Data augmentation and intelligent fault diagnosis of planetary gearbox using ILoFGAN under extremely limited samples

Mingzhi Chen, Haidong Shao, Haoxuan Dou, Wei Li, and Bin Liu, *Member*, *IEEE*

***Abstract*—Though the existing generative adversarial networks (GAN) have the potential for data augmentation and intelligent fault diagnosis of planetary gearbox, it remains difficult to deal with extremely limited training samples and effectively fuse the representative and diverse information. To tackle the above challenges, an improved local fusion generative adversarial network (ILoFGAN) is proposed. Time-domain waveforms are firstly transformed into the time-frequency diagrams to highlight the fault characteristics. Subsequently, a local fusion module is used to fully utilize extremely limited samples and fuse the local features. Finally, a new generator embedded with multi-head attention modules is constructed to effectively improve the accuracy and flexibility of the feature fusion process. The proposed method is applied to the analysis of planetary gearbox vibration signals. The results show that the proposed method can generate a large number of samples with higher similarity and better diversity compared with the existing mainstream GANs using 6 training samples in each type. The generated samples are used to augment the limited dataset, prominently improving the accuracy of the fault diagnosis task.**

***Index Terms*—intelligent fault diagnosis, extremely few samples, multi-head attention mechanism, ILoFGAN, planetary gearbox.**

I. Introduction

Planetary gearbox has been widely used in various machinery, including automobile, shield machine, and aircraft carrier. However, its performance will degrade during the long-term operation process. Distinguishing different fault modes is of great significance for failure prediction and maintenance decision-making [1-4]. Thanks to the rapid development of data-driven technology, many researchers have investigated intelligent fault diagnosis of planetary gearbox by mining the rich information provided by the collected vibration signals. Particularly, fault diagnosis based on deep learning has attracted more attention than traditional data-driven technology, due to its end-to-end feature learning ability [5-7]. However, the success of deep learning largely depends on the sufficiency of training samples. Unfortunately, with the increasing improvement of reliability and quality, the machinery works under healthy states most of time in the real industrial scenarios, which means that acquiring sufficient fault samples is difficult [8-10].

To address this challenge, researchers have successively developed data augmentation methods such as data sampling and data generation [11]. Data sampling techniques mainly include random over-sampling (RO-Sampling) [12], synthetic minority oversampling technique (SMOTE) [13], and adaptive synthetic (ADASYN) [14]. Although these methods can synthesize new fault samples, they focus on replication or interpolation techniques while failing to consider the effect of data distribution, which will lead to the lack of diversity.

Data generation technology aims to capture features from original data and generate new samples with a similar distribution. Typical representative algorithms include GAN [15], Wasserstein GAN (WGAN) [16], Wasserstein GAN with gradient penalty (WGAN-GP) [17], deep convolutional GAN (DCGAN) [18], auxiliary classier GAN (ACGAN) [19] and variational autoencoding GAN (VAEGAN) [20], which are gradually used in the field of machinery fault diagnosis [21-23]. Wang *et al.* [24] used the analog signals generated by WGAN to expand the unbalanced dataset and train stacked autoencoders to detect the health status of mechanical equipment. Wang *et al.* [25] used enhanced deep convolutional GAN to generate more fault samples and figure out the problem of imbalanced data, thus improving fault classification accuracy. Peng *et al.* [26] proposed the reinforcement auxiliary classification GAN, which stabilized the training process by using boundary-seeking loss and introduced cost-sensitive learning to alleviate data imbalance in the Tennessee Eastman dataset. Li *et al.* [27] proposed an auxiliary classier Wasserstein GAN with a gradient penalty to generate high-quality frequency samples. Dixit *et al.* [28] proposed a novel conditional auxiliary classifier ACGAN framework combined with model agnostic meta-learning, which is validated by bearing and air compressor datasets. Miao *et al.* [29] proposed an improved VAEGAN to complete the data augmentation task under the background of a few samples.

It can be witnessed from the literature that various GANs mentioned above have been studied successively in mechanical fault diagnosis. However, in the presence of extremely limited fault samples, the following problems still need to be addressed to further improve the diagnosis performance. (1) The inputs of generators applied in the GANs, as mentioned above, are mostly random noise or

This research is supported by the National Natural Science Foundation of China (No. 51905160), and the Natural Science Fund for Excellent Young Scholars of Hunan Province (No. 2021JJ20017). *(Corresponding author: Haidong Shao.)*

Mingzhi Chen, Haidong Shao, Haoxuan Dou, and Wei Li are with College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China (e-mail: 1297008453@hnu.edu.cn; hdshao@hnu.edu.cn; hxdou@hnu.edu.cn; liwei2020@hnu.edu.cn).

Bin Liu is with Department of Management Science, University of Strathclyde, Glasgow G1 1XQ, UK (email: b.liu@strath.ac.uk).

This is a peer reviewed, accepted author manuscript of the following research article: Chen, M., Shao, H., Dou, H., Li, W., & Liu, B. (Accepted/In press - October 2022). Data augmentation and intelligent fault diagnosis of planetary gearbox using ILoFGAN under extremely limited samples. IEEE Transactions on Reliability.

random noise with some labels, which fails to fully capture each training sample's information and make it hard to extract diverse and representative feature representation [24, 27, 29]. (2) The above GANs mostly generate the original time-domain vibration signals or the spectral signals while failing to involve extracting the correlation features of time-frequency spectrum of vibration signals. (3) In the above researches, the number of existing fault samples used to train GAN is generally more than 20 per category. Due to the increasingly expensive cost of labeled samples and the long-term accumulation of massive monitoring data, availability of fault samples will be limited in the practical engineering field. Therefore, we focus on the intelligent fault diagnosis method for planetary gearbox under extremely few samples.

To settle these problems, an improved local fusion GAN (ILoFGAN) is proposed for data augmentation of planetary gearbox under extremely limited samples. The comparison results from the experiments demonstrate that ILoFGAN can produce high-quality fault samples of the planetary gearbox by utilizing only a few time-frequency diagrams and remarkably improve diagnosis accuracy. Two quality evaluation indexes confirm the efficiency and superiority of the proposed method compared with mainstream GANs. The main contributions of this study are as follows:

1) To generate sufficient time-frequency diagrams of the planetary gearbox with high quality and diversity, ILoFGAN is used to fully extract value from the limited information and fuse the local feature of extremely few samples.
2) To assist the local fusion module to improve the precision and flexibility of local feature match, a novel generator embedded with multi-head attention modules is constructed for mining various key local features in the time-frequency diagrams.

The remainder of this article is organized as follows. The theoretical background of LoFGAN is provided in Section II. Section III demonstrates the details of the novel generator and the overall framework. Experimental study and evaluation of the proposed method are presented in Section IV. Conclusions are summarized in Section V.

II. Brief Introduction to LoFGAN

LoFGAN [30] was proposed in 2021, which added local fusion techniques into GAN to extract features. The basic structure is shown in **Fig. 1**. Dissimilar to traditional GANs, the generator ***G*** is composed of encoder ***E***, decoder ***H***, a local fusion module (***LFM***). Its input contains k images $X=\{x_1,\ldots,x_k\}$ in each batch. After images X through encoder ***E***, k feature vectors $\mathcal{F}=\boldsymbol{E}(X)$ are obtained. Then, ***LFM*** generates fusion feature vector $\mathcal{F}=\boldsymbol{LFM}(\mathcal{F},\alpha)$ by utilize $\mathcal{F}$ and random coefficient α, and $\mathcal{F}$ is fed into the decoder ***H*** to obtain the generated image $z=\boldsymbol{H}(\hat{F})$. Finally, the input images X and the generated image z are channeled into the discriminator ***D***.

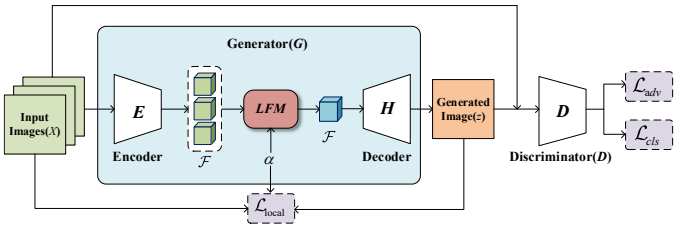


Fig. 1. Basic structure of LoFGAN.

Local fusion module (*LFM*): One of the feature vectors $\mathcal{F}=\boldsymbol{E}(X)$ will be randomly selected as the base feature $f_{base}\in\mathbb{R}^{w\times h\times c}$, where w, h and c represent the dimension of features in width, height and number of channels, respectively. The remaining (k-1) feature vectors will be used as the reference features $\mathbb{F}_{ref}$. Their purpose is to provide more relevant features for the base feature f_{base} to fusion features. Then $m=r\times w\times h$ local representations are randomly selected in f_{base} to obtain a local representation ϕ_{base}, where $r\in(0,1]$ is a selection ratio to decide the number of fused local representations. After that, the similarity map M of f_{ref} in $\mathbb{F}_{ref}$ and ϕ_{base} is established as follows:

$$M^{(i,j)}=s(\phi_{base}^{(i)},f_{ref}^{(j)}) \tag{1}$$

where $s(\cdot)$ is the cosine similarity function, $i\in\{1,\ldots,m\}$, $j\in\{1,\ldots,w\times h\}$. Using the similarity map M, the corresponding local feature representation ϕ_{ref} in f_{ref} that is most similar to ϕ_{base} can be found for each position. Then, a random coefficient vector $\alpha=[\alpha_1,\ldots,\alpha_k]$ is used to fuse the selected local feature representation ϕ_{ref} with the local representation ϕ_{base}. The local fusion representation ϕ_{fuse} is obtained as:

$$\phi_{fuse}^{(n)}=\alpha_{base}\cdot\phi_{base}^{(n)}+\sum_{j=1,\ldots,k,j\neq base}\alpha_j\cdot\phi_{ref}^{(n)}(j) \tag{2}$$

where $\sum_{j=1}^{k}\alpha_j=1$, $\alpha_j\geq 0$ and $n=1,\ldots,m$. Finally, the local fusion representation ϕ_{fuse} is replaced at the corresponding positions of f_{base}, so as to obtain a fusion feature vector $\mathcal{F}$, which is the outcome of ***LFM***. The algorithm process is sketched in **Fig. 2**.

Local reconstruction loss: Its concept is to duplicate the feature replacement process at the image level, which aim to constrain the generated image. Specifically, the positions where f_{base} and f_{ref} select ϕ_{base} and ϕ_{ref} respectively is adopted to map each position of the selected features back

to the original image size, and receive a roughly fused image $\boldsymbol{LFM}(X,\alpha)$. Then the generated image z is optimized with:

$$\mathcal{L}_{local} = \| z - \boldsymbol{LFM}(X,\alpha) \|_1 \tag{3}$$

where $\|\cdot\|_1$ is the first norm. By replacing the selected patches of the original image with the most similar patches in the reference images, the fusion image can be obtained to guide the network to generate more vivid images.

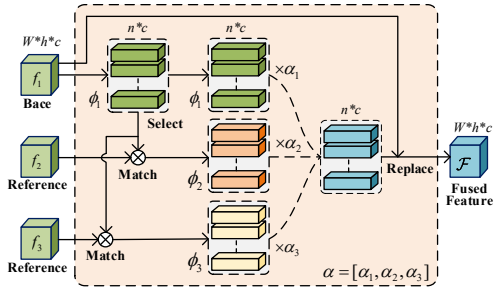


Fig. 2. Main feature fusion process of ***LFM***.

III. THE PROPOSED METHOD

A. Construction of novel generator imbedded with multi-head attention module

Compared with the shortcoming of the local receptive domain in convolution operation, the attention mechanism [31] can globally extract mutual information between any two locations in the input features. Generally, the query-key-value (QKV) mode is adopted to obtain the attention weight of the feature to effectively capture the global information, expressed in the following function:

$$\text{Attention}(Q,K,V) = \text{SoftMax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{5}$$

Instead of using d dimension queries, keys and values in single attention function, the multi-head attention (MHA) mechanism carries out linear learnable mapping of queries, keys and values for e times respectively to get the vectors of the corresponding d_k, d_k and d_v dimensions. Then, the attention function of each linear mapping result is calculated in parallel to obtain the output of d_k dimension. Finally, these values are concatenated together and linear mapping once again to obtain the final result:

$$\text{MultiHead}(Q,K,V) = \text{Concat}\left(\text{head}_1,\ldots,\text{head}_e\right)W^O \tag{6}$$

with

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \tag{7}$$

where $W_i^Q \in \mathbb{R}^{d_{\text{model}} \times d_k}$, $W_i^K \in \mathbb{R}^{d_{\text{model}} \times d_k}$, $W_i^V \in \mathbb{R}^{d_{\text{model}} \times d_v}$ and $W^O \in \mathbb{R}^{hd_v \times d_{\text{model}}}$ represent the parameter matrix of linear projection of queries, keys, values and global vector, respectively. Thus, the generator imbedded MHA module can focus on different subspace information from different locations simultaneously. **Fig. 3** display the structure of the MHA module.

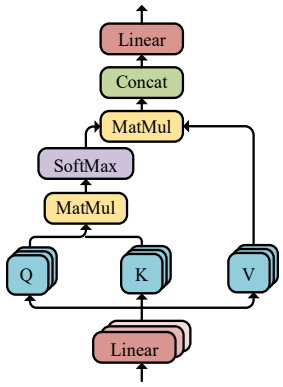


Fig. 3. Structure of MHA module

The generator of the proposed model mainly includes local fusion module, encoder and decoder. The encoder is composed of an MHA module and six convolution blocks. Each of these convolution blocks contains a convolution layer, leaky rectified linear unit (leaky-ReLU) activation function and batch normalization layer (BN). The encoder can quickly obtain the global information of the fault sample and pay more attention to the key local feature of the energy distribution in the time-frequency diagram, which facilitates ***LFM*** to find the local feature corresponding position between the basic image and the reference images more accurately. In addition, the multi-head mode assists the generator in mining various local features so that the feature vectors $\mathcal{F}$ contain more mutual information, which facilitates ***LFM*** to fuse local features in multiple ways and improve the diversity of the generated samples. The structure of the decoder is symmetric with the encoder, which includes an MHA module, two convolution blocks and four up sample convolution blocks. Therefore, the MHA module enables the decoder to focus more on the critical details of generated images to improve the quality of similarity.

B. Training process of ILoFGAN

The structure of the constructed ILoFGAN is displayed in **Fig. 4**, mainly composed of a generator and a discriminator. The discriminator uses four residual blocks as feature extractors, which contains two convolution layers, one average pooling layer, and a residual link. Finally, two full connection layers are used to evaluate the authenticity and classification of images respectively.

The following objective function along with the local reconstruction loss are utilized to guide the training process of generator ***G*** and discriminator ***D***:

$$\mathcal{L}_G = \mathcal{L}_{adv}^G + \lambda_{cls}^G \mathcal{L}_{cls}^G + \lambda_{local}\, \mathcal{L}_{local}^G \tag{7}$$

$$\mathcal{L}_D = \mathcal{L}_{adv}^D + \lambda_{cls}^D \mathcal{L}_{cls}^D \tag{8}$$

$$\mathcal{L}_{adv}^D = \max(0,1-\boldsymbol{D}(X)) + \max(0,1+\boldsymbol{D}(z)) \tag{9}$$

$$\mathcal{L}_{adv}^G = -\boldsymbol{D}(z) \tag{10}$$

$$\mathcal{L}_{cls}^D = -\log P(c(X)\,|\,X) \tag{11}$$

$$\mathcal{L}_{cls}^G = -\log P(c(X)\,|\,z) \tag{12}$$

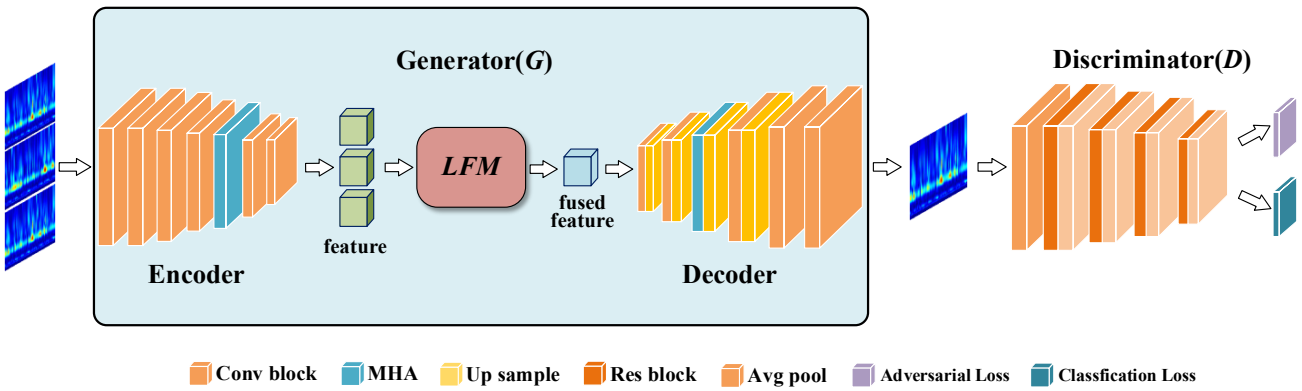


Fig. 4. Structure of ILoFGAN

where X represents the input images, $c(X)$ represents the category of the image, $z = \boldsymbol{G}(X,\alpha)$ represents the generated image, $\mathcal{L}_{adv}^{D}$ represents adversarial loss of discriminator, $\mathcal{L}_{adv}^{G}$ represents adversarial loss of generator, $\mathcal{L}_{cls}^{D}$ represents classification loss of discriminator, $\mathcal{L}_{cls}^{G}$ represents classification loss of generator, λ_{cls}^{D}, λ_{cls}^{G}, λ_{local} are the regularization parameters of the corresponding discriminator's classification loss, generator's classification loss and generator's local reconstruction loss, respectively. **Algorithm 1** list the main training process of the ILoFGAN.

C. Framework of the proposed fault diagnosis method

The framework of the proposed method is shown in **Fig. 5**, which mainly consists of the following steps:

- **Step 1:** Obtain the vibration signals of the planetary gearbox under various working conditions, and vibration signals of fault types are extremely few. Then transform them into corresponding time-frequency diagrams samples by continuous wavelet transform.
- **Step 2:** Establish ILoFGAN model with the MHA module embedded into the generator to improve the generation quality. Then train ILoFGAN with extremely few time-frequency diagrams.
- **Step 3:** Use the trained generator of ILoFGAN to produce a large number of generated samples for each fault type. Two evaluation indexes are adopted to assess the similarity and diversity of the generated samples.
- **Step 4:** The generated samples and the original samples are mixed and input into the convolutional neural network (CNN) for feature extraction and fault diagnosis.

IV. Experimental Verification

A. Fault data description of planetary gearbox

Two cases of planetary gearboxes are employed to verify the effectiveness of the proposed method.

Case 1: The experiment adopts vibration signals data from gearbox measurements from the University of Connecticut [32]. As shown in **Fig. 6**, the experimental equipment mainly includes the motor, motor controller, two-stage planetary gearbox and brake. The speed of the input shaft is measured by a tachometer, and the vibration signals of the gear are measured by a dSPACE system, whose sampling frequency is set to 20 kHz. The vibration signals dataset of various faults is collected by introducing 9 different working conditions on the input shaft, including healthy, missing tooth, root crack, spalling, and chipping tip with five different levels of severity. Each type has 104 samples, and each sample is a time-domain vibration signals containing 3600 sampling points.

Algorithm 1 Training process of the ILoFGAN

Inputs: the iteration number K, the steps to train the discriminator is T, the number of images in single set is k, the batch size is b, a random coefficient vector $\alpha = [\alpha_1,...,\alpha_k]$, hyperparameters λ_{cls}^{D}, λ_{cls}^{G}, λ_{local}, and the optimizer *Adam*.

Initialization: Initial discriminator parameter θ_D, generator parameter θ_G.

For K iterations **do:**

1. Sample a batch of sets $\{X^{(1)},...,X^{(b)}\}$ from real data distribution $p_r(X)$

   For T steps **do:**

   - Sample a batch of examples $\{z^{(1)},...,z^{(b)}\}$ from generated data distribution $p_g(z)$
   - Freeze generator, only update discriminator by ascending its stochastic gradient:

$$\mathcal{L}_{\text{adv}}^{D} = \max(0,1-\boldsymbol{D}(X)) + \max(0,1+\boldsymbol{D}(z))$$

$$\mathcal{L}_{\text{cls}}^{D} = -\log P(c(X)\,|\,X)$$

$$\theta_D \leftarrow Adam(\nabla_{\theta_D}(\mathcal{L}_{adv}^{D} + \lambda_{cls}^{D}\mathcal{L}_{cls}^{D}))$$

   End for

2. Freeze discriminator, only update generator by ascending its stochastic gradient:

$$\mathcal{L}_{adv}^{G} = -\boldsymbol{D}(z)$$

$$\mathcal{L}_{cls}^{G} = -\log P(c(X)\,|\,z)$$

$$\mathcal{L}_{local}^{G} = \| z - \boldsymbol{LFM}(X,\alpha) \|_1$$

$$\theta_G \leftarrow Adam(\nabla_{\theta_G}(\mathcal{L}_{adv}^{G} + \lambda_{cls}^{G}\mathcal{L}_{cls}^{G} + \lambda_{local}\mathcal{L}_{local}^{G}))$$

End for

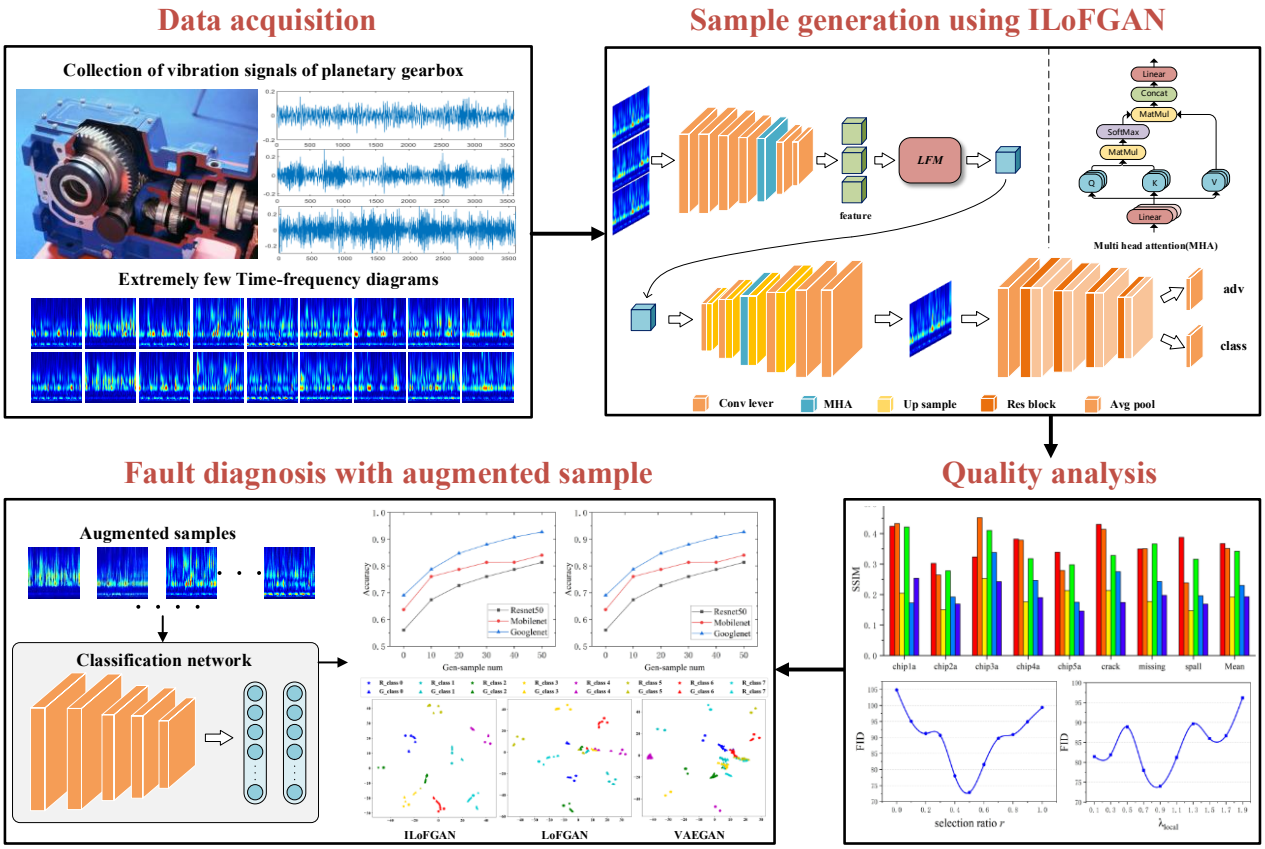


Fig. 5. Overall framework of the proposed fault diagnosis method

Case 2: The planetary gearbox dataset of this case was collected from the drivetrain dynamic simulator from Southeast University, China [33], as shown in **Fig. 7**. Two different working conditions are investigated with the rotating speed system load set to be either 20 HZ-0V or 30 HZ-2V. The sampling frequency is set to 2 kHz. In each condition, there are 5 different fault types of the gearbox, including chipping tip, missing tooth, root crack, surface wear and healthy. The experiment adopts the vibration signals of the motor vibration in the condition of 30 HZ-2V. There are 106 samples in each fault type, and each sample is a time-domain vibration signals containing 2000 sampling points.

Because the time-frequency diagrams can clearly describe the fluctuation of signal frequency with time and can also express the energy distribution, the continuous wavelet transform is adopted to transform each time-domain waveform into the corresponding time-frequency spectrum. Then, the time-frequency spectrums are visualized into an RGB three-channel time-frequency diagram with pixels of 64*64 in the form of a thermal diagram.

Finally, the original dataset of case 1 is transformed into an image dataset of 9 categories with 104 images in each category. The original dataset of case 2 is transformed into an image dataset of 5 categories with 106 images in each category. In order to verify the effectiveness of ILoFGAN in data augmentation under extremely few samples, only 6 images in each fault type are adopted to train the proposed model, as listed in Table I.

Fig. 6. Planetary gearbox fault simulation platform of case 1

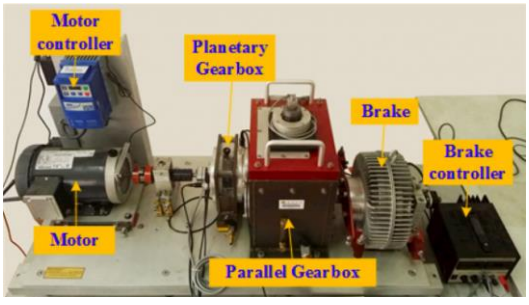


Fig. 7. Planetary gearbox fault simulation platform of case 2

TABLE I
DETAILS OF THE TIME-FREQUENCY DIAGRAM DATASETS OF GEARBOX

Diagnosis cases	Types of working condition	Data dimensions	Sample numbers Training	Testing
Case 1	healthy	[3,64,64]	**None**	100
	missing tooth (missing)	[3,64,64]	**6**	100
	root crack (crack)	[3,64,64]	**6**	100
	spalling (spall)	[3,64,64]	**6**	100
	chipping tip with severe 1 (chip1a)	[3,64,64]	**6**	100
	chipping tip with severe 2 (chip2a)	[3,64,64]	**6**	100
	chipping tip with severe 3 (chip3a)	[3,64,64]	**6**	100
	chipping tip with severe 4 (chip4a)	[3,64,64]	**6**	100
	chipping tip with severe 5 (chip5a)	[3,64,64]	**6**	100
Case 2	healthy	[3,64,64]	**None**	100
	chipping tip(chipped)	[3,64,64]	**6**	100
	missing tooth(miss)	[3,64,64]	**6**	100
	root crack(root)	[3,64,64]	**6**	100
	surface wear(surface)	[3,64,64]	**6**	100

B. Quantitative assessment of quality in generated samples

The main parameters of the proposed ILoFGAN are described as follows. The iteration number K is 100000, the steps to train the discriminator d is 1, the number of images in single set k is 3, the batch size b is 6, the selection ratio r is 0.4, and hyperparameters of loss functions are given as $\lambda_{\text{cls}}^{D} = 1$, $\lambda_{\text{cls}}^{G} = 1$, $\lambda_{\text{local}} = 0.7$. The fixed learning rate of 0.0001 is used for 50000 iterative training at first, and the linear decayed learning rate is used for 50000 iterative training. The gradient penalty regularization is used to enhance the stability of the training process. Finally, 50 generated samples are obtained for each type. The specific parameters of the model structure are listed in Table II.

TABLE II
MAIN PARAMETERS OF ILOFGAN

Modules	Descriptions	Operations	Activation functions
Encoder	Conv2D (5,3,16)	BN	Leaky-ReLU
	Conv2D (3,16,32)	BN	Leaky-ReLU
	Conv2D (3,32,64)	BN	Leaky-ReLU
	Conv2D (3,64,128)	BN	Leaky-ReLU
	MHA (channel=128, heads=8)	BN, dropout	——
	Conv2D (3,128,128)	BN	Leaky-ReLU
	Conv2D (3,128,128)	BN	Leaky-ReLU
Decoder	Conv2D (3,128,128)	BN, up sample	Leaky-ReLU
	Conv2D (3,128,128)	BN, up sample	Leaky-ReLU
	MHA (channel=128, heads=8)	BN, dropout, up sample	——
	Conv2D (3,64,128)	BN, up sample	Leaky-ReLU
	Conv2D (3,32,64)	BN, up sample	Leaky-ReLU
	Conv2D (3,16,32)	BN	Leaky-ReLU
	Conv2D (5,3,16)	BN	Tanh
Discriminator	Conv2D (5,3,64)	BN, up sample	Leaky-ReLU
	Res block (3,64,128)	BN, Avg Pool	Leaky-ReLU
	Res block (3,128,256)	BN, Avg Pool	Leaky-ReLU
	Res block (3,256,512)	BN, Avg Pool	Leaky-ReLU
	Res block (3,512,1024)	BN, Avg Pool	Leaky-ReLU
	Conv2D (1,1,1)	Avg Pool	——
	Conv2D (1,1,8)	Avg Pool	——

In addition, the performance of ILoFGAN is compared with some mainstream GANs, including LoFGAN, VAEGAN, WGAN_GP, ACGAN and DCGAN, respectively. All of these GANs are trained with 6 fault samples in each type and generate fault samples on the gearbox datasets of the 2 cases, as shown in **Fig. 8**. It can be clearly observed that the generated samples of ILoFGAN have the highest similarity with the original samples, which can generate fault samples with high quality. The generated samples of VAEGAN and WGAN_GP are similar to the real samples to some extent, which show transverse periodic repetition with low definition. ACGAN and DCGAN fail to generate samples due to their low model fitting ability. It is worth mentioning that compared with LoFGAN, ILoFGAN solves the generation defect in the local region of samples because the global attention weight imposed by the MHA module helps the ***LFM*** to accurately match local features between input samples, which improves the generation quality.

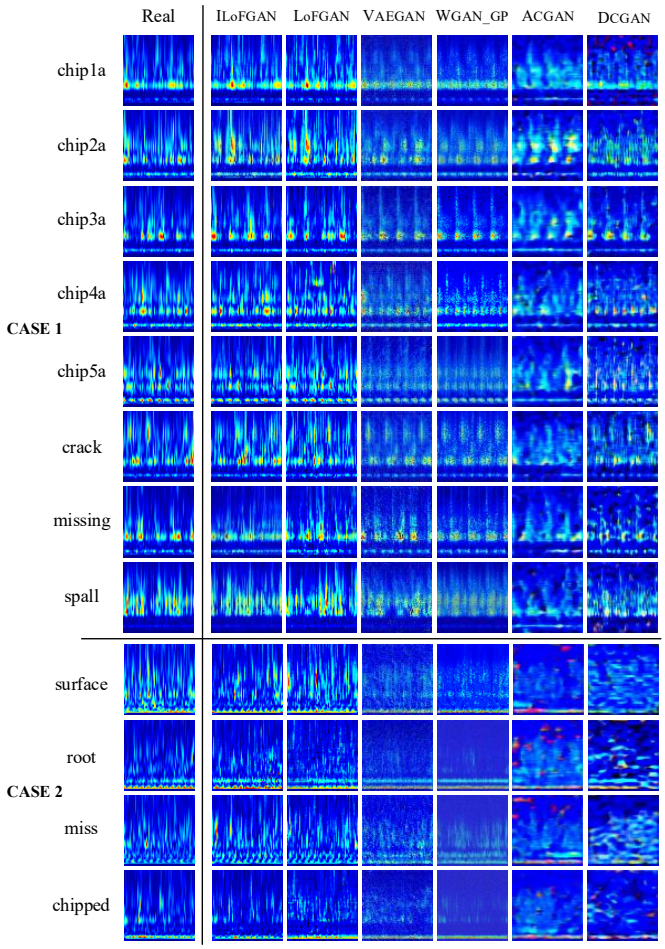


Fig. 8. Time-frequency diagrams generated by ILoFGAN and other GANs.

To fully prove the superiority of the proposed method, two quantitative indexes are calculated to evaluate the generation quality of the time-frequency diagrams given by the different GANs, including FID and SSIM. FID index [34] is used to evaluate the diversity and similarity of the generated samples,

and lower FID means better diversity and similarity. SSIM index [35, 36] can describe the similarity of two images in brightness, contrast and structure aspects with the range of [-1, 1]. When two samples are entirely identical, the value of SSIM is 1. The evaluation results of each GAN are presented in Table III.

TABLE III
THE TWO EVALUATION INDEXES OF THE GENERATED TIME-FREQUENCY DIAGRAMS OF THE DIFFERENT GANS

Different GANs	FID		SSIM (mean of all types)	
	Case 1	Case 2	Case 1	Case 2
ILoFGAN	**78.00** ↓	**123.01** ↓	**0.37** ↑	**0.3098** ↑
LoFGAN	140.98	164.36	0.35	0.3086
VAEGAN	152.79	266.14	0.19	0.11
WGAN_GP	104.96	170.22	0.34	0.24
DCGAN	190.42	398.55	0.23	0.26
ACGAN	208.75	319.96	0.15	0.24

From the FID index, we can observe that the FID of ILoFGAN are 78.00 and 123.01 respectively, which is far less than other GANs. For the SSIM index shown in **Fig. 9**, ILoFGAN also reports the best performance in most types in both cases, including chip2a, chip4a, chip5a, crack, spall, chipped, root and surface, and finally its mean SSIM values are also higher than the others in both cases. Comparing ILoFGAN and LoFGAN in particular, we can find that they almost have the same evaluation results in SSIM (mean of all types), but the FID of ILoFGAN is obviously better than LoFGAN. The result indicates that the ***LFM*** assists GANs to generate more vivid samples under the extremely limited dataset, and the multi-head attention modules increase the matching strategies of ***LFM*** in the process of local feature fusion by precisely mining more potential local features, which significantly improve the diversity of the generated samples. Therefore, it can be concluded that the proposed method is superior to other GANs in terms of generation diversity and similarity, and also in the generalibility for few-shot sample generation.

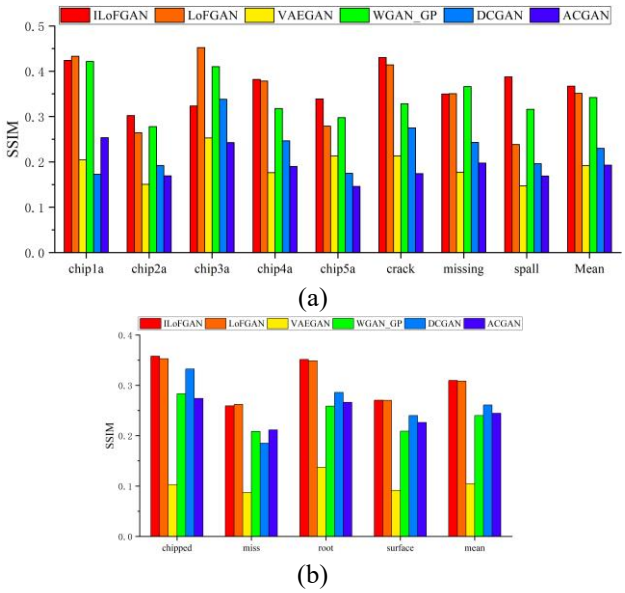


Fig.9. SSIM evaluation results of the different GANs of (a) case 1 (b) case 2

In order to further discuss the influence of the hyper-parameter setting in the ILoFGAN on the generation quality, several comparative experiments are conducted. We study the influence of the selection ratio r on the FID index of the generated samples of case 1. The selection ratio r decides the number of fused local representations in f_{base}. As shown in **Fig. 10(a)**, the FID will rise when r is either low or high. If r is low, only few local representations will be selected in ***LFM***, and the generator will output a large number of copies of the original samples, which results in a decreased generation diversity. On the contrary, when all local representations in f_{base} are selected to fuse with ϕ_{ref}, it will aggravate the uncertainty and instability of the fusion process, thus weakening the similarity of the generated samples. We discuss the regularization parameters of the local reconstruction loss λ_{local} in the same way, as shown in **Fig. 10(b)**. The local reconstruction loss utilizes the roughly fused image $\boldsymbol{LFM}(X,\alpha)$ to constrain the similarity of the generated samples. In addition, it can effectively improve the quality of the image generation, and the generated samples achieve the best quality when λ_{local} is close to 0.8. When λ_{local} is high, the dynamic balance of adversarial training will be affected and weaken the influence of other losses on the training process, which will reduce the generation quality.

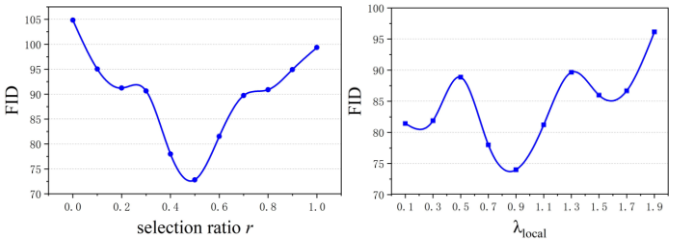


Fig.10. FID of samples generated by ILoFGAN change with (a) selection ratio r (b) λ_{local}

C. Fault diagnosis of planetary gearbox based on data augmentation

In this part, the generated samples of ILoFGAN mentioned above are used to augment limited datasets, aiming to improve the classification accuracy of fault diagnosis. Six comparative experiments of each case are conducted to prove the impact of the generated samples on fault diagnosis accuracy. The specific details of the dataset used for fault classification network are shown in Table IV. There are 70 healthy samples and 20 real samples of fault type in each sub-case, and the generated samples of fault type increase from 0 to 50 gradually. In addition, 30 samples are allocated in each type as the test set of the classification network.

For the sake of verifying the effectiveness of the ILoFGAN, Resnet50 [37], Mobilenet [38] and Googlenet [39] are used for fault diagnosis, respectively. The diagnosis accuracies of the two cases are displayed in **Fig. 11**, where the x-coordinate

represents the number of generated samples. They both show that when there is no generated sample, the accuracy of each classifier is low due to the limited samples in the dataset. With the increase of the generated samples, the accuracy of all classifiers increases, especially when the dataset is augmented with 10 or 20 generated images. It proves that the generated samples of ILoFGAN can assist feature extraction of CNN when the limited dataset is augmented, thus improving the accuracy of fault classification, which has the potential to solve the problem of few samples to some extent.

TABLE IV
CASES FOR FAULT DIAGNOSIS OF AUGMENTED PLANETARY GEARBOX DATASET

Experiment sub-cases	Generated sample number of each fault type	Real sample number of each fault type	Sample number of healthy type
sub-case 1	0	20	70
sub-case 2	10	20	70
sub-case 3	20	20	70
sub-case 4	30	20	70
sub-case 5	40	20	70
sub-case 6	50	20	70

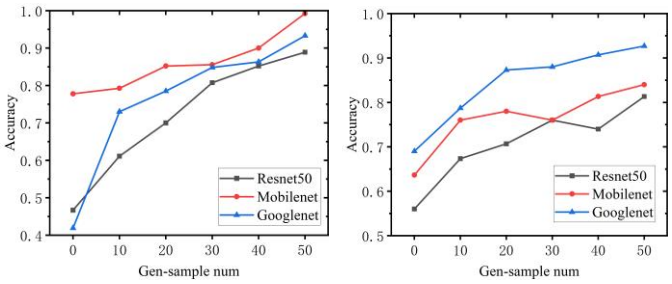


Fig. 11. Diagnosis accuracies of different classifiers based on data augmentation of (a) case 1 (b) case 2

Finally, the generated samples and real samples of case 1 are mixed at a ratio of 1:1, which are input into the same CNN to extract high-dimensional features. T-SNE algorithm is adopted for dimensionality reduction and visualization, as shown in **Fig.12**. We can see that the generated samples of ILoFGAN and the real samples are basically clustered together in each type, and the distribution of a single type is concentrated. As to LoFGAN, only part of generated samples and real samples show a good clustering effect, and other generated samples of different types are mixed together. VAEGAN, WGAN_GP, DCGAN and ACGAN hardly learn the distribution characteristics of the original dataset under extremely limited samples. It can be concluded that the generated samples of the proposed method have a more similar distribution to the real samples at the feature level.

V. CONCLUSION

To improve the accuracy of fault diagnosis for the planetary gearbox under extremely limited samples, ILoFGAN is proposed for the data augmentation task. The local fusion module is used to fully utilize extremely limited samples and fuse the local feature. Subsequently, the generator embedded with the MHA module is constructed, which can effectively improve the accuracy and flexibility of the feature fusion process.

Performance of the proposed method is illustrated by two cases of fault vibration signals of the planetary gearbox. The results show that ILoFGAN can generate sufficient fault samples of the planetary gearbox with higher quality and diversity, and effectively enhance the recognition accuracy of the classification network, which presents superior performance to most current GANs. The fault diagnosis research under few samples is more adaptive to the current industrial real scene. It has the potential to significantly reduce the cost of diagnosis, which is worth further exploration. In the future, we will further optimize the model to improve efficiency and combine fault diagnosis with the meta-learning strategy in limited samples of unknown types for further research.

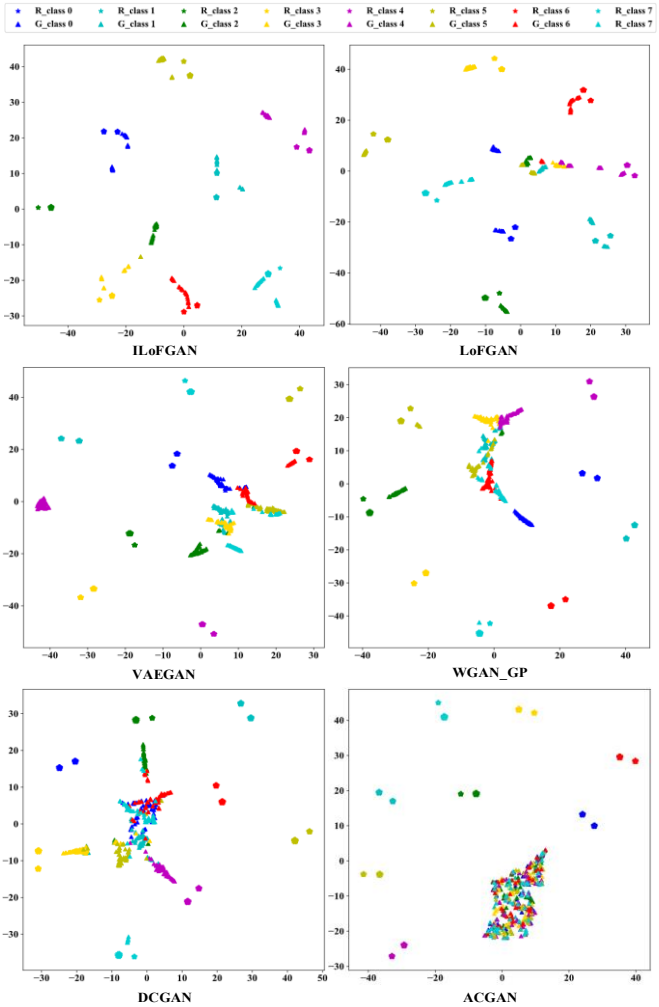


Fig. 12. T-SNE visualization results of real and generated samples.

REFERENCES

[1] Y. B. Li, X. Z. Wang, S. B. Si, and S. Q. Huang, "Entropy Based Fault Classification Using the Case Western Reserve University Data: A Benchmark Study," IEEE Transactions on Reliability, vol. 69, no. 2, pp. 754-767, Jun 2020.

[2] J. G. Miao, J. Y. Wang, and Q. Miao, "An Enhanced Multifeature Fusion Method for Rotating Component Fault Diagnosis in Different Working Conditions," IEEE Transactions on Reliability, vol. 70, no. 4, pp. 1611-1620, Dec 2021.

[3] Z. Y. He, H. D. Shao, P. Wang, J. Lin, J. S. Cheng, and Y. Yang, "Deep transfer multi-wavelet auto-encoder for intelligent fault diagnosis of gearbox with few target training samples," Knowledge-Based Systems, vol. 191, Mar 2020, Art no. 105313.

[4] H. D. Shao, M. Xia, J. F. Wan, and C. W. de Silva, "Modified Stacked Autoencoder Using Adaptive Morlet Wavelet for Intelligent Fault Diagnosis of Rotating Machinery," IEEE-ASME Transactions on Mechatronics, vol. 27, no. 1, pp. 24-33, Feb 2022.

[5] H. Wang, Z. Liu, and T. Ai. "Long-Range Dependencies Learning Based on Non-Local 1D-Convolutional Neural Network for Rolling Bearing Fault Diagnosis". Journal of Dynamics, Monitoring and Diagnostics, vol. 1, no. 3, pp. 148-159, Sep 2022.

[6] A. Q. He and X. N. Jin, "Deep Variational Autoencoder Classifier for Intelligent Fault Diagnosis Adaptive to Unseen Fault Categories," IEEE Transactions on Reliability, vol. 70, no. 4, pp. 1581-1595, Dec 2021.

[7] A. Dutta, R. Manral, P. Mitra, and R. Mall, "Hierarchically Localizing Software Faults Using DNN," IEEE Transactions on Reliability, vol. 69, no. 4, pp. 1267-1292, Dec 2020.

[8] C. Wang, J. Liu, and E. Zio. "A Modified Generative Adversarial Network for Fault Diagnosis in High-Speed Train Components With Imbalanced and Heterogeneous Monitoring Data". Journal of Dynamics, Monitoring and Diagnostics, vol. 1, no. 2, pp. 84-92, Apr 2022.

[9] W. Zhang, X. Li, X. D. Jia, H. Ma, Z. Luo, and X. Li, "Machinery fault diagnosis with imbalanced data using deep generative adversarial networks," Measurement, vol. 152, Feb 2020, Art no. 107377.

[10] X. Li, J. Cheng, H. D. Shao, K. Liu, and B. P. Cai, "A Fusion CWSMM-Based Framework for Rotating Machinery Fault Diagnosis Under Strong Interference and Imbalanced Case," IEEE Transactions on Industrial Informatics, vol. 18, no. 8, pp. 5180-5189, Aug 2022.

[11] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," Journal of Big Data, vol. 6, no. 1, Jul 2019, Art no. 60.

[12] H. X. Zhang and M. F. Li, "RWO-Sampling: A random walk over-sampling approach to imbalanced data classification," Information Fusion, vol. 20, pp. 99-116, Nov 2014.

[13] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," Journal of Artificial Intelligence Research, vol. 61, pp. 863-905, 2018.

[14] H. B. He, Y. Bai, E. A. Garcia, and S. T. Li, "ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning," in International Joint Conference on Neural Networks, Hong Kong, PEOPLES R CHINA, Jun 01-08 2008, in IEEE International Joint Conference on Neural Networks (IJCNN), 2008, pp. 1322-1328.

[15] I. Goodfellow et al., "Generative Adversarial Networks," Communications of the Acm, vol. 63, no. 11, pp. 139-144, Nov 2020.

[16] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein Generative Adversarial Networks," in 34th International Conference on Machine Learning, Sydney, AUSTRALIA, Aug 06-11 2017, vol. 70, in Proceedings of Machine Learning Research, 2017.

[17] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved Training of Wasserstein GANs," in 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, Dec 04-09 2017, vol. 30, in Advances in Neural Information Processing Systems, 2017.

[18] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," arXiv preprint arXiv:1511.06434, 2015.

[19] A. Odena, C. Olah, and J. Shlens, "Conditional Image Synthesis with Auxiliary Classifier GANs," in 34th International Conference on Machine Learning, Sydney, AUSTRALIA, Aug 06-11 2017, vol. 70, in Proceedings of Machine Learning Research, 2017.

[20] S. W. Liu, H. K. Jiang, Z. H. Wu, and X. Q. Li, "Rolling bearing fault diagnosis using variational autoencoding generative adversarial networks with deep regret analysis," Measurement, vol. 168, Jan 2021, Art no. 108371.

[21] S. Y. Shao, P. Wang, and R. Q. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," Computers in Industry, vol. 106, pp. 85-93, Apr 2019.

[22] Q. Li, L. Chen, C. Q. Shen, B. R. Yang, and Z. K. Zhu, "Enhanced generative adversarial networks for fault diagnosis of rotating machinery with imbalanced data," Measurement Science and Technology, vol. 30, no. 11, Nov 2019, Art no. 115005.

[23] J. R. Wang, B. K. Han, H. Q. Bao, M. Y. Wang, Z. Y. Chu, and Y. W. Shen, "Data augment method for machine fault diagnosis using conditional generative adversarial networks," Proceedings of the Institution of Mechanical Engineers Part D-Journal of Automobile Engineering, vol. 234, no. 12, pp. 2719-2727, Oct 2020.

[24] J. R. Wang, S. M. Li, B. K. Han, Z. H. An, H. Q. Bao, and S. S. Ji, "Generalization of Deep Neural Networks for Imbalanced Fault Classification of Machinery Using Generative Adversarial Networks," IEEE Access, vol. 7, pp. 111168-111180, 2019.

[25] R. G. Wang, S. H. Zhang, Z. Y. Chen, and W. H. Li, "Enhanced generative adversarial network for extremely imbalanced fault diagnosis of rotating machine," Measurement, vol. 180, Aug 2021, Art no. 109467.

[26] P. Peng, Y. Wang, W. J. Zhang, Y. Zhang, and H. M. Zhang, "Imbalanced Process Fault Diagnosis Using Enhanced Auxiliary Classifier GAN," in Chinese Automation Congress (CAC), Shanghai, PEOPLES R CHINA, Nov 06-08 2020, in Chinese Automation Congress, 2020, pp. 313-316.

[27] Z. X. Li, T. S. Zheng, Y. Wang, Z. Cao, Z. Q. Guo, and H. Y. Fu, "A Novel Method for Imbalanced Fault Diagnosis of Rotating Machinery Based on Generative Adversarial Networks," IEEE Transactions on Instrumentation and Measurement, vol. 70, 2021, Art no. 3500417.

[28] S. Dixit, N. K. Verma, and A. K. Ghosh, "Intelligent Fault Diagnosis of Rotary Machines: Conditional Auxiliary Classifier GAN Coupled With Meta Learning Using Limited Data," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2021, Art no. 3517811.

[29] J. G. Miao, J. Y. Wang, D. C. Zhang, and Q. Miao, "Improved Generative Adversarial Network for Rotating Component Fault Diagnosis in Scenarios With Extremely Limited Data," IEEE Transactions on Instrumentation and Measurement, vol. 71, 2022, Art no. 3500213.

[30] Z. Gu, W. Li, J. Huo, L. Wang, and Y. Gao, "Lofgan: Fusing local representations for few-shot image generation," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 8463-8471.

[31] Z. Y. Niu, G. Q. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," Neurocomputing, vol. 452, pp. 48-62, Sep 2021.

[32] P. Cao, S. Zhang, and J. Tang, "Pre-Processing-Free Gear Fault Diagnosis Using Small Datasets with Deep Convolutional Neural Network-Based Transfer Learning," IEEE Access, pp. 26241-26253, 2017.

[33] S. Shao, S. McAleer, R. Yan and P. Baldi, "Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning," IEEE Transactions on Industrial Informatics, vol. 15, no. 4, pp. 2446-2455, Apr 2019.

[34] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," arXiv preprint arXiv:1706.08500, 2017.

[35] W. Zhou, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on, 2003.

[36] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, 2004.

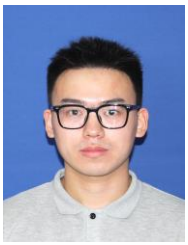
[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016 2016, pp. 770-778.

[38] A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017

[39] C. Szegedy, L. Wei, Y. Jia, P. Sermanet, and A. Rabinovich, "Going deeper with convolutions," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015 pp. 1-9.

Mingzhi Chen received the B.S. degree in Mechanical Design, Manufacturing, and Automation from Hunan University, Changsha, China, in 2022. He is currently pursuing the Ph.D. degree in Mechanical Engineering with the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China. His research interests include deep learning theories and machinery fault diagnosis.

Haidong Shao received the B.S. degree in Electrical Engineering and Automation and the Ph.D. degree in Vehicle Operation Engineering from Northwestern Polytechnical University, Xi'an, China, in 2013 and 2018, respectively. He is currently an Associate Professor in the College of Mechanical and Vehicle Engineering at Hunan University, Changsha, China. From 2019 to 2021, he was a Postdoctoral Fellow with the Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden. His current research interests include operation and maintenance, data mining, information fusion, and industrial internet. He won the World's Top 2% Scientists List, Elsevier Highly Cited Chinese Researchers, IOP Publishing Top Cited Author Award and IEEE ICCSE 2021 Best Paper Award. Thus far, he has published more than 60+ SCI-indexed papers, including 18 ESI Highly Cited Papers and 3 ESI Hot Papers. He has served as the Guest Editors of nine SCI-indexed journals and led various prestigious international conferences, including the Program Chairs of IEEE ICCSE 2021 and AMDS 2022; the Program Committees of CMVIT 2021, ICFPMCE 2021, CIMIA 2021, ICMACS 2021, ICCAIS 2021, IEECP 2022, and ICBDAI 2022; the Session Chairs of ICSMD 2021, ICAS 2021, IEEE Global Rel&PHM-2022 and IEEE ICPHM 2022.



Haoxuan Dou was born in Weifang, Shandong Province, China in 2001. He is now studying for the B.S. degree in Mechanical Design Manufacture and Automation at College of Mechanical and Vehicle Engineering, Hunan University, Changsha, China. He is interested in the applications of deep learning in advanced rotating mechanical fault diagnosis..

Wei Li received the B.S. degree in Mechanical Design, Manufacturing and Automation from Huazhong Agricultural University, Wuhan, China in 2020. He is now studying for the M.S. degree in Mechanical Engineering at College of Mechanical and Vehicle Engineering, Hunan University, Changsha, China. His research interests are deep learning theories and their applications in rotating machinery fault diagnosis.

Bin Liu received the B.S. degree in Automation from Zhejiang University, Hangzhou, China, in 2013, and the Ph.D. degree in Industrial Engineering from City University of Hong Kong, Hong Kong, China, in 2017. He is currently working as a lecturer in the Department of Management Science at University of Strathclyde, Glasgow, UK. Before that he was working as a Postdoctoral Fellow in University of Waterloo, Canada. His research interests include reliability and risk analysis, maintenance modeling, prognostic and health management, and decision making under uncertainty.