

Research Article

Machine-Learning-Based Intelligent Mechanical Fault Detection and Diagnosis of Wind Turbines

Qiang Gao ¹, Xinhong Wu ², Junhui Guo ¹, Hongqing Zhou ¹ and Wei Ruan ³

¹State Grid Zhejiang Electric Power Co., Ltd., Taizhou Power Supply Company, Taizhou 318000, China

²State Grid Zhejiang Integrated Energy Service Company, Hangzhou 310007, China

³College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China

Correspondence should be addressed to Wei Ruan; 2191724107@qq.com

Received 19 March 2021; Revised 3 April 2021; Accepted 15 April 2021; Published 27 April 2021

Academic Editor: Bo Yang

Copyright © 2021 Qiang Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wind power has gained wide popularity due to the increasingly serious energy and environmental crisis. However, the severe operational conditions often bring faults and failures in the wind turbines, which may significantly degrade the security and reliability of large-scale wind farms. In practice, accurate and efficient fault detection and diagnosis are crucial for safe and reliable system operation. This work develops an effective deep learning solution using a convolutional neural network to address the said problem. In addition, the linear discriminant criterion-based metric learning technique is adopted in the model training process of the proposed solution to improve the algorithmic robustness under noisy conditions. The proposed solution can efficiently extract the features of the mechanical faults. The proposed algorithmic solution is implemented and assessed through a range of experiments for different scenarios of faults. The numerical results demonstrated that the proposed solution can well detect and diagnose the multiple coexisting faults of the operating wind turbine gearbox.

1. Introduction

In recent decades, wind power generation has become one of the dominating sustainable and renewable energy sources and achieved high popularity and significant expansion [1]. Due to the increasing CO₂ emissions and the ever-increasing demand for energy, renewable power generation sources have been considered as one of the most important means to alleviate the urgent power demand in a low-carbon fashion. In practice, the deployment of wind turbines is generally located in some rural areas, remote sites, or even offshore. This brings about direct challenges for the inspection and maintenance of wind power generation infrastructures. Due to the harsh operating environment and extreme working conditions, wind turbines are prone to various faults, resulting in high maintenance efforts [2]. Gearbox, being a critical transmission component in wind turbines, shows a high fault rate in practice. The statistics provided in [3] indicated that 76% of the failures are observed at the bearings, and the faults in gears and other components are

17.1% and 6.9%, respectively. Therefore, efficient health condition monitoring is of paramount importance to reduce maintenance costs as well as downtimes for reliable system operation and low economic losses [4, 5].

In practice, the mechanical fault detection and diagnosis of wind turbines in large-scale wind farms is a nontrivial task due to the operating conditions and complex operational characteristics. In general, the manual condition monitoring methods were based on signal processing and analysis, which required skilled expert knowledge and consequently leading to time-consuming remedies [6]. In the literature, to effectively address the technical challenge of fault detection and analysis, the solutions of operational condition monitoring and online fault detection and diagnosis have been investigated (e.g., [7–10]).

It can be observed from the literature that the intelligent fault diagnosis process generally consists of three key steps: (1) signal acquisition by field sensors; (2) feature extraction through the signal preprocessing; and (3) fault classification using machine-learning algorithms, regarded collectively as

a typical pattern recognition problem. During the process of signal acquisition, various signals, including the measurements of vibration signals, acoustic emission signals, current, and infrared thermal, can be acquired from the field [11]. In particular, the condition monitoring and diagnosis through the analysis of vibration signals has been considered useful for the rotating machines as the vibration signal contains the most intrinsic characteristics related to health conditions [12]. During the process of the feature extraction, it is required to understand the characteristics of signals collected from various field sensors and generate the proper sensitive features to address specific diagnosis issues. As suggested by the existing studies (e.g., [13]), the extraction of intrinsic features from the available vibration signals can be generally carried out through the analysis in the time domain, frequency domain, and time-frequency domain. Time-domain features include some statistical parameters, for example, the root mean square, variance, kurtosis, skewness, impulse factor, and crest factor. For the analysis in the frequency domain, the fast Fourier transform (FFT) is a widely used technique for signal processing, while the short-time Fourier transform (STFT) and wavelet packet transform (WPT) is often adopted for the analysis of nonstationary signals [14]. Also, some other advanced techniques have been adopted for the time-frequency analysis, for example, local mean decomposition (LMD), ensemble empirical mode decomposition (EEMD), Hilbert–Huang transform (HHT), and sparse representation [15–18]. For fault classification, many machine-learning-based solutions have been developed in previous studies, such as support vector machine (SVM) [19], random forest (RF) [20], k -nearest neighbor (k -NN) [21], and artificial neural networks (ANN) [22]. The existing methods generally combine advanced signal processing techniques and machine-learning-based approaches to obtain the sensitive features from the collected field signals and make the diagnostic decisions efficiently, achieving remarkable results in various fault diagnosis applications. In [23], the statistical parameters were extracted and characterized from the measured vibration signals and the roller bearing operation conditions were analyzed based on an SVM-based approach to recognize the. The authors in [24] adopted the energy entropy of empirical mode decomposition (EMD) as the input feature and used the ANN for fault classification. The authors in [25] presented an algorithmic solution to calculate the multiscale permutation entropy of vibration signals that are further used as the inputs for the SVM classifier to carry out the bearing fault diagnosis. In [26], the authors proposed to integrate the ensemble empirical mode decomposition with the independent component analysis (ICA) for the classification of the component faults for carrying out the diagnosis of bearing failures in wind turbines.

The conventional machine-learning-based approaches have obtained great success in fault diagnosis. However, there are still a set of limitations that can be summarized as follows: (1) the feature extraction is generally needed for every specific fault diagnosis task. This demands expert knowledge of signal processing and is generally time-consuming; (2) some features may be redundant and it is very

difficult to select the most sensitive features without prior knowledge; and (3) the shallow architectures of learning models, for example, ANN, may potentially limit the capability of the classifiers to learn sophisticated nonlinear relationship during complicated fault scenarios. To overcome existing limitations, more advanced deep learning-based techniques need to be further investigated and assessed.

In recent years, deep learning-based techniques that can automatically and efficiently learn the high-level feature representations from the raw input data have been developed rapidly (e.g., [27, 28]). Also, a collection of deep learning methods, such as deep belief network (DBN) [29], sparse autoencoder (SAE) [30], denoising autoencoder (DAE) [31], sparse filtering [32], convolutional neural network (CNN) [33], and recurrent neural network (RNN) [34], have shown extraordinary feature learning capacity and significant diagnosis results in mechanical fault diagnosis. In [33], the study calculated 256 statistic features and reshaped them into a square matrix and then adopted 2D CNN for gearbox fault diagnosis. In [35], the authors applied various time-frequency methods including short-time Fourier transform, discrete wavelet transform, and S transform to generate time-frequency images and fed them into a CNN to evaluate the performance. The authors in [36] converted a 1D raw vibration signal into 2D vibration grey-scale image directly, utilized LeNet-5-based CNN for data-driven intelligent fault diagnosis of mechanical equipment, and resulted in remarkable accuracy. Besides, a one-dimensional CNN structure has also attracted significant attention due to its straightforward fitness in coping with time-series analysis. The solutions proposed in [37, 38] have adopted the 1D CNN for the fault detection and analysis of real-time signals through capturing the low-frequency features.

Based on the observation and analysis of the existing literature, there are still some challenges that are summarized as follows:

- (1) The early fault signal is very faint, easy to be interfered with by the environmental noise, even submerged. Thus, the fault signal may have a low signal-noise ratio, which is a challenge for fault prognostics.
- (2) Various faults may occur simultaneously which is referred to as multifault and must be taken into consideration.

This work attempts to address the aforementioned problems and proposed a deep learning-based intelligent fault detection and analysis solution of wind turbine gearboxes using the linear discriminant convolutional neural network. A linear discriminant criterion-based linear discriminating loss is introduced into the loss function to enhance the discriminative power of learned features during the training process. The enhanced discriminating power is a promising factor to strengthen the generalization capability and robustness against noise. The experimental results demonstrated that the proposed LDCNN solution can not

only efficiently extract the features from vibration measurements but also provide accurate fault detection and diagnosis in comparison with the existing solutions.

In this work, compared with the aforementioned existing solutions, the main technical contributions made in this work are described as follows:

- (1) A deep learning-based solution is proposed and implemented to extract and learn the various forms of fault features and characteristics from the vibration measurements.
- (2) The linear discriminant loss is introduced and integrated into the loss function to address the technical challenge of diagnosing slight mechanical faults under a noisy environment. This effectively minimizes the negative impact of noises in the process of fault detection and classification.
- (3) The proposed linear discriminant convolutional neural network (LDCNN) solution is extensively evaluated through experiments for a range of fault scenarios through experimental studies based on simulation signals, bearing testbed, and realistic measurements. It demonstrated that the proposed LDCNN can provide a diagnosis accuracy of 98.75%.

The rest of this work is organized as follows: the convolutional neural network is firstly overviewed in Section 2. Section 3 provides a detailed description of our proposed fault diagnosis scheme. Section 4 presents case studies and the analysis of experimental results. Finally, future research directions and conclusive remarks are given in Section 5.

2. Preliminaries and Models

2.1. Overview of CNN. The convolutional neural network (CNN) is considered a multilayer neural network (convolutional layer, pooling layer, and fully connected layer) with a deep supervised learning architecture and has demonstrated prominent performance in many pattern recognition tasks.

The convolutional layer performs a convolution of the input local regions with filter kernels and then generates the output features by the activation unit. Every kernel has the same size. We use W_i^l and b_i^l to denote the weights and bias of the i -th filter kernel in the layer l , respectively, and use a_j^l to denote the j -th region in layer l . The convolutional process can be described as follows:

$$a_{ij}^{l+1} = f(W_i^l \cdot a_j^l + b_i^l), \quad (1)$$

where a_{ij}^{l+1} represents the input of j -th neuron in frame i of layer $l+1$ and f denotes the activation function that may include the sigmoid function, rectified linear units (ReLU), and others.

The max-pooling layer is one of the most commonly used layers on CNN. Let M_j denote the j -th pooling window, the max-pooling transformation then can be formulated as follows:

$$p_i^l(j) = \max_{k \in M_j} (a_i^l(k)). \quad (2)$$

For classification tasks, the deep neural network is followed by a classifier, for example, a multilayer perceptron (MLP), a softmax classifier, or a support vector machine (SVM). Here, the softmax function-based classifier is adopted in the proposed solution. The output of the softmax classifier for a given dataset of k classes can be obtained based on the probability distribution function, as given in the following:

$$\sigma(x)_i = \frac{e^{W_i x + b_i}}{\sum_i^K e^{W_i x + b_i}}, \quad \text{for } i = 1, 2, \dots, k. \quad (3)$$

2.2. LDA-Based Distance Metric Learning. A reasonable distance metric between classes is usually helpful to improve the efficiency and accuracy in pattern recognition. Metric learning is widely used in machine-learning tasks due to its capability to learn suitable metrics from training data automatically. In terms of feature learning, metric learning can learn a new discriminative feature space by feature transformation [39]. Given the training dataset $D = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, with label $y_i \in \{1, 2, \dots, c\}$ and using D_i to represent the set of training samples in i -th class, the distance between two samples under projection is formulated in

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i, \mathbf{x}_j)^T \mathbf{W}^T \mathbf{W} (\mathbf{x}_i - \mathbf{x}_j)}. \quad (4)$$

Here, \mathbf{W} is the projection matrix and the symmetric matrix $\mathbf{W}^T \mathbf{W}$ is called a Mahalanobis distance metric. Distance metric learning aims to learn the matrix \mathbf{W} automatically.

The linear discriminant analysis (LDA) has been widely used in dimension reduction and pattern recognition applications [40]. It enables identifying the linear transformation matrix \mathbf{W} that is called the Fisher criterion. Let \mathbf{S}_b denote the interclass scatter matrix and \mathbf{S}_w denote the intraclass scatter matrix, as described by equations (5)–(7). Then, the Fisher criterion function can be expressed as equation (8):

$$\mathbf{S}_w = \sum_{i=1}^c \sum_{\mathbf{x} \in D_i} (\mathbf{x} - \boldsymbol{\mu}_c) (\mathbf{x} - \boldsymbol{\mu}_c)^T, \quad (5)$$

$$\mathbf{S}_b = \sum_{i=1}^c N_i (\boldsymbol{\mu}_i - \boldsymbol{\mu}) (\boldsymbol{\mu}_i - \boldsymbol{\mu})^T, \quad (6)$$

$$\boldsymbol{\mu}_i = \frac{1}{N_i} \sum_{\mathbf{x} \in D_i} \mathbf{x}, \quad \boldsymbol{\mu} = \frac{1}{N} \sum_{\mathbf{x} \in D} \mathbf{x}, \quad (7)$$

$$J_F(\mathbf{W}) = \max_{\mathbf{W}} \frac{\det(\mathbf{W}^T \mathbf{S}_b \mathbf{W})}{\det(\mathbf{W}^T \mathbf{S}_w \mathbf{W})}. \quad (8)$$

The linear transformation matrix W obtained by the fisher criterion could be used as a distance metric. As to deep learning tasks, better clustering of each class and separability between classes of the extracted features is promising to strengthen the diagnostic capability. Motivated by the metric learning and fisher criterion, this work introduces the linear discriminating loss to address fault diagnosis tasks that are described in detail in the following sections.

3. Proposed Method for Fault Diagnosis

This section developed a deep learning-based fault diagnosis solution for wind turbines, as illustrated in Figure 1. The overall procedure of the proposed linear discriminant convolutional neural network- (LDCNN-) based solution can be described as follows: firstly, the data argumentation is carried out to increase the sample numbers by sampling the raw signals with overlap. Through such data argumentation, a sufficient number of data samples can be obtained for the training process. Then, the LDCNN is designed based on one-dimensional CNN combining improved loss function to learn the highly abstracted features from the input data in an automatic fashion. Finally, the fault classification is carried out through a softmax classifier for different operating conditions with a range of fault scenarios. The implementation details of the proposed solution, including the data sample argumentation, feature extraction, and fault classification, are further explained in the following subsections.

3.1. Sample Acquisition. The raw signal measurements can be collected by the field accelerometers, and the overlap sampling is used to produce more samples for the training and testing processes, as illustrated in Figure 1. In this work, the data augmentation through slicing the original data with overlap is used to produce a sufficient number of training samples, as suggested in [33]. Here, the segment length and overlapping rate are set as 1024 and 0.3, respectively.

3.2. The Architecture of the Proposed Model. There are no deterministic criteria for the CNN structure and parameters for the fault diagnosis in the existing literature. To reduce the network complexity and enhance the efficiency of the proposed model, this work simplifies the network structure through rigorous experiments. In this work, the proposed model consists of three pairs of one-dimensional convolution and pooling layers and two fully connected layers. The wide kernels [33] are adopted in the first convolutional layer. The numbers of nodes of fully connected layers are set as 512 and 128, respectively.

3.3. Loss Function Design. In fact, the adopted loss function in the proposed solution combines both softmax loss and the proposed linear discriminant loss, and hence it can improve the classification efficiency. The target distribution and the estimated distribution are denoted as $p(x)$ and $q(x)$, respectively. The cross-entropy between them can be defined as

$$H(p, q) = - \sum_x p(x) \log q(x). \quad (9)$$

Specifically, combining equation (3), the softmax loss can be written as

$$L_s = - \sum_{i=1}^m \log \frac{e^{W_{y_i} x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j x_i + b_j}}. \quad (10)$$

To simultaneously minimize the intraclass and maximize the interclass variations of the extracted features, a linear discriminant loss function is proposed and implemented based on the distance metric defined in

$$L_{ld} = \frac{D_w}{D_b}, \quad (11)$$

Here, L_{ld} is defined as the ratio of intraclass variations and interclass variations, and the metric of intraclass and interclass variations can be calculated as follows:

$$D_w = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2, \quad (12)$$

$$D_b = \frac{1}{2} \sum_{i=1}^m \|x_{y_i} - c\|_2^2, \quad (13)$$

where x_i denotes the features generated from the top layer, c denotes the center of feature space for all samples, and c_{y_i} denotes the center of samples that belong to class y_i in the feature space. In this way, the weights of the top layer can be regarded as a linear transformation matrix and optimized in training.

The loss function can be written in the following form:

$$L = L_s + \alpha \cdot L_c. \quad (14)$$

Here, L_s is the softmax loss, L_c represents the linear discriminant loss for each epoch, and α is a hyperparameter used for balancing the two parts of the loss function.

3.4. Vibration Signal and LDCNN-Based Fault Diagnosis. Figure 1 shows that the sliced one-dimensional vibration signal segment is fed directly into the one-dimensional CNN for feature learning and classification. In the forward-propagation, features are extracted by successive convolutional and pooling layers followed by the fully connected layers, and the feature classification is carried out by using a softmax classifier. In the backpropagation, the model is optimized by minimizing the improved loss function based on the stochastic gradient descent (SGD) algorithm. After the training process, the testing dataset is used to evaluate the proposed solution, and the accuracy of fault detection and diagnosis is used as the main metric in the evaluation.

4. Experimental Assessment and Result Analysis

The developed fault detection and diagnosis solution is extensively assessed through a range of experiments. The experiments are carried out based on the 10-fold cross-

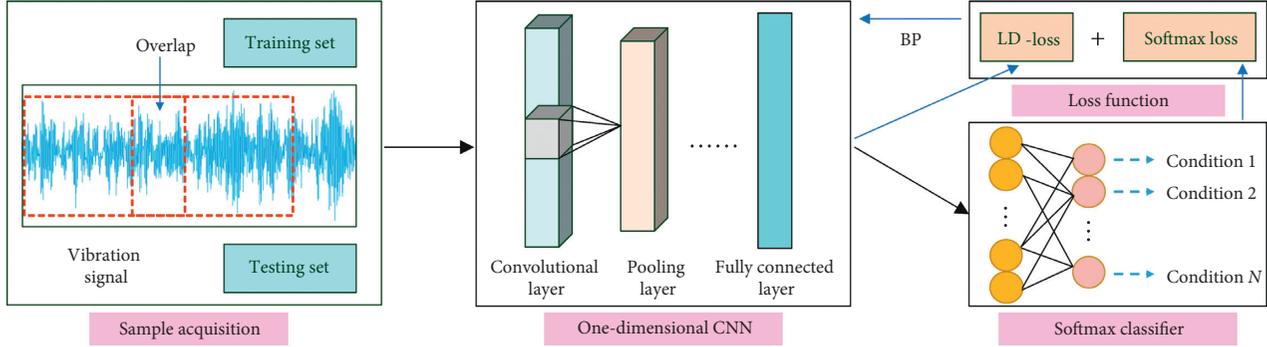


FIGURE 1: Flowchart of the proposed fault detection and diagnosis method.

validation and using a PC with the hardware configuration of Intel Core i5 CPU, 12 GB RAM, and GEFORCE GTX 940M GPU.

4.1. Case 1: Validation with Simulation Signals. This work firstly carries out the performance evaluation based on the simulated signals as adopted in [41]. The simulated signals can be expressed as

$$x(t) = \sum_i A_i s(t - T_i) + \sum_k B_k \cos(2\pi k f + \phi_k) + n(t). \quad (15)$$

Here, A_i is the amplitude of the i^{th} impulse excited by the defect and T_i is the time of its occurrence, B_k and ϕ_k are the amplitude and initial phase of the k^{th} harmonic caused by bearing imbalance or gear meshing; and $n(t)$ is the white noise in the measurement.

The vibration signals under both normal and fault scenarios are considered. Also, two vibration harmonics are generated under normal conditions. The simplified mathematical expressions are given by

$$\begin{aligned} s_1 &= 0.7 \sin(2\pi f_1 t), \\ s_2 &= 0.3 \sin(2\pi f_2 t), \end{aligned} \quad (16)$$

where $f_1 = 200$ Hz and $f_2 = 400$ Hz represent two gear vibration harmonic frequencies, t represents the time, and $t = n/f_s$ where n is sample point and f_s is sampling frequency which is kept 10 kHz in this experiment.

In addition, the signals for both outer-race and inner-race fault scenarios are simulated as follows [41, 42]:

$$\begin{aligned} s_3 &= 2e^{-200\pi t_1} \sin(2\pi f_{ro} t), \\ s_4 &= 1.5(\cos(2\pi f_o t) + 1) \cdot e^{-500\pi t_2} \sin(2\pi f_{ri} t), \end{aligned} \quad (17)$$

$$t_1 = \text{mod}\left(\frac{t, 1}{f_i}\right), \quad t_2 = \text{mod}\left(\frac{t, 1}{f_o}\right),$$

where $f_{ro} = 2000$ Hz and $f_{ri} = 3000$ Hz, $f_o = 30$ Hz, and $f_i = 150$ Hz represent the resonance frequency and fault characteristic frequency for the two kinds of faults, respectively. Here, $f_o = 20$ Hz is the bearing shaft rotating frequency. The simulated fault signals of the outer-race and inner-race fault are presented in Figure 2, respectively.

In simulations, the vibration signals can be generated using the aforementioned source signals based on (14), and four kinds of vibration signals are shown in Figure 3.

In this case, in total, 500 samples are used for each class and each sample consists of 400 data points, and hence the dataset contains 2000 (500×4) samples, where 70% of the samples are used for training and the remainder is used for the performance test. In experiments, different solutions are implemented as the comparison benchmarks: the 1D CNN [37], traditional 2D CNN [35], WDCNN with the wide kernel in the first layer [38], and two machine-learning models with 15 time-domain statistical features [43]. Figure 4 presents the numerical results of the evaluated solutions in terms of the detection accuracy based on the test datasets.

In Figure 4, the numerical results demonstrate that the proposed fault detection and diagnosis methods can efficiently detect and diagnose both single and multiple coexisting faults. The superiority of the proposed LDCNN-based solution is confirmed compared with other existing solutions.

4.2. Case 2: Validation Using the Bearing Testbed. The performance of the proposed algorithmic solution is further evaluated using the tested at the Case Western Reserve University Bearing Data Centre [44]. In detail, the CWRU testbed consists of a 2 hp motor, a torque transducer, and a dynamometer, as illustrated in Figure 5. This work considers three types of single mechanical faults, that is, inner race fault (IRF), outer race fault (ORF), and ball fault (BF), which are introduced to the test bearings using electrical discharge machining with different fault diameters. In this study, the bearing data with a fault diameter of 0.007 inches and 0.014 inches are selected and used for performance evaluation. Here, the bearing signals are collected from the drive end of the motor by the use of accelerometers with a sampling rate of 12 kHz under four different load conditions, that is, 0, 1, 2, and 3 hp, respectively.

The data of the aforementioned three fault categories with two fault diameters as well as normal operating condition data are directly adopted to evaluate our proposed solution. There are seven categories in total and for each health condition; 150 samples are generated in each load condition, where each sample consists of 1024 data points in

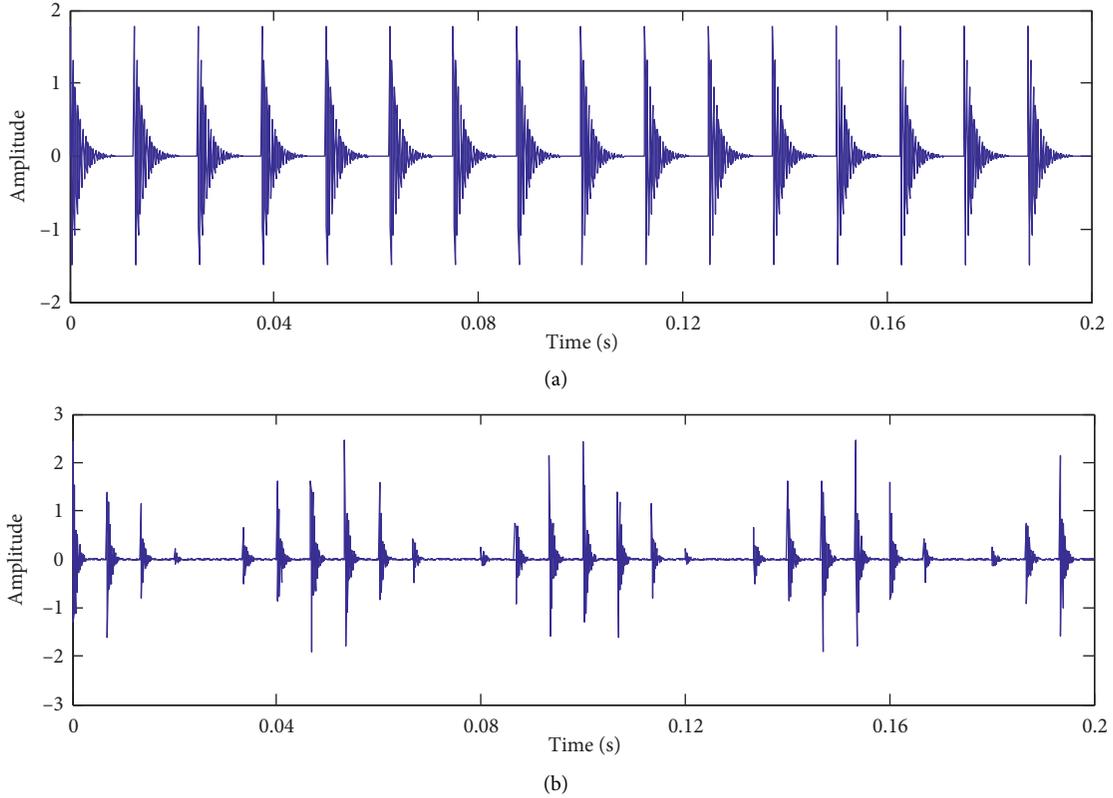


FIGURE 2: Fault signals in simulation experiments: (a) outer race fault; (b) inner race fault.

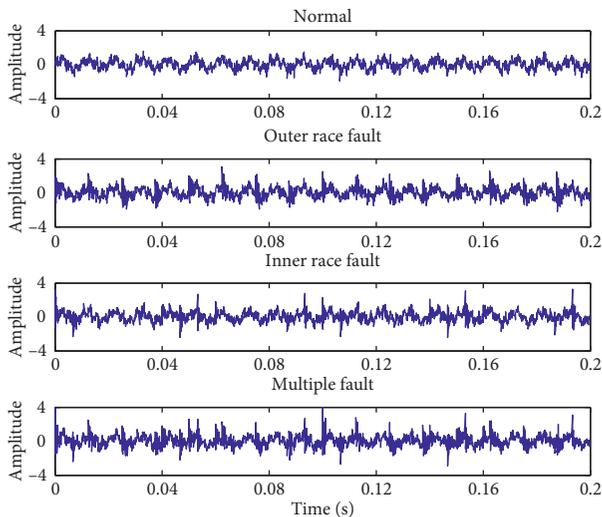


FIGURE 3: Vibration signals of different conditions in simulation experiments.

this work. Therefore, the constructed dataset contains 4200 samples in total, where 70 percent of them are used for training and 30 percent for testing. As in case 1, four deep learning models and two machine-learning models with manual feature extraction are applied and compared. Here, the adopted deep learning models are optimized by stochastic gradient descent (SGD) algorithm with a minibatch size of 64 samples. As a result, the 2D CNN-based method

achieved a high accuracy of 99.52%, while other deep learning models ended with 100%. In contrast, machine-learning-based methods resulted in an accuracy of 94.76% and 95.36%. Hence, the effectiveness and advantages of deep learning can be confirmed due to its higher accuracy despite the lack of manual feature extraction.

In addition, the effectiveness and robustness of the proposed solution are evaluated by adding the Gaussian white noise to the collected data samples. It can be observed that the noisy condition can bring about difficulty for fault detection and diagnosis, particularly in the presence of multiple simultaneously coexisting mechanical faults. Therefore, robustness against noise disturbance is essentially required and needs to be evaluated. In practice, noise-added data are generated for different signal-to-noise ratio (SNR) defined in

$$\text{SNR} = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right). \quad (18)$$

Here, P_{signal} and P_{noise} represent the powers of the original signal and the additive Gaussian white noise, respectively.

In this work, the data measurements are used in the training phase to train the intelligent models and the noise-added samples with various SNR are used for the testing phase. In the training stage, the hyperparameter of the loss function for LDCNN is set as $\alpha=0.2$. Figure 5 shows the experimental results of deep learning methods. Low SNR values represent greater noise power which, as a result,

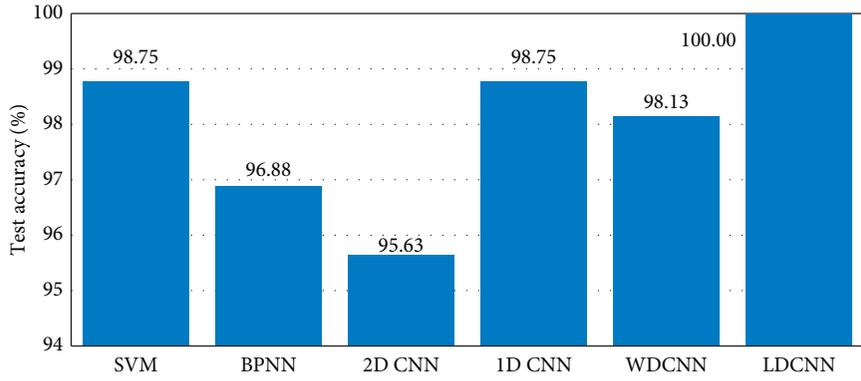


FIGURE 4: The mean detection accuracy of different machine-learning approaches.

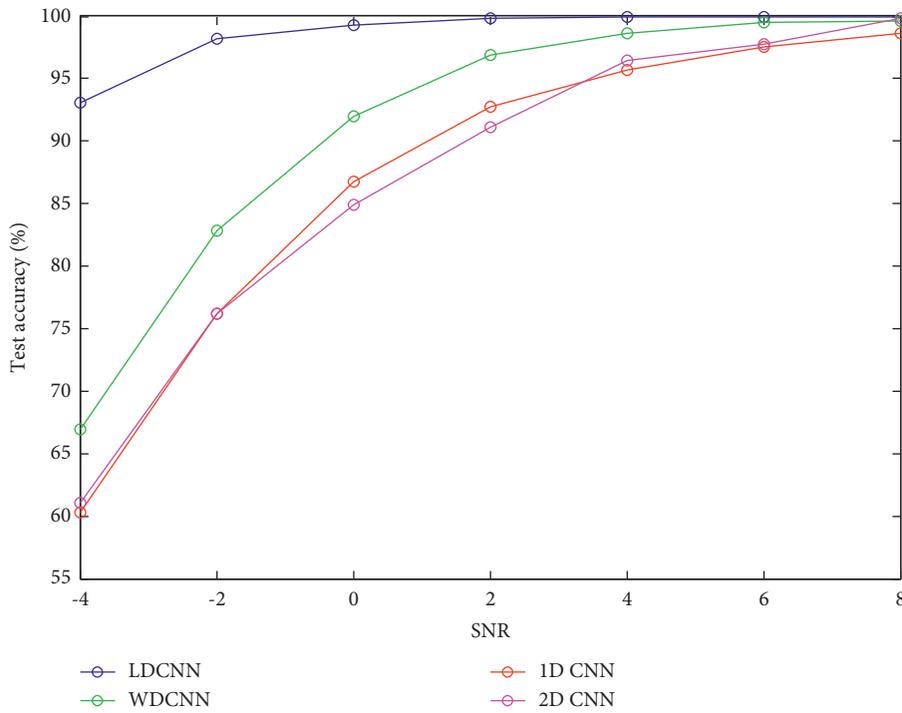


FIGURE 5: Performance of fault detection accuracy under different SNR conditions.

hinders the efficient fault diagnosis. It is evident from Figure 5 that the proposed LDCNN outperforms other intelligent methods with 100% and 93.07% testing accuracy for 8 dB and -4 dB SNR values, respectively.

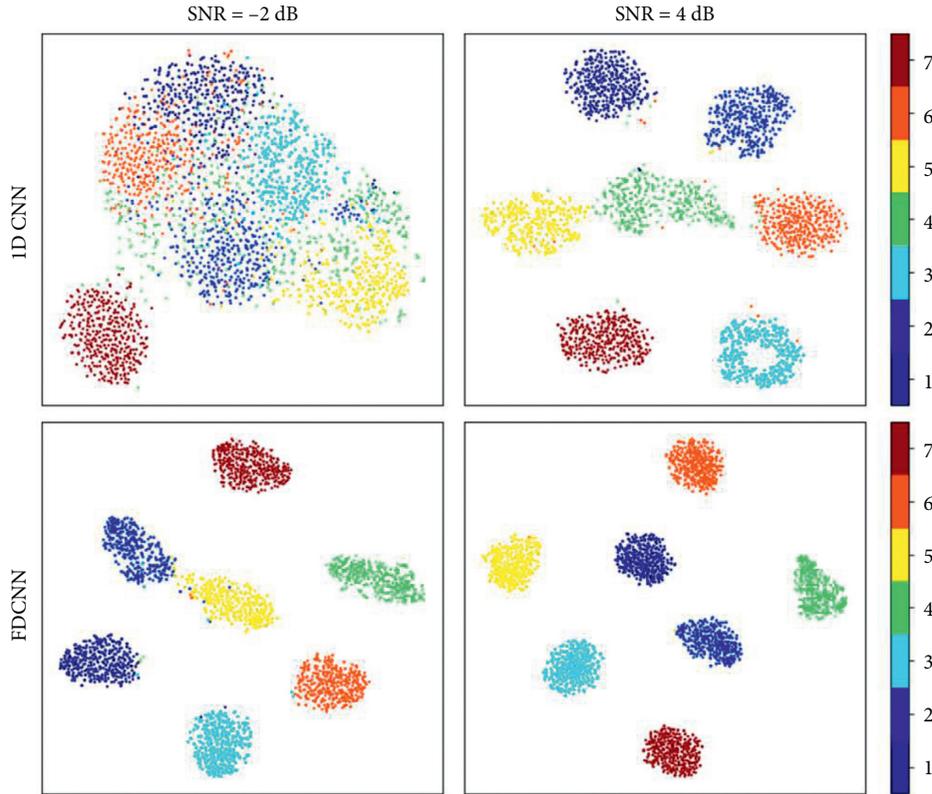
This work adopted the *t*-distributed Stochastic Neighbor Embedding (*t*-SNE) to represent the exacted features. Figure 6 illustrates the extracted high-dimensional features for 7 different conditions under two noisy conditions. The extracted features can be differentiated for SNR = 4 dB, whereas for SNR = -2 dB, features extracted by traditional 1D CNN are heavily overlapped, while features extracted by the proposed LDCNN are fairly distinguishable. This confirmed the performance of the proposed solution in terms of robustness under different noisy conditions.

It should be noted that although the effectiveness of the developed solution is confirmed by the numerical

experiments, the proposed solution for multifault analysis needs to be further evaluated. To assess its performance in multiple fault scenarios, this work utilizes the single fault signals to construct multiple fault vibration signals by employing mixing matrix [45] and nonlinear function [46] given by

$$X = \tan h(A \cdot S), \tag{19}$$

where *S* is the set of single fault signals and *A* and $\tan h(\cdot)$ are the linear mixing matrix and the nonlinear function, respectively. The nonlinear mixture is employed to mimic the vibration signals of realistic multiple faults. This work considers multifault signals based on only two types of faults to eradicate the combinational complexity. Specifically, any two of the three types of single faults with any fault size are selected to form various multiple faults to generate the

FIGURE 6: Extracted features visualization using t -SNE.

dataset. The obtained measurement datasets are given in detail in Table 1. There are in total 13 categories and 7800 samples used in this study. Here, 5460 randomly selected samples and the remaining 2340 samples are used as the training set and testing set, respectively.

Figure 7 provides the detection accuracy of the training dataset and testing dataset for deep learning-based solutions. It is clearly seen from the training and testing results that the proposed LDCNN and 1D CNN have a faster convergence rate as compared to 2D CNN. Moreover, the testing accuracy of LDCNN is higher than 1D CNN, which demonstrates the effectiveness of improved loss function. Considering the fact that the multifault data are manually constructed and have already introduced noises and errors, evaluating noise immunity by adding Gaussian white noises may not be consistent with realistic situations and valueless, which is left out for the present.

4.3. Case 3: Validation with Realistic Data. The proposed method is further evaluated through the adoption of realistic wind farm measurements. In this case, the realistic vibration signals are measured and collected from the operating wind turbines with the sampling rate at 25.6 kHz and the rotating speed of each selected wind turbine is about 1100 rpm. In addition to the bearing fault, the gearbox is also prone to gear failure caused by harsh operation state and extreme environmental conditions. Here, five different conditions are measured from a group of fault-diagnosed wind turbine gearboxes, including the bearing rolling ball fault (B), the

TABLE 1: The selected dataset.

Fault location	Diameter (inch)	Train/test	Label
None	0	420/180	1
Ball	0.007	420/180	2
	0.014	420/180	3
Inner race	0.007	420/180	4
	0.014	420/180	5
Outer race	0.007	420/180	6
	0.014	420/180	7
Ball and inner race	0.007/0.007	420/180	8
	0.007/0.014	420/180	9
Ball and outer race	0.007/0.007	420/180	10
	0.007/0.014	420/180	11
Inner and outer race	0.007/0.007	420/180	12
	0.007/0.014	420/180	13

inner raceway fault (I), the gear fault (G), and the multiple faults compounded by the ball and gear fault (B&G), and the inner race and gear fault (I&G), respectively. The measurements of the accelerometer installed at high-speed shaft bearing on the motor side are recorded. As a result, 300 measurement samples of each operating condition are obtained to set up the measurement dataset containing in total 1500 samples.

The performance of the proposed LDCNN is compared with the existing solutions, and the performance in terms of accuracy is provided in Table 2. It is evident from the numerical results that the developed LDCNN can achieve

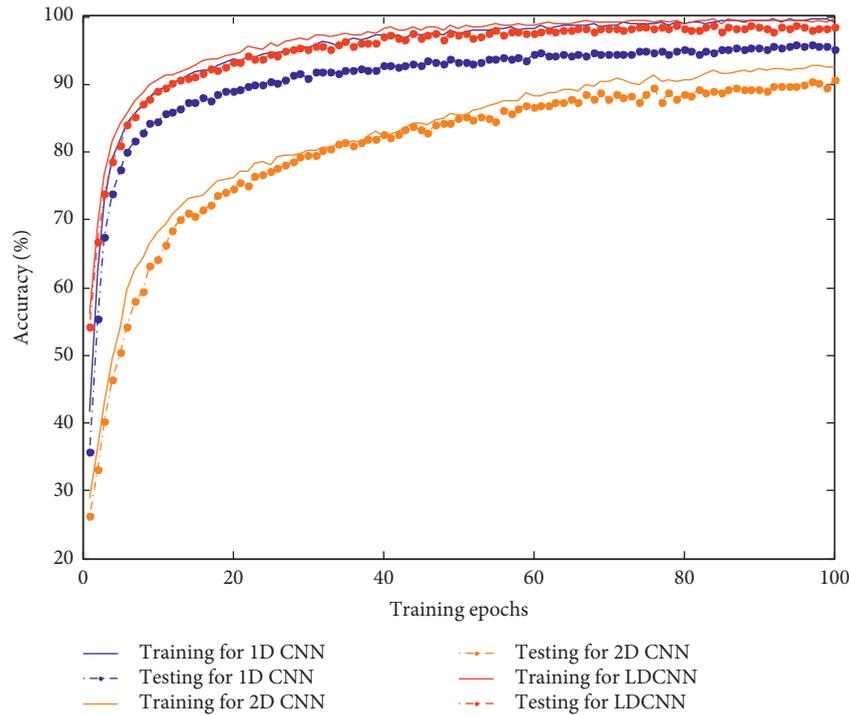


FIGURE 7: Accuracy curve during the training process.

TABLE 2: Method evaluation.

Methods	Mean accuracy (%)
SVM	75.49
BPNN	90.42
2-D CNN	93.57
1-D CNN	90.83
WDCNN	96.56
LDCNN	98.75

of one of the test results is presented in Figure 8. It is noteworthy from the confusion matrix that the multiple faults may be diagnosed as a single fault due to the variability in the size of each single fault. However, such fault detection cannot be ruled out as a misclassification case because multiple fault condition is also one of the fault conditions.

5. Conclusion and Remarks

In this paper, a linear discriminant CNN-based diagnostic solution is proposed for efficient detection and diagnosis of multiple coexisting mechanical faults in the operational wind turbines. The proposed solution is extensively assessed through simulation and experiments. In addition to the accurate performance of diagnosis for multiple faults, the noises immunity of the proposed algorithmic model is enhanced to provide excellent performance under the conditions with a low signal-to-noise ratio.

For future considerations, the following two research directions will be particularly examined. To further facilitate the proposed solution, the scalable data samples argumentation and accurate classification are required. Thus, the effectiveness and efficiency of the proposed solution need to be further validated and extensively assessed with massive data measurements of various other kinds of faults and operating conditions. A physical testbed that can simulate various kinds of bearing and gear faults should be established and used for further research. Also, more advanced data-driven optimization and machine-learning-based techniques can be developed and incorporated into the fault feature characterization and recognition.

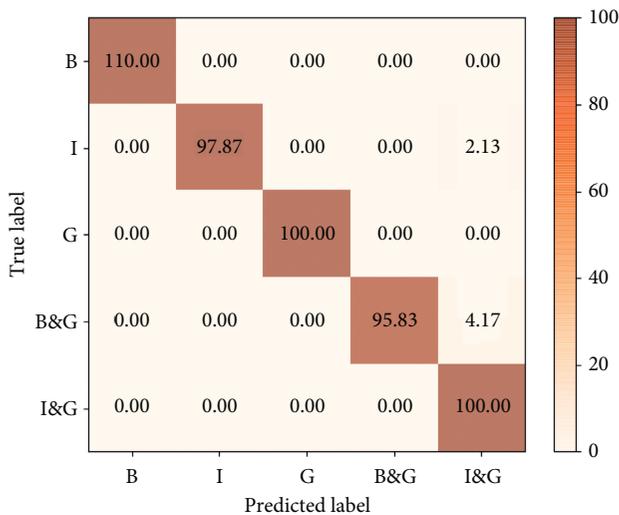


FIGURE 8: Confusion matrix of the proposed LDCNN solution.

the detection accuracy of 98.75%, which outperforms the compared machine-learning and deep learning-based methods. For the proposed solution, the confusion matrix

Data Availability

The experiments are supported by the data source from Case Western Reserve University, Cleveland, Ohio, America, available online at <http://csegroups.case.edu/bearingdatacenter/pages/bearing-information>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Science and Technology Project of State Grid Zhejiang Electric Power Co., Ltd. (research on the evaluation of the distributed renewable energy consumption based on the generation-network-demand coordination in the electricity market).

References

- [1] M. Nie and L. Wang, "Review of condition monitoring and fault diagnosis technologies for wind turbine gearbox," *Procedia CIRP*, vol. 11, pp. 287–290, 2013.
- [2] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis-part I: components and subsystems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6536–6545, 2015.
- [3] National Renewable Energy Laboratory (NREL), *Statistics Show Bearing Problems Cause the Majority of Wind Turbine Gearbox Failures*, National Renewable Energy Laboratory (NREL), Golden, CO, USA, 2015, <http://energy.gov/eere/wind/articles/statistics-show-bearing-problems-cause-majority-wind-turbine-gearbox-failures>.
- [4] X. M. Zhao, Q. H. Hu, Y. G. Lei, and M. J. Zuo, "Vibration-based fault diagnosis of slurry pump impellers using neighbourhood rough set models," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 224, no. 4, pp. 995–1006, 2010.
- [5] Y. G. Lei and Z. J. He, "Advances in applications of hybrid intelligent fault diagnosis and prognosis technique," *Journal of Vibration and Shock*, vol. 30, no. 9, pp. 129–135, 2011.
- [6] Z. Hameed, Y. S. Hong, Y. M. Cho, S. H. Ahn, and C. K. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: a review," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 1, pp. 1–39, 2009.
- [7] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis-part II: signals and signal processing methods," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6546–6557, 2015.
- [8] Z. Zhang, A. Verma, and A. Kusiak, "Fault analysis and condition monitoring of the wind turbine gearbox," *IEEE Transactions on Energy Conversion*, vol. 27, no. 2, pp. 526–535, 2012.
- [9] B. Lu, Y. Li, X. Wu, and Z. Yang, "A review of recent advances in wind turbine condition monitoring and fault diagnosis," in *Proceedings of the 2009 IEEE Power Electronics and Machines in Wind Applications*, pp. 1–7, Lincoln, NE, USA, June 2009.
- [10] W. Y. Liu, B. P. Tang, J. G. Han et al., "The structure healthy condition monitoring and fault diagnosis methods in wind turbines: a review," *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 466–472, 2014.
- [11] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: techniques and methods," *Renewable Energy*, vol. 46, pp. 169–178, 2012.
- [12] J. Zarei, M. A. Tadjedini, and H. R. Karimi, "Vibration analysis for bearing fault detection and classification using an intelligent filter," *Mechatronics*, vol. 24, no. 2, pp. 151–157, 2014.
- [13] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [14] H. O. A. Ahmed, M. L. D. Wong, and A. K. Nandi, "Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features," *Mechanical Systems and Signal Processing*, vol. 99, pp. 459–477, 2018.
- [15] H. Liu and M. Han, "A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings," *Mechanism and Machine Theory*, vol. 75, pp. 67–78, 2014.
- [16] X. Zhang, Y. Liang, J. Zhou, and Y. Zang, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," *Measurement*, vol. 69, pp. 164–179, 2015.
- [17] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, 2015.
- [18] L. Ren, W. Lv, S. Jiang, and Y. Xiao, "Fault diagnosis using a joint model based on sparse representation and SVM," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2313–2320, 2016.
- [19] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2560–2574, 2007.
- [20] D. Cabrera, F. Sancho, R.-V. Sánchez et al., "Fault diagnosis of spur gearbox based on random forest and wavelet packet decomposition," *Frontiers of Mechanical Engineering*, vol. 10, no. 3, pp. 277–286, 2015.
- [21] Z. Hu, L. X. Duan, and L. B. Zhang, "Application of improved KNNC method in fault pattern recognition of rolling bearings," *Journal of Vibration & Shock*, vol. 32, pp. 84–87+105, 2013.
- [22] L. Eren, A. Karahoca, and M. J. Devaney, "Neural network based motor bearing fault detection," in *Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference*, vol. 3, pp. 1657–1660, Como, Italy, May 2004.
- [23] X. Zhang, B. Wang, and X. Chen, "Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine," *Knowledge-Based Systems*, vol. 89, pp. 56–85, 2015.
- [24] Y. Yu, D. J. YuDejie, and C. Junsheng, "A roller bearing fault diagnosis method based on EMD energy entropy and ANN," *Journal of Sound and Vibration*, vol. 294, no. 1-2, pp. 269–277, 2006.
- [25] S.-D. Wu, P.-H. Wu, C.-W. Wu, J.-J. Ding, and C.-C. Wang, "Bearing fault diagnosis based on multiscale permutation entropy and support vector machine," *Entropy*, vol. 14, no. 8, pp. 1343–1356, 2012.
- [26] J. Wang, R. X. Gao, and R. Yan, "Integration of EEMD and ICA for wind turbine gearbox diagnosis," *Wind Energy*, vol. 17, no. 5, pp. 757–773, 2014.

- [27] S. Yin and O. Kaynak, "Big data for modern industry: challenges and trends [point of view]," *Proceedings of the IEEE*, vol. 103, no. 2, pp. 143–146, 2015.
- [28] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [29] H. Shao, H. Jiang, X. Zhang et al., "Rolling bearing fault diagnosis using an optimization deep belief network," *Measurement Science and Technology*, vol. 26, no. 11, p. 115002, 2015.
- [30] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," *Measurement*, vol. 89, pp. 171–178, 2016.
- [31] C. Lu, Z.-Y. Wang, W.-L. Qin, and J. Ma, "Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification," *Signal Processing*, vol. 130, pp. 377–388, 2017.
- [32] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3137–3147, 2016.
- [33] Z. Chen, C. Li, and R.-V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock and Vibration*, vol. 2015, no. 2, 10 pages, Article ID 390134, 2015.
- [34] T. de Bruin, K. Verbert, and R. Babuska, "Railway track circuit fault diagnosis using recurrent neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, pp. 523–533, 2017.
- [35] R. Chen, X. Huang, L. Yang, X. Xu, X. Zhang, and Y. Zhang, "Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform," *Computers in Industry*, vol. 106, pp. 48–59, 2019.
- [36] L. Wen, X. Li, L. Gao, and 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.
- [37] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016.
- [38] W. Zhang, G. Peng, C. Li et al., "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, 2017.
- [39] W. B. Li, J. Huo, Y. H. Shi et al., "Online deep metric learning," 2018, <http://arxiv.org/abs/1805.05510v1>.
- [40] Y. Chen, X. G. Zhao, and J. D. Han, "Distance metric learning with penalized linear discriminant analysis," in *Proceedings of the 2010 IEEE International Conference on Progress in Informatics and Computing*, pp. 170–174, Shanghai, China, December 2010.
- [41] M. Zhao, J. Lin, X. Xu, and X. Li, "multi-fault detection of rolling element bearings under harsh working condition using IMF-based adaptive envelope order analysis," *Sensors*, vol. 14, no. 11, pp. 20320–20346, 2014.
- [42] X. A. Yan, M. P. Jia, and L. Xiang, "Compound fault diagnosis of rotating machinery based on OVMD and a 1.5-dimension envelope spectrum," *Measurement Science and Technology*, vol. 27, no. 7, p. 075002, 2016.
- [43] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1693–1702, 2017.
- [44] Case Western Reserve University. Cleveland, OH, USA [Online]. Available: <http://csegroups.case.edu/bearingdatacenter/pages/bearing-information>.
- [45] H. Hong and M. Liang, "Separation of fault features from a single-channel mechanical signal mixture using wavelet decomposition," *Mechanical Systems and Signal Processing*, vol. 21, no. 5, pp. 2025–2040, 2007.
- [46] W. Li and H. Yang, "A maximum entropy based nonlinear blind source separation approach using a two-layer perceptron network," in *Proceedings of the 2013 10th IEEE International Conference on Control and Automation (ICCA)*, pp. 978–982, Hangzhou, China, June 2013.