Wind Turbine Gearbox Fault Diagnosis Based on Multi-sensor Signals Fusion

Yao Zhao, *Member*, *IEEE*, Ziyu Song, Dongdong Li, *Member*, *IEEE*, Rongrong Qian, and Shunfu Lin, *Member*, *IEEE*

***Abstract*—This paper proposes a novel fault diagnosis method by fusing the information from multi-sensor signals to improve the reliability of the conventional vibration-based wind turbine drivetrain gearbox fault diagnosis methods. The method fully extracts fault features for variable speed, insufficient samples, and strong noise scenarios that may occur in the actual operation of a wind turbine planetary gearbox. First, multiple sensor signals are added to the diagnostic model, and multiple stacked denoising auto-encoders are designed and improved to extract the fault information. Then, a cycle reservoir with regular jumps is introduced to fuse multidimensional fault information and output diagnostic results in response to the insufficient ability to process fused information by the conventional Softmax classifier. In addition, the competitive swarm optimizer algorithm is introduced to address the challenge of obtaining the optimal combination of parameters in the network. Finally, the validation results show that the proposed method can increase fault diagnostic accuracy and improve robustness.**

***Index Terms*—Wind turbine gearbox, fault diagnosis, multiple scenarios, deep learning, stacked denoising auto-encoder, cycle reservoir with regular jumps, feature fusion network.**

I. Introduction

Wind energy has become an important resource worldwide, and wind turbines have been widely researched and applied [1], [2]. Planetary gears are important components in the drive train of wind turbines, and have the advantages of compact structure, high power density, and high transmission efficiency [3]. However, they are prone to failure under dynamic load and frequently changing operating conditions, resulting in high maintenance costs. Accurate diagnosis of gearbox faults is essential to improve the safety, reliability, and economy of wind turbines.

Conventional fault diagnosis methods usually analyze the collected vibration signal via signal analysis methods to identify gearbox fault types. Feature analysis of vibration and current signals is conducted and an SVM-based feature fusion network for effective diagnosis of gearbox fault is constructed in [4]. However, gearbox fault signals are strongly nonstationary, which leads to poor distinguishability. Therefore, high-level expertise and signal-processing techniques are required for the above methods, while the poor generalizability can lead to missing information when applied in practice [5].

With the development of artificial intelligence technologies, intelligent diagnosis of planetary gears has been widely researched [6]. Deep learning has the advantages of high intelligence, high efficiency, and strong generalizability. Thus, it can meet the requirements of fault diagnosis for feature adaptive extraction and reduce the reliance on expert experience. It can also fully reveal the fault feature information of the original data. An improved method based on convolutional neural network (CNN) is proposed in [7] and the experimental results show that the improved network is more robust to external disturbances. In [8], a multiscale convolutional neural network (MSCNN) is proposed for the multi-scale features of a gearbox vibration signal, and the fault information at different scales is successfully extracted. A lifelong learning method for gearbox diagnosis is proposed in [9]. This has satisfactory robustness with incremental fault types. An improved CNN model structure is proposed in [10], where a new convolutional layer is added in front of the fully connected layer to mine the deeper features of the signal, and this improves the generalizability of the model. To overcome the shortcomings of traditional methods and improve fault diagnosis capability, a fault detection method based on the combination of multiple fractal spectra and a support vector machine is proposed in [11] where the classification and identification is made of normal signals and four kinds of sun wheel fault signals at different rotational speeds. Diagnostic results from

Received: October 30, 2023
Accepted: February 6, 2024
Published Online: July 1, 2024
Yao Zhao, Ziyu Song, Dongdong Li (corresponding author), and Shunfu Lin are with the College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China (e-mail: nihaozhaoyao@163.com, szy980614@outlook.com, shiepldd@163.com, shunfulin@shiep.edu.cn).
Rongrong Qian is with the Design Research and Development Center AECC Commercial Aircraft Engine Co., Ltd., Shanghai 200241, China (e-mail: rongrong_qian@126.com).
DOI: 10.23919/PCMP.2023.000241

different classifiers are fused to secure robustness in the presence of noise in [12].

The above research shows that deep learning methods have achieved remarkable success in intelligent fault diagnosis of planetary gears. However, the signal from a single sensor is used as input in most existing deep learning methods. As the planetary gear is in contact with the sun gear and the gear ring simultaneously, the vibration signal characteristics of the planetary gear are complex. In addition, planetary gears usually operate at low speed and high load. This increases noise interference in the vibration signal. Consequently, diagnostic accuracy and robustness may be compromised. Sensor failure may also directly lead to incorrect diagnostic results [13]. In recent years, many studies have been conducted on the introduction of other signals. An analysis method based on time domain signal characteristics is proposed in [14]. This is supported by a fast dynamic time regularization algorithm and a correlated kurtosis algorithm. The fusion of stator and rotor current signals of doubly-fed generators is used for the effective identification of gear faults in [15], whereas deep convolutional neural networks based on the fused data of vibration and current sensors are developed to achieve the detection of planetary gearbox faults in [16]. The above studies introduce the current signal as the input to the diagnostic model. However, the essence of fault information in a current signal is the feedback of torsional vibrations to the stator current, and these fault features are not prominent in the current signal with a fault in the gearing system [17].

Other signals have been used to diagnose faults in various types of rotating equipment over the last few years [18]–[20]. At present, multi-sensor signal fusion methods have been applied in several fields. A diagnosis method using multi-sensor fusion and CNN is proposed for mechanical fault diagnosis in [21]. This is suitable for diagnosis problems with multi-source signals and has advantages in the improvement of classification accuracy. Vibration signals and current signals are fused in most of the above research, while the fault information contained in the current signal is limited. Therefore, new sensor signals have to be introduced as input to the diagnostic model.

The autoencoder is improved and used as the fault extraction module to extract fault features in [22]. Generally, the feature extraction module can extract features better with sufficient training data. However, wind turbines often operate in strong noise conditions and the fault data is usually insufficient in real cases. So the feature extraction module needs to be improved to enhance the feature extraction capability in the case of strong noise and small samples. Also, very large computing resources are required for conventional feature extraction methods for multiple sensor signals. A CNN which is trained well by the denoising-classification neural network (DCNN) is proposed to develop a protection scheme [23]. In [24], bearing fault features from four sensors are extracted. However, the training time is longer than the feature extraction network with a single input signal. So it is necessary to improve the network to speed up the training process. A generative adversarial network is used to generate data for fault diagnosis with insufficient data and Softmax is used as a classifier to output the diagnosis results in [24], though the experiments conducted in the paper indicate that Softmax performs poorly in fusing the features extracted from the used signals. Therefore, a new classifier needs to be introduced to fuse the multi-dimensional features.

This paper proposes a wind turbine drivetrain gearbox fault diagnosis method based on a multi-sensor signals feature fusion network (MSFFN) for variable speed conditions, insufficient samples and strong noise scenarios. To fully extract the fault feature information and to minimize the missing information, signals from multiple sensors are used as input to the proposed diagnostic model. Then, a multi-sensor signals stacked denoising auto-encoder (MS-SDAE) is constructed and improved to extract the fault information from each sensor signal. Since the conventional classifier has insufficient ability to process fused information, a cycle reservoir with regular jumps (CRJ) is introduced to fuse the multisensory signal fault information and output diagnostic results. In addition, a CRJ based on competitive swarm optimizer (CSO) optimization is designed to address the difficulty for CRJs to obtain the optimal combination of parameters. The MSFFN model formulation for a wind turbine gearbox is proposed, and the proposed method is validated using data collected from multiple sensors in comparison with other diagnostic methods using single signal or multi-signal diagnostic methods.

The main contributions of this paper are:

1) Signals from various sensors are discussed and used as input to the diagnostic model. It is verified that the extra input signals can improve diagnostic accuracy.

2) SDAEs, as an efficient fault feature extraction model is developed in this paper. The activation function in the model is changed and the network structure is improved. These modifications can speed up the training process with more input signals.

3) CSO-CRJ is introduced into the diagnostic model as a classifier to replace the conventional Softmax classifier, and it is proven that the proposed classifier performs better than the conventional classifier at fusing fault features in multiple signals.

The remainder of this paper is organized as follows. Section II presents the theoretical background of the proposed multi-sensor signals feature fusion fault diagnosis method. In Section III, the MSFFN model formulation for a wind turbine gearbox is proposed based on the selected signals and deep learning methods. Section IV validates the effectiveness and superiority of

the proposed method in comparison with those using single signal or other multi-signals via experimental test results carried out on a gearbox test rig. Section V provides the conclusions.

II. Theoretical Background

A. Input Signal of the Diagnostic Model

Mechanical systems are dynamic systems, and the overall changes in the system cannot be accurately collected by a single sensor. Thus, here, more and different sensors are placed to fully collect the fault information of wind turbines.

Vibration signal is widely used in the field of gearbox fault diagnosis. A planetary gearbox dynamic model is developed to analyze the causes of gear vibration. The vibrations generated by internal and external excitations are considered and the differential equations for the dynamics can be established as:

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{C}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} = \boldsymbol{T} \tag{1}$$

where $\boldsymbol{M}$, $\boldsymbol{C}$, and $\boldsymbol{K}$ are the mass, damping, and stiffness matrices of the system, respectively; $\boldsymbol{x}$ and $\boldsymbol{T}$ are the vibration displacement vectors and external torque vectors at each point, respectively.

Based on (1), the dynamics of the single-stage planetary gear system is modeled as shown in Fig. 1. In Fig. 1, *s* represents the sun gear, *r* represents the gear ring, *c* represents the planetary frame, and p_n represents the *n*th planetary gear. The *x*-axis passes through the center of the circle of the first planetary gear, and the coordinate system rotates at the speed of the planet frame. The support stiffness and damping of each component are represented by a virtual spring-damper unit with centralized parameters, i.e., k_{ij} denotes the stiffness, c_{ij} denotes the damping, i denotes the component $(i = s, r, c, p_n)$, and j denotes the degrees of freedom $(j = x, y, u)$. k_{spn} and c_{spn} are the time-varying stiffness and damping constant of the sun-planetary gear meshing pair, respectively, whereas k_{rpn} and c_{rpn} are the time-varying stiffness and damping constant of the ring-gear meshing pair, respectively. $e_{spn}(t)$ and $e_{rpn}(t)$ are the transmission errors of the *n*th sun-planetary gear meshing pair and inner ring-gear meshing pair, respectively. x_{spn} and x_{rpn} are the relative displacements along the active line of the sun sun-planetary gear and the inner ring gear, respectively, and can be expressed as:

$$\begin{cases} x_{spn} = -x_s \sin\psi_{sn} + y_s \cos\psi_{sn} + u_s + u_{pn} - e_{spn}(t) \\ x_{rpn} = u_r + u_{pn} - e_{rpn}(t) \end{cases} \tag{2}$$

where $\psi_{sn} = \psi_n - \alpha_s$; $\psi_{rn} = \psi_n + \alpha_r$; α_s and α_r are the meshing angles of the sun-planetary gear and the inner ring gear, respectively; $\psi_n = 2\pi(n-1)/N$ represents the circumferential angle of the *n*th planetary gear around the sun gear; and *n* is the number of planetary gears.

Specifically, mesh stiffness is a function of angular displacement and can be approximated as a square wave when the gear is free of faults. However, when the gear fails, the meshing of faulty teeth causes a reduction of the meshing stiffness function, and the reduction in meshing stiffness results in abnormal vibration of the gears.

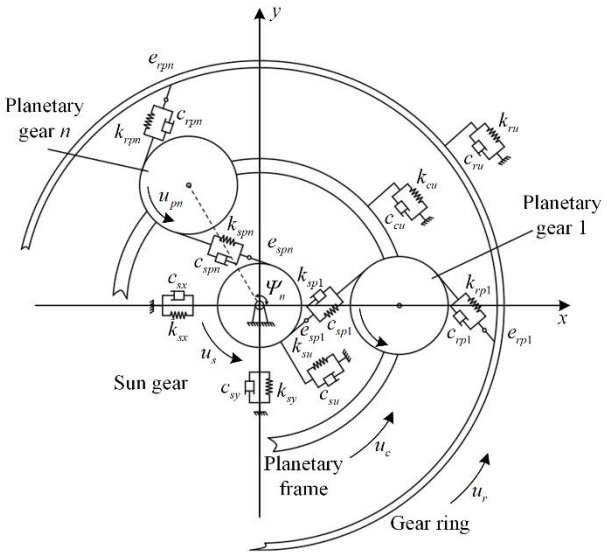


Fig. 1. Planetary gearbox dynamic model.

Variable speed operation of a wind turbine is considered in this paper. Two types of vibration signals, acceleration, and displacement, are selected as inputs. The acceleration signal is sensitive to high rotational speed and the displacement signal is sensitive to low rotational speed.

The vibration signal is the main input signal for the developed diagnostic model. However, because of the complex structure of the gearbox and continuous operation under variable speed and load conditions, the vibration of the gearbox is subjected to multiple excitation sources and complex transmission paths. The vibration signal is characterized by strong noise and multi-source coupled modulation. Therefore, non-vibration signals must be introduced as inputs to the diagnostic model.

Torque ripples caused by gear mesh can be observed at multiples of the mesh frequency, which is defined as the product of the number of teeth and the speed. These torque ripples can be expressed as friction torque, given as:

$$T_f = B\bar{\omega}_{\text{out}} + B\sum_n \omega_{\text{out},n}\cos(n\theta_{\text{out}} + \phi_{f,n}) \tag{3}$$

where T_f and B are the friction torque and friction coefficient, respectively; $\bar{\omega}_{\text{out}}$ is the averaged rotating speed of the output shaft; $\omega_{\text{out},n}$ and $\phi_{f,n}$ are the amplitude for the *n*th rotating speed of the output shaft and phase offset for the *n*th friction torque, respectively. The torque signal is used as the input to the diagnostic model because faults in the gearbox will cause torque ripples and thus affect the torque.

Torque variation and speed variation are closely related. Rotating machines are driven by output torque, which is electromagnetic torque amplified by the gearbox. When a local failure occurs in the planetary gearbox, a periodic shock torque T_{SB} is generated. For a wind turbine with a local gear failure, the motor speed is affected by the shock torque T_{SB} as follows:

$$\omega = \omega_0 - \frac{T_{SB}}{2\pi J}\sum_{n=1}^{\infty}\frac{1}{n}\sin(2\pi n f_c t) \tag{4}$$

where ω is the speed; J is the inertia; and f_c is the fault characteristic frequency. The speed signal is also used as an input signal to the model because gearbox faults also affect the motor speed.

In summary, in this paper, the acceleration, displacement, torque and speed signals are used as inputs of the diagnostic model, where the acceleration and displacement signals are contained in the vibration signal.

B. Denoising Auto Encoder

An autoencoder (AE) is a feature representation network based on a neural network, composed of input, hidden, and output layers, and is one of the typical models of deep learning. It reconstructs the input data through encoding and decoding operations to minimize the reconstruction error. In this way, the hidden layer representation of input data is obtained to achieve feature extraction.

The encoding stage of the AE is the process from input layer x to hidden layer h, which is specifically expressed as:

$$h = f_\theta(x) = \sigma(\boldsymbol{w}x + b) \tag{5}$$

where σ is a nonlinear active function, $\sigma(x) = 1/1+\mathrm{e}^{-x}$, parameter set $\theta = \{\boldsymbol{w}, b\}$; $\boldsymbol{w}$ is the weight matrix between the input and hidden layers; and b denotes the bias of the hidden layer. The decoding stage is the process of reconstructing the output layer z from the hidden layer h in a similar way, as:

$$z = g_{\theta'} = \sigma'(w'h + b') \tag{6}$$

where σ' is a nonlinear sigmoid active function; and $\theta' = \{w', b'\}$ is the parameter set. The internal parameters of the network are adjusted gradually by minimizing the reconstruction error function $L(x,z) = \|x - z\|^2$. The optimization method is the random gradient descent method, and the optimal parameter is expressed as:

$$\theta, \theta' = \arg\min_{\theta,\theta'} L(x, g_{\theta'}(f_\theta(x))) \tag{7}$$

A denoising auto-encoder (DAE) is a variant of AE, one which trains input data through noise pollution to increase the robustness of the network and prevent over-fitting [23]. As shown in Fig. 2, the input x is stochastically corrupted to $\tilde{x}$. The AE then maps it to h via the encoder f_θ and attempts to reconstruct x via the decoder g_θ, producing reconstruction output z. Reconstruction error is measured by the loss L (x, z), given as:

$$L(x,z) = -\sum_{k=1}^{d}[x_k \log z_k + (1-x_k)\log(1-x_k)] \tag{8}$$

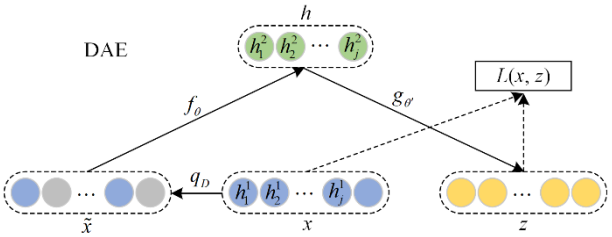


Fig. 2. Denoising auto-encoder architecture.

C. Cycle Reservoir with Regular Jumps Network

Reference [26] proposes a deterministic cycle reservoir with regular jumps (CRJ) network. This avoids the uncertainty caused by the randomly generated weight matrices in the reserve pool through the deterministic sparse connections of the CRJ network reserve pool neurons. The CRJ structure is shown in Fig. 3. The CRJ network consists of an input layer, a storage pool and an output layer. The hidden layer has two fixed positive values r_c and r_j, where the weights in the hidden layer are connected in a cycle r_c with two-way jumps r_j as given in Fig. 3. All the ring weights have the same value $r_c > 0$ and all jump weights also share the same value $r_j > 0$.

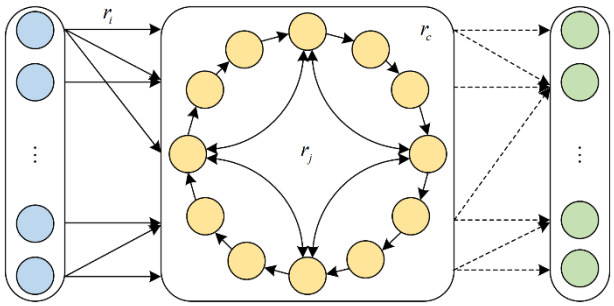


Fig. 3. CRJ architecture.

In the method proposed here, the memory pool of CRJ is used as the information processing medium to map the input feature vector from the low- to the high-dimensional state space and to train the connection weight of the network in the high-dimensional state space to achieve feature fusion. The network is represented as:

$$x(t+1) = f(Vs(t+1) + Wx(t)) \tag{9}$$

$$y(t+1) = Ux(t+1) \tag{10}$$

where t is the sample index; and f is the activation function of the hidden layer; V is the connection weights that link the input units with the hidden units; W is the connection weights between the hidden units; and U is the weights that connect hidden units to output units. The output weights U are computed using linear regression as:

$$U = (x^{\mathrm{T}}x + \lambda^2 \boldsymbol{I})^{-1}x^{\mathrm{T}}y \tag{11}$$

where $\boldsymbol{I}$ is the identity matrix; and $\lambda > 0$ is a regularization factor.

The loss function for evaluating the best readout weights for CRJ and ESN is the normalized mean square error (NMSE), measured as:

$$N_{\mathrm{MSE}}=\frac{\left\langle\left\|\hat{y}(t)-y(t)\right\|^{2}\right\rangle}{\left\langle\left\|y(t)-\left\langle y(t)\right\rangle^{2}\right\|\right\rangle} \tag{12}$$

where $\langle\cdot\rangle$ denotes the empirical mean; and $\|\cdot\|$ denotes the Euclidean.

D. Competitive Swarm Optimizer

A heuristic parameter optimization method is used in the training process of the above CRJ network. Since the setting of parameters has a large impact on network performance, a competitive particle swarm optimization algorithm is introduced to obtain the optimal combination of parameters and reduce the influence of human factors on the feature fusion network. The competitive particle swarm algorithm is a new particle swarm optimization algorithm proposed in recent years. It is simple to implement and fast to converge, and is widely used to solve single-objective optimization problems [27]. Each particle represents a candidate solution for the optimization problem, and the particles iterate until they find the optimal solution [28]. Each particle has its individual position and velocity, denoted as $X_p=\left\{x_p^1,\ x_p^2,\ \cdots,\ x_p^D\right\}$ and $V_p=\left\{v_p^1,\ v_p^2,\ \cdots,\ v_p^D\right\}$, respectively, while D is the number of decision variables of the optimization problem.

The powerful particle update mechanism is the advantage of the competitive particle swarm algorithm over traditional particle swarm algorithms. First, the particle swarm state of the kth iteration is denoted as $P(k)=\left\{X_1(k),X_2(k),\cdots,X_N(k)\right\}$. After calculating the fitness of all particles in $P(k)$, the group is randomly divided into two equal groups. Then the two groups of particles will compete with each other. The particle with better fitness will be the winner, directly entering $P(k+1)$, and the other particle will be the loser and will be updated based on the winner and transferred to $P(k)$. For the convenience of expression, the positions and speeds of the winner and loser are respectively denoted as $X_{p,w}(k)$, $V_{p,w}(k)$, $X_{p,l}(k)$, and $V_{p,l}(k)$. Then the update process of the loser is:

$$\begin{aligned}V_{p,l}(k+1)=&\,r_1(p,k)V_{p,l}(k)+r_2(p,k)(X_{p,w}(k)-\\&X_{p,l}(k)+\bar{\lambda}r_3(p,k)(\bar{X}(k)-X_{p,l}(k))\end{aligned} \tag{13}$$

$$X_{p,l}(k+1)=X_{p,l}(k+1)X_{p,l}(k)+V_{p,l}(k+1) \tag{14}$$

where $\mathrm{r}_1(p,k)$, $\mathrm{r}_2(p,k)$, and $\mathrm{r}_3(p,k)$ are random coefficients obeying a uniform distribution [0,1]; $\bar{X}_p(k)$ is the global average position of the particle population; and $\bar{\lambda}$ is the parameter controlling $\bar{X}_p(k)$.

III. Proposed Fault Diagnostic Method Based on Multi-sensor Information Fusion

In the fault diagnosis of planetary gearboxes, the signals collected by different sensors have different numerical information and fault features [16]–[18]. The existing deep learning models are mainly based on unidimensional feature extraction, which is prone to missing information, and the Softmax classifier, which has excellent classification performance and is widely used, performs poorly in the processing of fused information. Therefore, to make full use of the fault feature information, this paper proposes a multi-sensor signals feature fusion network, as shown in Fig. 4.

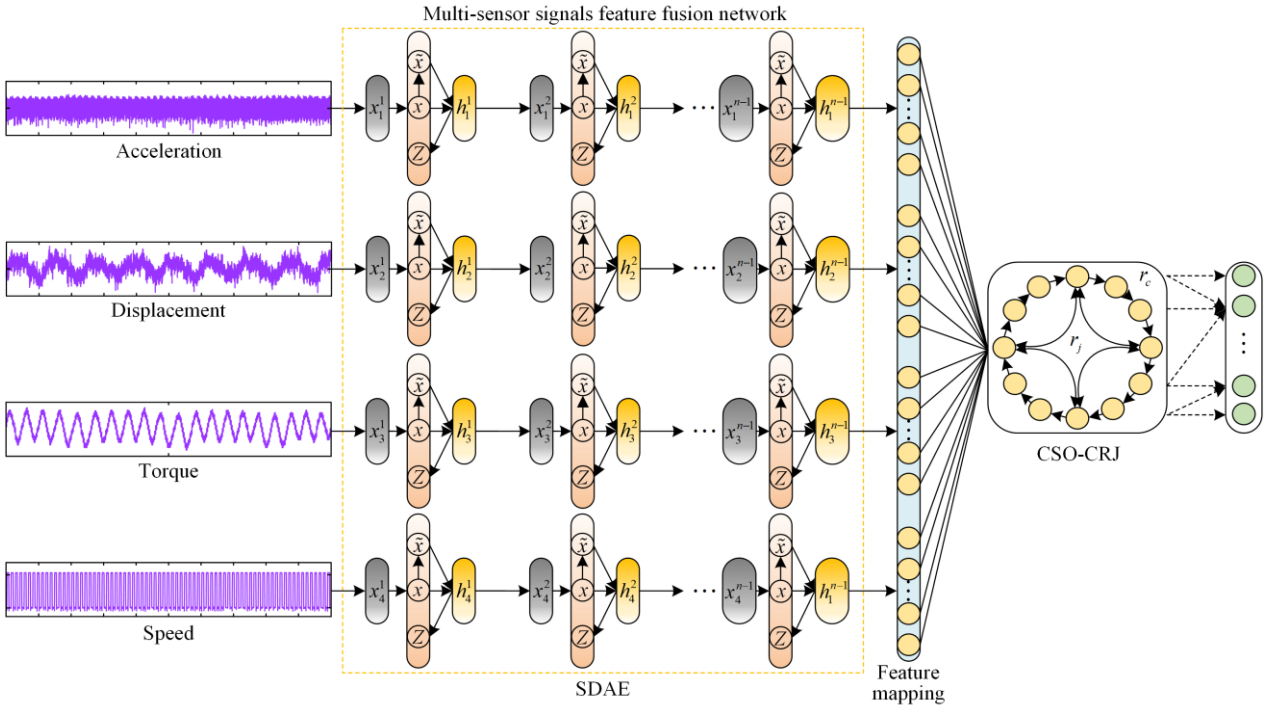


Fig. 4. Framework of MSFFN.

A multi-sensor signals stacked denoising auto-encoder (MS-SDAE) constructed by multiple SDAEs is used in the network to extract feature information from each sensor signal simultaneously. CSO-CRJ is then used to fuse information and output diagnostic results.

A. Improved Stacked Denoising Auto-encoders

Several DAEs are stacked to form an SDAE, as shown in Fig. 5. Each DAE uses the hidden layer output of the previous DAE as its original input data. After the noise has been added, the contaminated data is used for training. Feature learning is implemented by the stacked DAEs in the SDAE model. The activation function of traditional SE often uses a sigmoid function or rectified linear unit (ReLU). The sigmoid function is prone to gradient disappearance and consumes a lot of computational resource, whereas the sparse features extracted by ReLU negatively affect the decoder to generate samples and make the fit insufficient.

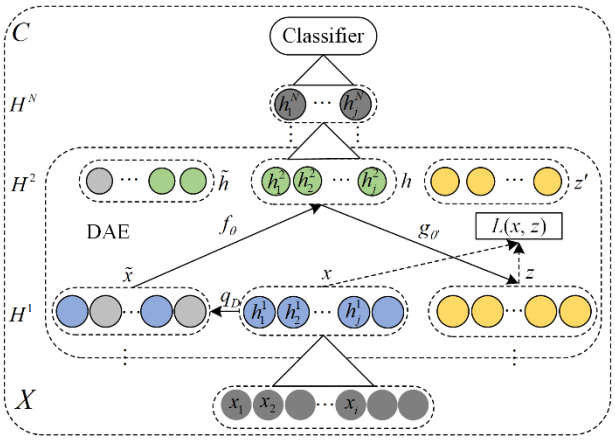


Fig. 5. Stacked denoising auto-encoder architecture.

The parametric rectified linear unit (PReLU) is:

$$\mathrm{PReLU}(x_i)=\begin{cases} x_i & (x_i>0) \\ a_i x_i & (x_i<0) \end{cases} \tag{15}$$

PReLU allows negative outputs, and has a very small number of additional parameters compared to ReLU. The risk of overfitting the network is not significant, while poor fit due to excessive feature sparsity can be avoided so that model generalization performance is not affected. Therefore, PReLU is used as the activation function.

Batch normalization (BN) is used to solve the problems of gradient explosion and gradient disappearance that tend to occur during the training process of SDAE. BN is performed by preprocessing the input with whitening transformation:

$$\bar{x}=\frac{x-E(x)}{\sqrt{\mathrm{var}(x)+\varepsilon}} \tag{16}$$

where $E(x)$ is the mean of one of the batches of inputs x; $\mathrm{var}(x)$ is the variance of that batch of data; and ε is a very small positive number. Whitening transformation tends to weaken the performance of the model. To improve the expressiveness of the model, a "scale and translate" operation is introduced as:

$$y=\lambda\bar{x}+\beta\equiv \mathrm{BN}_{\lambda,\beta}(x) \tag{17}$$

After "scale and translate", the network has strong nonlinear expressiveness. It avoids getting trapped at both ends of the nonlinear interval, which could make the network converge slowly.

A BN layer is added between each fully connected layer to speed up the SDAE network training to avoid overfitting, while PReLU is used to avoid the negative impact of over-sparse features on the decoder.

B. CSO-CRJ

As described in Section 1.2, the training process of CRJ is a heuristic parameter optimization process, which makes it difficult to obtain the optimal combination of parameters. In addition, the generalization performance of CRJ will be affected if the memory pool size, spectral radius, sparsity, and input variation factor are set manually.

The CSO algorithm has the advantages of simple operation, fewer parameters required, and easy implementation. Therefore, to achieve the information fusion of different feature domains efficiently, CSO-CRJ is proposed here, and the specific optimization steps are:

1) Initialization parameters, setting the maximum number of iterations, population size, and the key parameters of CRJ (storage pool size, spectral radius, sparsity, and input variation factor) as optimization variables.

2) Chaotic initialization of the initial position and velocity of the particle swarm using logistic mapping.

3) Calculating the root mean square error of the CRJ model corresponding to each particle in the training set according to (12) as the fitness value.

4) Making two groups of particles compete two by two and marking them as winners and losers respectively according to the magnitude of their fitness values.

5) Updating the velocity and position of the defeated particle according to (13) and (14), and calculating the fitness value, the global optimum, and the corresponding optimal solution of the updated particle.

6) If the maximum number of iterations or the fitness value is reached, the optimization ends and goes to step 7), otherwise it returns to step 4).

7) The optimal particle, i.e., the optimal network parameter combination, is obtained.

The feature vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \cdots, \boldsymbol{x}_i$ are obtained by feature extraction of SDAE. The feature vectors are completely concatenated to construct the input vector $\boldsymbol{x}$ for CRJ, while CSO is used to optimize the CRJ structural parameters. Finally, CSO-CRJ is trained and diagnostic results are generated. In summary, the fault diagnostic process based on MSFFN is shown in Fig. 6.

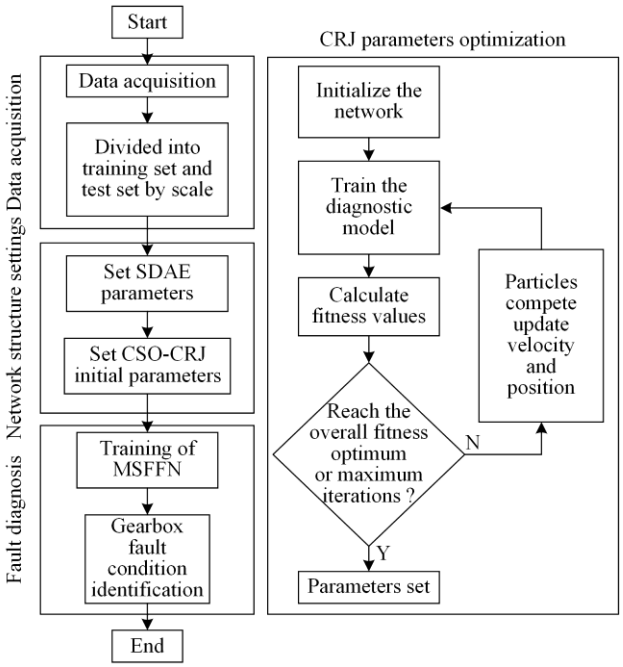


Fig. 6. Flowchart of the proposed method.

IV. Experimental Validation

A. Experimental Set-up

1) Data Preparing

The wind turbine gearbox fault diagnosis test bench is shown in Fig. 7, including the sensor locations. Figure 8 shows the planetary gear used in the experiments, including four states of normal condition (NC), chipped tooth fault (CTF), surface wear fault (SWF), and missing tooth fault (MTF). To collect the fault data, a planetary gear in the planetary gearbox is selected and replaced with the faulty planetary gear. The electric motor drives the planetary gear through torque encoded sensor, and then the speed is reduced through the parallel gearbox. The torque signal and speed signal used in this paper are directly from a fixed-position torque-encoded sensor. A displacement sensor is installed on the drive shaft to better respond to low-frequency vibration caused by unbalance, misalignment, loose contacts, etc. due to gear fault. The acceleration sensor is installed on the planetary gearbox. In this case, the collected signal contains more information about the fault.

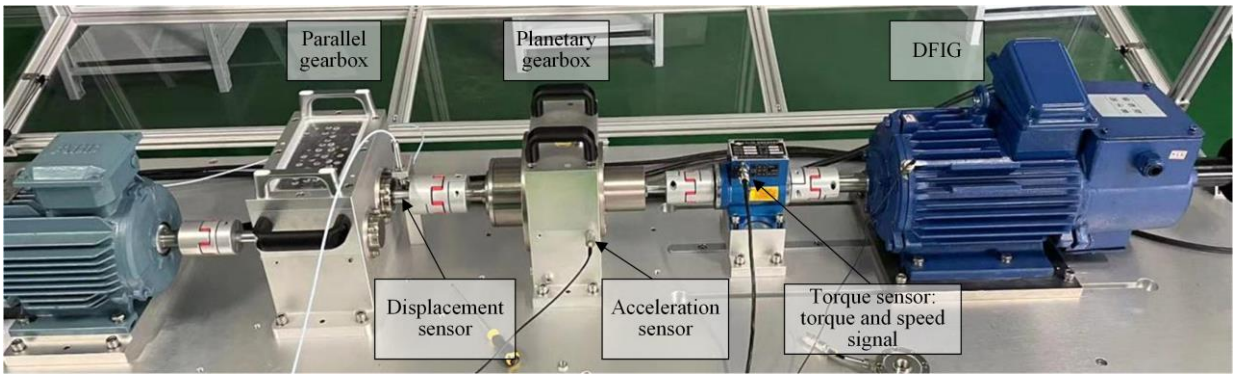


Fig. 7. Wind turbine fault diagnosis test bench and three types of sensors.

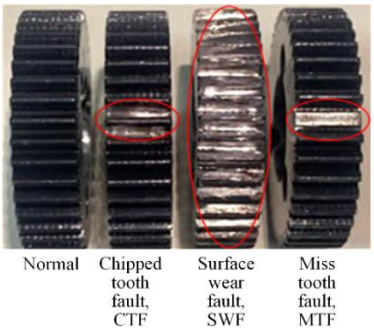
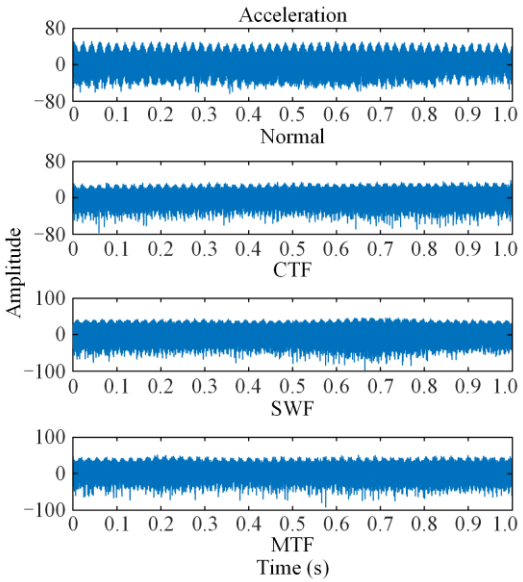


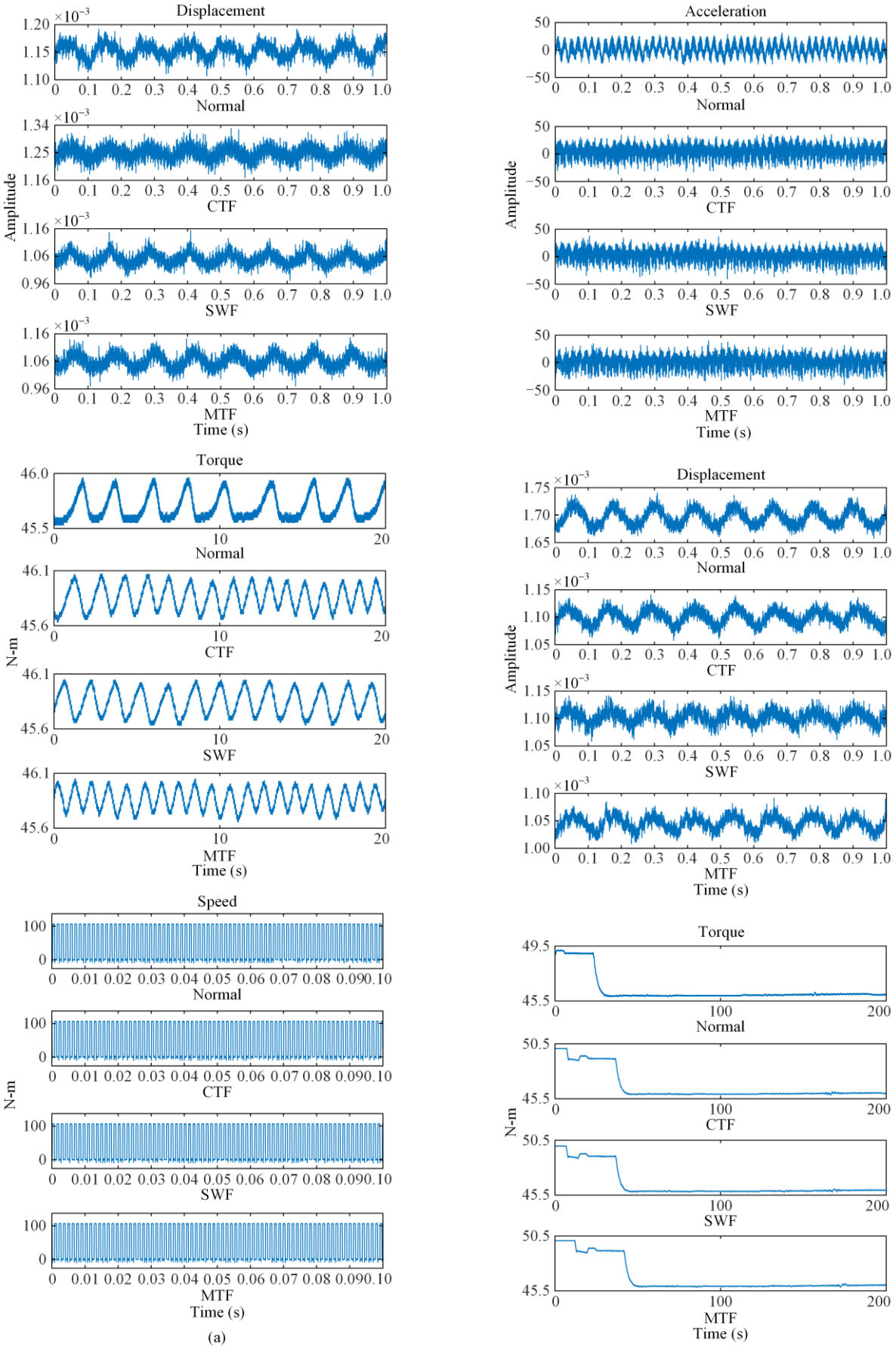
Fig. 8. Fault settings.

a) Constant Speed Experiment

The motor speed is set at 1500 rpm and the sensor sampling rate is set to 48 kHz. Four signals are collected from different health conditions of the planetary gearbox in four operating conditions, as shown in Fig. 9(a).

b) Variable Speed Experiment




Displacement
Acceleration
Torque
Displacement
Speed
Torque
Normal
CTF
SWF
MTF
Amplitude
N-m
Time (s)

(a)

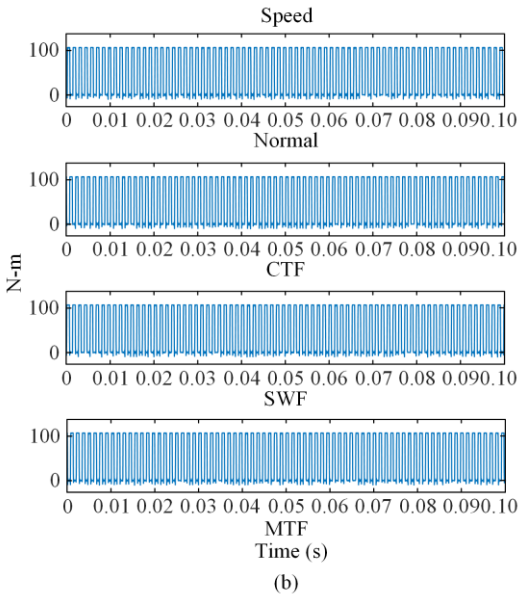


Fig. 9. Signal of planetary gearbox at normal and failure conditions. (a) Constant speed. (b) Variable speed.

Changing the load and adjusting the motor speed, the various signal waveforms in the variable speed condition are shown in Fig. 9(b), and the speed variation is shown in Fig. 10.

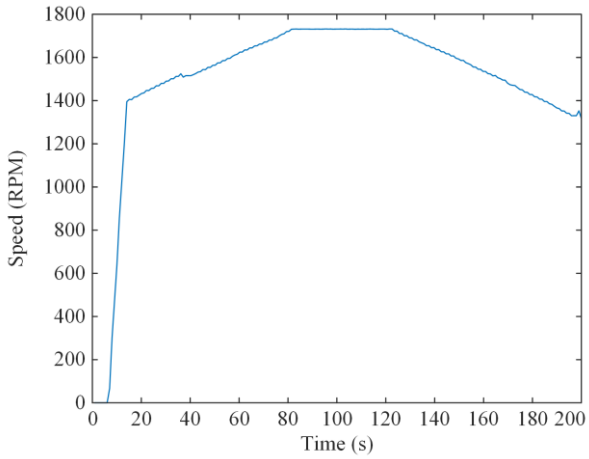


Fig. 10. Speed curve of planetary gearbox.

The signals generated by rotating machinery are periodic. Thus, to ensure that each sample data contains at least one complete failure cycle, and to facilitate the training of the network, each sample in the sample set contains 2048 sampling points and 500 samples for each working condition and speed, resulting in a total of 4000 samples.

2) Parameter Set

The number of nodes in the hidden layer of MS-SDAE for each sensor signal sample is shown in Table I, where the number of nodes in the first layer is the input dimension of each sample. The training process uses the Adam optimization algorithm, the initial learning rate is set to 0.01, and the number of training iterations is 200.

TABLE I
STRUCTURAL PARAMETERS OF MS-SDAE

Input	Layer size	Layers
Displacement	2048/1024/512/128/32/4	5
Acceleration	2048/1024/512/128/32/4	5
Speed	1024/512/128/32/4	4
Torque	2048/1024/512/128/32/4	5

B. Experimental Results

1) Fault Diagnosis in Case of Sufficient Samples

The training and test sets with sufficient samples are randomly divided in the ratio of 3:1, i.e., 1500 samples for constant speed training and 500 samples for constant speed testing, and 1500 samples for variable speed training and 500 samples for variable speed testing. The ①CNN [8], ②MSCNN [12], ③SDAE, ④deep enhanced fusion network (DEFN) [29], ⑤ MS-SDAE, ⑥ MS-SDAE-CRJ and ⑦ MS-SDAE-PSO-CRJ are used for comparison. MSCNN takes the 3 scales after coarse granulation of samples as input, and the individual CNN structure is the same as ①. SDAE takes the same multi-sensor signals as input, and the number of hidden layer nodes is the same as in Section IV. A.2). SDAE uses the original activation function and does not integrate the BN layer. DEFN extracts error information using a four-dimensional sparse self-encoder after feature enhancement of the input. The MS-SDAE structure parameters are the same as SDAE. A Softmax classifier is used for fault classification after the features are fully connected. The MS-SDAE-CRJ parameters are artificially set with a storage pool size of 300, a spectral radius setting of 1, a sparsity of 10%, and an input variation factor of 0.1.

Each experimental set is repeated 10 times to reduce the effect of randomness. The average classification accuracy and training time for the test set are shown in Table II, where accuracy is defined as the percentage ratio of samples that are correctly classified to the total number of samples in the test. As seen in Table II, each method is capable of reaching a high classification accuracy with sufficient training. The diagnostic results of the MSFFN model proposed in this paper are 98.54±0.43% for the constant speed condition and 98.23±0.49% for the variable-speed condition. Compared to MSCNN and DEFN, the proposed method has higher accuracy and lower standard deviation.

Compared to the conventional single-signal input deep learning network combined with the Softmax classifier, the multi-sensor signals feature extraction combined with the Softmax classifier is less effective in fault identification. Thus, it is concluded that the Softmax classifier is less capable of fusing information from different sensor signals. Therefore, in this paper, CRJ is introduced to perform feature fusion and output diagnostic results. The diagnostic results of ⑥ and ⑦ show that the structural

parameters of the CRJ greatly affect its performance. Appropriate optimization algorithms must be introduced because the structural parameters set by artificial experience tend to lead to insufficient generalization. The comparison of ⑦ with MSFFN shows that the feature fusion and classification performance of CSO-CRJ is better than that of PSO-CRJ, and the iteration speed of CSO is also faster.

TABLE II
COMPARISON OF FAULT DIAGNOSIS RESULTS

Method	Accuracy ± standard deviation (%)		Average training/testing time (s)
	Constant speed	Variable speed	
①	93.13±0.88	92.65±0.76	242.71/0.43
②	95.90±0.63	95.28±0.92	312.61/0.47
③	93.03±1.78	92.83±1.17	179.36/0.27
④	95.78±0.43	95.30±0.56	256.74/0.44
⑤	95.03±0.47	94.83±0.49	227.22/0.27
⑥	95.60±0.51	95.26±0.74	304.81/0.49
⑦	96.03±0.47	95.83±0.49	481.93/0.68
MSFFN	98.54±0.43	98.23±0.49	386.10/0.51

Training MSFFN takes more time than the conventional methods because more weights are used and the network is trained more often during optimization iterations. However, because of the improvement of the SDAE activation function and the BN layer, the testing time required for the model increases by a small margin, whereas the trained diagnostic model does not require large computing resources.

2) Fault Diagnosis with Insufficient Samples

In the actual operation of a wind turbine, the frequency of planetary gearbox failure is low. The fault diagnosis in daily work often faces the problem of insufficient samples. To simulate the small sample case, the number of samples in the training set is defined to be 100% when the samples are sufficient. For example, the network is trained with 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% of the total number of samples in the training set, while keeping the number of samples in the test set constant. Mean diagnostic results after 10 repetitions in each set are shown in Table III. In the case that the ratio is greater than 50%, the classification accuracy of each method tends to be stable, indicating that the training samples are sufficient at this rate.

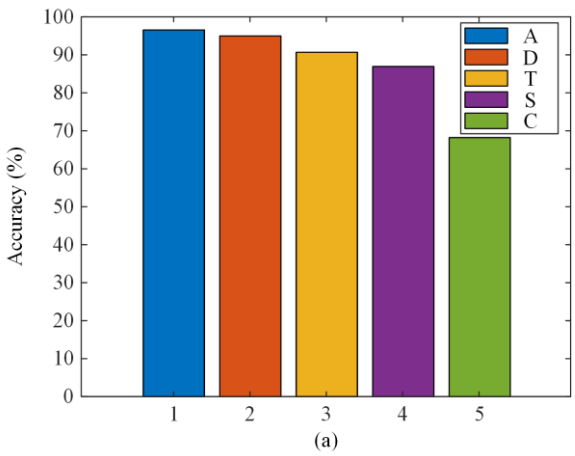
The classification capabilities of CNN, MSCNN and SSAE decrease significantly when the ratio of training samples is less than 50%. In contrast, the feature fusion networks DEFN, and MSFFN are robust to the reduction of samples. MSFFN performs better overall, with more than 90% classification accuracy even with only 10% training examples.

TABLE III
COMPARISON OF DIAGNOSTIC RESULTS WITH DIFFERENT SAMPLE RATIOS

Ratio (%)	Accuracy (%)				
	①	②	③	④	MSFFN
10	62.10	71.45	68.34	85.40	90.14
20	75.24	75.28	77.58	91.17	93.69
30	83.91	88.64	84.20	93.52	95.16
40	86.70	89.26	85.97	94.07	95.04
50	87.03	91.43	87.36	94.32	96.51
60	88.60	92.56	88.24	94.78	97.71
70	90.48	92.87	89.65	94.69	97.32
80	91.54	94.43	91.12	95.11	97.91
90	92.16	94.70	92.73	95.20	98.08
100	92.65	95.28	92.83	95.30	98.23

3) Fault Diagnosis with Different Input Signals

In this paper, torque and speed signals are added to the input to the diagnostic model to address the shortcomings of the diagnostic methods that use a single sensor signal as input. Diagnostic models trained with a sample ratio of 80% and 20% are used for validation. Input signals include vibration, current (C), torque (T), and speed (S), while vibration includes acceleration (A) and displacement (D). The diagnostic accuracy of MSFFN with a single vibration signal as input can reach approximately 95% with sufficient samples, as shown in Fig. 11(a). The diagnostic accuracies are 89.51% and 85.82% when the torque and speed signals are used as input, respectively. However, when the current signal is used as the input, the accuracy is only 67%. This indicates that the speed and torque signals have greater fault information than the current signal. However, with insufficient numbers of samples, the diagnostic accuracy of MSFFN with a single input signal shows varying degrees of degradation, as seen in Fig. 11(b). In comparison, Figs. 11 (c) and (d) show the results of fault diagnosis when multiple signals are used as inputs. As can be seen, the more the number of input signals, the higher the accuracy of the diagnosis. Despite the insufficient samples, the model with multiple signal inputs can still output more accurate diagnosis results.



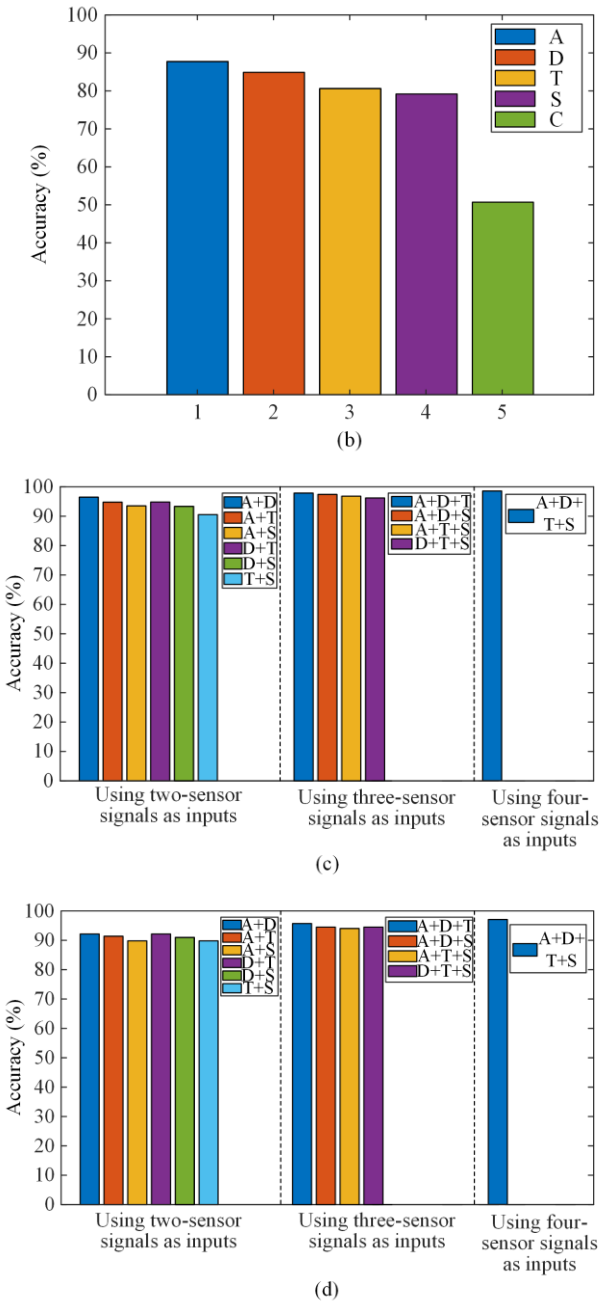


Fig. 11. Diagnostic accuracy of MSFFN when single signal and multiple signals are used as input. (a) Single signal and constant speed. (b) Single signal and variable speed. (c) Multiple signals and constant speed. (d) Multiple signals and variable speed.

In summary, this paper has verified the feasibility of adding torque and speed signals to vibration signals as the inputs to the diagnostic model.

4) Fault Diagnosis Under Strong Noise Scenario

The vibration signals are easily affected by external noise during actual wind turbine operation. As a result, conventional fault diagnosis models in practical applications lack generalizability. The robustness of MSFFN against noise is examined here with a fault diagnosis model trained with 50% of training samples. Gaussian white noise is added to the original constant and variable speed test set to simulate the external environment. Given that the diagnostic model does not have access to real-time noise information during training, the proposed training and testing scheme meets the actual operation scenarios. The noise content of the test sample is expressed by the signal-to-noise ratio as:

$$R_{\mathrm{SN}} = 10\lg\left(\frac{p_{\mathrm{signal}}}{p_{\mathrm{noise}}}\right) \tag{18}$$

The diagnostic accuracy of MSFFN with different noise levels is tested by mixing different levels of noise into the test set. The mean values of the diagnostic results are shown in Fig. 12. As seen, for test sets with SNR greater than 8 dB, each method performs well in diagnostic accuracy. However, the MSFFN has a significant advantage for test sets with signal-to-noise ratios below 8 dB. In addition, the reliability of traditional deep learning methods is greatly reduced in the strong noise scenario of −2 dB SNR of test samples, while MSFFN can still achieve more than 85% diagnostic accuracy. After adding strong noise to the test set, traditional deep learning methods perform poorly in the variable speed scenario, and the classification results are nearly randomly divided. The MSFFN diagnostic model, on the other hand, still can provide diagnostic accuracy of over 85% when there is strong noise.

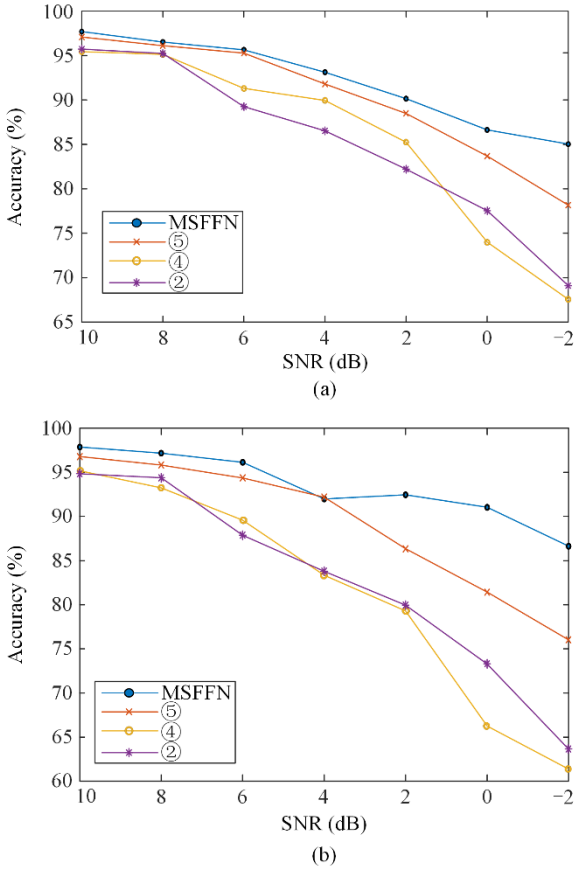


Fig. 12. Comparison of diagnostic accuracy at different signal-to-noise ratios. (a) Constant speed. (b) Variable speed.

In summary, the proposed method in this paper is more robust than conventional deep learning methods. Since noise does not affect the fault information in all signals, feature extraction from multi-sensor signals can thus significantly improve the feature extraction ability of deep learning methods when there is strong noise.

V. Conclusion

This paper has proposed a method for wind turbine gearbox fault diagnosis in variable speed conditions, insufficient samples, and strong noise scenarios by fusing the fault feature information from vibration, torque, and speed signals. The connection between signals and gearbox failure is analyzed. A multi-signal SDAE is developed and improved to output the probabilities of different gearbox fault types by using the features extracted from different sensory signals. For the fusion of fault features extracted from multiple signals, a CRJ network is designed to replace the conventional Softmax classifier. The key conclusions are:

1) Experiments are conducted on a wind turbine test rig to validate the effectiveness of the proposed gearbox fault diagnosis method. The results in sufficient sample scenarios show that the proposed network can achieve good diagnostic accuracy

2) For both constant and variable speed conditions, with sufficient samples, the average accuracy of MSFFN is 98.23%. Therefore, the proposed MSFFN in this paper can be applied to planetary gearbox fault diagnosis in variable speed conditions.

3) Experimental results of fault diagnosis show that MSFFN is robust when there are insufficient samples. The proposed method maintains high diagnostic accuracy when multiple signals are used as inputs in the case of insufficient samples. Compared with conventional deep learning methods, MSFFN is less sensitive to external noise. The results show that the developed network can be effectively applied to the fault diagnosis of a planetary gearbox of wind turbines in severe conditions.

Acknowledgment

Not applicable.

Authors' Contributions

Yao Zhao: the construction of the paper framework. Ziyu Song: the full text writing and software simulations. Dongdong Li: the construction of the network. Shunfu Lin: construction of the network. Rongrong Qian: collection of experimental data. All authors read and approved the final manuscript.

Funding

This work is supported by the Shanghai Rising-Star Program (No. 21QC1400200), the Natural Science Foundation of Shanghai (No. 21ZR1425400), and the National Natural Science Foundation of China (No. 52377111).

Availability of Data and Materials

Not applicable.

Declarations

Competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Authors' Information

Yao Zhao received the B.S. degree in automation from Anhui University, Hefei, China, in 2009, the M.S. degree in electrical engineering from Shanghai Maritime University, Shanghai, China, in 2011, and the Ph.D. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2016. He is currently an associate professor with the Shanghai University of Electric Power, Shanghai, China. His main research interests include electrical machines and power electronization of power system.

Ziyu Song received the B.Eng. degree in electrical automation from the Qingdao University of Science and Technology, Qingdao, China, in 2020. He is currently pursuing the M.S. degree in electrical engineering from the Shanghai University of Electric Power. His research interest includes the fault diagnosis of wind turbines.

Dongdong Li was born in Anhui, China, in 1976. He received the B.Sc. degree in electrical engineering from Zhejiang University, in 1998, and the Ph.D. degree in electrical engineering from Shanghai Jiao Tong University, in 2005. From 1998 to 2000, he was with Wuhu power Plant as an Electrical Engineer. He is currently a Professor and the Dean of the School of Electric Engineering, Shanghai University of Electric Power, Shanghai, China. His research interests include the analysis of electric power systems, new energy systems, smart grid, and the power electronization of power systems.

Rongrong Qian received the B.S. and Ph.D. degrees in automation from Anhui University and University of Science and Technology of China, Hefei, China, in 2009 and 2016, respectively. She is currently with the Design Research and Development Center, AECC Commercial Aircraft Engine Co., Ltd., as an Engineer. Her area of interest includes the control and health monitoring of the aircraft engine.

Shunfu Lin received his B.S. and Ph.D. degrees from the University of Science and Technology of China in

2002 and 2007, respectively. He is currently a professor and PhD supervisor. He is mainly engaged in research in the field of smart grid user side, fault diagnostic technology for electrical equipment.

REFERENCES

[1] M. Nadour, A. Essadki, and T. Nasser, "Improving low-voltage ride-through capability of a multimegawatt DFIG based wind turbine under grid faults," *Protection and Control of Modern Power Systems*, vol. 5, no. 4, pp. 1-13, Oct. 2020.

[2] J X. Ge, J. Qian, and Y. Fu *et al*., "Transient stability evaluation criterion of multi-wind farms integrated power system," *IEEE Transactions on Power Systems*, vol. 37, no. 4, pp. 3137-3140, Jul. 2022.

[3] M. Asr, M. Ettefagh, and R. Hassannejad *et al*., "Diagnosis of combined faults in rotary machinery by non-naive Bayesian approach," *Mechanical Systems and Signal Processing*, vol. 85, pp. 56-70, 2017.

[4] Y. Peng, W. Qiao, and F. Cheng *et al*., "Wind turbine drivetrain gearbox fault diagnosis using information fusion on vibration and current signals," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-11, 2021.

[5] D. Wei, T. Han, and F. Chu *et al*., "Weighted domain adaptation networks for machinery fault diagnosis," *Mechanical Systems & Signal Processing*, vol. 158, pp. 107744, 2021.

[6] R. Liu, B. Yang, and E. Zuo *et al*., "Artificial intelligence for fault diagnosis of rotating machinery: a review," *Mechanical Systems and Signal Processing*, vol 108, pp. 33-47, 2018.

[7] D. Huang, W. A. Zhang, and F. Guo *et al*., "Wavelet packet decomposition-based multiscale CNN for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Cybernetics*, vol. 53, no. 1, pp. 443-453, Jan. 2023.

[8] G. Jiang, H. He, and J. Yan *et al*., "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3196-3207, Apr. 2019.

[9] B. Chen, C. Shen, and D. Wang *et al*., "A lifelong learning method for gearbox diagnosis with incremental fault types," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-10, 2022.

[10] M. Chang and Y. Shen, "Wind turbine bearing fault diagnosis strategy based on improved convolutional neural network," *Power System Protection and Control*, vol. 49, no. 6, pp. 131-137, Apr. 2021. (in Chinese)

[11] D. Li, W. Zhou, and X. Zheng *et al*., "Fault diagnosis and research of wind turbine planetary gearboxes based on multiple fractal spectra and support vector machine," *Power System Protection and Control*, vol. 45, no. 11, pp. 43-48, Jun. 2017. (in Chinese)

[12] J. S. L. Senanayaka, H. Van Khang, and K. G. Robbersmyr, "Multiple classifiers and data fusion for robust diagnosis of gearbox mixed faults," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 8, pp. 4569-4579, Aug. 2019.

[13] A. G. Kavaz and B. Barutcu, "Fault detection of wind turbine sensors using artificial neural networks," *Journal of Sensors*, pp. 1-11, 2018.

[14] L. Hong and J. S. Dhupia, "A time-domain fault detection method based on an electrical machine stator current measurement for planetary gear-sets," in *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, *Wollongong*, *NSW*, *Australia*, 2013, pp. 1631-1636.

[15] N. Chai, M. Yang, and Q. Ni *et al*., "Gear fault diagnosis based on dual parameter optimized resonance-based sparse signal decomposition of motor current," *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3782-3792, Jul.-Aug. 2018.

[16] L. Jing, T. Wang, and M. Zhao *et al*., "An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox," *Sensors*, vol. 17, no. 2, pp. 414-419, Feb 2017.

[17] W. Zhou, T. G. Habetler, and R. G. Harley, "Stator current-based bearing fault detection techniques: a general review," in *2007 IEEE International Symposium on Diagnostics for Electric Machines*, *Power Electronics and Drives*, *Cracow*, *Poland*, 2007, pp. 7-10.

[18] Z. Feng and M. J. Zuo, "Fault diagnosis of planetary gearboxes via torsional vibration signal analysis," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 401-421, 2013.

[19] J. Yoon and D. He, "Planetary gearbox fault diagnostic method using acoustic emission sensors," *IET Science, Measurement & Technology*, vol. 9, no. 8, pp. 936-944, 2015.

[20] Q. Zeng, M. Zainab, and Y. Shao *et al*., "Planetary gear fault diagnosis based on instantaneous angular speed analysis," in *2017 23rd International Conference on Automation and Computing (ICAC)*, Sep. 2017, pp. 1-6.

[21] T. Xie, X. Huang, and S. K. Choi, "Intelligent mechanical fault diagnosis using multisensor fusion and convolution neural network," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3213-3223, May 2022.

[22] K. R. Shahi, P. Ghamisi, and B. Rasti *et al*., "Unsupervised data fusion with deeper perspective: a novel multisensor deep clustering algorithm," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 284-296, 2022.

[23] Z. Li, Z. Jiao, and A. He *et al*., "A denoising-classification neural network for power transformer protection," *Protection and Control of Modern Power Systems*, vol. 7, no. 4, pp. 1-14, Oct. 2022.

[24] 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, Jul. 2017.

[25] P. Vincent, H. Larochelle, and Y. Bengio *et al*., "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning*, pp. 1096-1103, Jul. 2008.

[26] A. Rodan and P. Tiňo, "Simple deterministically constructed cycle reservoirs with regular jumps," *Neural Computation*, vol. 24 no. 7, pp. 1822-1852, 2012.

[27] R. Cheng and Y. Jin, "A competitive swarm optimizer for large scale optimization," *IEEE Transactions on Cybernetics*, vol. 45, no. 2, pp. 191-204, Feb. 2014.

[28] W. Deng, J. Xu, and H. Zhao *et al*., "A novel gate resource allocation method using improved PSO-based QEA," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23. no. 3, pp. 1737-1745, Mar. 2020.

[29] S. R. Saufi, Z. A. B Ahmad, and M. S. Leong *et al*., "Gearbox fault diagnosis using a deep learning model with limited data sample," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6263-6271, Oct. 2020.