

## Article

# Multi-Mode Data Generation and Fault Diagnosis of Bearings Based on STFT-SACGAN

Hongxing Wang , Hua Zhu \* and Huafeng Li \*

State Key Laboratory of Mechanics and Control for Aerospace Structures, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; whx\_whx@nuaa.edu.cn

\* Correspondence: hzhu103@nuaa.edu.cn (H.Z.); lihuaf@nuaa.edu.cn (H.L.)

**Abstract:** To achieve multi-mode fault sample generation and fault diagnosis of bearings in a complex operating environment with scarce labeled data. Combining a semi-supervised generative adversarial network (SGAN) and an auxiliary classifier generative adversarial network (ACGAN), a semi-supervised auxiliary classifier generative adversarial network (SACGAN) is constructed in this paper. The network structure and the loss function are improved. A fault diagnosis method based on STFT-SACGAN is also proposed. The method uses a short-time Fourier transform (STFT) to convert one-dimensional time-domain vibration signals of bearings into two-dimensional time-frequency images, which are used as the input of SACGAN. Two multi-mode fault data generation and intelligent diagnosis cases for bearings are studied. The experimental results show that the proposed method generates high-quality multi-mode fault samples with high fault diagnosis accuracy, generalization, and stability.

**Keywords:** bearing fault diagnosis; generative adversarial network; short-time Fourier transform; multi-mode data generation



**Citation:** Wang, H.; Zhu, H.; Li, H. Multi-Mode Data Generation and Fault Diagnosis of Bearings Based on STFT-SACGAN. *Electronics* **2023**, *12*, 1910. <https://doi.org/10.3390/electronics12081910>

Academic Editors: Andrea Prati, Luis Javier García Villalba and Vincent A. Cicirello

Received: 28 March 2023

Revised: 14 April 2023

Accepted: 17 April 2023

Published: 18 April 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

With the vigorous development of modern industry, rotating machinery plays a crucial role in intelligent equipment, and its safety is widely concerned [1]. Bearings, as indispensable core transmission components in rotating machinery, under complex operating environments such as high speed and heavy load, is subject to various fault types, such as wear, corrosion, cracks, and deformation [2]. Faulty bearings directly affect the operational reliability of rotating machinery and cause potentially significant accidents, which lead to economic losses and even casualties [3]. Therefore, it is of great research value to carry out bearing fault diagnosis and predictive maintenance.

Traditional data-driven fault diagnosis methods aim to extract effective fault features from original signals to achieve high accuracy, which involves feature extraction and fault recognition [4]. For feature extraction, signal processing technologies are used to obtain effective features, such as the Fourier transform (FT) [5], the short-time Fourier transform (STFT) [6], and the wavelet transform (WT) [7], which convert the original signals in the time domain into the frequency domain or time-frequency domain. For fault recognition, machine learning methods, including k-nearest neighbor (KNN), support vector machines (SVM), and artificial neural networks (ANN), are employed to train classification models using the extracted features to recognize the fault types [8].

In recent years, deep learning-based fault diagnosis methods have gradually become a research hotspot relying on powerful fault feature learning capability and end-to-end diagnosis characteristics, including auto-encoder (AE), deep belief network (DBN), convolutional neural network (CNN), and recurrent neural network (RNN) [9,10]. Although the above models have shown effectiveness in fault diagnosis, they rely on a large amount of labeled data for supervised learning [11]. However, obtaining sufficient labeled fault

samples in practical engineering is difficult, which leads to serious overfitting problems with supervised models [12]. Therefore, training accurate and reliable deep learning-based fault diagnosis models using limited labeled fault samples is worth investigating.

The generative adversarial network (GAN) is one of the critical techniques of unsupervised generative models, which can generate new samples similar to existing samples while using only unlabeled samples. Variants of GAN have also received increasing attention in the field of fault diagnosis [13]. SGAN is a semi-supervised generative model improved by GAN, and some scholars have used SGAN to solve fault diagnosis problems in recent years. Pang et al. used CNN as the generator network and the discriminator network for SGAN for single- and compound-fault diagnosis of gears in different gearboxes [14]. Xing et al., based on SGAN, performed feature matching on the intermediate layers of the generator and the discriminator to fully extract the features of the fault samples to achieve bearing fault diagnosis with a small number of labeled samples [15]. Yang et al. converted one-dimensional vibration signals into two-dimensional images by variational mode decomposition as the input to SGAN, and the switchable normalization was used to replace batch normalization and used the trained discriminator to solve the problem of fault diagnosis of rolling bearings [16]. ACGAN is a supervised generative model improved by GAN, and ACGAN has been used to solve multi-mode sample generation and fault diagnosis problems in recent years. Lu et al. proposed a modified ACGAN to improve the training stability under noise interference by gradient penalty and to generate wind turbine primary bearing vibration signal data [17]. Li et al. proposed an improved ACGAN based on Bayesian optimization and the Wasserstein distance to improve fault feature extraction and generation ability for data enhancement and fault diagnosis of wind turbine planetary gearboxes [18]. Li et al. added a classifier to ACGAN to improve the compatibility of discrimination and classification for multi-mode data augmentation and fault diagnosis of bearings and gears [19].

The above studies show that SGAN is a semi-supervised learning mechanism, but it cannot generate multi-mode samples due to the lack of participation in category information. ACGAN can generate multi-mode samples using category labels as auxiliary information, but its supervised learning mechanism requires a large amount of labeled data. To achieve multi-mode fault sample generation and fault diagnosis for bearings with limited labeled data, SACGAN is constructed. Based on SGAN and ACGAN, the network structure and the loss function are improved, and the adversarial learning of the generator and the discriminator is carried out using a small amount of labeled data and a large amount of unlabeled data to improve the multi-mode fault sample generation ability of the generator and the recognition ability of the discriminator. A fault diagnosis method based on STFT-SACGAN is proposed, using STFT to convert one-dimensional time-domain vibration signals of bearings into two-dimensional time-frequency images as the input of SACGAN. To demonstrate the effectiveness and generalization of the proposed method, validation and analysis are carried out on two bearing datasets.

The remaining content of this paper is organized as follows: Section 2 describes the brief theories of GANs. Section 3 details the proposed SACGAN. Section 4 presents the fault diagnosis method based on STFT-SACGAN. Section 5 provides the experimental validation and analysis results. Finally, Section 6 concludes this paper and introduces future work.

## 2. The Brief Theories of GANs

### 2.1. GAN

GAN is an unsupervised generative model proposed by Goodfellow et al. in 2014, which is created based on zero-sum game theory and mainly consists of a generator  $G$  and a discriminator  $D$  [20]. The architecture of a basic GAN is shown in Figure 1. The generator obtains the generative sample  $G(z)$  from random noise  $z$ , and the discriminator is used to discriminate whether the input sample is a real sample  $x$  or a generative sample  $G(z)$ . Through iterative adversarial training, the performance of the generator and the

discriminator are simultaneously improved, eventually reaching Nash equilibrium. The loss function of GAN is defined as:

$$\min_G \max_D V(D, G) = E_{x \sim P_r(x)} [\log(D(x))] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \tag{1}$$

where  $P_r(x)$  and  $P_z(z)$  are the prior distributions of  $x$  and  $z$ , respectively,  $E_{x \sim P_r(x)}$  denotes the expectation of  $x$  from  $P_r(x)$ ,  $E_{z \sim P_z(z)}$  denotes the expectation of  $z$  from  $P_z(z)$ . The generator’s goal is to minimize  $V(D,G)$ , while the discriminator’s goal is to maximize  $V(D,G)$  in the adversarial training process.

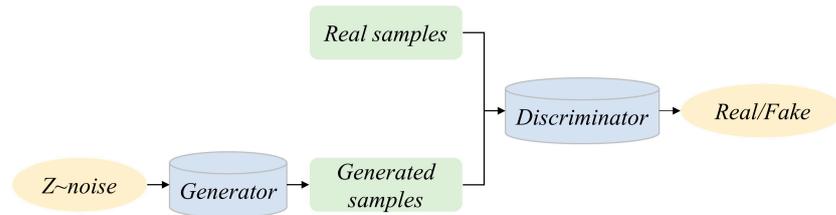


Figure 1. Architecture of GAN.

2.2. SGAN

SGAN is a semi-supervised generative model that Augustus Odena et al. proposed in 2016. SGAN extends GAN to semi-supervised scenarios by forcing the discriminator to output category labels [21]. The discriminator of traditional GAN uses a Sigmoid function as the output, while SGAN uses a Softmax function as the output. For a dataset with  $N$  categories, the discriminator extends the output to  $N + 1$ , including one discriminative output and  $N$  categorical output. The architecture of a basic SGAN is shown in Figure 2, and the input of the discriminator consists of three parts: labeled real samples, unlabeled real samples, and generated samples from the generator. The loss function of SGAN is defined as:

$$L_g = E_{z \sim P_z(z)} [\log(1 - P(y = N + 1|G(z)))] \tag{2}$$

$$L_{supervised} = E_{x,y \sim P_r(x,y)} [-\log P(y|x, y < N + 1)] \tag{3}$$

$$L_{unsupervised} = E_{x \sim P_r(x)} [-\log(1 - P(y = N + 1|x))] + E_{z \sim P_z(z)} [-\log(P(y = N + 1|G(z)))] \tag{4}$$

$$L_d = L_{supervised} + L_{unsupervised} \tag{5}$$

where  $L_g$  and  $L_d$  represent the loss functions of the generator and the discriminator, respectively.  $L_{supervised}$  and  $L_{unsupervised}$  denote the discriminator’s supervised loss function and the unsupervised loss function, respectively.

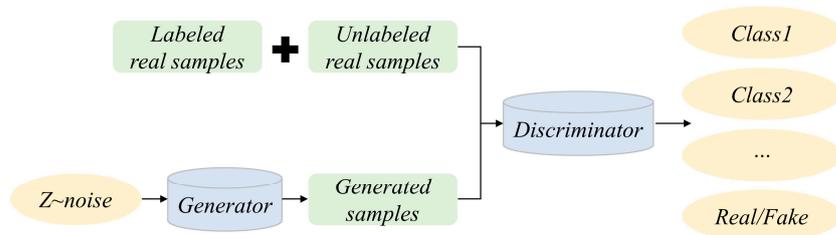


Figure 2. Architecture of SGAN.

2.3. ACGAN

ACGAN is a supervised generative model that Augustus Odena et al. proposed in 2017 [22]. The architecture of the basic ACGAN is shown in Figure 3. By embedding the category labels as auxiliary information into the random noise input, the generator is guided to

generate multi-mode samples. Unlike SGAN, the discriminator of ACGAN uses a Sigmoid function and a Softmax function as outputs to achieve discrimination and classification of the input samples. The loss function of ACGAN consists of two components:

$$L_{source} = E_{x \sim P_r(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(D(G(z, c_g)))] \tag{6}$$

$$L_{class} = E_{x \sim P_r(x)} [\log P(c = c_r | x)] + E_{z \sim P_z(z)} [\log P(c = c_g | G(z, c_g))] \tag{7}$$

where  $L_{source}$  denotes the loss function used to measure the validity of discriminating samples from  $x$ ,  $L_{class}$  denotes the loss function used to measure the validity of sample categories, and  $c_r$  and  $c_g$  are labels of  $x$  and  $G(z, c_g)$ , respectively.  $P(c = c_r | x)$  and  $P(c = c_g | G(z, c_g))$  are the conditional probability distributions of class labels  $x$  and  $G(z, c_g)$ , respectively. During adversarial training, the discriminator is trained to maximize  $L_{class} + L_{source}$ , while the generator's goal is to maximize  $L_{class} - L_{source}$ .

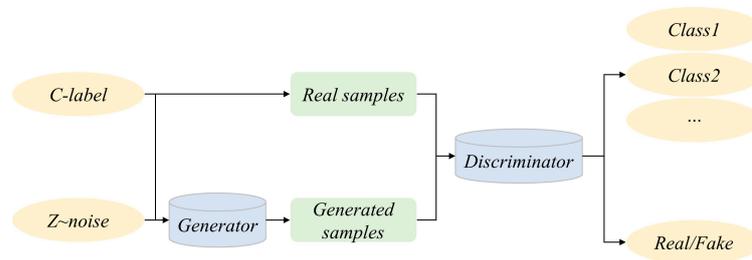


Figure 3. Architecture of ACGAN.

### 3. The Proposed SACGAN

SACGAN combines the respective features of SGAN and ACGAN and utilizes the semi-supervised learning mechanism of SGAN to improve generation, discrimination, and classification abilities by adversarial training using labeled and unlabeled data, thus transforming ACGAN from supervised learning to semi-supervised learning.

#### 3.1. Architecture of SACGAN

SACGAN refers to the structure of a Deep Convolutional Generative Adversarial Network (DCGAN) [23]. It is designed based on convolutional and deconvolutional networks; the architecture of SACGAN is shown in Figure 4. The generator mainly consists of embedding layers and two-dimensional deconvolutional layers; the input is a random noise vector with a category label vector, and the output is generated samples. The discriminator mainly consists of two-dimensional convolutional layers and fully connected layers, and the input includes labeled real samples, unlabeled real samples, and generated samples from the generator, and the output comprises discrimination results and classification results.

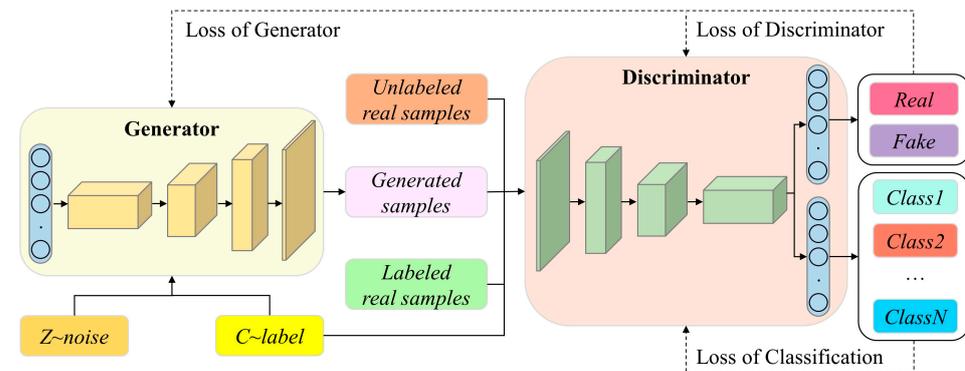


Figure 4. Architecture of SACGAN.

### 3.2. Structure of Generator

The generator’s input is a 200-dimensional Gaussian random noise vector  $z$  with a label vector  $c$ . The label vector is first embedded in the noise vector through an embedding layer, followed by deconvolution layers to generate fake samples whose sizes are  $64 \times 64$ . The generator contains five deconvolution layers, and the activation function of the first four layers is ReLU, which helps the generator achieve non-linear representations and makes the network easier for training. The activation function of the last layer is Tanh, which limits the output to  $[-1,1]$ , and BN is executed after each layer to speed up the training convergence and avoid overfitting. The specific structure is shown in Figure 5, and the detailed parameters are shown in Table 1.

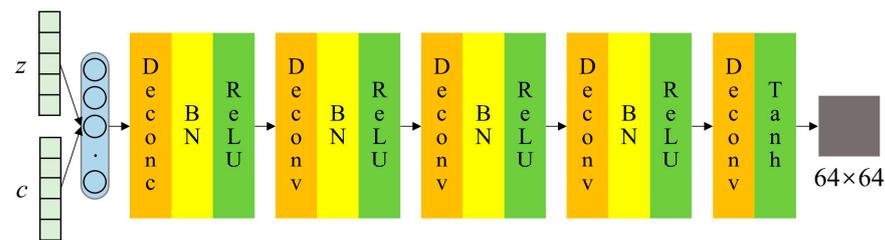


Figure 5. Structure of the generator.

Table 1. Parameters of the generator.

Layer Type	Kernel Size	Kernel Num	Strides	Output Size
Embedding	/	/	/	$200 \times 1 \times 1$
Deconv1	200	$3 \times 3$	$2 \times 2$	$200 \times 3 \times 3$
Deconv2	64	$3 \times 3$	$2 \times 2$	$64 \times 7 \times 7$
Deconv3	32	$3 \times 3$	$2 \times 2$	$32 \times 15 \times 15$
Deconv4	16	$3 \times 3$	$2 \times 2$	$16 \times 32 \times 32$
Deconv5	1	$4 \times 4$	$2 \times 2$	$1 \times 64 \times 64$

### 3.3. Structure of the Discriminator

The input of the discriminator is real samples and generated samples, and the output is a discrimination vector with a classification vector. The discriminator contains three convolutional layers, each with a leaky ReLU activation function, which allows the network to learn faster and prevent gradient disappearance. After each convolution, BN is executed first, followed by Dropout, which randomly drops neurons in the network to reduce the risk of overfitting and thus increase the generalization of the discriminator. The last layer contains two fully connected layers, using a Sigmoid function and a Softmax function as outputs for discriminating and classifying the input samples. The specific structure is shown in Figure 6, and the detailed parameters are shown in Table 2.

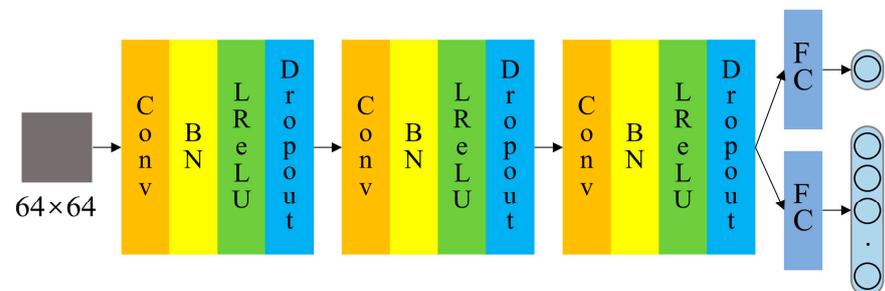


Figure 6. Structure of the discriminator.

**Table 2.** Parameters of the discriminator.

Layer Type	Kernel Size	Kernel Num	Strides	Padding	Output Size
Conv1	32	5 × 5	4 × 4	2 × 2	32 × 16 × 16
Conv2	64	5 × 5	2 × 2	2 × 2	64 × 8 × 8
Conv3	128	5 × 5	2 × 2	2 × 2	128 × 4 × 4
FC	/	/	/	/	1
	/	/	/	/	$K^1$

<sup>1</sup>  $K$  is the number of classification categories.

### 3.4. Loss Functions of SACGAN

SACGAN is constructed based on SGAN and ACGAN, so the loss functions of SACGAN combine the loss functions of SGAN and ACGAN. More specifically, the loss function of the generator of SACGAN is the same as that of ACGAN, and the loss function of the discriminator is referenced from SGAN, which consists of the supervised loss function of labeled data and the unsupervised loss function of unlabeled data. The specific definitions are as follows:

$$L_G = E_{z \sim P_{z(z)}} [\log P(c = c) | G(z, c)] - E_{z \sim P_{z(z)}} [\log(D(G(z, c)))] \quad (8)$$

$$L_D^U = E_{x \sim P_{r(x)}} [\log D(x)] \quad (9)$$

$$L_D^S = E_{x, y \sim P_{r(x, y)}} [\log D(x)] + E_{x, y \sim P_{r(x, y)}} [\log P(c = c | x)] + E_{z \sim P_{z(z)}} [\log(D(G(z, c)))] + E_{z \sim P_{z(z)}} [\log P(c = c) | G(z, c)] \quad (10)$$

$$L_D = \lambda_1 L_D^U + \lambda_2 L_D^S \quad (11)$$

where  $L_G$  and  $L_D$  represent the loss functions of the generator and the discriminator, respectively;  $L_D^U$  and  $L_D^S$  represent the supervised loss and the unsupervised loss of the discriminator, and  $\lambda_1, \lambda_2$  are the ratio factors;  $P_{r(x)}$  and  $P_{r(x, y)}$  are the prior distributions of labeled real samples and unlabeled real samples, respectively;  $c$  are labels of labeled real samples; and  $P(c = c | x)$  and  $P(c = c | G(z, c))$  are the conditional probability distributions of class labels of labeled real samples and generated samples, respectively.

## 4. Fault Diagnosis Based on STFT-SACGAN

### 4.1. Data Processing

STFT is a joint time-varying time-frequency analysis method for non-stationary signals that converts one-dimensional time-domain signals into two-dimensional time-frequency images for CNN processing: feature spectra containing both time-domain and frequency-domain information. It uses a fixed-length window function to translate over the time-domain signal, intercepting it and performing a Fourier transform to obtain a local set of spectra for each period [24]. Therefore, STFT is a 2D function of time and frequency, whose calculation formula is defined as:

$$STFT\{x(\tau)\}(\tau, w) = \int_{-\infty}^{\infty} x(t)g(t - \tau)e^{-j\omega t} dt \quad (12)$$

where  $x(t)$  represents the one-dimensional time-domain signal,  $t$  and  $w$  are the time and the frequency, respectively, and  $g(t - \tau)$  denotes a window function whose center is at the time  $\tau$ .

The time and frequency resolutions of the spectrum obtained by STFT depend on the length of the window function, with longer window lengths yielding a lower time resolution and a higher frequency resolution. Therefore, the window length should be

chosen reasonably according to the signal to be processed for better analysis [25]. The time and frequency resolutions are calculated as follows:

$$F = \begin{cases} N_w/2 + 1, & \text{if } N_w \text{ is even} \\ (N_w/1)/2, & \text{if } N_w \text{ is odd} \end{cases} \quad (13)$$

$$T = \left\lfloor \frac{N_x - N_o}{N_w - N_o} \right\rfloor \quad (14)$$

where  $N_x$  is the length of the signal to be processed,  $N_w$  is the length of the window function,  $N_o$  is the overlapped length during the translation of the window function, and  $\lfloor \cdot \rfloor$  represents the operation of rounding down.

This paper converts the original one-dimensional vibration signals of bearings into two-dimensional time-frequency images by STFT, which carry richer information and characterize a more complex structural distribution. According to Equations (13) and (14), the time-frequency image after STFT is a two-dimensional matrix of dimension  $F \times T$ . Let  $X$  denote the obtained matrix, and to speed up the convergence of the training process, each element of the matrix  $X$  is normalized into the interval  $[-1,1]$  according to Equation (16).

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1T} \\ \vdots & \vdots & \vdots \\ x_{F1} & \cdots & x_{FT} \end{bmatrix} \quad (15)$$

$$x_i^{normalized} = -1 + 2 \times \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}, i = 1, 2, \dots, N \quad (16)$$

#### 4.2. Fault Diagnosis Flow Based on STFT-SACGAN

A fault diagnosis method based on STFT-SACGAN is proposed, using the ability of STFT to process non-stationary signals. The original one-dimensional vibration signals of bearings are converted into two-dimensional time-frequency images by STFT as the input of SACGAN. The multi-mode fault sample generation ability of the generator and the recognition ability of the discriminator are improved by adversarial training. The flow of the proposed method is shown in Figure 7, and the main steps are as follows:

- (1) Vibration signals of bearings with various fault modes are collected, converted into time-frequency images by STFT, and normalized into the interval  $[-1,1]$ ;
- (2) The noise vector  $z$  and the label vector  $c$  are input to the generator to obtain generated samples;
- (3) Labeled real samples, unlabeled real samples, and generated samples from the generator are fed into the discriminator to obtain discrimination and classification results.
- (4) Calculate the loss of the generator and the discriminator;
- (5) Fix the generator's weight parameters and optimize the discriminator's weight parameters;
- (6) Fix the discriminator's weight parameters and optimize the generator's weight parameters;
- (7) Repeat steps 2–6 until the number of iterations is satisfied;
- (8) Save the trained model, use the generator to generate multi-mode fault samples, and use the discriminator for fault diagnosis of test signals.

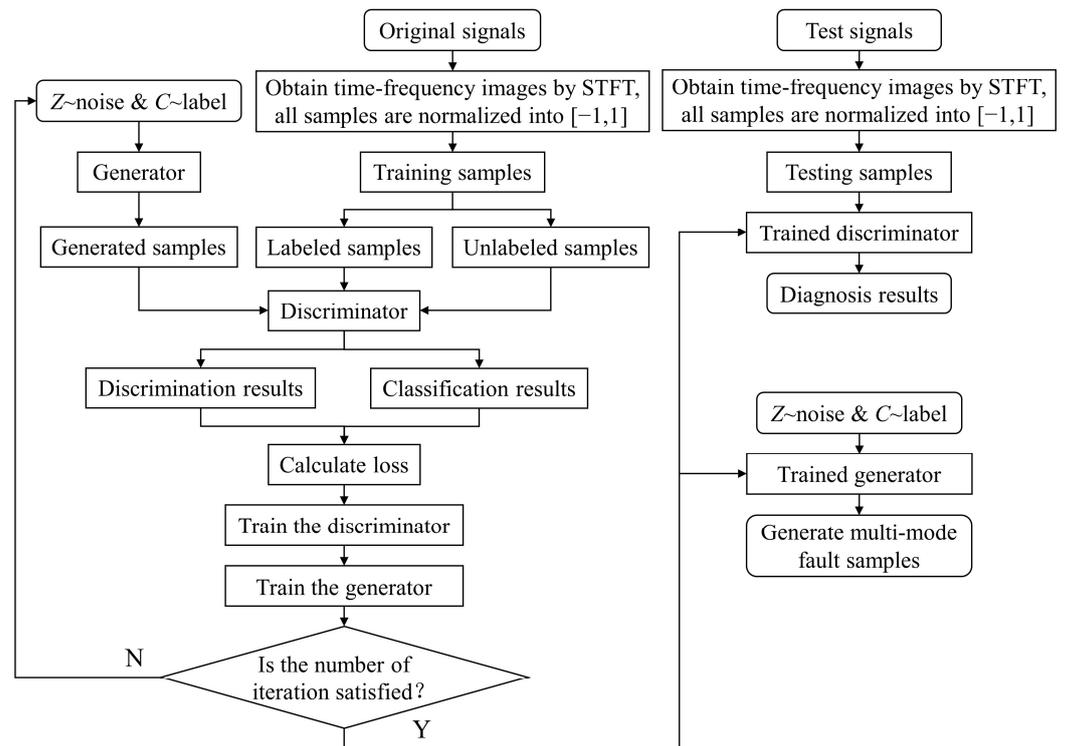


Figure 7. Fault diagnosis flowchart based on STFT-SACGAN.

### 5. Case Study

To demonstrate the effectiveness and generalization of the proposed method, validation and analysis are carried out on two bearing datasets. In those two case studies, the computer is a Core i5-9300H CPU @ 2.40 GHz with 16 GB of Ram and works in the Windows 64-bit operating system, and a GPU(GTX1650) with 4 GB of memory is added to improve the training speed. The programming language is Python 3.8.13, and the deep learning framework is Pytorch 1.10.1. During the training process, the generator and the discriminator of SACGAN use the Adam optimization algorithm [26], and the learning rate is 0.0005. To reduce the fluctuation during training, the exponential decay rates  $\beta_1$  and  $\beta_2$  are set to 0.5 and 0.999, respectively [27]. The batch size is  $K \times 10$  ( $K$  is the number of category labels), the dropout rate of the discriminator is 0.25, and the number of iterations is 200. In addition, the ratio factors of the loss function Equation (11) are set to 0.5 and 0.5, respectively.

#### 5.1. Case 1 Multi-Mode Data Generation and Fault Diagnosis of Bearings

##### 5.1.1. Introduction and Preprocessing of Bearing Data

In case 1, the Case Western Reserve University (CWRU)-bearing dataset is used to verify the effectiveness of the proposed method, and the test rig is shown in Figure 8 [28]. Specifically, ten kinds of bearing vibration data under 1797 rpm and 12 kHz sampling frequency are used for analysis. Bearing states include one normal state and nine faulty states, including three fault locations and three fault sizes, as listed in Table 3. The vibration signals from each state are randomly sampled 300 times over 1024 points as available samples, of which 100 samples from each state are used as testing samples. The remaining samples are divided into labeled and unlabeled training samples according to different ratios.

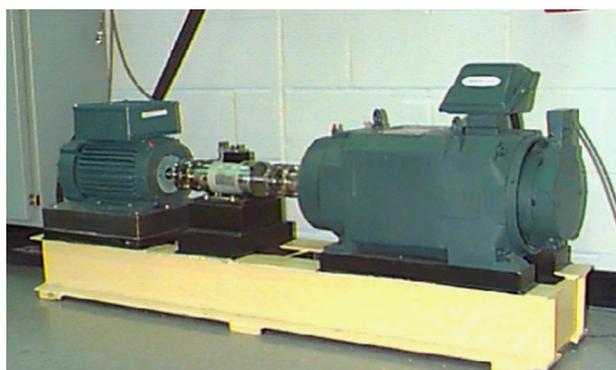


Figure 8. CWRU-bearing experimental platform.

Table 3. Details of CWRU-bearing samples.

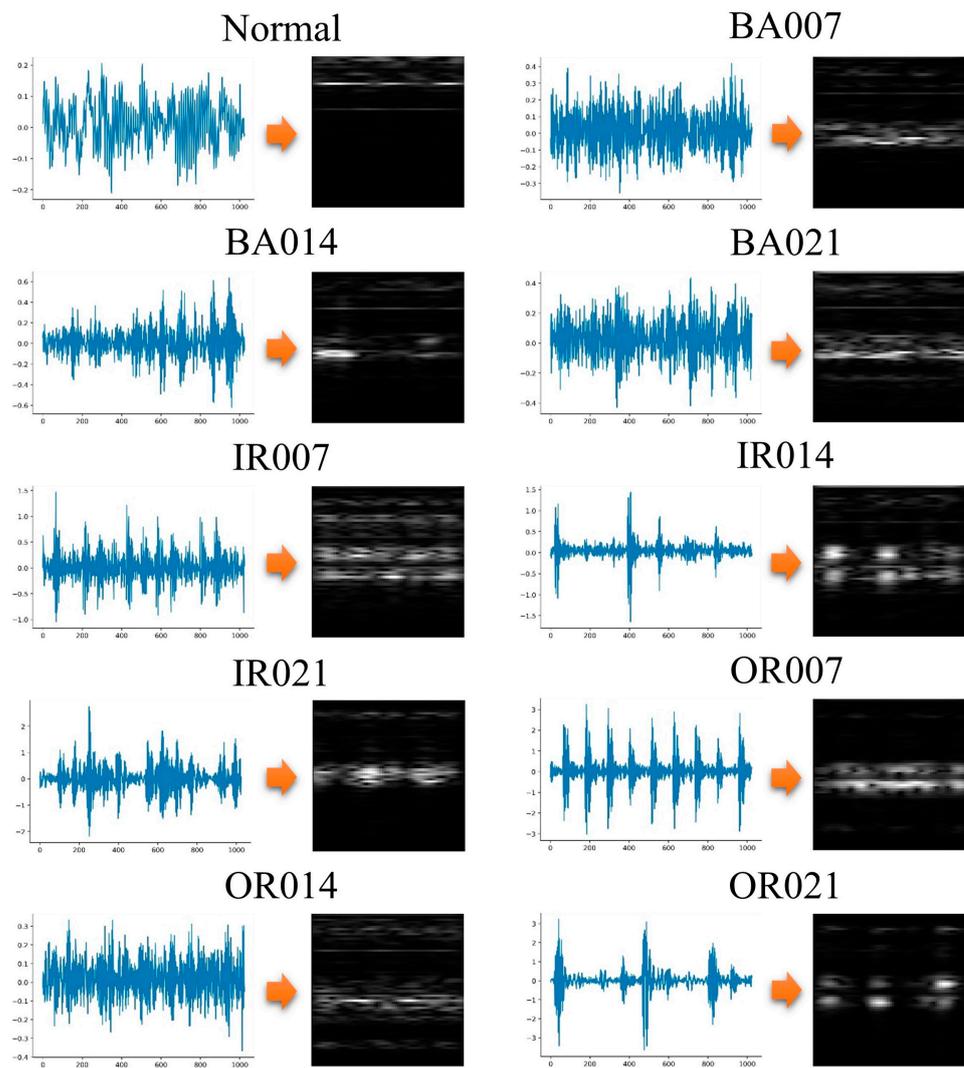
Fault Type	Fault Diameter	Class Labels
Normal	-	Normal
Rolling ball	0.007 inch	BA007
Rolling ball	0.014 inch	BA014
Rolling ball	0.021 inch	BA021
Inner race	0.007 inch	IR007
Inner race	0.014 inch	IR014
Inner race	0.021 inch	IR021
Outer race	0.007 inch	OR007
Outer race	0.014 inch	OR014
Outer race	0.021 inch	OR021

All samples are converted into time-frequency images by STFT. According to Section 3.1, the window function is Hanning window; the window function length  $N_w$  and the overlap length  $N_o$  are set to 256 and 250, respectively. The normalized time-frequency matrix is geometrically processed into a  $64 \times 64$  square matrix as the input of SACGAN to accommodate the training of SACGAN and save computation time. The original time-domain vibration signals and the corresponding two-dimensional time-frequency images obtained by conversion are shown in Figure 9. It can be seen that characteristic differences between different states are more significant than the original one-dimensional time-domain vibration signals after conversion.

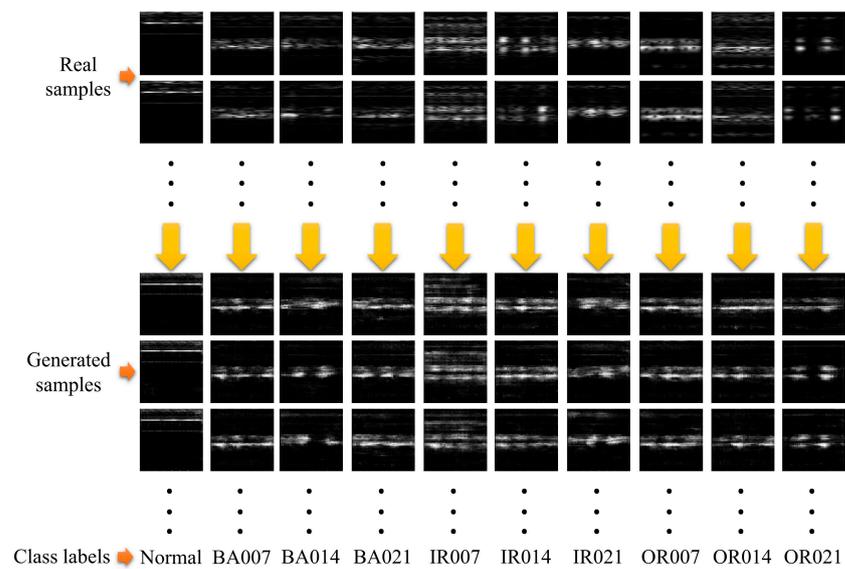
#### 5.1.2. Quality Evaluation and Comparison of Multi-Mode Generated Samples

The sample generation ability of the generator is usually verified by evaluating the quality of the generated samples. In this case study, the SAGAN is trained using training samples with a labeling ratio of 0.2 (40 labeled samples and 160 unlabeled samples of each class). The diagrammatic sketch of the real samples and the generated samples is shown in Figure 10. The generated multi-mode samples are highly similar to the real samples.

For a more objective evaluation, the structural similarity index measure (SSIM) is used to quantitatively evaluate the quality of the generated multi-mode samples. The SSIM aims to measure the similarity of two images from brightness, contrast, and structure. The range of SSIM values is the interval  $[-1,1]$ , and larger SSIM values indicate higher similarity between two images [29]. In contrast, the labeled training samples are also used to train the ACGAN, where the parameter structure of the ACGAN is the same as that of the SACGAN. Five pairs of the real samples and the generated samples of each class are randomly selected, and the average SSIM values of the ACGAN and the SACGAN are given in Figure 11. The results show that the average SSIM values of each class of the SACGAN are always larger than those of the ACGAN, meaning that the samples generated by the SACGAN are closer to the real samples.



**Figure 9.** Time-domain vibration signals for CWRU data and the corresponding time-frequency images obtained by conversion.



**Figure 10.** The diagrammatic sketch of generating samples for CWRU data.

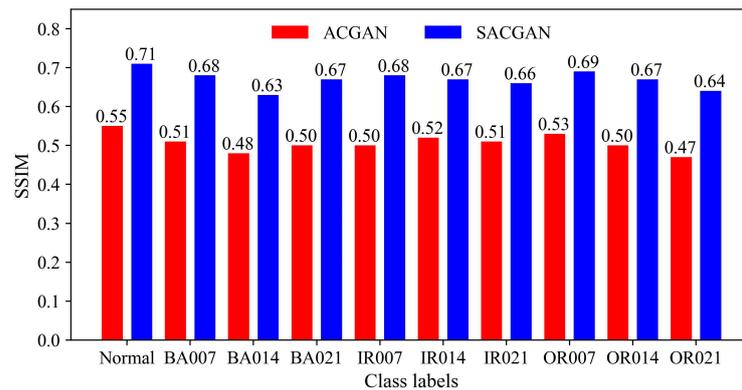


Figure 11. Quality evaluation of generated samples based on the SSIM for CWRU data.

### 5.1.3. Fault Recognition with Different Label Ratios

Each class’s training samples are divided into labeled and unlabeled samples in different ratios to train the SACGAN. The labeled samples are also used to train the ACGAN and the CNN model for comparison, where the parameter structure of the ACGAN is the same as the SACGAN. It is worth mentioning that to verify the effectiveness of the discriminator’s structure, the CNN model here is the discriminator of SACGAN that can perform supervised learning. Each group of comparative experiments is repeated ten times, and the average recognition accuracy of the test samples at different label ratios is given in Figure 12. The results show that all three models have high recognition accuracy, indicating the effectiveness of the discriminator structure; the ACGAN and the SACGAN effectively improve the recognition ability of the discriminator through adversarial learning; limiting the number of labeled training samples has a greater impact on the fault recognition ability of the ACGAN and the CNN and the SACGAN can weaken the influence of decreasing labeled training samples and has the highest recognition accuracy under different label ratios.

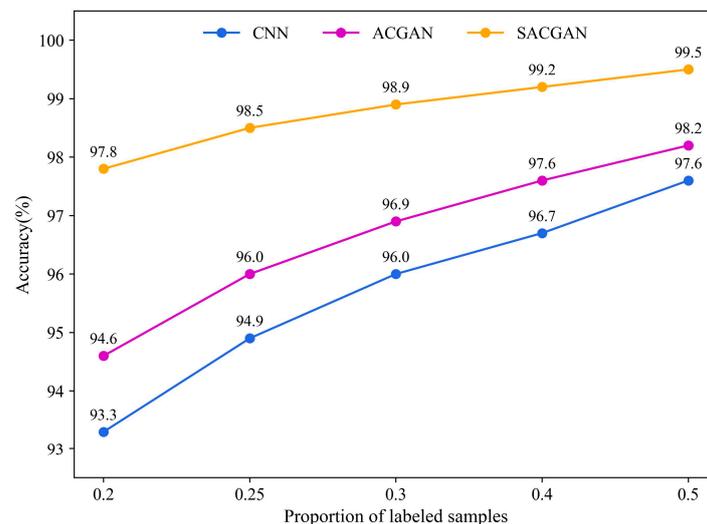
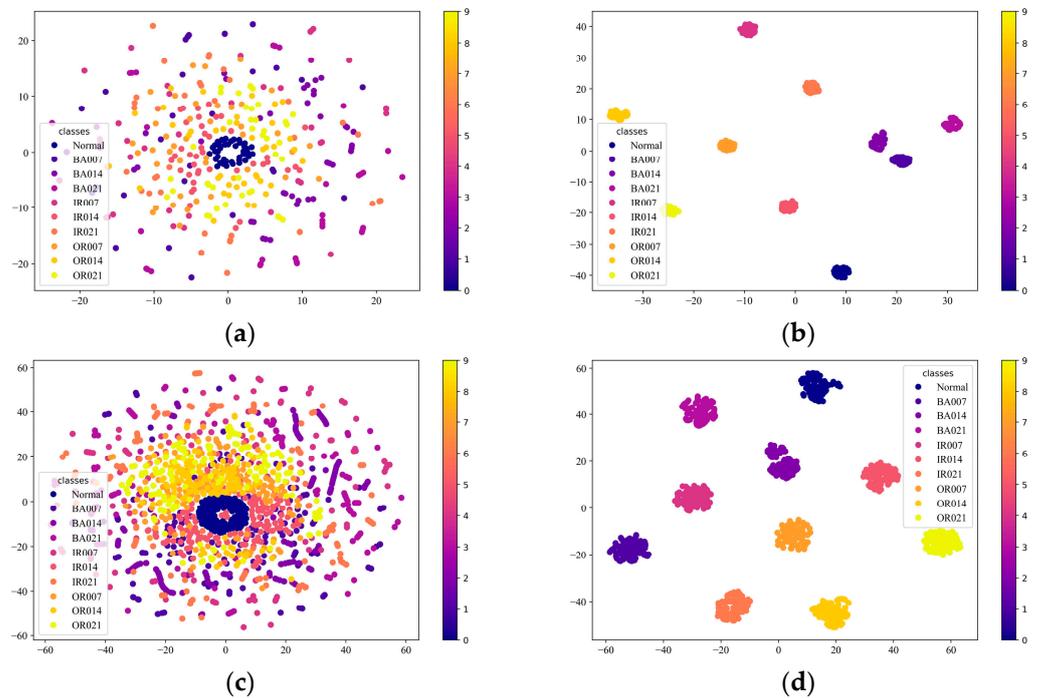


Figure 12. Recognition results with different label ratios for CWRU data.

The recognition ability of the model is closely related to its ability to extract effective features, and to observe the distribution of features extracted by the model, the t-SNE (t-Distributed Stochastic Neighbor Embedding) algorithm is used to visualize the distribution of extracted features [30]. The changes in the distribution of labeled training and testing samples with a labeling ratio of 0.2 are shown in Figure 13, which shows that the SACGAN can significantly improve the distribution of samples from a different class, reduce the

intra-class distance, and increase the inter-class distance, making it easier to recognize the differences between different classes.

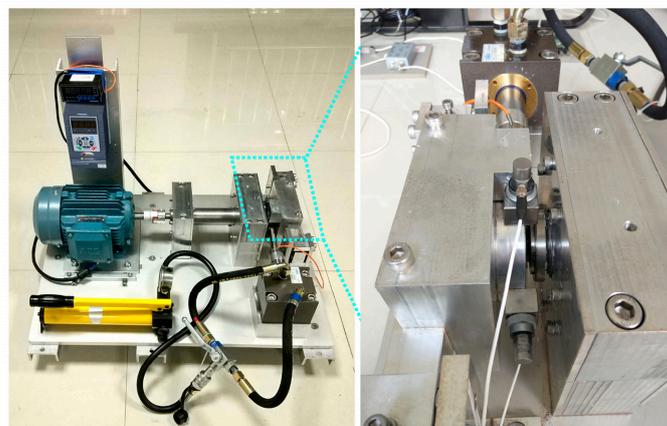


**Figure 13.** Visualization of feature distribution for CWRU data. (a) Labeled samples before training; (b) labeled samples after training; (c) test samples before testing; (d) test samples after testing.

## 5.2. Case 2 Multi-Mode Data Generation and Fault Diagnosis of Bearings

### 5.2.1. Introduction and Preprocessing of Bearing Data

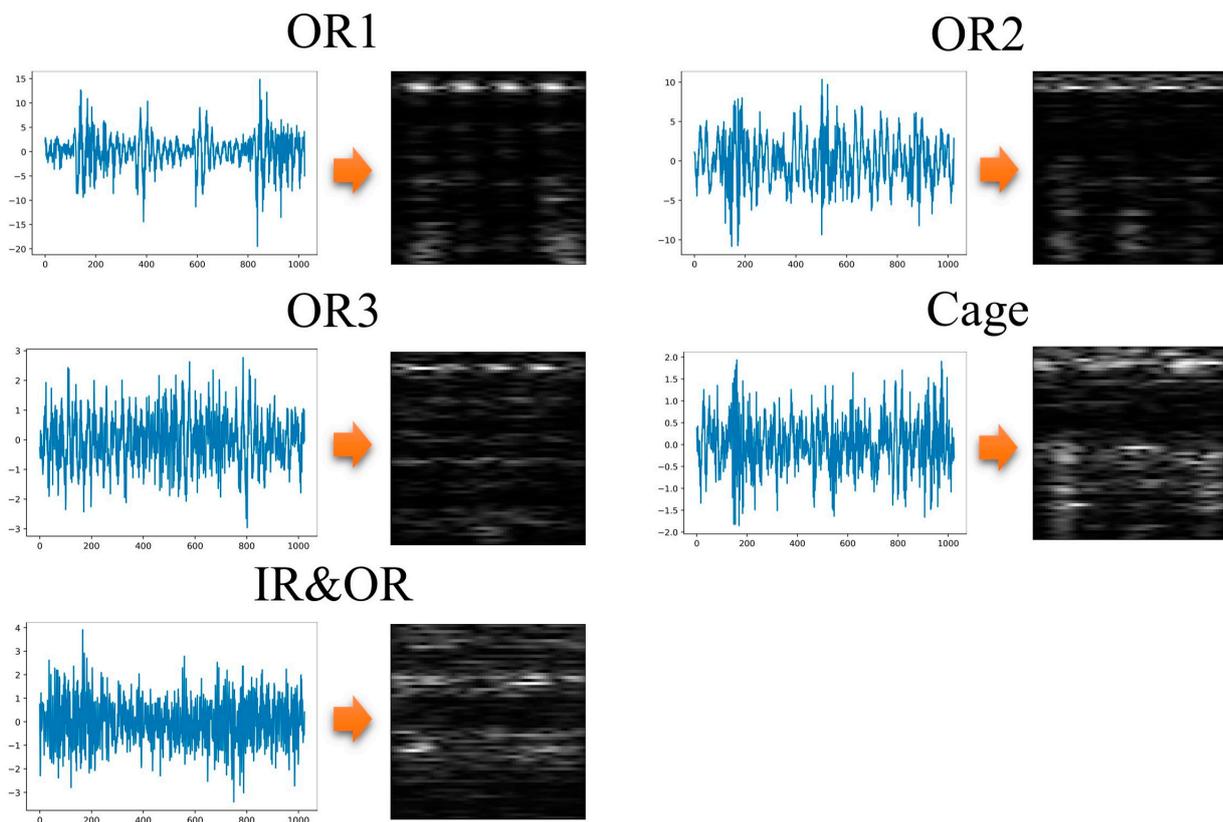
In case 2, Xi’an Jiaotong University and the Changxing Sumyoung Technology Co. (XJTU-SY) bearing dataset are used to verify the generalization ability of the proposed method, and the test rig is shown in Figure 14 [31], with a sampling frequency of 2.56 kHz. Five kinds of bearing vibration data with a speed of 2100 rpm and a radial force of 12 kN are selected for analysis, and the bearing states and class labels are shown in Table 4. Similar to Case 1, 1024 consecutive points are randomly sampled 300 times from the vibration signals of each state as available samples, and two-dimensional time-frequency images obtained by conversion are shown in Figure 15.



**Figure 14.** XJTU-SY-bearing experimental platform.

**Table 4.** Details of XJTU-SY-bearing samples.

Fault Type	Class Labels
Outer ring	OR1
Outer ring	OR2
Outer ring	OR3
Cage	Cage
Inner ring and outer ring	IR&OR



**Figure 15.** Time-domain vibration signals for XJTU-SY data and the corresponding time-frequency images obtained by conversion.

### 5.2.2. Multi-Mode Fault Sample Generation and Fault Recognition in Case 2

Similar to Case 1, the SAGAN is trained using training samples with a labeling ratio of 0.2 (40 labeled samples and 160 unlabeled samples of each class). Figure 16 shows the diagrammatic sketch of the real samples and the generated samples. The SSIM is also used to quantitatively evaluate the quality of the generated multi-mode samples. As shown in Figure 17, the SSIM values of each class of the ACGAN are always smaller than those of the SACGAN. The results further demonstrate the validity and superiority of the SACGAN in terms of generation. The comparative fault recognition experiments are the same as in Case 1, and the recognition accuracies of the test samples at different label ratios are shown in Figure 18. For feature-level evaluation, the results of feature distribution visualization are shown in Figure 19. The above results proved that the SACGAN has strong robustness and generalization ability and can be applied to different objects and operating environments.

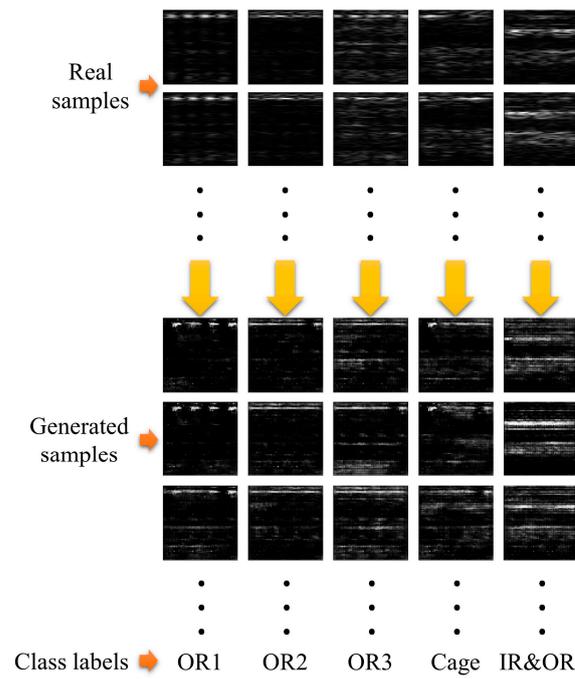


Figure 16. The diagrammatic sketch of generating samples for XJTU-SY data.

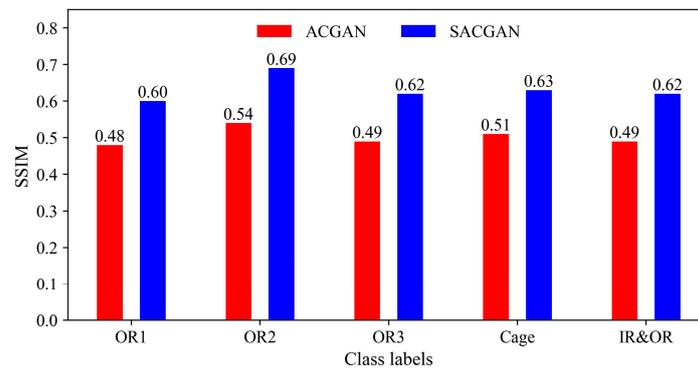


Figure 17. Quality evaluation of generated samples based on the SSIM for XJTU-SY data.

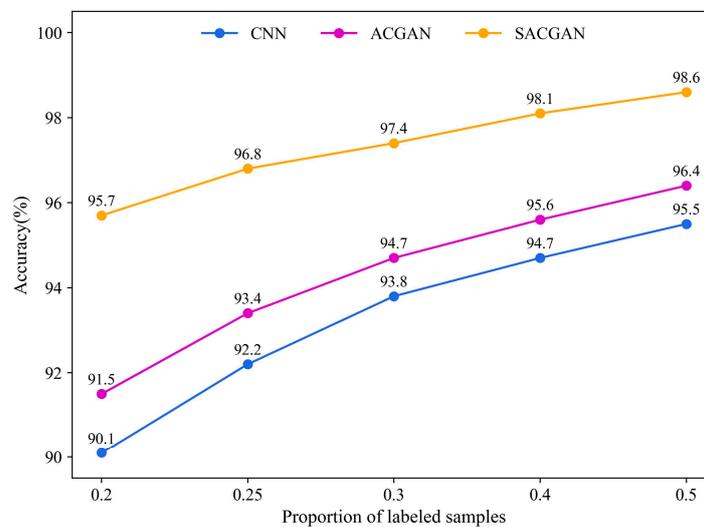
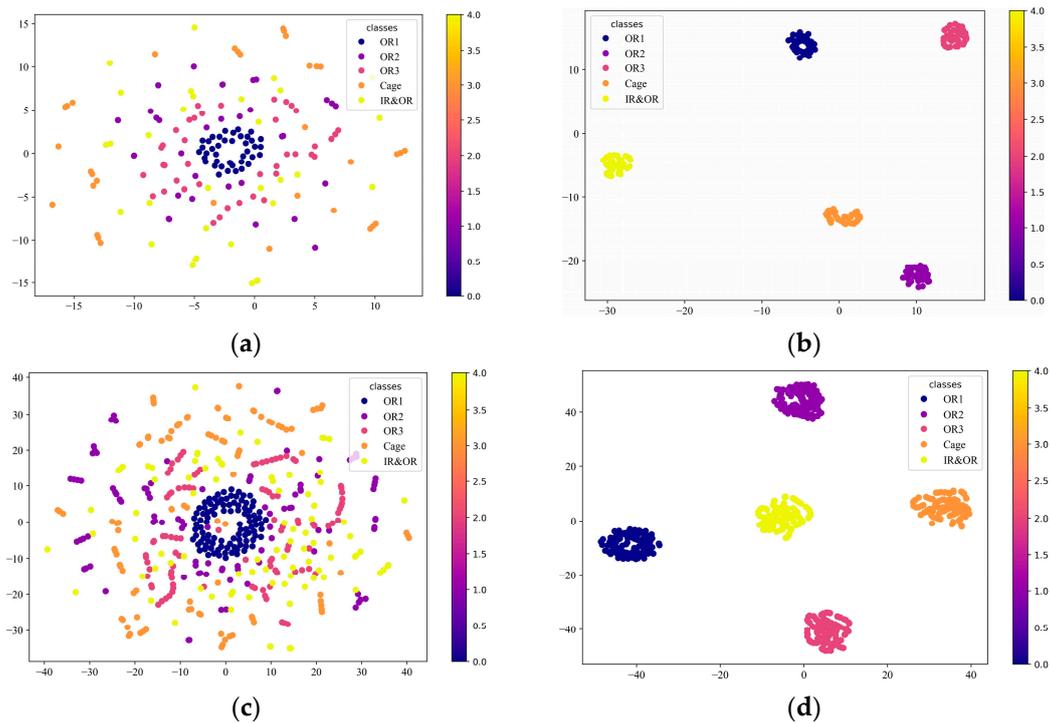


Figure 18. Recognition results with different label ratios for XJTU-SY data.



**Figure 19.** Visualization of feature distribution for XJTU-SY data. (a) Labeled samples before training; (b) labeled samples after training; (c) test samples before testing; (d) test samples after testing.

## 6. Conclusions

To address the problem of scarcity of labeled data in intelligent fault diagnosis of bearings, from the field of GAN, the SACGAN based on SGAN and ACGAN is constructed to improve the multi-mode fault sample generation ability of the generator and the recognition ability of the discriminator by using a small amount of labeled data with a large amount of unlabeled data for adversarial learning. Using the ability of STFT to process non-stationary signals, a fault diagnosis method based on STFT-SACGAN is proposed, and the original one-dimensional vibration signals of bearings are converted into two-dimensional time-frequency images by STFT as the input of the SACGAN. The effectiveness of the proposed method is verified on the CWRU bearing dataset, and the results show that the SACGAN can generate high-quality multi-mode fault samples and has excellent fault recognition ability. The proposed method is verified to have strong generalization and stability on the XJTU-SY bearing dataset.

Although the proposed method can perform multi-mode fault sample generation and fault diagnosis tasks well, it only involves data from a single source of information. In the future, we will explore how to use data from multi-source information fusion to generate high-quality, multi-source information fusion fault samples and further optimize the model to improve fault diagnosis efficiency. In addition, the study of fault diagnosis with limited labeled data is more suitable to current industrial practical scenarios. It can significantly reduce the cost of fault diagnosis, which is worth further exploration.

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

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant No. 52277055).

**Data Availability Statement:** The data used to support this study are available on the websites <https://engineering.case.edu/bearingdatacenter/download-data-file> and <https://biaowang.tech/xjtu-sy-bearing-datasets/>, accessed on 22 March 2023.

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

## References

1. Lei, Y.G.; Yang, B.; Jiang, X.W.; Jia, F.; Li, N.P.; Nandi, A. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587–106625. [CrossRef]
2. Sun, Y.J.; Wang, J.; Wang, X.H. Fault diagnosis of mechanical equipment in high energy consumption industries in China: A review. *Mech. Syst. Signal Process.* **2023**, *186*, 109833–109865. [CrossRef]
3. Zhao, Z.B.; Li, T.F.; Wu, J.Y.; Sun, C.; Wang, S.B.; Yan, R.Q.; Chen, X.F. Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study. *ISA Trans.* **2020**, *107*, 224–255. [CrossRef] [PubMed]
4. Cen, J.; Yang, Z.H.; Liu, X.; Xiong, J.B.; Chen, H.H. A Review of Data-Driven Machinery Fault Diagnosis Using Machine Learning Algorithms. *J. Vib. Eng. Technol.* **2022**, *10*, 2481–2507. [CrossRef]
5. Jalayer, M.; Orsenigo, C.; Vercellis, C. Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms. *Comput. Ind.* **2021**, *125*, 103378–103393. [CrossRef]
6. Przystupa, K.; Ambrozkiewicz, B.; Litak, G. Diagnostics of Transient States in Hydraulic Pump System with Short Time Fourier Transform. *Adv. Sci. Technol. Res. J.* **2020**, *14*, 178–183. [CrossRef]
7. Lepicka, M.; Górski, G.; Gradzka-Dahlke, M.; Litak, G.; Ambrozkiewicz, B. Analysis of tribological behaviour of titanium nitride-coated stainless steel with the use of wavelet-based methods. *Arch. Appl. Mech.* **2021**, *91*, 4475–4483. [CrossRef]
8. Kibrete, F.; Woldmichael, D.E. Applications of Artificial Intelligence for Fault Diagnosis of Rotating Machines: A Review. *Inst. Comput. Sci. Soc. Inform. Telecommun. Eng.* **2023**, *455*, 41–62.
9. Zhu, Z.Q.; Lei, Y.B.; Qi, G.Q.; Chai, Y.; Mazur, N.; An, Y.; Huang, X.H. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery. *Measurement* **2023**, *206*, 112346–112369. [CrossRef]
10. Zhao, Z.B.; Wu, J.Y.; Li, T.F.; Sun, C.; Yan, R.Q.; Chen, X.F. Challenges and Opportunities of AI-Enabled Monitoring, Diagnosis & Prognosis: A Review. *Chin. J. Mech. Eng.* **2021**, *34*, 16–44.
11. Zhang, T.C.; Chen, J.L.; Li, F.D.; Zhang, K.Y.; Lv, H.X.; He, S.L.; Xu, E.Y. Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions. *ISA Trans.* **2021**, *119*, 152–171. [PubMed]
12. Zhao, Z.B.; Zhang, Q.Y.; Yu, X.L.; Sun, C.; Wang, S.B.; Yan, R.Q.; Chen, X.F. Applications of Unsupervised Deep Transfer Learning to Intelligent Fault Diagnosis: A Survey and Comparative Study. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3525828–3525855. [CrossRef]
13. Pan, T.Y.; Chen, J.L.; Zhang, T.C.; Liu, S.; He, S.L.; Lv, H.X. Generative adversarial network in mechanical fault diagnosis under small sample: A systematic review on applications and future perspectives. *ISA Trans.* **2021**, *128*, 1–10. [CrossRef] [PubMed]
14. Pang, X.Y.; Wei, Z.H.; Tong, Y. Fault Diagnosis Method of Gear Based on SCGAN Network. *J. Vib. Meas. Diagn.* **2022**, *42*, 358–364.
15. Xing, X.S.; Guo, W. Intelligent diagnosis method for bearing with few labelled samples based on an improved semi-supervised learning-based generative adversarial network. *J. Vib. Shock* **2022**, *41*, 184–192.
16. Yang, Q.; Zhang, J.Y.; Wu, D.S.; Liu, Y.P. Fault Diagnosis for Rolling Bearings Based on Two-Dimensional Image and Switchable Normalization SGAN Network. *Bearing* **2021**, *8*, 39–46.
17. Lu, J.L.; Zhang, X.G.; Zhang, W.; Guo, L.Y.; Wen, R.T. Fault Diagnosis of Main Bearing of Wind Turbine Based on Improved Auxiliary Classifier Generative Adversarial Network. *Autom. Electr. Power Syst.* **2021**, *45*, 148–154.
18. Li, D.D.; Liu, Y.H.; Zhao, Y.; Zhao, Y. Fault Diagnosis Method of Wind Turbine Planetary Gearbox Based on Improved Generative Adversarial Network. *Proc. CSEE* **2021**, *41*, 7496–7507.
19. Li, W.; Zhong, X.; Shao, H.D.; Cai, B.P.; Yang, X.K. Multi-mode data augmentation and fault diagnosis of rotating machinery using modified ACGAN designed with new framework. *Adv. Eng. Inform.* **2022**, *52*, 101552–101567. [CrossRef]
20. He, W.P.; Chen, J.; Zhou, Y.; Liu, X.; Chen, B.Q.; Guo, B.L. An Intelligent Machinery Fault Diagnosis Method Based on GAN and Transfer Learning under Variable Working Conditions. *Sensors* **2022**, *22*, 9175. [CrossRef]
21. He, Y.; Tang, H.S.; Ren, Y.; Kumar, A. A semi-supervised fault diagnosis method for axial piston pump bearings based on DCGAN. *Meas. Sci. Technol.* **2021**, *32*, 125104–125122. [CrossRef]
22. Meng, Z.; Li, Q.; Sun, D.Y.; Cao, W.; Fan, F.J. An Intelligent Fault Diagnosis Method of Small Sample Bearing Based on Improved Auxiliary Classification Generative Adversarial Network. *IEEE Sens. J.* **2022**, *22*, 19543–19555. [CrossRef]
23. Gao, Y.D.; Piltan, F.; Kim, J.M. A Novel Image-Based Diagnosis Method Using Improved DCGAN for Rotating Machinery. *Sensors* **2022**, *22*, 7534. [CrossRef] [PubMed]
24. Li, H.; Zhang, Q.; Qin, X.R.; Sun, Y.T. Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network. *J. Vib. Shock* **2018**, *37*, 124–131.
25. Tao, H.F.; Wang, P.; Chen, Y.Y.; Stojanovic, V.; Yang, H.Z. An unsupervised fault diagnosis method for rolling bearing using STFT and generative neural networks. *J. Frankl. Inst.* **2020**, *357*, 7286–7307. [CrossRef]
26. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
27. Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv* **2015**, arXiv:1511.06434.

28. Smith, W.; Randall, R. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64–65*, 100–131. [[CrossRef](#)]
29. Setiadi, D. PSNR vs. SSIM: Imperceptibility quality assessment for image steganography. *Multimed. Tools Appl.* **2021**, *80*, 8423–8444. [[CrossRef](#)]
30. Pezzotti, N.; Lelieveldt, B.; Maaten, L.; Höllt, T.; Eisemann, E.; Vilanova, A. Approximated and User Steerable tSNE for Progressive Visual Analytics. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 1739–1752. [[CrossRef](#)]
31. Wang, B.; Lei, Y.G.; Li, N.P.; Li, N.B. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [[CrossRef](#)]

**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.