IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 21 August 2024, accepted 8 September 2024, date of publication 16 September 2024,
date of current version 30 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3461966

RESEARCH ARTICLE

Intelligent Fault Diagnosis of Wind Turbine Generator Bearings Using Acoustic Signals

BEI ZHAO[1,2,3,4], XIAOMENG LI[5,6,7], ZEDONG LI[2,3], MINLI YOU[2,3], AND FENG XU[2,3]

[1]Department of Physics, Xi'an Jiaotong University City College, Xi'an 710018, China
[2]Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China
[3]Bioinspired Engineering and Biomechanics Center (BEBC), Xi'an Jiaotong University, Xi'an 710049, China
[4]Engineering Research Center of Photovoltaic Technologies and Systems, Universities of Shaanxi Province, Xi'an 710049, China
[5]School of Computer and Network Engineering, Shanxi Datong University, Datong 037009, China
[6]Institute of Big Data Technology and Application, Shanxi Datong University, Datong 037009, China
[7]CRRC Yongji Electric Company Ltd., Xi'an 710018, China

Corresponding author: Xiaomeng Li (lxm096@sxdtdx.edu.cn)

This work was supported in part by the Key Project of Xi'an Jiaotong University City College under Grant 202002Z03, in part by the Potential Cultivation Project of Xi'an Jiaotong University City College under Grant 2023PY01, and in part by the Youth Innovation Team of Shaanxi Universities and the Opening Project of Chongqing Key Laboratory of Photo-Electric Functional Materials under Grant 202001.

ABSTRACT Wind turbine generator bearings are key components in wind power generators, where accurate and convenient fault diagnosis for product predelivery inspection and depot repair inspection is a major challenge to ensure their safe operation. In this study, an artificial-intelligence-based method was developed for bearings fault diagnosis using acoustic signals with convenient capture, collection, and transmission. Specifically, the running sound was used as the input signal; then, several machine learning models and deep learning models were used to analyze the acoustic signals. The performance of deep learning models using vibrational signals collected by acceleration sensors was investigated for comparison. Results show that an accuracy of 99.90% can be achieved by utilizing deep learning models. The developed method will be a powerful tool for accurate and convenient fault diagnosis because it can be easily deployed on embedded devices. Empirical results validate the effectiveness of our proposed method.

INDEX TERMS Acoustic applications, acoustic signal processing, artificial intelligence, convolutional neural networks, deep learning, fault diagnosis, generators, long short term memory, machine learning, recurrent neural networks.

I. INTRODUCTION

The energy revolution is being driven globally by growing energy and environmental challenges. Peaking CO_2 emissions by 2030 and achieving carbon neutrality by 2060 are China's commitments to the United Nations. Therefore, energy conservation, emission reduction, and development of renewable energy sources are essential. Wind energy as a reliable renewable energy source is growing worldwide, which is particularly important because of the global energy shortage, especially in China [1]. According to World Wind Energy Association statistics, the total capacity of all wind farms worldwide has reached 744 GW in 2020, generating 7% of the world's electricity demand. Wind energy is expected to play an increasingly vital role in the future national energy scene. The wind turbine (WT) power capacity and operational cost strongly depend on component failure and repair rates, so it is important to investigate fault diagnosis of WT generator bearings to improve the economic benefits. In addition, WTs with large volume and weight are generally located in remote areas with harsh environments, and the time and economic costs of installation, disassembly, and transportation are high; thus, strict predelivery checks are essential. Moreover, bearings are widely used in wind power gearboxes and generators, but they are easily damaged and have a great impact on cost and reliability [2]. For example, more than 50% of gearbox faults are caused by bearings, so bearing fault diagnosis is particularly important for predictive maintenance [3], [4], [5], [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Gerard-Andre Capolino.

The failure rate of WT generator bearings remains high due to the long continuous operation of mechanical equipment. Accurate fault diagnosis of WT generator bearings can save downtime and provide a clear direction for maintenance, thus reducing the loss of wind energy capacity and improving production efficiency. Therefore, it is essential to use WTs for high-precision and convenient diagnosis of bearing faults [7], [8], [9], [10].

Various signals have been used for traditional bearing fault diagnosis, with signals such as vibration, acoustic emission, mechanical strain, temperature, and torque being the most used. In the past decades, with the development of advanced technologies in various fields, the accuracy of fault diagnosis has significantly improved [11], [12], [13], [14], [15], [16], [17]. However, these conventional signals require complex data acquisition equipment and transmission systems, and their results are susceptible to environmental factors. To achieve high convenience and low software/hardware requirements, running acoustic signals can be used, which can be conveniently acquired and transmitted.

Different signal processing methods have been developed for fault diagnosis of rotating machinery, such as classic statistical analysis, envelope analysis, time-series analysis, Hilbert–Huang transform, fast Fourier transform (FFT) [18], [19], short-time Fourier transform [20], [21], [22], wavelet transform [23], [24], empirical mode decomposition [25], [26], and variational mode decomposition [27], [28]. However, the signal-to-noise ratio (SNR) of feature extraction decreases due to environmental interference, and it is still challenging to extract effective features from the operating acoustic signals of WTs. Also, traditional signal processing methods have certain limitations when applied in complex operating conditions and do not meet the requirements of practical applications. These traditional methods completely rely on exact values at the fault characteristics frequencies to identify the presence of a fault. As a result, it is important and challenging to choose specific features for characterizing the precise signals used for the classification.

Artificial intelligence (AI) technology has shown great potential in improving the accuracy of signal analysis and diagnosis and increasing the generalization ability. Integration of AI into fault diagnosis systems not only has been developed in theory but also has a successful application example [29], [30], [31], [32]. With the development of machine learning techniques, especially deep learning, it is possible to extract effective features from signals in noisy environments using AI models. Deep learning has achieved better results compared to shallow machine learning methods, but the application of deep learning to fault diagnosis is still under development. Several AI architectures were used for fault diagnosis in the previous studies [33]. However, most modern intelligent bearing fault diagnosis technologies raise the cost of the product due to the installation of new vibration or acoustic emission sensors [34]. To counter this issue, acoustic signals can be collected through built-in microphones in devices like smartphones. The technology that uses AI algorithms to analyze acoustic signals and extract useful information has been used successfully in several fields [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45]. However, machine learning algorithms and deep learning algorithms for fault diagnosis of bearings of WT generators using acoustic signals have not been well used. For instance, Manikanta et al. [35] developed a method for effective baby cry sound detection under different kinds of background sounds using deep learning. Lee et al. [37] successfully distinguished different conditions of a grinding wheel by analyzing the audio signals of the grinding process using deep learning. Xie et al. [41] studied snore detection using convolutional neural network (CNN) and recurrent neural network (RNN) and obtained high accuracy and sensitivity. Wang et al. [42] explored a sound signal analysis method based on deep learning and verifies it in a simulation environment. Tang et al. [44] proposed a method for AE event extraction using a short-time autocorrelation function (STAF) to accurately extract the pulse generated by the fault source. Guo et al. [45] proposed a hydraulic motor plunger wears fault diagnosis method based on the fusion of vibration and sound signal features recognized by LightGBM. Therefore, although the application of AI to WT generator bearing fault diagnosis has not been thoroughly explored, the analysis of acoustic signals and the extraction of useful information through AI methods have profound potential [46], [47], [48], [49], [50]. The use of acoustic signals for fault diagnosis, which allows data collection and fault diagnosis with just one embedded device and no additional sensors, has not been thoroughly investigated. Inspired by this, an intelligent acoustic signals-based fault diagnosis was implemented for WT generator bearing.

In this study, several AI methods such as machine learning and deep learning for acoustic fault diagnosis in WT generator bearing operation with the aim of clarifying their advantages and disadvantages were investigated. For comparison, the effectiveness of deep learning models studying vibrational signals was also investigated. Experimental results underscore the effectiveness of the proposed approach. In addition, this approach increases the ease of fault diagnosis processing.

The remainder of this paper is outlined as follows. Section II introduces the machine learning and deep learning models related to WT generator bearing fault diagnosis. The experimental details, the performance of the proposed method, and the comparison of the proposed method to conventional techniques are described in Section III. The conclusions and further works are presented in the final section.

II. RELATED THEORETICAL BASIS

A. MACHINE LEARNING

Machine learning is one of the most active areas of research and application in AI. In recent years, some of the major successes and advances in AI have been closely related to machine learning, such as self-driving cars,

face recognition, personalized recommendation systems, and machine translation. Intuitively, machine learning is about making machines mimic human learning. Machine learning gives computers or machines the ability to learn from data without being explicitly programmed. Machine learning can be categorized into various types; for example, according to the presence or absence of mentors, it can be categorized into supervised, semisupervised, and unsupervised machine learning algorithms, including support vector machines (SVMs), boosted decision trees, k-means, and different clustering methods. Machine learning algorithms use computational methods to ''learn'' information directly from data without relying on a predetermined equation as a model. Algorithms adaptively improve their performance as the number of samples available for learning increases.

B. CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network (CNN) is a biophysical model inspired by the neural mechanism of the visual nervous system. CNN can be thought of as a special type of multilayer perceptron or feedforward neural network with local sensing and weight sharing features, which can run universally on any embedded device. Since the advent of CNN in deep learning, it has been successfully applied to image and video classification and recognition, as well as object localization and detection. CNN is now the most popular approach for image identification and speech recognition.

A typical CNN architecture, as shown in Fig. 1, consists of an input layer, feature extraction blocks (contain one or more convolutional layers followed by a pooling layer), a fully connected layer, and a classification layer. The input layer is usually a numerical matrix, such as an image. From the perspective of feedforward neural networks, convolutional and pooling layers can be considered special hidden layers that can extract features. The fully connected layer computes the category of the input by transforming it into a one-dimensional array. The output layer outputs the recognition results by classifying the objects into their respective classes.

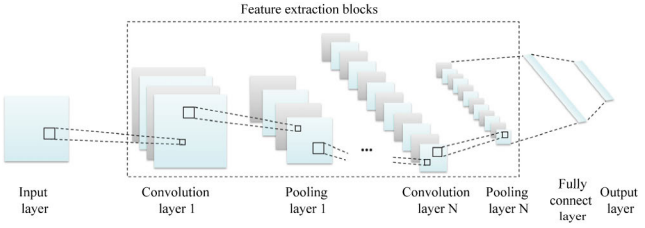


FIGURE 1. Typical architecture of common convolutional neural network.

The feature extraction block is the core component of CNN. Convolutional layers are also known as detection layers. The main purpose of convolutional layers is to learn the features of the input layers. The convolutional layer contains several convolutional kernels, which convolve the input through a set of convolutional kernel weights and output a feature map. All neurons in the same kernel share their weights, thus reducing the optimization time and complexity of the CNN. Pooling layers, also called down-sampling or subsampling layers, normally follow convolutional layers to obtain lower-resolution representations. The pooling layer combines semantically similar features to reduce the dimensionality and parameters of the network. The commonly used pooling operations are max pooling and average pooling, which determine the maximum and average values of each patch on the feature map generated by the convolutional layers, respectively.

C. RECURRENT NEURAL NETWORK

Recurrent neural network (RNN) is a neural network with feedback connection added in the feedforward neural network, which can produce the memory state of past data, so it is suitable for the processing of sequence data and the establishment of the dependence relationship of data from different time steps. A typical RNN architecture is shown in Fig. 2. The structure of an RNN across time can be described as a deep neural network with one layer per time step. Assume that at time t, the input vector of the network is x(t), its hidden vector is h(t), and its output vector is o(t). At time t, the recurrent neurons are input with information from both the previous layer x(t) and the recurrent neurons of the previous state h(t − 1). The output o(t) is influenced by the current input information and the information at time t − 1. These processes can be mathematically described by the following transition functions:

$$\begin{aligned} h(t) &= f(w_{hx}x(t) + w_{hh}h(t-1) + a) \\ o(t) &= g\left(w_{yh}h(t) + b\right) \end{aligned} \tag{1}$$

where f() and g() are activation functions, w_{hx} is the matrix of connection weights between the input layer and the hidden layer, w_{hh} is the matrix between the hidden layer and itself at adjacent time steps, w_{yh} is the matrix between the hidden layer and the output layer, and the vectors a and b are bias parameters of the hidden and output layers which allow each node to learn an offset.

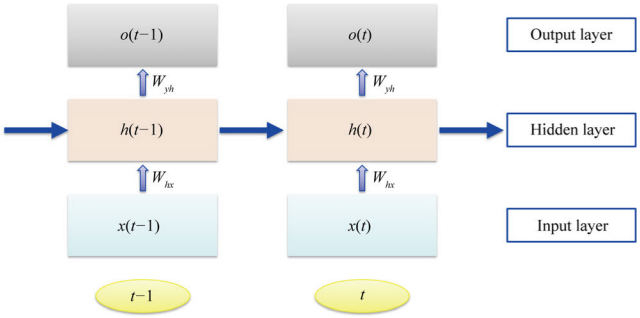


FIGURE 2. Typical architecture of common recurrent neural network.

Long short-term memory (LSTM) is an important improvement model of RNN. LSTM uses memory blocks to replace common hidden neurons used in traditional RNNs to ensure that the gradient will not vanish or explode after

crossing many time steps, which enables LSTM to overcome some difficulties such as a training optimal process approaching to a local extreme value encountered during training of traditional RNNs. The LSTM operation allows a network to learn long-term dependencies between time steps in time series and sequential data.

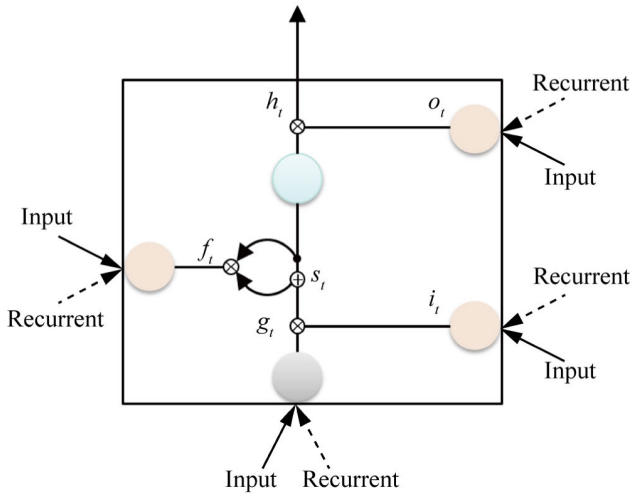


FIGURE 3. Typical architecture of a common long short-term memory block.

There is an internal state in the memory block that connects itself with a fixed weight. As shown in Fig. 3, an input node, an input gate, a forget gate, and an output gate are added to the memory block. The computational process of the memory block in the LSTM model at time t can be formulated as follows:

$$\begin{aligned} g_t &= \tanh(w_{xg}x_t + w_{hg}h_{(t-1)} + b_g) \\ i_t &= \sigma(w_{xi}x_t + w_{hi}h_{(t-1)} + b_i) \\ f_t &= \sigma(w_{xf}x_t + w_{hf}h_{(t-1)} + b_f) \\ o_t &= \sigma(w_{xo}x_t + w_{ho}h_{(t-1)} + b_o) \\ s_t &= s_{t-1} \otimes f_t + i_t \otimes g_t \\ h_t &= o_t \otimes \tanh(s_t) \end{aligned} \tag{2}$$

where g_t, i_t, f_t, and o_t are the output values of the input node, input gate, forget gate, and output gate, respectively; $h_{(t-1)}$ is the output value from the hidden layer at the previous time; w_{xg}, w_{xi}, w_{xf}, and w_{xo} are the weight values between the input layer x_t and the hidden layer h_t at time t; w_{hg}, w_{hi}, w_{hf}, and w_{ho} are the hidden layer weight values between time t and $t-1$; and b_g, b_i, b_f, and b_o are the biases of the input node, input gate, forget gate, and output gate, respectively. The hyperbolic tangent function tanh is a nonlinear transformation function. s_t is the internal state value at the current time t which can be replaced by a rectified linear unit (ReLU), and s_{t-1} is the internal state value at the previous time $t-1$. The symbol $\otimes$ is the pointwise multiplication operator.

D. DEEP NETWORK TRAINING METHODS

The learnable parameters (weights and biases) of different deep networks can be optimized using gradient descent methods. Stochastic gradient descent (SGD) is widely used in training deep learning models. A standard gradient descent algorithm updates the network parameters to minimize the loss function by taking small steps at each iteration in the direction of the negative gradient of the loss. The SGD with momentum (SGDM) optimizer, the root-mean-square propagation (RMSProp) optimizer, and the adaptive moment estimation (Adam) optimizer are commonly used solvers for training deep networks.

The SGD algorithm can oscillate along the path of steepest descent toward the optimum. The SGDM optimization algorithm adds a momentum term to the parameter update to reduce this oscillation. The SGDM algorithm update is

$$\theta_{n+1} = \theta_n - \alpha \nabla E(\theta_n) + \gamma(\theta_n - \theta_{n-1}) \tag{3}$$

where n is the iteration number, α is the learning rate, θ is the parameter vector, and $\nabla E(\theta)$ is the gradient of loss function.

The RMSProp optimization algorithm improves deep network training using learning rates that differ by parameter and can automatically adapt to the loss function being optimized. The RMSProp algorithm update is

$$\theta_{n+1} = \theta_n - \frac{\alpha \nabla E(\theta_n)}{\sqrt{v_n} + \varepsilon} \tag{4}$$

where v_n is the moving average, and ε is a small constant added to prevent division by zero.

The Adam optimization algorithm uses a parameter update with an added momentum term. It keeps an element-wise moving average of both the parameter gradients and their squared values. The Adam optimization algorithm uses the moving averages to update the network parameters as

$$\begin{aligned} \theta_{n+1} &= \theta_n - \frac{\alpha m_n}{\sqrt{v_n} + \varepsilon} \\ m_n &= \beta_1 m_{n-1} + (1-\beta_1)\nabla E(\theta_n) \\ v_n &= \beta_2 v_{n-1} + (1-\beta_2)[\nabla E(\theta_n)]^2 \end{aligned} \tag{5}$$

where β_1 and β_2 are decay rates.

III. EXPERIMENT AND ANALYSIS

Nowadays, smartphones are no longer just used as communication devices as they were when they first appeared. Smartphones are embedded with many lightweight and small sensors with high precision, such as light sensors, distance sensors, gravity sensors, gyroscopic sensors, acceleration sensors, GPS position sensors, magnetic field sensors, and temperature sensors. As the concept of pervasive computing and the Internet of things has deepened, the applications of smartphone sensors have received increasing attention, and the scope of research covers many areas such as behavior sensing, scene recognition, indoor localization, and environmental monitoring. Acoustic signals generated by rotating machinery contain rich information. For example, experienced workers can accurately determine the state of the machine by sound. Inspired by this, we believe that acoustic features can determine the type and location of the most critical component faults in WT generator bearings.

Smartphones have built-in microphones, so they can pick up WT generator acoustic signal easily and accurately.

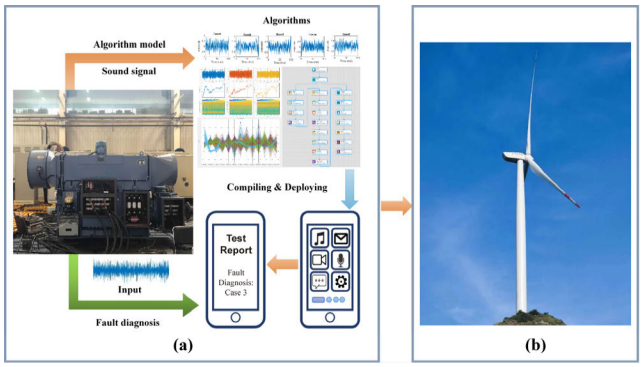


FIGURE 4. Schematic of the wind turbine generator bearing fault diagnosis using acoustic signals: (a) the fault diagnosis research process and (b) the prospects for future real-time application scenarios.

See Fig. 4 for a schematic of the WT generator bearing fault diagnosis based on acoustic analysis using the AI technique. The core processes include the following: (1) collecting bearing operation sounds, (2) building and training algorithmic models, (3) compiling and deploying algorithmic models on smart terminals, and (4) fault diagnostics. In this work, the WT generator was mounted on the test rig and operated at 1500 rpm, and a photograph of the scene drawing is shown in the left panel of Fig. 4.

An experimental data set obtained from CRRC Yongji Electric Laboratory was used to verify the viability and efficacy of the proposed method for bearing fault diagnosis in WT generators. The ball, outer, and inner race faults are the primary fault locations. Pitting and peeling faults are the two fault types that are most likely to occur in bearings. Four most common peeling off fault cases were selected because they account for more than half of the WT generator bearing faults in the CRRC Yongji plant fault statistics database. In this paper, a bearing test rig was used to collect the acoustic signals in the laboratory.

Five different WT generator bearing conditions were evaluated, and their corresponding acoustic signals were recorded. The type of test bearings used were NU214 roller bearings. The faults were inflicted using the electrical discharge engraver to simulate real peeling off fault. The inner race and outer race bearing peeling off faults are shown in Fig. 5(a). In this case study, five groups of acoustic datasets were generated by a commonly used smartphone, and the sampling rate was 48 kHz: (1) Case 1—the states of the drive-end bearing and the non-drive-end bearing are both normal; (2) Case 2—the condition of the drive-end bearing is normal, while the condition of the non-drive-end bearing inner race and outer race have peeling off fault; (3) Case 3—the condition of the outer race of the drive-end bearing have peeling off fault, and the condition of the non-drive-end bearing is normal; (4) Case 4—the condition of the drive-end bearing is normal, and the condition of the non-drive-end bearing outer race have peeling off fault; and (5) Case 5—the condition of the drive-end bearing inner race and outer race have peeling off fault, and the condition of the non-drive-end bearing is normal. In this acoustic signals-based fault diagnosis, a smartphone (Apple iPhone 13) is placed near the drive end of the running WT generator (approximately 2 m) to collect sound data. Then, the microphone recorded acoustic signal is analyzed in order to detect faults within WT generator bearings with no requirement for direct access to them. Five groups of sound datasets were sliced into pieces with the of length 100 ms (each contains 4,800 data points) for the following algorithms. Fig. 5(b) and (c) shows examples of time-domain plots of dataset frames and the spectrogram of dataset frames, respectively. Acoustic frames collected using a smartphone's microphone range from −1 to 1 and can be directly used in the following machine learning algorithms.

We proposed several machine learning algorithms and deep learning algorithms for the fault diagnosis of WT generator bearings. The detailed decomposition steps are shown in Fig. 6. In general, there is a need for feature engineering before building machine learning models. In contrast, deep learning models can be trained directly using the time- or frequency-domain features of raw acoustic signals.

A. MACHINE LEARNING ALGORITHMS USING TYPICAL STATISTICAL FEATURES

The first step in typical machine learning is to extract features of the WT acoustic signal. We extracted some commonly used statistical features. The basic statistical features include mean, standard deviation, RMS, and shape factor. The higher-order statistical features provide an insight into the system behavior through the fourth moment (kurtosis) and third moment (skewness) of the signal. Impulsive metrics are properties related to the peaks of the signal, including peak value, impulse factor, crest factor, and clearance factor. Some signal processing metrics were extracted such as SNR, total harmonic distortion (THD), and signal-to-noise-and-distortion ratio.

TABLE 1. Results of the machine learning models with commonly used statistical features.

	Tree (%)	Discriminant (%)	Naive Bayes (%)	SVM (%)	KNN (%)	Ensemble (%)	Shallow Neural Network (%)
T1	31.80	34.90	33.70	34.40	31.40	33.40	31.30
T2	30.70	30.10	35.90	32.50	31.00	31.90	31.90
T3	32.30	33.50	33.80	34.20	33.10	31.60	31.50
T4	32.00	31.50	34.20	35.10	31.40	33.60	30.60
T5	34.20	32.20	34.60	33.00	31.20	34.90	29.90
Mean	32.20	32.44	34.44	33.84	31.62	33.08	31.04
Max	34.20	34.90	35.90	35.10	33.10	34.90	31.90
Min	30.70	30.10	33.70	32.50	31.00	31.60	29.90
Std	1.27	1.84	0.89	1.06	0.84	1.35	0.79

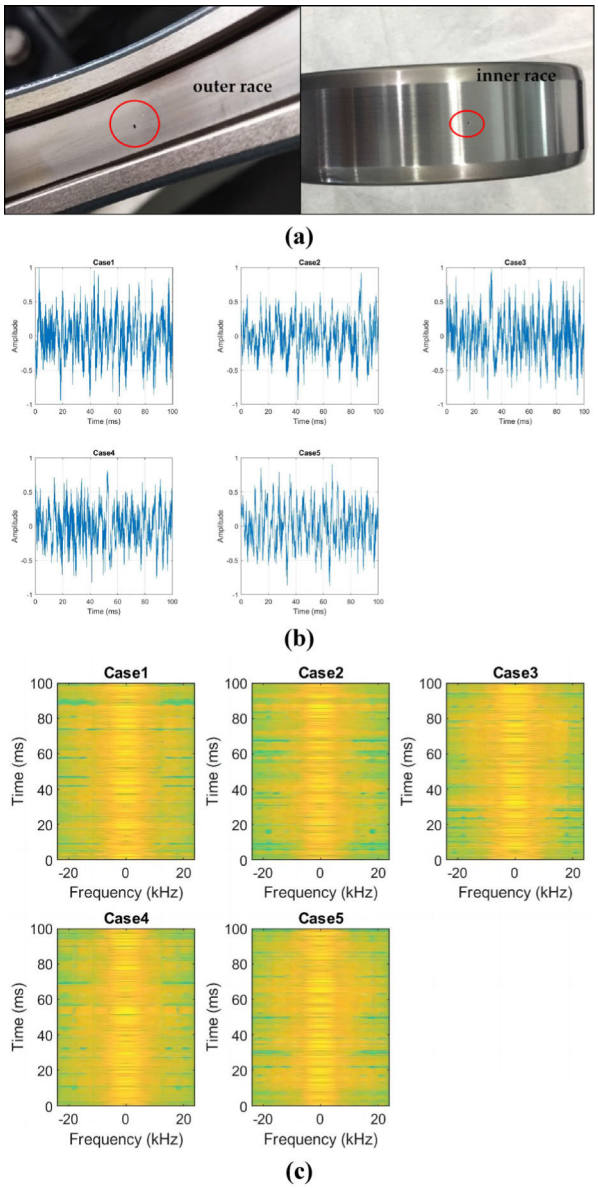


FIGURE 5. Examples of faulty bearings and acoustic signals: (a) faulty bearings of wind turbine generator with outer race fault and inner race, (b) time-domain plots of the dataset frames for five different case conditions and (c) spectrogram of the dataset frames for five different case conditions.

Then, the extracted features were sorted by importance using a one-way analysis-of-variance classification ranking method. The top six important features (skewness, THD, crest factor, impulse factor, clearance factor, and peak value) were selected as the features for the typical machine learning algorithm. After that, the selected six features were imported into several typical machine learning algorithmic models for training. Models including Tree, Discriminant, Naive Bayes, SVM, k-nearest neighbors (KNN), Ensemble, and Shallow Neural Network were trained in this work. We used a dataset of 1,000 samples for each condition and 25% of the training datasets were held out for validation. The result obtained

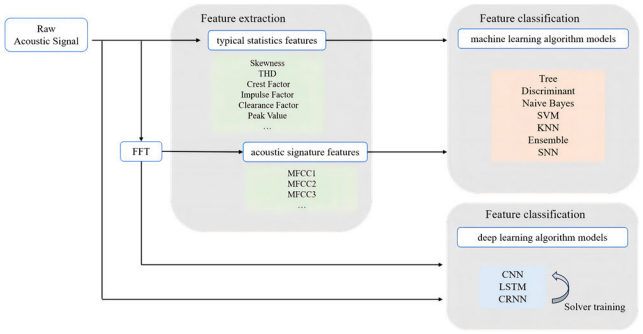


FIGURE 6. Overview of the proposed method. The raw acoustic signal is processed by machine learning algorithms and deep learning algorithms.

using different machine learning algorithmic models were shown in Table 1. We ran each machine learning algorithmic model five times, and the mean, maximum, minimum and the standard deviation of the accuracy were calculated. T1, T2, T3, T4, and T5 represent five tests in the experiment. Using typical statistical features, we achieved that the mean test accuracy is less than 35.00%, specifically 31.04% using Shallow Neural Network model and 34.44% using Naive Bayes model.

TABLE 2. Results of the machine learning models with acoustic signature features.

	Tree (%)	Discriminant (%)	Naive Bayes (%)	SVM (%)	KNN (%)	Ensemble (%)	Shallow Neural Network (%)
T1	90.20	98.50	97.50	98.20	97.90	97.80	97.40
T2	90.90	98.20	97.90	98.50	97.80	97.10	97.50
T3	90.70	98.40	97.80	98.40	97.00	97.80	97.20
T4	91.30	97.90	97.70	98.30	97.50	97.50	97.70
T5	92.00	98.30	96.90	97.40	97.60	97.50	97.30
Mean	91.02	98.26	97.56	98.16	97.56	97.54	97.42
Max	92.00	98.50	97.90	98.50	97.90	97.80	97.70
Min	90.20	97.90	97.50	97.40	97.00	97.10	97.20
Std	0.48	0.68	0.40	0.44	0.35	0.29	0.19

B. MACHINE LEARNING ALGORITHMS WITH ACOUSTIC SIGNATURE FEATURES

Using typical statistical features to train machine learning got unsatisfactory accuracy results. Then we extracted the acoustic signature of the WT acoustic signal. First, the signal was transformed by the FFT, and then the mel-frequency cepstral coefficient (MFCC) features were extracted. We also used a dataset of 1,000 samples for each condition and 25% of the training datasets were hold out for validation. We ran each machine learning model five times, and the mean, maximum, minimum and the standard deviation of the prediction accuracy were calculated (Table 2). The accuracy of either model is more than 90.00%, the maximum average accuracy

is 98.26% using Discriminant model and the minimum average accuracy is 91.02% using Tree model.

C. DEEP LEARNING ALGORITHMS

Deep learning is a machine learning method based on deep network, which is oriented to low-level data objects, adopts layer by layer abstraction mechanism, and finally forms high-level concepts. Unlike previously introduced machine learning algorithms, deep learning algorithms can directly extract high-dimensional features of data, thus requiring less hand-engineering of features.

1) CONVOLUTIONAL NEURAL NETWORK

The architecture of deep learning algorithm based on CNN starts with an input layer to input raw sound data and the corresponding FFT signal. Then, combined feature extraction blocks consist of a convolutional layer, a batch normalization layer, a ReLU layer and a max pooling layer are used to extract features. After that, an average pooling layer and a dropout layer were used to prevent overfitting. A combination of fully connected layer, softmax layer and output layer is used to output the classification results. The typical architecture of the CNN model with different combined feature extraction blocks is shown in Fig. 7, which consists of different network layers stacked one by one.

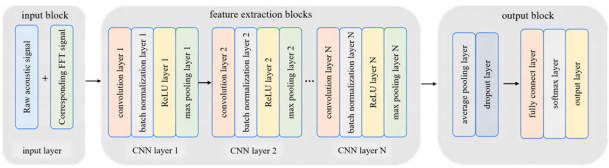


FIGURE 7. CNN structure diagram. The input block, the feature extraction blocks, and the output block are stacked.

Two thousand pieces of sound data for each type were used for the deep learning algorithm based on CNN, so 10,000 pieces of data were formed, where 60% of them were used for network training, 20% for validation during training for the network, and 20% for network accuracy testing. The max epoch w set to 10; the dropout rate is set to 50%; and the solver update the network learnable parameters using the SGDM algorithm.

The accuracy of the test results of CNN with various combined feature extraction block configurations is shown in Fig. 8. The accuracy of the fault diagnosis of WT generator bearings using CNN deep learning models can reach >99.00% when the number of feature extraction blocks exceed three, yielding the best test accuracies of 99.80% with six feature extraction blocks. Fig. 9 shows that good performance can be obtained from the CNN model without the feature extraction step in the above machine learning algorithm.

2) RECURRENT NEURAL NETWORK

Long Short-Term Memory (LSTM, as a representative RNN, is powerful at discovering the variation pattern underlying

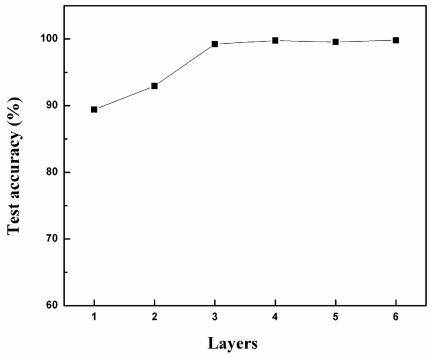


FIGURE 8. Relationship between test accuracy and the CNN layers. The six-layered CNN achieves the highest diagnostic accuracy.

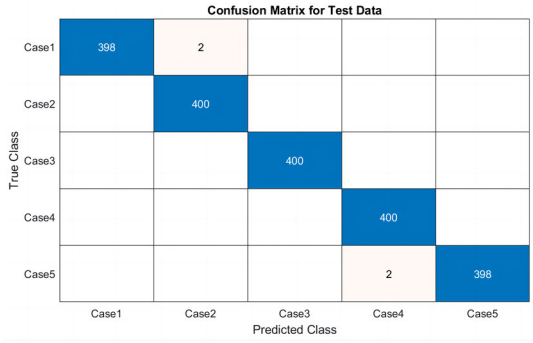


FIGURE 9. Confusion matrix for test data showing the classification results using the CNN model with six CNN layers.

time series. The architecture of deep learning algorithm based on RNN starts with a sequence input layer to input raw sound data and the corresponding FFT signal. An LSTM layer with 50 memory blocks was used to extract features and a dropout layer with 20% dropout rate was used to prevent overfitting. Finally, a combination of fully connected layer, softmax layer and output layer was used to output the classification results. The typical architecture of the RNN model is shown in Fig. 10.

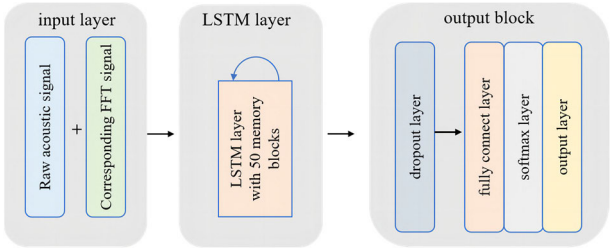


FIGURE 10. RNN structure diagram. The input layer, the LSTM layer, and the output block are stacked.

The Adam optimizer was used for training RNN deep networks. Sound data same as the CNN model were used. The relationship between test accuracy, elapsed time, and epochs is shown in Fig. 11. The test accuracy improved with the increase in the training epochs, indicating that achieving higher accuracy results in a longer elapsed time.

Of course, accuracy does not continue to increase with the number of training steps indefinitely. The test accuracy reached >90.00% when the training epochs was increased to 70. An out-of-memory error occurred when using a laptop with an Intel Core i5 processor and a 16-GB RAM with a GeForce MX450 GPU during training. The test accuracy is only 61.55% with 10 training epochs, and the best test accuracy is 92.10% with 70 training epochs.

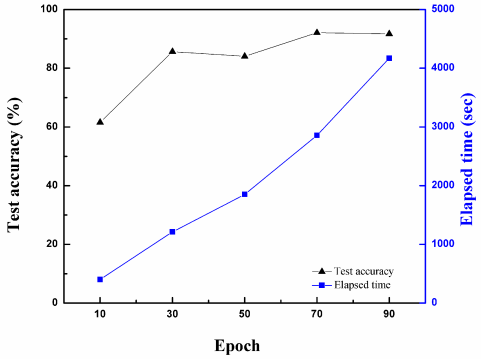


FIGURE 11. Relationship between test accuracy, elapsed time, and epochs. The optimizer with 70 training epochs achieves the highest diagnostic accuracy.

Fig. 12 is the confusion matrix of RNN model for test data with 70 training epochs, shows that good performance can be obtained from the LSTM model without the feature extraction step in the above machine learning algorithm. However, However, the accuracy of the algorithm is not as high as the CNN models.

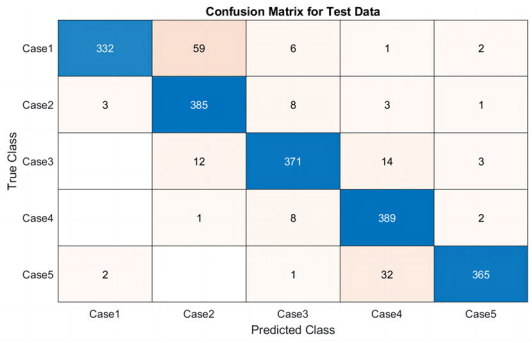


FIGURE 12. Confusion matrix for test data showing the classification results using the LSTM model with 70 training epochs.

3) CONVOLUTIONAL RECURRENT NEURAL NETWORK (CRNN)

Deep learning model combining CNN and RNN, termed as CRNN here, was also used in this work. The typical architecture of CRNN model used in this study is shown in Fig. 13. The architecture of the deep learning algorithm based on CRNN starts with a sequence input layer to input raw acoustic data with or without the corresponding FFT signal. Then, a folding layer was used to perform the convolutional operations on each time step independently. After that, combined feature extraction blocks same as CNN consist of a convolutional layer, a batch normalization layer, a ReLU layer and a max pooling layer were used to extract features. The sequence unfolding layer and flatten layer were used to restore the sequence structure of the input data after sequence folding. After that, an LSTM layer with different numbers of memory blocks was used and a dropout layer with 50% dropout rate was used to prevent overfitting. Finally, a combination of fully connected layer, softmax layer and output layer is used to output the classification results.

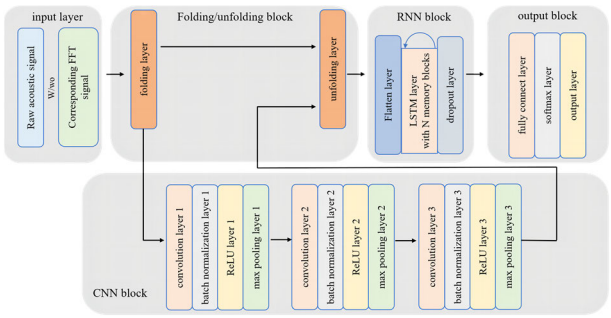


FIGURE 13. CRNN structure diagram. CRNN uses the folding/unfolding block to integrate the CNN and RNN blocks.

The solver updates the CRNN learnable parameters using the RMSProp optimization algorithm. The test accuracy for the CRNN models using one-dimensional signal by raw sound data (time domain) and two-dimensional signal constitute by raw sound signal and the corresponding FFT signal (time domain + frequency domain) with different memory blocks (hidden units) is shown in Fig. 14. The test accuracy using a two-dimensional signal is higher than that using a one-dimensional signal and reaches over 99.00% with memory blocks from 50 to 150. The best test accuracy is 97.55% using the one-dimensional signal with 100 memory blocks or with 150 memory blocks, and the best test accuracy is 99.90% using the two-dimensional signal with 100 memory blocks.

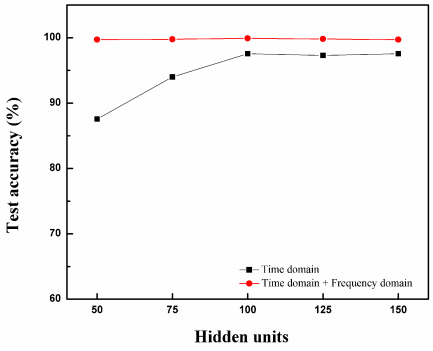


FIGURE 14. Relationship between test accuracy and the hidden units. The two-dimensional signal (time and frequency domains) achieves better diagnostic accuracy than the one-dimensional signal (time domain).

D. COMPARISON OF THE SAME DEEP LEARNING ALGORITHM USING VIBRATIONAL SIGNALS

Recently, new high-performance computing architectures have become a catalyst for the evolution of deep learning

techniques. With the support of big data in the enterprise, the fault diagnosis results of deep learning algorithms will be more accurate as the amount of data processed increases. Therefore, we believe that deep learning techniques will become mainstream for fault diagnosis and intelligent maintenance. The performance of three different deep learning algorithms using vibrational signals is investigated for comparison.

1) DATA ACQUISITION PROCESS

Incorporating the background of big data, industrial internet, and current smart industry trends, researchers have made significant progress by integrating bearing fault diagnosis with deep learning techniques based on vibrational signals. As a result, most new WT generators are fabricated with vibration sensors to monitor their operating conditions, and the data can be simultaneously transmitted back to the monitoring center of the wind farm.

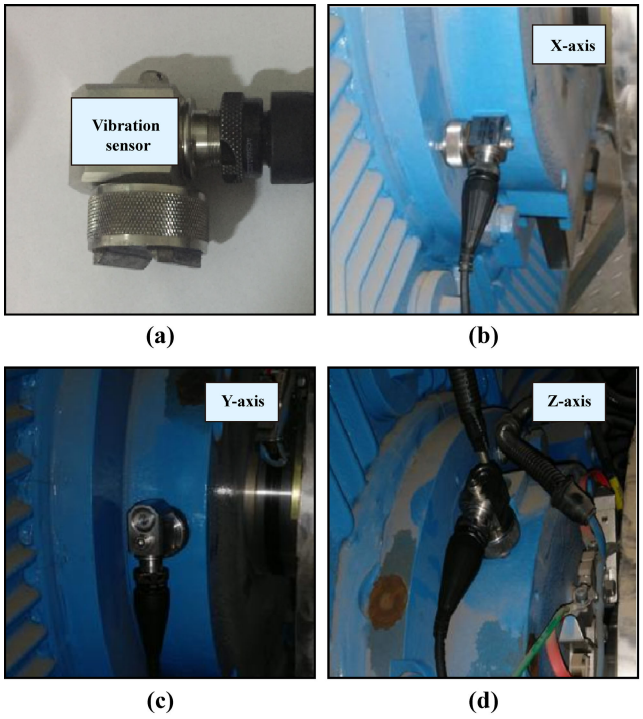


FIGURE 15. Vibration sensor and its different mounting orientations on the WT generator. Horizontal, vertical, and axial data were collected with the mount orientation of the (b) X-axis, (c) Y-axis, and (d) Z-axis, respectively.

To compare the performance of the vibrational and acoustic signals, the vibrational signal was collected using the same experimental parameters as above. An acceleration sensor was used to collect the vibration signal at a sampling frequency of 20 kHz. The sensor is mounted at the drive-end of the WT generator and by adjusting the mount orientation, horizontal, vertical, and axial data are collected, which are denoted as X-axis, Y-axis, and Z-axis, respectively, as shown in Fig. 15.

The same WT generator faulty bearings as in the previous acoustic signal experiment was employed. Sample time-domain frames of WT generator vibration signals for five different faults are shown in Fig. 16.

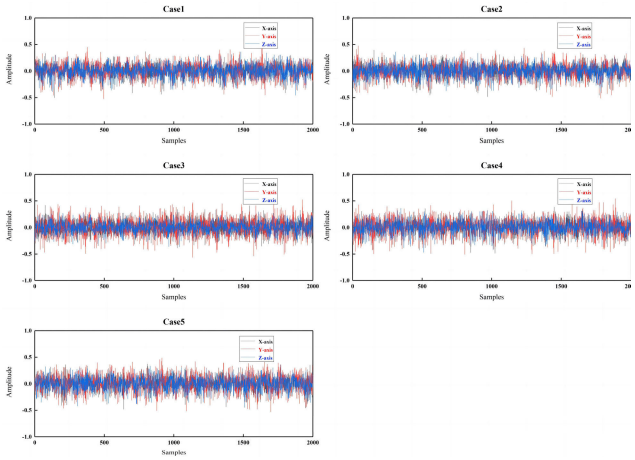


FIGURE 16. Typical frames of vibrational signals for different faulty bearings for five different case conditions.

2) DEEP LEARNING ALGORITHMS USING VIBRATIONAL SIGNALS

To compare the fault diagnosis performance of WT generator bearings with vibrational signal and acoustic signal, the accuracy of vibration sensors was evaluated under different mount orientations using three different deep learning algorithms as described above (CNN, RNN, and CRNN). Because there are vibration sensors available on the market that can measure vibration signals in three directions in synchrony, signals from three different mount orientations were combined to mimic a triaxial vibration sensor. To mitigate the impact of randomness, the average outcomes of five runs were reported. The mean accuracy test results for the five experiments are shown in Table 3.

TABLE 3. Results of the deep learning algorithms using vibrational signals.

Deep learning algorithm (%)	X-axis (%)	Y-axis (%)	Z-axis (%)	Three axis combination (%)
CNN	97.80	97.10	96.90	98.40
RNN	96.70	98.20	97.70	98.50
CRNN	99.50	99.10	99.20	99.60

From the fault diagnosis results using the vibrational signal, the accuracy of the WT generator bearing fault diagnosis based on the vibrational signal using the deep learning algorithm is more than 96%, and especially the accuracy of the CRNN deep learning algorithm is more than 99%, showing the best performance. There is a slight difference in the accuracy of deep learning algorithms using vibrational

signals with different mount orientations. Using a three-axis combination of vibration signals will slightly improve accuracy, but the price of a triaxial vibration sensor is considerably larger than that of a single-axis vibration sensor. Thus, from the perspective of high-cost performance, single-axis vibration sensors are more suitable for WT generator bearing fault diagnosis.

IV. CONCLUSION

In this study, machine learning algorithms and deep learning algorithms for the fault diagnosis of WT generator bearings are proposed. The proposed deep learning algorithms can efficiently and accurately determine the location and type of WT generator bearing faults. A summary of the test results of the different algorithms is shown in Fig. 17. Both machine learning algorithmic models using acoustic features and deep learning algorithmic models can achieve high accuracy.

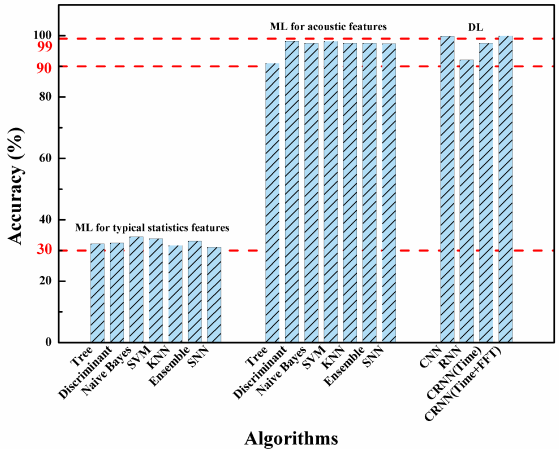


FIGURE 17. Summary of the test results of the different algorithms. Both machine learning algorithmic models using acoustic features and deep learning techniques achieve high diagnostic accuracy.

A reliable and high accuracy fault diagnosis of WT generator bearings can be obtained using the running sound as the input signal. The accuracy of typical machine learning algorithms is greatly affected by the type of the extracted features, where the accuracy is only less than 35.00% when the signals were analyzed by machine learning for typical statistical features but change the features as acoustic features, the accuracy rose to over 90.00%, meet the testing requirements basically. So that, MFCC features are more suitable than commonly used statistical features for the WT generator bearing operation sound signals. Even better promising results emerged after the leading of deep network, the accuracy has been further improved and reached 99.90% especially in CNN and CRNN, RNN take longer network training time and a lower accuracy. Such a high fault diagnosis accuracy of WT generator bearing provides protection for the product predelivery inspection and reliable evidence for depot repair products inspections, ensure safe equipment operation and production. In addition, convenience has been greatly improved because the algorithms can be compiled using C or C++ executable codes and can be easily deployed on embedded devices, such as a palmtop or a smartphone.

In general, most current intelligent bearing fault diagnosis technologies require the installation of new vibration sensors or acoustic emission sensors, resulting in increased product costs. Experimental results show that deep learning algorithms based on vibrational signals achieve high accuracy when applied to WT generator bearing fault diagnosis. Therefore, it is expected to be useful for field fault diagnosis and predictive maintenance. This study proposes the use of acoustic signals for fault diagnosis, which enables data collection and fault diagnosis using only a single smartphone without the need for additional sensors and reduces product costs while enabling real-time diagnosis. Smartphones and other embedded devices can readily implement the proposed techniques. As smartphones are ubiquitous, they can run the fault diagnosis algorithms by picking up acoustic signals through their built-in microphones.

In terms of the development process of fault diagnosis technology, intelligent maintenance and health management of equipment is the only way to achieve intelligent production. After years of development, more and more types of signals are monitored for equipment during fault diagnosis, and the amount of information obtained is increasingly rich, including vibration signals, acoustic emission signals, electrical signals, temperature, and pressure. Fault diagnosis algorithms have evolved from traditional ones such as FFT, envelope spectral analysis, power spectrum, and cepstrum to state-of-the-art popular AI-based methods. Equipment such as WT mostly operate in extremely complex environments (extreme temperature, elevated pressure, salt spray, humidity, etc.), which results in stronger randomness of state transition, and the mechanism modeling is quite difficult. In addition, fault diagnosis methods for large devices often suffer from low precision and accuracy issues caused by complex structures, weak signals, and environmental noise interference. How to select the most appropriate signal and algorithm is an essential direction of current research. Acoustic signals are used as input data in this work, and machine learning and deep learning techniques are used as fault diagnosis algorithms, which have advantages in terms of accuracy, low cost, and ease of deployment. As a result, it is expected to be widely used for fault detection in large and complex equipment components such as WTs or high-speed trains.

In conclusion, the proposed method can be used for factory inspection of new WT generator products and parts returned to the factory for maintenance, but it is not yet suitable for real-time diagnostics in the field. The research results of this study can provide a reference for the related research of application of deep learning in the field of bearing fault diagnosis. The adaptability of the algorithms in various noisy environments while the WT are deployed in remote areas and are exposed to harsh environments, suffering from constantly varying loads, and experiencing extreme operational temperature and humidity changes should be studied.

Conversely, the algorithms need further training to be suitable for different rotational speeds to expand the application scope. Finally, it is necessary to study the effectiveness and availability of the algorithm in more cases of bearing faults and implement the algorithms into expert diagnosis systems to reduce the workload of maintenance personnel and improve maintenance efficiency.

The proposed method is only trained and validated on a limited number of fault types. On the basis of the aforementioned analysis, future research works can proceed in the following manner. First, data collection and experimental analysis under different running conditions (such as operating speed or load) should be conducted. Second, more types of bearing faults, more locations and different levels of bearing failure will be investigated to achieve accurate judgments of bearing health for predictive maintenance. Finally, the use acoustic-signal-based fault diagnosis techniques will be explored for field detection, such as using unmanned aerial vehicles to collect field operation sounds, but how to extract weak fault features in complex industrial environments is a major challenge.

REFERENCES

[1] F. Blaabjerg and K. Ma, "Future on power electronics for wind turbine systems," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 1, no. 3, pp. 139–152, Sep. 2013.

[2] A. Kusiak and W. Li, "The prediction and diagnosis of wind turbine faults," *Renew. Energy*, vol. 36, no. 1, pp. 16–23, Jan. 2011.

[3] Z. Fu, Z. Zhou, and Y. Yuan, "Fault diagnosis of wind turbine main bearing in the condition of noise based on generative adversarial network," *Processes*, vol. 10, no. 10, p. 2006, Oct. 2022.

[4] J. Igba, K. Alemzadeh, C. Durugbo, and K. Henningsen, "Performance assessment of wind turbine gearboxes using in-service data: Current approaches and future trends," *Renew. Sustain. Energy Rev.*, vol. 50, pp. 144–159, Oct. 2015.

[5] K. Fischer, F. Besnard, and L. Bertling, "Reliability-centered maintenance for wind turbines based on statistical analysis and practical experience," *IEEE Trans. Energy Convers.*, vol. 27, no. 1, pp. 184–195, Mar. 2012.

[6] M. Cerrada, R.-V. Sánchez, C. Li, F. Pacheco, D. Cabrera, J. Valente de Oliveira, and R. E. Vásquez, "A review on data-driven fault severity assessment in rolling bearings," *Mech. Syst. Signal Process.*, vol. 99, pp. 169–196, Jan. 2018.

[7] T. Lin, Y. Zhu, Z. Ren, K. Huang, and D. Gao, "CCFT: The convolution and cross-fusion transformer for fault diagnosis of bearings," *IEEE/ASME Trans. Mechatronics*, vol. 29, no. 3, pp. 2161–2172, Jun. 2024.

[8] Y. Lin, L. Tu, H. Liu, and W. Li, "Fault analysis of wind turbines in China," *Renew. Sustain. Energy Rev.*, vol. 55, pp. 482–490, Mar. 2016.

[9] Z. Xue, Y. Huang, W. Zhang, J. Shi, and H. Luo, "Intelligent fault diagnosis of rolling bearings based on a complete frequency range feature extraction and combined feature selection methodology," *Sensors*, vol. 23, no. 21, p. 8767, Oct. 2023.

[10] Y. Wei, X. Wang, Y. Xu, and F. Fan, "Intelligent fault diagnosis of rotating machinery using composite multivariate-based multi-scale symbolic dynamic entropy with multi-source monitoring data," *Struct. Health Monitor.*, vol. 22, no. 1, pp. 56–77, Jan. 2023.

[11] Z. He, H. Shao, J. Cheng, Y. Yang, and J. Xiang, "Kernel flexible and displaceable convex hull based tensor machine for gearbox fault intelligent diagnosis with multi-source signals," *Measurement*, vol. 163, Oct. 2020, Art. no. 107965.

[12] X. Li, Y. Yang, H. Shao, X. Zhong, J. Cheng, and J. Cheng, "Symplectic weighted sparse support matrix machine for gear fault diagnosis," *Measurement*, vol. 168, Jan. 2021, Art. no. 108392.

[13] J. Yang and C. Delpha, "An incipient fault diagnosis methodology using local Mahalanobis distance: Detection process based on empirical probability density estimation," *Signal Process.*, vol. 190, Jan. 2022, Art. no. 108308.

[14] Y. Gu, R. Chen, K. Wu, P. Huang, and G. Qiu, "A variable-speed-condition bearing fault diagnosis methodology with recurrence plot coding and MobileNet-v3 model," *Rev. Sci. Instrum.*, vol. 94, no. 3, Mar. 2023, Art. no. 034710.

[15] D. B. G. Ke-Ming, L. I. Jin-Feng, J. I. N. Song, and Z. Zhen-Tao, "Bearing fault diagnosis method based on LMD and improved CNN," *Manuf. Autom.*, vol. 45, no. 1, pp. 216–220, 2023.

[16] H. Li, T. Liu, X. Wu, and S. Li, "Correlated SVD and its application in bearing fault diagnosis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 1, pp. 355–365, Jan. 2023.

[17] H. Zhang, P. Shi, D. Han, and L. Jia, "Research on rolling bearing fault diagnosis method based on AMVMD and convolutional neural networks," *Measurement*, vol. 217, Aug. 2023, Art. no. 113028.

[18] H. Zhou, P. Cheng, S. Shao, Y. Zhao, and X. Yang, "Intelligent bearing fault diagnosis method based on a domain aligned clustering network," *Meas. Sci. Technol.*, vol. 34, no. 4, Apr. 2023, Art. no. 044001.

[19] D. Cheng and K. I. Kou, "FFT multichannel interpolation and application to image super-resolution," *Signal Process.*, vol. 162, pp. 21–34, Sep. 2019.

[20] A. Stefani, A. Yazidi, C. Rossi, F. Filippetti, D. Casadei, and G.-A. Capolino, "Doubly fed induction machines diagnosis based on signature analysis of rotor modulating signals," *IEEE Trans. Ind. Appl.*, vol. 44, no. 6, pp. 1711–1721, Jun. 2008.

[21] I. Djurović and L. Stanković, "STFT-based estimator of polynomial phase signals," *Signal Process.*, vol. 92, no. 11, pp. 2769–2774, Nov. 2012.

[22] Y. Guo, A. Wang, and W. Wang, "Multi-source phase retrieval from multi-channel phaseless STFT measurements," *Signal Process.*, vol. 144, pp. 36–40, Mar. 2018.

[23] R. Klein, D. Ingman, and S. Braun, "Non-stationary signals: Phase-energy approach—Theory and simulations," *Mech. Syst. Signal Process.*, vol. 15, no. 6, pp. 1061–1089, Nov. 2001.

[24] H. Cui, Y. Qiao, Y. Yin, and M. Hong, "An investigation on early bearing fault diagnosis based on wavelet transform and sparse component analysis," *Struct. Health Monitor.*, vol. 16, no. 1, pp. 39–49, Jan. 2017.

[25] Z. He, J. Li, L. Liu, and Y. Shen, "Three-dimensional empirical mode decomposition (TEMD): A fast approach motivated by separable filters," *Signal Process.*, vol. 131, pp. 307–319, Feb. 2017.

[26] J. Guo, Z. Shi, D. Zhen, Z. Meng, F. Gu, and A. D. Ball, "Modulation signal bispectrum with optimized wavelet packet denoising for rolling bearing fault diagnosis," *Struct. Health Monitor.*, vol. 21, no. 3, pp. 984–1011, May 2022.

[27] X. Zhang and J. Zhou, "Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines," *Mech. Syst. Signal Process.*, vol. 41, nos. 1–2, pp. 127–140, Dec. 2013.

[28] A. Dibaj, M. M. Ettefagh, R. Hassannejad, and M. B. Ehghaghi, "Fine-tuned variational mode decomposition for fault diagnosis of rotary machinery," *Struct. Health Monitor.*, vol. 19, no. 5, pp. 1453–1470, Sep. 2020.

[29] X. Jiang, J. Wang, C. Shen, J. Shi, W. Huang, Z. Zhu, and Q. Wang, "An adaptive and efficient variational mode decomposition and its application for bearing fault diagnosis," *Struct. Health Monitor.*, vol. 20, no. 5, pp. 2708–2725, Sep. 2021.

[30] R. R. Shubita, A. S. Alsadeh, and I. M. Khater, "Fault detection in rotating machinery based on sound signal using edge machine learning," *IEEE Access*, vol. 11, pp. 6665–6672, 2023.

[31] Z. Meng, J. Zhu, S. Cao, P. Li, and C. Xu, "Bearing fault diagnosis under multisensor fusion based on modal analysis and graph attention network," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–10, 2023.

[32] D. Xu and C. Li, "Optimization of deep belief network based on sparrow search algorithm for rolling bearing fault diagnosis," *IEEE Access*, vol. 12, pp. 10470–10481, 2024.

[33] Y. Zhang, J. Hua, X. Fang, H. Zhang, J. He, and Q. Miao, "Trackside acoustic fault diagnosis of bearing based on Doppler knowledge embedded in domain adaptation network," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–13, 2024.

[34] M. Liu, Z. Cheng, N. Hu, Y. Yang, and J. Cheng, "Intelligent fault diagnosis of rotating machinery based on acoustic field signal and domain adaptation," in *Proc. Global Rel. Prognostics Health Manage. Conf. (PHM-Hangzhou)*, Hangzhou, China, Oct. 2023, pp. 1–6.

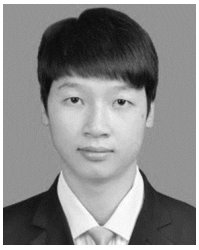
[35] K. Manikanta, K. P. Soman, and M. S. Manikandan, "Deep learning based effective baby crying recognition method under indoor background sound environments," in *Proc. 4th Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solution (CSITSS)*, Dec. 2019, pp. 1–6.
[36] T. Chen, D. Yu, B. Wu, and B. Xia, "Weak signals detection by acoustic metamaterials-based sensor," *IEEE Sensors J.*, vol. 21, no. 15, pp. 16815–16825, Aug. 2021.
[37] C.-H. Lee, J.-S. Jwo, H.-Y. Hsieh, and C.-S. Lin, "An intelligent system for grinding wheel condition monitoring based on machining sound and deep learning," *IEEE Access*, vol. 8, pp. 58279–58289, 2020.
[38] X. Zhao and M. Jia, "A novel unsupervised deep learning network for intelligent fault diagnosis of rotating machinery," *Struct. Health Monitor.*, vol. 19, no. 6, pp. 1745–1763, Nov. 2020.
[39] C. M. Vasile, "A review of deep learning in medical practice," *Res. Sci. Today*, vol. 21, no. 1, pp. 67–77, 2021.
[40] Y. Luo, W. Lu, S. Kang, X. Tian, X. Kang, and F. Sun, "Enhanced feature extraction network based on acoustic signal feature learning for bearing fault diagnosis," *Sensors*, vol. 23, no. 21, p. 8703, Oct. 2023.
[41] J. Xie, X. Aubert, X. Long, J. van Dijk, B. Arsenali, P. Fonseca, and S. Overeem, "Audio-based snore detection using deep neural networks," *Comput. Methods Programs Biomed.*, vol. 200, Mar. 2021, Art. no. 105917.
[42] Y. Wang and R. Liu, "Simulation of sound signal analysis model in complex environments based on deep learning algorithms," in *Proc. Int. Conf. Internet Things, Robot. Distrib. Comput. (ICIRDC)*, vol. 30, Rio De Janeiro, Brazil, Dec. 2023, pp. 666–670.
[43] R. Zaheer, I. Ahmad, D. Habibi, K. Y. Islam, and Q. V. Phung, "A survey on artificial intelligence-based acoustic source identification," *IEEE Access*, vol. 11, pp. 60078–60108, 2023.
[44] L. Tang, X. Wu, D. Wang, and X. Liu, "A comparative experimental study of vibration and acoustic emission on fault diagnosis of low-speed bearing," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.
[45] Y. Guo, M. Guo, Y. Shen, Y. Peng, and S. Zhao, "Research on radial rotor plunger wear fault monitoring method by fused sound vibration signal features," *IEEE Sensors J.*, vol. 24, no. 13, pp. 20896–20907, Jul. 2024.
[46] J. Geng, Y. Liao, L. Guo, X. Ma, and G. Wang, "Motor equipment fault diagnostics via multi-feature acoustic signal fusion," in *Proc. CAA Symp. Fault Detection, Supervision Saf. Tech. Processes (SAFEPROCESS)*, vol. 47, Yibin, China, Sep. 2023, pp. 1–6.
[47] M. A. Jarwar, S. A. Khowaja, K. Dev, M. Adhikari, and S. Hakak, "NEAT: A resilient deep representational learning for fault detection using acoustic signals in IIoT environment," *IEEE Internet Things J.*, vol. 10, no. 4, pp. 2864–2871, Feb. 2023.
[48] M. Wang, K. Huo, and Q. Jia, "Rolling bearings fault diagnosis method based on NRBO-SVM," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Tianjin, China, Aug. 2024, pp. 620–625.
[49] H. Geraei, E. A. V. Rodriguez, E. Majma, S. Habibi, and D. Al-Ani, "A noise invariant method for bearing fault detection and diagnosis using adapted local binary pattern (ALBP) and short-time Fourier transform (STFT)," *IEEE Access*, vol. 12, pp. 107247–107260, 2024.
[50] W. Liao, L. Wu, S. Xu, and S. Fujimura, "A novel approach for intelligent fault diagnosis in bearing with imbalanced data based on cycle-consistent GAN," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–16, 2024.

BEI ZHAO received the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 2018.

She is currently a Postdoctoral Researcher with the Bioinspired Engineering and Biomechanics Center (BEBC), Xi'an Jiaotong University. She is also a Lecturer with the Department of Physics Teaching, Xi'an Jiaotong University City College, Xi'an. Her current research interests include artificial intelligence, acoustic analysis, fault diagnosis, and deep learning.

XIAOMENG LI received the Ph.D. degree from Xi'an Jiaotong University, in 2016.

He is currently a Lecturer with the Department of Computer and Network Engineering, Shanxi Datong University. He is also an Engineer-in-Charge of motor fault diagnosis with CRRC Yongji Electric Company Ltd. His current research interests include computer networks and mechanical fault diagnosis.



ZEDONG LI received the Ph.D. degree from Xi'an Jiaotong University, in 2018.

He is currently an Associate Professor with the School of Life Science and Technology, Xi'an Jiaotong University. His current research interests include microfluidics devices for diagnostics including paper-based microfluidic chip real-time detection and the application of microfluidic chip body fluid nucleic acid detection.

MINLI YOU received the Ph.D. degree from Xi'an Jiaotong University, in 2019.

From 2017 to 2018, he was a Visiting Scholar with Washington University in St. Louis, St. Louis, MO, USA. He is currently an Associate Professor with the School of Life Science and Technology, Xi'an Jiaotong University. His research interests include ultrafast PCR and digital POCT.

FENG XU received the M.S. degree in mechanical engineering from Xi'an Jiaotong University, in 2004, and the Ph.D. degree in engineering from the University of Cambridge, Cambridge, U.K., in 2008.

From 2008 to 2011, he was a Research Fellow at Harvard Medical School and Harvard-MIT Health Science and Technology. He found the Bioinspired Engineering and Biomechanics Center, Xi'an Jiaotong University, and is the co-founding director of the center. He is currently a Professor with Xi'an Jiaotong University. His current research interests include advancing human health through academic excellence in education and research that integrates engineering, science biology, and medicine with a focus on bio-thermo-mechanics, engineering of cell microenvironment, and point-of-care technologies.