Bearing Diagnosis Accuracy Comparison Using Convolutional Neural Network with Time/Frequency Domain Signals

Da He[1], Wei Guo[1,2], Mao He[1]
1. School of Mechanical and Electronical Engineering,
University of Electronic Science and Technology of China (UESTC), Chengdu, China
2. Institute of Electronic and Information Engineering of UESTC in Guangdong, Guangdong, China
Email: gwuestc2013@163.com

***Abstract*—Deep learning is the most attractive topic in the field of machine learning and relevant applications. Owing to the strong learning ability of the convolutional neural network (CNN), it integrates the feature extraction from raw data and classification as a complete learning process and makes the bearing fault diagnosis intelligent. In the published results, the inputs of the CNN may be the raw temporal waveform of vibration, its processed waveform or converted 2D images. In this paper, focusing on the diagnosis accuracy of rolling bearings, a comparative study is conducted among the inputs using the raw temporal waveform, the frequency spectrum, and the envelope spectrum of a vibration signal. First, an appropriate classification model based on the CNN is constructed. Then, experimental data from bearing with real damages are collected and then transformed and converted into some small gray pixel images for training and testing the CNN model. Finally, the classification accuracies using three signals are compared. The results indicate that the diagnosis performances using the above three signals are close when the trained CNN models are stable; among them the model using the frequency spectrum of the vibration signal is a little better than the models using the other two signals, which may be a reference for further investigating the deep learning used in the field of bearing diagnosis.**

Keywords: Convolutional Neural Network; Fast Fourier Transform; Envelope Analysis; Fault Diagnosis; Rolling Bearing.

I. INTRODUCTION

Thanks to the change on the processing units, especially Graphics Processing Units (GPUs) with the increase of the computational ability and the decrease of the cost, the deep learning uses its deep architectures and efficient training to finish the abstraction from raw data, and thus has been a promising tool in the field of machine learning [1]. Some popular deep learning architectures, including the deep belief network, stacked autoencoder, deep Boltzmann machine, recurrent neural network, convolutional neural network (CNN) [2], and so on, are reported to be success on the performance improvement and extension on the range of applications, such as image recognition, natural language processing, and processing of big data. Some scholars also applied the deep learning to the fault classification of rotating machinery.

In bearing fault diagnosis, the deep learning mainly uses its deep architectures with many layers of processing units to directly extract complex features from raw data, which is quite different from the conventional feature selection in multi-domain and its reduction. In this paper, we focus on the improvement and applications of the CNN for bearing vibration analyses. Ince *et al*. [3] used 1D convolutional neural networks to fuse the feature extraction and classification phases into a single learning body and improved the efficiency of real-time motor condition monitoring. Janssens *et al*. [4] used the CNN to construct a feature learning model to autonomously learn useful features from data with various faults, which makes non-experts vibration analysis feasible.

Zhang *et al*. [5] directly chose raw data obtained under noisy and variable working load as inputs of the CNN and their intelligent diagnosis method achieve higher accuracy without denoising preprocessing. Yang *et al*. [6] used multisource raw vibration data as input of the CNN for feature extraction and classification. Shao *et al*. [7] used the compressed sensing before the improved convolutional deep belief network to reduce the vibration data amount for the network; then introduced Gaussian visible units to enhance the learning ability of the network. Pan *et al*. [8] proposed a deep learning network (LiftingNet) that involved more layers, i.e. the split layer, predict layer, and update layer, to classify the raw data containing large noise and randomness, in which the inputs of the network were time-domain vibration signal preprocessed by resampling technique.

Wang *et al*. [9] converted vibration signals from multi-sensor to images and the corresponding feature maps are used to train the CNN combining with the bottleneck layer optimization and increased features for bearing fault recognition. Wen *et al*. [10] used the CNN based on LeNet-5 to extract features from raw data automatically, in which the raw vibration signal was separated into small segments by sequence and then converted into some gray pixel images. Li *et al*. [11] used S-transform algorithm to automatically convert raw data

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

to 2D time-frequency matrix and its coefficients was inputted to the CNN for fault classification.

Jia *et al*. [12] proposed the deep normalized convolutional neural network to improve the classification of imbalanced bearing data and pointed out that kernels works as filters in the CNN and are complex with the increase of the network depth. Abdeljaber *et al*. [13] utilized compact 1D convolutional neural networks to identify, quantify, and localize bearing damage, which can be further used to handle new damage scenarios with utmost accuracy. Guo *et al*. [14] proposed the deep convolutional transfer learning network to learn complex features in raw data and then recognize the health states of machines, so that the limited labeled data and unlabeled data from other machines can be classified.

Zhao *et al*. [15] summarized the relevant work of the deep learning on machine health monitoring and compared some popular deep learning networks on the performances of fault classification. Hoang and Kang [1] also reviewed the study using the deep learning and its applications on bearing fault diagnosis.

According to the above literature review, the inputs of the CNN mainly come from raw vibration signals from single or multi-sensors, or their corresponding preprocessed signals or converted images. The preprocessing and converting of temporal waveforms aim to provide rich information about the tested machine components and help to improve the classification accuracy. The signals in time-domain and its frequency spectra in frequency-domain also play important roles in the signal analysis of bearing fault diagnosis. Inspired by these two ideas, the time/frequency domain signals are chosen as inputs of the LeNet-5 CNN model for bearing fault diagnosis. Without the time-consuming preprocessing, the raw temporal waveform of a bearing, its frequency spectrum obtained by using the fast Fourier transform, and its envelope spectrum obtained by using the envelope analysis are individually converted by using a signal-to-image conversion method [10]. As a result, a large amount of small gray pixel images forms plentiful training samples to improve the learning ability of the CNN. Experimental vibration data from real bearing damages are applied to the CNN model to compare their performances on the diagnosis accuracy.

The remaining of this paper is organized as follows. Section II briefly introduces relevant theory about the CNN and the LeNet-5 model used in this paper. Section III introduces the experimental data and their analysis results, and then compares the classification results obtained by using three source data. Finally, Section IV draws the conclusions in this paper.

II. FAULT CLASSIFICATION BASED ON CNN

In this section, brief introduction on CNN and the CNN model used in the paper are briefly introduced.

A. Brief Introduction of CNN [1]

The convolutional neural network is a multi-layer structure and consists of several convolutional layers, pooling layers and fully connected layers. The function of convolutional layers and pooling layers corresponds to the feature extraction and reduction, and the fully connected layers are used for classification.

1) Convolutional layer

The convolutional layer convolves the input data with filter kernels, then output the feature maps with activation unit, which now often is Rectified Linear Unit (ReLU). Each filter uses the same kernel which means the weight is shared, and thus the local feature of the input local region is extracted. One filter corresponds to one frame in the next layer, and the number of frames also represents the depth of this layer. The convolutional process in each kernel can be described:

$$y_i^{l+1}(j) = \boldsymbol{K}_i^l * \boldsymbol{X}^l(j) + b_i^l \tag{1}$$

where $\mathbf{K}_i^l$ and b_i^l represent the weight matrix and bias of i-th kernel in the l-th layer, respectively; $\mathbf{X}^l(j)$ denotes the j-th local region matrix in the l-th layer, and $y_i^{l+1}(j)$ represents the output of the l-th layer and the input of the (l+1)-th layer.

After the convolutional process described above, the ReLU is used as activation function to intensify the convergence of the model.

2) Pooling layer

After a convolutional layer, the pooling layer is arranged to conduct the down-sampling to merges features of a matrix into one, and thus reduces the sampling rate of the whole architecture. The most commonly used pooling layer is the max-pooling layer, which extracts the maximum data of the input matrix and is represented as:

$$\boldsymbol{P}_i^{l+1}(j) = \max_{\substack{(j-1)W+1<t_x\le jW \\ (j-1)H+1<t_y\le jH}} \left\{ \boldsymbol{Q}_i^l(t) \right\} \tag{2}$$

where W and H represent the width and height of the input matrix that is going to be pooled, respectively; $Q_i^l(t)$ denotes the t-th neuron in the i-th frame of layer l; and $P_i^{l+1}(j)$ represents the output of the pooling operation.

3) Classification stage

This stage is composed of two fully connected layers to realize the classification. Based on this, the Softmax function is used to accord the output logits of three neurons with the probability distribution of three conditions, and is described as:

$$f(z_j) = \frac{e^{z_j}}{\sum_{k=1}^{3} e^{z_k}} \tag{3}$$

where z_j denote the input logits of the j-th neuron of the output layer.

B. LeNet-5 architecture of the CNN

The classic LeNet-5 proposed by LeCun in1998 [2] is a classical release of CNN and successfully used for handwritten and machine-printed character recognition, whose property is its minimal data preprocessing. In this paper, it is used as a fault classifier and finally give the result of bearing fault diagnosis. The normal LeNet-5 architecture consists of 7 layers. According to available experimental data to be analyzed in the next section, the parameters and structure are set as follows:

1) Layer 1 – the first convolutional layer. The input of this layer is an image of 32x32 pixels and uses convolutional kernels of 5x5x64 size and the stride is 1x1. The padding method is same padding, which means when coming to the edge of images, convolutional kernels automatically fill zeros in the empty positions. Besides, the activation function used in this layer is ReLU. The output of layer one is a 32x32 feature map.

2) Layer 2 – the first pooling layer. It takes the feature map from layer one as input and uses the 2x2 max-pooling strategy. The output of this layer is a feature map of 16x16 size.

3) Layer 3 – the second convolutional layer. Here, the feature map from layer two is convoluted with convolutional kernels of 5x5x128 size with the identical stride with layer one. Its output is a 16x16 feature map.

4) Layer 4 – the second pooling layer. It uses the same parameters as layer two does, and its output is an 8x8 feature map.

5) Layer 5 – one full connection layer. It is a full connection layer, and its input is of 8x8x128 and the output is of 1024 size.

6) Layer 6 – dropout layer. Data are randomly dropped out in a certain probability in this layer in order to relieve the tendency of overfitting.

7) Layer 7 – Softmax layer. The Softmax function is used as a classifier and output the final classification result.

As for other parameters in this model, the cross entropy is to evaluate the loss, the learning rate is set to a small value of 0.00001, and Adam is chosen for the optimizer.

III. Experiments and Results

A. Description on Experimental Data

To train the LeNet-5 model, a large amount of vibration signal data of bearing tests are needed. Lessmeier *et al*. [17] conducted two types of experiments on rolling bearings, i.e. artificial damage and real damage, the latter of which were generated during the accelerated life time tests. The used apparatus is shown in Figure 1. Vibration data were collected from bearings with various damage extents and working conditions. The specification on the used experimental data are summarized in Table I. The state descriptions on three data sets are shown in Table II.

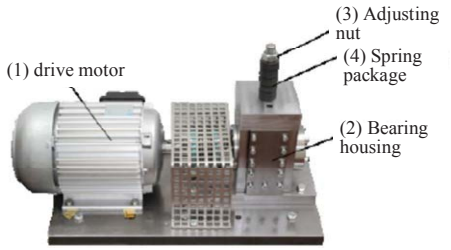

Figure 1. Apparatus for accelerated life time test [17].

TABLE I. Specification on The Experiments

Operation parameter	Value
Rotational Speed [rpm]	900
Load Torque [Nm]	0.7
Radial Force [N]	1000
Sampling Frequency [Hz]	64000
Time Span [s]	4.0

TABLE II. Test Bearings With Real Damages Caused by Accelerated Lifetime Test

Bearing code	K004	KA04	KI04
State/Damage	Health	Fatigue; Pitting	Fatigue; Pitting
Bearing element	\	Outer race (OR)	Inner race (IR)
Extent of damage	\	1	1
Characteristic of damage	\	Single point	Single point

Three experimental data sets, including K004, KA04, and KI04, are chosen for the diagnosis comparison of the CNN, which correspond to the healthy bearing, the bearing with the outer race (OR) defect, and the bearing with the inner race (IR) defect, respectively. During the tests, the tested bearing (type 6203) ran at 900 rpm with a load torque of 0.7 Nm and a radial force on the bearing of 1000 N applied by a spring-screw mechanism. The sampling frequency is set to 64kHz. The time span of each data set is 4 seconds.

B. Experimental Time/Frequency Domain Signals

Parts of vibration data sets are shown in Figures 2, 3, and 4. Each experiment collected 4 seconds of data. In order to show the impulses in vibration signals, only 0.1s of 4s data are shown here. The temporal waveform in Figure 2(a) was collected from a healthy bearing (code K004). Its frequency spectrum and envelope spectrum in the range of 0-10kHz are shown in Figures 2(b) and 2(c), respectively. Similarly, for the bearing with outer race defect (code KA04), one temporal waveform, its frequency spectrum and envelope spectrum in 0-10kHz are shown in Figures 3(a), 3(b), and 3(c), respectively. The time-domain and frequency-domain signals for the bearing with the inner race defect KI04 are shown in Figure 4.

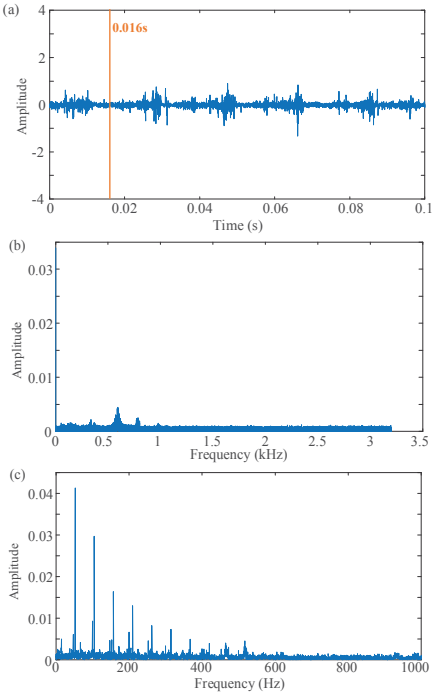

Figure 2. Experimental Data K004: the temporal waveform in (a), its frequency spectrum in (b), and its envelop spectrum in 0-10kHz in (c).

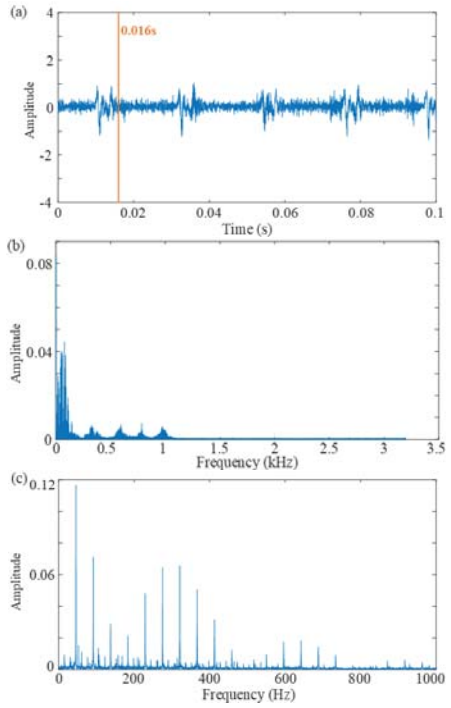

Figure 3. Experimental Data KA04: the temporal waveform in (a), its frequency spectrum in (b), and its envelop spectrum in 0-10kHz in (c).

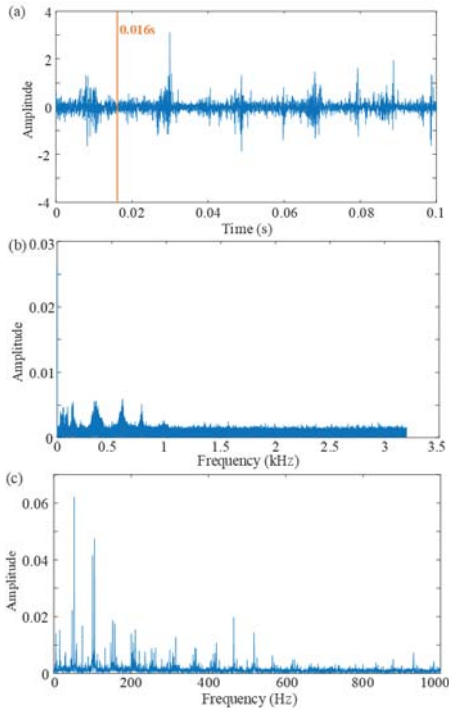

Figure 4. Experimental Data KI04: the temporal waveform in (a), its frequency spectrum in (b), and its envelop spectrum in 0-10kHz in (c).

Comparing the temporal waveforms of three vibration signals in Figures 2-4, the first 0.016s of 0.1s signal contains one set of impulses in the vibration signal and thus is chosen as one sample unit. Therefore, for one signal with 4 second, the temporal waveform was separated into 250 segments (4/0.016), and each segment includes 1024 sample points (0.016×64000). Using the signal-to-image conversion method [10], one gray pixel image with the size of 32x32 can be obtained. For 4s of one data set, 250 images are then available for learning and testing.

During the accelerated life time tests, keeping the same working condition for each bearing code, 20 data sets were collected. Accordingly, for each bearing code, there are 5000 images that are obtained from the signal in time domain and used as inputs of the CNN, in which the first 4750 images are used for training the network, and the remaining 250 images are used to test its generalization performance. In similar way, if using the frequency or envelope spectrum as inputs, there are also 4750 images for training and 250 images for testing.

Some images from the bearing code KA04 are shown in the following three figures. Figure 5 shows two gray pixel images of the raw signal in 0 - 0.016s and in 2.0s - 2.016s. Without any preprocessing on the raw signal, these two images are quite different though they have the same time length. Figures 6 and 7 show the gray pixel images for the same time span by applying the FFT and envelope analysis, respectively.

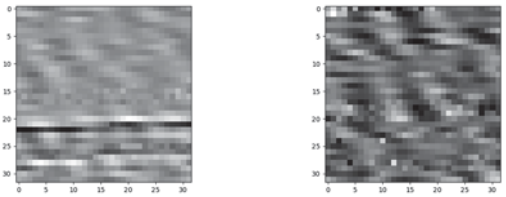

Figure 5. Gray pixel images of time-domain signal from bearing code KA04: the images corresponding to 0s-0.016s (Left) and 2.0s-2.016s (Right).

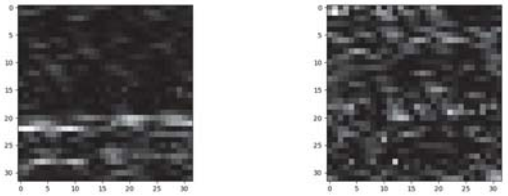

Figure 6. Gray pixel images of the frequency spectrum from bearing code KA04: the images corresponding to 0s-0.016s (Left) and 2.0s-2.016s (Right).

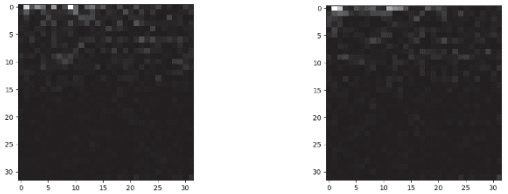

Figure 7. Gray pixel images of the envelope spectrum from bearing code KA04: the images corresponding to 0s-0.016s (Left) and 2.0s-2.016s (Right).

C. *Results and Comparisons*

Using the above images converted from the time-domain or frequency-domain signal, the CNN model is individually trained to learn the features in the signals. The corresponding training accuracies for 1-1500 iterations are shown in Figure 8. In this figure, the red, green and blue curves correspond to the results obtained by using the original signal (raw temporal waveform), the frequency spectrum by using the FFT, and the envelope spectrum, respectively. The classification accuracy is set as the ratio of the number of correctly classified samples to the total number of testing samples.

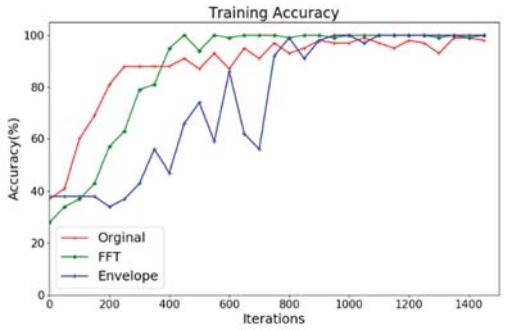

Figure 8. Training accuracies of the CNN model with three time-domain and frequency-domain signals.

According to Figure 8, some observations are listed here.

1) No matter which signal is used, after 1500 iterations in the training process, the classification accuracy is close to 100%.

2) Using the raw signal, the accuracy first reaches to a higher value at 200 iterations, while the results obtained using the frequency-domain signals are still relatively low. However, after that, the accuracy using raw signal still fluctuates.

3) Using the frequency or envelope spectrum, the accuracies almost reach to 100% after 1000 iterations. Comparing these results, the former shows better performance in the training stage, and almost has 100% accuracy after 500 iterations. The latter shows a fluctuation in the range of 500-1000 iterations.

Based on the above observations, details on the classification accuracy after 1500, 500, and 1000 iterations are shown in the following tables. Table III shows the final classification accuracy of three signals after 1500 iterations. Details on the classification after 500 and 1000 iterations are shown in Tables IV - IX.

TABLE III. CLASSIFICATION ACCURACIES OF CNN MODELS TRAINED BY THREE SIGNALS (1500 ITERATIONS)

Sources of training data	**Classification accuracy**
Raw vibration signals	96.4%
Frequency spectra of raw signals	98.9%
Envelope spectra of raw signals	97.9%

Based on these tables, some observations are summarized as follows:

1) As shown in Table III, when the accuracy in the training process is close to 100% after 1500 iterations, it indicates that the learning of the CNN is almost finished, and the corresponding trained CNN models have similar generalization ability. Applying to 250 testing images, the final classification accuracies for three signals are all larger than 96%.

TABLE IV. CONFUSION MATRIX OF CNN TRAINED BY RAW VIBRATION SIGNALS (500 ITERATIONS)

		Classification		
		Health	OR defect	IR defect
Real class	Health	192 76.8%	0 0.0%	58 23.2%
	OR defect	1 0.4%	240 96.0%	9 3.6%
	IR defect	28 11.2%	4 1.6%	218 87.2%

Note: OR – outer race; IR – inner race.

TABLE V. CONFUSION MATRIX OF CNN TRAINED BY RAW VIBRATION SIGNALS (1000 ITERATIONS)

		Classification		
		Health	OR defect	IR defect
Real class	Health	227 90.8%	0 0.0%	23 9.2%
	OR defect	0 0.0%	248 99.2%	2 0.8%
	IR defect	19 7.6%	2 0.8%	229 91.6%

TABLE VI. CONFUSION MATRIX OF CNN TRAINED BY FREQUENCY SPECTRA (500 ITERATIONS)

		Classification		
		Health	OR defect	IR defect
Real class	Health	192 73.0%	0 8.8%	58 18.2%
	OR defect	1 13.4%	240 74.1%	9 12.5%
	IR defect	28 21.9%	4 9.0%	218 69.1%

TABLE VII. CONFUSION MATRIX OF CNN TRAINED BY FREQUENCY SPECTRA (1000 ITERATIONS)

		Classification		
		Health	OR defect	IR defect
Real class	Health	232 92.8%	4 1.6%	14 5.6%
	OR defect	1 0.4%	246 98.4%	3 1.2%
	IR defect	4 1.6%	2 0.8%	244 97.6%

TABLE VIII. CONFUSION MATRIX OF CNN TRAINED BY ENVELOPE SPECTRA (500 ITERATIONS)

		Classification		
		Health	OR defect	IR defect
Real class	Health	128 51.2%	54 21.6%	68 27.2%
	OR defect	90 36.0%	93 37.2%	67 26.8%
	IR defect	106 42.4%	61 24.4%	83 33.2%

TABLE IX. CONFUSION MATRIX OF CNN TRAINED BY ENVELOPE SPECTRA (1000 ITERATIONS)

		Classification		
		Health	OR defect	IR defect
Real class	Health	187 74.8%	14 5.6%	49 19.6%
	OR defect	3 1.2%	218 87.2%	29 11.6%
	IR defect	15 6.0%	19 7.6%	216 86.4%

2) Confusion matrix of the classification using the CNN model with the time or frequency domain signal shows the details about the classification results for each bearing state. Table IV is chosen as one example to explain the meanings of the classification results. In this table, the first row represents the identified labels obtained by using the CNN model, and the first

column represents real labels of testing samples. The values in the second row indicate that 192 of 250 (192+58=250) testing samples from a healthy bearing are correctly classified and its accuracy is then 76.8%; 58 of 250 samples are wrongly classified into the bearing with IR defect, and its ratio is 23.2%. For the third row, 240 of 250 (96.0%) testing samples are correctly classified as the bearing with OR defect. For the last row, 218 of 250 (87.2%) testing samples from the bearing with IR defect are correctly classified. 28 and 4 samples in 250 testing samples with real label 'IR defect' are wrongly identified as Health and OR defect, respectively.

3) Comparing the diagonal elements in Tables IV and V, the classification accuracies for health, OR defect, and IR defect after 1000 iterations are much higher than those after 500 iterations, the former of which are 76.8%, 96%, and 87.2%, respectively, and the latter of which are 90.8%, 99.2%, and 91.6%, respectively. Similar conclusions can be drawn by comparing Tables VI and VII, or comparing Tables VIII and IX. Therefore, no matter which signal is used, the classification results using the model after 1000 iterations are better than those using the model after 500 iterations. It is because that the trained model after 500 iterations is still not stable and the learning process is not finished.

4) Comparing the diagnosis accuracies of using three signals after 1000 iterations, the classification results with the frequency spectrum are the best and the total accuracy is (232+246+244)/750 = 93.9%, while the results with the envelope spectrum are not so good and the total accuracy is (187+218+216)/750 = 82.8%. Although the results using three signals after 1500 iterations shown in Table III are close, using the second type of signal is still slightly better than the others.

IV. CONCLUSIONS

In this paper, a comparative study based on the CNN in the field of bearing fault diagnosis is conducted. Although many methods are developed for applying the CNN to the bearing fault diagnosis, the inputs of the CNN may be the raw vibration signal, or its processed or converted forms. In this paper, three types of signals commonly used in the vibration signal analysis, including the raw temporal waveform, its frequency and envelope spectra are inputted into the constructed CNN model based on LeNet-5. The experiment results of bearings with real damages show the classification accuracies when using three signals. When the trained CNN models are stable, the classification results using three signals are close and their accuracies are higher than 96%. The trained CNN with the time-domain signal quickly reaches to a higher classification accuracy and less learning time is used, but the model performance still fluctuates after that. As for the use of the frequency-domain signals, i.e. the frequency and the envelope spectra, the CNN using the former signal shows better results on the training accuracy, training time and testing accuracy. In the future work, development and optimization on the CNN are considered to further quantify the fault severity of bearings under variable working conditions.

ACKNOWLEDGMENTS

The work described in this paper is fully supported by the Natural Science Foundation of Guangdong Province under Grant 2017A030313278, the Fundamental Research Funds for the Central Universities (No. ZYGX2018J045), and the National Natural Science Foundation of China (Nos. 71771038, 51537010 and 61833002).

REFERENCES

[1] D.T. Hoang, H.J. Kang, "A survey on Deep Learning based bearing fault diagnosis," Neurocomputing, vol. 335, pp. 327-335, 2019.

[2] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

[3] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1D convolutional neural networks," IEEE Transactions on Industrial Electronics, vol. 63, no. pp. 7067-7075, 2016.

[4] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. Van de Walle, S. Van Hoecke, "Convolutional neural network based fault detection for rotating machinery," Journal of Sound and Vibration, vol. 377, pp. 331-345, 2016.

[5] W. Zhang, C.H. Li, G.L. Peng, Y.H. Chen, Z.J. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," Mechanical Systems and Signal Processing, vol. 100, pp. 439-453, 2018.

[6] H.B. Yang, J.A. Zhang, L.L. Chen, H.L. Zhang, S.L. Liu, "Fault diagnosis of reciprocating compressor based on convolutional neural networks with multisource raw vibration signals," Mathematical Problems in Engineering, vol. 2019, 6921975, 2019.

[7] H.D. Shao, H.K. Jiang, H.Z. Zhang, W.J. Duan, T.C. Liang, S.P. Wu, "Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing," Mechanical Systems and Signal Processing, vol. 100, pp. 743-765, 2018.

[8] J. Pan, Y.Y. Zi, J.L. Chen, Z.T. Zhou, B. Wang, "LiftingNet: a novel deep learning network with layerwise feature learning from noisy mechanical data for fault classification," IEEE Transactions on Industrial Electronics, vol. 65, no. 6, pp. 4973-4982, 2018.

[9] H.Q. Wang, S. Li, L.Y. Song, L.L. Cui, "A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals," Computers in Industry, vol. 105, pp. 182-190, 2019.

[10] L. Wen, X.Y. Li, L. Gao, Y.Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," IEEE Transactions on Industrial Electronics, vol. 65, no. 7, pp. 5990-5998, 2018.

[11] G.Q. Li, C. Deng, J. Wu, X.B. Xu, X.Y. Shao, Y.H. Wang, "Sensor data-driven bearing fault diagnosis based on deep convolutional neural networks and S-transform," Sensors, vol. 19, no. 12, paper no. 2750, 2019.

[12] F. Jia, Y.G. Lei, N. Lu, S.B. Xing, "Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization," Mechanical Systems and Signal Processing, vol. 110, pp. 349-367, 2018.

[13] O. Abdeljaber, S. Sassi, O. Avci, S. Kiranyaz, A.A. Ibrahim, M. Gabbouj, "Fault detection and severity identification of ball bearings by online condition monitoring," IEEE Transactions on Industrial Electronics, vol. 66, no. 10, pp. 8136-8147, 2019.

[14] L. Guo, Y.G. Lei, S.B. Xing, T. Yan, N.P. Li, "Deep convolutional transfer learning network: a new method for intelligent fault diagnosis of machines with unlabeled data," IEEE Transactions on Industrial Electronics, vol. 66, no. 9, pp. 7316-7325, 2019.

[15] R. Zhao, R.Q. Yan, Z.H. Chen, K.Z. Mao, P. Wang, R.X. Gao, "Deep learning and its applications to machine health monitoring," Mechanical Systems and Signal Processing, vol. 115, pp. 213-237, 2019.

[16] R.B. Randall, J. Antoni. Rolling element bearing diagnostics — A tutorial. Mechanical Systems and Signal Processing, 2011, 25: 485-520.

[17] C. Lessmeier, J.K. Kimotho, D. Zimmer and W. Sextro, "Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification," European conference of the prognostics and health management society, 2016.