

## Article

# Research on an Improved Auxiliary Classifier Wasserstein Generative Adversarial Network with Gradient Penalty Fault Diagnosis Method for Tilting Pad Bearing of Rotating Equipment

Chunlei Zhou <sup>1,2</sup> , Qingfeng Wang <sup>1,2,\*</sup> , Yang Xiao <sup>1,2</sup>, Wang Xiao <sup>3</sup> and Yue Shu <sup>4</sup>

<sup>1</sup> School of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029, China; zhoucl@buct.edu.cn (C.Z.); xiao0619yang@buct.edu.cn (Y.X.)

<sup>2</sup> Beijing Key Laboratory of Health Monitoring and Self-Recovery of High-End Machinery Equipment, Beijing University of Chemical Technology, Beijing 100029, China

<sup>3</sup> Western Branch of National Pipe Network Group United Pipeline Company Ltd., Urumqi 830013, China; xiaowang@pipechina.com.cn

<sup>4</sup> Hefei General Machinery Research Institute Company Ltd., Hefei 230031, China; shuyuecr@mail.ustc.edu.cn

\* Correspondence: wangqf2422@buct.edu.cn

**Abstract:** The research on fault diagnosis methods based on generative adversarial networks has achieved fruitful results, but most of the research objects are rolling bearings or gears, and the model test data are almost all derived from laboratory bench test data. In the industrial Internet environment, equipment-fault diagnosis is faced with the characteristics of large amounts of data, unbalanced data samples, and inconsistent data file lengths. Moreover, there are few research results on the fault diagnosis of rotor systems composed of shafts, impellers or blades, couplings, and tilting pad bearings. There are still shortcomings in the operational risk evaluation of rotor systems. In order to ensure the reliability and safety of rotor systems, an Improved Auxiliary Classifier Wasserstein Generative Adversarial Network with Gradient Penalty (IACWGAN-GP) model is constructed, a fault diagnosis method based on IACWGAN-GP for tilting pad bearings is proposed, and an intelligent fault diagnosis system platform for equipment in an industrial Internet environment is built. The verification results of engineering case data show that the fault diagnosis model based on IACWGAN-GP can adapt to any length of sequential data files, and the automatic identification accuracy of early faults in tilting pad bearings reaches 98.7%.

**Keywords:** fault diagnosis; risk evaluation; improved auxiliary classifier Wasserstein generative adversarial network with gradient penalty; industrial Internet; rotor system



**Citation:** Zhou, C.; Wang, Q.; Xiao, Y.; Xiao, W.; Shu, Y. Research on an Improved Auxiliary Classifier Wasserstein Generative Adversarial Network with Gradient Penalty Fault Diagnosis Method for Tilting Pad Bearing of Rotating Equipment. *Lubricants* **2023**, *11*, 423. <https://doi.org/10.3390/lubricants11100423>

Received: 28 August 2023

Revised: 18 September 2023

Accepted: 24 September 2023

Published: 2 October 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Centrifugal compressors, steam turbines, flue gas turbines, expanders and other high-speed rotating machinery are widely used in petrochemical, coal chemical, metallurgical and other industrial fields. In the event of blade fracture, rotor imbalance, rubbing, surges and other faults [1], minor faults may cause equipment failure and production interruption, and serious faults may cause machine damage and fatal accidents, leading to huge economic losses or social impact to enterprises [2,3]. Timely and automatic identification of equipment failure types, to take control and to take preventive measures, is of great significance for reducing or avoiding economic losses in enterprises and preventing catastrophic failures of rotating machinery [4].

Oil-whirl faults caused by improper assembly-clearance and contact areas between tilting pad bearing and shaft is one of the most common faults in rotating equipment. Oil whirl failure refers to severe fluctuations or vibrations in the lubricating oil film, which usually occurs when the lubricating oil film cannot be stably maintained on the surface

of mechanical parts. This failure may cause serious harm to mechanical equipment and systems. The specific hazards include: (1) Friction and wear increase: oil whirl will lead to the instability of the lubricating oil film, so that the contact area between the mechanical parts increases; friction and wear will increase accordingly. Long-term friction and wear can lead to damage and the shortened life of parts. (2) Energy loss: oil whirl will cause abnormal contact between mechanical parts, which will lead to energy loss, thus affecting the efficiency and performance of the mechanical system. (3) Vibration and noise: Oil whirl can cause the vibration of mechanical parts, and then produce noise. These vibrations and noises not only affect the normal operation of mechanical equipment, but also may affect the surrounding environment and the health of workers. (4) Heat accumulation: oil whirl may lead to local energy concentration, resulting in excessive heat accumulation. This may lead to overheating of lubricating oil and further aggravate the damage to mechanical parts. (5) System failure: If the problem caused by oil whirl is not solved in time, it may lead to the failure of the components of the mechanical system, which in turn affects the normal operation of the entire equipment. This may require expensive maintenance and downtime. In addition, due to the shaft misalignment, rotor imbalance, surges, rubbing and other faults of rotating equipment may also occur at the same time; it is challenging to accurately identify the early faults of tilting pad bearings.

The research on fault diagnosis methods based on artificial intelligence has achieved fruitful results. Zhong et al. [5] proposed a rolling bearing fault diagnosis method based on a convolutional autoencoder and nearest-neighbor algorithm, which was verified experimentally by using the experimental data set published by CWRU under different working conditions. Mohiuddin et al. [6] proposed an improved AlexNet-based intelligent fault diagnosis method for rolling bearings, which was verified experimentally using the data of different working conditions and a different signal-to-noise ratio of the experimental data set published by CWRU. Cui et al. [7] proposed a method for fault diagnosis of rolling bearings under the condition of sample imbalance based on CNN, and used the conventional rolling bearing-fault data set collected in the laboratory for experimental verification. Zhang et al. [8] proposed a CNN-based multi-channel data fusion neural network for rolling bearing fault diagnosis, using bearing data collected by eight vibration sensors on the SB25 aero-engine bearing bench test for experimental verification of the model. Shen et al. [9] proposed an improved Gray Wolf optimizer algorithm based on a support vector machine and swarm intelligence optimization algorithm for rolling bearing fault diagnosis. The proposed algorithm was verified experimentally using the experimental data set published by CWRU and the data obtained from the mechanical transmission bearing life-cycle test platform independently developed by Nanjing Agricultural University. Huang et al. [10] proposed a rolling bearing fault-detection method based on an improved Gray Wolf algorithm to optimize multi-stable stochastic resonance parameters, and conducted experimental verification using the published experimental data sets of CWRU and MFPT. Tian et al. [11] proposed a CNN-LSTM bearing fault diagnosis model based on hybrid particle swarm optimization, and conducted experimental verification using the experimental data set disclosed by CWRU. However, most of the research objects of these research results are rolling bearings, and the model test data are almost all derived from laboratory bench test data. Moreover, there are few research results on the fault diagnosis of rotor systems composed of shafts, impellers or blades, couplings, and tilting pad bearings. There are still shortcomings in the operational risk evaluation of rotor systems.

The traditional fault diagnosis method for rotating machinery relies on the experience and knowledge of external experts, and relies on a spectrum analysis diagram, Bode diagram, Nyquist diagram and other analysis toolboxes in the condition-monitoring and analysis software to carry out a one-by-one manual analysis. This not only has a low efficiency of fault diagnosis and analysis, but also has a great lag, which often leads to untimely early fault-detection. Industrial Internet-enabled equipment management technology has developed rapidly in China. The accumulated equipment state-aware data has laid a foundation for intelligent fault diagnosis based on artificial intelligence and big data

analysis. This data-driven, deep learning intelligent fault diagnosis method [12–14] makes full use of the advantages of industrial big data and greatly reduces the dependence of the model on external experts. It has gradually become a development trend for equipment fault diagnosis technology in the industrial Internet environment [15,16].

During the service life-cycle of rotating equipment, the time of fault-free operation for equipment is far greater than that of fault operation, which determines that the data samples of normal-state perception for equipment are significantly more abundant than those of fault-state perception. The data samples are of a long-tail distribution type and have the characteristics of low value density [17]. For specific rotating equipment, it is impossible to go through all the faults such as rotor imbalance, axis misalignment, rubbing, oil film whirl, surges and so on in the service life-cycle. Some equipment will not even have any kind of fault in the whole life-cycle. The lack of equipment-fault sample data is one of the challenging problems faced by fault diagnosis technology based on artificial intelligence and big data [18].

When model-training samples are insufficient, a generative adversarial network is considered as one of the effective methods for solving the problem of data imbalance [19]. Generative Adversarial Networks (GAN) [20] are a deep learning model that is one of the most promising approaches to unsupervised learning over complex distributions in recent years. The model produces a fairly good output through game learning between (at least) two modules in the framework: the generative model and the discriminative model. GAN models generally use deep neural networks as G and D. A good GAN application needs to have a good training method; otherwise the output may not be ideal due to the freedom of the neural network model. To improve a GAN's data generation capabilities and optimize the training process, a Deep Convolution Generative Adversarial Network (DCGAN) based on a deep Convolutional Neural Network (CNN) and generated high-resolution images have been proposed [21]. However, as the training time of the model increases, some filters of the model will collapse and oscillate, resulting in mode collapse.

In order to solve the problem of GAN pattern collapse, a Wasserstein generative adversarial network (WGAN) model was constructed to overcome the problem by improving the stability of training [22]. Aiming at the problem that WGAN adopts weight-clipping to solve the problem of Lipschitz constraints that can easily cause gradient disappearance or gradient explosion and slow model convergence, an improved Wasserstein GAN training method (WGAN-GP) has been proposed by Gulrajani et al. [23]. By using a gradient penalty instead of weight-clipping to solve the problem of Lipschitz constraints, gradient disappearance or gradient explosion during model training can be avoided, and the problem of slow convergence of WGAN also can be solved. GAN, DCGAN, WGAN, and WGAN-GP are all unsupervised learning models that generate samples without category labels and cannot generate multiple types of samples using the same model.

In order to enhance the performance of GAN, an Auxiliary Classifier GAN (ACGAN) supervised learning model, which adds category labels to the generator and discriminator, as well as a classifier to the output part of the discriminator, have been proposed [24]. This ACGAN model realizes that the generated samples all have a corresponding category label. The ACGAN model is improved based on DCGAN, so ACGAN still has the problem of model collapse. A Parallel Classification Wasserstein Generative Adversarial Network with Gradient Penalty (PCWGAN-GP) has been proposed by Yu et al. [25]. By feeding healthy samples into the PCWGAN-GP model, the model will produce various failure samples of good quality, which can gradually expand the unbalanced data set until equilibrium is reached.

PCWGAN-GP is an unsupervised learning model, which needs to be constructed and trained independently for each fault type to obtain a balanced data set. This undoubtedly increases the workload of model construction and increases the time of model training. An ACWGAN-GP model based on a gradient penalty and auxiliary classifier has been built by Li et al. [26], which can generate good-quality samples from an unbalanced training set, and has used the balanced data set for training Multilayer Perceptron (MLP), CNN, Support

Vector Machine (SVM) and other classifiers for fault diagnosis. Cao et al. [27] constructed a fault diagnosis model based on ACWGAN-GP and homogeneous superposition ensemble learning, which significantly improved the classification accuracy and stability of the model.

ACWGAN-GP combines the advantages of WGAN-GP and ACGAN, so that the model has the ability to generate multi-class label samples while overcoming the problems of pattern collapse and gradient disappearance. As a supervised learning model, ACWGAN-GP still needs a complete variety of fault label sample training data sets. Obviously, engineering application scenarios are not always able to meet such needs. Furthermore, the length of a single device state-aware data file is different, and the ACWGAN-GP model can only adapt to a single data file type, which cannot meet the needs of engineering applications. Therefore, the application of the ACWGAN-GP model for equipment full-fault diagnosis needs improvements in the model structure, so that it cannot only meet the data function of generating complete fault samples, but also automatically adapt to different equipment-state perception data.

Rotating equipment is generally composed of shafts, impellers or blades, comb seals, couplings, tilting pad bearings and other components. Among them, oil whirl faults caused by improper assembly-clearance and contact areas between tilting pad bearings and shafts are the most common. It is challenging to accurately identify the early faults of tilting pad bearings, because shaft misalignment, rotor imbalance, surges, rubbing and other faults of rotating equipment may also occur at the same time. Aiming at the engineering status of unbalanced data samples of rotating equipment, this paper studies an improved auxiliary classifier Wasserstein generative adversarial network with a gradient penalty for fault diagnosis of tilting pad bearings. The contributions of this paper are listed as follows:

(1) An improved auxiliary classifier Wasserstein generative adversarial network with gradient penalty is developed, in which the input data-length adaptive layer is added before the 2D convolution layer of the discriminator.

(2) A fault diagnosis method based on IACWGAN-GP for tilting pad bearings is proposed, which is able to accurately identify the early faults of tilting pad bearing oil whirl despite the interference of shaft misalignment, rotor imbalance, surges, rubbing and other faults that may occur simultaneously in rotating equipment.

(3) The application of an IACWGAN-GP-based fault diagnosis model in an industrial Internet environment via cloud-integrated prediction and health management systems, which includes a cyber-physical system layer, network layer and application layer, is proposed. The application layer consists of micro-service systems such as early fault warning, health evaluation and fault diagnosis.

## 2. Theoretical Foundations

### 2.1. Generative Adversarial Network

As shown in Figure 1, a GAN is composed of two neural networks: the generator  $G$  and the discriminator  $D$ . In order to trick the discriminator  $D$ , the generator must learn the data distribution of real samples and create fake samples using random noise. The discriminator's task is to tell the real samples from the fake samples that have been generated. The performance of  $G$  and  $D$  is continuously improved until Nash equilibrium is reached in these two adversarial trained neural networks. The GAN's loss function can be described as follows:

$$Loss = E_{X \sim P_{data}} [\log D(x)] + E_{Z \sim P_Z} [\log(1 - D(G(z)))] \quad (1)$$

where  $E_{S \sim P_{data}}$  and  $E_{Z \sim P_Z}$  stand for the expectation of  $x$  from the real data distribution  $P_{data}$  and  $z$  sampled from the random noise prior distribution  $P_Z$ , respectively.  $D(x)$  represents the discriminant result when the input of the discriminator is real data  $x$ ,  $G(z)$  represents the generated data of the generator, and  $D(G(z))$  represents the discriminant result when the input of the discriminator is the generated data  $G(z)$ .



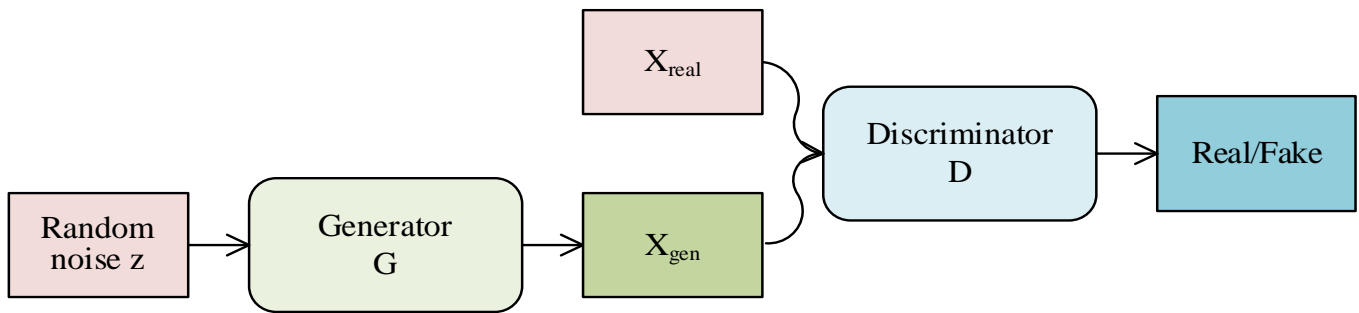


Figure 1. Architecture of GAN.

The binary minimax problem that describes the optimization procedure for  $D$  and  $G$  is represented by the following equation:

$$Goal = \underset{G}{\operatorname{argmin}} \underset{D}{\operatorname{max}} Loss \quad (2)$$

## 2.2. Auxiliary Classifier Generative Adversarial Network

ACGAN is a modified model of GAN with the structure shown in Figure 2. Unlike GAN, ACGAN can use label information to generate samples of specified types and to identify and classify the input samples. Specifically, the generator  $G$  generates new samples  $G(z, y)$  using random noise  $z$  and label  $y$ , while the discriminator  $D$  needs not only to determine the real or fake nature of the input samples, but also to classify the input samples. During the adversarial training, the generating sample capability and recognition capability of ACGAN are continuously optimized. Eventually, the model has a strong ability to generate new samples with corresponding labels.

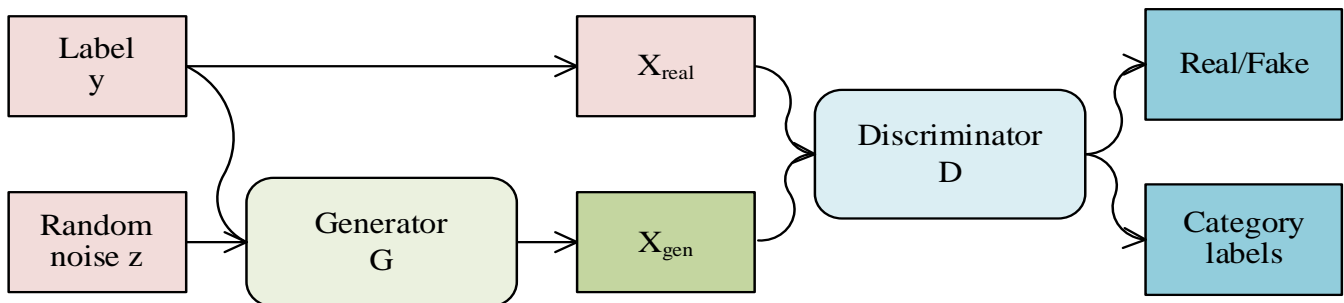


Figure 2. Architecture of ACGAN.

Since ACGAN needs to process the source and class label information of the input samples, the loss function of ACGAN contains two components, defined as follows:

$$L_{source} = E_{X \sim P_{data}} [\log D(x)] + E_{Z \sim P_z} [\log(1 - D(G(z, y)))] \quad (3)$$

$$L_{class} = E_{X \sim P_{data}} [\log P(Y = y | X_{real})] + E_{Z \sim P_z} [\log P(Y = y | X_{generated})] \quad (4)$$

where  $G(z, y)$  represents the generated sample when the generator inputs are random noise  $z$  and sample label  $y$ ,  $P(Y = y | X_{real})$  represents the conditional probability distribution of the real sample, and  $P(Y = y | X_{generated})$  represents the conditional probability distribution of the generated sample.

The objective function of the D is to maximize  $L_{source} + L_{class}$ , and the objective function of the G is to maximize  $L_{source} - L_{class}$ , shown as follows:

$$L_D = E_{X \sim P_{data}} [\log D(x)] + E_{Z \sim P_Z} [\log(1 - D(G(z, y)))] + E_{X \sim P_{data}} [\log P(Y = y | X_{real})] + E_{Z \sim P_Z} [\log P(Y = y | X_{generated})] \quad (5)$$

$$L_G = E_{X \sim P_{data}} [\log D(x)] + E_{Z \sim P_Z} [\log(1 - D(G(z, y)))] - E_{X \sim P_{data}} [\log P(Y = y | X_{real})] - E_{Z \sim P_Z} [\log P(Y = y | X_{generated})] \quad (6)$$

### 2.3. Wasserstein Distance and Gradient Penalty

GANs have attracted the attention of many researchers because of their powerful sample generation capability. But there are gradient-disappearance and model-collapse problems leading to unstable training. To solve these problems, a lot of attempts have been made by many researchers. However, the problem of GAN training instability was not solved until the proposal of WGAN.

Arjovsky et al. [22] proposed WGAN and credited the objective function, which took the form of J-S divergence, as the cause of the unstable training of GAN. They then suggested utilizing the Wasserstein distance (WD) rather than the J-S divergence in the WGAN. The following is an expression of the WGAN's objective function:

$$\min_G \max_{D \in \Omega} E_{X \sim P_{data}} [D(x)] - E_{Z \sim P_Z} [D(G(z))] \quad (7)$$

where  $\Omega$  denotes the set of 1-Lipschitz functions that take values in the range  $[-\omega, \omega]$ .

Although the WGAN training process is faster and more stable than the original GAN, the quality of the generated samples is occasionally unsatisfactory. The issue, according to Gulrajani et al. [23], was caused by the weights in WGAN being restricted in order to enforce the Lipschitz constraint on the discriminator. Therefore, they introduced a gradient penalty to propose WGAN-GP. The following defines the WGAN-GP's loss function and objective function:

$$L = E_{Z \sim P_Z} [D(G(z))] - E_{X \sim P_{data}} [D(x)] + \varphi E_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (8)$$

$$\min_G \max_D E_{X \sim P_{data}} [D(x) - E_{Z \sim P_Z} [D(G(z))]] - \varphi E_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (9)$$

where  $\hat{x} = \varepsilon x + (1 - \varepsilon)G(z) \sim P_{\hat{x}}$  and the random numbers  $\varepsilon \sim U(0, 1)$ ,  $\varphi$  represents the penalty factor. Without meticulously adjusting the hyperparameters, WGAN-GP performs better than WGAN and achieves steady training.

## 3. The Proposed Method

### 3.1. Building an Improved Auxiliary Classifier Wasserstein Generative Adversarial Network with Gradient Penalty

In order to overcome the limitation of the input data length on the neural network model, the IACWGAN-GP model has been designed. In order to avoid the problem of model collapse and gradient disappearance during the training process, Wasserstein distance and a gradient penalty are introduced into the loss function of the model. The model introduces category labels in the generator and discriminator, and introduces an auxiliary classifier in the discriminator, so that the model has the ability to generate multi-class label samples and a sample classification ability. The generator uses three 2D deconvolution layer results, and the discriminator uses three 2D convolution layer structures. Before the first convolutional layer of the discriminator, an input data length adaptive layer designed in this paper is added, so that the model can automatically adapt to different device-status sensing data, and improve the applicability and generalization of the model to various types of data. The architecture of the IACWGAN-GP model is shown in Figure 3.

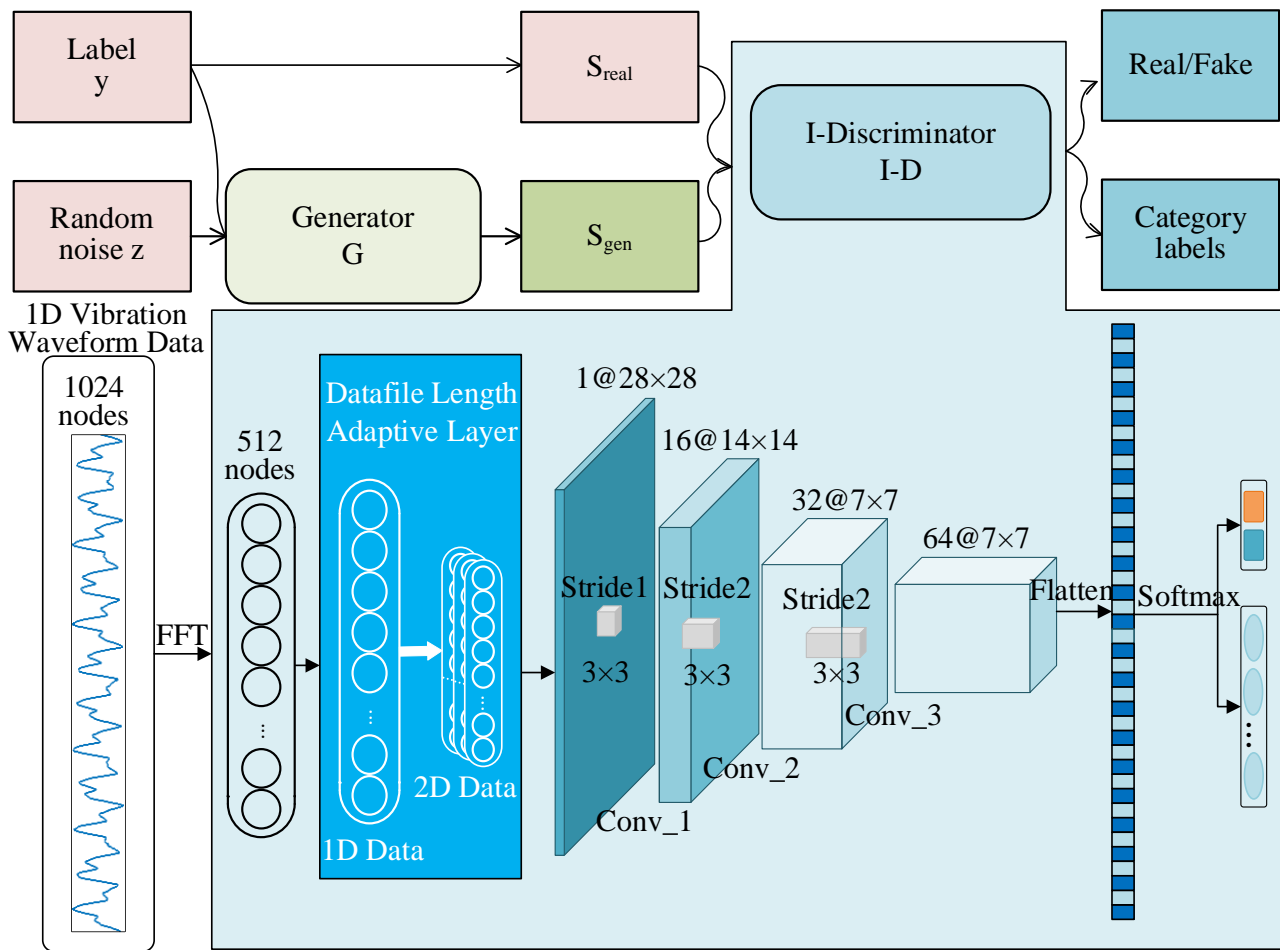


Figure 3. Architecture of IACWGAN-GP.

The 2D CNN has superior performance in feature extraction and classification compared to the 1D CNN [28], so both the generator and discriminator of IACWGAN-GP constructed in this paper use 2D convolutional structures. The vibration signal of rotating machinery is a 1D signal, which cannot be directly convolved in 2D. Therefore, it is necessary to convert 1D data into 2D data, which requires the data length to be a square number. However, the length of engineering case data often does not meet this requirement. For example, the 1D vibration engineering case data with data length of 1024 is halved to 512 after Fourier transform. Since the input data dimension of 2D convolution requires 2D, 1D data needs to be converted into 2D data, and the data length is usually taken as a square number, such as 784(28 × 28), 1024(32 × 32), etc. However, the length of the spectrum data after Fourier transform (512) is not a square number. To overcome the limitation of the neural network model on the input data length, an input adaptive learning framework is designed. Specifically, the Input Adaptive Layer (IAL) is designed before the first convolutional layer of the discriminator, as shown in Figure 3.

A 1D signal of length  $m$  is defined as:

$$S_{input} = [a_1 \ a_2 \ \cdots \ a_m] \quad (10)$$

where  $S_{input}$  indicates the 1D input data, and  $a_i$  indicates the value of node  $i$  of the 1D waveform data.

In order to convert 1D data with length of non-square number  $m$  into 2D data with target size, multiply  $S_{input}$  with size  $(1, m)$  by a weight matrix with size  $(m, n)$ , the input

data can be converted into 1D data with length of square number  $n$ , and the result can be corrected with deviation, as shown in Equation (11).

$$S = F_A(S_{input} \cdot K + b) \quad (11)$$

where  $K$  is the weight matrix created by the kernel layer, as shown in Equation (12);  $b$  is the deviation vector, created by the layer, as shown in Equation (13);  $F_A(\cdot)$  is the activation function of the neural network.

$$K = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1n} \\ k_{21} & k_{22} & \cdots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{m1} & k_{m2} & \cdots & k_{mn} \end{bmatrix} \quad (12)$$

$$b = [b_1 \quad b_2 \quad \cdots \quad b_n] \quad (13)$$

By substituting Equations (12) and (13) into Equation (11), the value of the transformed  $i$ -th node is as follows:

$$S_i = F_A\left(\sum_{j=1}^m a_j \times k_{ji} + b_i\right) \quad (14)$$

Since  $n$  is a square number, it is easy to transform 1D data of length  $(1, n)$  into 2D data of shape  $(\sqrt{n}, \sqrt{n})$ :

$$S_{output} = F_R(S) \quad (15)$$

where  $F_R(\cdot)$  is the Reshape function.

Equations (10)–(15) are the derivation process of IAL. IAL can be defined as follows:

$$S_{2D} = IA(S_{1D}) \quad (16)$$

where  $IA(\cdot)$  is the input adaptive function whose input is 1D data of arbitrary length and output is 2D data of desired size.

With the introduction of IAL, 1D data whose data length is not a square number can be easily converted into 2D data required for 2D convolution models, enabling input data-length adaption and improving the applicability and generalization of neural network models to various types of data.

Bringing Equation (8) into Equation (5) and Equation (6), the objective function of IACWGAN-GP is expressed as follows:

$$L_D = E_{x \sim P_X}[\log D(x)] - E_{\hat{x} \sim P_G}[\log D(\hat{x})] - \varphi E_{\tilde{x} \sim P_{\tilde{x}}}[(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2] + E_{x \sim P_X}[\log P(Y = y|x)] \quad (17)$$

$$L_G = E_{\hat{x} \sim P_G}[\log D(\hat{x})] + E_{\hat{x} \sim P_{\tilde{x}}}[\log P(Y = y|\hat{x})] \quad (18)$$

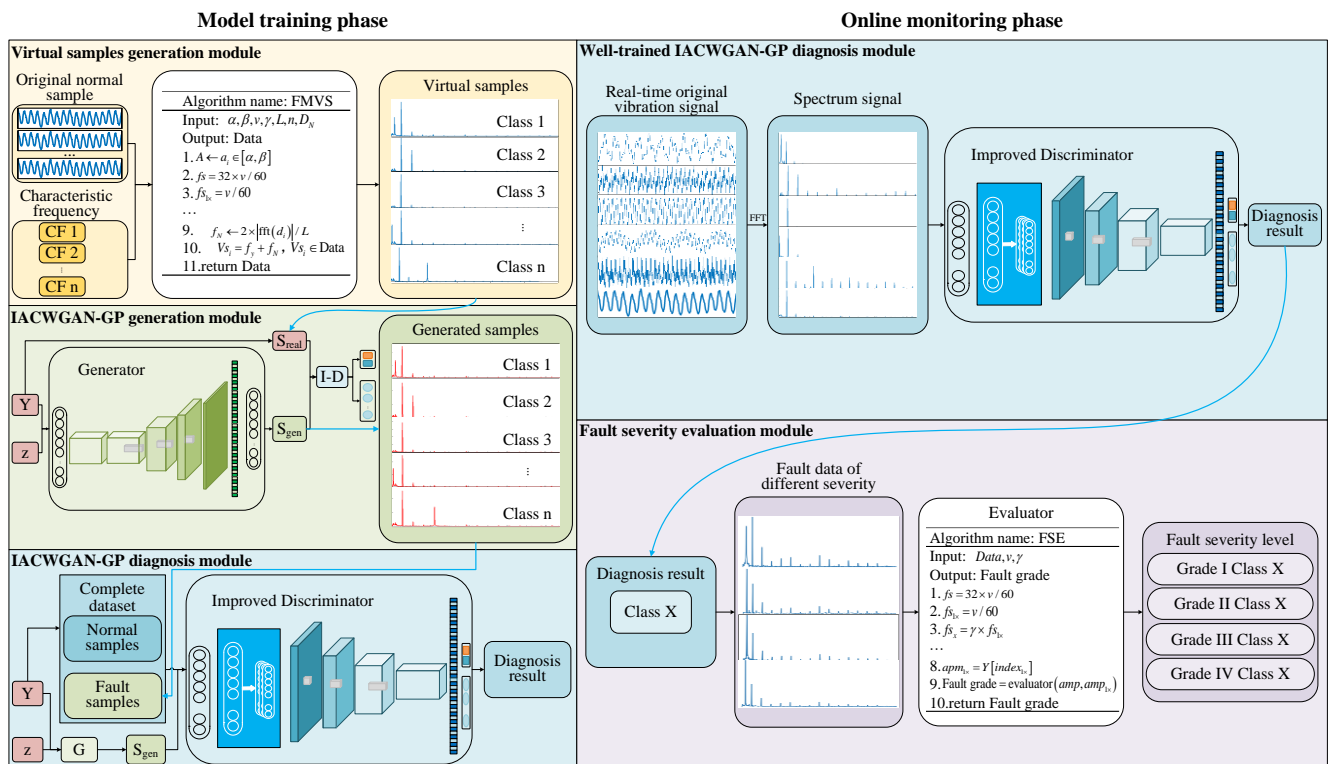
where  $\tilde{x} = \varepsilon x + (1 - \varepsilon)\hat{x} \sim P_{\tilde{x}}$  is the random sample got by interpolating between  $x$  and  $\hat{x}$ ,  $\varphi$  is the gradient penalty factor.

Equations (17) and (18) are the objective functions of IACWGAN-GP. Combining the advantages of WGAN-GP and ACGAN, the model can generate multi-class label samples, and overcome the problems of pattern collapse and gradient disappearance.

### 3.2. Establishing Fault Diagnosis Model Based on IACWGAN-GP

The architecture of the IACWGAN-GP-based fault diagnosis model is shown in Figure 4. The model is divided into a model training phase and an online monitoring phase. The model training phase includes a virtual sample generation module, an IACWGAN-GP generation module, and an IACWGAN-GP diagnosis module. Using the virtual sample generation module to generate virtual samples of full fault types, the IACWGAN-GP

generation module is trained. The well-trained IACWGAN-GP generation module is used to generate the required number and required type of fault samples, which, together with the normal samples of the equipment, form a complete data set to train the IACWGAN-GP diagnosis module. The well-trained diagnosis module can be used for fault diagnosis in the online monitoring phase. The online monitoring phase includes a fault diagnosis module and a fault-severity evaluation module. The fault diagnosis module uses the IACWGAN-GP diagnosis module with good training in the model training phase as the classifier. The real-time original vibration signal is transformed into a frequency domain signal by FFT and then input into the classifier to obtain the real-time fault diagnosis result. After the fault diagnosis results are obtained, the fault-severity evaluation module enables the fault-severity evaluator of the corresponding fault type, and the frequency domain signal is input into the evaluator to obtain the real-time fault-severity evaluation results.



**Figure 4.** Architecture of fault diagnosis model based on IACWGAN-GP.

### 3.2.1. Virtual Sample Generation Module

#### Fault Virtual Sample Definition

As the characteristics of frequency domain signals are more prominent than those of time domain signals, many researchers in the field of fault diagnosis in rotating machinery employ frequency domain signals as the input to their models [29,30]. The frequency domain signal can be obtained by Fourier transforming the vibration signal in the time domain of the equipment, and studying the spectrum of the equipment fault signal shows that the spectrum contains the fault-characteristic frequencies that coincide with the fault mechanism, and these fault-characteristic frequencies belong to the common characteristics of the fault. Besides fault-characteristic frequencies, other frequency information in the spectrum responds to some private characteristics of the equipment, such as working conditions, environmental noise, etc. Studying the spectrum of the normal-state signal of the equipment reveals that the information in the spectrum matches with the private features in the fault signal. In other words, the normal-state signal of the equipment can reflect the private characteristics of the equipment. As shown in Figure 5, the overlapping parts of the fault data belong to the common features, and those other than the overlapping



parts belong to the private features. Take surge and oil whirl as an example for illustration, as shown in Figure 6.

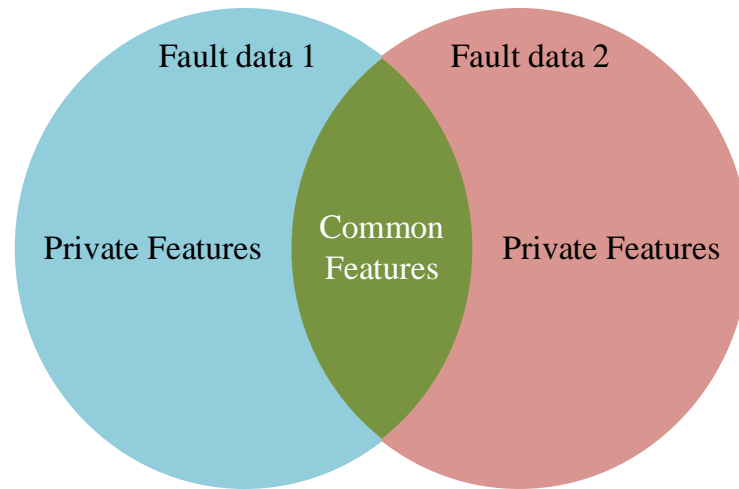


Figure 5. Fault data Venn diagram.

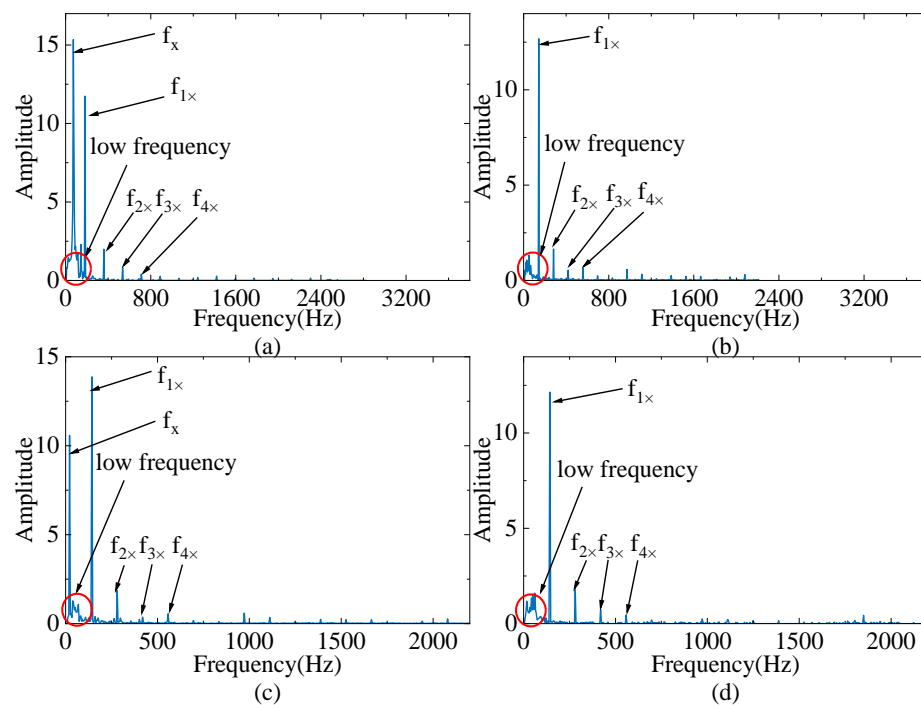


Figure 6. Spectrum of equipment fault and normal data. (a) The frequency spectrum of the oil whirl. (b) The frequency spectrum of the normal-state data of the equipment. (c) The frequency spectrum of the surge fault data. (d) The frequency spectrum of the normal-state data of the equipment.

Figure 6a shows the frequency spectrum of the oil whirl, the speed of the device is 10,600 rpm, in which  $f_x = 66.25$  Hz is the fault characteristic frequency of the oil-whirl fault. Figure 6b shows the frequency spectrum of the normal-state data of the equipment. Comparing with Figure 6a, it can be found that except the fault feature frequency, the private features such as the low-frequency component, the 1st-order frequency  $f_{1x}$ , the 2nd-order frequency  $f_{2x}$ , the 3rd-order frequency  $f_{3x}$ , and the 4th-order frequency  $f_{4x}$  are all included in the spectrum. Figure 6c shows the frequency spectrum of the surge fault data, the speed of the device is 8300 rpm, in which  $f_x = 17.29$  Hz is the fault characteristic frequency of the surge fault. Figure 6d shows the frequency spectrum of the normal-state

data of the equipment. Comparing with Figure 6c, it can be found that excepts the fault feature frequency, the private features such as the low-frequency component, the 1st-order frequency  $f_{1\times}$ , the 2nd-order frequency  $f_{2\times}$ , the 3rd-order frequency  $f_{3\times}$ , and the 4th-order frequency  $f_{4\times}$  are all included in the spectrum.

The Pearson Correlation Coefficient (PCC) [31] is extensively used to measure the degree of correlation between two variables, and the value of PCC is between  $-1$  and  $1$ ; the closer to  $\pm 1$  the higher the correlation. The PCC of the fault and normal data, the PCC of the fault data with the fault characteristic frequency removed, and the normal data are calculated as shown in Table 1. The PCC value of the oil whirl and normal data of the device is 0.81824, and the PCC value of the oil whirl with the fault characteristic frequency removed and normal data is 0.90919. The PCC value of the surge and normal data of the device is 0.78202, and the PCC value of the surge with the fault characteristic frequency removed and normal data is 0.98130.

**Table 1.** PCC of fault data and normal data.

|           | The PCC of the Fault Data and the Normal Data | The PCC of the Fault Data with the Fault Characteristic Frequency Removed and the Normal Data |
|-----------|---|---|
| Oil whirl | 0.81824                                       | 0.90919   |
| Surge     | 0.78202                                       | 0.98130   |

Through the comparative study of spectrograms of fault and normal data, as well as the comparative study of a PCC index calculation for fault and normal data, the above results show that fault data can be described as a collection of common features and private features. That means the fault feature frequency characterizing the common features of faults, and the normal data characterizing the private features of the equipment, can be composed of fault data. Based on this result, it provides an idea for constructing a fault virtual sample. Based on the normal data spectrum of the equipment, the fault characteristic frequencies are superimposed to form the fault virtual samples.

#### Fault Mechanism-Based Virtual Sample Generation

Based on the results of Fault Virtual Sample Definition, when the target device can only provide the normal-state data of the equipment, or when the fault type space is incomplete, a fault virtual sample can be generated based on normal data and fault mechanism features frequency. In this paper, we propose a FMVS generation algorithm with the following pseudocode:

As shown in Algorithm 1,  $\alpha$  and  $\beta$  represent the upper and lower bounds of the amplitude range of virtual sample fault mechanism features frequency; the amplitude of the fault mechanism features frequency takes a value within  $[\alpha, \beta]$ , which can characterize the severity of the fault;  $v$  represents the speed of the device;  $\gamma$  represents the fault characteristic frequency parameters, which are determined based on the fault mechanism;  $L$  represents the length of the normal data of the device;  $n$  represents the number of data sets of normal data and virtual fault samples;  $D_N$  represents the normal data of the device.

First, calculate the fundamental frequency  $f_{s_{1\times}}$  of the target device, and the amplitude of the fundamental frequency can be obtained by Fourier transform on the normal data of the target device. Then, calculate the fault characteristic frequency  $f_c$ , and the amplitude of the fault characteristic frequency is determined from the amplitude of the fundamental frequency as well as  $\alpha$  and  $\beta$ . Next, a sine function is used to fit the virtual signal containing only the fault characteristic frequencies, and the virtual signal is Fourier transformed to obtain the virtual signal spectrum  $f_y$  with frequency  $f_c$  and amplitude within  $[\alpha, \beta]$ . Finally, the virtual signal spectrum  $f_y$  plus the normal signal spectrum  $f_N$  of the target device can obtain the fault virtual sample  $Vs_i$  of the target device. Repeating the above process, the virtual sample *Data* with full fault types can be generated.

**Algorithm 1** Pseudocode for FMVS algorithm

---

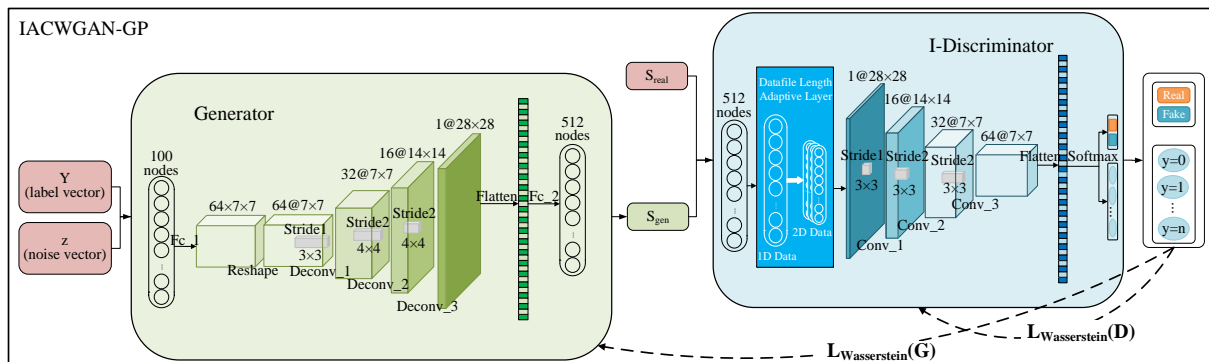
Input:  $\alpha, \beta, \nu, \gamma, L, n, D_N$   
Output: *Data*

1.  $A \leftarrow a_i \in [\alpha, \beta]$
2.  $f_s = 32 \times \nu / 60$
3.  $f_{s_{1 \times}} = \nu / 60$
4.  $f_c = \text{round}(\gamma \times f_{s_{1 \times}} \times L / f_s)$
5. for  $i \leftarrow 0$  to  $n$
6.    $y = a_i \times \sin(2\pi \times f_c \times x)$
7.    $f_y \leftarrow 2 \times |\text{fft}(y)| / L$
8.    $d_i \in D_N$
9.    $f_N \leftarrow 2 \times |\text{fft}(d_i)| / L$
10.  $V_{s_i} = f_y + f_N, V_{s_i} \in \text{Data}$
11. return *Data*

---

**3.2.2. IACWGAN-GP Generation Module and Diagnosis Module**

The IACWGAN-GP generation module and the diagnosis module adopt the same neural network structure, as shown in Figure 7. The IACWGAN-GP generation module mainly uses the sample generation ability of the generator, and the IACWGAN-GP diagnosis module mainly uses the sample classification ability of the discriminator. It makes full use of the performance of IACWGAN-GP and reduces the difficulty of constructing the model.



**Figure 7.** Structure of IACWGAN-GP neural network.

The number of fault virtual samples constructed in Section 3.2.1 is limited, and to avoid the pattern collapse problem caused by training classification networks with limited data sets [32], an IACWGAN-GP generation module is constructed for generating fault mechanistic feature spectrum samples. Using the virtual samples generated in the Section 3.2.1 training generation module, with supervised learning and the powerful sample generation capability of IACWGAN-GP, the prior knowledge in Section 3.2.1 is no longer required, and only the fault sample labels and the number of required samples need to be input for generating the corresponding fault mechanism feature spectrum samples with corresponding fault types and numbers. Compared with the FMVS generation algorithm in Section 3.2.1, the number of inputs to the model is reduced from seven to two. The IACWGAN-GP diagnosis module is trained with a complete fault data set consisting of fault mechanism feature spectrum samples generated using generation module and equipment normal-state samples. The well-trained model can be used for real-time online full-fault-type fault diagnosis and identification.

**3.2.3. Fault Severity Evaluation Module**

In an actual engineering scenario, the faults of rotating equipment often experience a gradual development process from weak to strong. Therefore, in addition to determining

the type of failure of the equipment, it is also necessary to evaluate the severity of the failure of the target equipment. After the spectrum analysis of the vibration signal, the fault-type of the equipment can be known from the fault characteristic frequency, and the severity of the fault can be judged by analyzing the amplitude ratio of the fault-characteristic frequency to the fundamental frequency. Based on this method, this paper has constructed a Fault Severity Evaluation (FSE) module.

The pseudo-code of FSE algorithm is shown in Algorithm 2.  $D$  represents the data to be measured;  $v$  represents the speed of the device;  $\gamma$  represents the fault characteristic frequency parameters, which are determined based on the fault mechanism;  $G$  represents the result of fault severity evaluation. Firstly, the device sample frequency  $fs$  and the fundamental frequency  $fs_{1\times}$  are calculated from the speed  $v$ . Secondly, the fault characteristic frequency is calculated by  $\gamma$  and  $fs_{1\times}$ , and the Fourier transform is performed on the data to be measured  $D$ , and the vibration signal in the time domain is transformed into the frequency domain signal. Then, the amplitude of the fault characteristic frequency  $amp$  and the amplitude of the fundamental frequency  $amp_{1\times}$  are found in the spectrum signal, and the fault severity of the fault signal to be measured is evaluated based on the amplitude ratio of the fault characteristic frequency to the fundamental frequency. Finally, the output fault grade  $G$  is used as the evaluation result of the fault severity.

---

**Algorithm 2** Pseudocode for FSE algorithm

---

Input:  $D, v, \gamma$

Output:  $G$

1.  $fs = 32 \times v / 60$
  2.  $fs_{1\times} = v / 60$
  3.  $fs_x = \gamma \times fs_{1\times}$
  4.  $f_x, Y = \text{FFT}(D, fs)$
  5.  $index = \text{find}(\text{round}(f_x) == \text{round}(fs_x))$
  6.  $index_{1\times} = \text{find}(\text{round}(f_x) == \text{round}(fs_{1\times}))$
  7.  $amp = Y[index]$
  8.  $amp_{1\times} = Y[index_{1\times}]$
  9.  $Fault\ grade = \text{evaluator}(amp, amp_{1\times})$
  10. return  $G$
- 

### 3.3. Fault Diagnosis Method Based on IACWGAN-GP

In this paper, a fault diagnosis method based on IACWGAN-GP for tilting pad bearings is proposed, which is able to accurately identify the early faults of tilting pad bearing oil whirl despite the interference of shaft misalignment, rotor imbalance, surges, rubbing and other faults that may occur simultaneously in rotating equipment. The specific steps are as follows:

Step 1: The virtual sample generation module uses the normal data of the equipment and the characteristic frequency of the fault mechanism to generate the virtual sample of the full fault types.

Step 2: Use the generated full fault type virtual samples to train the IACWGAN-GP generation module.

Step 3: The IACWGAN-GP generation module is used to generate the fault mechanism-characteristic spectrum samples of different fault types, and the fault-type complete data set is formed with the normal-state data of the equipment.

Step 4: Use the complete data set to train the IACWGAN-GP diagnostic module.

Step 5: The IACWGAN-GP diagnosis module is used to realize the intelligent diagnosis and identification of full fault types.

Step 6: The fault-severity evaluation module is used to evaluate the severity of the fault.

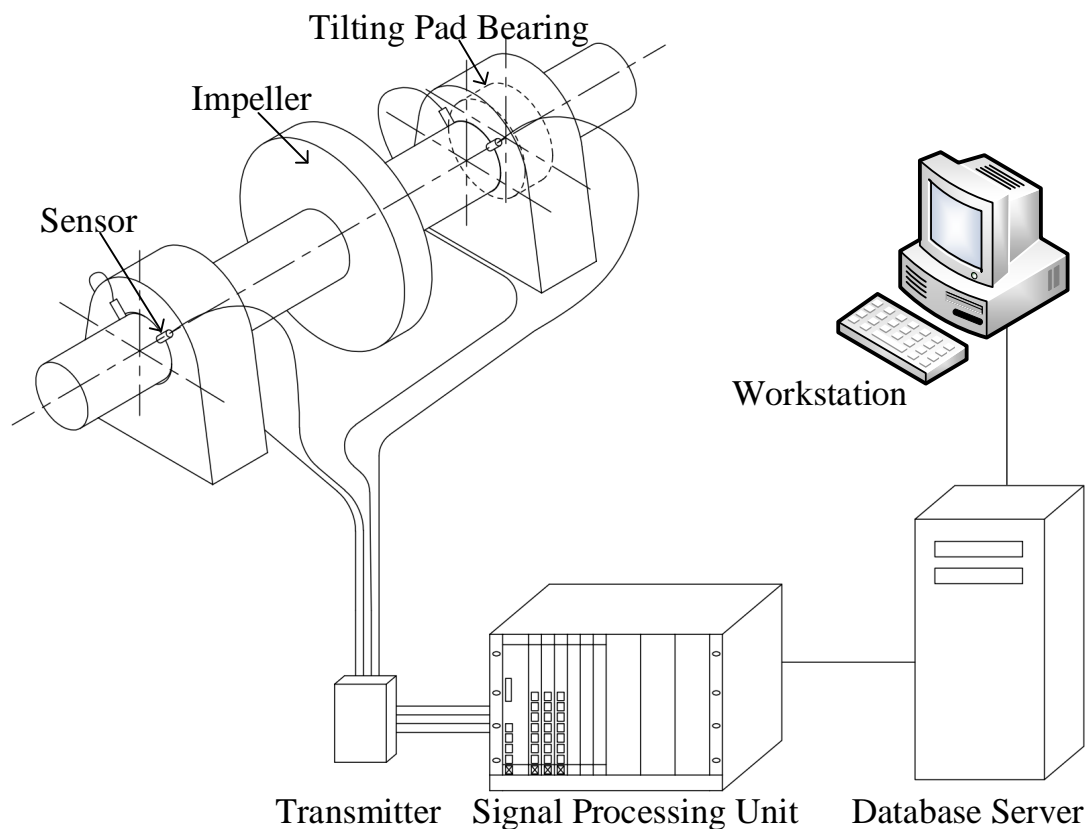
Among these, steps one to four are the model training phase, and steps five to six are the online monitoring phase.

#### 4. Experiments and Analysis of Results

In this paper, model validation and comparison experiments are conducted using the rotating equipment fault case-data of petrochemical enterprises to verify the effectiveness of the rotating machinery fault diagnosis method based on IACWGAN-GP. The proposed IACWGAN-GP-based fault diagnosis model is used to generate fault mechanism feature frequency spectra samples of different fault types and perform fault diagnosis. Subsequently, the validity of the fault virtual sample generation method and the fault diagnosis model is verified by the fault diagnosis accuracy.

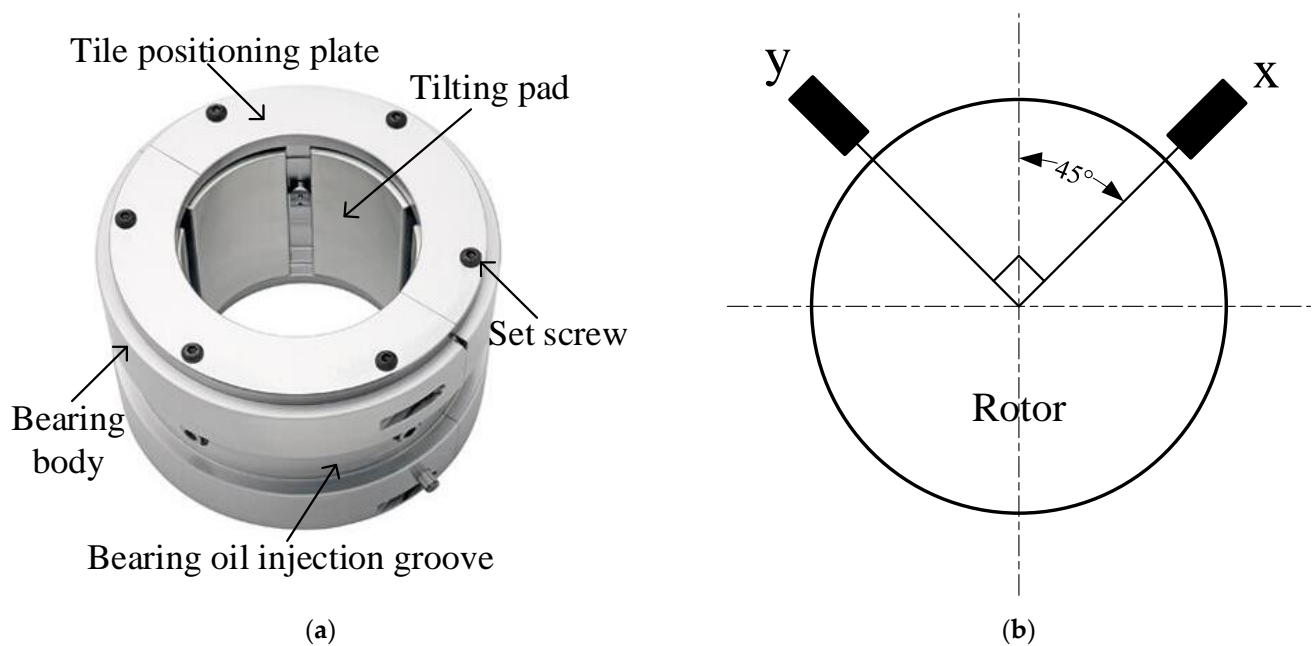
##### 4.1. Rotating Equipment Condition Monitoring System

The rotating equipment condition-monitoring system is composed of shaft, impeller, bearing seat, tilting pad bearing, thrust pad bearing, shaft vibration, shaft displacement, temperature measurement unit, transmitter, signal processing unit, database server and workstation, as shown in Figure 8. The main research object of this paper is the tilting pad bearing, as shown in Figure 9a. Its characteristic is that the bearing surface is composed of multiple tiles. These tiles can be tilted in the bearing seat, so as to adapt to different working conditions during the rotation process. It has the advantages of strong adaptability, load distribution, vibration suppression, adaptability to non-uniform deformation and long service life. In this study, a non-contact eddy current displacement sensor is used to collect the condition-monitoring data for the rotor bearing system. The installation method for the eddy current sensor is shown in Figure 9b. The x-direction and y-direction eddy current sensors are arranged in a 45° angle direction, and the angle between the two sensors is 90°. To avoid energy leakage during data analysis, a synchronous whole-cycle sampling method is used, i.e., the rotor is sampled 32 times for each rotation week, and the number of rotation weeks is 32, so there are 1024 total sampling points in a single sample, and each set of data files (100 × 1024) includes 100 samples of fault data.



**Figure 8.** Rotating equipment condition monitoring system structure.





**Figure 9.** (a) Tilting pad bearings. (b) Installation method of non-contact eddy current sensor.

#### 4.2. Engineering Case Data Validation

##### 4.2.1. Rotor System Condition Monitoring Data Acquisition

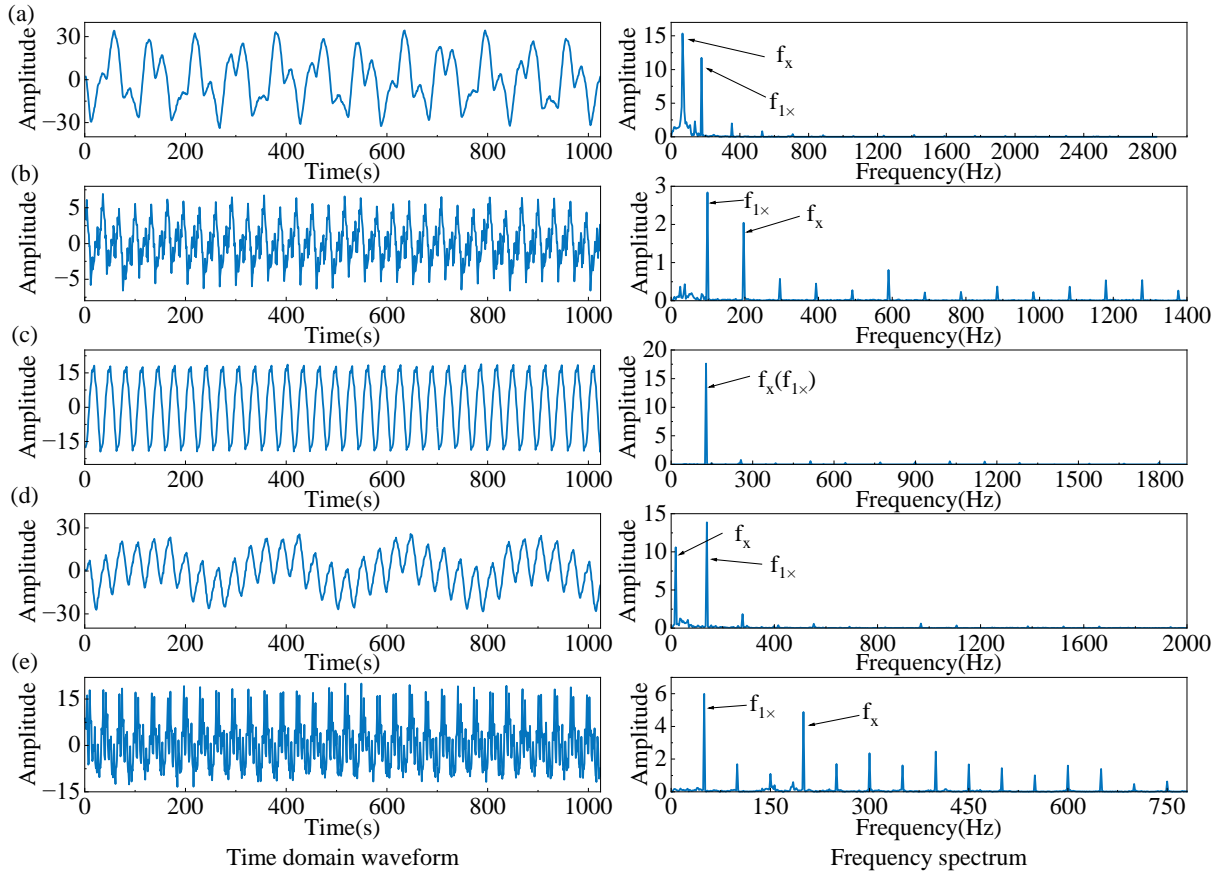
Real engineering case data from mixed refrigerant compressor units, axial flow main air units, syngas compressor units and turbogenerator units have been collected. The types of failures include oil whirl, shaft misalignment, rotor imbalance, surges, and rubbing, which are five types of typical rotor system failures, as shown in Table 2. With the oil-whirl fault data from the syngas compressor units, equipment speed is 10,600 rpm, fault code is Class I. The shaft misalignment fault data comes from the axial main air units; the equipment speed is 5900 rpm, the fault code is Class II. The rotor imbalance fault data comes from the syngas compressor units; the equipment speed is 7700 rpm, the fault code is Class III. The fault data of surges comes from the mixed refrigerant compressor units; the equipment speed is 8300 rpm, the fault code is Class IV. The rubbing data comes from the turbogenerator units; equipment speed is 3000 rpm, fault code is Class V. The normal state of the equipment data code is Class VI.

**Table 2.** Engineering case data introduction.

| Fault Code | Units-Equipment                               | Fault Type         | Unit Code | ID   | Speed/rpm |
|------------|---|--------------------|-----------|------|-----------|
| Class I    | Syngas compressor units—Steam turbine         | Oil whirl          | 170***443 | 0008 | 10600     |
| Class II   | Axial main air units—Shaft system structure   | Shaft misalignment | 171***945 | 0013 | 5900      |
| Class III  | Syngas compressor units—Compressor            | Rotor imbalance    | 170***397 | 0033 | 7700      |
| Class IV   | Mixed refrigerant compressor units—Compressor | Surge              | 200***002 | 0009 | 8300      |
| Class V    | Turbogenerator units—Steam turbine            | Rubbing            | 140***131 | 0015 | 3000      |

Figure 10 shows the time domain waveform and spectrum of the engineering case data. Figure 10a shows the time domain waveform and spectrogram of Class I fault data with fault mechanism characteristic frequency  $f_x = 0.375f_{1x}$ ; Figure 10b shows the time domain waveform and spectrogram of Class II fault data with fault mechanism characteristic frequency  $f_x = 2f_{1x}$ ; Figure 10c shows the time domain waveform and spectrogram of Class III fault data with fault mechanism characteristic frequency  $f_x = f_{1x}$ ; Figure 10d shows the time domain waveform and spectrogram of Class IV fault data with fault mechanism characteristic frequency  $f_x = 0.125f_{1x}$ ; Figure 10e shows the time domain

waveform and spectrogram of Class V fault data with fault mechanism characteristic frequency  $f_x = 4f_{1x}$ . Class I to Class V data are from different units with different operating conditions, so the private characteristics of the fault samples are not the same, but all of them have obvious characteristic frequencies of fault mechanisms. The typical fault characteristic frequencies of rotor system [33] are shown in Table 3.



**Figure 10.** Time domain waveforms and spectrograms of engineering case data. (a) Class I. (b) Class II. (c) Class III. (d) Class IV. (e) Class V.

**Table 3.** Rotor system typical fault characteristic frequency.

|                          | Class I          | Class II                       | Class III              | Class IV                         | Class V                   |
|--------------------------|------------------|--------------------------------|------------------------|----------------------------------|---------------------------|
| Characteristic frequency | $\leq 0.5\times$ | $2\times$ significantly higher | $1\times$              | 1~30 Hz low frequency components | $1/n\times$ and $n\times$ |
| Accompanying frequency   | $1\times$        | $1\times$ , High harmonics     | Smaller high harmonics | $1\times$                        | $1\times$                 |

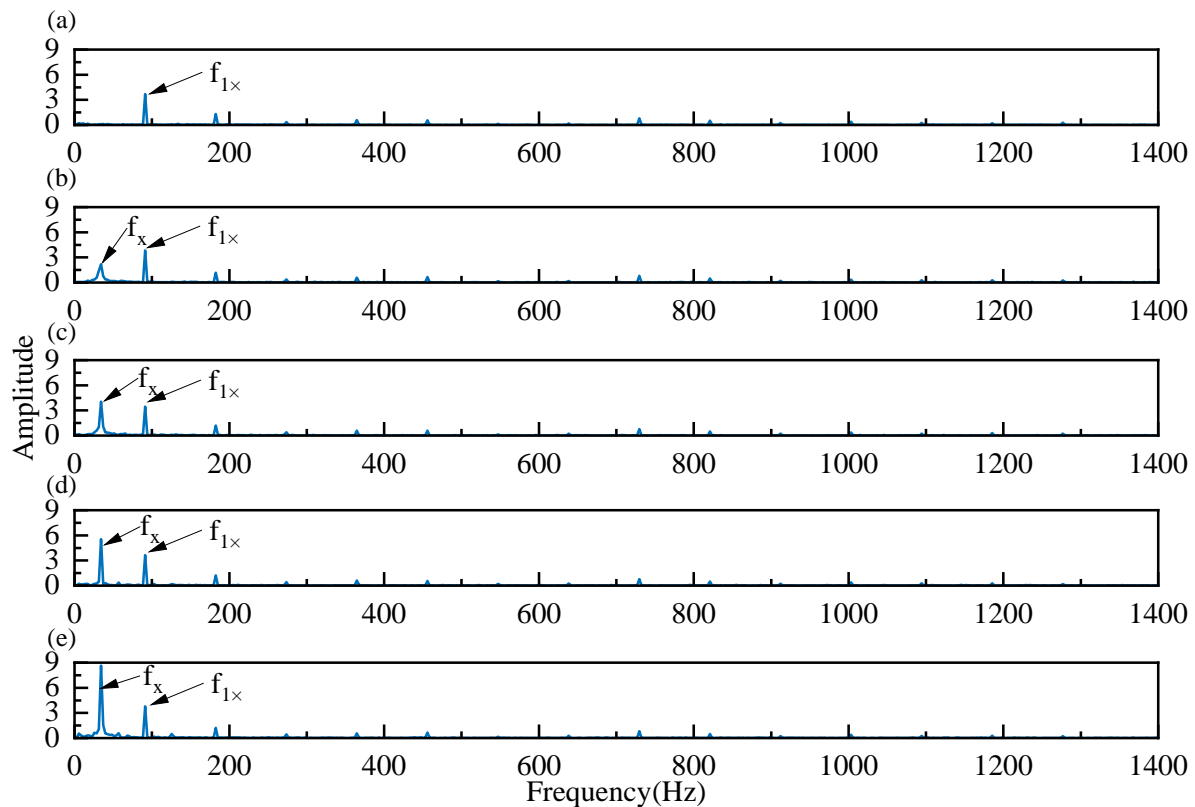
As shown in Table 4, the data sets A to E contain equipment normal-state data (Class VI) and a class of real fault data, and data set F contains equipment normal-state data and five classes of real fault data across equipment and operating conditions. To confirm the efficiency of the proposed approach, the engineering case data are used for three purposes in this research: (1) to compare with the generated fault mechanism feature spectrum samples to verify the similarity between the virtual samples and the real samples; (2) as the test set for fault diagnosis to calculate the fault diagnosis accuracy of the models; (3) the data set F is used as the test data of cross-device and cross-condition research to verify the robustness and versatility of the proposed method.

**Table 4.** Introduction to the data sets.

| Dataset | Class I | Class II | Class III | Class IV | Class V | Class VI |
|---------|---------|----------|-----------|----------|---------|----------|
| A       | ✓       | ×        | ×         | ×        | ×       | ✓        |
| B       | ×       | ✓        | ×         | ×        | ×       | ✓        |
| C       | ×       | ×        | ✓         | ×        | ×       | ✓        |
| D       | ×       | ×        | ×         | ✓        | ×       | ✓        |
| E       | ×       | ×        | ×         | ×        | ✓       | ✓        |
| F       | ✓       | ✓        | ✓         | ✓        | ✓       | ✓        |

Note: ✓ means the corresponding real fault data are available, × means the corresponding real fault data are not available.

As shown in Figure 11, the spectrum analysis of the oil whirl fault data of different fault severity for the same equipment shows that the equipment has experienced a shift from normal state to fault state, and the fault has gradually developed from weak to strong. Figure 11a is in the normal state, and the fault characteristic frequency  $f_x$  of oil whirl has not yet appeared. Figure 11b has a fault characteristic frequency  $f_x$  with a relatively small amplitude. Figure 11b–e, as the fault gradually develops from weak to strong, the amplitude of the fault characteristic frequency  $f_x$  gradually increases.



**Figure 11.** Oil whirl spectrum with different fault severity. (a) The normal state. (b) Grade I oil whirl. (c) Grade II oil whirl. (d) Grade III oil whirl. (e) Grade IV oil whirl.

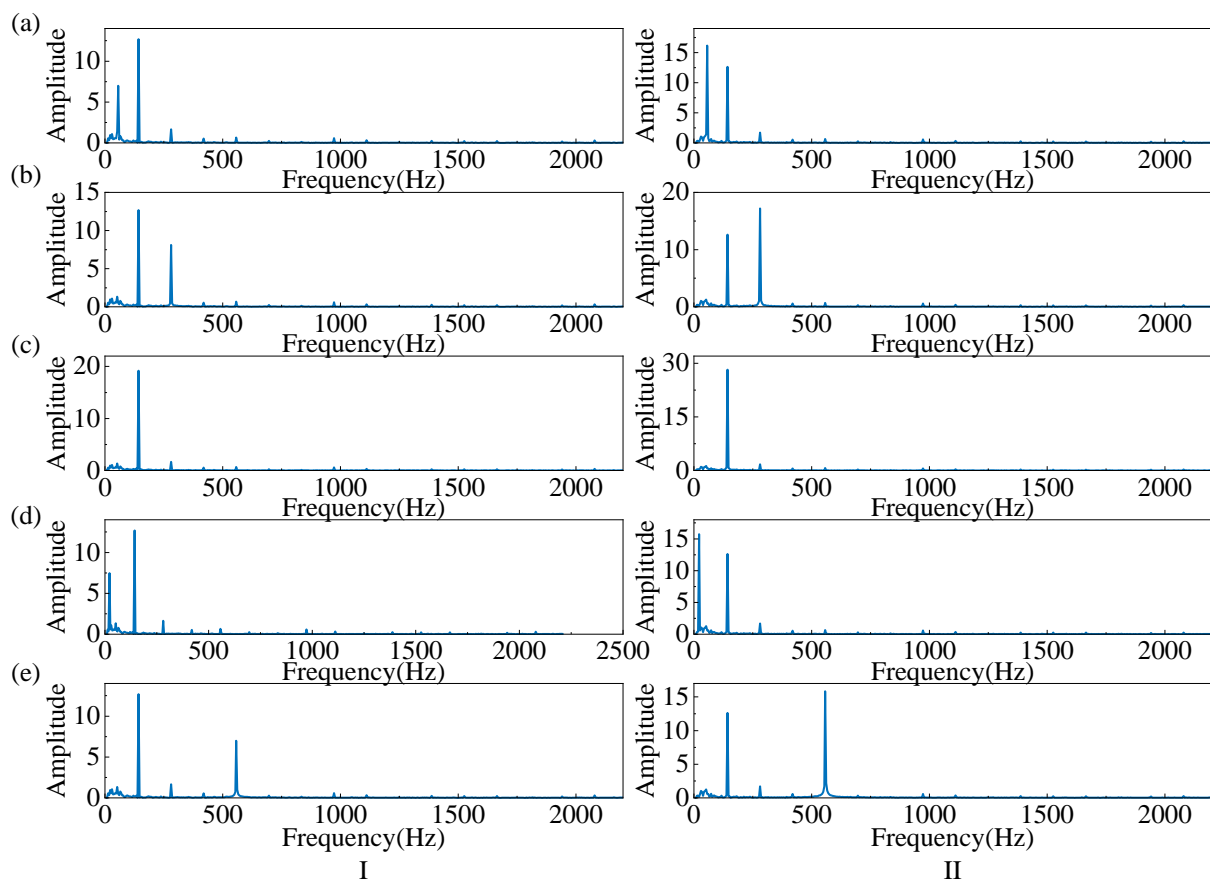
#### 4.2.2. Rotor System FMVS Generation

Take data set A as an example; the device has Class I failure. Show the construction process of the proposed fault diagnosis method based on IACWGAN-GP with this device as an example, and verify the effectiveness of the method. In the engineering application scenario where the target device can only provide normal data, the fault virtual samples are generated based on the normal-state data of the device and fault mechanism feature frequencies, generating a total of five fault types' complete virtual samples for the target device from Class I to Class 5.

As shown in Figure 6b, the fundamental frequency of the device is 176.7 Hz with the amplitude of 11.82, which can also be calculated by the speed  $f_{s_{1\times}} = v/60 = 10600/60 = 176.7$  Hz. The FMVS algorithm process is shown in Algorithm 1. In the table,  $\alpha$  and  $\beta$  are the upper and lower bounds of the value domain of the fault virtual sample characteristic frequency amplitude,  $\alpha$  is 0.5 times the fundamental frequency amplitude,  $\beta$  is 1.2 times the fundamental frequency amplitude, and the amplitude of the fault characteristic frequency is taken within [5.91, 14.18], which can represent the different severity of the fault. Speed  $v = 10600$  rpm. The fault characteristic frequency parameter  $\gamma$  is taken according to the fault mechanism, and the value of  $\gamma$  is shown in Table 5. The characteristic frequency of the fault mechanism of Class I is 0.375 times the fundamental frequency, usually less than 0.5 times the fundamental frequency. Fault mechanism characteristic frequency of Class II is  $2\times$ . Fault mechanism characteristic frequency of Class III is  $1\times$ . Fault mechanism characteristic frequency of Class IV is the low frequency component of 0.125 times the fundamental frequency, usually in the range of 1 to 30 Hz. Fault mechanism characteristic frequency of Class V is  $4\times$ . The data length  $L$  of the normal state data of the equipment is 1024 according to the engineering case data. The number of data groups of normal data and fault virtual samples is set to 100 here. The generated fault virtual samples are shown in Figure 12. Figure 12 I indicate the first group of fault virtual samples with the smallest fault characteristic frequency amplitude, and Figure 12 II indicate the 100th group of fault virtual samples with the largest fault characteristic frequency amplitude.

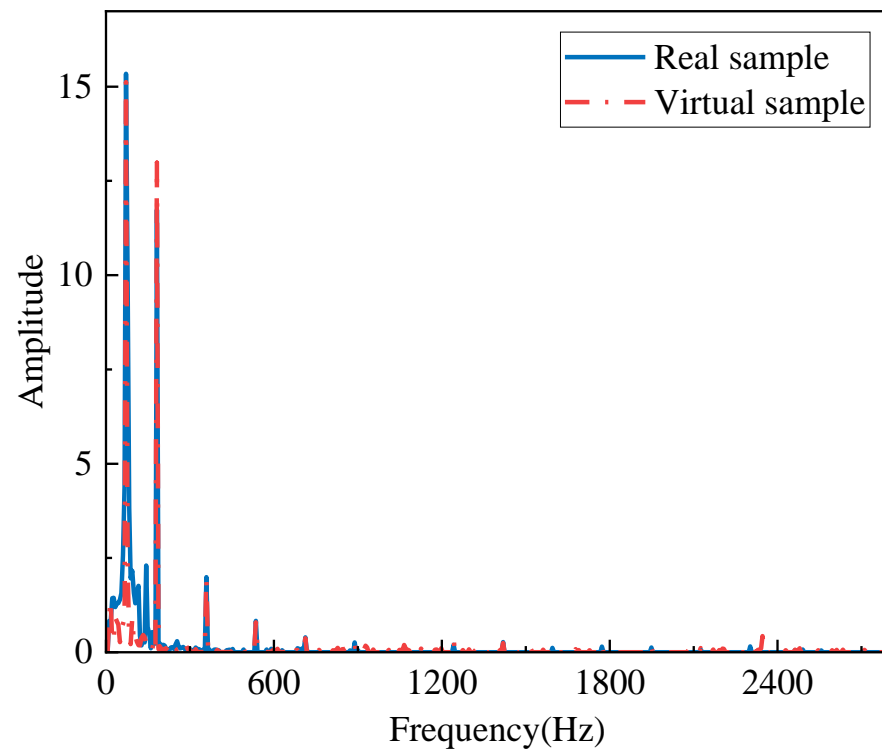
**Table 5.** Fault characteristic frequency parameters.

|          | Class I | Class II | Class III | Class IV | Class V |
|----------|---------|----------|-----------|----------|---------|
| $\gamma$ | 0.375   | 2        | 1         | 0.125    | 4       |



**Figure 12.** Fault virtual samples. (a) Class I. (b) Class II. (c) Class III. (d) Class IV. (e) Class V.

As shown in Figure 13, since the device can provide Class I real fault data, the real fault data can be compared with the generated virtual samples of faults. The real and virtual samples are compared by drawing them under the same coordinate system and it is found that the overlap is very high. The PCC of the real and virtual samples is calculated to be 0.98367, showing that the generated fault virtual samples' data distribution closely resembles that of real fault samples.



**Figure 13.** Comparison of real and virtual sample.

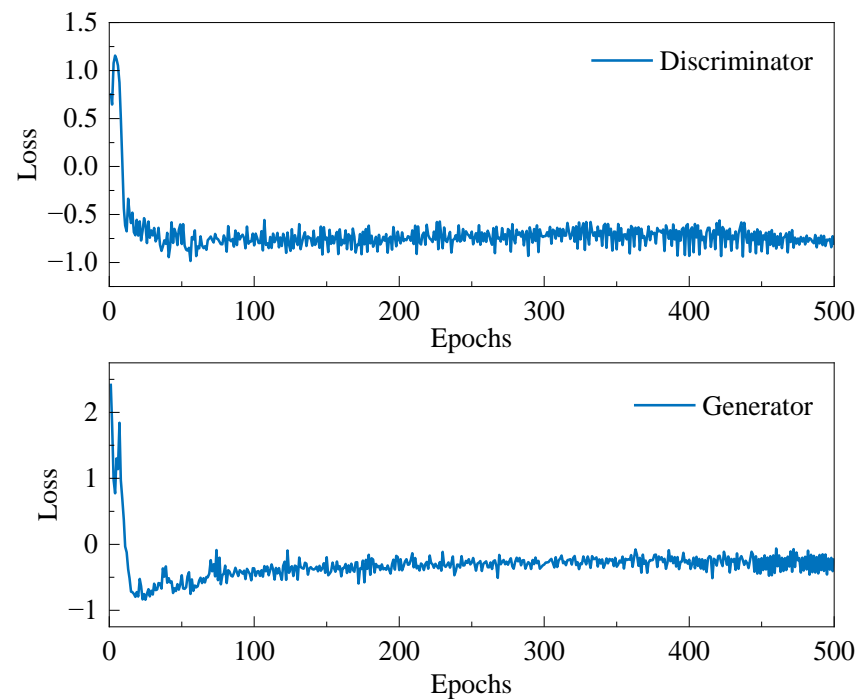
#### 4.2.3. IACWGAN-GP-Based Fault Diagnosis Model Training

The label set  $Y = (y_1, y_2, y_3, \dots, y_k)$  and the random noise vector  $Z = (z_1, z_2, z_3, \dots, z_m)$  are fed to the generator to produce the fake samples  $\hat{X} = G(Z, Y)$ . Then the generated samples  $\hat{X}$  is mixed with the real samples  $X$  and used as the input to the discriminator for true-false discrimination and classification. The generators and discriminators are trained alternately until Nash equilibrium is reached.

The fault virtual samples generated in Section 4.2.2 are used as the training set, training the IACWGAN-GP generation module. To avoid overfitting problems during network training, a dropout layer is added after each layer of the discriminator network and the ratio is set to 0.5. A ReLU activation function is used for both generator activation functions. The discriminator convolutional layer activation functions all use LeakyReLU, and the last two fully connected layers' activation functions use sigmoid and softmax, respectively. The model uses the Adam optimizer, and the learning rates of the generator and discriminator are set to  $10^{-4}$  and  $2 \times 10^{-4}$ , respectively. The batch size is 32. Since Wasserstein distance and a gradient penalty are introduced in the model, it will not cause the problem of the generator gradient disappearing because the discriminator accuracy is too high, so the generator is optimized once for every five instances of discriminator optimization.

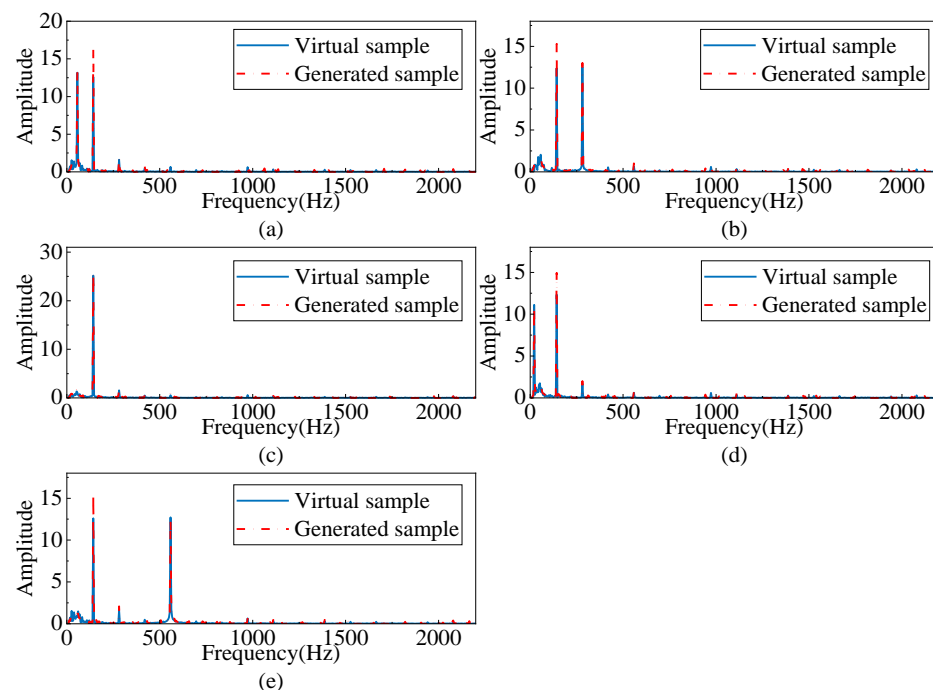
During the training of the model, the values of the loss function are recorded to characterize the performance of the model, as shown in Figure 14. As the number of iterations increases, both the loss of the discriminator and the loss of the generator drop sharply in the beginning stage and stabilize at about 400 iterations, showing that the model is well trained and can be used for samples generation.





**Figure 14.** Loss of discriminator and generator.

To compare the similarity of the generated samples and the fault virtual samples, they are drawn under the same coordinate system as shown in Figure 15. Obviously, these samples are very similar. To further evaluate the quality of the generated samples, PCC and Cosine Similarity (CS) are calculated to measure the similarity between the generated and virtual samples [25]. The results are shown in Table 6, and PCC and CS of both virtual and generated samples exceed 0.9, further showing that the generated samples are highly similar to the virtual samples.



**Figure 15.** Comparison of virtual and generated sample. (a) Class I. (b) Class II. (c) Class III. (d) Class IV. (e) Class V.

**Table 6.** Similarity of virtual and generated samples.

|     | Class I | Class II | Class III | Class IV | Class V |
|-----|---------|----------|-----------|----------|---------|
| PCC | 0.96407 | 0.94688  | 0.99522   | 0.94555  | 0.93213 |
| CS  | 0.96489 | 0.94816  | 0.99528   | 0.94655  | 0.93426 |

An IACWGAN-GP generation module is used to generate 200 samples for each fault type, which are combined with the device's normal-state data to form a full fault types data set to be used as the training set for the IACWGAN-GP diagnosis module. The constructed complete data set is shown in Table 7 A\*, and the real data are used as test data to calculate the fault diagnosis accuracy of the fault diagnosis model. The fault diagnosis model is tested using real Class I and Class VI data. The test experiments are performed 10 times and averaged, and the fault diagnosis accuracy is 100% for both fault types.

**Table 7.** Full fault types data sets.

| Dataset | Class I | Class II | Class III | Class IV | Class V | Class VI |
|---------|---------|----------|-----------|----------|---------|----------|
| A*      | √○      | ○        | ○         | ○        | ○       | √        |
| B*      | ○       | √○       | ○         | ○        | ○       | √        |
| C*      | ○       | ○        | √○        | ○        | ○       | √        |
| D*      | ○       | ○        | ○         | √○       | ○       | √        |
| E*      | ○       | ○        | ○         | ○        | √○      | √        |

Note: √ indicates that the corresponding real fault data is available, ○ indicates the fault virtual sample, \* indicates Full fault types data sets.

Using data set A as an example, Section 4.2.2 shows the process of the IACWGAN-GP-based fault diagnosis model in constructing full fault types dataset A\*. Data set A contains only real Class I fault data and Class VI normal data. In order to more comprehensively show the correlation between the generated fault mechanism feature spectrum samples of the proposed method and real fault samples, the above process is repeated using data set B to data set E. Based on the normal-state data of the equipment in each data set and the fault mechanism feature frequencies, a total of five fault mechanism feature spectrum samples from Class I to Class V are generated using the IACWGAN-GP generation module. The constructed complete fault data sets are shown in Table 7.

As shown in Figure 16, the virtual samples are highly similar to the real samples. The PCC and CS between the virtual sample and the real sample are calculated, and as shown in Table 8, the PCC and CS values of the virtual sample and the real sample are both higher than 0.9, showing that they are highly positively correlated.

**Table 8.** Similarity of virtual sample and real sample.

|     | Class I | Class II | Class III | Class IV | Class V |
|-----|---------|----------|-----------|----------|---------|
| PCC | 0.92215 | 0.96864  | 0.99851   | 0.98367  | 0.93462 |
| CS  | 0.92346 | 0.96911  | 0.99839   | 0.98401  | 0.93543 |

The IACWGAN-GP generation module is trained using the above generated virtual samples to obtain the generated fault mechanism feature spectrum samples. As shown in Figure 17, the generated samples and real samples are very comparable to one another. PCC and CS of generated and real samples are calculated, and as shown in Table 9, these values are higher than 0.88, showing a high positive correlation between them.

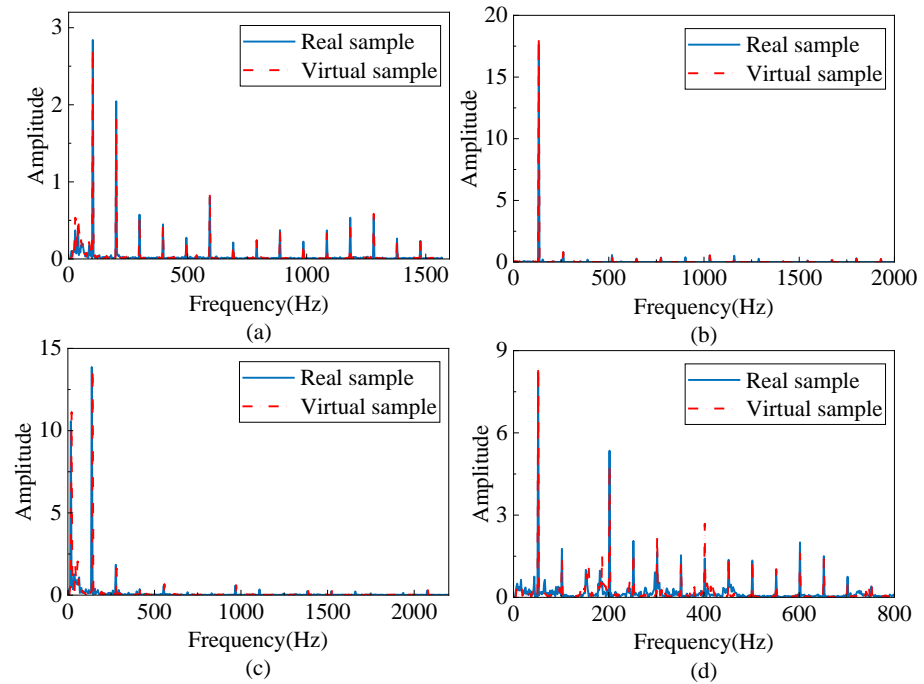


Figure 16. Comparison of virtual and real sample. (a) Class II. (b) Class III. (c) Class IV. (d) Class V.

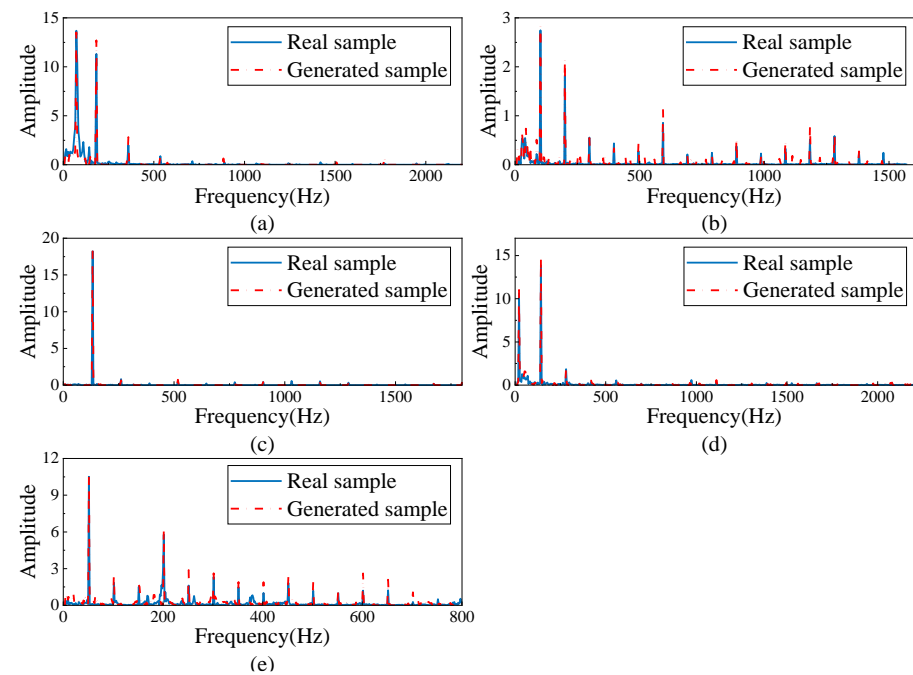


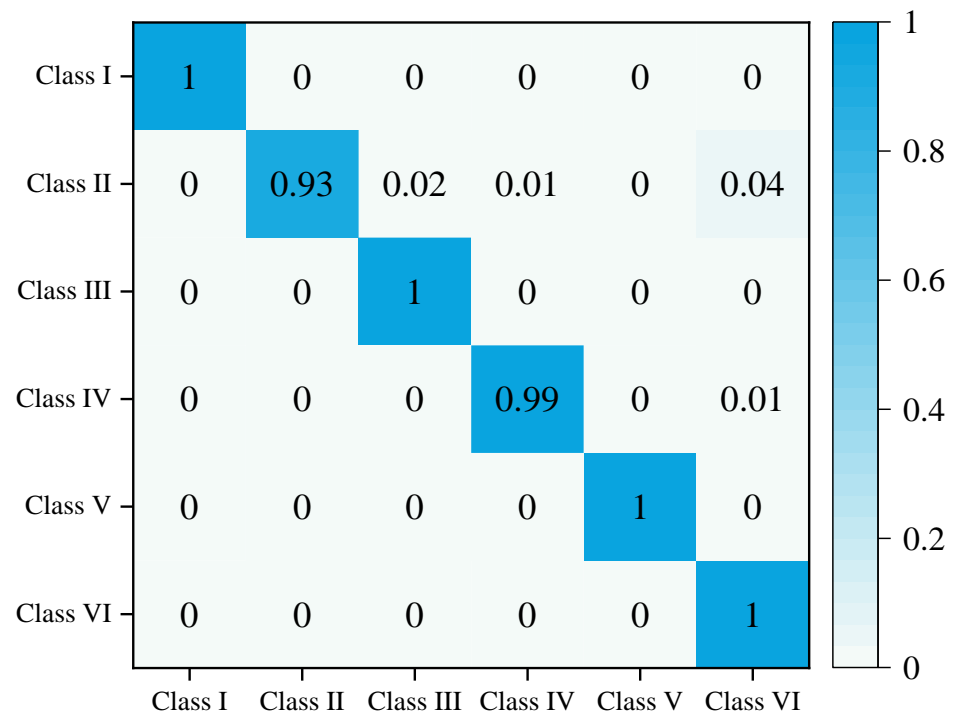
Figure 17. Comparison of generated and real sample. (a) Class I. (b) Class II. (c) Class III. (d) Class IV. (e) Class V.

Table 9. Similarity of generated sample and real sample.

|     | Class I | Class II | Class III | Class IV | Class V |
|-----|---------|----------|-----------|----------|---------|
| PCC | 0.88128 | 0.95451  | 0.99654   | 0.98000  | 0.92743 |
| CS  | 0.88285 | 0.95466  | 0.99637   | 0.98045  | 0.92856 |

#### 4.2.4. IACWGAN-GP-Based Fault Diagnosis Model Test and Methods Comparison

The IACWGAN-GP diagnosis module is trained using data sets  $A^*$  to  $E^*$ , respectively, and the well-trained fault diagnosis module is tested using real Class I to Class VI data, respectively, and all test experiments are conducted 10 times and averaged; the results as shown in Figure 18. The diagnosis accuracy of shaft misalignment faults is 93%, the diagnosis accuracy of oil-whirl faults is 99%, the diagnosis accuracy of the remaining fault types is 100%, and the overall average fault diagnosis accuracy is 98.7%.



**Figure 18.** Fault diagnosis accuracy.

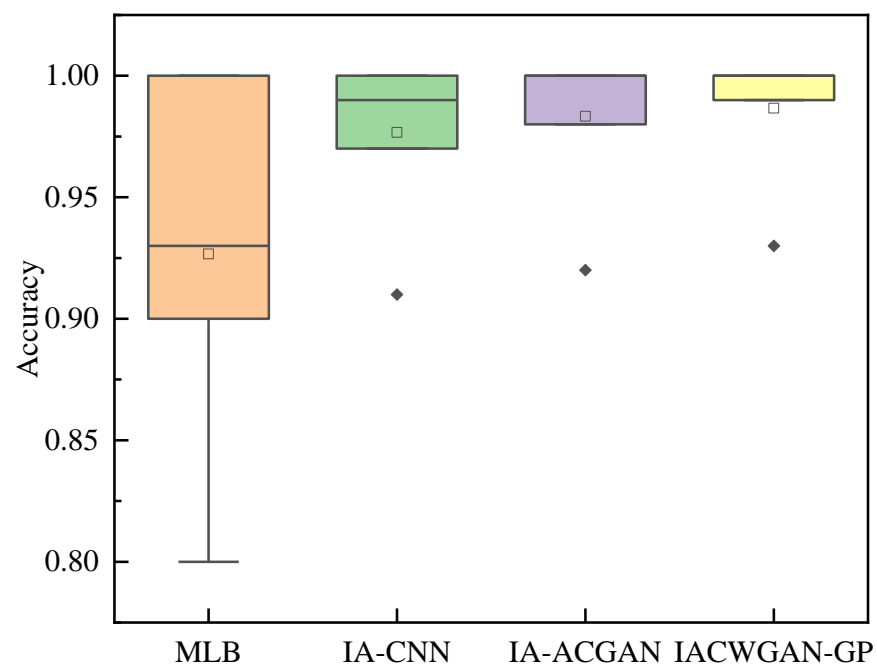
To demonstrate the efficacy of the fault mechanism feature spectrum samples generated by the proposed method, the Classifiers MLP, CNN, and ACGAN are trained separately using the complete fault data sets in Table 7, and then the real data in the data sets is used as test data to calculate the fault diagnosis accuracy of the classification networks and compare them with the IACWGAN-GP diagnosis module classification methods. Since both CNN and ACGAN have convolution operations, the input adaption layer is introduced before the first convolution layer to lift the limitation on the input data length. The structures and parameters of the three Classifiers MLP, IA-CNN, and IA-ACGAN are shown in Table 10.

The classifiers MLP, IA-CNN, and IA-ACGAN are trained using the complete data sets  $A^*$  to  $E^*$  in Table 7, respectively, and then the real data in the data sets are used as test data to calculate the fault diagnosis accuracy of the above classifiers, the results as shown in Figure 19. In the case of providing only the normal data of the equipment and using the method proposed in this paper to generate a complete fault data set to train the classifiers, the fault diagnosis accuracy of MLP, which has a very simple network structure, reaches 0.927, and the highest fault accuracy is the IACWGAN-GP, which reaches 0.987, reflecting the effectiveness of the method proposed in this paper.

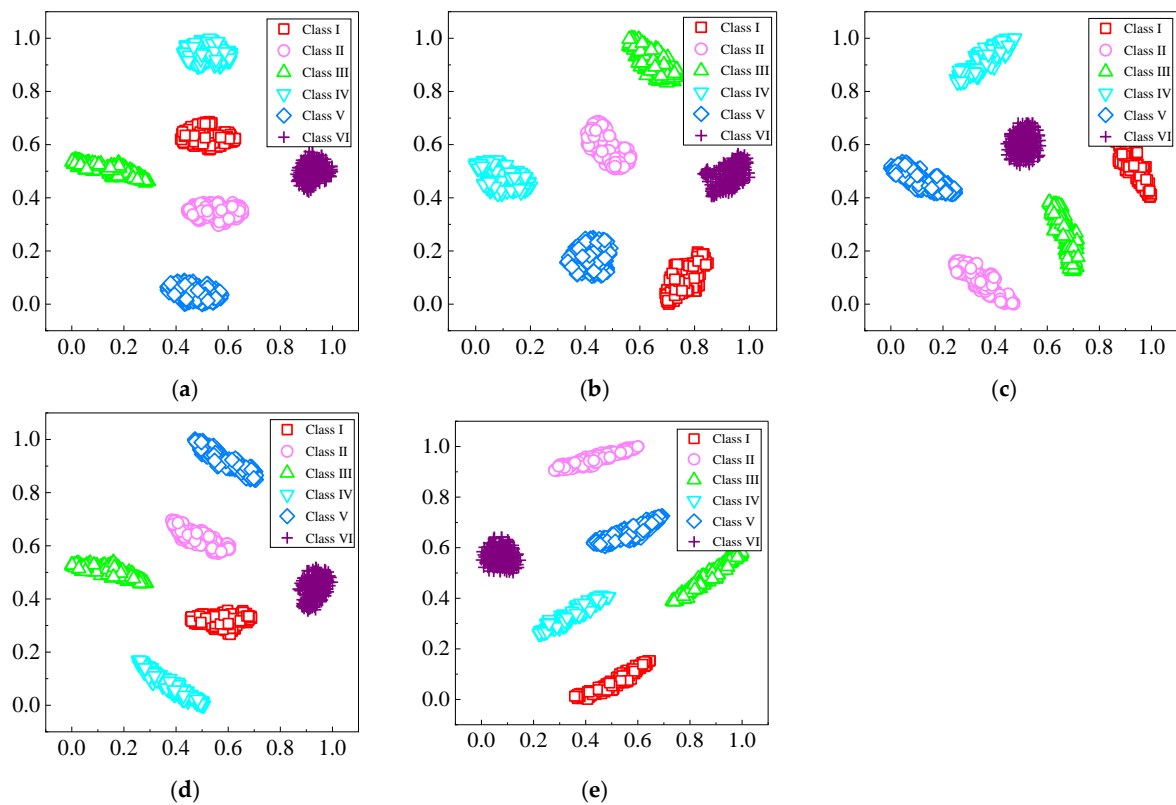
A T-SNE is used to downscale and feature-visualize the complete data sets  $A^*$  to  $E^*$  in Table 7, as shown in Figure 20. From the figure, it can be seen that the same classes of fault data are clustered together, while different classes of fault data are far away from each other, so all four classification networks in Table 11 achieve more than 92% fault diagnosis accuracy.

**Table 10.** Structure and parameters of the classifiers.

| Model                       | Structure                                     | Activation Function | Parameters                    |
|-----------------------------|---|---------------------|-------------------------------|
| MLP                         | Flatten                                       |                     |                               |
|                             | Dense_I                                       | ReLU                | 2048                          |
|                             | Dense_II                                      | ReLU                | 2048                          |
|                             | Dense_III                                     | ReLU                | 2048                          |
|                             | Flatten                                       |                     |                               |
|                             | Dense_IV                                      | ReLU                | 128                           |
|                             | Dense_V                                       | Softmax             | 6                             |
| IA-CNN                      | IAL   |                     |                               |
|                             | Conv2D_I (BatchNormalization/Dropout = 0.5)   | ReLU                | 32@(3, 3)                     |
|                             | MaxPooling2D                                  |                     | (2, 2)                        |
|                             | Conv2D_II (BatchNormalization/Dropout = 0.5)  | ReLU                | 64@(3, 3)                     |
|                             | MaxPooling2D                                  |                     | (2, 2)                        |
|                             | Conv2D_III (BatchNormalization/Dropout = 0.5) | ReLU                | 128@(3, 3)                    |
|                             | Flatten                                       |                     |                               |
|                             | Dense (Dropout = 0.5)                         | ReLU                | 128                           |
|                             | Dense   | Softmax             | 6                             |
| IA-ACGAN<br>(Discriminator) | IAL   |                     |                               |
|                             | Conv2D_I (BatchNormalization/Dropout = 0.5)   | LeakyReLU           | 16@(3, 3)<br>Strides = (1, 1) |
|                             | Conv2D_II (BatchNormalization/Dropout = 0.5)  | LeakyReLU           | 32@(3, 3)<br>Strides = (2, 2) |
|                             | Conv2D_III (BatchNormalization/Dropout = 0.5) | LeakyReLU           | 64@(3, 3)<br>Strides = (2, 2) |
|                             | Flatten                                       |                     |                               |
|                             | Dense   | Sigmoid             | 1                             |
|                             | Dense   | Softmax             | 6                             |

**Figure 19.** Average diagnosis accuracy of comparison methods on engineering case data.





**Figure 20.** Full fault type data set T-SNE feature dimension reduction visualization. (a) Data set A<sup>\*</sup>, (b) Data set B<sup>\*</sup>, (c) Data set C<sup>\*</sup>, (d) Data set D<sup>\*</sup>, (e) Data set E<sup>\*</sup>.

**Table 11.** Comparison of fault diagnosis accuracy.

| Method           | MLP   | IA-CNN | IA-ACGAN | IACWGAN-GP |
|------------------|-------|--------|----------|------------|
| Average accuracy | 0.927 | 0.977  | 0.983    | 0.987      |

In order to verify the robustness and versatility of the proposed method, the data set A<sup>\*</sup> in Table 7 is used to train MLP, IA-CNN, IA-ACGAN and IACWGAN-GP, respectively. The cross-device and cross-condition data set F in Table 4 is used as the test set to calculate the fault diagnosis accuracy, respectively. As shown in Table 12, the proposed method achieves the highest fault diagnosis accuracy of 98% in the case of cross-equipment and cross-working conditions. However, the fault diagnosis accuracy of the comparison methods MLP, IA-CNN and IA-ACGAN are 60.8%, 31.8% and 77.7%, respectively. The method proposed in this paper has higher robustness.

**Table 12.** Comparison of fault diagnosis accuracy across datasets.

| Method           | MLP   | IA-CNN | IA-ACGAN | IACWGAN-GP |
|------------------|-------|--------|----------|------------|
| Average accuracy | 0.608 | 0.318  | 0.777    | 0.980      |

#### 4.2.5. Fault Severity Evaluation

The oil-whirl data of different fault severity shown in Figure 11 amounts to a total of 24,434 sets of data, and the length of each set of data is 1024. The amplitude ratio of oil-whirl fault-characteristic frequency to fundamental frequency is shown in Figure 21a. The data is evaluated for fault severity, and the results are shown in Figure 21b. Due to the large fluctuation of the amplitude ratio of engineering data, the evaluation results of fault severity also fluctuate. In order to reduce the fluctuation of the fault severity evaluation

results, a custom windowing function is used to correct the results. The custom windowing function is defined as follows:

$$G_c^i = \text{mode}(G^{i-\tau} : G^i) \tag{19}$$

where  $G^i$  is the fault severity evaluation result of group i data,  $G_c^i$  is the fault severity evaluation result after correction of group ith data,  $\tau$  is the window scale factor, mode is the mode function, and the mode of the evaluation result within the range of the window scale factor is taken as the correction value of the fault severity evaluation result.

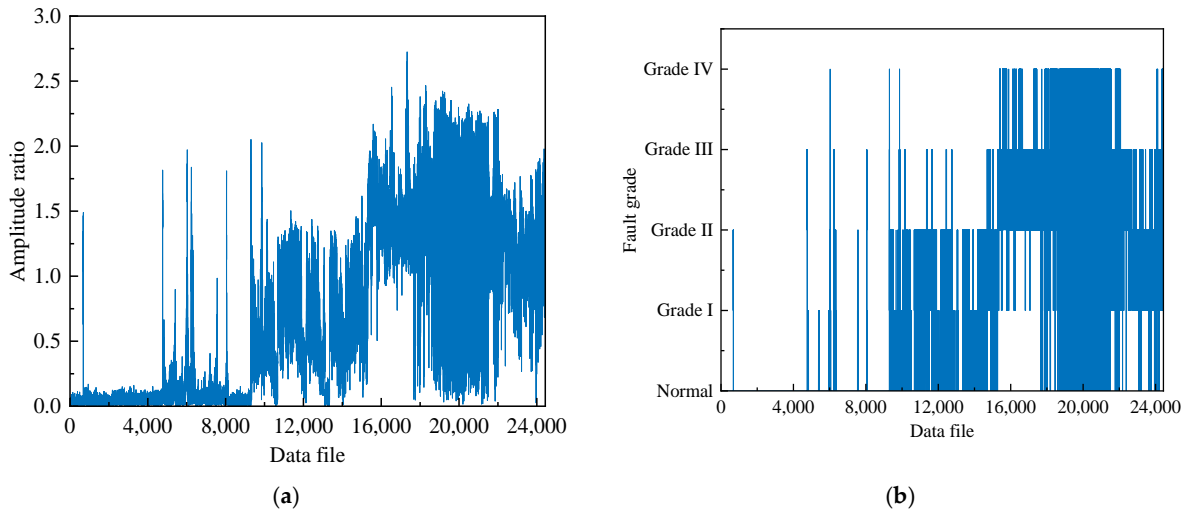


Figure 21. (a) Amplitude ratio of oil-whirl fault data. (b) The results of fault severity evaluation.

The fault severity evaluation results corrected by the custom windowing function are shown in Figure 22. With the increase of the window scale factor  $\tau$ , the fluctuation of the evaluation results decreases, and the development trend of the oil whirl fault severity of the equipment is more obvious.

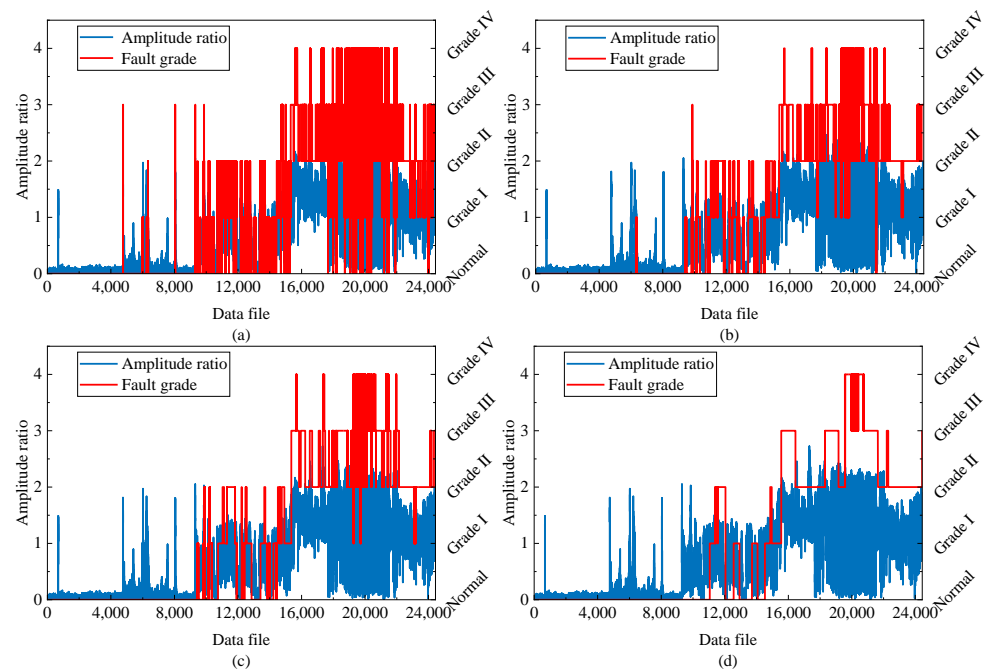
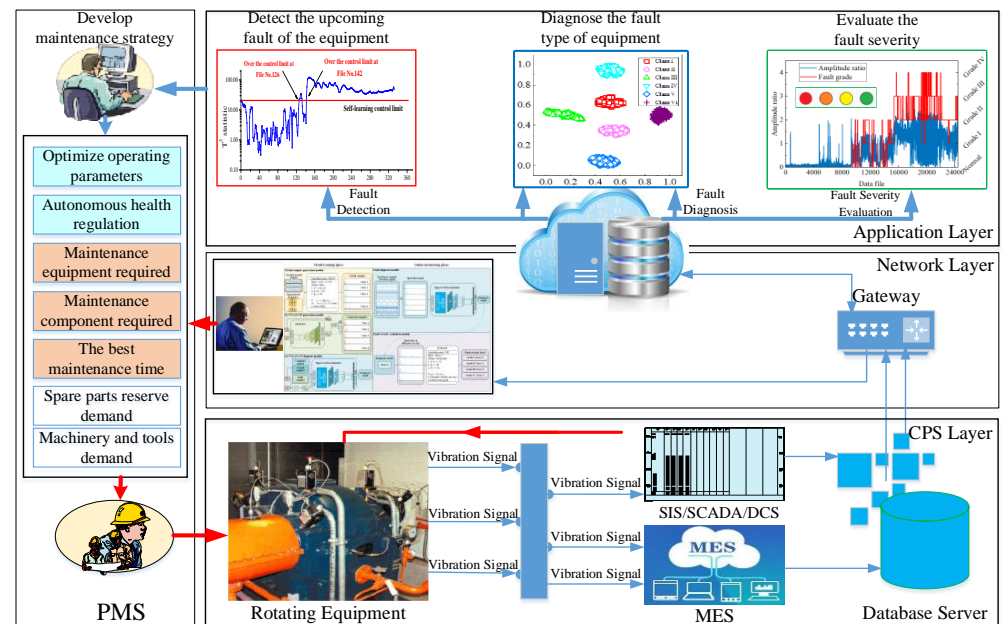


Figure 22. Correction results of fault severity evaluation. (a)  $\tau = 10$ . (b)  $\tau = 50$ . (c)  $\tau = 100$ . (d)  $\tau = 500$ .

#### 4.3. Application of IACWGAN-GP-Based Fault Diagnosis Model in an Industrial Internet Environment

The application of an IACWGAN-GP-based fault diagnosis model in an industrial Internet environment via a cloud-integrated prediction and health management system, which includes a cyber-physical system layer, network layer and application layer, is proposed, as shown in Figure 23. The application layer consists of micro-service systems such as an early fault warning, health evaluation and fault diagnosis.



**Figure 23.** The intelligent fault diagnosis system platform of cloud integrated equipment under the industrial Internet environment.

##### 4.3.1. Cyber-Physical System Layer

The cyber-physical system consists of rotating equipment, Bentley 3500 system, Safety Interlocking System (SIS), Supervisory Control and Data Acquisition (SCADA), Distributed Control System (DCS), Manufacturing Execution System (MES), Condition Monitoring System (CMS), Predictive Maintenance System (PMS) and Intranet communication network.

The function design of a cyber-physical system is mainly for real-time perception of rotating equipment status, dynamic control of abnormal operations, unit fault safety interlock protection and unit fault predictive maintenance control.

##### 4.3.2. Network Layer

The intelligent edge device in the industrial Internet environment is not only a traditional controller or gateway, it can be an edge server. Collecting edge data, intelligent computing and operable decision feedback are three essential capabilities of the edge end.

China's rotating equipment condition-monitoring system (such as S8000, SG8000, etc.) mostly adopts a synchronous full-cycle sampling method: the rotor samples 32 times every rotation cycle, continuously samples 32 cycles, and obtains a data file with 1024 sampling points. A turbine centrifugal compressor unit with a speed of 6000 rpm has a total of 12 vibration displacement measuring points. Using the aforementioned full-cycle sampling method, the vibration waveform data generated every day will reach 12.35 GB. Due to the low value density of the original vibration waveform data, it is unnecessary and unrealistic to upload all 12.35 GB data to the cloud. The original vibration waveform data only performs the fault detection model operation at the edge end. It does not need to upload the original vibration waveform data to the cloud, and only needs to upload the result of the fault detection model operation to the cloud. This not only helps to reduce the computing

load of the cloud server and the pressure on network bandwidth occupation caused by data uploading to the cloud, but it also enables fast and efficient potential fault-warning and predictive maintenance decisions.

#### 4.3.3. Application Layer

The cloud design adopts an industrial micro-service architecture, which mainly includes modular components such as fault detection, fault diagnosis, and health evaluation.

Fault detection micro-service is used to detect the occurrence of potential faults in rotating equipment. Fault diagnosis micro-service is used to identify the potential fault type and fault location of rotating equipment. Health assessment micro-service is used to evaluate the severity of rotating equipment failure.

Intelligentization is the essential feature of modular components of industrial micro-services. Industrial micro-service modular components such as fault detection, fault diagnosis, and health evaluation provide knowledge for making maintenance decisions and optimizing maintenance tasks.

### 5. Conclusions

Aiming at the engineering status of unbalanced data samples for rotating equipment, this paper studies an improved auxiliary classifier Wasserstein generative adversarial network with gradient penalty for fault diagnosis of tilting pad bearings. The work can be summarized as follows:

(1) An improved auxiliary classifier Wasserstein generative adversarial network with gradient penalty is developed, in which the input data length adaptive layer is added before the 2D convolution layer of the discriminator. It overcomes the limitation of neural networks on the length of input data and improves the applicability and generalization of neural networks to various types of data.

(2) A fault diagnosis method based on IACWGAN-GP for tilting pad bearings is proposed, which is able to accurately identify the early faults of tilting pad bearing oil whirl despite the interference of shaft misalignment, rotor imbalance, surges, rubbing and other faults that may occur simultaneously in rotating equipment. This method can identify oil whirl faults as they develop from weak to strong, and evaluate the grade of the fault. The engineering case-data verification results show that, with only normal data of the equipment, the model can achieve an accuracy of 98.7% in spotting upcoming faults. Train Multilayer Perceptron, CNN and Auxiliary Classifier GAN fault diagnosis models using full-fault virtual samples, and the accuracy of the models reach 92.7%, 97.7%, and 98.3%, separately. The proposed method and three comparison methods are tested by using cross-device and cross-condition engineering case data sets. The fault diagnosis accuracy of the proposed method and the three comparison methods are 98%, 60.8%, 31.8% and 77.7%, respectively, and the proposed method shows better robustness.

(3) The application of an IACWGAN-GP-based fault diagnosis model in an industrial Internet environment, via a cloud-integrated prediction and health management system, which includes cyber-physical system layer, network layer and application layer, is proposed. The application layer consists of micro-service systems such as early fault warning, health evaluation and fault diagnosis.

In this paper, the typical fault diagnosis of rotor systems is studied, and the proposed fault diagnosis method has a high fault diagnosis accuracy and robustness. However, the engineering case data involved in this paper only contains the data of a single type of fault. When the equipment has multiple faults at the same time, the proposed method can only draw a diagnosis conclusion for one of the faults. In addition, when a fault outside the fault category included in the training data set occurs, the proposed method will draw a similar diagnostic conclusion based on the similarity of fault characteristics between the unknown fault and the known fault, which may lead to incorrect diagnosis results.

Future studies will collect more complex fault engineering case data, use the virtual sample generation module in the proposed method to generate complex virtual fault

samples based on the normal data of equipment and the fault characteristic frequency of complex faults, and use the engineering case data for experimental verification. In addition, the knowledge base of fault mechanisms except typical faults will be extended, and the range of fault categories in the training set of the model will be expanded to solve the fault diagnosis problem of unknown faults to a certain extent. The proposed method can also be applied to the fault diagnosis of various types of bearings, gears and other key components, providing support for fault prediction and health management of rotating equipment.

**Author Contributions:** Conceptualization, C.Z.; methodology, C.Z. and Q.W.; software, C.Z.; validation, C.Z.; formal analysis, C.Z.; resources, W.X. and Y.S.; data curation, C.Z.; writing—original draft preparation, C.Z.; writing—review and editing, C.Z., Q.W. and Y.X.; visualization, C.Z.; supervision, W.X. and Y.S.; project administration, Q.W.; funding acquisition, W.X. and Y.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by PipeChina, grant number WZXGL202107, and State Key Laboratory of Compressor Technology, grant number SKL-YSJ202101.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The authors do not have permission to share data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Li, C.; Yun, M.; Rui, Z.; Wei, W. Rotor system fault diagnosis based on semi-supervised spectrum kernel clustering. *J. Vib. Meas. Diagnosis* **2016**, *36*, 562–567.
2. Xiong, H.; Peng, Y.; Hu, Y.; Zhang, L.; Li, Y. Vibration fault signal analysis and diagnosis of flue gas turbine. *Eng. Fail. Anal.* **2022**, *134*, 105981. [[CrossRef](#)]
3. Nath, A.; Sharma, A.; Udmale, S.; Singh, S. An Early Classification Approach for Improving Structural Rotor Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–13. [[CrossRef](#)]
4. Liu, Z.; Zhang, C.; Dong, E.; Wang, R.; Li, S.; Han, Y. Research Progress and Development Trend of Prognostics and Health Management Key Technologies for Equipment Diesel Engine. *Processes* **2023**, *11*, 1972. [[CrossRef](#)]
5. Zhong, Z.; Liu, H.; Mao, W.; Xie, X.; Cui, Y. Rolling Bearing Fault Diagnosis across Operating Conditions Based on Unsupervised Domain Adaptation. *Lubricants* **2023**, *11*, 383. [[CrossRef](#)]
6. Mohiuddin, M.; Islam, M.S.; Islam, S.; Miah, M.S.; Niu, M.-B. Intelligent Fault Diagnosis of Rolling Element Bearings Based on Modified AlexNet. *Sensors* **2023**, *23*, 7764. [[CrossRef](#)]
7. Cui, W.; Ding, J.; Meng, G.; Lv, Z.; Feng, Y.; Wang, A.; Wan, X. Fault Diagnosis of Rolling Bearings in Primary Mine Fans under Sample Imbalance Conditions. *Entropy* **2023**, *25*, 1233. [[CrossRef](#)]
8. Zhang, W.-T.; Liu, L.; Cui, D.; Ma, Y.-Y.; Huang, J. An Anti-Noise Convolutional Neural Network for Bearing Fault Diagnosis Based on Multi-Channel Data. *Sensors* **2023**, *23*, 6654. [[CrossRef](#)]
9. Shen, W.; Xiao, M.; Wang, Z.; Song, X. Rolling Bearing Fault Diagnosis Based on Support Vector Machine Optimized by Improved Grey Wolf Algorithm. *Sensors* **2023**, *23*, 6645. [[CrossRef](#)]
10. Huang, W.; Zhang, G. Bearing Fault-Detection Method Based on Improved Grey Wolf Algorithm to Optimize Parameters of Multistable Stochastic Resonance. *Sensors* **2023**, *23*, 6529. [[CrossRef](#)]
11. Tian, H.; Fan, H.; Feng, M.; Cao, R.; Li, D. Fault Diagnosis of Rolling Bearing Based on HPSO Algorithm Optimized CNN-LSTM Neural Network. *Sensors* **2023**, *23*, 6508. [[CrossRef](#)] [[PubMed](#)]
12. Xiao, Y.; Wang, Q.; Wang, S.; Chen, W. Research on a Multisource Domain Improved Fault Diagnosis Method of the Rotor System. *IEEE Access* **2022**, *10*, 85399–85415. [[CrossRef](#)]
13. Wang, S.; Wang, Q.; Xiao, Y. Research on rotor system fault diagnosis method based on vibration signal feature vector transfer learning. *Eng. Fail. Anal.* **2022**, *139*, 106424. [[CrossRef](#)]
14. Yu, J.; Wen, Y.; Yang, L.; Zhao, Z.; Guo, Y.; Guo, X. Monitoring on Triboelectric Nanogenerator and Deep Learning Method. *Nano Energy* **2022**, *92*, 106698. [[CrossRef](#)]
15. Ruan, D.; Song, X.; Gühmann, C.; Yan, J. Collaborative Optimization of CNN and GAN for Bearing Fault Diagnosis under Unbalanced Datasets. *Lubricants* **2021**, *9*, 105. [[CrossRef](#)]
16. Wu, Z.; Yan, H.; Zhan, X.; Wen, L.; Jia, X. Gearbox Fault Diagnosis Based on Optimized Stacked Denoising Auto Encoder and Kernel Extreme Learning Machine. *Processes* **2023**, *11*, 1936. [[CrossRef](#)]
17. Lei, Y.; Jia, F.; Kong, D.; Lin, J.; Xing, S. Opportunities and challenges of intelligent fault diagnosis of machinery under big data. *J. Mech. Eng.* **2018**, *54*, 94–104. [[CrossRef](#)]

18. Luo, J.; Zhu, L.; Li, Q.; Liu, D.; Chen, M. Imbalanced Fault Diagnosis of Rotating Machinery Based on Deep Generative Adversarial Networks with Gradient Penalty. *Processes* **2021**, *9*, 1751. [[CrossRef](#)]
19. Ruan, D.; Chen, X.; Gühmann, C.; Yan, J. Improvement of Generative Adversarial Network and Its Application in Bearing Fault Diagnosis: A Review. *Lubricants* **2023**, *11*, 74. [[CrossRef](#)]
20. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, BC, Canada, 8–13 December 2014; Volume 2, pp. 2672–2680.
21. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.
22. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 214–223.
23. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A. Improved training of Wasserstein GANs. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 5767–5777.
24. Odena, A.; Olah, C.; Shlens, J. Conditional Image Synthesis With Auxiliary Classifier GANs. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 2642–2651.
25. Yu, Y.; Guo, L.; Gao, H.; Liu, Y. PCWGAN-GP: A New Method for Imbalanced Fault Diagnosis of Machines. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3515711. [[CrossRef](#)]
26. Li, Z.; Zheng, T.; Wang, Y.; Cao, Z.; Guo, Z.; Fu, H. A Novel Method for Imbalanced Fault Diagnosis of Rotating Machinery Based on Generative Adversarial Networks. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3500417. [[CrossRef](#)]
27. Cao, Z.; Li, Z.; Zhang, J.; Fu, H. A Homogeneous Stacking Ensemble Learning Model for Fault Diagnosis of Rotating Machinery With Small Samples. *IEEE Sens. J.* **2022**, *22*, 8944–8959. [[CrossRef](#)]
28. Gao, Z.; Shi, F.; Zhang, B.; Su, Y. Analysis of convolutional neural networks for defect type recognition. *Appl. Acoust.* **2022**, *41*, 301–309.
29. Miao, J.; Wang, J.; Zhang, D.; Miao, Q. Improved Generative Adversarial Network for Rotating Component Fault Diagnosis in Scenarios With Extremely Limited Data. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3500213. [[CrossRef](#)]
30. Du, G.; Jiang, T.; Wang, J. Improved Multi-Bandwidth Mode Manifold for Enhanced Bearing Fault Diagnosis. *Chin. J. Mech. Eng.* **2022**, *35*, 14. [[CrossRef](#)]
31. Prion, S.; Haerling, K. Making Sense of Methods and Measurement: Pearson Product-Moment Correlation Coefficient. *Clin. Simul. Nurs.* **2014**, *10*, 587–588. [[CrossRef](#)]
32. Allahyani, M.; Alsulami, R.; Alwafi, T. DivGAN: A diversity enforcing generative adversarial network for mode collapse reduction. *Artif. Intell.* **2023**, *317*, 103863. [[CrossRef](#)]
33. Sheng, Z. *Equipment Condition Monitoring and Fault Diagnosis Technology and Application*; Chemical Industry Press: Beijing, China, 2003.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.