HASAN, M.J., SOHAIB, M. and KIM, J.-M. 2019. 1D CNN-based transfer learning model for bearing fault diagnosis under variable working conditions. In Omar, S., Haji Suhaili, W.S. and Phon-Amnuaisuk, S. (eds.) *Computational intelligence in information systems: proceedings of the 2018 Computational intelligence in information systems conference (CIIS 2018), 16-18 November 2018, Brunei*. Advances in intelligent systems and computing, 888. Cham: Springer [online], pages 13-23. Available from: https://doi.org/10.1007/978-3-030-03302-6_2

1D CNN-based transfer learning model for bearing fault diagnosis under variable working conditions.

HASAN, M.J., SOHAIB, M. and KIM, J.-M.

2019

1D CNN-Based Transfer Learning Model for Bearing Fault Diagnosis Under Variable Working Conditions

Md Junayed Hasan, Muhammad Sohaib, and Jong-Myon Kim*

Department of Electrical and Computer Engineering,
University of Ulsan, Ulsan, South Korea
junhasan@gmail.com, md.sohaibdurrani@gmail.com,
jmkim07@ulsan.ac.kr

Abstract. Classical machine learning approaches have made remarkable contributions to the field of data-driven techniques for bearing fault diagnosis. However, these algorithms mainly depend on distinct features, making the application of such techniques tedious in real-time scenarios. Under variable working conditions (i.e., various fault severities), the acquired signals contain variations in the signal amplitude values. Therefore, the extraction of reliable features from the signals under such conditions is important because it could discriminate the health conditions of the bearings. In this paper, a transfer learning approach based on a 1D convolutional neural network (CNN) and frequency domain analysis of the vibration signals is presented to solve the problem. Transfer learning enables the developed model to utilize information obtained under a given working condition to diagnose faults under other working conditions. The proposed approach has a classification accuracy of 99.67% when tested with the data acquired from the bearings with various fault severities. We also observe that a frequency spectrum enhances the performance of the transfer learning-based fault diagnosis model.

Keywords: Bearing fault identification; Convolutional neural network Transfer learning; Vibration signals

1 Introduction

Bearing faults are some of the most prominent reasons for induction machine failure [1]. Incipient fault diagnosis of the bearings can avoid financial loss, reduce machine down time due to a failure, and improve plant production capacity [2, 3]. Bearing fault diagnosis can be divided into model-driven and data-driven techniques [4, 5]. Data driven techniques seem more promising for bearing fault diagnosis [1]. In such techniques, distinct features associated with the different health states of the bearing are extracted from signals without knowing the detailed architecture of the machinery. Later, the extracted features are provided to a sophisticated machine learning algorithm to classify the data into different classes.

However, the challenge in such a technique is to extract suitable features from the signals. Feature extraction is a tedious process, and the quality of the extracted features affects the performance of the classical machine learning algorithms [6–8]. Moreover, the same types of data and feature space under a constant working condition are required for such methods [1]. Data collected from the bearing in the form of signals is a variable of working conditions, i.e., variation in a working condition effects the data values. In the case of bearings, fault severity, motor loads, and speeds are responsible for inconstant working conditions. The performance of the classical bearing fault diagnosis techniques deteriorates under variable working conditions due to the data variation. In addition, dealing with a massive amount of data makes the fault diagnosis task more difficult and effects the performance of a developed model. To address such problems, many deep learning techniques have been used in the field bearing fault diagnosis [2, 6, 9–11]. However, these methods suffer when dealing with big data in variable working conditions. To solve this issue, Zhang et al. [4] developed transfer learning using a convolutional neural network. The transfer learning process has been described in detail by using raw vibration signals for training and testing. However, many important points are left out, which are useful for fault classification. In another study by Wen et al. [1] a transfer learning model with an auto encoder was developed to solve the problem. However, the dimensionality reduction process through an auto encoder requires many additional efforts as well.

The main contribution of the current work is: (1) a transfer learning-based approach for fault diagnosis of the bearing with various fault severities, (2) a suggestion of a bearing fault diagnosis model that does not require well engineered features, and (3) introduction of a frequency spectrum calculation as a preprocessing step of the vibrations signals to explore the potential information and enhance the performance of the proposed model. To project a real-time scenario, we have created two different working conditions by using publicly available bearing data [12]. Both the working conditions have different fault severities and number of classes.

The rest of the paper is described as follows. Section 2 defines the proposed methods with the background, while Sect. 3 elaborates on the experimental setup and results of the experiment. Finally, Sect. 4 draws conclusions.

2 Proposed Method

2.1 Skimming Frame

In this study, vibration signals acquired from a bearing under two different working conditions are used for evaluation of the proposed bearing fault diagnosis model. In transfer learning, a network is first trained using the training data, and then the knowledge gained during the training is transferred to the target domain while testing the network. Hence, to create two working conditions, a huge amount of data is required. Thus, to create a significant number of samples, a skimming frame algorithm is used [4]. For each tiny step, one sample will be created.

If the total length of a vibration signal is “L”, then the total number of the samples “n” will be:

$$n = \left(\frac{L - l}{s}\right) + 1. \quad (1)$$

In (1) “l” is the length of a single frame that needs to be selected based on the experimental requirement. The step size is considered as “s”. The dimension of the vibration signal “L” is fixed. A complete process of the skimming frame algorithm is illustrated in Fig. 1.

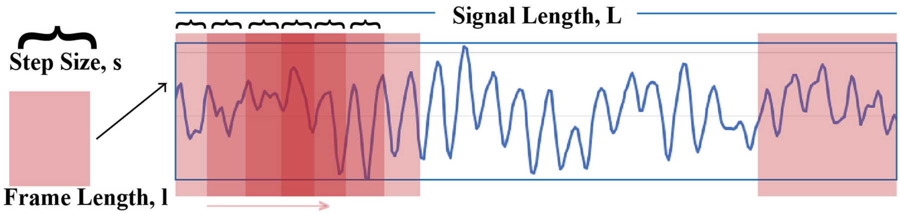

Fig. 1. Skimming frame process for generating samples.

2.2 Fast Fourier Transform

In this research, vibration signals are used in the experiments. It is easier to extract useful information from frequency domain by using Fast Fourier Transformation (FFT) [2]. If the number of samples is v, then FFT will consider $v * \log(v)$ operations. Figure 2 shows the FFT calculation process of the raw vibration signals. The FFT of each sample is calculated, and according to the Nyquist theorem, half of the data points in the frequency spectrum of each sample are considered for use further in the experiments.

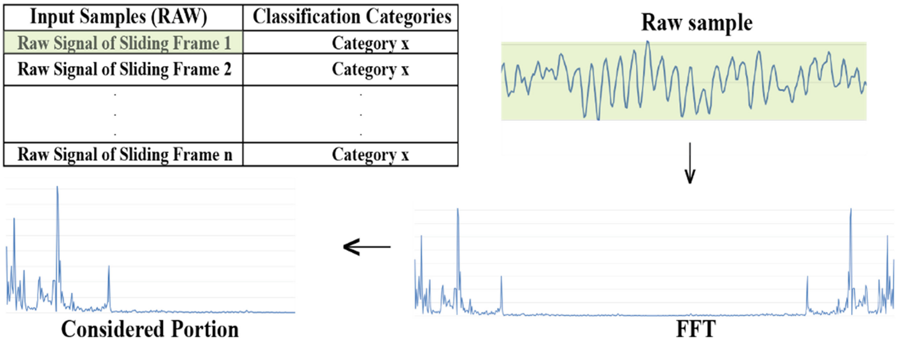

Fig. 2. FFT process for preprocessing the RAW signal.

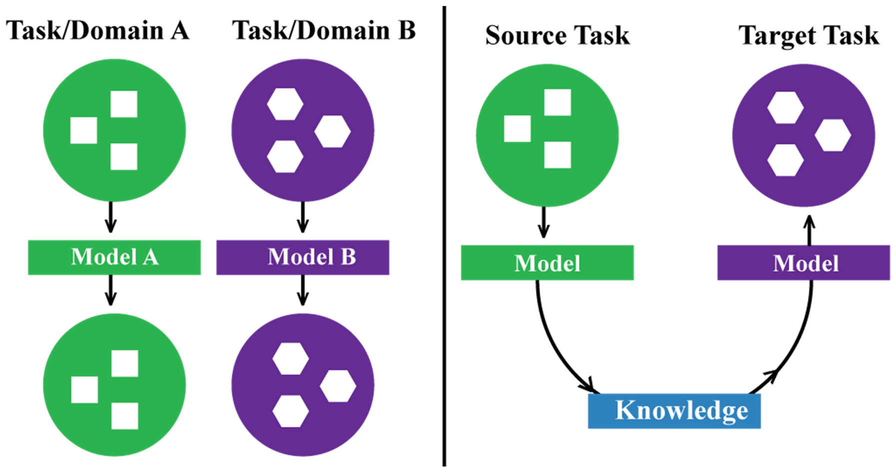

Fig. 3. (a) A traditional machine learning approach is presented which has two independent models for two different tasks; (b) a transfer learning approach is shown where the two tasks are denoted as source task and target task and knowledge is transferred between the tasks.

2.3 Transfer Learning

Developing a new network for a new task is a cumbersome and challenging process that adds an overhead to the fault diagnosis process. On the other hand, transfer learning makes the fault diagnosis process easy and robust [13]; it extracts valuable information from one working condition (task A) and applies that knowledge to the target working condition (task B) [4]. In general, transfer learning allows for learning from a vast number of past experiences, and then transferring that knowledge into different environments. In this research, first, the network learns distinct characteristics from massive amounts of source data in one working condition, and then passes that knowledge to a target task that has the same type of data but acquired under a different working condition. The main idea behind transfer learning is identifying a remarkable improvement regarding target scenarios from the knowledge that is gathered from the source scenario. Figure 3 illustrates the idea of transfer learning.

2.4 Overview of 1D CNN

A Convolution Neural Network (CNN) is one of the most effective supervised machine learning approaches [9]. In the current work, 1D CNN has been used in a transfer learning-based approach for machine fault diagnosis. CNN has a hierarchical architecture composed of convolutional, subsampling, and fully connected layers. 1D CNN is the same as conventional 2D CNN; the only difference is the dimension of the input data and filters used in the multiple layers. In 1D CNN, if the training data is $M = [m_1, m_2, \ldots\ldots m_j]$, then the number of training sample is j. Moreover, a target vector is $N = [n_1, n_2, \ldots\ldots n_j]$ which is associated with $M \cdot K$ layers constitute a CNN, then each layer in the network has F^K features, which are used in the convolution and subsampling processes [4]. If the sigmoid activation function is $\sigma(.)$, the weight matrix between input and hidden layer is w_1, the weight matrix between hidden and output layer is w_2, the bias vector of the hidden layer is b_1 and the output layer is b_2, then the feed forward process will be denoted as follows:

$$P = \sigma(w_1 M + b_1) \tag{2}$$

$$N = \sigma(w_2 P + b_2) \tag{3}$$

$$N = \sigma(w_2 P + b_2) \tag{4}$$

CNN is like an ordinary neural network that has adjustable weights and biases. In 1D CNN if the input signal length is $L^{'}$, then the input size is $(1 * L^{'} * 1)$. As a result, the convolution layer will calculate the output of the neurons. In the proposed model, 32 1D filters have been considered in the convolution layer. After the convolution layer, a dropout layer is added to the network to avoid overfitting. The third layer in the network is max-pooling, which subsamples the data. The max-pooling layer is followed by a fully connected layer, which computes the class scores. Each neuron of the fully connected layer is connected to all the neurons in the previous layer. A stochastic gradient descent algorithm is used for the fine tuning of the network. A softmax classifier is added to the network for data classification. Figure 4 describes the overall architecture of the network.

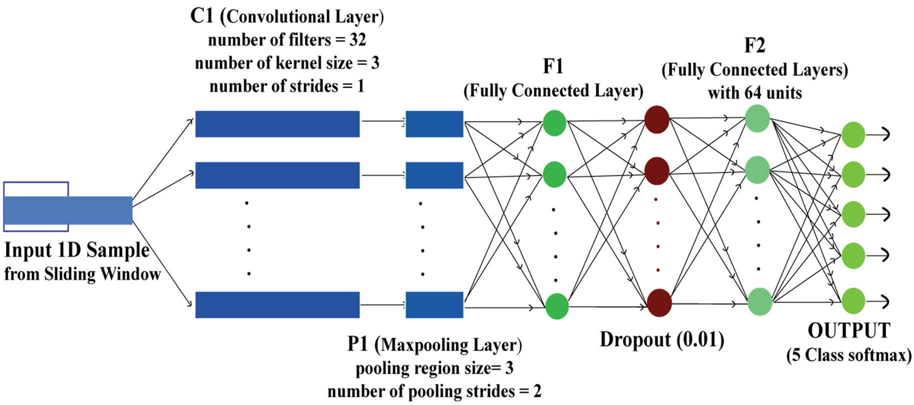

Fig. 4. Detail architecture of 1D CNN.

2.5 Details of the Proposed Approach

The proposed transfer learning-based bearing fault diagnosis consists of two phases. The first phase is the training phase of the network and the second is the testing phase of the model by transferring the domain knowledge. In this research, two working conditions are considered. First, the network is trained on the first working condition.

Then, the acquired knowledge is transferred to the second working condition. Figure 5 illustrates an overall workflow of the proposed model. The steps of the training phase are as follows:

1. Required source data is collected for the first working condition.
2. The skimming frame technique is applied to generate multiple samples.
3. To each sample, a label is assigned, as is necessary for supervised learning.
4. Fast Fourier Transform (FFT) calculation of the raw vibration samples is performed to obtain the frequency spectra of the signals.

Finally, the network (1D CNN) is trained using all samples from the training data, and the weights of the trained model are saved for later use during the testing phase.

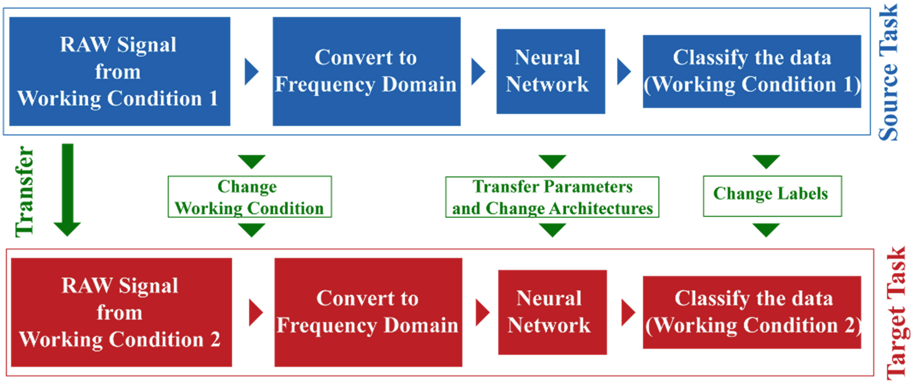

Fig. 5. Block diagram of the proposed model.

The Steps of the transferring phase are as follows:

1. As the working conditions are different, the data may contain atypical features and labels.
2. Labels are defined for each sample.
3. The FFT is calculated.
4. 2% of data is utilized to build the network architecture according to the data labels and dimensionality. The input size should be like the trained network, but the output will be different because of additional new labels.
5. The parameters of the main network will be unchanged.
6. Finally, the classification accuracy will be calculated.

The proposed model is evaluated by comparing the results with that of a fault diagnosis model where no preprocessing step is considered. Moreover, the results of the proposed model are also compared with the model where transfer learning is not used.

3 Experimental Setup

3.1 Dataset

To evaluate the proposed model, publicly available seeded fault bearing dataset by Case Western Reverse University was considered throughout the experiments. Drive end bearings were seeded with faults on the inner raceways, outer raceways, and rolling elements having a diameter of 0.007, 0.014, and 0.021 in. with the help of an electro-discharge machine as shown in Fig. 6.

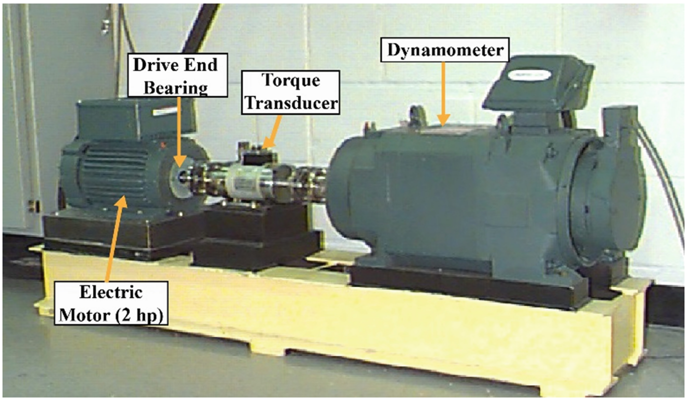

Fig. 6. Experimental set up by Case Western Reserve University [12].

Table 1. Details of the considered working conditions

Field name	Working condition 1	Working condition 2
Fault diameter	0.007 in.	0.021 in.
Motor load	0 horse power (hp)	0 horse power (hp)
Motor speed	1797 rounds/minute	1797 rounds/minute
Number of fault types	3	5
Number of labels	4	6
Name of the labels	Normal Inner fault Ball fault Outer fault at center	Normal Inner fault Ball fault Outer fault at center Outer fault at orthogonal Outer fault at oppositely
Utilized for	Source task	Target task

Thus, the dataset consists of an inner race fault, ball fault, and outer race fault signals under variable working conditions. Variable length vibration signals were recorded via an accelerometer attached to the housing of the motor close to the drive end bearing with a sampling data rate of 12,000 Hz. A dynamometer was used to generate different motor loads, ranging from 0 to 3 horsepower (hp). The placement and load zone on bearings create an impact on the vibration data of the outer race fault as it is stationary [12]. Details about the dataset and different working conditions are given in Table 1.

3.2 Result Analysis

The network is trained on the data from the first working condition. Table 1 shows that there are four types of signals associated with the health of the bearings for the first working condition. The four health states contain one normal and three faulty conditions. For each type, 3980 samples are considered. During the source training phase, the network gives 100% learning accuracy, while using 0.02% data for validation. The weights of the trained network are stored and transferred to the second working condition. For working condition 2, new labels are added to the dataset; thus, the output architecture of the network is modified. Therefore, from this working condition, 0.02%data is used to adjust the network architecture according to the new data. After the

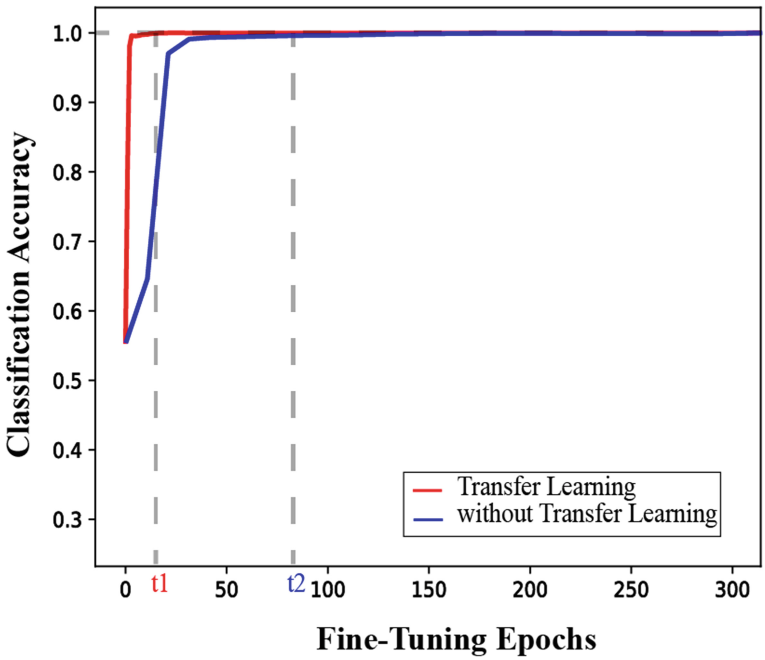

Fig. 7. Classification accuracy comparison of second working model (transfer learning-based model vs. without transfer learning).

network adjustment, the model is tested with the data of the second working condition. In addition, a separate CNN is developed and tested with the data of the second working condition without using the transfer learning technique, and the results are compared with that of the proposed model.

From Fig. 7, we see that both experiments can achieve around 100% accuracy. It is observed that by adopting a transfer learning technique in the fault diagnosis model, better performance can be achieved for the second working model at an earlier stage. After 16 fine tuning epochs (t1), the network can give almost 100% accuracy. On the other hand, without using transfer learning technique, after 80 fine-tuning epochs (t2), the network exhibits similar accuracy. Hence, the two approaches can achieve same performance in a different time span. In Table 2, the label-wise performance is given. For normal and ball fault condition, the improvement is 0%. The significance of using transfer leaning in the proposed model can be observed in case of the inner fault with an accuracy improvement of 4.8%. Overall, a 1.23% improvement in the average classification accuracy is achieved through a transfer learning-based approach for bearing fault diagnosis. In short, through transfer learning, better performance can be achieved at an earlier stage as compared to the conventional network.

Table 2. Comparison of classification accuracies of different approaches

Labels	Transfer learning	Without transfer learning	Improvements
Normal	100%	100%	0%
Inner fault	100%	95.2%	4.8%
Ball fault	100%	100%	0%
Outer fault at center	98.46%	98.13%	0.33%
Outer fault at orthogonal	100%	98.24%	1.76%
Outer fault at oppositely	99.54%	99.10%	0.44%
Overall	**99.67%**	**98.45%**	**1.23%**

To whether the preprocessing of the vibration signals has any impact on the performance of the proposed model, comparisons are made with the results of a 1D CNN that is trained on raw vibration signals. The training and testing approach of the network is identical to the proposed scheme: (1) train with the raw data of the first working condition, save, and then transfer the knowledge; (2) then test the model with the data from the second working condition. It is evident from Fig. 8 that with the same experimental set up, a preprocessing step can improve average classification accuracy of the model by at least 16%.

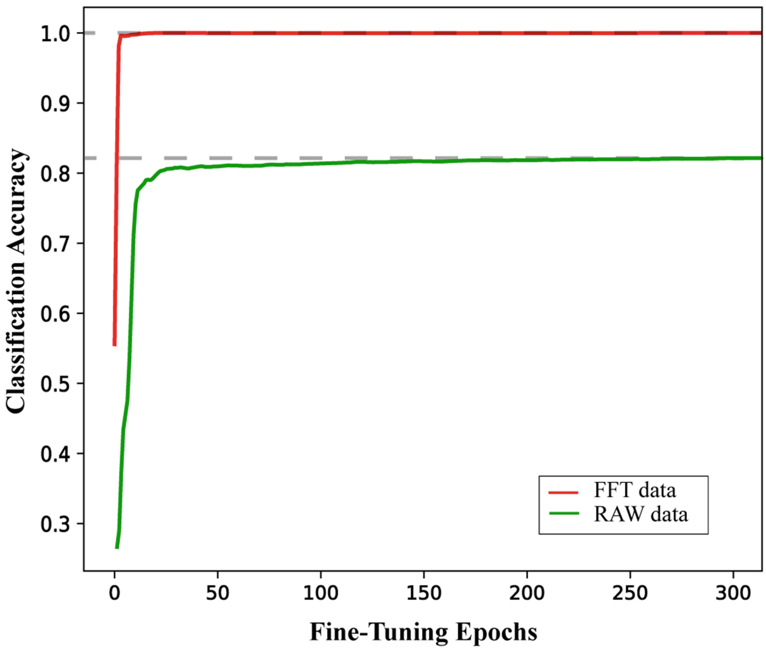

Fig. 8. Classification accuracy of the proposed model using preprocessed data (FFT data) vs. RAW data.

4 Conclusion

This paper proposed a novel fault diagnosis approach for the bearings of a rotary machine using a 1D convolutional neural network (CNN)-based transfer learning technique. Moreover, a preprocessing step was added into the fault diagnosis pipeline, which consists of a frequency spectrum calculation of the vibration signals. The addition of a preprocessing step in the proposed model enhanced the overall performance by extracting useful information. The results demonstrated that the proposed method achieves a 99.67% accuracy by using vibration signals acquired form the bearings under different working conditions. The frequency spectrum calculation of the vibration signals improved the average accuracy of the proposed model by 16%. In the future, the presented work can be extended to build an automated online fault detection system.

Acknowledgement. This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP), and the Ministry of Trade, Industry, & Energy (MOTIE) of the Republic of Korea (Nos. 20181510102160, 20162220100050, 20161120100350, 20172510102130). It was also funded in part by The Leading Human Resource Training Program of Regional Neo industry through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, and future Planning (NRF-2016H1D5A1910564), and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1D1A3B03931927).

References

1. Wen, L., Gao, L., Li, X.: A new deep transfer learning based on sparse auto-encoder for fault diagnosis. IEEE Trans. Syst. Man Cybern. Syst. 1–9 (2017)
2. Ince, T., Kiranyaz, S., Eren, L., Askar, M., Gabbouj, M.: Real-time motor fault detection by 1-D convolutional neural networks. IEEE Trans. Ind. Electron. **63**, 7067–7075 (2016)
3. Eren, L., Devaney, M.J.: Bearing damage detection via wavelet packet decomposition of the stator current. IEEE Trans. Instrum. Meas. **53**, 431–436 (2004)
4. Zhang, R., Tao, H., Wu, L., Guan, Y.: Transfer learning with neural networks for bearing fault diagnosis in changing working conditions. IEEE Access. **5**, 14347–14357 (2017)
5. Gao, Z., Cecati, C., Ding, S.X.: A survey of fault diagnosis and fault-tolerant techniques Part I: fault diagnosis. IEEE Trans. Ind. Electron. **62**, 3768–3774 (2015)
6. Worden, K., Staszewski, W.J., Hensman, J.J.: Natural computing for mechanical systems research: a tutorial overview. Mech. Syst. Signal Process. **25**, 4–111 (2011)
7. Kang, M., Islam, M.R., Kim, J., Kim, J.M., Pecht, M.: A hybrid feature selection scheme for reducing diagnostic performance deterioration caused by outliers in data-driven diagnostics. IEEE Trans. Ind. Electron. **63**, 3299–3310 (2016)
8. Sohaib, M., Kim, C.-H., Kim, J.-M.: A hybrid feature model and deep-learning-based bearing fault diagnosis. Sensors **17**, 2876 (2017)
9. Malek, S., Melgani, F., Bazi, Y.: One-dimensional convolutional neural networks for spectroscopic signal regression. J. Chemom., 1–17 (2017)
10. Appana, D.K., Ahmad, W., Kim, J.-M.: Speed invariant bearing fault characterization using convolutional neural networks. In: Phon-Amnuaisuk, S., Ang, S.-P., Lee, S.-Y. (eds.) Multi-disciplinary Trends in Artificial Intelligence, pp. 189–198. Springer, Cham (2017)
11. Jia, F., Lei, Y., Lin, J., Zhou, X., Lu, N.: Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mech. Syst. Signal Process. **72–73**, 303–315 (2016)
12. Case Western Reserve University: Bearing Data Center Website. http://csegroups.case.edu/bearingdatacenter/pages/download-data-file
13. Michalik, T., Polska, O.: How effective is Transfer Learning method for image classification, vol. 12, pp. 3–9 (2017)