Journal of Dynamics, Monitoring and Diagnostics, 2024, 3, 1-10
https://doi.org/10.37965/jdmd.2024.508

ISTP
RESEARCH ARTICLE

Fault Diagnosis of Linear Guide Rails Based on SSTG Combined with CA-DenseNet

Yanping Wu,[1] Juncai Song,[2] Xianhong Wu,[1] Xiaoxian Wang,[3] and Siliang Lu[1]
[1]School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China
[2]School of Internet, Anhui University, Hefei 230039, China
[3]School of Electronic Information Engineering, Anhui University, Hefei 230601, China

(Received 01 January 2024; Revised 30 January 2024; Accepted 19 February 2024; Published online 19 February 2024)

Abstract: Monitoring the status of linear guide rails is essential because they are important components in linear motion mechanical production. Thus, this paper proposes a new method of conducting the fault diagnosis of linear guide rails. First, synchrosqueezing transform (SST) combined with Gaussian high-pass filter, termed as SSTG, is proposed to process vibration signals of linear guide rails and obtain time-frequency images, thus helping realize fault feature visual enhancement. Next, the coordinate attention (CA) mechanism is introduced to promote the DenseNet model and obtain the CA-DenseNet deep learning framework, thus realizing accurate fault classification. Comparison experiments with other methods reveal that the proposed method has a high classification accuracy of up to 95.0%. The experimental results further demonstrate the effectiveness and robustness of the proposed method for the fault diagnosis of linear guide rails.

Keywords: CA-DenseNet; fault diagnosis; linear guide rails; SSTG

I. INTRODUCTION

Linear guide rails are important components of mechanical production and are typically used in conjunction with ball bearings in sliding blocks to produce reciprocating linear motion. Due to their excellent load-bearing rigidity, they are widely applied in various fields, such as transportation, industrial robots, and automated production lines [1–3]. Linear guide rails typically operate in open working environments. Thus, the continuous linear motion and complex working conditions easily lead to the wear of the guide rail and ball bearing inside the slider [4], leading to the faults of the entire linear motion unit and eventually causing significant losses. Therefore, monitoring the status of linear guide rails is a highly necessary procedure.

Current studies on linear guide rail status monitoring have mainly relied on data-driven methods [5,6]. However, such methods require massive amounts of data to study the features of fault signals [7]. Generally, these methods can extract fault signal features in the frequency domain or time-frequency domain [8] and then use intelligent classification models to distinguish different types of faults. However, in using these methods, certain challenges remain regarding fault feature extraction and intelligent classification recognition.

In terms of fault feature extraction, various approaches have been developed. For example, Qin *et al.* used wavelet packets to perform the decomposition of rail bogie fault signals and extract model features, successfully diagnosing the rail bogie fault [9]. You *et al.* used the method based on ensemble empirical mode decomposition (EEMD) to denoise the vibration signals of the escalator guide rail, the main engine, and the main drive shaft, successfully extracting the signal features. Furthermore, the improved envelope spectrum analysis method was used to extract the characteristic frequencies and amplitudes of the fault signal, thus forming the feature vector and successfully diagnosing the fault location of the escalator [10]. Chommuangpuck *et al.* analyzed spectrum, peak factor, and variance of signals from linear guide rail, wear ball, and missing ball signals, successfully extracting frequency and time-domain features, establishing different fault indicators corresponding to different types of faults, and distinguishing ball wear and missing faults [11]. Borowiecki *et al.* proposed a second-order sinusoidal model to analyze railway track vibration signals, effectively extracting its amplitude and phase features [12]. Ni *et al.* proposed the partitioned edge features (PEF) algorithm to effectively extract the rail surface features and establish a uniform background in the visual inspection of the rail, effectively extracting the rail surface damage information [13]. Zhang *et al.* designed a feature fusion algorithm. The algorithm had a three-branch structure, which effectively overcomes the missing detection of the guide surface [14].

In terms of intelligent classification recognition, Kim *et al.* used a one-dimensional (1D) convolutional neural network (CNN) to classify and visualize vibration signals, successfully differentiating between healthy guide rails and delamination faults [1]. Ye *et al.* introduced a semantic segmentation network, including a fully convolutional segmentation module and a symmetric mapping module, which can detect the surface faults of guide rails [15]. Niu *et al.* embedded the adaptive pyramid graph (APG) module into a specific module to obtain feature correlations, successfully introducing the residual module to enhance the feature expression and detect surface defect faults [16]. Chen designed the BoTNet 50 network to extract the fault image features of the guide rail surface and used CUFuse model to detect the constructed fault dataset, finally realizing the fault diagnosis of guide rail surfaces [17]. Liu *et al.* proposed a multipopulation state optimization algorithm (MPVHGA) with high iteration and no fixed iteration

Corresponding authors: Juncai Song (email: songjuncai@ahu.edu.cn); Siliang Lu (e-mail: silianglu@ahu.edu.cn).

number limitations, successfully solving the problem of rail crack faults [18].

However, the above methods are only suitable for the state monitoring of linear guide rails with simple fault. In fact, it is difficult to use the above methods to achieve the fault grade classification of linear guide rails and ball bearings. In addition, current research on various types of faults of ball bearings is relatively scarce, along with studies on different types and degrees of faults of the guide and the ball bearing. To solve the problem, a method combined with synchrosqueezing transform (SST) and Gaussian high-pass filter, termed as SSTG, is used to process and vibration signal. SSTG has high time-frequency aggregation, which can effectively capture the time-frequency information with no obvious rules, such as ball wear and loss, and visualize it as an image. Different lines on the image represent different time-frequency information and the details such as lines on the image are enhanced to extract vibration information features. Then a coordinate attention (CA) mechanism combined with the DenseNet model is proposed to detect the state of linear guide rails.

The main work in this paper includes the following: proposing the SSTG algorithm, wherein 1D vibration signals are transformed into 2D time-frequency images, and the details of these images are enhanced. The DenseNet model is used to classify the generated images by SSTG, and the CA mechanism is introduced to improve the attention to the space and position information of feature matrices during training. Finally, the test accuracy reached 95.0%.

The innovations and novelties of this paper are as follows. First, the SSTG algorithm is proposed, in which 1D vibration signals are transformed into 2D time-frequency images by SST, after which Gaussian high-pass filtering is carried out on the images to enhance the lines and spots of the images and make different features more prominent. Second, the CA mechanism is introduced into the DenseNet model. The dense connections of the model enable it to realize feature reuse in the training process, thus retaining the dimension information of the feature matrices. This work introduces the CA mechanism to DenseNet, paying attention to the spatial position information of the feature matrices. By combining DenseNet and the CA mechanism, the information loss in the training process is reduced, and the accuracy of the model is improved.

The rest of this work is organized into sections. Section II describes the composition of the linear guide model and the researched faults. Section III introduces the basic theory of SSTG and the CA-DenseNet model. Section IV establishes the experimental platform to collect vibration signals in various fault states. Then, comparison and robustness experiments are designed to verify the superiority of proposed method. Finally, Section V concludes this work.

II. LINEAR GUIDE RAIL MODULE

Linear guide rails are applied in linear servo motion scenarios and are typically used in conjunction with sliding blocks. A large number of balls in the sliding block are closely arranged and cooperate with the guide rail. Thus, rolling friction occurs between the ball and the guide rail during the operation of the sliding block. All of these can ensure that the moving operation process is smooth and efficient. The linear guide rail investigated in this work is shown in Fig. 1.

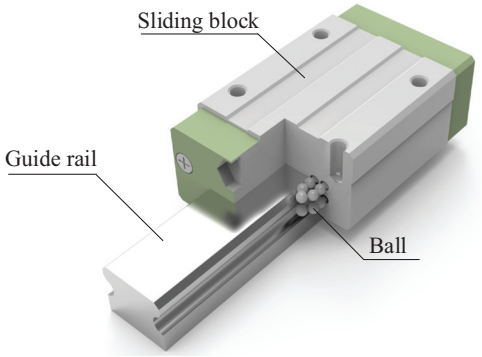

Fig. 1. Components of the linear guide module.

Table I. Parameters of the linear guide rail and sliding block

Parameter	Guide rail	Sliding block
Type	g×150	hg20
Length (mm)	1000	77.5
Width (mm)	20	44
Height (mm)	17.5	30

The diameter of the ball bearing installed in the sliding block is 3.8 mm, and the number of balls in each sliding block is 128. The detailed main parameters about the guide rail and the sliding block are shown in Table I.

Linear guide rails often work in open and exposed environments. Thus, they are greatly affected by the external environment. When the environmental noise and vibration surrounding the guide rails are more pronounced, the guide rails and the ball bear long-term load and reciprocating motion, making them vulnerable to wear and damage, and even the failure of the missing ball. In addition, improper installation, long service life, external impact on the guide module, and other factors can easily cause the ball to fall off in the sliding block, eventually resulting in the failure of the entire linear unit. In view of the above situations, this work not only studies the basic faults such as guide wear, ball wear, and ball missing but also researches how to distinguish the different fault severity degrees of these basic faults. The main fault types of linear guide rails are shown in Table II.

Table II. Fault types and labels

Types	Fault degrees	Labels
Ball wear	One ball	A0
	Two balls	A1
	Three balls	A2
Ball missing	One ball	A3
	Two balls	A4
	Three balls	A5
Normal	\	A6
Mixed fault	One ball wear with one ball missing	A7
Guide rail wear	100 mm	A8
	200 mm	A9

III. PROPOSED METHOD

The proposed method consists of multiple signal processing, image processing, and deep learning methods, which will be introduced in detail in the following sections.

A. SST METHOD

SST can represent signals in both the time and frequency domains [19], which in turn can enhance time-frequency concentration. Thus far, SST has been successfully applied in seismic signal detection [20], weather analysis [12], oil exploration [21], and fault diagnosis of gearbox [22] occasions. By specially rearranging the time-frequency coefficients from 2D wavelet transformations, it can enhance the time-frequency representation. All of these ensure the more concentrated distribution of time-frequency energy [23]. As such, the time-frequency information of images generated by SST becomes clearer [24]. The detailed theoretical analysis is as follows:

$$WTZ(a,b) = \frac{1}{\sqrt{a}} \int_R Z(t)\varphi(\frac{t-b}{a})dt \tag{1}$$

where $\varphi(t)$ is the mother wavelet, a is the scale factor representing the scaling transformation of the $\varphi(t)$, b is the translation factor representing the time translation of the $\varphi(t)$, and $WTZ(a, b)$ is the continuous wavelet transform result of vibration signal $Z(t)$.

The frequency domain expression of $WTZ(a, b)$ is expressed as follows:

$$WTZ_f(a,b) = \frac{1}{2\pi\sqrt{a}} \int \hat{Z}(\omega)\hat{\varphi}(\omega)e^{ib\omega}d\omega \tag{2}$$

where ω represents frequency, and $\hat{Z}(\omega)$ and $\hat{\varphi}(\omega)$ represent the Fourier transforms of $Z(t)$ and $\varphi(t)$, respectively.

The instantaneous frequency $\omega_f(a,b)$ of the signal Z(t) is calculated as:

$$\omega_f(a,b) = -i[\frac{\partial}{\partial b} WTZ_f(a,b)]/WTZ_f(a,b) \tag{3}$$

Finally, at the estimated instantaneous frequency $\omega_f(a,b)$, the coefficients obtained from the continuous wavelet transformation are rearranged, with the rearrangement formula given by:

$$I = \int WTZ_f(a,b)a^{-\frac{2}{5}}\delta[\omega(a,b) - \omega]d\omega \tag{4}$$

where I is the final result of the SST and $\delta[\cdot]$ is the Dirac function.

B. GAUSSIAN HIGH-PASS FILTER

After processing the original vibration signals of linear guide rails via SST, they are converted into 2D time-frequency images. The high time-frequency resolutions of the SST result in a concentrated display of time-frequency information in these images, which are characterized by differently distributed lines and spots [25]. Gaussian high-pass filtering is an image processing method that effectively enhances the details of the image, improves the quality and reliability of the image, and is widely used in medical image processing [26]. This method is also used in the fault diagnosis of rolling element bearing fields [27]. The main feature of Gaussian high-pass filtering is its use of Gaussian kernels to carry out the convolution operation on the image, and such kernels are generated by the Gaussian function. Equation (5) is the 2D space Gaussian kernel calculation function [28]:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{5}$$

where x and y denote the plane coordinates, G (x, y) is the Gaussian distribution function, and σ is the Gaussian filter parameter, which can control the passing high-frequency components. Using Equation (5), the required convolution kernel can be obtained, and then the input image is convolved with the convolution kernel. Finally, the Gaussian high-pass filtering image is obtained using Equation (6):

$$H(x,y) = G(x,y) * I(x,y) \tag{6}$$

where I (x, y) is the input image and H (x, y) is the output image after Gaussian high-pass filtering. After the enhancement of images via Gaussian high-pass filtering, the SSTG method can be executed.

C. SSTG

SST algorithm converts the time-domain signal to the time-frequency domain and visualizes the time-frequency image. Different lines and colors on the image represent different time-frequency information. In this paper, the Gaussian high-pass filtering algorithm is innovatively introduced to enhance the detailed features of time-frequency images such as lines and colors. The specific processes are as follows:

According to Equations (1)–(4), the vibration signal $Z(t)$ is processed by SST to obtain the time-frequency image I.

The Gaussian kernel is calculated according to Equation (5).

The Gaussian kernel and the input image are convolved according to Equation (6) to obtain the SSTG output image H.

D. DenseNet MODEL

The DenseNet model employs a dense connection mechanism, wherein each layer is formed by concatenating the outputs of all preceding layers in the channel dimension [29]. All these can achieve the effects of feature reuse, suppress gradient vanishing, enhance feature propagation, and reduce model parameters. With such advantages, this method has been successfully applied in semantic segmentation [30,31], computer vision [32], and fault diagnosis [33–35]. This model mainly consists of dense blocks.

The dense block structure is contained by multiple bottlenecks, and a dense connection is used between multiple bottlenecks. This consists of several convolutional layers, batch normalization (BN) layers, and ReLU activation function. Before the convolution operation, normalization is performed using the BN layer, and nonlinearization is achieved by using the ReLU activation function. The following passages describe the layer composition of each bottleneck and the dense connections among multiple bottlenecks.

1) BN LAYER. BN layer normalization can accelerate the convergence process and enhance the generalization ability of DenseNet. Its expression is given by:

$$x'l = \gamma \times \frac{xl - \mu}{\sqrt{h^2 + \varepsilon}} + \beta \tag{7}$$

where x_l is the input to the BN layer, x'_l is the output of the BN layer, μ and h are the mean and variance of feature data in the same channel of the same batch samples, respectively, ε is a nonzero correction coefficient, and β and γ are the parameters learned from backpropagation during the model training process.

2) ReLU ACTIVATION FUNCTION. The ReLU activation function does not have a saturation region; it is sparse with simple calculations and fast convergence. It can effectively prevent the gradient dispersion problem. The theoretical analysis is as follows:

$$\mathrm{ReLU}(x'l) = \begin{cases} x'l, & x'l \geq 0 \\ 0, & otherwise \end{cases} \tag{8}$$

3) DENSE CONNECTIONS. Multiple bottleneck structures are densely connected, with the input of each layer representing the concatenation of all preceding layers inputs. Figure 2 illustrates the dense connection process between multiple bottlenecks, while Fig. 3 shows the internal structure of the bottleneck.

The calculation formula for the densely connected bottleneck structure is as follows:

$$xl = Fl([x0, x1, \ldots xl-1]) \tag{9}$$

where l represents the l^{th} layer of the model, $F_l(\cdot)$ represents the nonlinear function of each layer, which includes BN, ReLU, convolution, and pooling operations, and x_l represents the output features of the l^{th} layer.

DenseNet is composed of different dense blocks, which are connected by transition structures between adjacent blocks. The transition layer structure consists of the convolutional layer and average pooling layer and is used to connect adjacent dense blocks while reducing the size of the feature matrices and compressing the model simultaneously. Dense connections are used between each dense block, wherein the output feature matrices of each layer in the channel dimension represent the concatenation of all previous layers' outputs. This step enhances the flow of information between layers, thus achieving the purpose of feature reuse. The ratio of the number of bottlenecks in each dense block is 6:12:24:16.

However, DenseNet only performs feature matrix concatenation at the channel level, neglecting space and position information. Therefore, the CA mechanism is introduced to maintain interests in the space and position information of the feature matrix during the training process.

E. CA MECHANISM

CA attention mechanism enables the model to focus on both the channel information and spatial information of the feature matrices during the training process, often applied in tasks, such as object detection and instance segmentation [36]. For an input matrix, the CA module can decompose it into feature matrices along the width and height directions, as shown in Fig. 4.

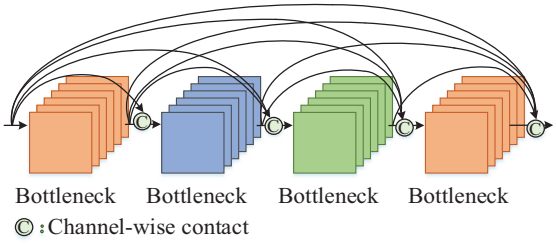

Fig. 2. Densely connected structure.

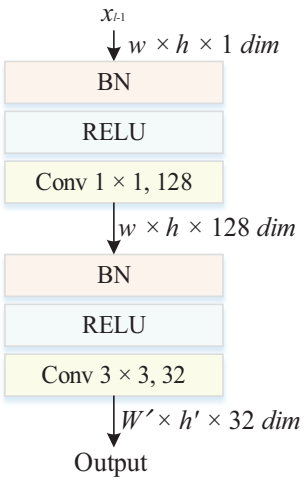

Fig. 3. Structure of the bottleneck.

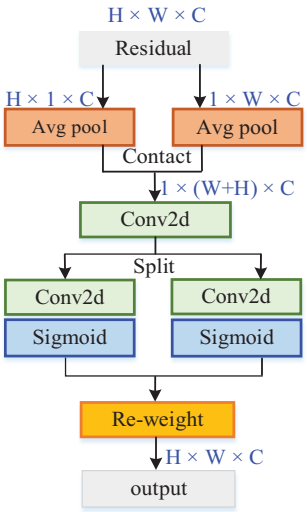

Fig. 4. CA attention mechanism.

The input feature matrix is subjected to average pooling operations along the height and width directions to obtain the H×1×C and 1×W×C feature matrixes, respectively. These matrixes are then transposed and concatenated, followed by convolution operations to extract features and reduce dimensions.

The convolved matrix is divided into matrixes along the width and height directions. This ultimately multiplies them to output an H×W×C feature matrix and incorporates space and position information.

F. FAULT DIAGNOSIS PROCESS BASED ON CA-DenseNet

In this work, the CA attention mechanism is added after each transition structures. An image of size 224×224×3 was used as input, and the feature matrices of multiple channels were obtained after dense block processing. The dense

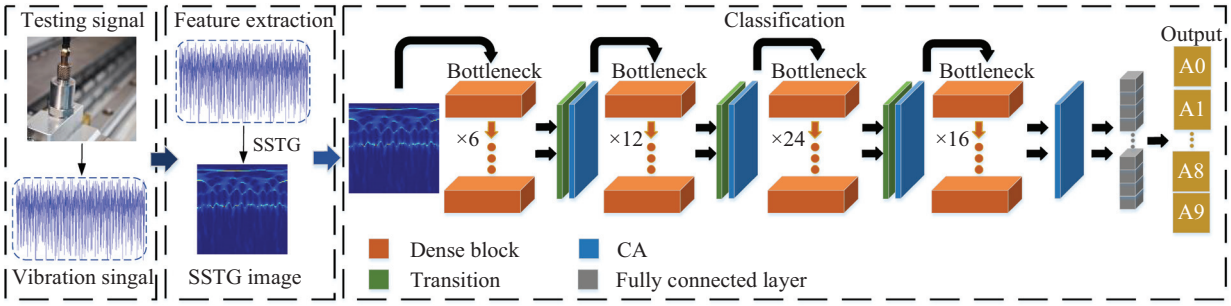

Fig. 5. General process of the fault diagnosis of linear guide rails.

connection structure enhanced the information interaction between the channels of the feature matrix. The transition structure maintains the characteristics of these matrices while reducing the size of the feature matrix and then inputs it into CA. CA performs pooling, contact, convolution, and other operations on these feature matrices along the height and width directions, respectively, and forms a new feature matrix containing the location information of different directions. Finally, good classification results are obtained. The fault diagnosis process based on CA-DenseNet is shown in Fig. 5.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL PLATFORM INTRODUCTION

The experiment platform is shown in Fig. 6.

As can be seen, there are four sliding blocks fixed under the platform. The rotating motor is connected to the ball screw structure through the coupler, and then the rotating motion of the motor is converted into the linear motion of the cargo platform. This experimental platform consists of the following parts: (1) rotating motor, (2) linear guide rail, (3) rotary motor driver, (4) pulse controller, (5) CA-YD-1182 acceleration sensor, (6) USB4431 NI data acquisition card, (7) 5 kg load, and (8) a computer. The detailed procedure is described below.

(1) Step 1-Faults presetting:
Several steps must be carried out to present different type of faults as ball wear faults, ball missing faults, guide rail wear faults, and mixed fault. (a) Ball wear faults: three balls are removed from the sliding blocks one by one to create wear faults, after which the worn balls are loaded back into the sliding block for experiment. (b) Ball loss faults: three balls are removed from the sliding block one by one, and the experiment is conducted in the case of one, two, and three missing balls. (c) Guide rail wear faults: the total length of the guide rail is 1000 mm. Wear damage is carried out in the middle part of the guide rail, and the wear length is set to 100 mm and 200 mm. Experiments were carried out under these two wear conditions separately. (d) Mixing fault: this work investigates the mixed fault when one ball was missing and the other ball was worn at the same time. A mixed fault of balls is created by removing one ball in the sliding block while wearing another ball. All of these presetting faults are as shown in Fig. 7.

(2) Step 2-Motor control:
Connect the controller to the driver, which is connected to the rotating motor. Then, control the speed and travel distance of the cargo platform using the pulse controller. Next, load the 5 kg load on the platform.

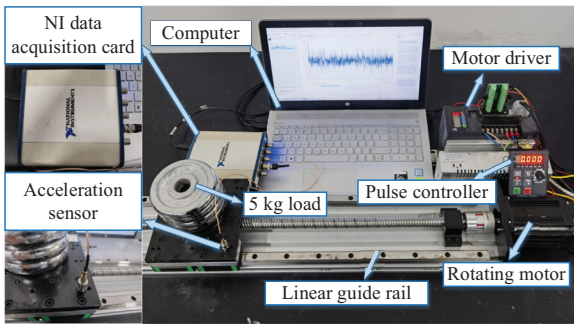

Fig.6. Experiment platform of linear guide rail.

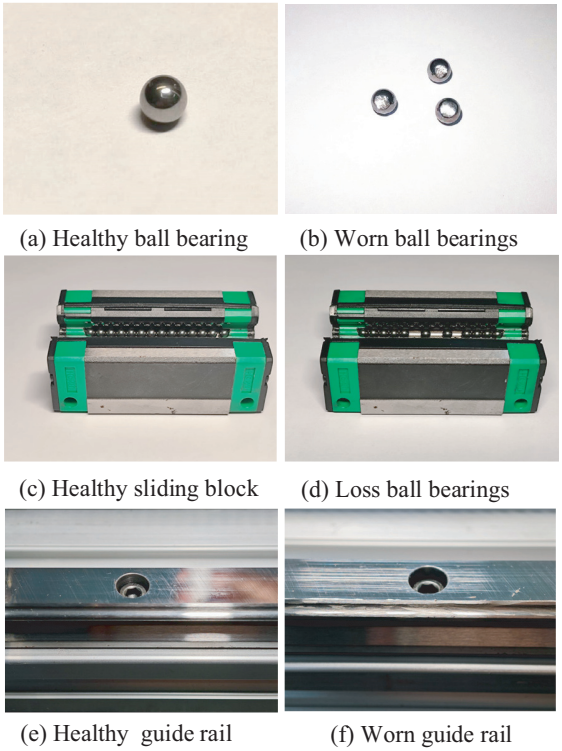

Fig. 7. Basic fault settings.

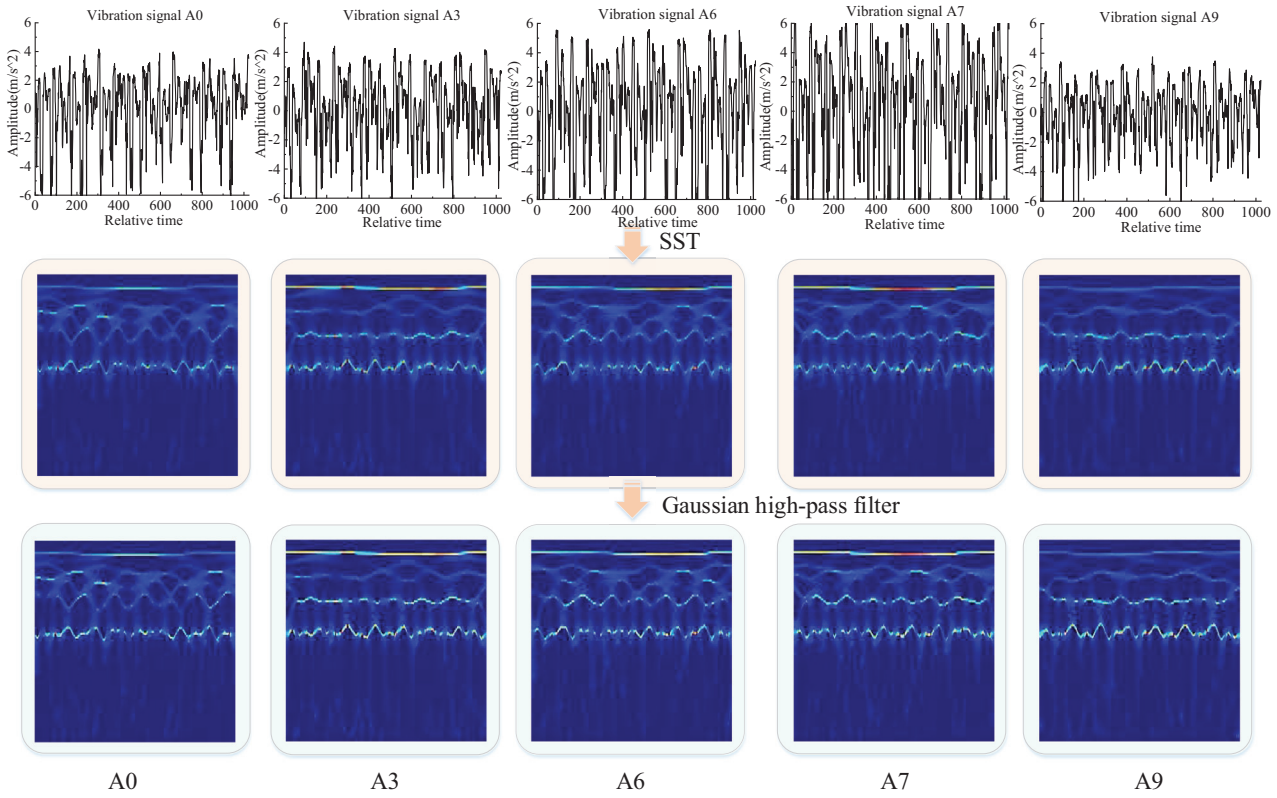

Fig. 8. Time-frequency images of SSTG.

(3) Step 3-Vibration signal collecting:
Fix the accelerometer to the load platform with bolts, ensuring that it is positioned perpendicular to the moving direction. Connect the accelerometer to the NI data acquisition card. This data acquisition card is linked to a computer to upload the collected vibration signals to the computer in real time. The sample frequency is set to 6000 Hz, the velocity is set to 600 mm/s, and the period of linear motion is set to 5 times of round trips. The pulse controller and data acquisition program are initiated, and the vibration signals are recorded into the computer.

B. DATASET CONSTRUCTION OF LINEAR GUIDE RAIL BASED ON SSTG

Once the original vibration signals are collected, SST is applied to conduct signal processing and 1D signals are converted into 2D images. Every set of 1024 points in the vibration signals is converted into a 224×224×3 image. These time-frequency images are then subjected to Gaussian high-pass filtering, which attenuates or filters out the low-frequency components. The entire conversion process is shown in Fig. 8.

Table III. Dataset introduction

Dataset	Length	Size	Numbers
Train	1024	224×224×3	180
Validation	1024	224×224×3	60
Test	1024	224×224×3	60

Upon completion of the SSTG transformation, each type of fault yields 300 images for a total of 3000 images for 10 types of faults. Around 60% of the images for each type of fault are used to build the train set, 20% are used for the validation set, and 20% are used for test set. The dataset introduction is shown in Table III.

C. FAULT CLASSIFICATION RESULTS BASED ON CA-DenseNet

This experiment used NVIDIA RTX 3060 GPU to train this model, after which the deep learning environment is established using Pytorch 1.12.1.

This model's learning rate was set to 0.005, and the optimizer used was stochastic gradient descent (SGD). The training process was conducted over 100 rounds, and the model's training accuracy increased with each round, ultimately converging at 96.4%, as shown in Fig. 9. A confusion matrix was created using the test results to further demonstrate the model performance. The results are shown in Fig. 10.

D. COMPARISON WITH OTHER METHODS

In order to verify the necessity of signal processing, raw vibration signals are directly fed into each model, and to validate the superiority of the proposed method, in this paper, the images generated by SSTG and SST were input into the CA-DenseNet, DenseNet, ResNet, ShuffleNet, and ConvNext models for training. Figure 11(a)–(e) shows the training classification accuracy of SSTG, SST images, and 1D signals on different classification models. The mean and standard deviation of the five test results are shown in Table IV.

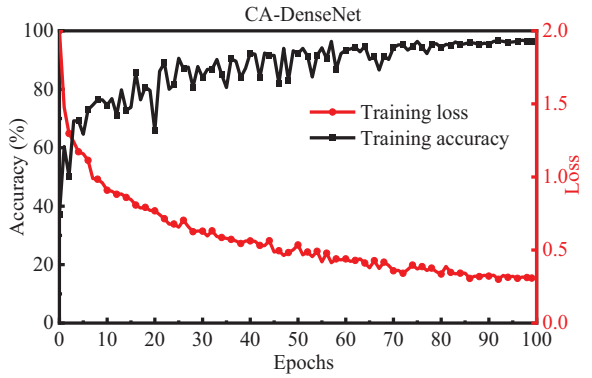

Fig. 9. Training results of the CA-DenseNet.

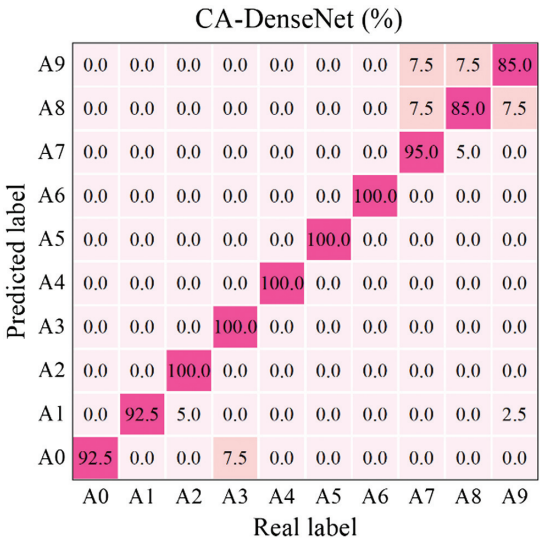

Fig. 10. Confusion matrix diagram of the CA-DenseNet.

It can be seen from Table IV and Fig. 11 that the proposed method that combines the CA-DenseNet and SSTG achieves the best performance, namely, the highest average test result (95.0%) and the lowest standard deviation (52.8%). In comparison, the accuracy of 1D signals on each model are much lower than the classification accuracy of time-frequency images on each model, which proves the necessity of time-frequency processing. When the images are input to the model. the test results of the images generated by SSTG on the CA-DenseNet model was 1.9% higher than those of the images generated by SST, 1.4% higher on the DenseNet model, 1.6% higher on the ResNet model, 11.1% higher on the ShuffleNet model, and 4.7% higher on the ConvNeXt model. The main reasons are as follows. First, there is no obvious difference between the original vibration signals, and it is difficult for the model to extract the corresponding features, so the classification accuracy is low. Second, the results indicate that the images

Table IV. Test results of different inputs on different models

	Average test results (%)			Standard deviation (%)		
Model	**SSTG**	**SST**	**1D**	**SSTG**	**SST**	**1D**
CA-DenseNet	**95.0**	93.1	81.4	**52.8**	78.6	95.8
DenseNet	93.6	91.9	79.8	66.3	41.1	68.5
ResNet	93.4	92.1	81.5	139.3	132.9	115.7
ShuffleNet	83.9	83.3	79.5	142.0	119.0	84.5
ConvNext	90.3	89.4	88.8	70.2	65.6	68.9

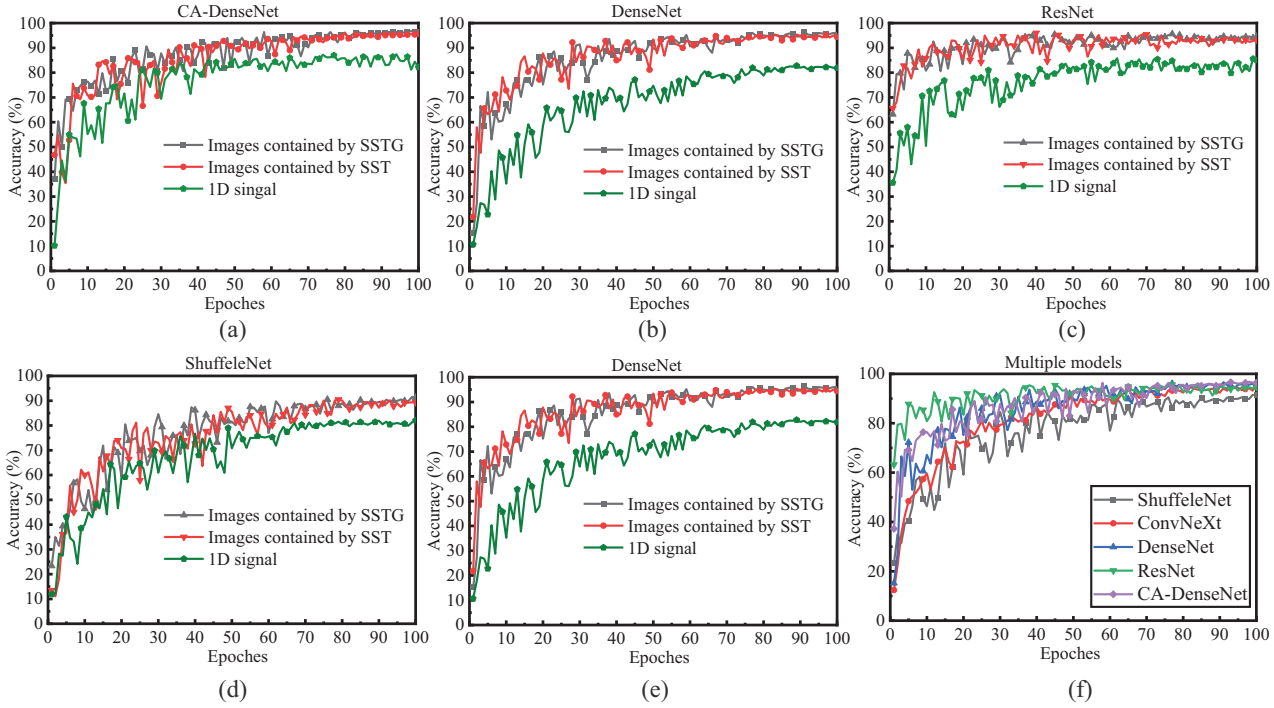

Fig. 11. Comparing experimental results: (a), (b), (c), (d), and (e) The training results of images generated by SSTG, images generated by SST and 1D signal on CA-DenseNet, DenseNet, ResNet, ShuffeleNet, and ConvNeXt, respectively; (f) the training results of images generated by SSTG on different models.

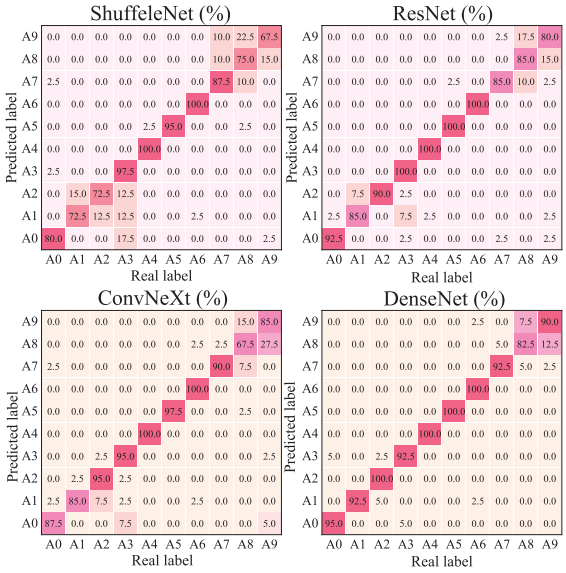

Fig. 12. Confusion matrixes of different classification models.

generated by SSTG have more easily extractable features than those generated by SST, likely due to the high time-frequency characteristics of SST, which makes the time-frequency information of vibration signals processed by SST concentrated and presents clearer lines. The Gaussian high-pass filtering primarily filters out low-frequency information, retaining high-frequency information, which enhances the detailed line features in the images, benefiting the feature extraction process in intelligent classification.

To validate the superiority of the proposed model, this work uses different models to train images of SSTG. Figure 11(f) shows the training process. The training accuracy of the proposed model is higher than those of the comparison models, possibly because of the following reasons. (1) The DenseNet model's dense connection structure allows the model to extract high-dimensional features and preserve initial features, thus achieving feature reuse during training. (2) The CA attention mechanism enables the model to focus not only on channel information but also on the space and position information of feature matrices, ultimately improving classification accuracy.

Confusion matrices are made for the test results of the images generated by SSTG on the Densenet, ResNet, ShuffeleNet, and ConvNeXt, as shown in Fig. 12. The confusion matrices reveal that the proposed method has the highest test accuracy among all models tested.

However, these models all indicate that classification errors are mainly concentrated in labels A8 and A9. This may be due to the fact that labels A8 and A9 correspond to different degrees of rail wear faults. There are no obvious differences in their time-frequency information after SSTG processing. Thus, it is difficult to distinguish them in the intelligent classification model.

Table V. Results of CA-DenseNet ablation experiments

Model	Accuracy (%)
Original model	95.0
×·CA	93.6
×·Dense block	48.4
×·CA & Dense block	40.6

E. ABLATION EXPERIMENT

It can be seen from section III that the model proposed in this paper is mainly composed of dense blocks, CA mechanism. In order to explore the contribution of different modules to classification accuracy, ablation experiment is conducted in this section. The experimental test results of model are shown in Table V. The '×' denotes 'elimination'.

By removing CA, the accuracy of the model decreases by 1.4%, which is consistent with the comparative experimental results in Section IV, indicating that CA module has a certain role in improving the accuracy of the model. After dense block is removed, the model test accuracy drops dramatically, because CA-DenseNet is mainly composed of dense block, and it plays a dominant role in the model test results.

F. ROBUSTNESS ANALYSIS

Signal-to-noise (SNR) ratio refers to the ratio between a signal and noise, which is used to describe the quality and reliability of a signal in the presence of noise [37]. Noises of SNR 30 and 40 dB are added to the original signals, and the time-frequency images containing noise information are obtained by the SSTG processing of the noise signals. These images are used as the input of multiple models, average accuracy and standard deviation of five tests results of each model are shown in Table VI. Figure 13 shows the accuracy of different models in different noise environments.

After adding 40 dB and 30 dB noise, the classification accuracies of the CA-DenseNet model decreased by 1.2% and further decreased by another 1.0%, respectively. Other models are more significantly affected by the noise influence, with a larger decrease in model classification accuracy. Compared with other models, the proposed model is less affected by noise, as it yields the lowest standard deviation for all the noise interferences in Table VI.

Table VI. Comparison of the test results of different models in noisy environments

	Noiseless		SNR = 40 dB		SNR = 30 dB	
Model	Average test results (%)	Standard deviation (%)	Average test results (%)	Standard deviation (%)	Average test results (%)	Standard deviation (%)
CA-DenseNet	95.0	**52.8**	93.8	**66.9**	92.8	**63.5**
DenseNet	93.6	66.3	92.2	70.1	90.0	68.5
ResNet	93.4	139.3	91.8	128.9	89.9	130.1
ShffeleNet	83.9	142.0	80.5	135.8	77.8	125.67
ConvNext	90.3	70.2	88.5	68.5	86.2	65.3

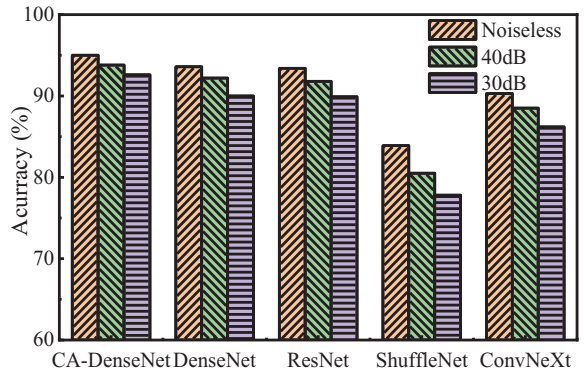

Fig. 13. Test results of different models in noisy environments.

Thus, the experimental results demonstrate the good robustness of the proposed method.

V. CONCLUSIONS

This paper proposes a fault diagnosis method for linear guide rails and ball bearings based on SSTG and CA-DenseNet. The following conclusions can be drawn from the experimental data analysis. (1) As a new signal processing method, SSTG can convert vibration signals into time-frequency images. Introducing Gaussian high-pass filtering can enhance the features of image details and eliminate the impact of manual feature extraction. Furthermore, it can realize fault feature visual enhancement. (2) Using the CA-DenseNet classification model with an attention mechanism is effective, because its dense connection structure enables feature reuse. By adding the CA module to each dense block, the model can focus on both the channel and spatial position information of the feature matrices during training, thus demonstrating the superiority of the proposed model in comparative experiments.

This proposed method is practical for diagnosing faults in linear guide rails and can also provide insights for diagnosing other faults.

Acknowledgments

This work was partly supported by the following organizations: National Natural Science Foundation of China (Grant Nos. 52375522, 52207036, and 62203010), the Anhui Provincial Natural Science Foundation (Grant Nos. 2308085Y03 and 2208085QE167), the Project of the Outstanding Young Talents in Colleges and Universities of Anhui Province (Grant No. gxyqZD2022006), the College Natural Science Research Key project of Anhui Education Department (Grant No. KJ2021A0018), and the University Outstanding Youth Research Project of Anhui Province (Grant No. 2022AH030016).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

REFERENCES

[1] M. S. Kim, J. P. Yun, and P. Park, "An explainable convolutional neural network for fault diagnosis in linear motion guide," *IEEE Trans. Ind. Inf.*, vol. 17, no. 6, pp. 4036–4045, Jun. 2021.

[2] V. C. Tong, S. W. Kwon, and S. W. Hong, "Modeling of moving table with linear roller guides subjected to geometric errors in guide rails," *Int. J. Precis. Eng. Manuf.*, vol. 21, no. 10, pp. 1903–1919, Oct. 2020.

[3] X. F. Ni, Z. J. Ma, J. W. Liu, B. Shi, and H. L. Liu, "Attention network for rail surface defect detection via consistency of Intersection-over-Union(IoU)-guided center-point estimation," *IEEE Trans. Ind. Inf.*, vol. 18, no. 3, pp. 1694–1705, Mar. 2022.

[4] S. Su, J. Qu, Y. Cao, R. Li, and G. Wang, "Adversarial training Lattice LSTM for named entity recognition of rail fault texts," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 11, pp. 21201–21215, 2022.

[5] Y. J. Chen, M. Rao, K. Feng, and M. J. Zuo, "Physics-Informed LSTM hyperparameters selection for gearbox fault detection," *Mech. Syst. Signal Process.*, vol. 171, p. 108907, May 2022.

[6] Y. Chen, M. Rao, K. Feng, and G. Niu, "Modified varying index coefficient autoregression model for representation of the nonstationary vibration from a planetary gearbox," *IEEE Trans. Instrument. Meas.*, vol. 72, art. on. 3511812, 2023.

[7] J. C. Yao and T. Han, "Data-driven lithium-ion batteries capacity estimation based on deep transfer learning using partial segment of charging/discharging data," *Energy*, vol. 271, art. no. 127033, May 2023.

[8] S. Lu, R. Yan, Y. Liu, Q. Wang, "Tacholess speed estimation in order tracking: a review with application to rotating machine fault diagnosis," *IEEE Trans. Instrument. Meas.*, vol. 68, pp. 2315–2332, Jul. 2019.

[9] N. Qin, J. H. Du, Y. M. Zhang, D. Q. Huang, and B. Wu, "Fault diagnosis of multi-railway high-speed train bogies by improved federated learning," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 7184–7194, Jun. 2023.

[10] F. Q. You, D. A. L. Wang, G. H. Li, and C. H. Chen, "Fault diagnosis method of escalator step system based on vibration signal analysis," *Int. J. Control Automat. Syst.*, vol. 20, no. 10, pp. 3222–3232, Oct. 2022.

[11] P. Chommuangpuck and T. Wanglomklang, "Fault detection and diagnosis of linear bearing in auto core adhesion mounting machines based on condition monitoring," *Syst. Sci. Control Eng.*, vol. 9, no. 1, pp. 290–303, 2021.

[12] R. Borowiecki, V. A. Kravchinsky, M. van der Baan, and R. H. Herrera, "The synchrosqueezing transform to evaluate paleoclimate cyclicity," *Comput. Geosci.*, vol. 175, art no. 105336, Jun. 2023.

[13] X. Ni, H. Liu, Z. Ma, C. Wang, and J. Liu, "Detection for rail surface defects via partitioned edge feature," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 6, pp. 5806–5822, 2022.

[14] J. Zhang, J. Chen, S. Wang, and L. Wang, "Rail surface defect detection through bimodal RSDINet and three-branched evidential fusion," *IEEE Trans Instrument. Meas.*, vol. 72, art. no. 2508714, 2023.

[15] J. Ye, E. Stewart, Q. Chen, C. Roberts, A. M. Hajiyavand, and Y. Lei, "Deep learning and laser-based 3-D pixel-level rail surface defect detection method," *IEEE Trans. Instrument. Meas.*, vol. 72, art. no. 2513612, 2023.

[16] M. Niu, Y. Wang, K. Song, Q. Wang, Y. Zhao, and Y. Yan, "An adaptive pyramid graph and variation residual-based anomaly detection network for rail surface defects," *IEEE Trans. Instrument. Meas.*, vol. 70, art. no. 5020013, 2021.

[17] Z. Chen, Q. Wang, Q. He, T. Yu, M. Zhang, and P. Wang, "CUFuse: camera and ultrasound data fusion for rail defect detection," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 11, pp. 21971–21983, 2022.

[18] M. M. Liu, R. P. Gao, J. Zhao, Y. R. Wang, and W. Shao, "A multi-population state optimization algorithm for rail crack fault diagnosis," *Meas. Sci. Technol.*, vol. 33, art. no. 055014, May 2022.

[19] K. Wang, Y. Shang, and Y. Lu, "An improved second-order multi-synchrosqueezing transform for the analysis of non-stationary signals," *J. Dyn. Monit. Diagnost.*, vol. 2, pp. 183–189, 2023.

[20] G. W. Zhang and J. H. Gao, "Inversion-driven attenuation compensation using synchrosqueezing transform," *IEEE Geosci Remote Sens. Lett.*, vol. 15, no. 1, pp. 132–136, Jan. 2018.

[21] X. D. Jiang, J. X. Cao, Z. D. Su, J. C. Fu, and S. C. Shi, "Hydrocarbon detection based on empirical mode decomposition, teager-kaiser energy, and the cepstrum," *Front. Earth Sci.*, vol. 10, art. no. 923173, Sept. 2022.

[22] D. Zhang and D. Feng, "Wind turbine planetary gearbox fault diagnosis via proportion-extracting synchrosqueezing chirplet transform," *J. Dyn. Monit. Diagnost.*, vol. 2, pp. 177–182, 2023.

[23] X. Mao, "A concentrated time-frequency method for reservoir detection using adaptive synchrosqueezing transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[24] M. H. Wei, J. W. Yang, D. C. Yao, J. H. Wang, and Z. S. Hu, "Fault diagnosis of bearings in multiple working conditions based on adaptive time-varying parameters short-time Fourier synchronous squeeze transform," *Meas. Sci. Technol.*, vol. 33, art. no. 124002, Dec. 2022.

[25] W. Wiratama, Y. Ahn, I. Marzuki, and D. Sim, "Adaptive Gaussian low-pass pre-filtering for perceptual video coding," *IEIE Trans. Smart Process. Comput.*, vol. 7, no. 5, pp. 366–377, 2018.

[26] N. Duan, J. Wang, T. Zhao, W. Du, X. Guo, and J. Wang, "A Novel adaptive fault diagnosis method for wind power gearbox," *IEEE Access*, vol. 9, pp. 11226–11240, 2021.

[27] Y. Liu, Z. S. Jiang, H. Z. Huang, and J. W. Xiang, "A TEO-based modified Laplacian of Gaussian filter to detect faults in rolling element bearing for variable rotational speed machine," *IET Sci. Meas. Technol.*, vol. 15, no. 2, pp. 193–203, Mar. 2021.

[28] X. B. Liu, W. B. Mei, and H. Q. Du, "Detail-enhanced multimodality medical image fusion based on gradient minimization smoothing filter and shearing filter," *Med. Biol. Eng. Comput.*, vol. 56, no. 9, pp. 1565–1578, Sept. 2018.

[29] W. P. Guo, Z. M. Xu, and H. B. Zhang, "Interstitial lung disease classification using improved DenseNet," *Multimed. Tools Appl.*, vol. 78, no. 21, pp. 30615–30626, Nov. 2019.

[30] S. Brahimi, N. Ben Aoun, A. Benoit, P. Lambert, and C. Ben Amar, "Semantic segmentation using reinforced fully convolutional densenet with multiscale kernel," *Multimed. Tools Appl.*, vol. 78, no. 15, pp. 22077–22098, Aug. 2019.

[31] B. L. Chen, T. S. Zhao, L. P. Zhou, J. Yang, J. H. Liu, and L. Q. Lin, "An optimized segmentation scheme for ambiguous pixels based on improved FCN and DenseNet," *Circ. Syst. Signal Process.*, vol. 41, pp. 372–394, Jan. 2022.

[32] D. P. Chowdhury, S. Bakshi, C. Pero, G. Olague, and P. K. Sa, "Privacy preserving ear recognition system using transfer learning in industry 4.0," *IEEE Trans. Ind. Inf.*, vol. 19, no. 5, pp. 6408–6417, May 2023.

[33] C. H. Qian, Q. S. Jiang, Y. H. Shen, C. R. Huo, and Q. K. Zhang, "An intelligent fault diagnosis method for rolling bearings based on feature transfer with improved DenseNet and joint distribution adaptation," *Meas. Sci. Technol.*, vol. 33, art no. 025101, Feb. 2022.

[34] Y. Zhang, N. Qin, D. Huang, B. Wu, and Z. Liu, "High-accuracy and adaptive fault diagnosis of high-speed train bogie using dense-squeeze network," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2501–2510, 2022.

[35] Y. J. Zhou, X. Y. Long, M. W. Sun, and Z. Q. Chen, "Bearing fault diagnosis based on Gramian angular field and DenseNet," *Math. Biosci. Eng.*, vol. 19, no. 12, pp. 14086–14101, 2022.

[36] R. H. Shang et al., "Hyperspectral image classification based on pyramid coordinate attention and weighted self-distillation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, art. no. 5544316, 2022.

[37] H. R. Fang, J. L. An, H. Liu, J. W. Xiang, B. Zhao, and F. Dunkin, "A lightweight transformer with strong robustness application in portable bearing fault diagnosis," *IEEE Sens. J.*, vol. 23, pp. 9649–9657, May 2023.