fault signals, the samples of each kind of signals are fixed in a fixed length, and the amplitude matrix corresponds to the time-frequency diagram in a short time interval. Under the condition of constant speed, the samples of the same kind of signal should be basically the same without considering the influence of noise and random factors (Janssens et al., 2016). Under the condition of variable speed, the value of gear meshing frequency, the characteristic frequency of bearing fault, the frequency of impact and the amplitude of each frequency component change with time, but in general, they are similar. Since convolutional neural network is invariant to translation, scaling and torsion in image recognition, convolution neural network is used to identify time-frequency images under variable speed (Deutsch et al., 2017).

From time zero to t_1, input shaft speed is increased from 0 to n_1, draw a demarcation point at t_0 (t_0 is a time point between time zero to t_1), and the time-frequency map with signal corresponding with time ($0\sim t_0$) is set as the training samples, a total of 500 training samples is thus created; while time-frequency map corresponding to the signal from $t_0\sim t_1$ set as test samples, which is a total of 500 test samples (Li et al., 2017). The time frequency conversion method is used to adjust the size of the time frequency graph to be 32 * 32, so that it can be used as the input of the convolution neural network. In the image recognition, there are eight kinds of time frequency graphs in different states. Each class has 1000 samples, and the training sample and the test sample each account for 50%.

Three deep learning algorithms represented by convolutional neural network, deep belief network and stacked auto-encoder are used to recognize time-frequency images under acceleration.

The convolutional neural network parameters are set as: 6 convolution kernels in the first layer, the size of which is 5 x 5, and 12 convolution kernels in the second layer, size 3 * 3; cut pooling area of sample layer to the size of 2 * 2, with the method average pooling; batch size=5; iteration 10 times.

The parameters selection of deep belief network is as follows: the number of nodes in input layer is 1024, the number of nodes in output layer is 8, and the number of nodes in three hidden layers is: 1000, 800 and 500. The pre-training stage with unlabeled data were stratified training for each RBM, 100 iterations; fine-tuning stage, classification model combining a pre-trained neural network and softmax classifier, use tag data through the back-propagation algorithm to fine tune the entire network, 50 iterations.

The parameters of stacked auto-encoder are as follows: the number of nodes in input layer is 1024, the number of nodes in output layer is 8, and the number of nodes in three hidden layers is: 400, 200 and 50, respectively. The pre-training phase is iterated 100 times. Due to the relationship between computing time and the operation structure of the algorithm itself, the number of times of reverse fine tuning is selected to calculate the correct rate of classification once every 20 times until the iteration reaches 200 times.

Each experiment was repeated 5 times, and the average value of the 5 classification results was taken as the final classification accuracy rate. Then the relationship between classification accuracy rate of the three classes of deep learning algorithm and the number of iterations can be shown in figure 2-4.

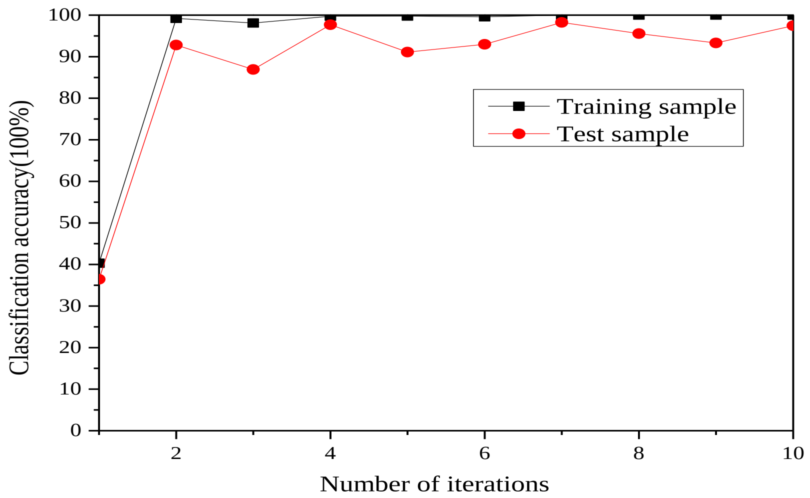

Figure 2. Classification accuracy of convolution neural network

Figure 2 is the training and test classification accuracy of the convolution neural network when the time frequency image is identified at rising speed. When the number of iterations increased from 1 to 2, the classification accuracy of the training samples and the test samples increased significantly. After 4 iterations, the training accuracy rate is over 99%. After 6 iterations, the training accuracy reaches 99.9%, indicating that the fitting effect of convolutional neural network on training samples is very good. Increasing the number of iterations is no longer meaningful(Sohaib et al., 2017). For test samples, after 4 iterations, the correct rate of test is over 90%, but the accuracy does not increase as the number of iterations increases, but fluctuates slightly. When the iteration is 7 times, the correct rate reaches the maximum value of 98%.

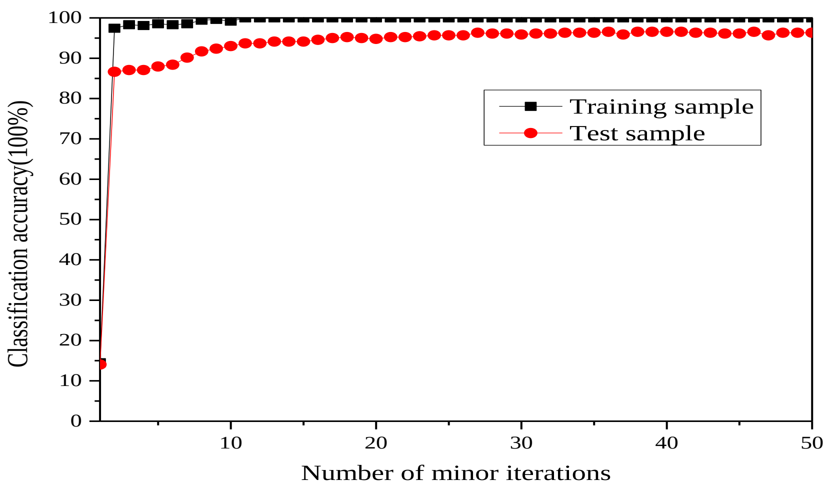

Figure 3. Classification accuracy of deep belief network

The depth belief network underwent two processes— pre-training and fine tuning, and figure 3 is the change curve of the correct rate of both training and testing samples as the number of iterations and fine-tuning increases. As can be seen from the figure, when the number of iterations is increased from 1 to 2, the correct rate of training is raised from 15.01% to 97.7%, and the correct rate of test went up from 13.98% to 86.78%. With the increase of the number of fine-tuning and iterations, the classification accuracy of training and testing continues to increase slightly. When the number of iterations reached 12, the correct rate of training reached nearly 100%, and the correct rate was 94.2% at this time. After 27 iterations, the test accuracy is maintained over 96%, and the maximum value is reached at 40 times, which is 97.11%.

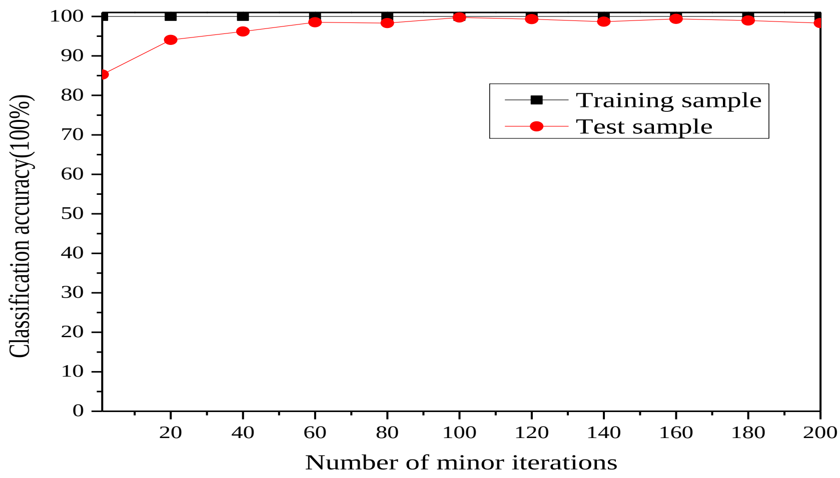

Figure 4. Classification accuracy of stacked auto-encoder

The stacked auto-encoder is also pre-trained and fine-tuned, and Figure 4 is the change curve for the training and testing correct rate with the number of iteration and fine-tuning. Because all the training samples are input at one time, the fitting of the network is for all training samples, so after pre-training, the correct rate of training is 100%. The value of the transverse coordinates at 0 is the correct rate of classification without a fine tune. At this time, the pre-training of 100 iterations has reached, and the correct rate is 84.52%. After 60 times of fine tuning, the accuracy of the test is maintained over 98%, and the maximum value is reached at 120 times, which is 99.53%.

Generally speaking, when the number of iterations or the number of fine-tuning iterations reaches a certain degree, the training accuracy of the three deep learning algorithms can basically reach 100%, indicating that the three algorithms can basically match the training samples. Moreover, the accuracy rates of the highest test classification of the three algorithms have all reached over 96%, which indicates that these three kinds of deep learning algorithms can better classify and identify time frequency diagram in transmission failure under the condition of acceleration (Gan et al., 2016).

The maximum test accuracy of the three algorithms is compared. The highest accuracy is stacked auto-encoder, at 99.43%, followed by convolution neural network, 98%, and the maximal accuracy of deep belief network, 96.71%. In terms of the stability of the method, the stability of the deep belief network and the stacked auto-encoder are better, while the convolution neural network is slightly more fluctuant. Considering the computation time needed for the three algorithms to achieve the maximum accuracy, the time for single operation of convolutional neural network is 88.9s, which is about 1/5 of that of deep belief network, and 1/11 that of stacked auto-encoder.

From the above analysis we can know that the convolution neural network can effectively identify time-frequency image of gearbox fault under acceleration, although the test accuracy rate fluctuates, not as stable as belief network and stacked auto-encoder, but the time cost is much lower than the other two algorithms.

The shallow machine learning algorithm supported vector machine is also used to identify the time frequency image, the parameters of support vector machine are: the use of algorithm in libsvm (Deutsch et al., 2017). first adjust the format of the input data, with radial basis function as the kernel function, through cross validation and grid search select for support vector machine penalty factor C and kernel function parameter g. The optimal parameters are trained on the whole training set, creating support vector machine model (Abid et al., 2017). Table 2 shows the highest test and training classification accuracy of support vector machine and three deep learning algorithms.

Table 2. The highest training and accuracy rate of the four algorithms

	Algorithm type	Highest training accuracy rate (100%)	Highest test accuracy rate (100%)
	Convolutional Neural Network	100	98
	Deep Belief Network	100	97.11
	Stacked Auto-Encoder	100	99.53
	Support Vector Machine	98.975	45.025

It can be seen from table 2 that the support vector machine with optimized parameters can recognize time frequency images with the highest training classification accuracy rate as 98.975%, which indicates that the training process is successful. While the highest test accuracy is only 45.025%, far below the three deep learning algorithms, not only indicates that deep learning algorithm is better overall than the shallow algorithm supported vector machine, also that the support vector machine is not suitable for time-frequency image recognition against speed rise.

4 CONCLUSION

This paper takes the vehicle transmission as the research object, and applies the convolution neural network, which is a kind of deep learning, to the fault identification in automobile transmission. The time-frequency analysis and convolution neural network are combined to diagnose the gearbox fault under variable speed. The continuous wavelet transform is used to get the time-frequency vibration signals under variable speed conditions, and then the convolutional neural network is used to identify these images. The results show that under the rising speed, the test accuracy reached 98%, while under lifting speed conditions, the highest correct rate in the test reached 96%, revealing that under complex condition, convolutional neural network can be effectively applied to classification recognition and intelligent diagnosis of vibration signal time-frequency image. Comparison of the performance in time-frequency image recognition variable speed conditions, using convolutional neural network under deep belief networks and stacked auto-encoder was made. The results show that the convolution neural network test accuracy is slightly lower than the stacked auto-encoder, and the classification results slightly fluctuateds, but the computation time of convolutional neural network is much lower than the other two deep learning algorithms.

5 REFERENCES

►Guo, X., Chen, L., Shen, C., (2016). Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. Measurement, 93, 490-502.

►Tarolli, J. G., Naes, B. E., Butler, L., et al (2017). Fully convolutional neural network for removing background in noisy images of uranium bearing particles. Analyst, 142(9), 1499.

►Tarolli, J. G., Naes, B. E., Willingham, D., et al (2017). Combining a Convolutional Neural Network and Watershed Segmentation for Identifying U-Bearing Particles in Secondary Ion Mass Spectrometry Images. Microscopy & Microanalysis, 23(S1), 118-119.

►Tra, V., Kim, J., Khan, S. A., et al (2017). Bearing Fault Diagnosis under Variable Speed Using Convolutional Neural Networks and the Stochastic Diagonal Levenberg-Marquardt Algorithm. Sensors, 17(12), 2834.

►Sobie, C., Freitas, C., Nicolai, M., (2018). Simulation-driven machine learning: Bearing fault classification. Mechanical Systems & Signal Processing, 99, 403-419.

►Janssens, O., Slavkovikj, V., Vervisch, B., et al (2016). Convolutional Neural Network Based Fault Detection for Rotating Machinery. Journal of Sound & Vibration, 377, 331-345.

►Deutsch, J., He, M., He, D., (2017). Remaining Useful Life Prediction of Hybrid Ceramic Bearings Using an Integrated Deep Learning and Particle Filter Approach. Applied Sciences, 7(7), 649.

►Li, S., Liu, G., Tang, X., et al (2017). An Ensemble Deep Convolutional Neural Network Model with Improved D-S Evidence Fusion for Bearing Fault Diagnosis. Sensors, 17(8), 1729.

►Sohaib, M., Kim, C. H., Kim, J. M., (2017). A Hybrid Feature Model and Deep-Learning-Based Bearing Fault Diagnosis. Sensors, 17(12), 2876.

►Gan, M., Wang, C., Zhu, C., (2016). Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings. Mechanical Systems & Signal Processing, 73(2), 92-104.

►Deutsch, J., He, D., (2017). Using Deep Learning-Based Approach to Predict Remaining Useful Life of Rotating Components. IEEE Transactions on Systems Man & Cybernetics Systems, 99, 1-10.

►Abid, A., Khan, M. T., Khan, M. S., (2017). Multidomain Features-Based GA Optimized Artificial Immune System for Bearing Fault Detection. IEEE Transactions on Systems Man & Cybernetics Systems, 99, 1-12.