Computer Modeling in Engineering & Sciences

Tech Science Press

DOI: 10.32604/cmes.2023.025307

ARTICLE

A Novel Motor Fault Diagnosis Method Based on Generative Adversarial Learning with Distribution Fusion of Discrete Working Conditions

Qixin Lan, Binqiang Chen* and Bin Yao

School of Aerospace Engineering, Xiamen University, Xiamen, 361005, China

*Corresponding Author: Binqiang Chen. Email: cbq@xmu.edu.cn

Received: 04 July 2022 Accepted: 26 September 2022

ABSTRACT

Many kinds of electrical equipment are used in civil and building engineering. The motor is one of the main power components of this electrical equipment, which can provide stable power output. During the long-term use of motors, various motor faults may occur, which affects the normal use of electrical equipment and even causes accidents. It is significant to apply fault diagnosis for the motors at the construction site. Aiming at the problem that signal data of faulty motor lack diversity, this research designs a multi-layer perceptron Wasserstein generative adversarial network, which is used to enhance training data through distribution fusion. A discrete wavelet decomposition algorithm is employed to extract the low-frequency wavelet coefficients from the original motor current signals. These are used to train the multi-layer perceptron Wasserstein generative adversarial model. Then, the trained model is applied to generate fake current wavelet coefficients with the fused distribution. A motor fault classification model consisting of a feature extractor and pattern recognizer is built based on perceptron. The data augmentation experiment shows that the fake dataset has a larger distribution than the real dataset. The classification model trained on a real dataset, fake dataset and combined dataset achieves 21.5%, 87.2%, and 90.1% prediction accuracy on the unseen real data, respectively. The results indicate that the proposed data augmentation method can effectively generate fake data with the fused distribution. The motor fault classification model trained on a fake dataset has better generalization performance than that trained on a real dataset.

KEYWORDS

Motor fault diagnosis; data augmentation; wavelet decomposition; generative adversarial network; civil and building engineering

1 Introduction

Electrical equipment is widely used in civil and building engineering. The application of large electrical equipment can greatly improve construction ability and efficiency. The motor is one of the main power components of this electrical equipment, which can convert electrical energy into mechanical energy. During the long-term service of electrical equipment, aging, wear, and fatigue damage of motor parts may cause various motor faults. Motor faults will change the operation state of equipment and cause adverse effects on normal construction. Therefore, in the service of engineering

equipment, engineers need to conduct health monitoring and implement maintenance for motors regularly. It has great significance for ensuring safety and improving the efficiency of construction.

Motor monitoring is to obtain the signal of the running motor, then analyze and extract the signal characteristics, and finally judge the health status of the motor according to the characteristics. The traditional methods of motor monitoring are to set thresholds for motor running parameters and then determine the real motor states according to the relationship of both the parameters and the thresholds [1]. These methods can detect many kinds of motor faults related to the macroscopic characteristics of the operating parameters (current value, voltage value, temperature, etc.) [2]. However, multiple motor faults may exhibit the same macroscopic characteristics, which makes it difficult to determine the specific fault type. In addition, some motor faults can't be shown in the macroscopic characteristics in the early stage of development which leads to the failure of detection.

With the development of information technology, fault diagnosis methods based on detailed features are more and more popular in motor monitoring [3,4]. By analyzing detailed features and deep features of motor running signal, these methods can identify small changes in motor states. In motor condition monitoring, compared with vibration, torque, sound and other signals, the current signal can be obtained directly from the motor control system, which has the characteristics of easy acquisition and low acquisition cost [5]. At the same time, the current signal is directly related to the input and output characteristics of the motor and contains abundant motor state information [6]. Machine learning technology has been used in motor condition monitoring by many researchers. Motor fault diagnosis method based on machine learning technology has a strong ability for feature extraction and data classification. Some researchers use neural networks to extract deep features from motor signals to achieve motor fault classification [7].

In the fault diagnosis based on neural network, researchers often encounter the situation that the experiments can only obtain data under a few working conditions because of the expensive cost on time and economic. At the same time, the data in the actual working environment is usually not easy to obtain due to environmental constraints. So researchers can only obtain a small amount of valid data and the feature distribution of samples is relatively narrow, which corresponds to limited working conditions. Deep learning model based on neural network needs a large amount of data for training to get the optimal parameters. Thus, this kind of training data is not conducive to the training of neural network model. Some studies have proposed that data augmentation methods can be used to generate fake data, which has similar feature distribution compared with the real data. And the fake data can be used for the training of neural network models. The generative adversarial network was used to generate realistic one-dimensional raw data [8]. Five data augmentation methods were selected for training a contrastive representation learning network [9]. In order to increase the number of data samples, Meng et al. [10] divided a single sample into multiple monomers and then recombined the monomers. Wang et al. [11] enhanced the resolution of the original sample for data augmentation.

In order to solve the above problems and promote the application of machine learning technology in motor fault diagnosis based on data augmentation and neural network technology, this paper proposes a motor fault diagnosis method which is suitable for a small dataset and a narrow range of data characteristics. The method constructs a current data generation model based on an improved generative adversarial network to generate the fake training data. The discrete wavelet decomposition algorithm is used to decompose the current data, and the low frequency wavelet coefficients are selected to construct datasets. A motor fault classification model based on multi-layer perceptron neural network is designed, which is applied to classify the motor current data and realize the motor fault diagnosis. The main contributions of this paper are as follows:

(1) Aiming at the characteristics of limited computing resources in the building engineering site, an improved generative adversarial network suitable for the small dataset is constructed. The network can generate fake data with fused distribution under virtual working conditions, which increases the diversity of the training dataset.

(2) A classification model for motor fault diagnosis is designed for deep feature extraction and classification. The low frequency coefficients of wavelet decomposition of motor current data are used as training data. This model can be trained on a personal computer, which is conducive to the deployment and implementation in various building engineering sites.

(3) This paper presents a motor fault diagnosis method based on motor current data, which realizes end-to-end fault diagnosis. This method does not need a large number of previous experimental data. It can obtain a diagnostic model with good generalization performance by using a small amount of current data under discrete working conditions.

The rest of this paper is organized as follows: Section 2 summarizes the research status of related technologies and methods. Then the motor fault diagnosis method based on data augmentation technology and neural network technology is described in Section 3. In Section 4, experiments of data augmentation and fault classification are carried out, and experimental results are shown and discussed. Finally, the conclusion is presented in Section 5.

2 Related Work

2.1 Data Augmentation Methods

In computer vision research, many studies use image augmentation methods such as image space transformation and noise superposition, to generate fake images to support the training of large neural network models [12]. For time series data, the simplest data augmentation method is window overlap. But this method does not generate new feature distributions and does not increase the diversity of data. It only superimposes the data features of several samples [13,14]. The noise superposition is to add the background noise of working conditions to the original sample, which increases the diversity of data and improves the generalization performance of the trained model. However, this method does not enrich the original basic feature information and cannot effectively improve the performance of the model [15]. Amplitude offset and time warping methods are also widely used in time series data tasks and have achieved some results, but they are the linear transformations of original data [16]. Yu et al. proposed a data augmentation method which is composed of several data augmentation strategies. The research results show that the method has achieved good results on small sample datasets [17].

In recent years, data augmentation technology based on the generative adversarial network has attracted much attention. This method can learn the feature distribution of real data through unsupervised training. The trained generator can randomly generate fake data, which is similar with the original data. The method can generate data with the new feature on the basis of the main feature distribution of original real data. According to the characteristics of application background, researchers have designed many types of generative adversarial networks. A deep generative adversarial network was proposed to address the imbalanced data problem be explicitly creating additional training data [18]. Cabrera et al. [19] artificially balanced the dataset based on improved generative models with a dissimilarity-based model selection.

2.2 Motor Fault Diagnosis Methods Based on Current Signal

2.2.1 Diagnosis Methods Based on Time Domain and Frequency Domain Feature of Current

The stator phase current collected by the current sensor contains rich motor state information in the time domain. Researchers observe the current signal characteristics in the time domain, and construct the mapping relationship between signature and motor fault to realize the motor fault diagnosis.

Yang et al. [20] used the covariance of three-phase current to diagnose and locate transistor open-circuit faults by analyzing the correlation between them. Naseri et al. [21] employed three Kalman filters to estimate the three-phase motor currents, and used averaged normalized residual signals as diagnostic criteria for the detection of open-switch faults. Kou et al. [22] presented an open-circuit fault diagnosis approach based on Concordia transform and random forests technique. Eldeeb et al. [23] implemented a mathematical morphological gradient technique in the online diagnosis of stator's interturn failure. Salehifar et al. [24] defined the cross-correlation factor as the fault detection index, which is the similarity between estimated current and real current. The estimated current is formed by a model-based observation. Yan et al. [25] calculated the fault symptom variables by using average current Park's vector method, and then applied fuzzy logic approach to process the faulty symptom variables.

The method based on time domain analysis can process some simple motor current signals and obtain signature components. But when the current signal contains a lot of noise, the applicability of time-domain analysis method is poor. Transforming the current signal from the time domain to the frequency domain or time-frequency domain can greatly reduce the amount of data, which is conducive to highlighting the signal's key frequency information.

Potamianos et al. [26] used discrete wavelet transform to analyze a monitoring system's measured output current waveform. Zaman et al. [27] used Wavelet Packet Decomposition to extract the features by evaluating energy eigenvalues and feature coefficients at decomposition levels from stator current signals. Shabbir et al. [28] employed Discrete Fast Fourier Transform to obtain the harmonic spectra and chose the fundamental and 5th harmonic of stator current as signature frequency components. Malik et al. [29] extracted the features by empirical mode decomposition, and used Modified-Fuzzy-Q-learning technique for turbine fault diagnosis. Rabbi et al. [30] performed the multiresolution analysis of nonstationary current signals to detect and extract the sideband frequency components based on wavelet packet decomposition.

The frequency domain or time-frequency domain analysis method of current can accurately locate the signature components corresponding to motor faults, which helps understand the evolution process of faults.

2.2.2 Diagnosis Methods Based on Machine Learning Technology

The analysis methods of the current signal based on machine learning technology get the optimal signal feature representation mode through repeated iteration, and then classifies the extracted features to realize the fault diagnosis. Kou et al. [22] applied Concordia transform to process the fault samples, and the transformed fault samples were used to train the fault diagnosis classifier based on random forests. Long et al. [31] employed an attention mechanism and improved the AdaBoost multi-classification classifier in motor fault diagnosis. In processing complex signals, artificial neural networks usually have stronger feature extraction ability than ordinary machine learning models. In

recent years, many researchers have applied the artificial neural network model in motor fault diagnosis based on motor current.

Skowron et al. [32] proved the possibility of using a convolutional neural network in the real-time diagnostic system with the high accuracy of incipient stator winding fault detection and classification. Munikoti et al. [33] used three kinds of statistical machine learning methods to address motor fault detection and identification and revealed that the convolutional network performs better than support vector machine and recurrent networks. Masrur et al. [34] developed and trained a structured neural network system to detect and isolate the most common types of faults. An unsupervised domain-share convolutional neural network is constructed to extract the domain invariant features for efficient fault transfer diagnosis of machines from steady speeds to time-varying speeds [35]. Hussain et al. [36] trained and tested three deep learning models MLP, 1DCNN, and LSTM for the purpose of healthy and unhealthy conditions of the induction motor.

The diagnosis method based on the machine learning model has obvious advantages in the representation of complex signal features through appropriate training and testing. In recent years, researchers have proposed a large number of effective fault diagnosis methods based on big industrial data and neural networks. These methods play an important role in equipment health monitoring in the fields of energy, aerospace and so on [37].

3 Proposed Methodology

In the application of motor fault diagnosis based on the neural network model, researchers often only obtain the current data of the motor under a few working conditions. This kind of data has centralized characteristics and small distribution range, which makes it difficult to train a neural network model with good generalization performance. In the actual operating environment of the motor, the speed, load and ambient temperature change at any time, which will cause the current signal of the motor to have complex characteristics. If the limited experimental data is used to train the fault diagnosis model, the model is used to diagnose motors in the actual environment, which must show very poor performance [38].

This paper proposes a motor fault diagnosis method based on generative adversarial learning that aims at the above problems. Firstly, the wavelet decomposition algorithm is used to extract the low-frequency wavelet coefficients of the current data, which forms a real dataset. Then a multi-layer perceptron Wasserstein generative adversarial network is designed to generate fake data, which has more complex data characteristics and wider data distribution than real data. The motor fault classification model consists of an extractor and recognizer to realize the motor fault classification. The architecture of the proposed method is shown in Fig. 1.

3.1 Preprocessing of Current Signal Based on Wavelet Decomposition

The main components of motor stator current are low frequency signals. When a motor fault occurs, the overall low-frequency characteristics of the current signal will change. And the fault characteristic frequency is near the power frequency [39]. Therefore, the fault types of the motor can be judged by analyzing the frequency characteristics of the low-frequency component in the current signal [40].

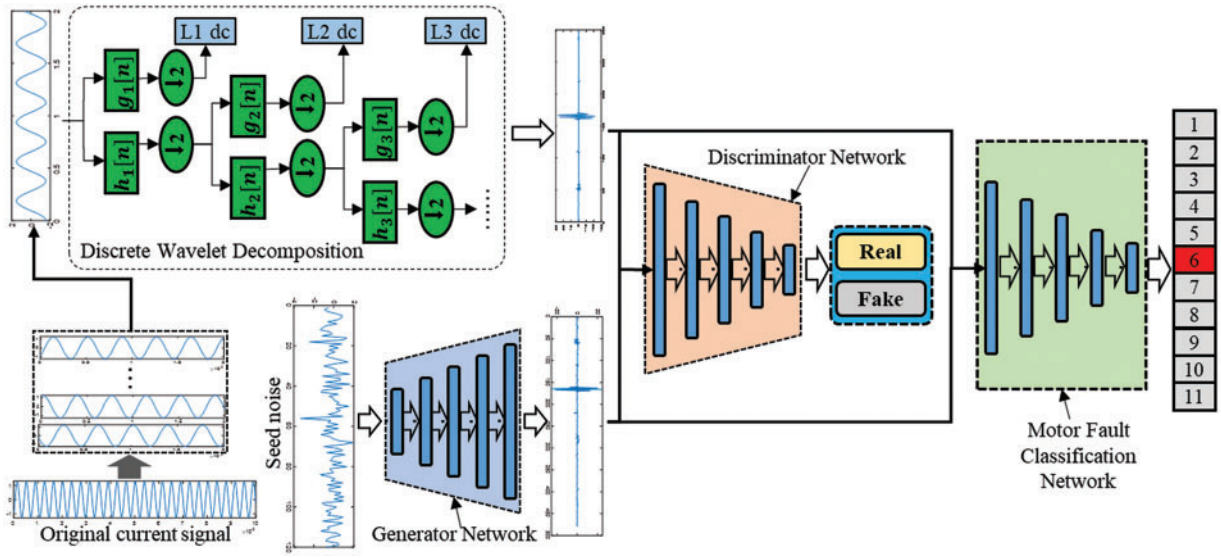

Figure 1: Architecture of the proposed motor fault diagnosis method

The discrete wavelet decomposition algorithm can transform time domain data to frequency domain data. The frequency domain data decomposed by the algorithm has a dynamic resolution, which meets the requirement of signal decomposition [41]. The discrete wavelet decomposition algorithm can be expressed as Eq. (1), where $x(n)$ represents time-domain discrete signal and $\varphi_{j,k}(n)$ represents dyadic wavelet.

$$\begin{aligned} D_{\varphi}X(j,k) \triangleq \left(DWT_{\varphi}x\right)(2^{j},k2^{j}) &= \sum_{n=-\infty}^{+\infty} x(n)\,\overline{\varphi_{j,k}(n)} \\ &= 2^{-\frac{j}{2}} \sum_{n=-\infty}^{+\infty} x(n)\,\overline{\varphi(2^{-j}n-k)}\,(j,k,n\in Z) \end{aligned} \tag{1}$$

In this paper, the Mallet algorithm is used to realize discrete wavelet decomposition of motor current signal, which is suitable for analyzing motor current data [42]. It is assumed that the number of data points of the discrete current signal is N, and the highest frequency of the signal is Fs. The method comprises the following steps: firstly, perform wavelet decomposition on the first level: constructing a high-pass filter $g_1[n]$ with a filtering frequency of $\left[\frac{Fs}{2}, Fs\right]$ and a low-pass filter $h_1[n]$ with a filtering frequency of $\left[0, \frac{Fs}{2}\right]$, where n represents the serial number, and then performing two times down-sampling on the filtered data to obtain low-frequency wavelet coefficients and high-frequency wavelet detail coefficients (L1 dc), respectively; secondly, perform decomposition on the second level: keeping the high-frequency coefficients unchanged, and decomposing the low-frequency coefficients again to obtain low-frequency and high-frequency coefficients, respectively; after that, the decomposition is carried out m times to obtain Wavelet coefficients in frequency domain interval of $\left[\frac{Fs}{2}, Fs\right], \left[\frac{Fs}{4}, \frac{Fs}{2}\right], \ldots, \left[\frac{Fs}{2^{m}}, \frac{Fs}{2^{m-1}}\right], \left[0, \frac{Fs}{2^{m}}\right]$

The characteristics of motor faults mainly distribute in the low-frequency wavelet coefficients of the current signal. Therefore, the high frequency coefficients will be discarded, and the multiple layers of low frequency wavelet coefficients are spliced into one-dimensional data for motor fault classification.

3.2 Data Augmentation Method of Current Signal

3.2.1 Multilayer Perceptron Wasserstein Generative Adversarial Network

A generative adversarial network was proposed in recent years. By using adversarial learning between the generator and the discriminator, the generator can learn the data distribution of the training dataset. In this paper, a multi-layer perceptron Wasserstein generative adversarial network is proposed, which can be expressed as Eq. (2) [43], where G is the discriminator model for judging the type of sample, and D is the generator model for generating fake data, and $\boldsymbol{z}$ is the seed noise, $\boldsymbol{x} \in \mathbb{R}^n$ are real samples, $\widetilde{\boldsymbol{x}} \in \mathbb{R}^n$ are generated samples, $\boldsymbol{y} \in \mathbb{R}^2$ are the decision values of the discriminator, $\boldsymbol{\theta}^{\mathrm{G}}$ and $\boldsymbol{\theta}^{\mathrm{D}}$ are the generator and discriminator model parameters, respectively.

$$\begin{cases} \mathrm{G}: \widetilde{\boldsymbol{x}} = \mathrm{G}\left(\boldsymbol{z}, \boldsymbol{\theta}^{\mathrm{G}}\right) \\ \mathrm{D}: \boldsymbol{y} = \begin{bmatrix} \mathrm{D}(\boldsymbol{x}, \boldsymbol{\theta}^{\mathrm{D}}) \\ \mathrm{D}\left(\widetilde{\boldsymbol{x}}, \boldsymbol{\theta}^{\mathrm{D}}\right) \end{bmatrix} \end{cases} \tag{2}$$

The structure of the network is shown in Fig. 1, in which the discriminator is composed of multi-layer perceptron neural networks, and the activation functions of the former four layers are ReLu, and the last layer does not use the activation function. The generator is also composed of multi-layer perceptron neural networks, in which the activation functions of the former four layers are ReLu, and the last layer does not have an activation function, and a discarded layer is added after the third layer to prevent the model from overfitting. The discriminator and generator have the same framework of neural networks. Suppose there are m neurons in the layer $l-1$ and n neurons in the layer l of the networks. The weight coefficients matrix $\boldsymbol{W}^l$ and the bias matrix $\boldsymbol{b}^l$ of the layer l are expressed as Eq. (3).

$$\boldsymbol{W}^l = \begin{pmatrix} w_{11}^l & w_{12}^l & \cdots & w_{1m}^l \\ w_{21}^l & w_{22}^l & \cdots & w_{2m}^l \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}^l & w_{n2}^l & \cdots & w_{nm}^l \end{pmatrix}, \boldsymbol{b}^l = \begin{pmatrix} b_1^l \\ b_2^l \\ \vdots \\ b_n^l \end{pmatrix} \tag{3}$$

The output vector of the layer $l-1$, linear output vector of the layer l, and output vector of the layer l are expressed as Eq. (4).

Then the output of the layer l can be expressed in matrix as Eq. (5), where σ is the activation function.

From this, the forward propagation algorithm of multi-layer perceptron can be obtained:

Step 1: initialize $\boldsymbol{a}^l = \boldsymbol{x}_0, l = 1$;

Step 2: let $l = l + 1$, and calculate $\boldsymbol{a}^l$;

Step 3: repeat Step 2, until $l = L$, and output $\boldsymbol{a}^L$, where L is the number of layers of multi-layer perceptron.

$$\boldsymbol{a}^{l-1} = \begin{pmatrix} a_1^{l-1} \\ a_2^{l-1} \\ \vdots \\ a_m^{l-1} \end{pmatrix}, \boldsymbol{z}^l = \begin{pmatrix} z_1^l \\ z_2^l \\ \vdots \\ z_n^l \end{pmatrix}, \boldsymbol{a}^l = \begin{pmatrix} a_1^l \\ a_2^l \\ \vdots \\ a_n^l \end{pmatrix} \tag{4}$$

$$\boldsymbol{a}^l = \sigma(\boldsymbol{z}^l) = \sigma(\boldsymbol{W}^l \boldsymbol{a}^{l-1} + \boldsymbol{b}) \tag{5}$$

3.2.2 Loss Function

Traditional generative adversarial learning uses Kullback-Leibler Divergence to measure the distribution distance between generated data and real data. Many studies have found that Kullback-Leibler Divergence leads to some problems, such as gradient disappearance, training gradient instability and pattern collapse in the training process [44], which makes the training of generative adversarial model very difficult.

Wasserstein distance is used in the loss function of the generative adversarial model, which is often used to measure the distance of two distributions. The loss function of the discriminator can be presented as Eq. (6) [45], where E is the expectation function, P_{r} is the distribution of real data, P_{g} is the distribution of fake data.

$$\mathrm{V}(\mathrm{G}, \mathrm{D}) = \max_{\mathrm{D} \in 1-\mathrm{Lipschitz}} \{\mathrm{E}_{x-P_{\mathrm{r}}}[\mathrm{D}(x)] - \mathrm{E}_{x-P_{\mathrm{g}}}[\mathrm{D}(x)]\} \tag{6}$$

The 1-Lipschitz constraint condition in Eq. (6) is equivalent to that the gradient of the discriminator is less than 1 everywhere, which can be expressed as Eq. (7). Then, the loss function can be expressed as Eq. (8). The $\lambda \int_x \max(0, ||\nabla_x \mathrm{D}(x)|| - 1)\mathrm{d}x$ in the equation is regard as a regular term. When the gradient of the discriminator is greater than 1, $\lambda \int_x \max(0, ||\nabla_x \mathrm{D}(x)|| - 1)\mathrm{d}x$ will be deducted from $\max_{\mathrm{D}}\{E_{x-P_{\mathrm{data}}}[\mathrm{D}(x)] - E_{x-P_{\mathrm{G}}}[\mathrm{D}(x)]\}$ to ensure that the gradient is less than 1.

$$\mathrm{D} \in 1 - \mathrm{Lipschitz} \Leftrightarrow ||\nabla_x \mathrm{D}(x)|| \leq 1 \text{ for all } x \tag{7}$$

$$\mathrm{V}(\mathrm{G}, \mathrm{D}) \approx \max_{\mathrm{D}}\{\mathrm{E}_{x-P_{\mathrm{data}}}[\mathrm{D}(x)] - \mathrm{E}_{x-P_{\mathrm{G}}}[\mathrm{D}(x)] - \lambda \int_x \max(0, ||\nabla_x \mathrm{D}(x)|| - 1)\mathrm{d}x\} \tag{8}$$

This requires that the gradient corresponding to all inputs of discriminator D should be less than 1. In order to approximately satisfy the condition, it is necessary to define a penalty sample space, which is smaller than the whole data space. The distribution of the data from the defined space is P_{penalty}. So the problem is transformed into that the gradients $\nabla_x \mathrm{D}(x)$ corresponding to x in P_{penalty} are all less than 1. Further, the cost function of the discriminator can be modified as Eq. (9) [46], where λ is the gradient penalty coefficient.

$$\mathrm{V}(\mathrm{G}, \mathrm{D}) \approx \max_{\mathrm{D}}\{\mathrm{E}_{x-P_{\mathrm{r}}}[\mathrm{D}(x)] - \mathrm{E}_{x-P_{\mathrm{g}}}[\mathrm{D}(x)] - \lambda \mathrm{E}_{x-P_{\mathrm{penalty}}}[(||\nabla_x \mathrm{D}(x)|| - 1)^2]\} \tag{9}$$

Some research has found that it is better to set the P_{penalty} to the space between the spatial distribution of generated data and the spatial distribution of real data [47]. In this paper, the gradient descent method is used to solve the maximum of the loss function of the discriminator. Therefore, the

loss function can be finally expressed as shown in Eq. (10).

$$L_D = E_{x\sim P_g}[D(x)] - E_{x\sim P_r}[D(x)] + \lambda E_{x\sim P_{penalty}}[(||\nabla_x D(x)||_2 - 1)^2] \quad (10)$$

The loss function of the generator can be expressed as Eq. (11).

$$L_G = E_{x\sim P_g}[D(x)] \quad (11)$$

3.2.3 Training Method

In this research, the generator and the discriminator have trained alternately. In every cycle, the discriminator is trained for k steps first, and then the generator is trained for only one step. The purpose of the training of the generator is to enable the generator to generate fake data similar to real data. At this time, the discriminator is used to identify the authenticity of the generated data. The purpose of the training of discriminator is to enable the discriminator to measure the similarity between the real data and fake data more accurately. At this time, the generator is used to generate fake data. So the model parameters will be updated circularly until the two models converge to the predetermined target based on the training method. It is found that using the gradient descent algorithm based on momentum will make the loss of the discriminator jitter significantly, and using the RMSProp algorithm will not cause this problem [48]. So the generative adversarial model is trained based on the RMSProp optimizer.

3.3 Motor Fault Classification Method

One-dimensional wavelet coefficients of the motor current signal obtained by wavelet decomposition contain abundant motor operating information. In this paper, the motor fault classification model is a multi-layer perception neural network as shown in Fig. 1. The layers before the last layer of the network are used as a feature extractor to extract deep features. The last layer of the network is used as a pattern recognizer to classify the features into different faults. ReLu activation function is used for every layer of the feature extractor. And Softmax activation function is used for pattern recognizer to output the probability of various fault types. Cross-entropy loss function is used as the model loss function, and Adam optimizer is used as the learning rate optimizer.

4 Experiment

4.1 Dataset

The motor stator current dataset used in this paper contains 11 kinds of motor faults, which is simulated by artificially manufactured motor components defects. The experiment is carried out in laboratory environment under 8 kinds of simulated working condition. The data duration of each working condition is 10 s. The electric current's signals are measured with a current probe and an oscilloscope on two phases denoted φ_l, with $l = 1, 2$. The range of measurement is 5 A with an accuracy of 1% at a sampling frequency of 100 kHz [49]. The fundamental frequency of the motor current is 28 Hz. The maker of the open dataset only discloses the current data of the motor, but does not give the specific fault type and working condition description. This research focuses on the augmentation and classification of motor current data without prior knowledge, which makes the method have better applicability. In other words, the study realizes the classification of various current data by directly analyzing the statistical characteristics of data.

Fig. 2 shows the motor current wave and its single-side amplitude spectrum. On the whole, the time-domain waveform of the current signal is a regular sinusoidal curve, and does not have obvious fault characteristics. From the local view, irregular curves and a variety of structural features can be

observed from the time-domain waveform. It can be found that although there are differences in motor current signal waveforms of different faults, these differences are relatively small. Therefore, it is difficult to distinguish the specific motor state directly from the motor current waveform. There is a prominent peak in the frequency spectrum, and the center frequency of the peak is equal to the fundamental frequency of the current. This indicates that the signal energy in the current signal is concentrated around the fundamental frequency. But there are some differences in the shape of this peak in the current spectrum under various faults, which indicates that the frequency characteristics of various faults are different. Specifically, these differences are mainly reflected in the side frequencies near the main frequency. Some studies have pointed out the specific relationship between the side frequency of the motor current signal and motor fault, and motor fault can be judged according to the position and amplitude of side frequency [50,51]. The amplitude of the high frequency component in the spectrogram is very small.

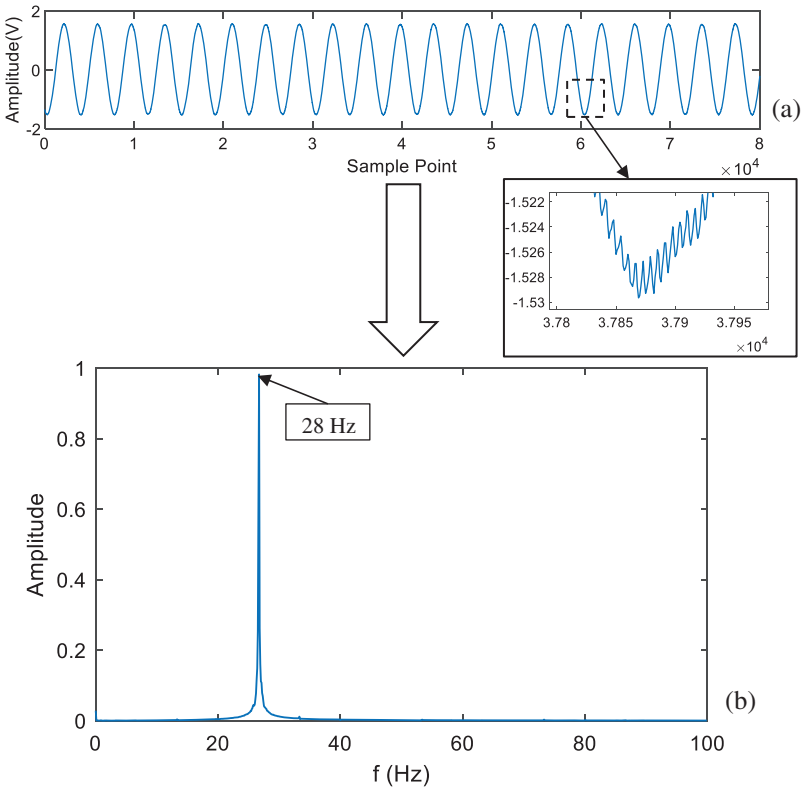

Figure 2: Motor current signal (a) and its single-side amplitude spectrum (b)

4.2 Experiments Setup

4.2.1 Setup of Data Augmentation

The effective signal components in the current signal are mainly distributed around the fundamental frequency of the current. In order to reduce the amount of sample data, the current signal is filtered first. The filtering method based on the Fourier transform is used in this study. First, Fourier transform is applied to the original signal to obtain the signal spectrum. Then, set the signal frequency band to be filtered to zero on the spectrum. Finally, the inverse Fourier transform is performed on the spectrum to obtain the filtered signal. The effective frequency of the signal is reduced from 50,000 to

5000 Hz. In the experiment, the current data is truncated with 0.25 s, and then the truncated data is decomposed using a discrete wavelet decomposition algorithm. The coefficients of the motor current signal are shown in Fig. 3.

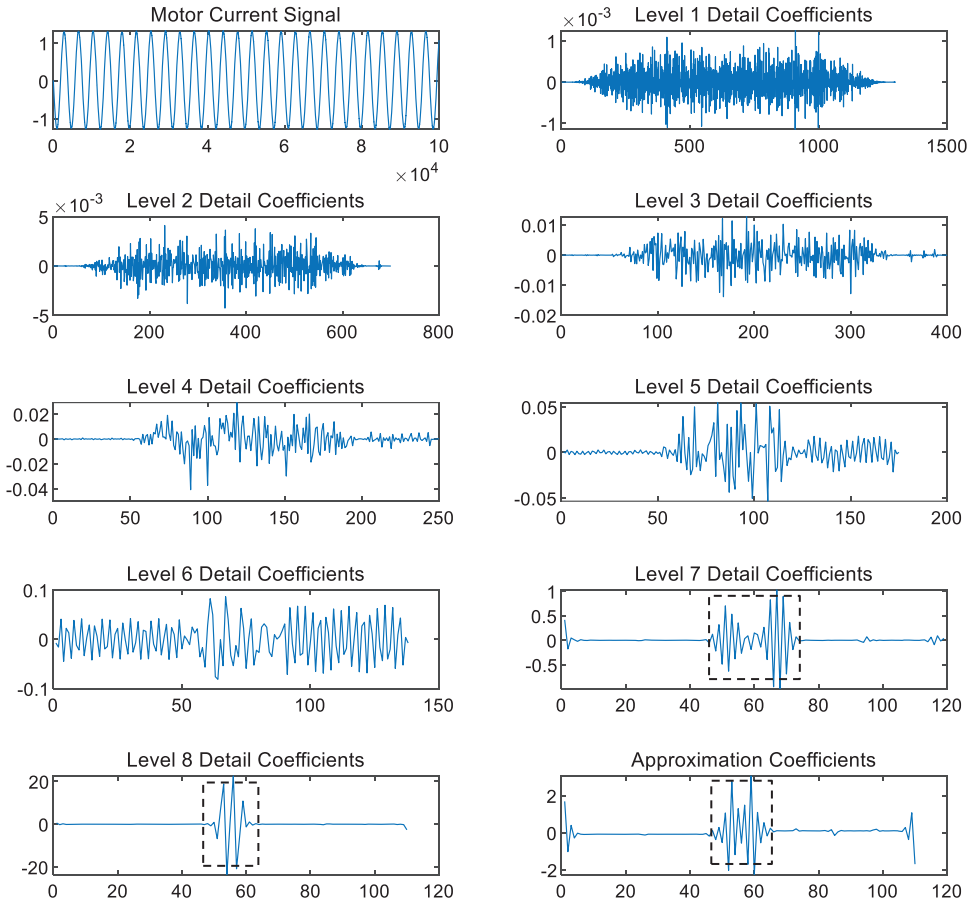

Figure 3: Discrete wavelet decomposition coefficients of current signal

It can be seen that the amplitudes of the former five layers of detail wavelet coefficients are all less than 0.05. And no obvious signal characteristic components can be observed in them, which indicates that these signal components are noise. The wavelet coefficients of the last three layers have large amplitudes, and significant signal components can be observed from them (shown in the wireframe). The signal frequency ranges corresponding to the detail coefficients of level 6–8 are 78–156, 39∼78, 19.5∼39, respectively. The frequency range of the approximation coefficients is 0∼19.5. This research focuses on the analysis of the current signal components, which is in the range of five times the fundamental frequency of the current. The last 3 layers of detail coefficients and the approximation coefficients are extracted to form the training data.

In this paper, a corresponding generative adversarial model is established for each kind of motor fault, which is trained using the training dataset of this motor fault. The training algorithm of the proposed model is shown in Table 1. Python and PyTorch machine learning architecture are used to build the model. This research chooses the number of neurons of the neural network based on the principle of minimizing the model training time. Therefore, on the basis of ensuring the performance of

the model, the perceptron network is constructed with as few neurons as possible to reduce the amount of computation. The number of neurons in the first layer of the network is first determined based on the complexity of training data, and the number of neurons in the following layers is calculated layer by layer with a magnification of 0.5. The generator model consists of three layers of perceptron neural networks, and the number of neurons in each layer is 120, 240, and 480, respectively. The discriminator model consists of four layers of perceptron neural networks, and the number of neurons in each layer is 120, 60, 30 and 1, respectively. The number of training rounds for the model was set to 12,000, and the learning rate was set to 1 e–4. In each round, the discriminator parameters are updated 5 times and the generator parameters are updated 1 time.

Table 1: Algorithm of the proposed data augmentation model

Algorithm: multi-layer perceptron Wasserstein generative adversarial model training algorithm.
All experiments in the paper use the following values $n_{\text{critic}} = 5$, $\alpha = 0.0001$, $\lambda = 0.3$.

Require: The gradient penalty coefficient λ, the number of critic iterations per generator iteration n_{critic}, RMSProp hyper parameters α, the batch size b.
Require: Initial discriminator parameters θ_{D}, initial generator parameters θ_{G}.
1: while θ_{G} has not converged **do**
2: **for** $i = 1, \ldots, n_{\text{critic}}$ **do**
3: **for** $j = 1, \ldots, m$ **do**
4: Sample real data $x \sim P_{\text{r}}$, latent variable $z \sim P_{\text{l}}$
5: $\widetilde{x} \leftarrow \text{G}_{\theta_{\text{G}}}(z)$
6: $\text{L}_{(j)} \leftarrow \text{D}_{\theta_{\text{D}}}(\widetilde{x}) - \text{D}_{\theta_{\text{D}}}(x) + \lambda \left(\left|\left| \nabla_{\tilde{x}} \text{D}_{\theta_{\text{D}}}(\widetilde{x}) \right|\right|_2 - 1 \right)^2$
7: **end for**
8 : $\theta_{\text{D}} \leftarrow \text{RMSProp}\left(\nabla_{\theta_{\text{D}}} \frac{1}{b} \sum_{j=1}^{b} \text{L}_{(j)}, \theta_{\text{D}}, \alpha \right)$
9: **end for**
10: Sample a batch of latent variables $\{z_{(j)}\}_{j=1}^{b} \sim P_{\text{l}}(z)$.
11 : $\theta_{\text{G}} \leftarrow \text{RMSProp}\left(\nabla_{\theta_{\text{G}}} \frac{1}{b} \sum_{j=1}^{b} - \text{D}_{\theta_{\text{D}}}(\text{G}_{\theta_{\text{G}}}(z)), \theta_{\text{G}}, \alpha \right)$
12: end while

Fig. 4 shows the fake wavelet coefficients of fault 11 generated by the generator in different model training rounds. It can be seen from the figure that as the training round increases, the fake data gradually shows some data characteristics of the real data. When the round is more than 10,000, the data tends to be stable. It indicates the model has converged. By observing the training process of different fault samples, it can be found that the model in this paper can quickly capture the main signal characteristics of the signal, as shown in the 2000 round signal in Fig. 4. But the detailed features are difficult to obtain, which needs more than 10,000 training rounds. This shows that the signal details under multiple working conditions are complex. In order to improve the distribution of fake datasets and improve the efficiency of model training, this research adjusts the hyper parameters according to the test results on the validation dataset.

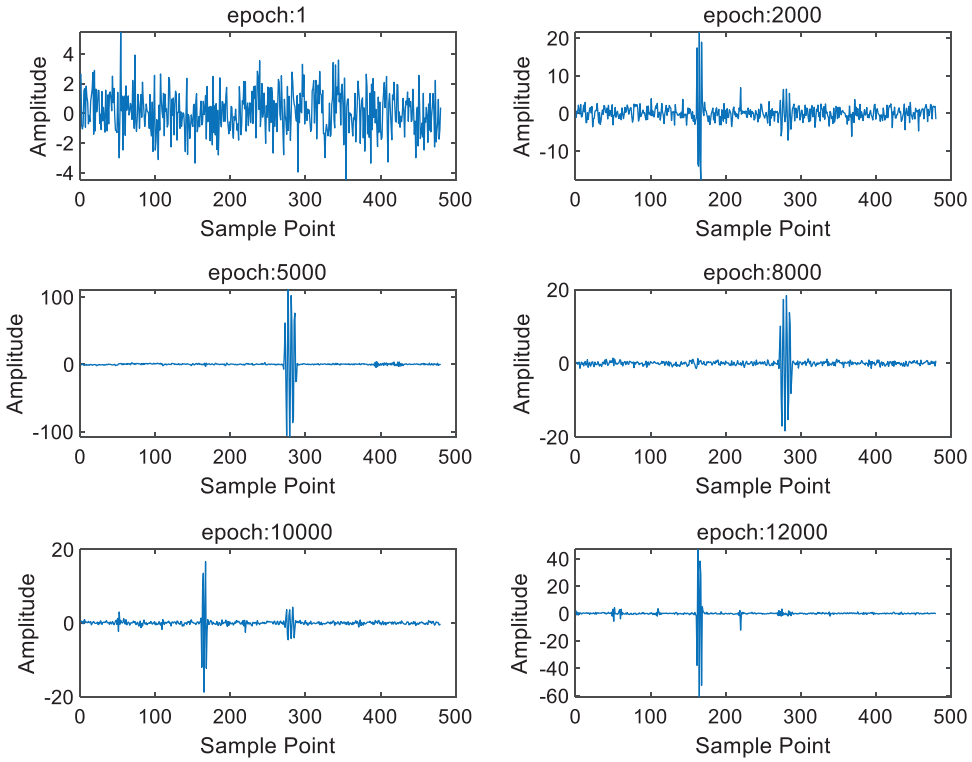

Figure 4: Fake wavelet decomposition coefficients in different training epoch

After the trained multi-layer perceptron Wasserstein generative adversarial model is obtained, the generator is used to generate fake wavelet coefficients. For different faults (a, b, c), the real and fake wavelet coefficients under the same working condition are shown in Fig. 5. It can be seen from the figure that the wavelet coefficients under different faults have different characteristics. The fake data and real data under the same fault have similar characteristics. The wavelet coefficients of different faults have some similar characteristics. These characteristics mainly come from the inherent physical characteristics of the motor system and environmental noise, such as the current noise of the power grid and the inherent vibration of the motor. In addition, the wavelet coefficients of the low-frequency part are taken as the characteristic signal, and the wavelet coefficients of the high-frequency part are discarded, which leads to fewer characteristic components of the signal.

4.2.2 Setup of Motor Fault Classification

In this paper, Python and PyTorch machine learning architecture are used to build a multi-layer perceptron neural model for motor fault classification. The model consists of five layers of perceptron neural networks, and the number of neurons in each layer is 240, 120, 60, 30 and 11, respectively. The model is trained with the real dataset, fake dataset and combined dataset under 5 kinds of working conditions, respectively. And then, the models are tested with real data under another 3 kinds of working conditions. The combined dataset consists of real data and fake data. The composing proportion of the dataset will affect the performance of the classification model. An experiment about the relationship between the proportion of fake data in the combined dataset and the performance of motor fault classification model was carried out to obtain the best proportion.

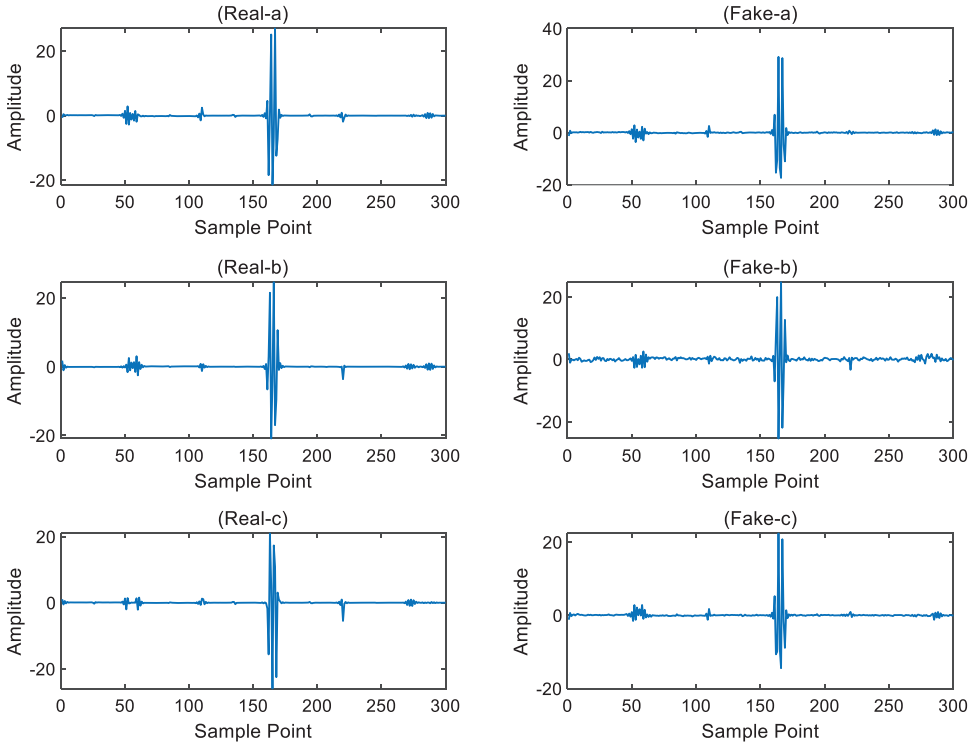

Figure 5: Real and fake wavelet decomposition coefficients under different faults (a, b, c)

In order to determine the best number of model training rounds, this manuscript takes the number as a variable to explore the relationship between the rounds and training loss and test accuracy. In the experiments, when the number of training rounds reaches about 2500, the training loss and test accuracy of the model do not change, which indicates that the model has converged to the optimal state. Therefore, the number of rounds of model training is set to 2500. The experimental loss function is the cross-entropy loss function, the optimizer is the Adam optimizer, and the learning rate is set to 1 e−4.

To validate the effectiveness of the designed method, comparisons among the proposed method and other four methods are conducted. The four methods include random vector functional link network (RVFLN) [52], convolutional neural network-based detector (CNND) [32], dynamic time warping (DTW) [53], SVM multi-class classification (SVMMCC) [54]. RVFLN uses fast learning technology to learn faulty knowledge from the historical dataset based on feature extraction and feature selection. CNND is a convolutional neural network model. DTW is a time domain based method, which is used to suppress the supply frequency component and highlight the sideband components based on the introduction of a reference signal. SVMMCC uses one against all strategies to classify motor faults.

4.3 Results and Discussion

4.3.1 Result

The result of the experiment about composing proportion of the combined dataset is shown in Fig. 6. The result shows that the classification model has the best performance when the composing

proportion is set to about 0.5. Based on the optimal proportion, the training loss of the models is obtained, as shown in Fig. 7. It can be found that the training loss values of the models trained with the fake dataset and combined dataset are 0.11 and 0.15, respectively, when the number of training epoch is 100. The model trained with the real dataset has a training loss of 0.08 when the number of training epoch is 18. The prediction accuracy of the three trained models on unseen real data is shown in Table 2. The model trained on the combined dataset achieves the best test accuracy of 90.1%.

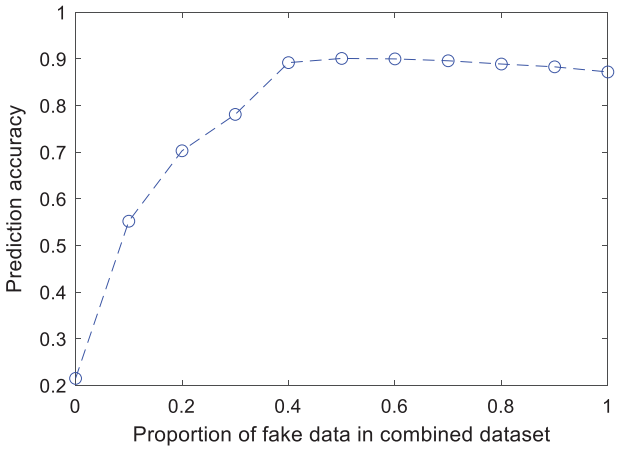

Figure 6: Prediction accuracy of the proposed classification model trained with combined dataset

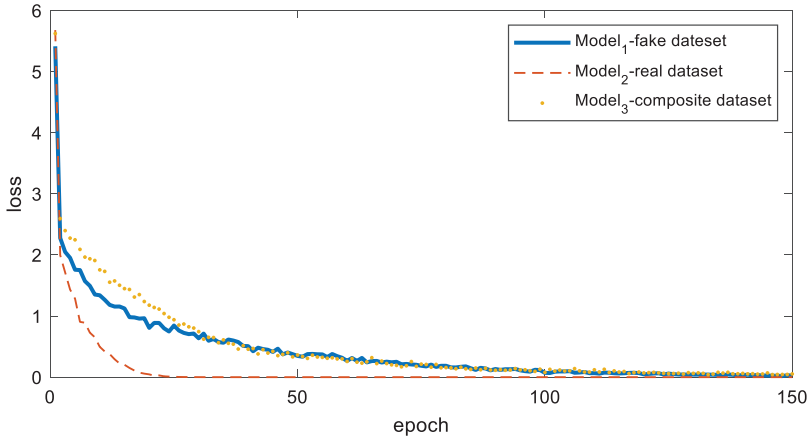

Figure 7: Training loss of motor fault classification models

Table 2: Prediction accuracy of the trained motor fault classification models

Training dataset	$Model_1$ -fake dataset	$Model_2$ -real dataset	$Model_3$ -combined dataset
Prediction accuracy	87.2%	21.5%	90.1%

Further, this study conducted experiments to compare the generalization performance of the four models described in Section 4.2.2 and the proposed model. The test accuracy of the models is shown in Table 3. Compared with other methods, the proposed method achieves the highest test accuracy of 89.7%. This shows the generalization performance of the proposed method is better than the other four methods.

Table 3: Test accuracy of the five motor fault classification methods

Method	RVFLN	CNND	DTW	SVMMCC	Proposed method
Prediction accuracy	72.0%	86.7%	73.2%	84.3	89.7%

In order to compare the distribution of fake data and real data, the datasets are input into the trained fault classification model, respectively. And the output of the feature extractor is taken out, which is reduced to 11 dimensions by multi-layer neural networks. Then the TSNE algorithm, which is based on unsupervised learning, is used to reduce the dimension of the output to obtain two-dimensional data and realize data visualization. The current data in different periods correspond to different motor rotor phases. After the samples are transformed by the multi-layer neural networks and the TNSE, the samples will be mapped to certain positions in the two-dimensional space. The distributions of the two datasets are shown in Figs. 8 and 9, where the data of different motor fault categories are represented by colors and marker symbols.

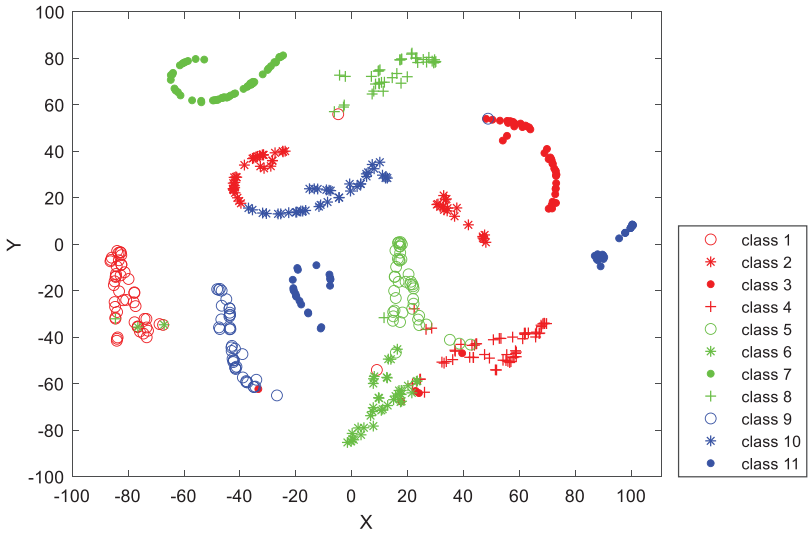

Figure 8: Distribution of the real dataset

As shown in Fig. 8, some points' classification is wrong. For example, the data of fault category 6 is classified as category 4, category 5 and category 1. In addition, it can be seen from the figure that some data of fault category 2 and fault category 10 are in similar spatial positions, but the data points do not overlap. This shows that the two kinds of data have similar characteristics.

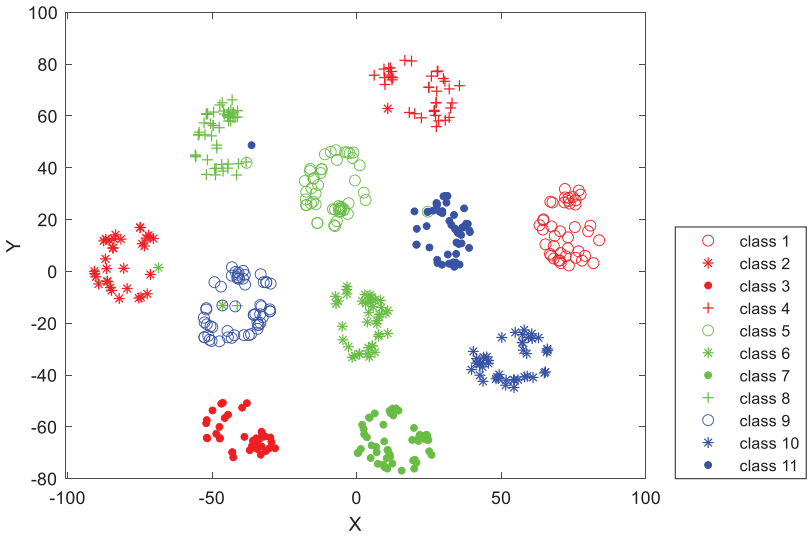

Figure 9: Distribution of the fake dataset

4.3.2 Discussion

As shown in Fig. 8, the data points of different fault categories show curve distribution. The points are highly aggregated and linearly distributed, which indicates that the characteristics of the data lack diversity. As shown in Fig. 9, the fake data points are clustered into several point clouds. It can be easily inferred that each cloud corresponds to one type of motor fault. This shows that the deep feature vector extracted in this paper can be well separable. The data points are highly dispersed and have a large distribution range, which indicates that the data cover a larger characteristic range. In addition, the TSNE models of the real and fake datasets are trained separately. Based on different TSNE models, the visual characteristics of the dataset may be different. In Figs. 8 and 9, there is no clear corresponding relationship between the positions of specific fault types.

As can be seen from Fig. 7, the classification model trained with a real dataset has a fast convergence rate, which indicates that the feature diversity of the dataset is poor, and the model can quickly fit the data features. The model trained with the fake dataset or combined dataset has a slow convergence rate, which indicates that the characteristics of datasets are more complex and the training is more difficult.

From the prediction accuracy of the models on the test dataset, it can see that the model trained by the real dataset has a very low accuracy on unknown working condition data. The model trained by the fake dataset and combined dataset achieves higher accuracy, which shows that the fake data generated by the multi-layer perceptron Wasserstein generative adversarial model has a larger distribution range than the real data. In this research, the combined dataset consists of real data and fake data. The real data has outstanding fault features. The fake dataset contains more diversity than the real dataset. But the data augmentation will change data characteristics, which weakens fault feature expression. Therefore, it can be inferred that the combined dataset has diversity and outstanding fault features at the same time, which help the classification model achieve better performance.

The data augmentation model and fault classification model proposed in this paper belong to small neural networks. Both models can be run directly on personal computers. This makes the method easier to implement in the building engineering environment.

5 Conclusion and Future Works

In order to solve the problem that the current signal of the motor lack diversity, a current data augmentation model and a motor fault classification model are proposed in this paper. The data augmentation model uses the wavelet coefficients of the current signal under discrete working conditions to generate the fake dataset. The fault classification model trained on the fake dataset and combined dataset achieves 87.2% and 90.1% prediction accuracy on unseen real data, respectively. The data augmentation model can effectively improve the diversity of training samples. The motor fault classification model trained on the combined dataset has better generalization performance.

The effectiveness of regarding generated data as a training dataset has been verified in the proposed diagnosis framework under a small sample diagnosis scenario of motor fault diagnosis. In the future, the research works on fake data based diagnosis could be extended to: (1) Based on generative adversarial learning, establishing effective diagnosis methods under other small sample scenarios, such as only part of the time or location node data can be obtained, or only part of the attribute data can be observed. (2) Exploring the method on other equipment, such as gearboxes and bearings. (3) Exploring the possibility of combining other data generation methods with generative adversarial methods to propose a new data enhancement method, which has wider application scope and can further improve the quality of generated data.

Funding Statement: This work is supported by the National Key Research and Development Program of China (No. 2020YFB1713503); the Fundamental Research Funds for the Central Universities (No. 20720190009) and 2019 Industry-University-Research Cooperation Project of Aero Engine Corporation of China (No. HFZL2019CXY02).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Hernandez-Solis, A., Carlsson, F. (2010). Diagnosis of submersible centrifugal pumps: A motor current and power signature approaches. *EPE Journal, 20(1),* 58–64. DOI 10.1080/09398368.2010.11463749.
2. Park, B. G., Lee, K. J., Kim, R. Y., Kim, T. S., Ryu, J. S. et al. (2011). Simple fault diagnosis based on operating characteristic of brushless direct-current motor drives. *IEEE Transactions on Industrial Electronics, 58(5),* 1586–1593. DOI 10.1109/TIE.2010.2072895.
3. Choi, S., Haque, M. S., Bin Tarek, M. T., Mulpuri, V., Duan, Y. et al. (2018). Fault diagnosis techniques for permanent magnet AC machine and drives–A review of current state of the art. *IEEE Transactions on Transportation Electrification, 4(2),* 444–463. DOI 10.1109/tte.2018.2819627.
4. Shao, H. D., Lin, J., Zhang, L. W., Galar, D., Kumar, U. (2021). A novel approach of multi-sensory fusion to collaborative fault diagnosis in maintenance. *Information Fusion, 74,* 65–76. DOI 10.1016/j.inffus.2021.03.008.
5. Niu, G., Widodo, A., Son, J. D., Yang, B. S., Hwang, D. H. et al. (2008). Decision-level fusion based on wavelet decomposition for induction motor fault diagnosis using transient current signal. *Expert Systems with Applications, 35(3),* 918–928. DOI 10.1016/j.eswa.2007.08.024.
6. Gonzalez-Jimenez, D., Del-Olmo, J., Poza, J., Garramiola, F., Madina, P. (2021). Data-driven fault diagnosis for electric drives: A review. *Sensors, 21(12)*. DOI 10.3390/s21124024.
7. Alauddin, M., Khan, F., Imtiaz, S., Ahmed, S. (2018). A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems. *Industrial & Engineering Chemistry Research, 57(32),* 10719–10735. DOI 10.1021/acs.iecr.8b00936.

8. Shao, S. Y., Wang, P., Yan, R. Q. (2019). Generative adversarial networks for data augmentation in machine fault diagnosis. *Computers in Industry, 106,* 85–93. DOI 10.1016/j.compind.2019.01.001.
9. Peng, T. Y., Shen, C. Q., Sun, S. L., Wang, D. (2022). Fault feature extractor based on bootstrap your own latent and data augmentation algorithm for unlabeled vibration signals. *IEEE Transactions on Industrial Electronics, 69(9),* 9547–9555. DOI 10.1109/TIE.2021.3111567.
10. Meng, Z., Guo, X. L., Pan, Z. Z., Sun, D. Y., Liu, S. (2019). Data segmentation and augmentation methods based on raw data using deep neural networks approach for rotating machinery fault diagnosis. *IEEE Access, 7,* 79510–79522. DOI 10.1109/ACCESS.2019.2923417.
11. Wang, X. Y., Chu, Z. Y., Han, B. K., Wang, J. R., Zhang, G. W. et al. (2020). A novel data augmentation method for intelligent fault diagnosis under speed fluctuation condition. *IEEE Access, 8,* 143383–143396. DOI 10.1109/ACCESS.2020.3014340.
12. Shorten, C., Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6,* 60. DOI 10.1186/s40537-019-0197-0.
13. Zhang, W., Peng, G. L., Li, C. H., Chen, Y. H., Zhang, Z. J. (2017). A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors, 17(2)*. DOI 10.3390/s17020425.
14. Ullah, I., Hussain, M., Qazi, E. U., Aboalsamh, H. (2018). An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert Systems with Applications, 107,* 61–71. DOI 10.1016/j.eswa.2018.04.021.
15. Salamon, J., Bello, J. P. (2017). Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Processing Letters, 24(3),* 279–283. DOI 10.1109/LSP.2017.2657381.
16. Li, X., Zhang, W., Ding, Q., Sun, J. Q. (2020). Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. *Journal of Intelligent Manufacturing, 31(2),* 433–452. DOI 10.1007/s10845-018-1456-1.
17. Yu, K., Lin, T. R., Ma, H., Li, X., Li, X. (2021). A multi-stage semi-supervised learning approach for intelligent fault diagnosis of rolling bearing using data augmentation and metric learning. *Mechanical Systems and Signal Processing, 146*. DOI 10.1016/j.ymssp.2020.107043.
18. Zhang, W., Li, X., Jia, X. D., Ma, H., Luo, Z. et al. (2020). Machinery fault diagnosis with imbalanced data using deep generative adversarial networks. *Measurement, 152*, 107377. DOI 10.1016/j.measurement.2019.107377.
19. Cabrera, D., Sancho, F., Long, J. Y., Sanchez, R. V., Zhang, S. H. et al. (2019). Generative adversarial networks selection approach for extremely imbalanced fault diagnosis of reciprocating machinery. *IEEE Access, 7,* 70643–70653. DOI 10.1109/ACCESS.2019.2917604.
20. Yang, H. W., Zhou, Y., Zhao, J. (2020). Current covariance analysis-based open-circuit fault diagnosis for voltage-source-inverter-fed vector-controlled induction motor drives. *Journal of Power Electronics, 20(2),* 492–500. DOI 10.1007/s43236-020-00043-5.
21. Naseri, F., Schaltz, E., Lu, K. Y., Farjah, E. (2020). Real-time open-switch fault diagnosis in automotive permanent magnet synchronous motor drives based on kalman filter. *IET Power Electronics, 13(12),* 2450–2460. DOI 10.1049/iet-pel.2019.1498.
22. Kou, L., Liu, C., Cai, G. W., Zhou, J. N., Yuan, Q. D. et al. (2020). Fault diagnosis for open-circuit faults in npc inverter based on knowledge-driven and data-driven approaches. *IET Power Electronics, 13(6),* 1236–1245. DOI 10.1049/iet-pel.2019.0835.
23. Eldeeb, H. H., Berzoy, A., Saad, A. A., Zhao, H. S., Mohammed, O. A. (2020). Differential mathematical morphological-based online diagnosis of stator interturn failures in direct torque control drive systems. *IEEE Transactions on Industry Applications, 56(6),* 6272–6285. DOI 10.1109/TIA.2020.3019779.
24. Salehifar, M., Salehi Arashloo, R., Moreno-Eguilaz, M., Sala, V., Romeral, L. (2015). Observer-based open transistor fault diagnosis and fault-tolerant control of five-phase permanent magnet motor drive for application in electric vehicles. *IET Power Electronics, 8(1),* 76–87. DOI 10.1049/iet-pel.2013.0949.

25. Yan, H., Xu, Y., Cai, F., Zhang, H., Zhao, W. et al. (2019). PWM-VSI fault diagnosis for a pmsm drive based on the fuzzy logic approach. *IEEE Transactions on Power Electronics, 34(1),* 759–768. DOI 10.1109/tpel.2018.2814615.
26. Potamianos, P. G., Mitronikas, E. D., Safacas, A. N. (2014). Open-circuit fault diagnosis for matrix converter drives and remedial operation using carrier-based modulation methods. *IEEE Transactions on Industrial Electronics, 61(1),* 531–545. DOI 10.1109/tie.2013.2240639.
27. Zaman, S. M. K., Liang, X., Li, W. (2021). Fault diagnosis for variable frequency drive-fed induction motors using wavelet packet decomposition and greedy-gradient max-cut learning. *IEEE Access, 9,* 65490–65502. DOI 10.1109/access.2021.3076149.
28. Shabbir, M. N. S. K., Liang, X., Chakrabarti, S. (2021). An anova-based fault diagnosis approach for variable frequency drive-fed induction motors. *IEEE Transactions on Energy Conversion, 36(1),* 500–512. DOI 10.1109/tec.2020.3003838.
29. Malik, H., Almutairi, A. (2021). Modified fuzzy-q-learning (MFQL)-based mechanical fault diagnosis for direct-drive wind turbines using electrical signals. *IEEE Access, 9,* 52569–52579. DOI 10.1109/access.2021.3070483.
30. Rabbi, S. F., Little, M. L., Saleh, S. A., Rahman, M. A. (2017). A novel technique using multiresolution wavelet packet decomposition for real time diagnosis of hunting in line start IPM motor drives. *IEEE Transactions on Industry Applications, 53(3),* 3005–3019. DOI 10.1109/tia.2016.2633541.
31. Long, Z., Zhang, X., Zhang, L., Qin, G., Huang, S. et al.(2021). Motor fault diagnosis using attention mechanism and improved adaboost driven by multi-sensor information. *Measurement, 170,* 108718. DOI 10.1016/j.measurement.2020.108718.
32. Skowron, M., Orlowska-Kowalska, T., Wolkiewicz, M., Kowalski, C. T. (2020). Convolutional neural network-based stator current data-driven incipient stator fault diagnosis of inverter-fed induction motor. *Energies, 13(6),* 1475. DOI 10.3390/en13061475.
33. Munikoti, S., Das, L., Natarajan, B., Srinivasan, B. (2019). Data-driven approaches for diagnosis of incipient faults in DC motors. *IEEE Transactions on Industrial Informatics, 15(9),* 5299–5308. DOI 10.1109/tii.2019.2895132.
34. Masrur, M. A., Chen, Z., Murphey, Y. (2010). Intelligent diagnosis of open and short circuit faults in electric drive inverters for real-time applications. *IET Power Electronics, 3(2),* 279–291. DOI 10.1049/iet-pel.2008.0362.
35. Cao, H. R., Shao, H. D., Zhong, X., Deng, Q. W., Yang, X. K. et al. (2022). Unsupervised domain-share cnn for machine fault transfer diagnosis from steady speeds to time-varying speeds. *Journal of Manufacturing Systems, 62,* 186–198. DOI 10.1016/j.jmsy.2021.11.016.
36. Hussain, M., Memon, T. D., Hussain, I., Memon, Z. A., Kumar, D. (2022). Fault detection and identification using deep learning algorithms in induction motors. *Computer Modeling in Engineering & Sciences, 133(2),* 435–470. DOI 10.32604/cmes.2022.020583.
37. Mushtaq, S., Islam, M. M. M., Sohaib, M. (2021). Deep learning aided data-driven fault diagnosis of rotatory machine: A comprehensive review. *Energies, 14(16)*, 5150. DOI 10.3390/en14165150.
38. He, Z. Y., Shao, H. D., Jing, L., Cheng, J. S., Yang, Y. (2020). Transfer fault diagnosis of bearing installed in different machines using enhanced deep auto-encoder. *Measurement, 152*, 107393. DOI 10.1016/j.measurement.2019.107393.
39. Campos-Delgado, D. U., Espinoza-Trejo, D. R. (2011). An observer-based diagnosis scheme for single and simultaneous open-switch faults in induction motor drives. *IEEE Transactions on Industrial Electronics, 58(2),* 671–679. DOI 10.1109/tie.2010.2047829.
40. Zhang, Z., Wu, Y., Qi, S. (2020). Diagnosis method for open-circuit faults of six-phase permanent magnet synchronous motor drive system. *IET Power Electronics, 13(15),* 3305–3313. DOI 10.1049/iet-pel.2019.1594.

41. Aktas, M. (2012). A novel method for inverter faults detection and diagnosis in pmsm drives of hevs based on discrete wavelet transform. *Advances in Electrical and Computer Engineering, 12(4),* 33–38. DOI 10.4316/aece.2012.04005.
42. Guo, M. F., Yang, N. C., You, L. X. (2018). Wavelet-transform based early detection method for short-circuit faults in power distribution networks. *International Journal of Electrical Power & Energy Systems, 99,* 706–721. DOI 10.1016/j.ijepes.2018.01.013.
43. Wang, C. Y., Xu, C., Yao, X., Tao, D. C. (2019). Evolutionary generative adversarial networks. *IEEE Transactions on Evolutionary Computation, 23(6),* 921–934. DOI 10.1109/TEVC.2019.2895748.
44. Lu, Y., Wang, S. G., Zhao, W. T., Zhao, Y. (2019). WGAN-based robust occluded facial expression recognition. *IEEE Access, 7,* 93594–93610. DOI 10.1109/ACCESS.2019.2928125.
45. Panaretos, V. M., Zemel, Y. (2019). Statistical aspects of wasserstein distances. *Annual Review of Statistics and Its Application, 6,* 405–431. DOI 10.1146/annurev-statistics-030718-104938.
46. Yinka-Banjo, C., Ugot, O. A. (2020). A review of generative adversarial networks and its application in cybersecurity. *Artificial Intelligence Review, 53(3),* 1721–1736. DOI 10.1007/s10462-019-09717-4.
47. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A. (2017). Improved training of wasserstein gans. *31st Annual Conference on Neural Information Processing Systems (NIPS)*, Long Beach, California, USA.
48. Zhou, C. S., Zhang, J. S., Liu, J. M. (2018). Lp-WGAN: Using Lp-norm normalization to stabilize wasserstein generative adversarial networks. *Knowledge-Based Systems, 161,* 415–424. DOI 10.1016/j.knosys.2018.08.004.
49. Bator, M., Dicks, A., Mnks, U., Lohweg, V. (2012). Feature extraction and reduction applied to sensorless drive diagnosis. *22nd Workshop Computational Intelligence*, Dortmund, North Rhine-Westphalia, Germany. DOI 10.13140/2.1.2421.5689.
50. Freire, N. M. A., Estima, J. O., Marques Cardoso, A. J. (2014). A new approach for current sensor fault diagnosis in pmsg drives for wind energy conversion systems. *IEEE Transactions on Industry Applications, 50(2),* 1206–1214. DOI 10.1109/tia.2013.2271992.
51. Espinoza-Trejo, D. R., Campos-Delgado, D. U., Bossio, G., Barcenas, E., Hernandez-Diez, J. E. et al. (2013). Fault diagnosis scheme for open-circuit faults in field-oriented control induction motor drives. *IET Power Electronics, 6(5),* 869–877. DOI 10.1049/iet-pel.2012.0256.
52. Gou, B., Xu, Y., Xia, Y., Deng, Q. L., Ge, X. L. (2020). An online data-driven method for simultaneous diagnosis of igbt and current sensor fault of three-phase pwm inverter in induction motor drives. *IEEE Transactions on Power Electronics, 35(12),* 13281–13294. DOI 10.1109/TPEL.2020.2994351.
53. Zhen, D., Wang, T., Gu, F., Ball, A. D. (2013). Fault diagnosis of motor drives using stator current signal analysis based on dynamic time warping. *Mechanical Systems and Signal Processing, 34(1–2),* 191–202. DOI 10.1016/j.ymssp.2012.07.018.
54. Widodo, A., Yang, B. S., Gu, D. S., Choi, B. K. (2009). Intelligent fault diagnosis system of induction motor based on transient current signal. *Mechatronics, 19(5),* 680–689. DOI 10.1016/j.mechatronics.2009.02.002.