Intelligent Fault Diagnosis of Gearbox Based on Vibration and Current Signals: A Multimodal Deep Learning Approach

Guoqian Jiang, Jingyi Zhao, Chenling Jia, Qun He, Ping Xie, and Zong Meng
School of Electrical Engineering
Yanshan University
Qinhuangdao 066004, China
jiangguoqian@ysu.edu.cn; zjy261@163.com; jcl1294516297@163.com; hq@ysu.edu.cn; pingx@ysu.edu.cn; mz@ysu.edu.cn

***Abstract*—This paper proposes a new intelligent fault diagnosis approach based on multimodal deep learning to fuse vibration and current signals to diagnose wind turbine gearbox faults. The proposed method typically consists of modality-specific feature learning network and feature fusion network, specifically based on a popular deep learning model named deep belief networks (DBNs). First, two individual DBNs are designed to learn fault-related features directly from raw vibration signals and current signals, respectively. Then, the learned vibration-based features and current-based features are further fused by a third DBN to output the final diagnosis results. The proposed approach is verified on a wind turbine drivetrain gearbox test rig. The experimental results demonstrate that the proposed approach outperformed the compared methods based on single sensor and data-level fusion in terms of diagnostic accuracy, which attributes to the complementary diagnosis information from vibration signals and current signals.**

***Index Terms*—wind turbines, gearbox, fault diagnosis, multi-modal deep learning, information fusion**

I. Introduction

GEARBOX is a critical component in a wind turbine [1], and it will often suffer from various faults owing to the severe working environments of wind turbines, as a result, often leading to frequent unexpected downtime and huge economical losses. Currently, there is an urgent need to develop advanced fault diagnosis systems to accurately and automatically identify the health state of the gearbox.

The conditions of a gearbox can be reflected by different sensor signals, including vibration [2], acoustic [3], current [4] and temperature [5], etc. Vibration signals have been considered as the most commonly used way for fault diagnosis because of its ability to accurately indicate the health status of rotating machinery. To extract effetive features for fault diagnosis, a series of signal preprocessing algorithms, including wavelet transform [6], empirical mode decomposition [7], multiscale representation [8] and Hibert-Huang transform [9], have been widely used and achieved successful results. However, the measured vibration signals are easily affected by ambient noise, leading to low signal-to-noise ratio and even false diagnosis results. Consequently, it is still a challenging task to extract useful and fault-related features from the raw vibration signal [10]. Moreover, the installiation position of accelerometer will significantly affect the dignosis accuracy [11]. Differently, in a recent study [12], Lu *et al.* has proven that the failure of a gearbox will induce the changes of the current signal of the generator adjacent to it, and therefore the current signal can be used as an alternative fault diagnosis way for wind turbine drivetrain gearbox. Compared with the vibration sensor, the current sensor is easy to mount and current signal is easier to obtain. However, due to the presence of dominant current fundamental component and harmonic component as well as electrical noise, it is still a major challenge to effectively extract the fault-related features contained in current signals.

From the analysis above, we can draw the conclusion that the information obtained from a single sensor obviously has its limitations. The fault characteristics reflected by a single sensor can not well represent the running state of a certain component due to the complexity of the equipment itself and its operation conditions. This uncertainty and limitation will inevitably lead to the reduction of the fault diagnosis performance. Multi-sensor information fusion is a great option to address these problems, which has been applied to some other fields [13]–[15]. It has been shown that different sensors usually contain complementary fault diagnosis information. By fusing these different information, the final diganosis decisions can make full use of the strengths of the individual sensor information, overcome their respective weaknesses and get a better diagnotic performance [16]. This makes multi-sensor information fusion have unique advantages in solving the uncertainty and limitation of fault diagnosis based on single sensor information. Information fusion has been widely recognized as three levels on data, feature and decision [17]. Feature level fusion is a process of target fusion by pre-processing sensor data to obtain data feature information and correlating the acquired features. Compared with other fusion level, feature level fusion uses more features and has

This work is partly supported by the Natural Scientific Foundation of China (Grants No. 61803329), the Natural Scientific Foundation of Hebei Province, China (Grants No. F2018203413 and F2016203421), the Key Research and Development Program of Hebei Province (19214306D), and China Postdoctoral Science Foundation (2018M640247).

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

less information loss. Recent researches have shown that fault diagnosis employing multiple sensors information fusion techniques can enhance diagnosis accuracy and reliability. C. Li *et al.* [18] proposed a deep random forest fusion approach to fuze acoustic and vibration signals to diagnosis different health conditions of a two-stage parallel gearbox. M. Xia *et al.* [19] designed a convolutional neural network model to fuse mutiple channel vibration signals from sensors at different locations. M. Ma *et al.* [20] developed a new deep coupled autoencoder model to learn joint features from vibration and acoustic multimodal signals.

Recently, deep learning, as a prevalent method in the machine learning community, has been successfully applied to image recognition, speech recognition, information retrieval and other fields [21]–[23]. As a classical model of deep learning, deep belief network (DBN) has drawn widespread attention for its superior feature extraction ability and unique training methods [24], and recently it has found wide applications in the fault diagnosis field. For example, P. Tamilselvan *et al.* [25] first applied DBN to aircraft engine fault diagnosis in 2013. M. Gan *et al.* [26] designed a DBN-based two-layer hierarchical diagnosis network to identify bearing faults. H. Shao *et al.* [27] designed a novel convolutional DBN with Gaussian visible units to learn the representative features for electric locomotive bearing fault diagnosis. The above studies using DBN to perform fault diagnosis mainly focus on the feature extraction from a single sensor.

To better fuse the vibration signals and current signals, motivated by previous studies, this paper proposes a multimodal deep learning based fault diagnosis approach for wind turbine gearboxes. The proposed method designed three different DBN models to perform the feature learning and fusion. Two individual DBNs are first used to learn and extract the features from raw vibration signals and current signals, respectively. Then, the learned vibration and current based features are fused by the third DBN to realize the health state classification. The proposed approach can learn the complementary and rich diagnosis information from vibration signals and current signals and therefore enhance fault diagnosis performance.

The remainder of this paper is organized as follows. Section II gives a brief introduction of DBN. Section III details the proposed Multimodal DBN method for intelligent fault diagnosis. Section IV presents the gearbox fault diagnosis experiments and the corresponding evaluation results. Section V finally draws a conclusion.

II. DBN

DBNs are stacked by Restricted Boltzmann Machines (RBM) structurally. The neurons in each layer of RBM are divided into visible layer neurons and hidden layer neurons, where the former is used to accept input and then the hidden layer neurons output extracted features. A greedy layer-by-layer training scheme is utilized to train the network. The overall training process of the DBNs consists of the following two stages.

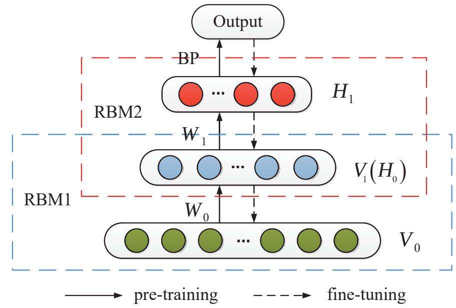

Figure 1. Schematic architecture of DBN

A. Pre-training

Each RBM network is trained separately in an unsupervised way to initialize of the network parameters. As shown in Figure 1, taking two layers of RBM as an example, the training data is first used to train RBM1, which is composed of the visible layer V_0 and the hidden layer H_0. The visible layer V_0 receives the original feature vector. Learning the parameter $\mathbf{W}_0$ between the visible layer and the hidden layer first during the training procedure. After training RBM1, continue to train RBM2 in an unsupervised manner. As can be seen from Figure 1, the visible layer V_1 of RBM2 is the hidden layer H_0 of RBM1.

Supposing that the visible layer V contains n dominant neurons and the hidden layer H contains m recessive neurons after training. Then, the energy of RBM as a system can be calculated as:

$$E(v,h|\theta) = -\sum_{i=1}^{n} a_i v_i - \sum_{i=1}^{m} b_j h_j - \sum_{i=1}^{n}\sum_{i=1}^{m} v_i h_j \omega_{ij} \tag{1}$$

where v_i and a_i represent the state and the bias of the i-th dominant neuron, h_j and b_j represent the state and the bias of the j-th recessive neuron, ω_{ij} denotes the connection weight between dominant neuron i and recessive neuron j. Based on this energy function, the joint probability distribution of (v,h) can be obtained as:

$$P(v,h|\theta) = \frac{e^{-E(v,h|\theta)}}{Z(\theta)} \tag{2}$$

$$Z(\theta) = \sum_{v,h} e^{-E(v,h|\theta)} \tag{3}$$

where $Z(\theta)$ is the normalization factor. The values are distributed to the interval (0,1) to maintain the sparsity. All the hidden layers are added to obtain the likelihood function of the joint probability distribution $P(v,h|\theta)$:

$$P(v|\theta) = \frac{1}{Z(\theta)} \sum_{h} e^{-E(v,h|\theta)} \tag{4}$$

The learning task of RBM is to find the parameters θ to fit the training data. By using the contrast divergence algorithm, we can calculate the partial derivatives of the connection

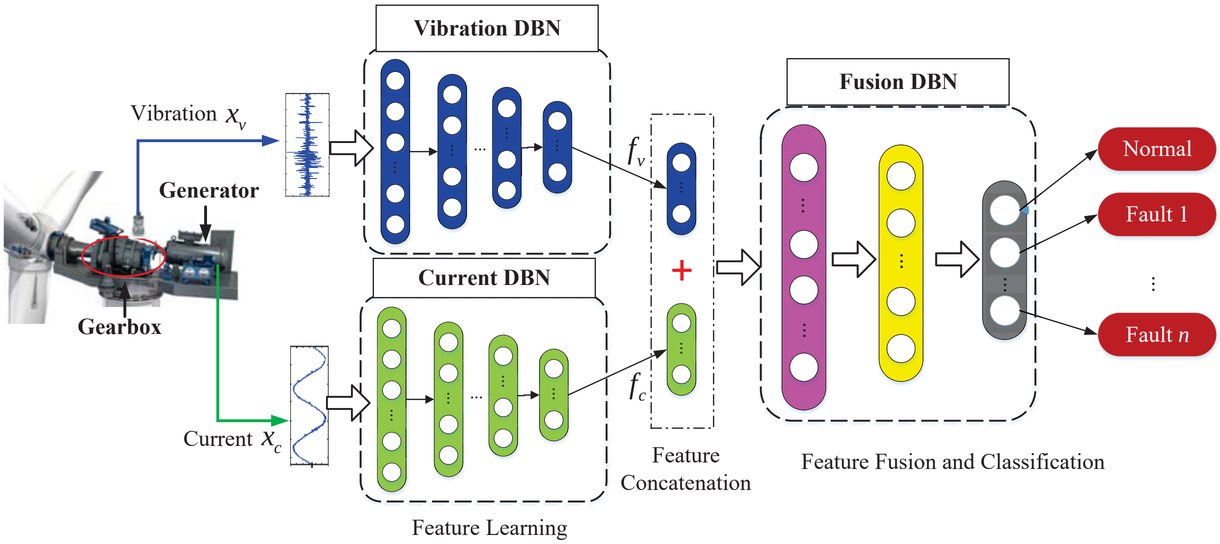

Figure 2. Proposed fault diagnosis approach based on multimodal deep belief networks with vibration and current signals.

weight ω_{ij}, the bias a_i and the bias b_j, which are respectively expressed as:

$$\frac{\partial \log P(v|\theta)}{\partial \omega_{ij}} =< v_i h_j >_{data} - < v_i h_j >_{model} \tag{5}$$

$$\frac{\partial \log P(v|\theta)}{\partial a_i} =< v_i >_{data} - < v_i >_{model} \tag{6}$$

$$\frac{\partial \log P(v|\theta)}{\partial b_j} =< h_j >_{data} - < h_j >_{model} \tag{7}$$

where $< \bullet >_{data}$ represents the mathematical expectation with respect to the data set and $< \bullet >_{model}$ represents the mathematical expectation defined by the model.

B. Fine-tuning

Although each layer of RBM can obtain the initial parameters after pre-training without supervision, they are not the optimal parameters. In this case, the parameters of each layer need to be further optimized. The process of fine-tuning is to train the top-level softmax network in a supervised manner using back-propagation algorithm. The features learned by the RBM network are combined for classification. At the same time, the error information is reversely transmitted to all RBM networks and the parameters between the RBM layers are fine-tuned. Then the obtained results are the optimal parameters of the DBN network after supervised fine-tuning process.

III. Proposed Approach

The flowchart of the proposed framework is shown in Figure 2. It aims to learn different and complementary information hidden in the gearbox vibration signals and the generator current signals through a multimodal DBN architecture, which consists of modality-specific feature learning network and feature fusion and classification network.

A. Modality-specific Feature Learning Network

Different sensor modalities usually contain different diagnosis information related to fault patterns. Vibration signals and current signals measured from the gearbox and the generator, respectively, can both represent the health state of the underlying machine from two different aspects with complementary information, which provides a possible way for information fusion. However, each sensor signal has its own complex characteristics. In order to extract useful fault features, in this paper, two individual DBN networks are used for vibration and current feature learning. As shown in Figure 2, **Vibration DBN** and **Current DBN** are used to directly learn the hidden layer fault feature from the vibration signal and the current original signal, respectively. Taking the vibration signal as an example, $X_v \in R^{p\times q}$, p is the sample size and q is the sliding window width. The relationship between X_v and feature vector F_v can be found through DBN1 and it can be expressed as follows:

$$F_v = f_{DBN_1}(X_v) \tag{8}$$

Similarly, the relationship between current feature vector F_c and current signal X_c is:

$$F_c = f_{DBN_2}(X_c) \tag{9}$$

Notice that the hidden layer number n of **Vibration DBN** and **Current DBN** can be determined according to the specific signal. The number of hidden layers n in **Vibration DBN** and **Current DBN** can be different, and the number of neurons in the hidden layer can be also different. In this paper, for its simplicity the values of n in both DBNs are set the same value.

B. Feature Fusion and Classification Network

Considering different modalities provide complementary discriminability among different health conditions, the aim of the stage is to fuse the learned features for vibration

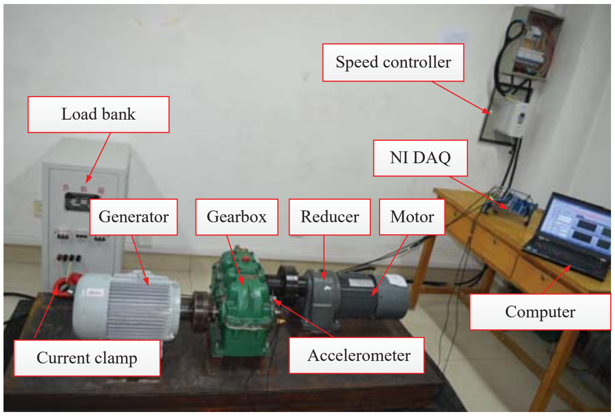

Figure 3. Wind turbine gearbox experimental setup.

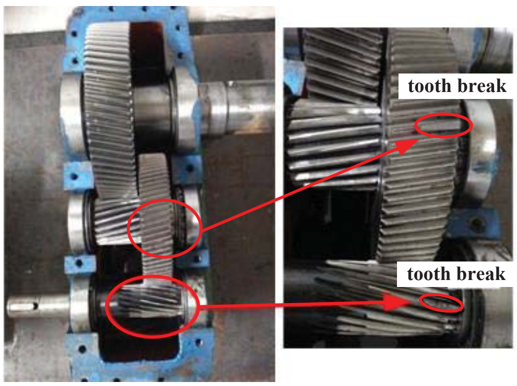

Figure 4. The structure of the gearbox

and current sensor modalities. After the feature extraction of single-mode sensor signals, the fault features of n-th hidden layer in each single-mode signals can be concatenated as a joint feature vector (F_v, F_c). Then a **Fusion DBN** is designed to learn unified featured hidden in the vibration and current modalities, and the learned unified feature vector $F_{c,v}$ can be represented as:

$$F_{v,c} = f_{DBN_3}(F_v, F_c) \tag{10}$$

In the fusion feature network, **Fusion DBN** is trained in an unsupervised, layer-by-layer way and then it is fine-tuned through supervised training using back-propagation algorithm. The softmax network is utilized to identify the health conditions of the gearbox and given the final diagnosis results.

IV. Experimental validation

A. Experiment setup and data description

In this section, we conducted experiments on a wind turbine drivetrain gearbox test rig to validate the proposed approach. Figure 3 illustrates the experimental setup, which mainly consists several modules: an induction motor of a 3kW rated power coupled with a speed reducer of a decreasing ratio of 40, a two-stage parallel gearbox of a total increasing ratio of 20, a permanent magnet synchronous generator of a 3kW rated power and a load bank. The motor is operated by a speed controller. The reducer is used to emulate the dynamics of the rotor, and to reduce the output speed of the motor to a lower speed. The gearbox is responsible for increasing the speed to to drive the generator. The load bank is employed to consume the generated power.

In our experiment, the gearbox radial vibration is measured on the gearbox near the high-speed shaft using an accelerometer. And one phase generator current is acquired via a current clamp . The input speed of the high-speed shaft is about 300 r/min. The sampling frequency of vibration signal and current are set as 5 kHz and 1 kHz, respectively.

Figure 4 shows the structure of the tested gearbox. In this experiment, one third of the tooth width of one gear on the high-speed shaft is removed to simulate single-shaft tooth break fault. On this basis, 1/2 of the tooth width of the gear is removed to simulate the double-shaft tooth broken fault. Therefore, three health conditions are considered, including normal condition (N), single-gear tooth broken (B) fault and double-gear tooth broken (DB) fault. Vibration and current data are acquired under three health conditions for 30 seconds. We use the sliding window with different size of 2000 and 4500 to partition the current signal and vibration signal into several segments, respectively, where each segment is considered as one sample. In this study, totally 200 samples are obtained for both vibration signals and current signals.

B. Experimental Results

For performance evaluation, the dataset is split into both sets: the training set and the testing set. Specifically, 60% of the total samples are randomly selected for training and and the rest for testing. All training and testing sets are normalized to limit its range within (0,1). As shown in Figure 2, the vibration-based and current-based fault features of the last hidden layer are first learned through **Vibration DBN** and **Current DBN**. Finally, both learned fault features are concatenated and input into **Fusion DBN** to obtain the final fault classification results. In this study, 10 random experiments are repeated to calculate the average results for performance evaluation.

Before the model training, all parameters, including model structure, learning rate, momentum and the number of epochs, for **Vibration DBN**, **Current DBN** and **Fusion DBN** need to be determined, and the corresponding parameters are summarized in Table I. For each DBN model, its structure is denoted with an input layer, several hidden layers and an output layer. For example, in **Vibration DBN**, its input layer contains 4500 neurons corresponding to the length of each vibration sample, and its output layer contains 3 neurons corresponding to three health conditions of the gearbox. In additions, **Vibration DBN** contains five hidden layers with the number of neurons of 2000, 1000, 500, 250 and 100, respectively. All DBN models are trained using the stochastic gradient decent (SGD) method.

In order to highlight the superiority of the proposed method, it is compared with two single sensor based methods with vibration signals and current signals. The results are shown in Table II, where the reported results are the averaged over

TABLE I
PARAMETERS INFORMATION FOR THE PROPOSED MODEL.

model	**Vibration DBN**	**Current DBN**	**Fusion DBN**
structure	4500-2000-1000-500-250-100-3	2000-1000-500-250-100-3	200-100-45-12-3
learning rate	0.001	0.001	0.001
momentum	0.9	0.6	0.6
number of epochs	100	100	100

TABLE II
COMPARATIVE RESULTS OF SINGLE SENSOR-BASED METHODS AND THE PROPOSED METHOD.

Methods	Vibration	Current	Proposed method
Classification accuracy (%)	64.28	88.5	98.69

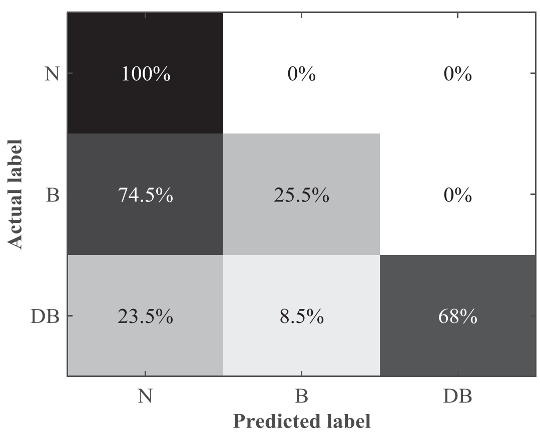

Figure 5. Confusion matrix obtained based on **Vibration DBN** with vibration signals only.

10 experiments. From Table II, the proposed method significantly outperformed both single sensor based methods, with the classification accuracy of 98.69%. This result suggested that fusion of different sensor signals can greatly improve the classification performance which may attributes to the complementary diagnosis information contained different sensor signals. Specifically, the classification accuracy of single vibration signals is 64.28% and the classification accuracy of the original current signal is 88.5% as it is less affected by the environment with a higher signal-to-noise ratio than vibration signals.

In order to demonstrate the detailed information of classification results, we use the confusion matrix to visualize the fault classification accuracy of three different heath conditions of the gearbox, and the results of Vibration DBN with vibration signals only, Current DBN with current signals only and the proposed multimodal DBN method, are shown in Figures 5-7. It can be seen from Figure 5 that the vibration signals under normal conditions can be completely identified. Only 25.5% in the data samples of single-gear tooth break are predicted correctly, and the remaining 74.5% are classified as the normal condition, which greatly affects the overall classification accuracy. However, 68% of the double-gear tooth broken fault are correctly predicted, while 23.5% and 8.5% of the samples are classified as the normal condition and the

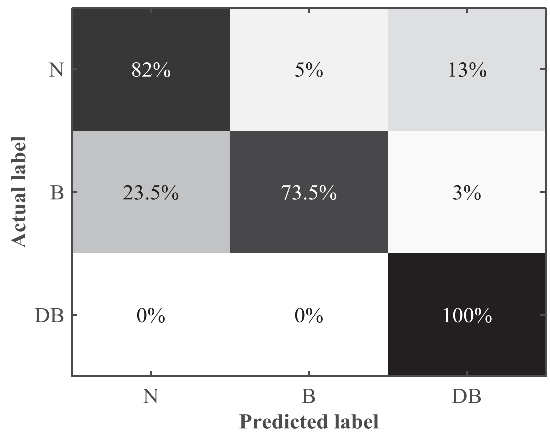

Figure 6. Confusion matrix obtained based on **Current DBN** with current signals only.

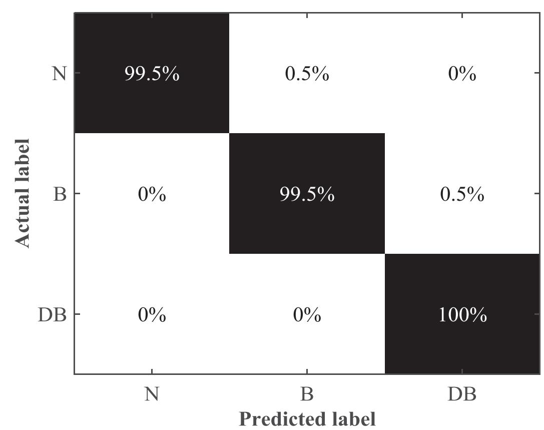

Figure 7. Confusion matrix obtained based on the proposed multimodal method with the fusion of vibration signals and current signals.

single-gear tooth broken fault condition, respectively. As can be seen from Figure 6, **current DBN** misclassifies 13% and 5% of normal samples as the double-shaft tooth broken and single-shaft tooth broken conditions, respectively. Moreover, 23.5% of single-gear tooth broken are classified as the normal ones, and the remaining 3% are classified as the double-gear tooth broken condition. Compared with Figure 5-6, the classification accuracy for each condition of the proposed multimodal DBN is obviously increased. The classification accuracy of N and B is 99.5% and that of DB reaches 100%.

V. CONCLUSIONS

This paper has presented a multimodal deep learning approach to fuse the gearbox vibration and the generator current signals for the wind turbine gearbox fault diagnosis. The proposed method can first learn useful fault-related features automatically from the vibration signals and current signals through the designed **Vibration DBN** and **Current DBN**, respectively. Then the learned vibration and current features are further fused through the designed **Fusion DBN** to output the final diagnosis result. The proposed method can effectively can take full advantage of the complementary diagnosis information contained in different sensor signals and enhance fault diagnosis performance. It has been shown that the proposed method achieved better classification performance compared with those single sensor signal based method, which can provide a more reliable diagnosis result in practical applications.

References

[1] G. Helbing and M. Ritter, "Deep learning for fault detection in wind turbines," *Renewable and Sustainable Energy Reviews*, vol. 98, pp. 189 – 198, 2018.

[2] Y. Lei, D. Kong, J. Lin, and M. J. Zuo, "Fault detection of planetary gearboxes using new diagnostic parameters," *Measuremente & Technology*, vol. 23, no. 5, pp. 55 605–55 614(10), 2012.

[3] J. Hou, W. Jiang, and W. Lu, "Application of a near-field acoustic holography-based diagnosis technique in gearbox fault diagnosis," *Journal of Vibration & Control*, vol. 19, no. 1, pp. 3–13, 2013.

[4] D. Lu, W. Qiao, and X. Gong, "Current-based gear fault detection for wind turbine gearboxes," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1453–1462, 2017.

[5] P. Guo, D. Infield, and X. Yang, "Wind turbine generator condition-monitoring using temperature trend analysis," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 124–133, 2011.

[6] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal Processing*, vol. 96, no. 5, pp. 1–15, 2014.

[7] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems & Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.

[8] G. Q. Jiang, P. Xie, X. Wang, M. Chen, and Q. He, "Intelligent fault diagnosis of rotary machinery based on unsupervised multiscale representation learning," *Chinese Journal of Mechanical Engineering*, vol. 30, no. 6, pp. 1–11, 2017.

[9] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilberthuang transform, support vector machine, and regression," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, 2014.

[10] G. Jiang, H. He, P. Xie, and Y. Tang, "Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis," *IEEE Transactions on Instrumentation & Measurement*, vol. 66, no. 9, pp. 2391–2402, 2017.

[11] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Transactions on Instrumentation & Measurement*, vol. 66, no. 7, pp. 1693–1702, 2017.

[12] D. Lu, X. Gong, and W. Qiao, "Current-based diagnosis for gear tooth breaks in wind turbine gearboxes," in *Energy Conversion Congress and Exposition*, 2012, pp. 3780–3786.

[13] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013.

[14] H. Seraji and N. Serrano, "A multisensor decision fusion system for terrain safety assessment," *IEEE Transactions on Robotics*, vol. 25, no. 1, pp. 99–108, 2009.

[15] X. Liu, L. Ma, and J. Mathew, "Machinery fault diagnosis based on fuzzy measure and fuzzy integral data fusion techniques," *Mechanical Systems & Signal Processing*, vol. 23, no. 3, pp. 690–700, 2009.

[16] J. Hang, J. Zhang, and M. Cheng, "Fault diagnosis of wind turbine based on multisensors information fusion technology," *Iet Renewable Power Generation*, vol. 8, no. 3, pp. 289–298, 2014.

[17] B. S. Yang and K. J. Kim, "Application of dempstershafer theory in fault diagnosis of induction motors using vibration and current signals," *Mechanical Systems & Signal Processing*, vol. 20, no. 2, pp. 403–420, 2006.

[18] C. Li, R. V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vsquez, "Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals," *Mechanical Systems & Signal Processing*, vol. 76-77, pp. 283–293, 2016.

[19] M. Xia, T. Li, L. Xu, L. Liu, and C. W. De Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 101–110, 2017.

[20] M. Ma, C. Sun, and X. Chen, "Deep coupling autoencoder for fault diagnosis with multimodal sensory data," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1137–1145, 2018.

[21] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[22] Y. Bengio, "Deep learning of representations for unsupervised and transfer learning," *Workshop on Unsupervised & Transfer Learning*, vol. 7, pp. 1–20, 2012.

[23] H. Larochelle, M. Mandel, R. Pascanu, and Y. Bengio, "Learning algorithms for the classification restricted boltzmann machine," *Journal of Machine Learning Research*, vol. 13, no. 1, pp. 643–669, 2012.

[24] L. Jing, T. Wang, M. Zhao, and P. Wang, "An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox," *Sensors*, vol. 17, no. 2, p. 414, 2017.

[25] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," *Reliability Engineering & System Safety*, vol. 115, no. 7, pp. 124–135, 2013.

[26] M. Gan, C. Wang, and C. Zhu, "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*, vol. 72-73, no. 2, pp. 92–104, 2016.

[27] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2727–2736, 2017.