

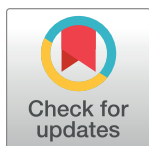
## RESEARCH ARTICLE

# A fault diagnosis method based on Auxiliary Classifier Generative Adversarial Network for rolling bearing

Chunming Wu<sup>1,2</sup>, Zhou Zeng<sup>2\*</sup>

**1** Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin, China, **2** Department of Electrical Engineering, Northeast Electric Power University, Jilin, China

\* [1143879389@qq.com](mailto:1143879389@qq.com)



## OPEN ACCESS

**Citation:** Wu C, Zeng Z (2021) A fault diagnosis method based on Auxiliary Classifier Generative Adversarial Network for rolling bearing. PLoS ONE 16(3): e0246905. <https://doi.org/10.1371/journal.pone.0246905>

**Editor:** Tao Song, Polytechnical Universidad de Madrid, SPAIN

**Received:** October 14, 2020

**Accepted:** January 27, 2021

**Published:** March 1, 2021

**Copyright:** © 2021 Wu, Zeng. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are available from the Case Western Reserve University Bearing Data Centre website (<http://csegroups.case.edu/bearingdatacenter>). Other data are uploaded as [Supporting Information](#).

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Rolling bearing fault diagnosis is one of the challenging tasks and hot research topics in the condition monitoring and fault diagnosis of rotating machinery. However, in practical engineering applications, the working conditions of rotating machinery are various, and it is difficult to extract the effective features of early fault due to the vibration signal accompanied by high background noise pollution, and there are only a small number of fault samples for fault diagnosis, which leads to the significant decline of diagnostic performance. In order to solve above problems, by combining Auxiliary Classifier Generative Adversarial Network (ACGAN) and Stacked Denoising Auto Encoder (SDAE), a novel method is proposed for fault diagnosis. Among them, during the process of training the ACGAN-SDAE, the generator and discriminator are alternately optimized through the adversarial learning mechanism, which makes the model have significant diagnostic accuracy and generalization ability. The experimental results show that our proposed ACGAN-SDAE can maintain a high diagnosis accuracy under small fault samples, and have the best adaptation performance across different load domains and better anti-noise performance.

## 1 Introduction

As a common part of rotating machinery, rolling bearing may cause great economic loss if it breaks down in the working process [1]. Therefore, it is of great significance for the normal operation of the machine to diagnose the rolling bearing effectively [2]. At present, vibration signal analysis is one of the widely employed and effective techniques for machinery fault diagnosis and health monitoring [3]. In essence, machinery fault diagnosis can be regarded as a pattern recognition problem, which includes data acquisition, feature extraction and fault classification. The diagnostic performance largely depends on the effectiveness of feature extraction and classification methods. Thus, traditional fault diagnosis methods based on vibration are generally: signal processing methods based on time domain, frequency domain and time-frequency domain. These methods include time domain statistics, short-time Fourier transform [4], wavelet transform [5], Empirical Mode Decomposition (EMD) [6], Hilbert-Huang

Transform (HHT) [7] and other variants [8–11]. Then, these extracted features are fed into some shallow machine learning algorithms, including Artificial Neural Network (ANN) [12], Support Vector Machine (SVM) [13], cluster analysis [14], etc. However, the fault feature representation extracted from the above methods is usually designed manually and requires a lot of professional knowledge and manpower [15]. At the same time, most of the methods are limited to the domain and can not be well extended to other new fault diagnosis fields. Instead, deep learning can effectively solve the above problems by modeling the high-level representation of data and predicting/classifying patterns through a layered architecture of multiple non-linear processing units [16].

Since Hinton et al. [17] proposed unsupervised layer by layer training combined with supervised fine-tuning method, deep learning theory has become a hot spot in the field of machine learning and artificial intelligence, and has made brilliant achievements in computer vision, speech recognition and other fields. Some experts and scholars have also applied for deep learning theory to the field of mechanical fault diagnosis. Chen et al. [18] proposed a bearing fault diagnosis method based on Deep Belief Network (DBN). By using the advantages of DBN automatic feature extraction and classification, the original vibration signal is directly studied and stratified training, and the fault diagnosis results are automatically given. Considering the multi-scale characteristics inherent in vibration signals of a gearbox, Jiang et al. [19] proposed a new multi-scale convolutional neural network (MSCNN) architecture that can perform multi-scale features extraction and classification simultaneously. Due to the time-series characteristics of wind turbine vibration signals, Lei et al. [20] adopted the Long Short-Term Memory (LSTM) model to realize the end-to-end fault diagnosis of wind turbines.

Although the above fault methods can achieve good results in some specific aspects, there is still room for improvement: (i) Most of the improvements in the traditional depth model to make the model satisfy better diagnostic accuracy for specific data sets, but it may not be suitable for practical fault diagnosis tasks. (ii) In fact, the machinery generally runs under normal working conditions for a long time, and the sensor can collect enough positive samples, while the negative samples collected under fault conditions are seriously unbalanced compared with the positive samples. Therefore, the diagnosis performance of the unbalanced data set under small sample conditions is very poor. (iii) At the same time, considering the cross domain adaptive problem with the actual variable load conditions and the influence of high background noise pollution, the diagnostic performance of the model will be further deteriorated.

Goodfellow et al. [21] proposed a Generative Adversarial Network (GAN) in 2014. Due to its powerful performance, GAN has made great achievements in the field of image processing. Radford et al. [22] proposed a novel Deep Convolution Generative Adversarial Network (DCGAN) based on the GAN, which is stable in the training process and can generate high-quality images. The application of GAN to the field of mechanical fault diagnosis provides a new perspective. Shao et al. [23] applied GAN as a data set enhancement technique to fault diagnosis for small fault sample sizes, and achieved good results. Han et al. [24] proposed a novel deep adversarial convolutional neural network (DACNN) framework. By introducing adversarial learning into CNN, it helps to make feature representation more robust and enhance the generalization ability of the training model. Zhao et al. [25] proposed an improved Wasserstein GAN fault diagnosis method based on K-means and applied it to aero-engine fault diagnosis. Wasserstein GAN with gradient penalty was used to make the model converge process faster.

Aiming at the problems that the data unbalances caused by small fault sample size, cross domain adaptive problem under variable load and the influence of high background noise pollution in the fault diagnosis of rolling bearing, we proposed a novel fault diagnosis method of rolling bearing which combines ACGAN and SDAE. Different from the traditional GAN, this

paper introduces the variant structure ACGAN with auxiliary classification labels. In detail, we use one-dimensional convolutional neural network (1D-CNN) as a generator, and use category labels as auxiliary information to enhance the original GAN, improve the generation effect of the generator, and generate high-quality labeled artificial samples to expand the number of fault samples. The SDAE is used as a discriminator to identify the authenticity and fault of the input samples. SDAE can automatically extract features with better robustness by adding noise to samples for sample reconstruction. At the same time, in the process of simulating the generation of false data, the generator is helpful to understand the distribution of original data, and the adversarial learning is used as a cross domain regularizer to learn the universal and domain-invariant features of data.

The rest of this paper is organized as follows: in section 2, we will introduce the theoretical background of the relevant methods. Section 3 details our proposed ACGAN-SDAE method. In section 4, some experiments are performed to evaluate our method and other methods. Finally, section 5 summarizes the full text.

## 2 The theoretical background of related methods

### 2.1 SDAE

SDAE [26] consists of multiple Denoising Auto Encoder (DAE) [27] by stacking. Like the standard Auto Encoder (AE), DAE consists of encoder and decoder. But different from standard AE, DAE enhances the robustness of extracted features by adding impairment noise to the input data, thereby enhancing its anti-noise ability. The encoder compresses the input data in the high-dimensional space to obtain the encode vectors in the low-dimensional space. The decoder reconstructs the encode vectors to obtain the original input data without noise. The structure principle of standard DAE is shown as Fig 1.

Given an unlabeled rolling bearing faults sample training set  $\{x^m\}_{m=1}^M$ , the noise  $qD$  is added to the training sample  $x^m$  before coding to get the sample with noise  $\tilde{x}^m$ .

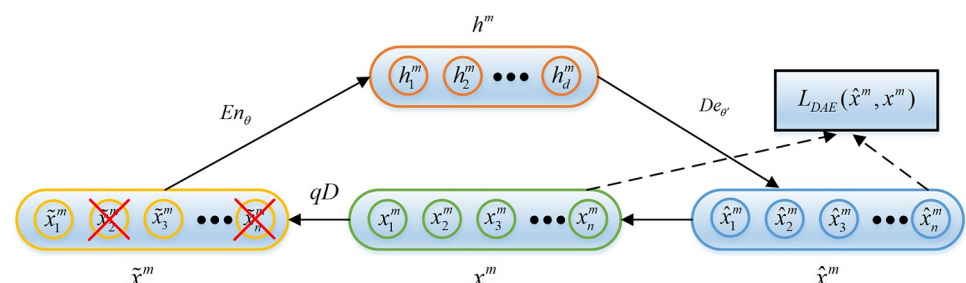
$$\tilde{x}^m \sim qD(\tilde{x}^m | x^m) \quad (1)$$

where  $qD$  is a binomial random hidden noise.

The coding network encodes the sample with noise  $\tilde{x}^m$ . In the coding process, the encode function  $En_\theta$  maps the samples  $\tilde{x}^m$  to the encode vectors  $h^m$ .

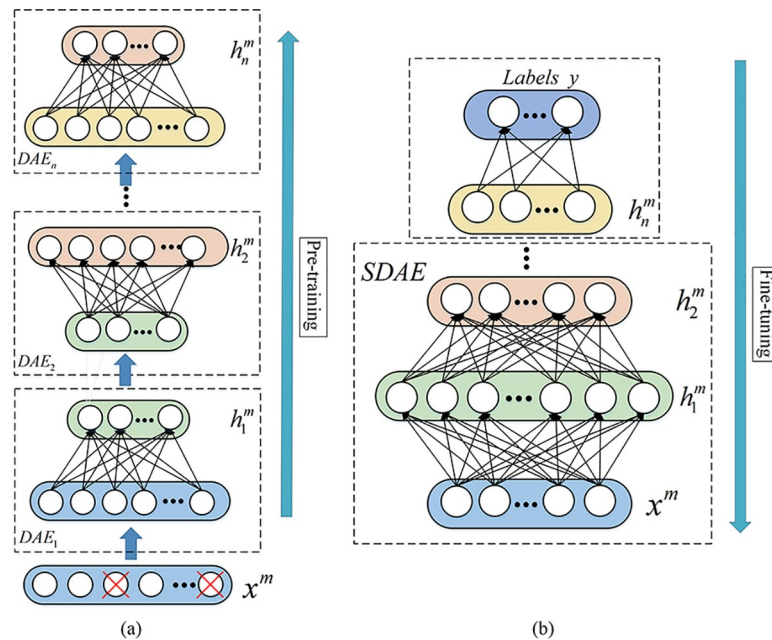
$$h^m = En_\theta(\tilde{x}^m) = \sigma(w_e \tilde{x}^m + b_e) \quad (2)$$

where  $\sigma$  is the sigmoid activation function and  $\sigma = 1/(1+\exp(-x))$ ,  $\theta$  is the parameters set of the encoding network and  $\theta = \{w_e, b_e\}$ ,  $w_e$  and  $b_e$  are the weight matrix and bias vector of the coding network respectively.



**Fig 1. The structure principle of DAE.**

<https://doi.org/10.1371/journal.pone.0246905.g001>



**Fig 2. The structure principle of SDAE.** (a) Pre-training, (b) Fine-tuning.

<https://doi.org/10.1371/journal.pone.0246905.g002>

The decoding network reversely transforms the encode vectors  $h^m$  into the reconstructed representation  $x^m$  of  $\hat{x}^m$  by decode function  $De_{\theta'}$ .

$$\hat{x}^m = De_{\theta'}(h^m) = \sigma(w_d h^m + b_d) \quad (3)$$

where  $\theta'$  is the parameters set of the decoding network and  $\theta' = \{w_d, b_d\}$ ,  $w_d$  and  $b_d$  are the weight matrix and bias vector of the decoding network respectively.

DAE aims to complete the training of the whole network by optimizing the parameters set  $\Theta = \{\theta, \theta'\}$  to minimize the reconstruction error  $L_{DAE}(\hat{x}^m, x^m)$  between  $\hat{x}^m$  and  $x^m$ .

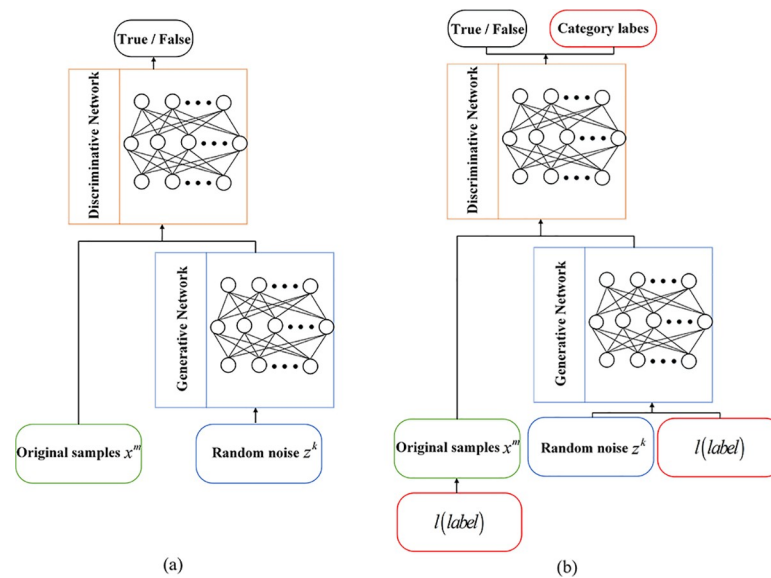
$$L_{DAE}(\hat{x}^m, x^m) = \arg \min_{\Theta} \frac{1}{M} \sum_{m=1}^M \|\hat{x}^m - x^m\|^2 \quad (4)$$

where  $\Theta$  is a set of parameters for DAE and  $\Theta = \{\theta, \theta'\}$ ,  $M$  is the sample size.

SDAE constructs deep networks by stacking multiple DAEs, and extracts deep features through unsupervised learning. SDAE training steps includes pre-training and fine-tuning, as shown in Fig 2. The pre-training step trains each layer of DAE through unsupervised layer-by-layer greedy learning to extract sample fault features. The encode vector of the hidden layer of the previous layer of DAE is used as the input to next layer of DAE, and repeat this process until the last layer of  $DAE_n$  is trained and the encoding vector  $h_n^m$  is obtained. Finally, supervised fine-tuning is carried out by using labeled sample data and adding a Softmax classifier at the top of the network.

## 2.2 GAN

The structure of GAN is inspired by game theory, regular GAN consists of two parts: generator  $G$  and discriminator  $D$  (as shown in Fig 3A). Among them,  $x^m$  is sampled from the original sample and  $z^k$  is the input of the generator. The generator aims to capture the potential distribution of real data samples  $x^m$  and generate realistic generated data  $G(z^k)$  from Gaussian random noise vector  $z^k$  in an attempt to deceive the discriminator. Instead, the purpose of the discriminator is to distinguish whether the input data is real data  $x^m$  or generated data  $G(z^k)$ .



**Fig 3.** The structure of (a) regular GAN, and (b) ACGAN.

<https://doi.org/10.1371/journal.pone.0246905.g003>

GAN continuously optimizes the generation ability of  $G$  and the discrimination ability of  $D$  through the adversarial learning mechanism, and finally they reach the Nash equilibrium. The optimization process is a minimax two-player game that can be formulated as:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x^m)] + E_{z \sim p_z(z)} [\log (1 - D(G(z^k)))] \quad (5)$$

where  $p_{data}(x)$  is the real data distribution,  $p_z(z)$  is the prior distribution of the noise vector  $z$ ,  $E_{x \sim p_{data}(x)}$  is the expected value of the real data distribution of  $x$ , and  $E_{z \sim p_z(z)}$  is the expected value of  $z$  sampled from the noise.

In the process of training, one part is fixed and the parameters of the other network are updated. Training  $D$  maximizes  $\log D(x^m)$  and training  $G$  minimizes  $\log(1 - D(G(z^k)))$ . The generator defines a probability distribution function  $p_g$ , and GAN expects  $p_g$  to converge to the real data distribution  $p_{data}$  through alternating iteration. If and only if  $p_g = p_{data}$  reaches Nash equilibrium, GAN can well estimate the actual distribution of real samples and generate new samples to expand the training fault sample set.

### 2.3 ACGAN

Odena et al. [28] proposed a variant architecture of regular GAN to achieve accurate classification of images in the MNIST dataset. This variant architecture is called Auxiliary Classifier Generative Adversarial Network (ACGAN) by adding category labels to generator and discriminator (as shown in Fig 3B). According to research, when GAN is allowed to process additional information, the original generation task of the model will be completed better. Therefore, high quality generated samples can be generated by using auxiliary category label information.

For the generator, there are two inputs: random noise vector  $z$  and label classification information  $c$ . And the generated data is  $X_{fake} = G(c, z)$ . For the discriminator, it is necessary to judge whether the data source is the probability distribution of the real data and the probability distribution of the data source for the classification label, so that the discriminator can not only identify the data source but also distinguish various fault categories. Therefore, the

objective function of ACGAN consists of two parts, as shown in the following formula:

$$L_s = E_{x \sim P_{data}} [\log P(\text{source} = \text{real} | x_{real})] + E_{z \sim P_z} [\log P(\text{source} = \text{fake} | x_{fake})] \quad (6)$$

$$L_c = E_{x \sim P_{data}} [\log P(\text{class} = c | x_{real})] + E_{z \sim P_z} [\log P(\text{class} = c | x_{fake})] \quad (7)$$

The first part  $L_s$  is a cost function for the truthfulness of the data, and the second part  $L_c$  is a cost function for the accuracy of data classification. In the training process, the optimization direction is to train the discriminator to maximize  $L_s + L_c$ , and the generator to minimize  $L_s - L_c$ . The corresponding physical meaning is that the discriminator is called upon to distinguish between real data and generated data as far as possible and to classify the data effectively. For the generator, the generated data are considered as real data as possible and the data can be classified effectively.

### 3. The proposed fault diagnosis method

In this paper, aiming at the data unbalance problem caused by the small fault sample size in the actual rolling bearing fault diagnosis, the cross domain adaptive problem with variable load conditions and the influence of high background noise pollution, a novel ACGAN-SDAE fault diagnosis method is proposed.

#### 3.1 Fault diagnosis model of ACGAN-SDAE

In this paper, by combining ACGAN and SDAE, we proposed an ACGAN-SDAE fault diagnosis method. The overall structure of the model as shown in Fig 4. In detail, an one-dimensional convolutional neural network (1D-CNN) [29] is used as the generator, and use category labels as auxiliary information to enhance the original GAN, improve the generation effect of the generator, and generate high-quality labeled artificial samples to expand the number of fault samples. By adding category labels, the generator can generate specific conditional data and make model training more stable. The SDAE is used as a discriminator to distinguish the authenticity and fault category of the input samples.

#### 3.2 Training of discriminator

The four-layer structure of the generator SDAE is 1024-800-200-10 (1), as shown in Fig 5. The generated samples  $\{x_{fake}^k\}_{k=1}^K$  are labeled as 0, and the corresponding real category labels are  $y_{fake}^k$ . The real samples  $\{x_{real}^m\}_{m=1}^M$  are labeled as 1, and the corresponding real category labels are  $y_{real}^m$ . Then input them together to the discriminator SDAE for authenticity discrimination and fault identification. SDAE adds two label classifiers in the top layer. The sigmoid function is used to predict the sample source, and output the corresponding authenticity labels  $d_{real}^m, d_{fake}^k$ . The softmax function is used to predict the fault category labels  $c_{real}^m, c_{fake}^k$ .

ACGAN-SDAE completes the training of discriminator by minimizing the error of authenticity labels and fault category labels through (10).

$$L_c = -\frac{1}{M} \sum_{m=1}^M [y_{real}^m \ln c_{real}^m + (1 - y_{real}^m) \ln (1 - c_{real}^m)] - \frac{1}{K} \sum_{k=1}^K [y_{fake}^k \ln c_{fake}^k + (1 - y_{fake}^k) \ln (1 - c_{fake}^k)] \quad (8)$$



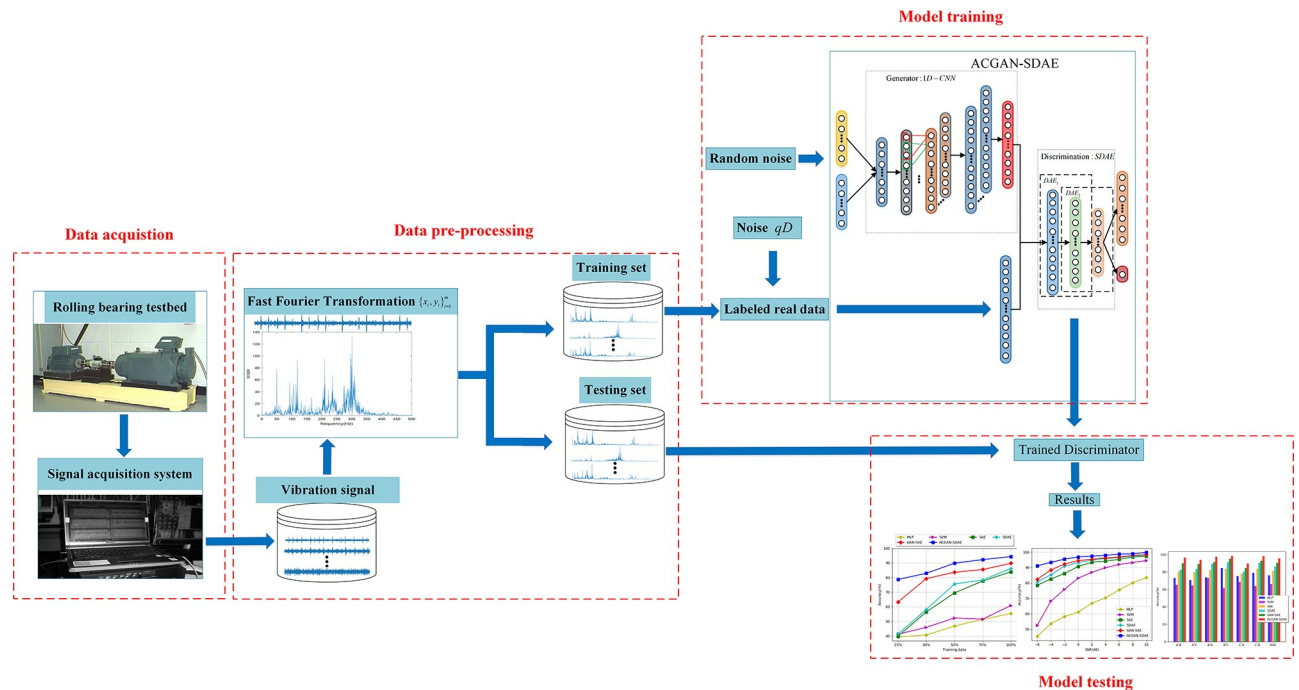


Fig 4. The overall architecture of ACGAN-SDAE fault diagnosis model.

<https://doi.org/10.1371/journal.pone.0246905.g004>

$$L_d = -\frac{1}{M} \sum_{m=1}^M \ln d_{real}^m - \frac{1}{K} \sum_{k=1}^K \ln(1 - d_{fake}^k) \quad (9)$$

$$L_D = \arg \min_{\Theta_D} (L_c + L_d) \quad (10)$$

where  $L_D$  is the loss function of the discrimination in ACGAN-SDAE,  $\Theta_D$  is the parameter set,  $L_c$  is the cross-entropy loss error of the category label,  $L_d$  is the cross-entropy loss error of the authenticity label.

### 3.3 Training of generator

The generator adopts 1D-convolution operation, in which the first layer is the input layer, which is a combination of Gaussian random noise vector input and category input, and contains two up-sampling layers of size 2. Then two layers of convolution are performed separately, and each layer uses batch normalization, and its momentum is 0.8. The kernel size of the first 1D convolutional layer is 16, has 16 feature maps, and uses the Rectified Linear Unit(ReLU) activation function. The second 1D fusion layer has a kernel size of 8, including 1 feature map, and uses the hyperbolic tangent function as the activation function. Its network structure is shown as in Fig 6. The output of the generator is one-dimensional data sample.

The generated sample  $\{x_{fake}^k\}_{k=1}^K$  are labeled as 1 and input it to discriminator SDAE for authenticity verification. Then complete the training of the generator by minimizing the (12).

$$L_g = -\frac{1}{K} \sum_{k=1}^K \ln d_{fake}^k \quad (11)$$

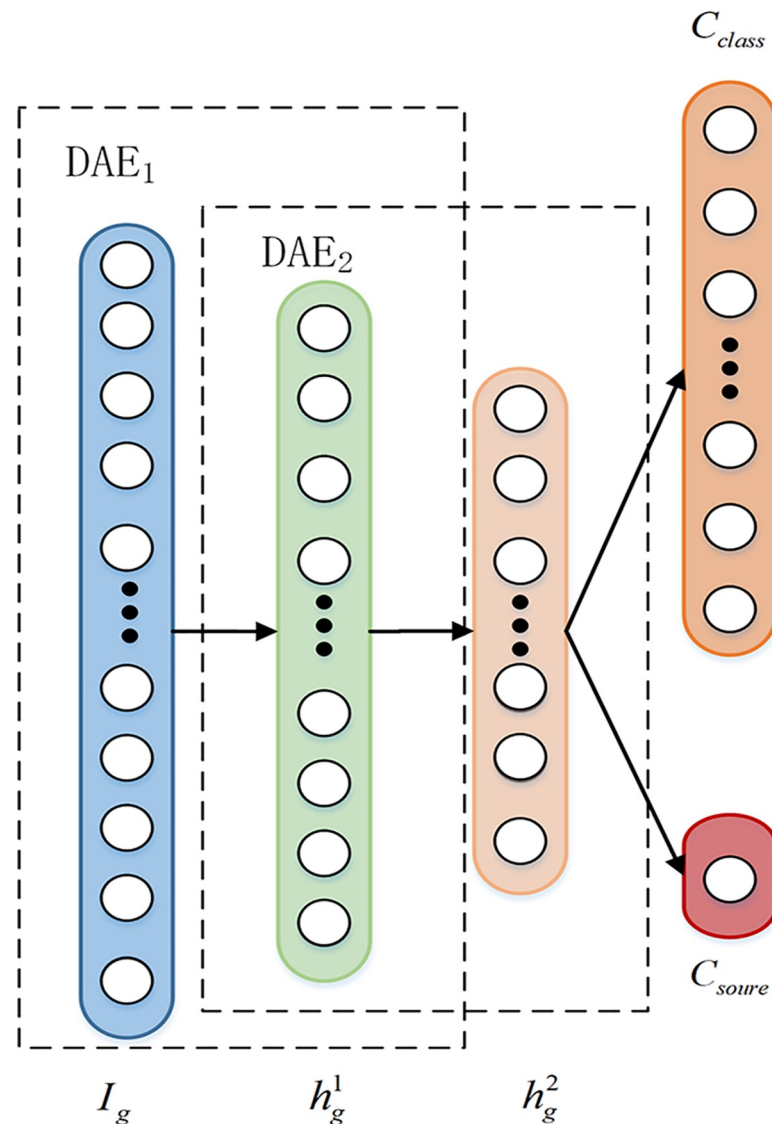


Fig 5. Network structure of discriminator.

<https://doi.org/10.1371/journal.pone.0246905.g005>

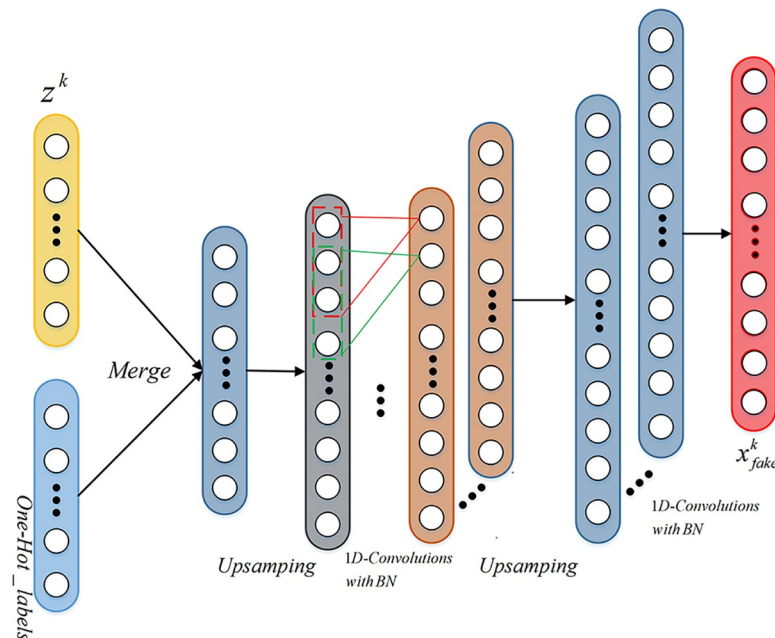
$$L_G = \operatorname{argmin}_{\Theta_G} (L_c + L_g) \quad (12)$$

where  $L_G$  is the loss function of the generator in ACGAN-SDAE,  $\Theta_G$  is the parameter set,  $L_g$  is the cross-entropy loss error of the authenticity label.

### 3.4 Adversarial training mechanism of model

The model realizes the adversarial training mechanism by alternately optimizing the generator and the discriminator. Through the zero-sum game between the them, its optimization goal has been passed as a minimum-maximization problem. Based on the above loss function, the ADAM optimizer is used for training, the learning rate of the discriminator is 0.001, the learning rate of the generator is 0.002, and updates parameters iteratively. The training process can be divided into three steps.





**Fig 6. Network structure of generator.**

<https://doi.org/10.1371/journal.pone.0246905.g006>

*Step 1:* Firstly, generator generates fake samples from Gaussian random noise of potential space with class labels.

*Step 2:* Then, the generated samples and the original samples are input into the generator SDAE for authenticity identification and fault classification. By training the above loss function, the labels and parameters in the discriminator are able to be updated.

*Step 3:* After training the discriminator, at this stage, the discriminator is set to be untrainable and its parameters are frozen. In this stage, only the parameters in generator can be updated, and generator can be trained to generate more realistic fake samples. After a period is completed, the training process starts again from *Step 1*.

Through the above multiple alternating optimization iterations, until the generator and discriminator reach Nash equilibrium, the training of the whole model is accomplished.

### 3.5 Implementation of fault diagnosis algorithm

The steps of fault diagnosis in this paper are mainly divided into three parts: data acquisition and pre-processing, model training, fault identification. The algorithm flow chart is shown in Fig 7.

1. **Data acquisition and pre-processing:** In the rolling bearing feature extraction process, considering the complexity of the original vibration signal, the spectral signal is used as the input signal of the model. Firstly, a sensor is used to collect the original vibration signal of the rolling bearing, and the frequency spectrum sample  $\{x_i, y_i\}_{i=1}^m$  is obtained through the Fast Fourier Transform (FFT), which is divided into training set and testing set.
2. **Model training:** The training set is input into the ACGAN-SDAE model. Through the adversarial training of generator and discriminator, the training of the whole model is completed by using ADMA optimizer, alternating optimization generator and discriminator until the they reaches Nash equilibrium.

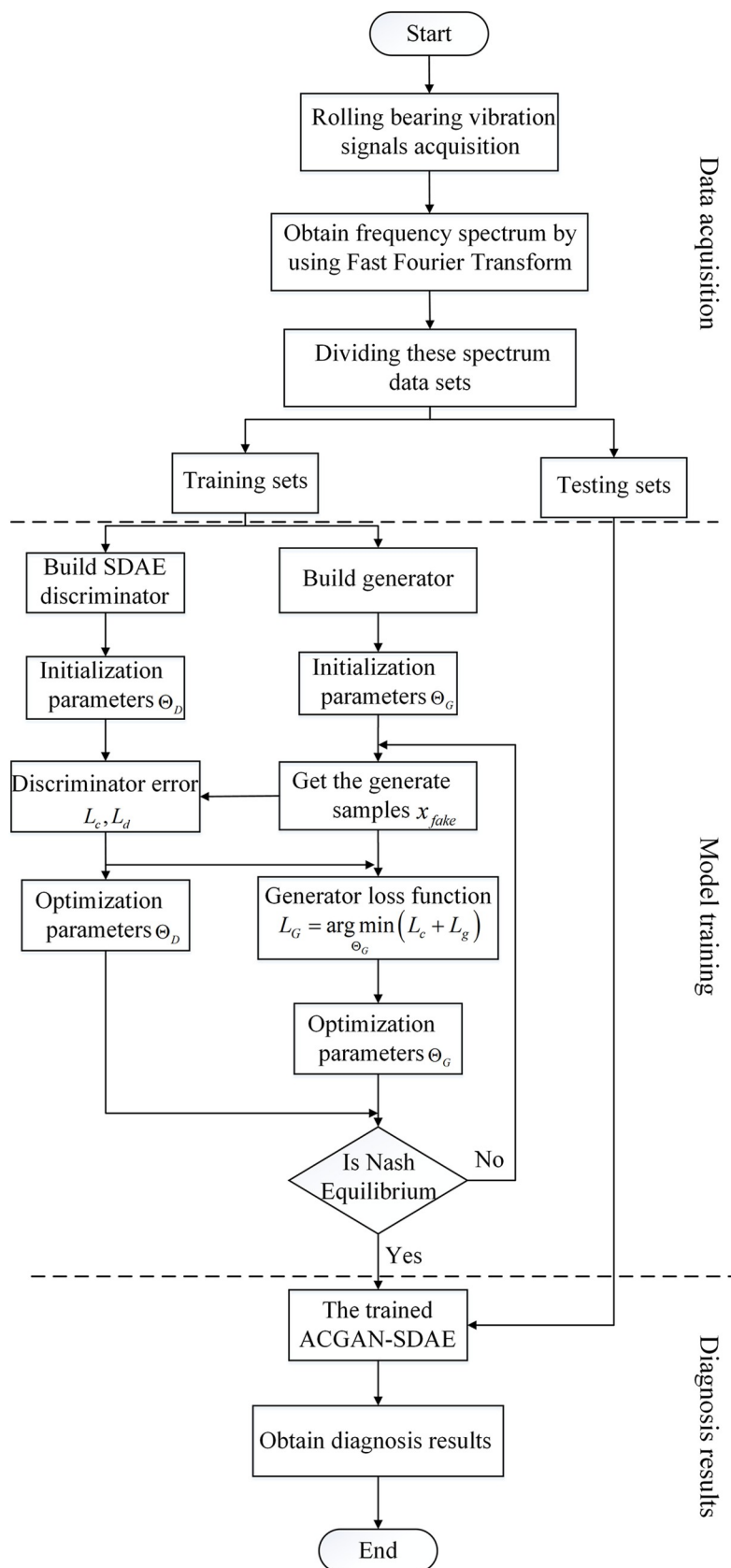
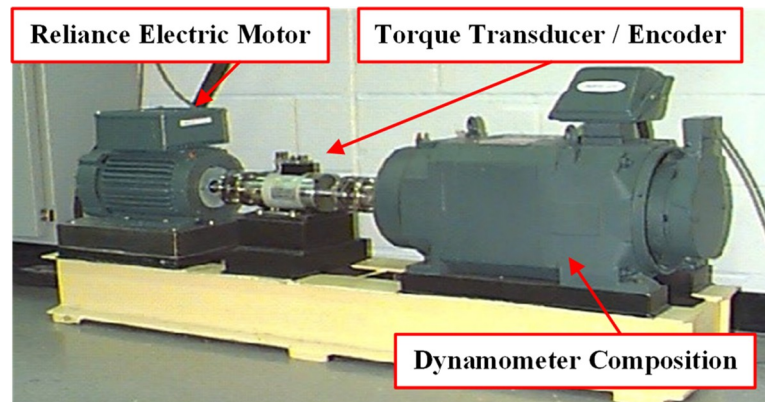


Fig 7. The flowchart of ACGAN-SDAE fault diagnosis method.

<https://doi.org/10.1371/journal.pone.0246905.g007>



**Fig 8. Experiment platform.**

<https://doi.org/10.1371/journal.pone.0246905.g008>

3. **Fault identification:** The testing set is input into the trained discriminator SDAE, and output the diagnosis results.

## 4. Experimental results and analysis

### 4.1 Dataset description

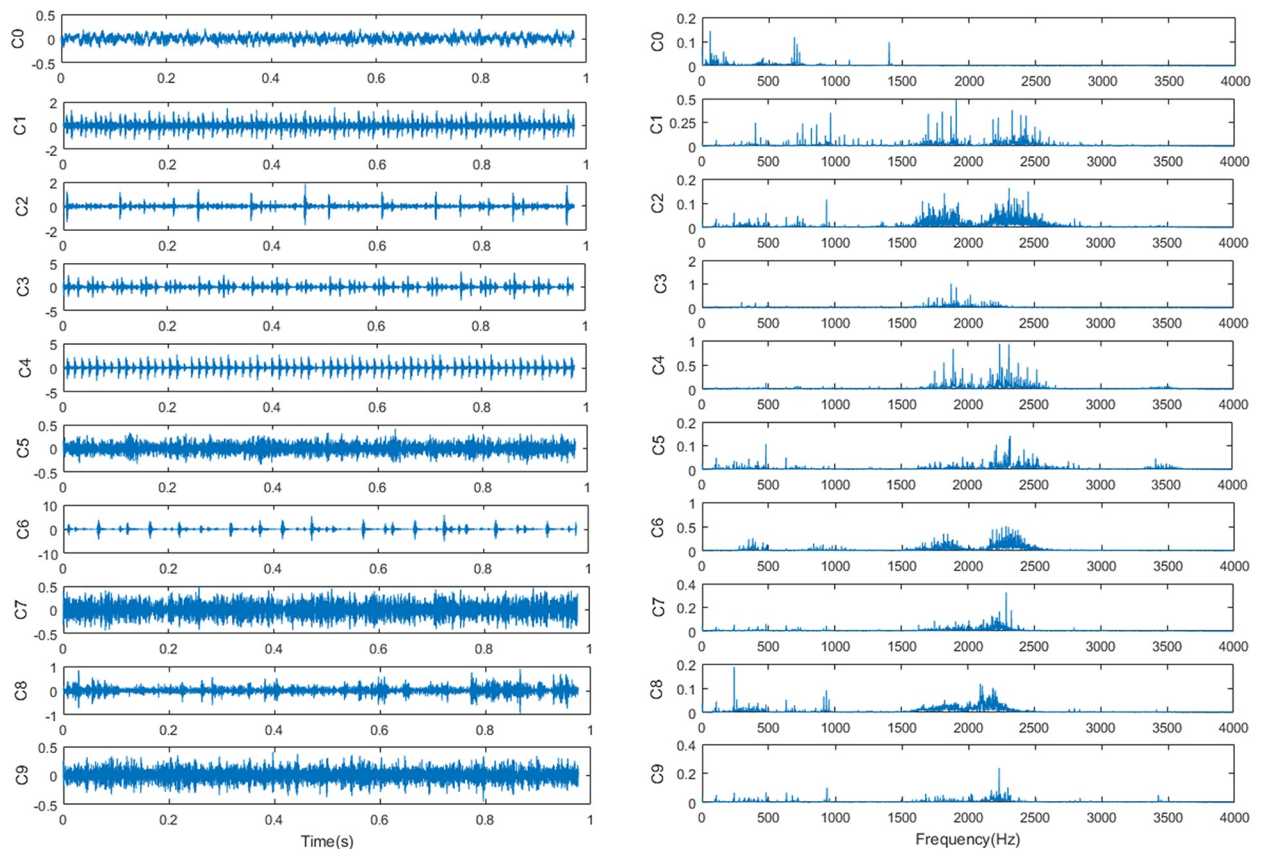
The CWRU rolling bearing data set is obtained by the Electrical Engineering Laboratory of Case Western Reserve University [30]. It is an open data set and widely used in fault diagnosis and it can be obtained through the website: <https://csegroups.case.edu/bearingdatacenter>. The experimental platform is shown in Fig 8. The vibration data used in this study was collected from the driving end of the motor at three speeds of 1750rpm, 1772rpm and 1797rpm. And its sampling frequency was 12kHz. The bearing with fault was machined by EDM, which caused different degrees of damage to the inner race, outer race and roller of the bearing. The damage diameter included 0.007 inches, 0.014 inches and 0.021 inches, with a total of 9 damage states. Therefore, the data contains four health states: 1) Normal condition (Normal) 2) Inner race fault (IF) 3) Outer race fault (OF) 4) Roller fault (RF). Typical time-domain waveform and frequency spectra of the original vibration signals in the 10 health conditions are shown in Fig 9.

In the experiment, FFT is used to preprocess the original signal to obtain spectrum samples, and 1024 data points are used for diagnosis each time. There are three datasets are set in the experiment, as shown in Table 1. Datasets A, B and C are data sets under the load of 1 hp, 2 hp and 3 hp respectively. Each data set contains 6600 training samples and 1000 testing samples.

### 4.2 Experiments settings

In this paper, three groups of experiments are set up to verify the effectiveness and robustness of the proposed ACGAN-SDAE model from unbalanced fault sample size, different signal-to-noise ratio and across different load domains. Therefore, we design three kinds of experiment data settings:

**4.2.1 Unbalanced fault sample size dataset.** Only part of the data is used for training. Supervised learning needs a huge number of training data to achieve good performance. However, in fact, we usually cannot get enough fault samples to train a deep learning model. Therefore, it is necessary to study the robustness of different models under small fault data. In our experiments, the unbalance rate of training samples in each state mode of data set A is 100%, 40%, 20%, 10% and 5% for comparison experiments.



**Fig 9. The original vibration signals in the 10 health conditions.** (a) the time-domain waveform and (b) the corresponding frequency spectra.

<https://doi.org/10.1371/journal.pone.0246905.g009>

**4.2.2 Different signal-to-noise ratio dataset.** In actual fault diagnosis, the sample signal usually contains a lot of noise, which makes the diagnosis performance of the model unsatisfactory. Therefore, in our experiments, Gaussian noise with different signal-to-noise ratio (SNR) from -6 to 10dB is added to data set A to test the recognition rate of the model, so as to verify the anti-noise performance of the model. SNR is defined as follows:

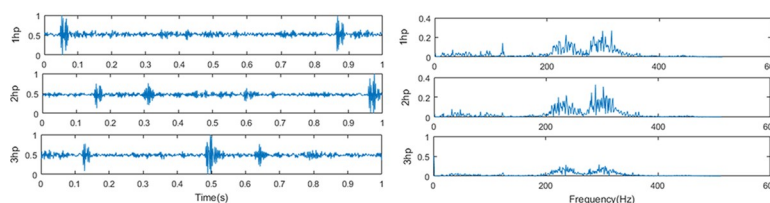
$$SNR_{dB} = 10 \log_{10} \left( \frac{P_{signal}}{P_{noise}} \right) \quad (13)$$

where  $P_{signal}$  and  $P_{noise}$  are the power of signal and noise respectively.

**Table 1. Description of experimental dataset.**

Fault type	Normal	IF	IF	IF	OF	OF	OF	RF	RF	RF	Load
Fault number	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	
Damage diameter (inch)	0	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021	
A	train	660	660	660	660	660	660	660	660	660	1
	test	100	100	100	100	100	100	100	100	100	
B	train	660	660	660	660	660	660	660	660	660	2
	test	100	100	100	100	100	100	100	100	100	
C	train	660	660	660	660	660	660	660	660	660	3
	test	100	100	100	100	100	100	100	100	100	

<https://doi.org/10.1371/journal.pone.0246905.t001>



**Fig 10. The diagnostic signal with an inner race fault size of 0.014 inches under different loads.** (a) the time-domain waveform and (b) the corresponding frequency spectra.

<https://doi.org/10.1371/journal.pone.0246905.g010>

**4.2.3 Across different load domains dataset.** Across different load domains problem is also called cross load domain adaptive problem. Fig 10 shows the time-domain waveform and frequency spectrum of the diagnostic signal with an inner race fault size of 0.014 inches under different loads. It can be seen from Fig 10 that under different loads, the time-domain and frequency spectrum features of the vibration signal are very different, which will cause the classifier to fail to correctly classify the extracted features, thereby reducing the fault recognition rate. Therefore, it is of great practical significance to use the diagnostic model trained with the data under a load to diagnose the vibration signal when the load changes. The ACGAN-SDAE model will be trained using samples with loads of 1hp, 2hp and 3hp respectively, and the signals under the other two loads will be used as the test set for testing. Detailed description of the across different load domains data is shown in Table 2.

All the network models used in this paper are trained under the Ubuntu 16.04 operating system and Keras deep learning framework. The CPU model is Intel(R) Core(TM) i7-8700, 16GB memory, the graphics card model is NVIDIA GeForce GTX 1080 Ti. The algorithm is implemented through Python3.6 programming language.

### 4.3 Experimental design and result analysis

**4.3.1 Noise factor selection of ACGAN-SDAE.** The diagnostic performance of ACGAN-SDAE is mainly affected by the discriminator SDAE, and the selection of noise factor  $\rho$  directly affects the performance of SDAE. The noise factor of SDAE is expressed as the random zeroing ratio of input data. The noise factor is too large or too small will make the SDAE performance decline. If the noise factor is too large, the input signal will be masked by the noise, so that the SDAE can not extract rich fault features. If the noise factor is too small, the sample damage is too little, which leads to poor anti-noise performance. In this section, we have optimized the noise factor  $\rho$ .

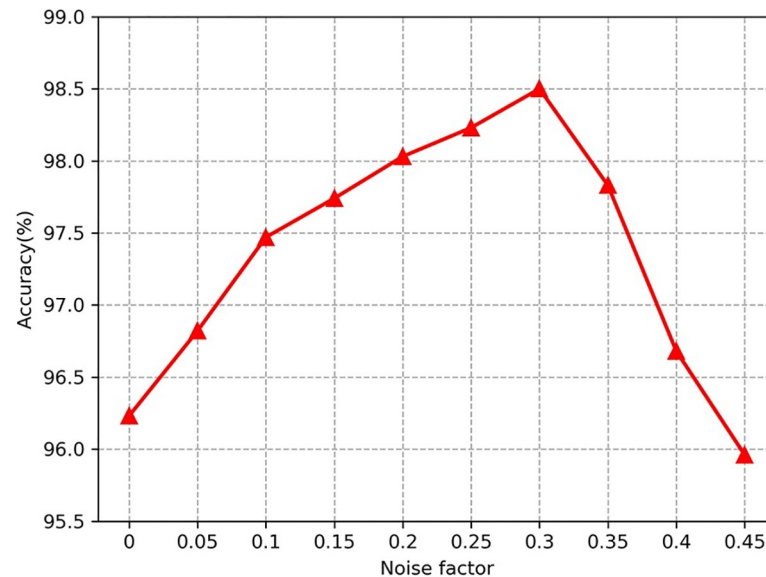
Here, ten noise factors of different sizes are studied to determine the best noise factor. To eliminate the effect of contingency, ten tests were carried out, and the average values were taken as the results, the specific results are shown in Fig 11. It can be seen that with the increase of noise factor, the diagnostic accuracy presents an "arch" distribution, which is consistent with the previous analysis. The highest diagnostic accuracy is 98.5% when the noise factor is 0.3. Therefore, the noise factor is finally selected as  $\rho = 0.3$ .

**Table 2. Description of across different load domains data.**

Dataset types	Training set	Testing set	
Description	labeled signals under one signal load	unlabeled signals under another load	
Dataset	Training set A	Testing set B	Testing set C
	Training set B	Testing set C	Testing set A
	Training set C	Testing set A	Testing set B

<https://doi.org/10.1371/journal.pone.0246905.t002>





**Fig 11. Diagnosis accuracy of ACGAN-SDAE under different noise factors.**

<https://doi.org/10.1371/journal.pone.0246905.g011>

**4.3.2 Comparison of diagnostic performance under different fault sample sizes.** In this section, unbalanced fault sample size data set will be used for comparison experiments, and the  $\text{SNR} = -4\text{dB}$  Gaussian noise will be added to data set to simulate the operating environment under noise interference. We compared the diagnostic accuracy of our method with GAN-SAE, SAE, SDAE, MLP and SVM respectively. Among them, the generator of GAN-SAE uses BPNN with a 256-512-1024 structure, and the discriminator uses SAE, which structure is 1024-512-256-10. Both SAE and SDAE are 4-layer with the structure of 1024-512-256-10. The noise factor of SDAE is 0.3. The inputs of GAN-SAE, SAE and SDAE are all spectrum samples. The kernel function of SVM adopts Radial Basis Function (RBF), and the penalty factor and related parameters of the kernel function are optimized through the method of cross validation. The number of hidden nodes set by MLP is 50. And 59 artificial extracted features [31] are selected from time domain and frequency domain as inputs to shallow model SVM and MLP.

In order to decrease the influence of randomness, repeat the experiment ten times and use the average value as the final diagnosis result. The diagnosis results of different models under different fault sample sizes are given in Fig 12 and Table 3. As can be seen from Fig 12 and Table 3, with the increase of the proportion of fault samples, the diagnosis accuracy of deep network model is significantly improved, while the diagnosis accuracy of shallow model has little change. This is because the deep network model can mine rich information from big data, which significantly improve the diagnostic accuracy of the model. In addition, the diagnostic accuracy of GAN-SAE is higher than the SAE and SDAE, which indicate that GAN can improve the generalization ability of the model under limited fault sample size. Under different fault sample sizes, ACGAN-SDAE obtains better results than other comparison methods. In addition, although only 25% of the fault samples were used in our method, the accuracy can reach 78.67%, which proved the good robustness of our method. ACGAN-SDAE combines the generator and discriminator with the adversarial learning mechanism, and use category labels as auxiliary information to enhance the original GAN, improve the generation effect of the generator, and generate high-quality labeled artificial samples to expand the number of fault samples, which greatly improves the fault diagnosis performance in the case of small samples.



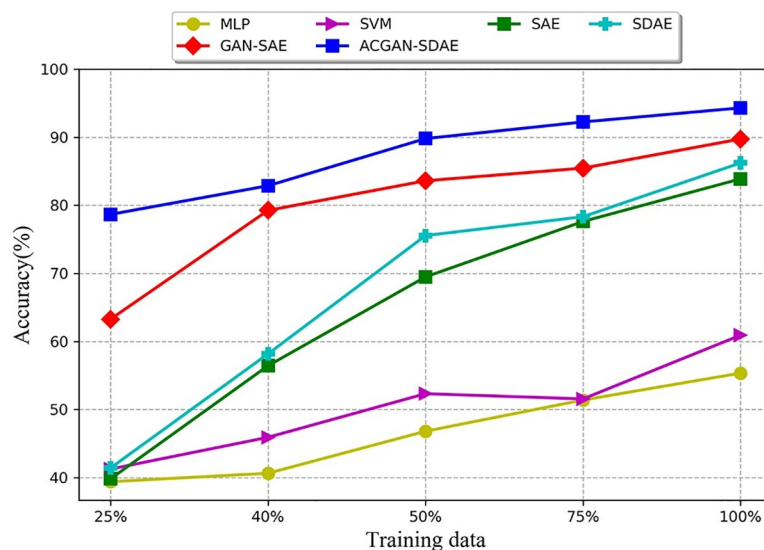


Fig 12. Diagnosis results of different models in different fault sample proportions.

<https://doi.org/10.1371/journal.pone.0246905.g012>

**4.3.3 Comparison of diagnostic performance under different SNRs and across different load domains.** In this section, different SNRs data set are used to verify ACGAN-SDAE anti-noise performance. The results are shown in Fig 13 and Table 4. As can be seen from Fig 13 and Table 4, with the decrease of SNRs, the diagnostic performance of different models will also decline. This is because with the increase of noise intensity, the sample damage is more serious, which makes it difficult for the model to extract effective features. In addition, although the diagnostic results of SVM are better than MLP, the anti-noise performance of both models is very weak. For example, in the environment with weak noise  $SNR = 4dB$ , the diagnostic accuracy of both models is less than 90%. In contrast, the diagnostic accuracy of ACGAN-SDAE under different SNRs is higher than 90%, and even in the environment with  $SNR = -6dB$ , it is 8.76% and 10.48% higher than GAN-SAE and SDAE, respectively. Benefiting from the adversarial learning mechanism and the denoising principle, ACGAN-SDAE have the best anti-noise robustness in a strong noise environment.

Next, we used across different load domains data set to simulate fault diagnosis under variable working conditions, and further test domain adaptation performance of ACGAN-SDAE. The results are shown in Fig 14 and Table 5. As can be seen from Fig 14 and Table 5, the average accuracy of SVM and MLP is less than 80%, the performance of GAN-SAE is better than SAE and SDAE, and the average diagnostic accuracy can reach 90.61%. GAN-SAE can learn more sample features through adversarial training. ACGAN-SDAE helps to understand the original data distribution during the generator's simulation of the generation of fake data. The

Table 3. Diagnosis accuracy of different models in different fault sample proportions (%).

Method	25%	40%	50%	75%	100%
MLP	39.40	40.64	46.81	51.40	55.35
SVM	41.25	45.92	52.33	51.56	60.92
SAE	39.82	56.44	69.49	77.63	83.89
SDAE	41.41	58.18	75.56	78.33	86.23
GAN-SAE	63.28	79.26	83.61	85.47	89.75
ACGAN-SDAE	78.67	82.88	89.81	92.25	94.32

<https://doi.org/10.1371/journal.pone.0246905.t003>

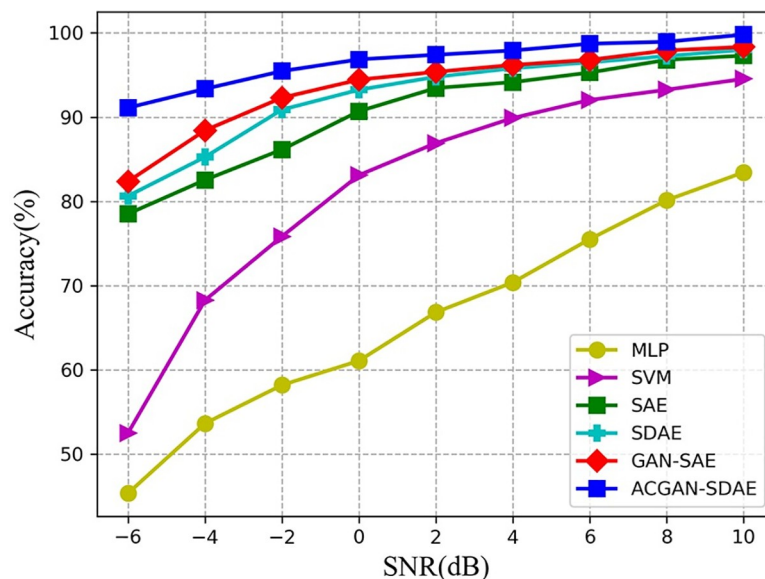


Fig 13. Diagnosis results of six diagnosis models under different SNRs.

<https://doi.org/10.1371/journal.pone.0246905.g013>

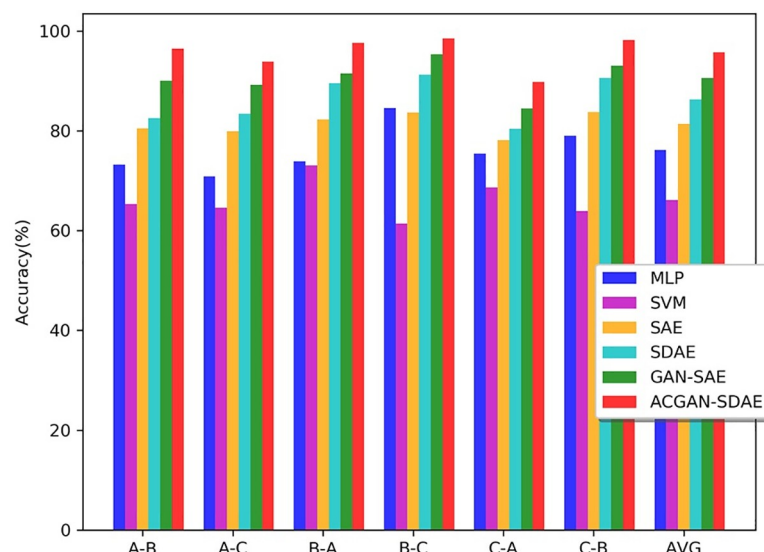
adversarial learning can be used as a cross domain regularizer, and the universal and domain-invariant features of data can be better learned. As a result, the model has significant cross domain adaptive ability, with the average accuracy reaching the highest 95.75%. In addition, we can observe that the diagnostic accuracy of all models from C to A and A to C is significantly lower than that of other cross domain conditions. This result is consistent with intuition. When there are big differences between the two working conditions, the diagnostic performance of the model is poor, that is, domain adaptability of the model are not good.

**4.3.4 Feature extraction and generate samples visual analysis of ACGAN-SDAE.** In order to better understand the feature extraction ability of ACGAN-SDAE, by using t-SNE [32] dimension reduction technology, the features of dataset A at the input layer and the last hidden layer of the discriminator SDAE were reduced to two dimensions and visualized. It can be seen from Fig 15 that the original feature distribution of the input signal is scattered before feature extraction, and different categories are mixed with each other. It is difficult to distinguish them. Through the feature extraction of the discriminator SDAE, we can clearly see that the features of the same fault type have been well aggregated, and the features of different fault types have been well separated. This indicates that ACGAN-SDAE has excellent feature extraction capabilities and can effectively distinguish various fault types. Fig 16 shows the training loss of confrontation and classification in each epoch. It can be observed that with the increase

Table 4. Diagnosis accuracy of six diagnosis models under different SNRs (%).

Method	SNR(dB)								
	-6	-4	-2	0	2	4	6	8	10
MLP	45.32	53.62	58.17	61.06	66.82	70.37	75.51	80.14	83.45
SVM	52.48	68.25	75.83	83.11	86.92	89.90	92.04	93.26	94.55
SAE	78.51	82.54	86.15	90.70	93.46	94.16	95.30	96.81	97.31
SDAE	80.64	85.26	90.89	93.25	94.79	95.81	96.52	97.26	98.01
GAN-SAE	82.36	88.43	92.31	94.45	95.38	96.20	96.81	97.92	98.35
ACGAN-SDAE	<b>91.12</b>	<b>93.36</b>	<b>95.47</b>	<b>96.86</b>	<b>97.41</b>	<b>97.92</b>	<b>98.27</b>	<b>98.96</b>	<b>99.80</b>

<https://doi.org/10.1371/journal.pone.0246905.t004>



**Fig 14. Diagnosis accuracy of different models under across different load domains.**

<https://doi.org/10.1371/journal.pone.0246905.g014>

of the number of epochs, the training loss of the generator and the discriminator gradually converges and finally stays around the Nash equilibrium, and the classification loss also tends to be convergent and stable.

After training, generated samples from ACGAN-SDAE are obtained. Fig 17 shows the frequency spectrum of the original samples and the corresponding generated samples under nine fault conditions. From the figure shown, we can see that the original samples and the corresponding generated samples are highly similar, that is, the samples are different but the distributions are similar. Therefore, ACGAN-SDAE can effectively learn the data distribution by adding auxiliary category label information to generate high-quality generated samples similar to the original samples, so as to expand the number of fault samples, and further improve the robustness of the model.

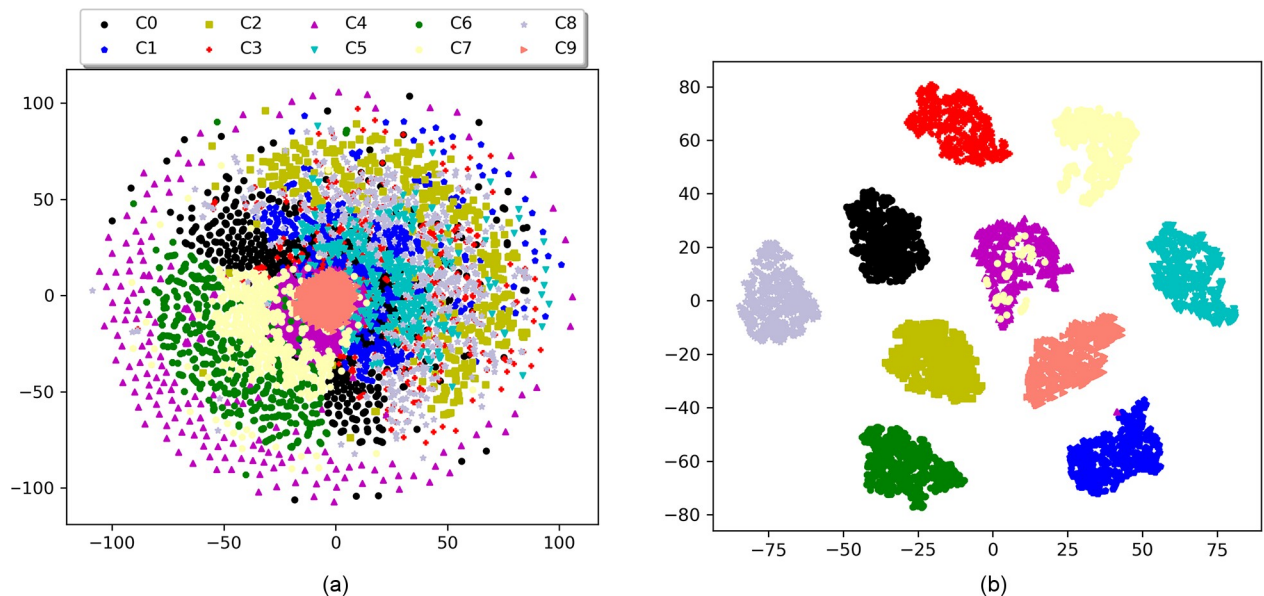
## 5. Conclusions and future work

In this paper, a novel ACGAN-SDAE fault diagnosis method is proposed to solve the problem of data unbalance caused by small sample size, cross domain adaptive problem under variable load and the influence of high background noise pollution in the fault diagnosis of rolling bearing. Through the analysis of the experimental results, the following conclusions can be drawn:

**Table 5. Diagnosis accuracy of different models under across different load domains (%).**

Method	A-B	A-C	B-A	B-C	C-A	C-B	AVG
MLP	73.24	70.89	73.86	84.61	75.47	79.05	76.18
SVM	65.31	64.55	73.05	61.44	68.65	63.92	66.15
SAE	80.53	79.92	82.33	83.65	78.11	83.73	81.37
SDAE	82.51	83.47	89.58	91.28	80.40	90.62	86.31
GAN-SAE	90.02	89.25	91.50	95.31	84.52	93.06	90.61
ACGAN-SDAE	<b>96.47</b>	<b>93.90</b>	<b>97.62</b>	<b>98.53</b>	<b>89.81</b>	<b>98.22</b>	<b>95.75</b>

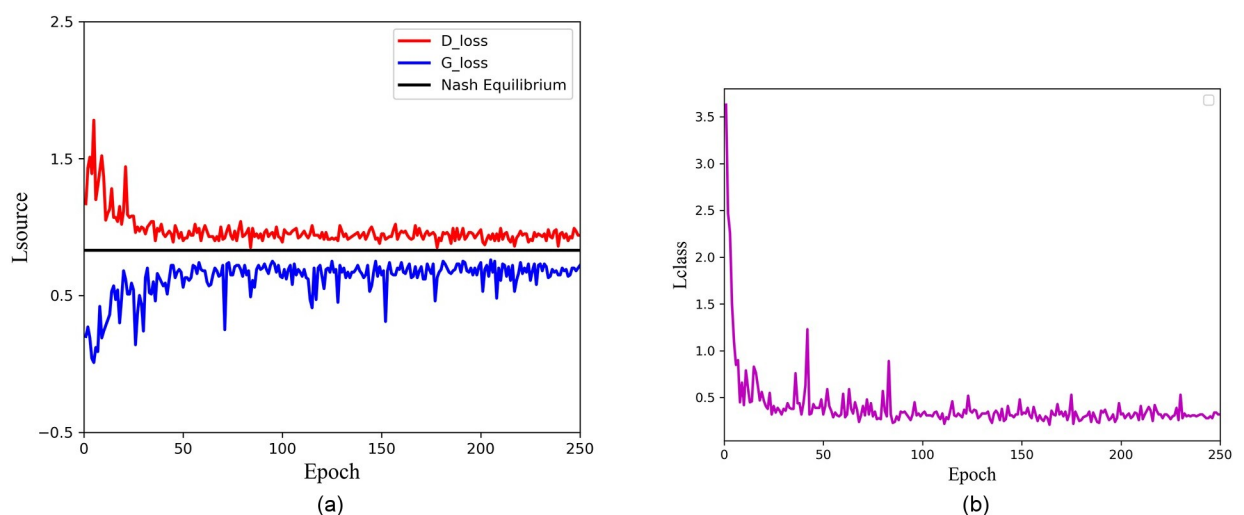
<https://doi.org/10.1371/journal.pone.0246905.t005>



**Fig 15. Feature visualization via t-SNE.** (a) features of original input data, (b) features extracted by ACGAN-SDAE.

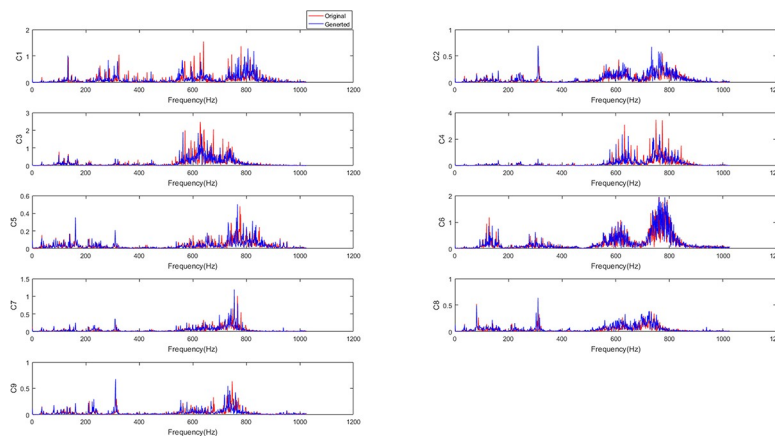
<https://doi.org/10.1371/journal.pone.0246905.g015>

1. ACGAN can adaptively learn the data distribution to generate high-quality artificial samples by adding auxiliary category label information, so as to expand the number of training fault samples and improve the fault feature extraction ability of the model under the condition of small fault samples.
2. SDAE can be used as discriminator to automatically extract features with better robustness, which makes the model have stronger anti-noise capability. At the same time, in the process of simulating the generation of fake data, the generator is helpful to understand the distribution of original data, and the adversarial learning is used as cross domain regularizer to learn the universal and domain-invariant features of data, which makes the model have significant cross domain adaptive ability.



**Fig 16. ACGAN-SDAE training loss.** (a) adversarial loss, (b) classification loss.

<https://doi.org/10.1371/journal.pone.0246905.g016>



**Fig 17. The spectrum of original samples and corresponding generated samples under nine fault conditions.**

<https://doi.org/10.1371/journal.pone.0246905.g017>

- Compared with other fault diagnosis models (GAN-SAE, SDAE, SAE, MLP and SVM), ACGAN-SDAE has better diagnosis performance and stronger robustness.

At present, spiking neural P systems (in short, SNP systems) [33] is in full swing, and it is commonly used in the power system fault diagnosis field [34]. SNP systems are a type of membrane computing model, which is abstracted from the neurophysiological behavior of biological neurons sending out electronic pulses along synapses. SNP systems are a distributed and parallel computing model in which neurons work in parallel. Therefore, in the future work, we will apply SNP systems and variant structure to mechanical fault diagnosis to solve the uncertainty and incomplete problems in mechanical fault.

## Supporting information

**S1 Data.**  
(RAR)

## Acknowledgments

The authors would like to thank the anonymous reviewers for their critical and constructive comments, their thoughtful suggestions have helped improve this paper substantially.

## Author Contributions

**Conceptualization:** Chunming Wu.

**Data curation:** Zhou Zeng.

**Formal analysis:** Chunming Wu.

**Funding acquisition:** Chunming Wu.

**Methodology:** Zhou Zeng.

**Software:** Zhou Zeng.

**Supervision:** Chunming Wu.

**Writing – original draft:** Zhou Zeng.

**Writing – review & editing:** Chunming Wu.

## References

1. Hui K H, Ooi C S, Lim M H, Leong M S, Al-Obaidi S M. An improved wrapper-based feature selection method for machinery fault diagnosis. *PLoS One*. 2017; 12(12): 10. <https://doi.org/10.1371/journal.pone.0189143> PMID: 29261689
2. Lu C, Wang Y, Ragulskis M, Cheng Y J. Fault Diagnosis for Rotating Machinery: A Method based on Image Processing. *PLoS One*. 2016; 11(10): 22. <https://doi.org/10.1371/journal.pone.0164111>
3. Zhang L, Lang Z-Q. Wavelet Energy Transmissibility Function and Its Application to Wind Turbine Bearing Condition Monitoring. *IEEE Transactions on Sustainable Energy*. 2018; 9(4): 1833–1843. <https://doi.org/10.1109/TSTE.2018.2816738>
4. Wang WJ, McFadden PD. EARLY DETECTION OF GEAR FAILURE BY VIBRATION ANALYSIS .1. CALCULATION OF THE TIME-FREQUENCY DISTRIBUTION. *Mechanical Systems and Signal Processing*. 1993; 7(3): 193–203. <https://doi.org/10.1006/mssp.1993.1008>
5. Seshadrinath J, Singh B, Panigrahi BK. Incipient Turn Fault Detection and Condition Monitoring of Induction Machine Using Analytical Wavelet Transform. *IEEE Transactions on Industry Applications*. 2014; 50(3): 2235–2242. <https://doi.org/10.1109/tia.2013.2283212>
6. Lu S, Wang J, Xue Y. Study on multi-fractal fault diagnosis based on EMD fusion in hydraulic engineering. *Applied Thermal Engineering*. 2016; 103: 798–806. <https://doi.org/10.1016/j.applthermaleng.2016.04.036>
7. Espinosa A G, Rosero J A, Cusido J, Romeral L, Ortega J A. Fault Detection by Means of Hilbert-Huang Transform of the Stator Current in a PMSM With Demagnetization. *IEEE Trans Energy Convers*. 2010; 25(2): 312–318. <https://doi.org/10.1109/tec.2009.2037922>
8. Sadeghian A, Ye Z M, Wu B. Online Detection of Broken Rotor Bars in Induction Motors by Wavelet Packet Decomposition and Artificial Neural Networks. *IEEE Transactions on Instrumentation and Measurement*. 2009; 58(7): 2253–2263. <https://doi.org/10.1109/tim.2009.2013743>
9. Lei Y G, He Z J, Zi Y Y. Application of the EEMD method to rotor fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*. 2009; 23(4): 1327–1338. <https://doi.org/10.1016/j.ymssp.2008.11.005>
10. Wang Y X, Markert R, Xiang J W, Zheng W G. Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. *Mechanical Systems and Signal Processing*. 2015; 60–61: 243–251. <https://doi.org/10.1016/j.ymssp.2015.02.020>
11. Teng W, Ding X, Cheng H, Han C, Liu Y, Mu H. Compound faults diagnosis and analysis for a wind turbine gearbox via a novel vibration model and empirical wavelet transform. *Renewable Energy*. 2019; 136: 393–402. <https://doi.org/10.1016/j.renene.2018.12.094>
12. Wanderley Neto E T, da Costa E G, Maia M J A. Artificial Neural Networks Used for ZnO Arresters Diagnosis. *IEEE Transactions on Power Delivery*. 2009; 24(3): 1390–1395. <https://doi.org/10.1109/tpwr.2009.2013402>
13. Zhu X T, Xiong J B, Liang Q. Fault Diagnosis of Rotation Machinery Based on Support Vector Machine Optimized by Quantum Genetic Algorithm. *IEEE Access*. 2018; 6: 33583–8. <https://doi.org/10.1109/access.2018.2789933>
14. Song L, Yan R. Bearing fault diagnosis based on Cluster-contraction Stage-wise Orthogonal-Matching Pursuit. *Measurement*. 2019; 140: 240–253. <https://doi.org/10.1016/j.measurement.2019.03.061>
15. Shao H D, Jiang H K, Zhao H W, Wang F A. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing*. 2017; 95: 187–204. <https://doi.org/10.1016/j.ymssp.2017.03.034>
16. Yuan J, Tian Y. A Multiscale Feature Learning Scheme Based on Deep Learning for Industrial Process Monitoring and Fault Diagnosis. *IEEE Access*. 2019; 7: 151189–151202. <https://doi.org/10.1109/access.2019.2947714>
17. Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*. 2006; 313(5786): 504–507. <https://doi.org/10.1126/science.1127647> PMID: 16873662
18. Chen Z Y, Zeng X Q, Li W H, Liao G L. Machine Fault Classification Using Deep Belief Network. *Proceedings of 2016 IEEE International Instrumentation and Measurement Technology Conference; USA*. New York: IEEE; 2016. p.831–836. <https://doi.org/10.1109/I2MTC.2016.7520473>
19. Jiang G, He H, Yan J, Xie P. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox. *IEEE Transactions on Industrial Electronics*. 2019; 66(4): 3196–3207. <https://doi.org/10.1109/tie.2018.2844805>
20. Lei J H, Liu C, Jiang D X. Fault diagnosis of wind turbine based on Long Short-term memory networks. *Renewable Energy*. 2019; 133: 422–432. <https://doi.org/10.1016/j.renene.2018.10.031>



21. Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. *Proceedings of 28th Conference on Neural Information Processing Systems (NIPS)*; Canada. Montreal: Neural information processing systems foundation; 2014. p.2672–2680.
22. Radford A, Metz L. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *Proceedings of 4th International Conference on Learning Representations*; Puerto rico. San Juan: ICLR; 2016.
23. Shao S Y, Wang P, Yan R Q. Generative adversarial networks for data augmentation in machine fault diagnosis. *Computer Ind.* 2019; 106: 85–93. <https://doi.org/10.1016/j.compind.2019.01.001>
24. Han T, Liu C, Yang W, Jiang D. A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. *Knowledge-Based Systems.* 2019; 165: 474–487. <https://doi.org/10.1016/j.neucom.2016.01.120>
25. Zhao Z, Zhou R, Dong Z. Aero-engine faults diagnosis based on K-means improved Wasserstein GAN and relevant vector machine. *Proceedings of the 38th Chinese Control Conference*; China. Guangzhou: IEEE; 2019. p.4795–4800. <https://doi.org/10.23919/ChiCC.2019.8865682>
26. Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P-A. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research.* 2010; 11: 3371–408. <https://doi.org/10.1016/j.mechatronics.2010.09.004>
27. Vincent P, Larochelle H, Bengio Y, Manzagol P A. Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th International Conference on Machine Learning*; Finland. Helsinki: ACM; 2008. p.1096–1103. <https://doi.org/10.1145/1390156.1390294>
28. Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier GANs. *Proceedings of 34th International Conference On Machine Learning, ICML 2017*; Australia. Sydney: IMLS; 2017. p.4043–4055.
29. Abdeljaber O, Avci O, Kiranyaz S, Gabbouj M, Inman DJ. Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *Journal of Sound and Vibration.* 2017; 388: 154–170. <https://doi.org/10.1016/j.jsv.2016.10.043>
30. These Data Comes From Case Western Reserve University Bearing Data Center, Dec. 2018, [online] Available from: <https://csegroups.case.edu/bearingdatacenter>
31. Rauber T W, Boldt F d A, Varejao F M. Heterogeneous Feature Models and Feature Selection Applied to Bearing Fault Diagnosis. *IEEE Transactions on Industrial Electronics.* 2015; 62(1): 637–646. <https://doi.org/10.1109/tie.2014.2327589>
32. van der Maaten L, Hinton G. Visualizing Data using t-SNE. *Journal of Machine Learning Research.* 2008; 9: 2579–2605.
33. Ionescu M, Paun G, Yokomori T. Spiking Neural P Systems. *Fundamenta Informaticae.* 2011; 71(2): 279–308. <https://doi.org/10.1109/BICTA.2010.5645192>
34. Tu M, Wang J, Peng H, et al. Application of Adaptive Fuzzy Spiking Neural P Systems in Fault Diagnosis of Power Systems. *Chinese Journal of Electronics.* 2014, 23(1): 87–92. <https://doi.org/10.3233/JAE-131740>