

ORIGINAL ARTICLE

Open Access



Intelligent Diagnosis Method for Typical Co-frequency Vibration Faults of Rotating Machinery Based on SAE and Ensembled ResNet-SVM

Xiancheng Zhang¹, Xin Pan^{1,2*} , Hao Zeng¹ and Haofu Zhou³

Abstract

Intelligent fault diagnosis is an important method in rotating machinery fault diagnosis and equipment health management. To deal with co-frequency vibration faults, a type of typical fault in rotating machinery, this paper proposes a fault diagnosis method based on the stacked autoencoder (SAE) and ensembled ResNet-SVM. Furthermore, the time- and frequency-domain features of several co-frequency vibration faults are summarized based on the mechanism analysis and calculated using actual vibration data. To realize and validate the high-precision diagnosis method of rotating equipment with co-frequency faults proposed in this study, the following three criteria are required: First, to improve the effectiveness and robustness of the ensembled model and the sliding window using data augmentation, adding noise, autoencoder (AE) and SAE methods are analyzed in terms of principle and practical effects. Second, ResNet is used as the feature extractor for the ensembled ResNet-SVM model. Feature extraction is carried out twice, and the extracted co-frequency fault features are more comprehensive. Finally, the data augmentation method and ensemble ResNet-SVM are combined for fault diagnosis and compared with other methods. The experimental results show that the accuracy of the proposed method can exceed 99.9%.

Keywords Co-frequency vibration, Data argumentation, Ensembled ResNet-SVM, High precision fault diagnosis

1 Introduction

The co-frequency fault is the most common type of fault in rotating machinery, and its reliable diagnosis has practical engineering significance. Owing to the similarity of the co-frequency fault, misdiagnosis frequently

occurs, which leads to serious consequences. Therefore, the highly accurate diagnosis of co-frequency faults is a critical problem that needs to be solved [1–3]. Meanwhile, with the development and gradual improvement of deep learning frameworks, intelligent fault diagnosis based on deep learning methods has become a research hotspot in recent years. At the same time, the application of deep learning in rotating machinery is also increasing [4–7]. The co-frequency faults of rotating machinery mainly include imbalance, misalignment, and looseness faults [8–10]. Ma et al. [11] employed ensemble learning to identify various faults in a rotor-bearing system, which included three methods: a convolutional residual network, deep confidence network, and deep autoencoder (AE) [12]. On the other hand, Guo et al. applied a combination model of continuous wavelet transform and

*Correspondence:

Xin Pan
panxin@mail.buct.edu.cn

¹ Beijing Key Laboratory of Health Monitoring and Self-recovery for High-end Mechanical Equipment, Beijing University of Chemical Technology, Beijing 100029, China

² Engineering Research Center of Chemical Safety (Ministry of Education), Beijing University of Chemical Technology, Beijing 100029, China

³ Engineering Management Department of China Three Gorges New Energy (Group) Co., Ltd., Beijing 101199, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

convolutional neural networks (CNN) to co-frequency fault diagnosis [13].

Deep learning and machine learning frameworks have influenced various fields attributed to high performance and low expertise. Based on the Internet of Things (IoT), data-driven fault diagnosis and health status monitoring of machinery have adopted machine learning and deep learning for more than a decade [14], especially with particular success in gearbox and bearing fault diagnosis [15, 16]. The traditional fault diagnosis method uses only a support vector machine (SVM) for single-domain analysis. In contrast, to analyze the inner features of the fault signal, Yan et al. [17] combined time-domain statistical features, frequency-domain features, and signal modal decomposition features. For the adverse operating conditions of bearings, batch-normalized CNNs eliminate the differences in feature distribution owing to data imbalance and are used to improve diagnostic accuracy [18]. Transferring methods from the image recognition field to the fault diagnosis field can be performed in two ways. First, the model is trained on a non-faulty dataset but evaluated on a faulty dataset. For example, the ResNet-50 model is trained in ImageNet and then transferred the trained model to the time-domain vibration signal test dataset of the signal converted into RGB images [19]. Second, the model training and testing processes are applied to faulty datasets. Similarly, based on the ResNet model, He et al. [20] converted the vibration signals into gray images using ResNet-50 and ResNet-101 in multiple concurrent fault diagnoses of parallel shaft gearboxes. To the authors' knowledge, the use of ResNet for fault feature extraction is mainly based on layers 50 and 101 instead of ResNet-18, due to higher recognition accuracy of the ResNet network as the number of layers increases [21].

In practical engineering, the normal operation equipment data account for most of the data, and the fault data only account for a small portion or even less. Hence, it is necessary to perform data augmentation using vibration data [22–24]. In image recognition, speech recognition, machine translation, etc., data augmentation is extensively used to improve the diagnostic accuracy of algorithms and model generalization. Gong et al. [25] transformed a vibration signal into pictures and enhanced the images through geometric transformation, random cropping, and average blurring. In recent years, generative neural network (GAN)-based data augmentation methods have become a hot research topic. Shao et al. [26] designed auxiliary classifier GAN structures based on one-dimensional convolutional layers and used this architecture to generate sufficient real samples on the induction motor dataset. Azamfar et al. [27] proposed the MoGAN oversampling technique for learning the

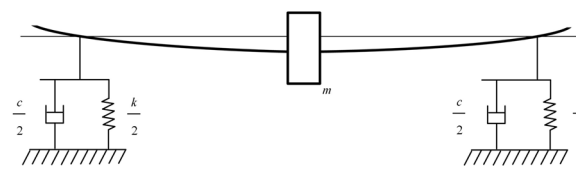


Figure 1 Typical rotor dynamics model

joint distribution of the minority and majority samples in adversarial learning. The discriminator in adversarial learning not only discriminates between the truth and falsity of the generated samples but also acts as a multi-classifier for faults. In addition to the GAN method, AE-based technology has been used in data augmentation to diagnose faults. Liu et al. [28] combined variational AE and GAN to learn the high-level features of rolling bearings and improved the effectiveness and robustness of the model via a joint analysis of the discriminator and depth repentance. Tang et al. [29] extended the AE method to various operating conditions by adding an adaptive module and establishing an adaptive transmission AE structure. The AE-based method has a more straightforward structure and is easier to train than the GAN method. There are three points to consider when using the GAN and AE. First, the generated samples can be similar but not identical to the input samples. Second, the same samples will be generated differently after applying the generator framework. Finally, this method can reduce the effect of data imbalance on the input data.

This paper is organized as follows: Section 2 introduces the mechanism and features of co-frequency vibration faults. Section 3 describes the work related to intelligent diagnosis and data augmentation, and the ResNet-SVM algorithm is introduced in detail. Section 4 verifies the performance of the ensemble classifier using actual co-frequency fault data, and conclusions are presented in Section 5.

2 Co-frequency Fault Mechanism and Features

2.1 Mechanism Analysis of Co-frequency Vibration Faults

The dynamics model for the rotor-bearing system is illustrated in Figure 1, and the rotor dynamics equation can be expressed by the following equation:

$$M\ddot{X} + C\dot{X} + KX = F(t), \quad (1)$$

where M , C and K are the mass, damping, and stiffness of the rotor system, X is the displacement response of the rotor system; $F(t)$ is the excitation force of the rotor. For a rotor-bearing system with n degrees of freedom, $X \in \mathbf{R}^n$. The simplified 2-dimensional rotor dynamic model shows that the displacement response and

excitation force can be decomposed into the horizontal and vertical directions.

Rotor imbalance faults are the most common and major vibration faults among rotor faults. The rotor system, including the rotor accessories, has a certain imbalance during processing and manufacturing, which develops into a fault state under both sudden and gradual circumstances. In general, unbalanced mass m' and eccentricity e are the main consequences of vibration exceeding the limits. Therefore, the resulting excitation force is given by the following equation:

$$F_1 = m'e\omega^2, \tag{2}$$

where ω is the rotor speed. From Eq. (2), it can be concluded that the spectrum of the rotor unbalance fault is dominated by the rotational frequency.

In the case of well-aligned rotors, the coupling transmits only circumferential forces such that multiple rotors rotate steadily together between or between the rotor and the prime mover. The axial and radial alternating forces caused by misalignment failures lead to excessive vibrations of the system in both directions. Therefore, the additional axial and radial forces can damage the bearing prematurely. The misalignment forms include angular misalignment, parallel misalignment, and comprehensive misalignment.

$$\begin{cases} F_{2x} = 2m\Delta \sin(2\omega - \varphi'), \\ F_{2y} = 2m\Delta \cos(2\omega - \varphi'), \end{cases} \tag{3}$$

where Δ is the product of the misalignment and speed. Eq. (3) shows that the characteristic frequency of the misalignment fault is dominated by twice its rotational frequency.

Most of the looseness faults occur in the bearing connection, bolts, or other connections at the base or foot of the equipment. Looseness causes large gaps in the joint surface and simultaneously leads to low damping and insufficient joint stiffness, resulting in excessive machinery vibration. Consequently, a very small degree of imbalance and misalignment that already exists in the rotor is amplified at this stage. Therefore, the spectrogram of the loosening fault is still theoretically dominated by the rotation frequency combined with other harmonics.

$$k = \begin{cases} k_1 & |x| \leq \Delta', \\ k_2 & |x| > \Delta'. \end{cases} \tag{4}$$

where, Δ' is the amount in the gap, and the stiffness of the rotor system varies with the change of the gap. In addition, the rotational frequency is dominant in the spectrogram and is accompanied by other multipliers.

2.2 Analytical Methods Based on Mechanistic Features

The collected raw signals are discrete in the time domain. The statistical features of time-domain signals include the root mean square (RMS), peak-to-peak, variance, kurtosis, skewness, and form factor. However, the time-domain features of vibration signals often do not entirely meet the fault diagnosis requirements, so the barycenter frequency and RMS frequency in the frequency domain features are introduced. In addition, parameters such as kurtosis, skewness, and form factor should not differ between the co-frequency faults based on this principle. However, these parameters are still considered in Figure 2.

The data collected in Section 4 and the equations in Table 1 are used to select different fault data with the same length (16384) for feature statistics. After each parameter is calculated 10 times, the mean value is obtained, as shown in Figure 2. The statistical figure shows that, although the parameters differ in terms of co-frequency faults, the feature values they contain can't be characterized. Therefore, the vibration signal analysis method is unsuitable for co-frequency fault diagnosis.

3 Data Augmentation and Intelligent Diagnostic Algorithms

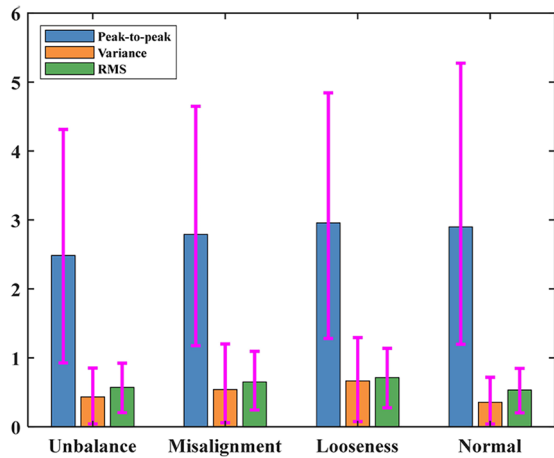
3.1 Data Augmentation

During the equipment operation, the fault state is insignificant. Even for some sudden faults such as shaft breakage, the vibration data are small. Hence, it is necessary to perform data augmentation on the fault data. The dataset after data augmentation not only enhances the model generalization, but also alleviates the data imbalance caused by the large difference in data sizes between different types of faults. Typical data augmentation methods used in rotating machinery fault diagnosis include adding noise, sliding windows, AE, and SAE methods.

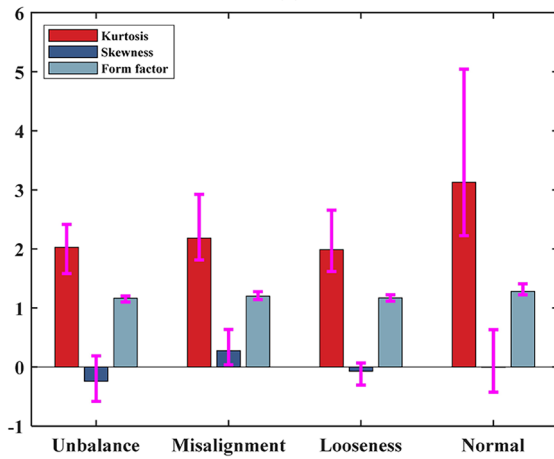
3.1.1 Sliding Window

The sliding window method is a frequently used data augmentation method for fault diagnosis. The principle here is to select different starting points from the raw vibration data and intercept the same length of data consecutively or intercept from the exact location but with different sampling. The repetition of the original dataset using the fixed window and continuous sliding interception, enhances the order of magnitude enough to satisfy the training model requirements. The sliding window process is expressed by Eq. (5)

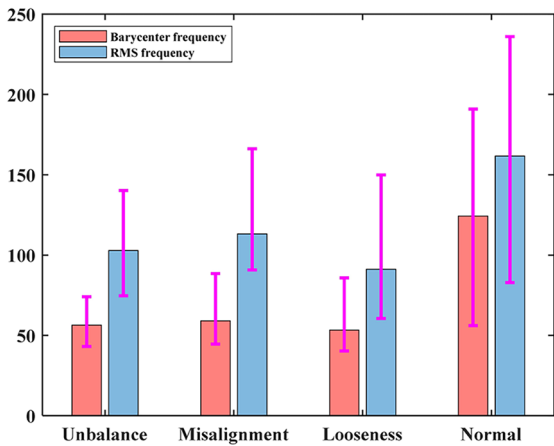
$$X * win = [x_1 \ x_2 \ \dots \ x_n] * win = \begin{bmatrix} x_{i+1} & x_{i+2} & \dots & x_{i+L} \\ \dots & \dots & \dots & \dots \\ x_{j+1} & x_{j+2} & \dots & x_{j+L} \\ \dots & \dots & \dots & \dots \\ x_{N+1} & x_{N+2} & \dots & x_{N+L} \end{bmatrix}, \tag{5}$$



(a)



(b)



(c)

Figure 2 Statistical chart of fault signal characteristics of different faults

Table 1 Common vibration signal characteristics

Domain	Parameter name	Expression
Time-domain	Peak-to-peak	$x_{p-p} = \max(x) - \min(x)$
	RMS	$x_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$
	Variance	$x_{var} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$
	Skewness	$x_{skew} = E \left[\left(\frac{x - \mu}{\sigma} \right)^3 \right]$
	Kurtosis	$x_{kurt} = E \left[\left(\frac{x - \mu}{\sigma} \right)^4 \right]$
	Form factor	$x_{w-f} = \frac{x_{rms}}{ \bar{x} }$
Frequency-domain	RMS frequency	$F_{rms} = \sqrt{\frac{1}{k-1} \sum_{k=1}^k [A(k) - \bar{A}]}$
	Barycenter frequency	$F_{b-f} = \frac{\sum_{k=1}^k f(k)A(k)}{\sum_{k=1}^k A(k)}$

$$N = \left\lfloor \frac{n - i}{L} \right\rfloor, \tag{6}$$

where the window size of the sliding signal $X(t)$, L , $*win$, refers to the sliding window operation, i is the starting point of the sliding window, and N groups of length L samples are obtained. The number of groups N can be calculated using Eq. (6), where $\lfloor a \rfloor$ denotes a rounded-down value of a .

3.1.2 Adding Noise

The vibration signal is affected by the current, environment and other factors [30]. To increase the diversity of the samples, a straightforward method is to add noise to the collected signal, such as Gaussian white noise. Given that the input signal $X(t) = [x_1(t), x_2(t), \dots, x_n(t)]$ is divided into n small segments, where the intercepted signal is x_i , after adding noise it is $\bar{x}_i^1(t)$, and the noise signal is $n_i(t)$, as shown in Eq. (7).

$$\bar{x}_i^1(t) = x_i(t) + n_i(t). \tag{7}$$

However, for the co-frequency fault, adding noise is unsuitable because the characteristic frequency of the fault is already clear in Section 2, so the spectrum is relatively single, while adding noise disrupts the original clear spectrum.

3.1.3 AE and SAE

An AE is an efficient neural network structure for learning input data features using unsupervised learning methods, which are commonly applied in the pre-training process of neural networks[31]. Furthermore, it is

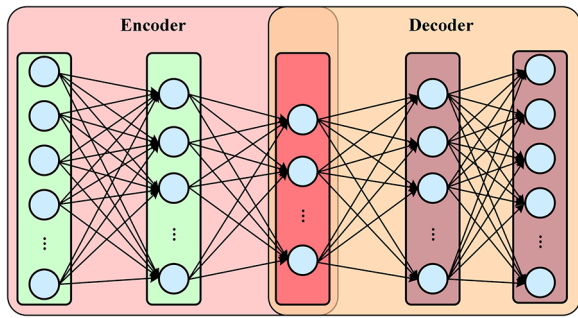


Figure 3 Typical AE network structure

also used as a generator for its encoder-decoder structure; its typical structure is shown in Figure 3.

The crucial information is extracted in the encoder part, and the input data dimensionality is reduced simultaneously. Subsequently, the decoder part receives the reduced-dimensionality data and achieves the data generation function by reducing the loss of the input layer data. Notably, excessive hidden layers lead to overfitting. In this case, the AE turns into a numerical mapping rather than discovering the underlying features. Therefore, only four hidden layers were present in the encoder structure in the data augmentation process described in Section 4.2.

The vibration signal from the coding layer to the decoding layer is expressed using Eq. (8).

$$\bar{x} = f(g(x)), \tag{8}$$

where $f(x)$ and $g(x)$ denote the decoding and encoding functions, respectively, both of which result from

layer-by-layer network superposition. Taking the decoding function as an example, it can be expressed by Eq. (9).

$$f(x + 1) = W^l f(x) + b^l, \tag{9}$$

where W^l and b^l are the weight and bias of the l -th layer.

To further improve the similarity between the generated and input data as well as the robustness and stability of the model, each layer of the network can be trained greedily layer by layer to achieve model pre-training. Each layer is trained, and then the entire structure, that is, from the input layer to the output layer, is back-propagated to train and is fine-tuned. The pre-training process is the initialization of the model because a better initialization places the parameters to be learned in a more "suitable" position. The SAE method can learn higher-level features of the input data than AE and a higher number of features from the original data, and the features are more similar. I and II in Figure 4 show the training process for each layer and the overall model, respectively.

3.2 SVM and ResNet

The SVM method is one of the most popular classification methods in machine learning. The main idea of the SVM classifier is to find the hyperplane in the categories using different methods. The location of the hyperplane is determined by the nearest sample, that is, support vectors. Two approaches can be considered for the linearly indistinguishable categories. First, samples can exist between the support vector and hyperplane, that is, a soft margin. The margin size and classification accuracy are balanced by the penalty factor C . The appropriate C

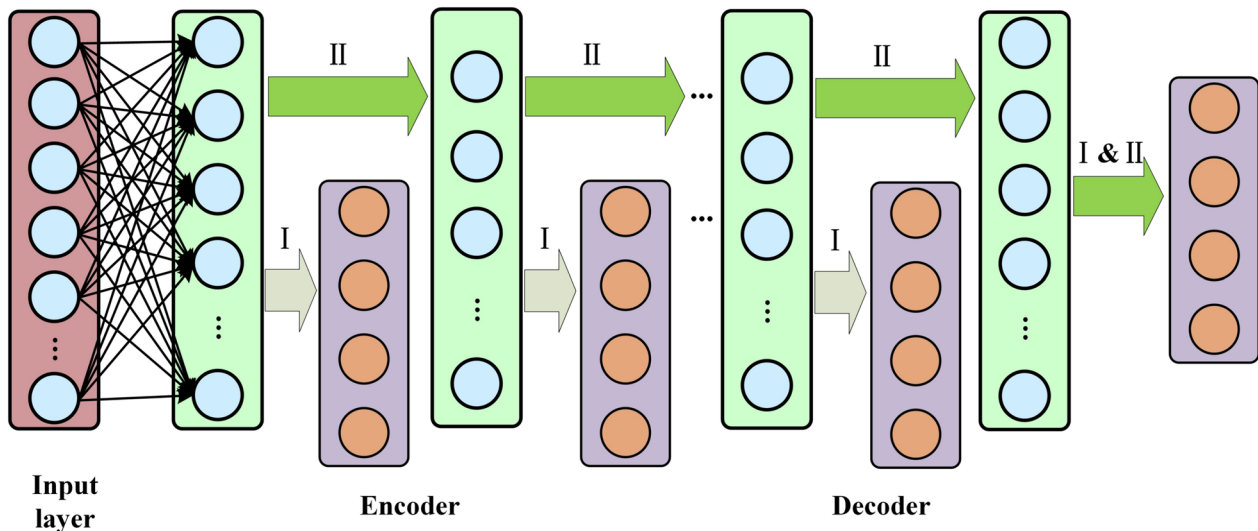


Figure 4 Typical SAE network structure

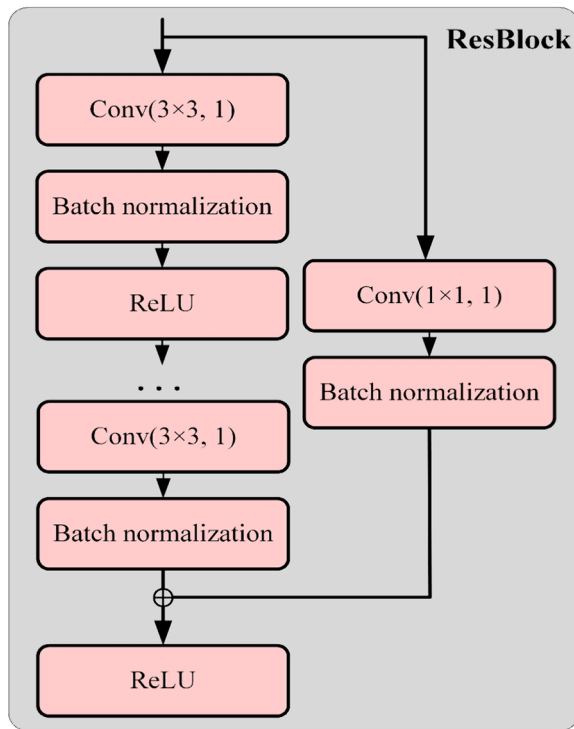


Figure 5 Residual block schematic

value is usually obtained by cross-validation, and, the smaller the C , the softer the margin and the more noise in the data. Second, the feature vector is the inner product, and then the low-dimensional feature data space is elevated to a high-dimensional space by kernel function transformation. Theoretically, increasing the low-dimensional indistinguishable data to a sufficiently high dimension can achieve absolute linear differentiation. Common kernel functions include the Gaussian kernel function, polynomial kernel function, radial basis function (RBF) kernel function, and linear kernel function. In the practical training of SVM, the optimal hyperplane is determined by combining the two methods and calculating the square of the classification error.

ResNet is a residual structure proposed to degrade the accuracy of multilayer CNN models, as shown in Figure 5. The right side of the figure is the "shortcut," and the left side is the standard CNN structure, that is, convolution-pooling-activation. The CNN structure focuses on the convolutional layer, and the convolutional computation process of the input data is shown in Figure 6. The convolution kernel Q slides over the entire area to obtain convolution values from left to right and from top to bottom. Subsequently, the values in the sliding direction are arranged to obtain the convolution result for the entire input. The sliding step and size of the convolution kernel are hyperparameters, where the convolution kernel size is generally 3×3 or 5×5 . The smaller the convolution kernel, the more localized the features, that is, the smaller the receptive field. Thus, convolution kernel Q , continuously updated with Q by backward gradient propagation, can also be regarded as the weight of data P . After the pooling layer receives the convolution result, further downsampling is performed. In the pooling layer, different kernels are used, but the kernel only finds the average or the maximum convolution result within the kernel range, that is, average pooling and maximum pooling. In this process, the kernel size of the pooling layer is hyperparametric and parameter-free, and the data dimensionality is reduced, further improving the global feature extraction. The activation function further improves the ability to fit the nonlinear features of the model.

The shortcut structure on the right side uses a convolution kernel size of 1×1 , maintaining the same channels between the input and output data, and utilizes number multiplication on the input data, which can improve the nonlinear representation of the model. For dataset x fed into the residual block, the result $F(x)$ is output via the left side and the result x is output via the right side. Because the multiplication operation does not change the data structure, both the results are superimposed to activate the output. Therefore, the residual block learned is $F(x) - x$. Because of this network structure design, a shortcut structure can increase the gradient of the deep network, thus greatly reducing the gradient explosion and dispersion.

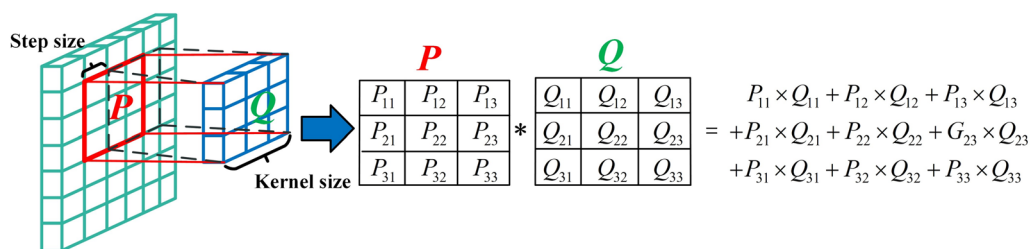


Figure 6 Convolution calculation process

Therefore, the gradient descent direction of the model is more obvious, and the model is more expressive. After superimposing multiple layers of residual blocks, global average pooling layer is applied. For the j -th category, the output is normalized using Eq. (10) Softmax function as follows.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, j = 1, 2, \dots, k. \quad (10)$$

3.3 The Proposed Ensembled ResNet-SVM Algorithm

This study proposes a novel ensembled network architecture for rotating machinery co-frequency vibration faults, known as the ensembled ResNet-SVM architecture. This architecture can be divided into three parts. The first is the data fusion. A variety of fault signals are collected through the data collector and sensors. To prevent gradient dispersion or explosion in the network caused by excessive data fluctuation, the raw data should be fused into a multidimensional dataset after normalization processing. The second is the feature extractor. Multidimensional data information is fed into the ResNet structure. It outputs the initial features, that is, the first feature extraction step, and then starts the SVM training with the second feature extraction step. The input data size is $1 \times 32 \times 32$ (channel \times height \times width), first transformed into $64 \times 32 \times 32$ through a convolutional layer; next, the transformed data each time flow through the residual block, doubling the data channel, and changing the width and height by half; finally, the size of the output feature is $512 \times 1 \times 1$. In the training ResNet stage, the residual block is followed by connecting the global average pooling and fully connected layer to output the classification probability. After calculating the loss using Softmax and cross-entropy, the parameters are propagated for further iterative optimization. The third is the SVM classification. The trained ResNet is loaded and the fully connected layer in the network is removed. ResNet reduces the input data dimension to 512. The data after feature extraction are then entered the SVM multi-classifier. The appropriate parameters are adjusted in the SVM to determine the fitting support vectors. While testing the SVM, the support vectors and parameters are directly loaded. This part comprises the second feature extraction and the classification result output. The structures of Parts 2 and 3 are shown in Figure 7.

According to the above analysis, the proposed ResNet-SVM algorithm has the following advantages.

- (1) The features obtained after two feature extractions are more comprehensive, and this method is helpful for exploring the features of co-frequency faults.
- (2) When the model parameters need to be updated, the proposed integrated model, which can be retrained only for the classifier part, that is, by increasing the support vectors of the SVM, saves time and computing resources compared with the iterative training of neural networks.
- (3) This method can be applied to the fault diagnosis of a small number of fault samples based on the SAE.

4 Experimental Validation

The data augmentation model and ensemble ResNet-SVM model were validated on the experimental rig. All network architectures were based on Python 3.8, using the open-source learning library Pytorch 1.8 as the back-end learning and deep learning engine to build the network. All training and testing of the networks were performed on workstations using an Intel(R) Xeon(R) W-2255 CPU, 64 GB RAM, and NVIDIA GTX Geforce 2080 Ti GPU with Windows 10 as the operating system. The experimental flow is illustrated in Figure 8.

4.1 Experiment Preparation and Data Collection

The test rig validated the experiments in the following figures. Figure 9 shows the schematic illustration, and Figure 10 shows the actual experimental pictures. The first-order natural frequency of the rotor test rig was approximately 45 Hz, and the fixed speed frequency was 30 Hz. In the experiment, an eddy current sensor was used as the phase signal, and four acceleration sensors were placed in the horizontal and vertical directions of the bearing position. Different faults were alternately experimented on the test rig, and the design of different faults is as follows. An unbalanced fault was simulated by adding small studs on the disk in the middle of the rotor; a misalignment fault was set at the coupling by twisting the knob to adjust the misalignment level, and the loose fault was simulated by changing the tightening degree of the bolts at the bearing position away from the motor end. The BH7000 data collector was used, the sampling frequency was set to 25600 Hz, the sampling point was 16384, and a sliding window processed the collected data to perform the first data augmentation process. The specific magnitudes of the faults, raw vibration data size, and divided dataset size are listed in Table 2.

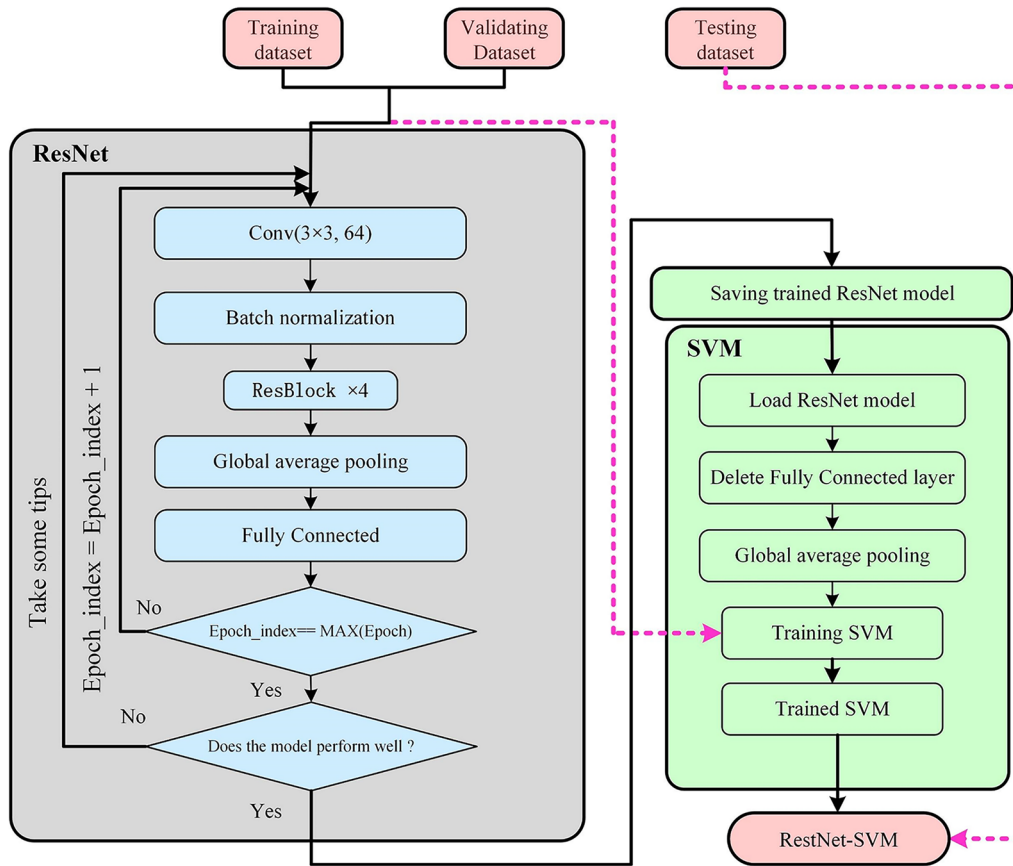


Figure 7 ResNet-SVM actual training flow chart

4.2 Experimental Comparison between Different Data Augmentation Methods

As described in Section 3.1, both AE and SAE methods were applied for data augmentation in the experiments. The two methods used the same encoder-decoder structure, with the number of neurons varying from 1024-512-256-128-64-128-256-512-1024. In addition, both data augmentation models have model loss as the detection object in each model, and use the learning rate reduction strategy to overcome the model falling into a local minimum. The dataset divided in Section 4.1 was used to train the model for each fault separately with the data augmentation method. The training process is shown in Figure 11, and the training results are shown in Figures 12 and 13.

In the model training process, the network layers, activation functions, optimizer, and optimizer parameters were maintained the same for both encoders, as listed in Table 3. The loss of the AE and SAE models decreased with increasing iteration time and gradually stabilized in the later stage. Only the first 20 epochs are shown in Figure 11 to facilitate observation of the downward trend.

In terms of the performance of the two encoders from the perspective of loss reduction, Eq. (11) is used to calculate the average loss reduction amplitude on the dataset, where l_{in} represents the initial loss, and l_s represents the loss after stabilization.

$$\begin{cases} l_i = \frac{l_{in} - l_s}{l_{in}} \\ l_m = \frac{\sum_{i=1}^n l_i}{n} \end{cases} \quad (11)$$

The loss reduction for each fault was first calculated for each encoder, followed by the average loss reduction for all encoders of the same type. For the SAE, because the greedy algorithm is used to train each layer first, the average loss reduction is only trained for the whole model at the end when calculating the average loss reduction. The average reduction in the AE training set loss was 65.8%, the validation set average reduction loss was 47%, while the training set loss average reduction and validation set loss of SAE were 93.6% and 85.2%, respectively. Meanwhile, the average number of iterations of the AE model during the model training

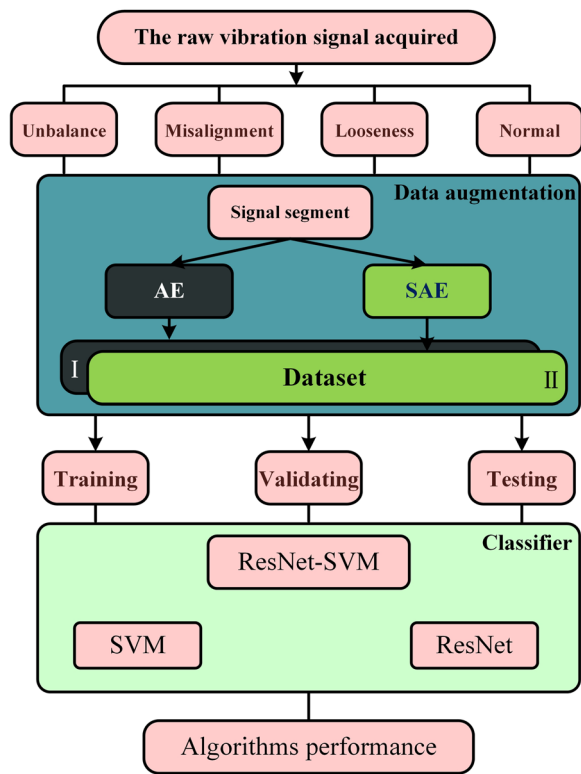


Figure 8 Experimental flow chart

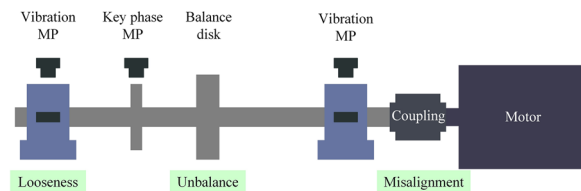


Figure 9 Schematic illustration of the rotor-bearing system with the co-frequency fault simulation test rig

process was approximately 10 epochs to reach stability, whereas the SAE model only required 5 epochs. In contrast to the data augmentation results, the output of the SAE model can simulate not only the overall pattern in the fault feature image but also the details of the grain in the image, as shown in Figure 13. From the above perspectives, the SAE model can generate more detailed and full-scale fault features, and the results for the same fault type are more similar after data augmentation. However, the training process of the SAE model is more complicated, the training time is longer, the model is sensitive to parameter adjustment, and it requires more effort to tune parameters; therefore, the AE model is still retained in the following experiments.

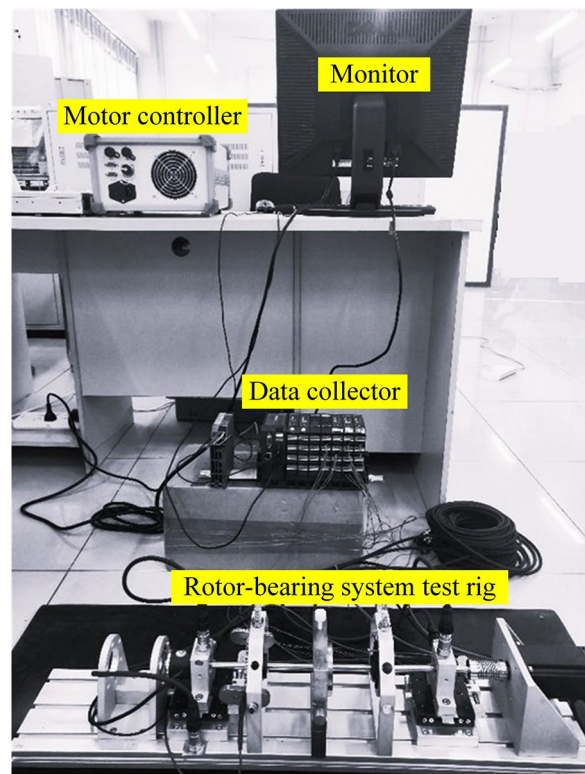


Figure 10 Actual images of the rotor-bearing system with the co-frequency failure simulation test rig

4.3 Fault Diagnosis Methods Validation

After AE and SAE, each fault data point was expanded to $1.5 \times$ the original, and 80% of the overall dataset was divided as the training set, leaving 20% as the validation set. The batch size of the validation set was 128. After model training, the entire validation dataset was used as the test set to test model performance.

To verify and evaluate the accuracy of the proposed ensemble ResNet-SVM for co-frequency fault diagnosis, the algorithm is viewed from the perspective of the precision ratio and recall ratio for different faults in this study. The F1 score was obtained from the harmonic average of precision and recall. For the binary classification results the fault precision ratio, recall ratio, and F1 score were calculated using Eq. (12) as follows:

$$\begin{cases} P = \frac{TP}{TP+FP}, \\ R = \frac{TP}{TP+FN}, \\ F1score = 2 \frac{P \cdot R}{P+R}, \end{cases} \quad (12)$$

where TP represents the number of true positives, FP represents the number of false positives, and FN represents the number of false positives. For the multi-classification

Table 2 Basic parameters for different faults and fault dataset

Fault	Magnitude	Raw data size	Data size after sliding window
Unbalance	2 g/ $\angle 0^\circ$	4*47*16384	3008*1024
Misalignment	0.75 mm	4*64*16384	4096*1024
Looseness	1 turn	4*43*16384	2752*1024
Normal	\	4*25*16384	1600*1024

results, the F1 score of each class calculated and the average were calculated.

4.3.1 SVM and ResNet Fault Diagnosis Validation

For contrast tests, the fault diagnosis performance of the SVM and ResNet models must be evaluated. With scikit-learn1.0.1, the size of the penalty coefficient C

and the selection of the kernel function greatly influence the performance and training time of the model. The RBF was chosen as the kernel function to provide a better fit for nonlinear vibration coupling.

The penalty coefficients were set to series values, and the F1 score was used to evaluate the classification ability of the model. The evaluation process is illustrated in Figure 14. As illustrated, the diagnostic performance of the model gradually increased as the penalty coefficient increased for both datasets, and the inflection points of the two lines had the same position and trend. When the penalty coefficient C reached 0.04, the performance of the SVM on both datasets gradually became the same and stabilized using this value as the threshold. The penalty factor for the SAE dataset was larger than that for the AE dataset at the same F1 score, indicating that the SAE data have less noise. The average

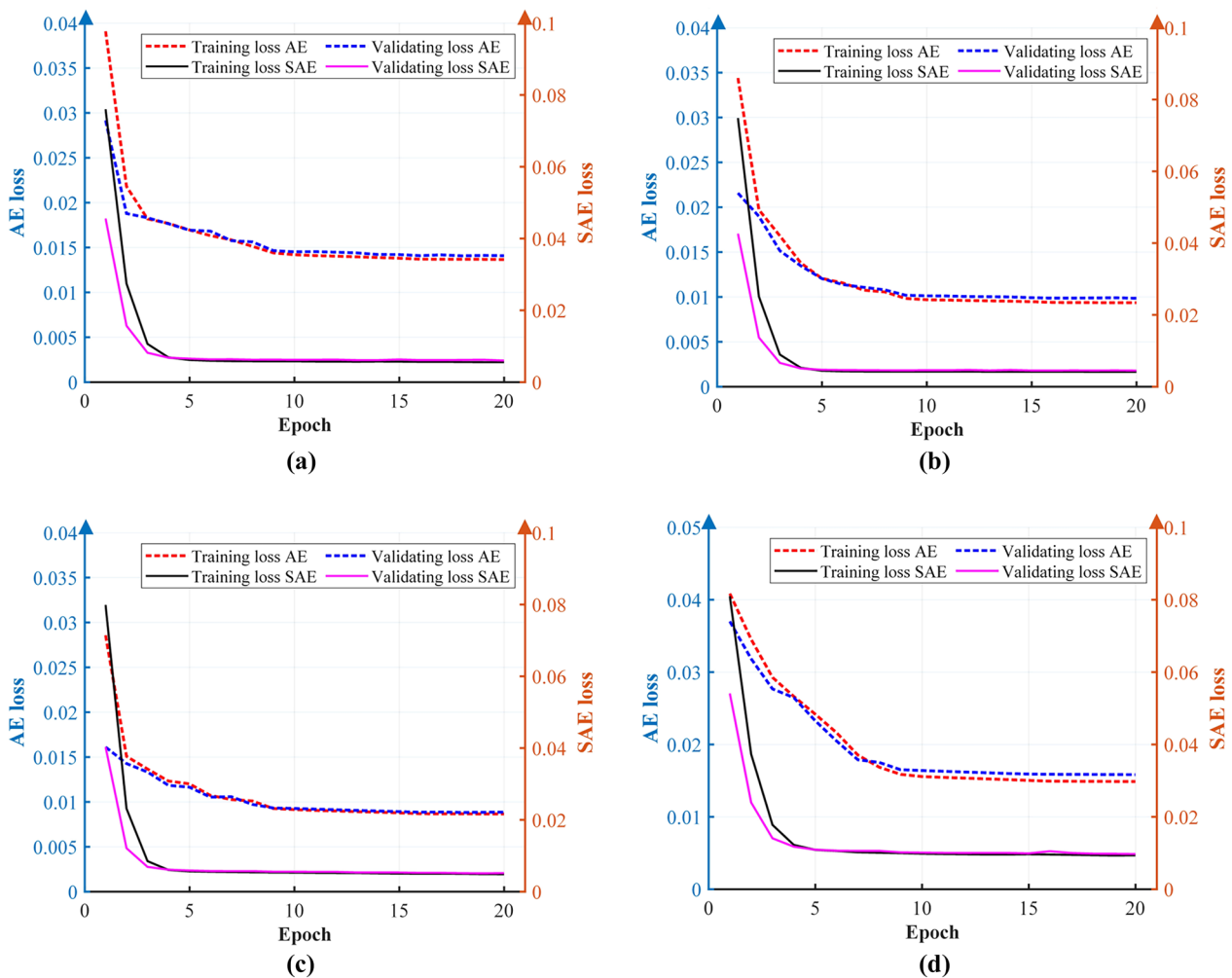


Figure 11 Training process for different faults using AE and SAE data argumentation methods (a. unbalanced, b. misaligned, c. looseness, d. normal)

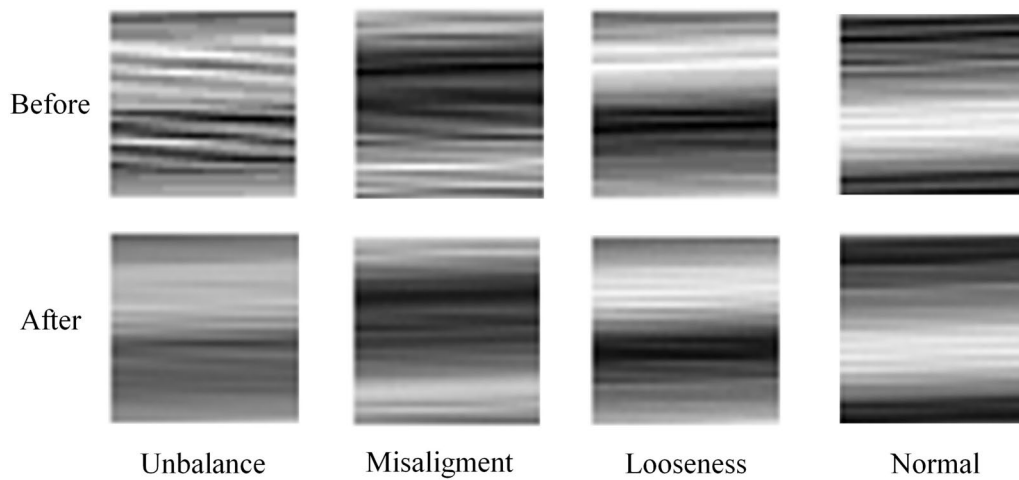


Figure 12 AE data augmentation results

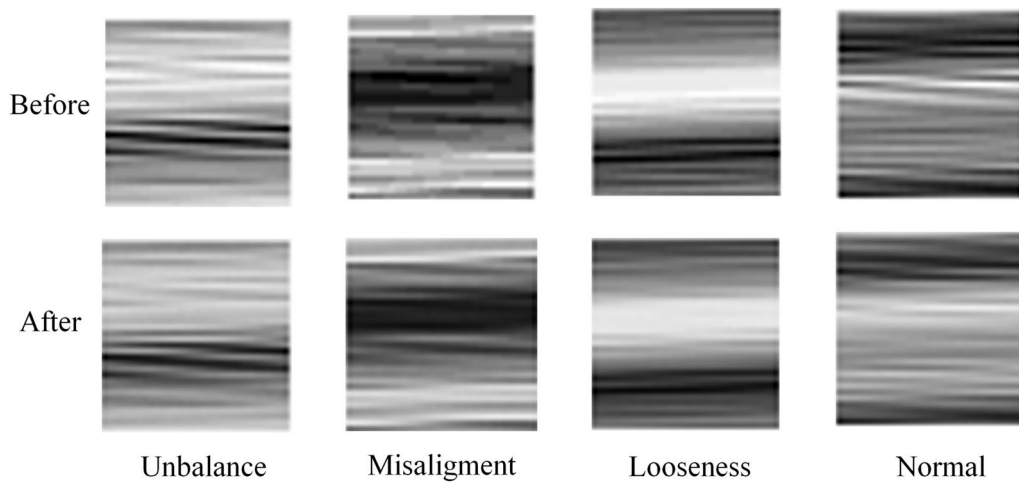


Figure 13 SAE data augmentation results

diagnostic accuracy of the SVM for different faults was 98.65% for the AE dataset and 98.80% for the SAE dataset when the penalty coefficient was set as the threshold (Figure 15).

It is necessary to fine-tune the ResNet-18 input channel to 1. The optimizer used was stochastic gradient descent (SGD) with an initial learning rate of 0.01. To avoid the gradient dispersion phenomenon, the momentum parameter was set to 0.9. To reduce the current gradient at each gradient update, a weight decay factor was

introduced to prevent the model from overfitting, which was set to $1e-3$ in this experiment, and the batch size was 128. The obtained loss drop and accuracy increase curves during the training process are shown in Figure 16. In the model training process, the training curve of the AE dataset had jitters; in contrast, the SAE training curve was smooth. The test results obtained using both datasets are shown in Figure 17. ResNet demonstrated an average diagnostic accuracy of 99.82% for different faults on the

Table 3 Model training parameters

Parameter name	AE	SAE
Batch size	32	
Optimizer	Adam	
initial learning rate	0.001	
Momentum	0.9	
Weight decay	1e-5	
Epoch	50	
Number of epochs per layer training	\	50
Batch size per training layer	\	32
Learning rate per training layer	\	0.001
Optimizer per training layer	\	Adam

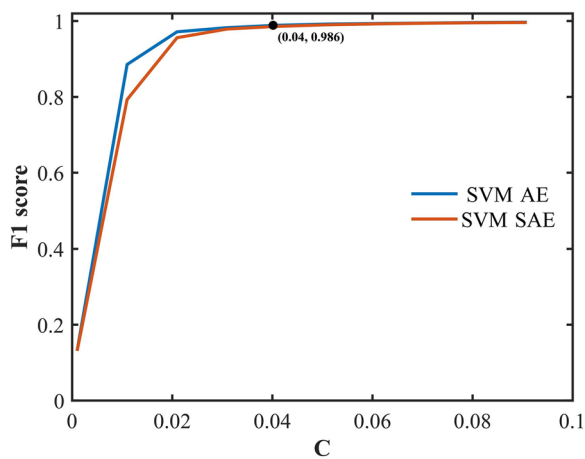


Figure 14 SVM performance of different data enhancement methods with different penalty factor C

AE datasets and 99.90% on the SAE datasets, which was 0.1% higher.

4.3.2 The Validation of Ensembled ResNet-SVM Fault Diagnosis

The ResNet part adopts the previously trained part of the ensemble ResNet-SVM model. After extracting features from the ResNet part, the SVM part must reconsider the choice of the penalty coefficient. The selection process for the SVM parameters, as previously described, is shown

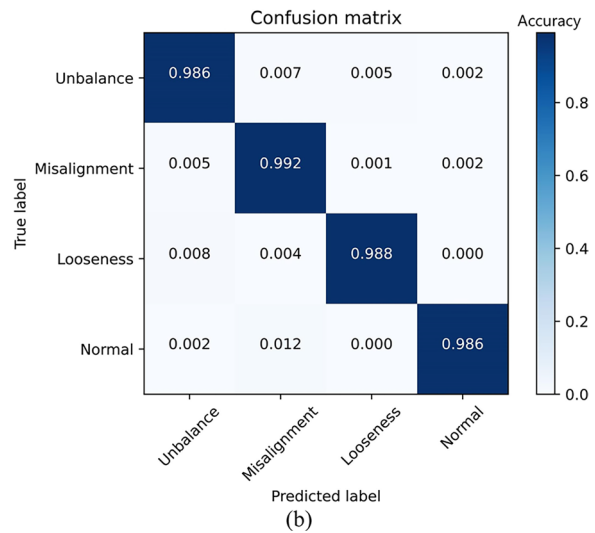
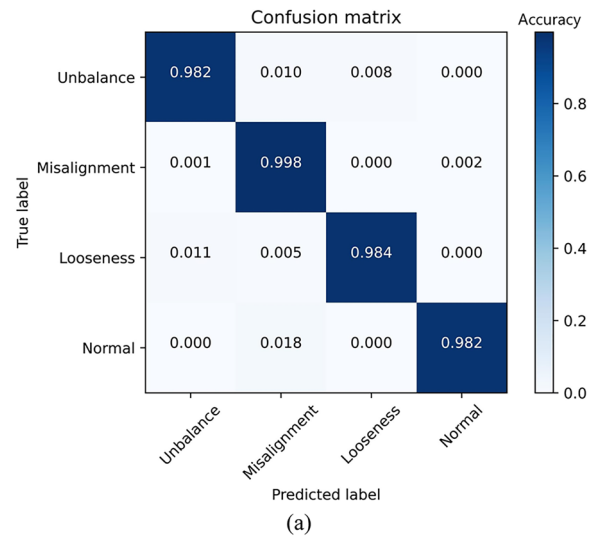


Figure 15 Confusion matrix of SVM in different data sets (a. AE, b. SAE)

in Figure 18. As can be seen, when the penalty coefficient of the SAE dataset is small, the F1 score reaches 0.999.

Although there is a slight increase in the later period, the F1 score of the AE dataset still starts below 0.2, indicating that the features extracted from the SAE dataset are more full-scale after ResNet feature extraction. After reaching the stabilization point, the penalty coefficients

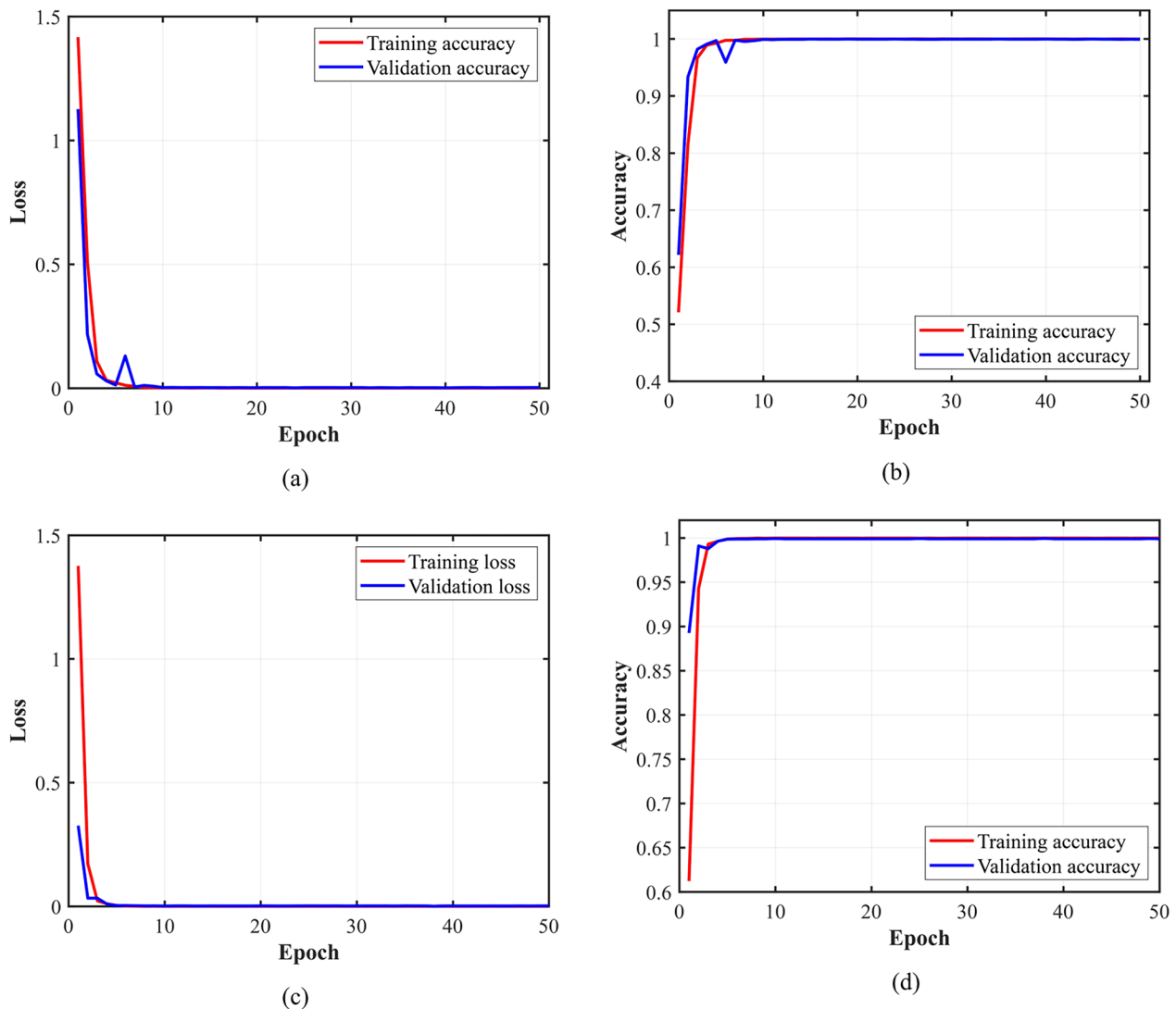


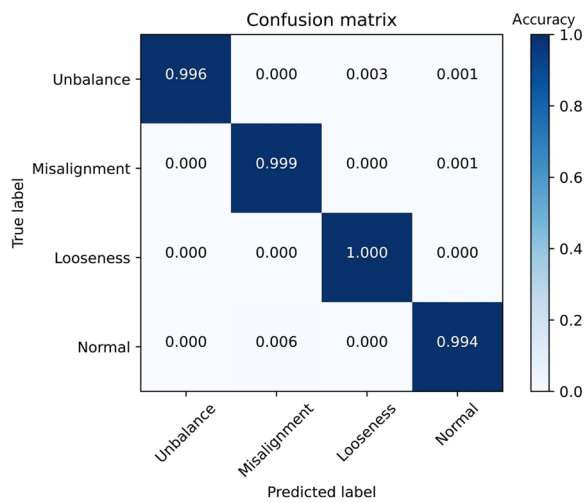
Figure 16 ResNet loss and accuracy plots using different datasets (**a.** Loss curve for AE dataset, **b.** Accuracy curve for AE dataset, **c.** Loss curve for SAE dataset, **d.** Accuracy curve for SAE dataset)

were adopted as 0.01 for the AE dataset and 0.03 for the SAE dataset. The tested confusion matrix of the ResNet-SVM model after thresholding the corresponding datasets separately is shown in Figure 19. The diagnostic accuracy of the ensemble model was 100% for different faults on the SAE dataset, whereas that of the AE dataset was 99.8%.

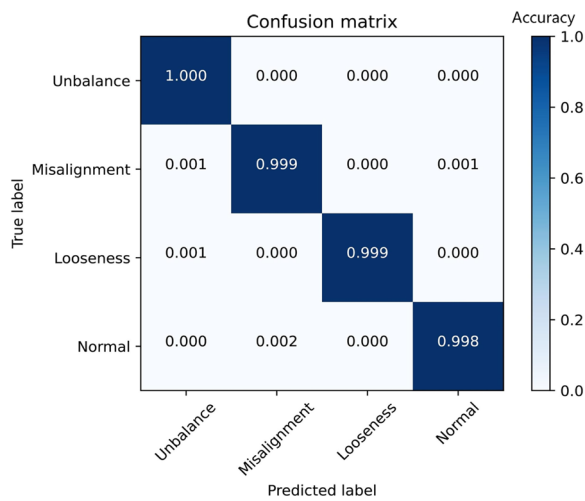
4.4 Experimental Conclusions

Compared with the cross-validation process of the SVM penalty coefficient, when the ensemble model adopts the SAE dataset, the F1 score is over 99.9%, and the remaining method starts at approximately 0.1. SAE and AE were used as the data-augmented datasets in

the experiments, and three models, SVM, ResNet, and ResNet-SVM, were used for the experimental comparison. The before-and-after images obtained from the experiments show that both the methods are effective. However, the model training process was smoother and the dataset had less noise when using the SAE dataset. To validate the diagnostic effectiveness of the three models adequately, the average results after five tests using the two datasets are shown in Figure 20. In terms of the overall diagnostic results, the diagnostic accuracy of the SVM, ResNet, and ensemble ResNet-SVM models increased sequentially. The diagnostic accuracy was 0.1%–0.01% higher using the SAE dataset for the same diagnostic



(a)



(b)

Figure 17 Confusion matrix for ResNet using different test sets (a. AE, b. SAE)

method, particularly using the method based on SAE and ResNet-SVM, with an accuracy of up to 99.9%.

5 Conclusions

For practical purposes, it is important to diagnose co-frequency faults with high accuracy. In this study, we propose an ensemble ResNet-SVM model and regard ResNet as a feature extractor, while SVM performs further feature extraction and final classification. Moreover, to improve the generalization ability of the model, AE and SAE were used as fault data generators in the experiments and the two data augmentation methods were

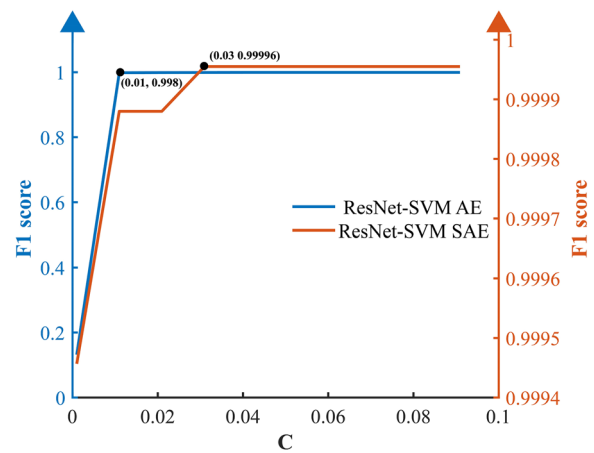
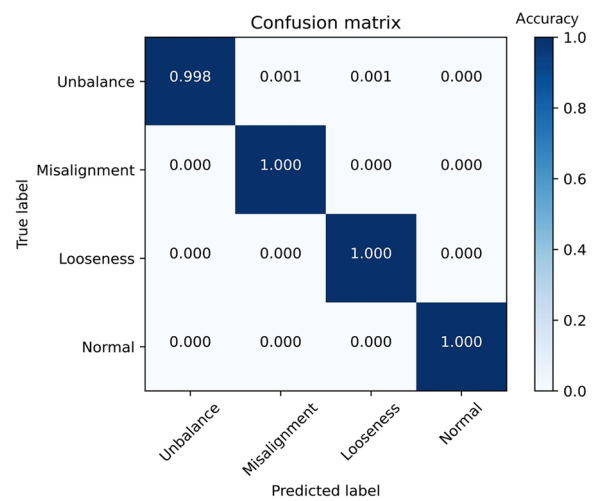
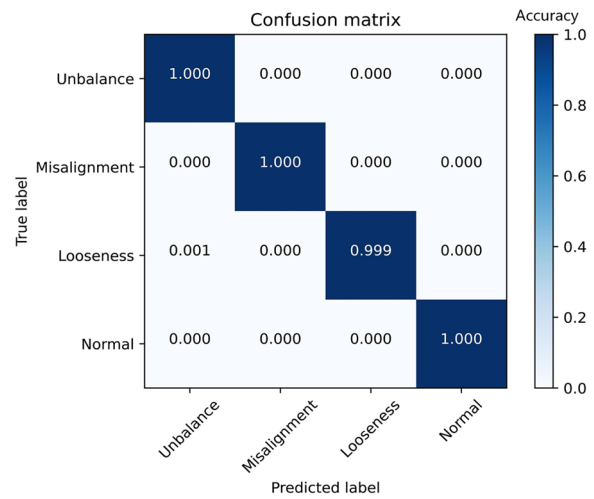


Figure 18 Penalty coefficient selection in ensemble model SVM part



(a)



(b)

Figure 19 Confusion matrix for ensemble models in different test sets (a. AE, b. SAE)

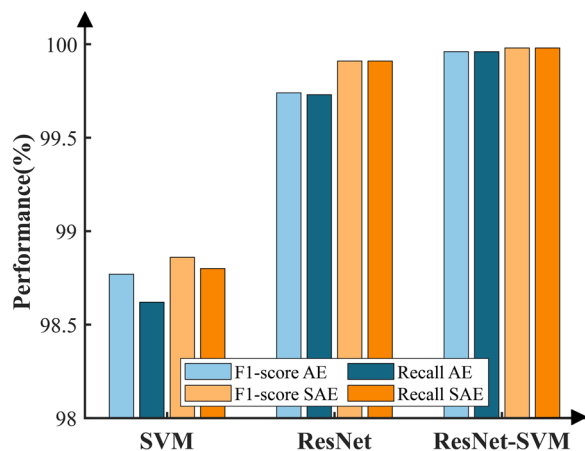


Figure 20 Performance of different fault diagnosis algorithms using different data augmentation methods

compared in the experiment. The following conclusions can be drawn from the experimental results.

- (1) The analysis of the experimental vibration data reveals that the time- and frequency-domain parameters of the co-frequency faults contain each other. If the traditional characteristic parameters are used to identify faults, the feasibility is low; therefore, an intelligent algorithm is necessary.
- (2) Comparing the performance of AE and SAE with three types of fault diagnosis methods, these two data augmentation methods are effective. However, with SAE data augmentation, the loss of intelligent diagnosis methods falls more smoothly and with less jitter, and the diagnostic accuracy obtained for the same diagnosis method is 0.1%–0.01% higher on an average.
- (3) A high-precision co-frequency fault diagnosis can be achieved using the SAE and ensemble ResNet-SVM fault diagnosis model. The average diagnostic accuracy can reach 99.9%.

Although the method proposed in this study combining the advantages of ResNet and SVM parts obtains promising results, the loss of the two parts in ensemble model can't be completely fused owing to the different loss functions. Further research will be needed to backpropagate the fused loss.

Acknowledgements

Not applicable.

Author Contributions

XZ: conceptualization, methodology, investigation, writing original draft; XP: writing—review & editing, supervision, funding acquisition, in charge of the whole trial; HZ and HZ: Methodology, Validation. All authors read and approved the final manuscript.

Funding

Supported by National Natural Science Foundation of China (Grant No. 51875031), Beijing Municipal Natural Science Foundation (Grant No. 3212010).

Availability of Data and Materials

The datasets supporting the conclusions of this article are included within the article.

Declarations

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Received: 14 March 2023 Revised: 25 February 2024 Accepted: 22 May 2024

Published online: 15 July 2024

References

- [1] X Zhu, D Hou, P Zhou, et al. Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images. *Measurement*, 2019, 138: 526-535.
- [2] S Lu, R Yan, Y Liu, et al. Tacholeless speed estimation in order tracking: A review with application to rotating machine fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 2019, 68(7): 2315-2332.
- [3] R M Souza, E G Nascimento, U A Miranda, et al. Deep learning for diagnosis and classification of faults in industrial rotating machinery. *Computers & Industrial Engineering*, 2021, 153: 107060.
- [4] Y Lei, B Yang, X Jiang, et al. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mechanical Systems and Signal Processing*, 2020, 138: 106587.
- [5] Z Shang, W Li, M Gao, et al. An intelligent fault diagnosis method of multi-scale deep feature fusion based on information entropy. *Chinese Journal of Mechanical Engineering*, 2021, 34: 58.
- [6] H Pan, W Sun, Q Sun, et al. Deep learning based data fusion for sensor fault diagnosis and tolerance in autonomous vehicles. *Chinese journal of Mechanical Engineering*, 2021, 34: 72.
- [7] Z Shang, Z Zhao, R Yan, Denoising fault-aware wavelet network: A signal processing informed neural network for fault diagnosis. *Chinese Journal of Mechanical Engineering*, 2023, 36: 9.
- [8] M Gohari, A M Eydi. Modelling of shaft unbalance: Modelling a multi discs rotor using k-nearest neighbor and decision tree algorithms. *Measurement*, 2020, 151: 107253.
- [9] N Wang, D Jiang. Vibration response characteristics of a dual-rotor with unbalance-misalignment coupling faults: Theoretical analysis and experimental study. *Mechanism and Machine Theory*, 2018, 125: 207-219.
- [10] A G Nath, A Sharma, S S Udmale, et al. An early classification approach for improving structural rotor fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 2020, 70: 1-13.
- [11] S Ma, F Chu. Ensemble deep learning-based fault diagnosis of rotor bearing systems. *Computers in Industry*, 2019, 105: 143-152.
- [12] Z Yang, D Gjorgjevikij, J Long, et al. Sparse autoencoder-based multi-head deep neural networks for machinery fault diagnostics with detection of novelties. *Chinese Journal of Mechanical Engineering*, 2021, 34: 54.
- [13] S Guo, T Yang, W Gao, et al. A novel fault diagnosis method for rotating machinery based on a convolutional neural network. *Sensors*, 2018, 18(5): 1429.
- [14] R Zhao, R Yan, Z Chen, et al. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 2019, 115: 213-237.

- [15] W Hong, W Cai, S Wang, et al. Mechanical wear debris feature, detection, and diagnosis: A review. *Chinese Journal of Aeronautics*, 2018, 31(5): 867-882.
- [16] M Cerrada, R V Sánchez, C Li, et al. A review on data-driven fault severity assessment in rolling bearings. *Mechanical Systems and Signal Processing*, 2018, 99: 169-196.
- [17] X Yan, M Jia. A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing. *Neurocomputing*, 2018, 313: 47-64.
- [18] B Zhao, X Zhang, H Li, et al. Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions. *Knowledge-Based Systems*, 2020, 199: 105971.
- [19] L Wen, X Li, L Gao. A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Computing and Applications*, 2020, 32: 6111-6124.
- [20] Z He, Y Yang, D Liang. A multi-concurrent fault diagnosis scheme for the parallel shaft gearbox based on resnet neural network and image recognition approach. *2021 China Automation Congress (CAC)*, Yunnan, China, August 12-13, 2022: 6123-6127.
- [21] K He, X Zhang, S Ren, et al. Deep residual learning for image recognition. *CVPR*, Las Vegas, USA, June 26-July 1, 2016: 770-778.
- [22] R Bai, Q Xu, Z Meng, et al. Rolling bearing fault diagnosis based on multi-channel convolution neural network and multi-scale clipping fusion data augmentation. *Measurement*, 2021, 184: 109885.
- [23] T Zhang, J Chen, J Xie, et al. SASLN: Signals augmented self-taught learning networks for mechanical fault diagnosis under small sample condition. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70(99): 1-11.
- [24] X Li, W Zhang, Q Ding, et al. Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. *Journal of Intelligent Manufacturing*, 2020, 31: 433-452.
- [25] W Gong, H Chen, Z Zhang, et al. A novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved CNN-SVM and multichannel data fusion. *Sensors*, 2019, 19(7): 1693.
- [26] S Shao, P Wang, R Yan. Generative adversarial networks for data augmentation in machine fault diagnosis. *Computers in Industry*, 2019, 106: 85-93.
- [27] M Zareapoor, P Shamsolmoali, J Yang. Oversampling adversarial network for class-imbalanced fault diagnosis. *Mechanical Systems and Signal Processing*, 2021, 149: 107175.
- [28] S Liu, H Jiang, Z Wu, et al. Rolling bearing fault diagnosis using variational autoencoding generative adversarial networks with deep regret analysis. *Measurement*, 2021, 168: 108371.
- [29] Z Tang, L Bo, X Liu, et al. An autoencoder with adaptive transfer learning for intelligent fault diagnosis of rotating machinery. *Measurement Science and Technology*, 2021, 32(5): 055110.
- [30] W A Smith, R B Randall. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*, 2015, 64: 100-131.
- [31] D Bank, N Koenigstein, R Giryes. Autoencoders. *arXiv preprint, arXiv:2003.05991*, 2020.

Haofu Zhou, born in 1993, currently serves as the Business Manager at the Engineering Management Department of China Three Gorges New Energy (Group) Co., Ltd. He Received his masters degree from Beijing University of Technology, China, in 2023. His research interests include vibration monitoring and fault diagnosis.

Xiancheng Zhang, born in 1997, is currently a graduate student at Beijing University of Chemical Technology, China. He received his bachelors degree from Beijing University of Chemical Technology, China, in 2020. His research interests include fault intelligent diagnosis.

Xin Pan, born in 1987, is currently an associate professor and master supervisor at Beijing Key Laboratory of Health Monitoring and Self-recovery for High-end Mechanical Equipment, Beijing University of Chemical Technology, China. He received his PhD degree from Beijing University of Chemical Technology, China, in 2015. His research interests include vibration monitoring, fault diagnosis and active control.

Hao Zeng, born in 1999, is currently a graduate student at Beijing University of Chemical Technology, China. He received his bachelors degree from Beijing University of Chemical Technology, China, in 2021. His research interests include artificial intelligence neural network.