Multi-branch convolutional neural network with generalized shaft orbit for fault diagnosis of active magnetic bearing-rotor system

Yan Xunshi [a,b,c,*], Zhang Chen-an [d], Liu Yang [d,e]

[a] Institute of Nuclear and New Energy Technology, Tsinghua University, China
[b] The Key Laboratory of Advanced Reactor Engineering and Safety, Ministry of Education, China
[c] Collaborative Innovation Center of Advanced Nuclear Energy Technology, China
[d] State Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, China
[e] School of Engineering Science, University of Chinese Academy of Sciences, China

A B S T R A C T

Fault diagnosis based on vibration signals in active magnetic bearing-rotor systems is an important research topic. However, it is difficult to obtain discriminative features to represent faults due to the nonlinear and non-stationary characteristics of the vibration signals and diverse sources of failures. Hence, this paper proposes a novel end-to-end learning mechanism of multi-sensor data fusion to learn fault representation based on the structural characteristics of active magnetic bearings. Taking the five displacement sensors of active magnetic bearing as signal sources, generalized shaft orbits are constructed and converted into discrete 2D images. Based these 2D images, a multi-branch convolutional neural network is designed to achieve high discriminative features and fault types. The experiments are performed on the rig supported by active magnetic bearings, and the effectiveness of the proposed algorithm is verified, proving it suitability in cases with changing rotating speeds and sample lengths.

1. Introduction

Active Magnetic Bearing (AMB) has been developed in the past 40 years, using electromagnetic force as the supporting force to achieve non-contact support with the rotor. Due to the characteristics of non-lubrication and friction-free, AMBs are widely used in many industrial products. In the main circulator of the high-temperature gas-cooled reactor nuclear power plant, the usage of AMBs avoids the contamination of the circulating medium helium with lubricating oil [1]; Due to its active control capability, AMB can suppress the rotor vibration at the bending frequency and enable the operation of high-speed motors above the critical speed; AMB can produce a large bearing capacity and are suitable in large-scale rotors or rotating machinery with complex aerodynamic characteristics and high loads [2]. AMBs are generally applied in cases of high safety, high speed, and heavy load; therefore, once the machines malfunction, it will cause serious consequences and losses. Automatic fault diagnosis research on the AMB-Rotor system (AMB-R) is required to provide maintenance decisions and failure warnings.

The working principle of AMB is as follows. The displacement sensors of the AMB measure the position of the rotor and transmit the displacement signals to the controller. Compared with reference signals, the controller computes the deviations and adjusts them in the feedback system and then sends commands to the power amplifiers. According to the commands, the power amplifiers change currents in the electromagnet to generate an electromagnetic force, drawing the rotor back to the center position [3]. Due to complexity of AMB-R, automatic fault diagnosis methods have rarely appeared in the literatures. The research on rotor dynamics based on the failure of AMB-R is helpful in understanding the failure mechanism but is hardly applied in a direct manner to the automatic judgment of failure. At present, the automatic diagnosis method of AMB-R commonly used in engineering is mainly based on ISO14839 [4]. According to the amplitude of signals threshold collected by the sensors, the state of AMB-R is judged as normal or faulty, which is easily implemented and linked to a distributed control system. However, this method is prone to noise interruptions, lacks ease in issuing early warnings and provides virtually no possibility to distinguish the type of faults.

The AMB-R is a typical type of rotating machinery. In recent years, the combination of signal processing and machine learning has developed rapidly in the intelligent fault diagnosis of rotating machinery. Zhao [5] performed wavelet transform on the vibration signal, converted the transformed data into a two-dimensional form and used the

---

data as input for a deep residual neural network to analyze the fault type of the rolling bearings. Ou [6] calculated the Laplacian energy of the fault signals as features, used the weighted *K* nearest neighbor method with the Mahalanobis distance metric to diagnose the bearing faults. Zhao [7] applied the Hilbert–Huang transform to the vibration signal as the fault feature and combined it with a deep convolutional neural network to realize the fault classification of a planetary gearbox. Lei [8] selected features with small intra-class dispersion and large inter-class dispersion from 43 sets of time-domain, frequency-domain and time–frequency-domain features, and classified faults together with an adaptive neuro-fuzzy inference system. In general, based on signal processing technology, the manually designed features merge the domain knowledge but have certain limitations in terms of fully reflecting the difference in characteristics between fault types.

Different from the methods of manually designed features using signal processing technology, deep neural networks can automatically learn features with informative semantics from the raw data even without domain knowledge, and the representation learning mechanism for obtaining fault features is called end-to-end learning. Shao [9] input the original vibration signal fragments into a series of deep autoencoders with different activation functions, aggregated the output results of the hidden layers in the autoencoders to obtain the fault diagnosis results of rolling bearings. Wen [10] divided the time-domain signal captured by the vibration sensor into a series of equal-length segments serving as a row of pixels in a two-dimensional image, and LeNet5 was used to identify the image to achieve bearing fault classification. Jiang [11] sampled vibration signals according to different sampling rates to form three signal sequences with different sampling rates, and used a Convolutional Neural Network(CNN) to classify the faults under varying working conditions. Wang [12] captured the fragments of the original signals by sliding windows, and adopted a bidirectional gated recurrent unit network to obtain two hidden layers that were combined as a feature to predict the wear of CNC machine tools. Hu [13] compressed the signal into the coefficients of the basic vectors by using compressed sensing technology, and intercepted the fragments from the compressed signal to form a 2D image to achieve fault classification results by a multiscale network. Zhang [14] reshaped the acoustic signals into a 2D image, and viewed each point in the image as a vertex on the graph that was related to its 8 neighbors by edges, transforming the image into graph, and a graph convolutional neural network was employed for further feature learning. Generally, it is necessary to reconstruct the data form of the original signals when using end-to-end learning to learn an effective feature representation under the conditions of the limited samples and varying working conditions.

Compared with signals captured by a single sensor, those signals captured by multiple sensors provide more information of the operating states of the system. Through an effective fusion mechanism, a more distinguishable fault representation can be obtained. Duan [15] proposed that the multi-sensor data fusion methods in fault diagnosis of rotating machinery can be divided into data level, feature level and decision level. Safizadeh [16] installed two acceleration and load sensors in the bearing-rotor system, extracted time-domain and frequency-domain features, classified faults according to the K-nearest neighbor and used the waterfall model for data fusion. Yan [17] regarded the sampling points captured by multiple sensors at the same time as vibration points in a high-dimensional space, and acquired a histogram description of faults by calculating the distance from all vibration points to the zero center. Wang [18] placed the multiple signals side by side, cut signals into several segments along the time axis, and connected them in a 2D matrix which was input into a CNN to achieve classification results. Te [19] captured spatial–temporal relationship information between multiple sensors, and the series of signals were converted into a symbol sequence. The symbol transfer matrix between different sequences was constructed, and the result was merged as a 2D image input to a CNN. Xia [20] directly took each signal as a raw of an image, and the multiple signals were constructed into a 2D

image and classified by CNN. Jin [21] used various types of sensors to capture signals, performed Fast Fourier Transform (FFT) and joined feature vectors input to a 1D CNN. Liu [22] calculated low-dimensional features of signals captured by each sensor and then combined features into high-dimensional vectors. Deep auto-encoder and full-connected deep neural network were used to separately reduce the feature dimensionality and classify faults. Zhao [23] placed the signals captured by different sensors side by side, cut them into several consecutive segments, and extracted the time- and frequency-domain features of the corresponding segments. These features were directly connected, and the characteristics of the time segment sequence were used as inputs to a Recurrent Neural Network(RNN) to assess the health states of the mechanical system. Wang [24] implemented a CNN to extract features from multi-level signals captured by vibration and current sensors and used RNN to fuse the information in one framework. Although the fusion algorithms have achieved good results, it still needs to install additional sensors in the system, which increases the complexity of the system. Moreover, the contribution of the installation positions of the sensors is usually neglected resulting in information loss.

The shaft orbit is considered as a representation of the information fusion of two sensors placed vertically in the same plane and contains substantial intuitive information about the operation state of the rotor. Shi [25] extracted the high-resolution spectrum of the signal to obtain a pure shaft orbit that removes interference terms, calculated moment and curve features, and determined misalignment, imbalance, and rub-impact failure; Yan [26] divided the area covered by the shaft orbit into sub-regions such as grids, sectors, and counted the number of sampling points of the shaft orbit in sub-regions to construct a histogram as features, and used the improved AdaBoost algorithm to classify the faults. Wen [27] obtained the entropy values on each sub-regions the points of shaft orbit falled and used fuzzy C-means clustering algorithm as a tool for fault discrimination. Wu [28] divided the five types of shaft orbits into 15 subcategories associated with the severity of the faults, and utilized the fine-tuned LeNet5 to classify the faults according to the images corresponding to the shaft orbits. The requirements for the sensor installation locations limit the application of shaft orbits. Moreover, features based on the shaft orbit are manually extracted and are easily susceptible by noise interference, such that preprocessing steps are necessary to remove noises.

Intelligent fault diagnosis is essentially a pattern recognition problem, and feature extraction is the key problem. Compared with mechanical bearings, AMBs are mechatronic equipment; therefore, the sources of AMB-R failures are complex, which may come from rotors, couplings, gear boxes in mechanical components, or from power amplifiers and sensors, or be generated due to the lack of robustness of the control algorithm. The fault phenomenon of AMB-R is highly complicated, and the nonlinearity and non-stationarity of the fault signals are prominent and difficult to identify, which presents more great challenges for the fault description of the AMB-R than that of traditional rotating machinery supported by mechanical bearings.

To overcome the abovementioned problems, the manuscript proposes a fault diagnosis algorithm based on the generalized shaft orbit combined with the characteristics of AMB-R to achieve multi-sensor fusion, and implements end-to-end learning with CNN to classify fault types. The main contributions are threefold as follows:

First, the signals captured by the five displacement sensors integrated in AMB are used for analysis, which avoids installing additional vibration sensors and reduces the complexity of the system. Moreover, based on the special positional relationship between the sensors, the traditional concept of shaft orbit is expanded, and the generalized shaft orbits by any two sensors in AMB are constructed, which provide sensor coupling information for multi-sensor fusion.

Second, the Multi-branch CNN(MCNN) is designed to fuse the information of multiple sensors without the need to manually extract features, and to obtain a high-level representation of the fault in an end-to-end learning manner. To our best knowledge, it is the first time that
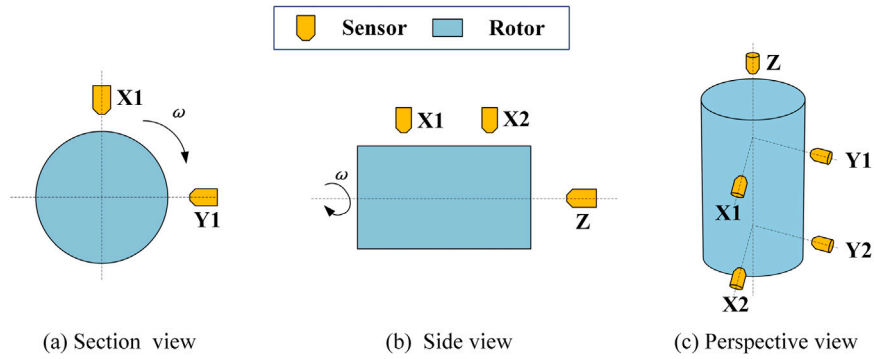
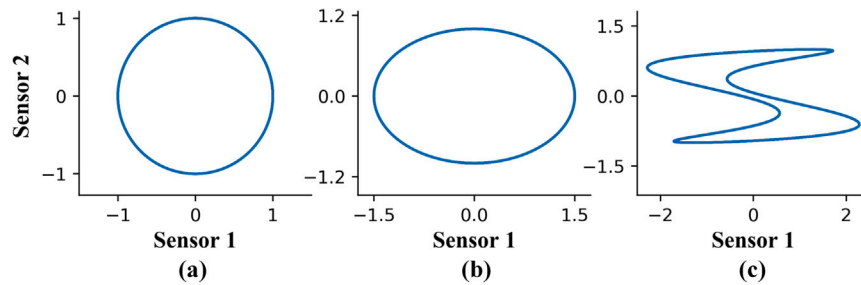**Fig. 1.** Sensors arrangement in AMB.

Fig. 2. The simulation of shaft orbits.

deep learning methods have been introduced into the fault diagnosis research of AMB-R.

Third, it is verified that the proposed algorithm can be used in a variety of working conditions through experiments, such as speed changes, and can also be utilized for training samples with different lengths, which improves the application scope of the algorithm.

The manuscript is organized as follows. Section 2 introduces the Shaft Vibration Image (SVI). MCNN is detailed in Section 3. A series of experiments are performed on the test rig in Section 4. Section 5 presents the summary and outlook.

2. Shaft vibration image

The five displacement sensors of AMB have two functions: one is to feedback deviation information in the control loop, and the other is to monitor the operating state of AMB-R. Fig. 1 shows positions of sensors, which are named X1, Y1, X2, Y2 and Z. Z is an axial displacement sensor, and the others are radial sensors. X1/Y1 and X2/Y2 are both installed orthogonally. X1/X2, Y1/Y2 are located in the same axis line along the rotor. According to the definition of traditional shaft orbit, two shaft orbits are obtained by the signals captured by X1/Y1, X2/Y2. Obviously, they do not fully show the coupling relationship between the five sensors, leading to information loss of the fault representation. Hence, the generalized shaft orbit constructed by the signals is captured by any two sensors in AMB and its discretized data form will be given to describe faults comprehensively.

2.1. Generalized shaft orbit

Shaft orbit is a two-signal combination requiring two vibration sensors placed orthogonally in the same plane perpendicular to the axis. It not only reflects the motion orbit of the geometric center of the rotor, but also fuses the information of the two sensors and embodies the correlation between the two channels of signals. As the shape of the shaft orbit varies under different health states of rotating machinery, shaft orbit is an effective tool to describe faults and often employed in monitoring.

Essentially, shaft orbit is a kind of Lissajous figure and reflects the combination result of two signals. Under normal circumstances, two vibration sensors S_1, S_2 capture two signals x_1, x_2 denoted as:

$$
\begin{aligned}
x_1 &= A_1 sin(\omega t + \varphi) \\
x_2 &= A_2 sin(\omega t + \varphi + \theta)
\end{aligned}
\tag{1}
$$

where ω is the rotational angular frequency, and θ is the phase difference of the two signals, i.e., the angular difference of the S_1, S_2 positions relative to the axis center.

When S_1, S_2 are placed orthogonally in the same plane perpendicular to the axis, i.e., $\theta = \pi/2$, x_1, x_2 are synthesized into a shaft orbit.

If $A_1 = A_2 = A$:

$$
x_1{}^2 + x_2{}^2 = A^2
\tag{2}
$$

and the shape of the shaft orbit in Eq. (2) is a circle (see Fig. 2(a)) in the coordinate system.

If $A_1 \neq A_2$:

$$
\frac{x_1{}^2}{A_1{}^2} + \frac{x_2{}^2}{A_2{}^2} = 1
\tag{3}
$$

and the synthesized orbit in Eq. (3) is an ellipse (see Fig. 2(b)).

When an abnormality occurs, the shaft orbit is no longer a circle. For example, if there is a large number of triple fundamental components in x_1 as follows:

$$
\begin{aligned}
x_1 &= A_1 sin(\omega t + \varphi_1) + A_2 sin(3\omega t + \varphi_2) \\
x_2 &= A_1 sin(\omega t + \varphi_1 + \pi/2)
\end{aligned}
\tag{4}
$$

the shaft orbit is distorted compared to that of a normal circle (see Fig. 2(c)).

However, the shaft orbit has strict requirements on the positions of the two sensors. If the position limitation is relaxed, the two signals x_1 and x_2 captured by any arbitrary sensors in AMB-R can also synthesize an orbit G which is defined as "generalized shaft orbit" (GSO) as Eq. (5):
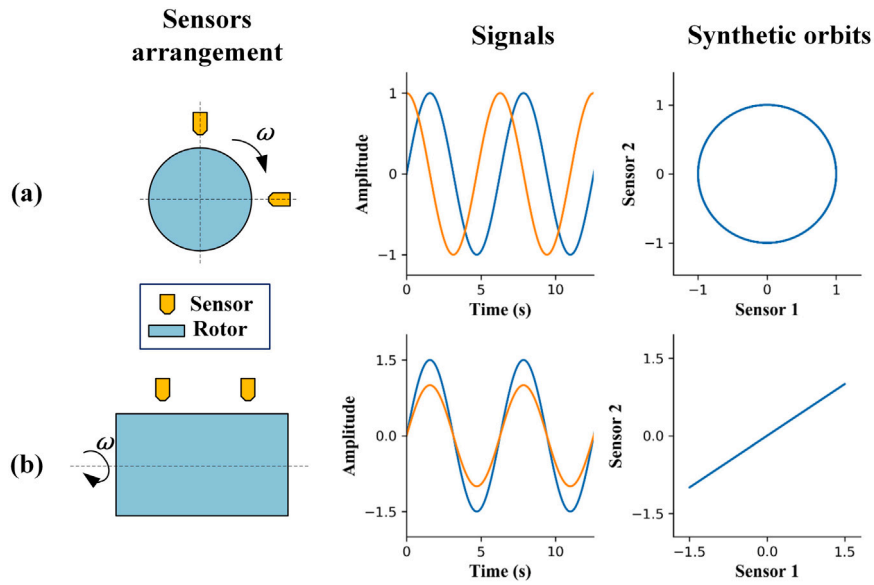
$$
G = x_1 + j x_2
\tag{5}
$$

Fig. 3. Sensors arrangements, simulated signals and their synthetic orbits. (a) Conventional shaft orbit (b) GSO.
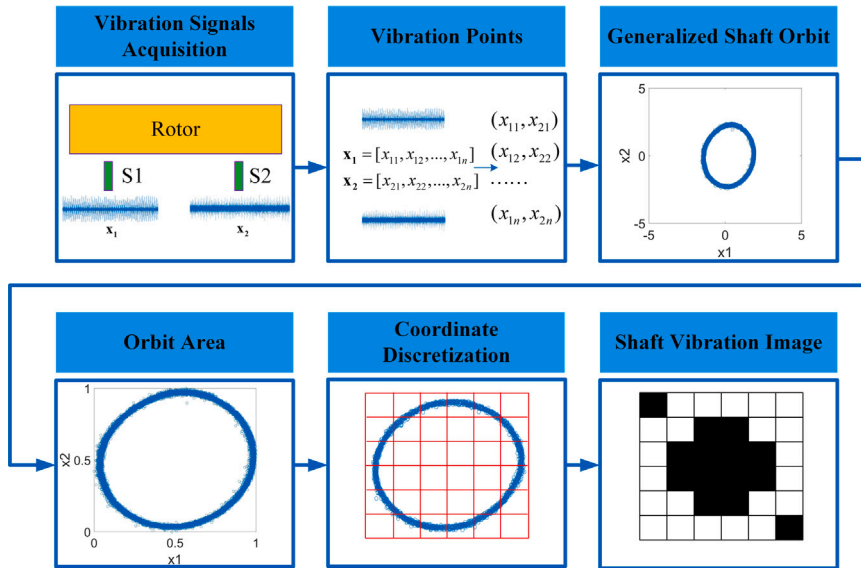

Fig. 4. The formation process of the SVI.

where $j = \sqrt{-1}$.

Analogous to the conventional shaft orbit, the amplitudes of any two signals make up coordinate points that are connected into GSO in the Cartesian coordinate system. Taking X1 and X2 in AMB as an example, the two sensors are located on the same line along the axial direction of the rotor, i.e., $\theta = 0$. In the normal state, the synthetic GSO is

$$x_1 = \frac{A_1}{A_2} x_2 \tag{6}$$

As Eq. (6) indicates, the GSO is a straight line. When a failure occurs in the system, the GSO is distorted and no longer a straight line.

Fig. 3 shows the comparison of the shaft orbit and a kind of GSO, including their sensor arrangements and simulated signals. It can be concluded that shaft orbit is a special case of GSO. From the shaft orbit to the GSO, it extends the usage of the shaft orbit and explores the coupling relationship in more combinations of sensors, providing abundant information for diagnosing faults.

2.2. Shaft vibration image

The GSO contains discriminative information and can be adopted as a basis for determining the health states of AMB-R. To input GSOs into machine learning models, it needs to transform GSO into a discretized representation as follows.

First, the signals $x_1 = [x_{11}, x_{12}, \ldots, x_{1n}]$, $x_2 = [x_{21}, x_{22}, \ldots, x_{2n}]$ are captured by two sensors S_1, S_2 fixed at arbitrary positions. A series of vibration points $(x_{1t}, x_{2t}), t = 1, 2, \ldots, n$ are drawn in the coordinate system X_1-X_2 to form GSO G.

Then, x_1, x_2 are mapped to the range of [0,1] separately and the coverage of G in the coordinate system is normalized in two directions, implying G is constrained to a square area named the orbit area. The orbit area normalizes the amplitude of the signals and eliminates the difference in the absolute magnitude of various GSOs. It should be noted that the absolute magnitude of the amplitude is also an important criterion, and faults can be identified according to a simple threshold in some cases, which is out of our scope. The relationship between the amplitudes of different signals is mainly considered in this paper.
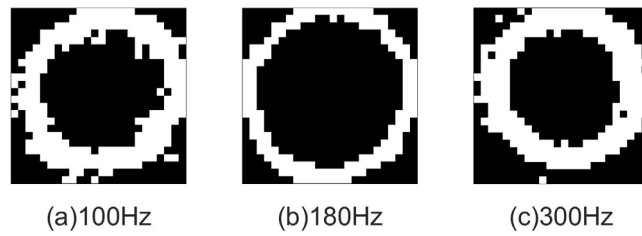
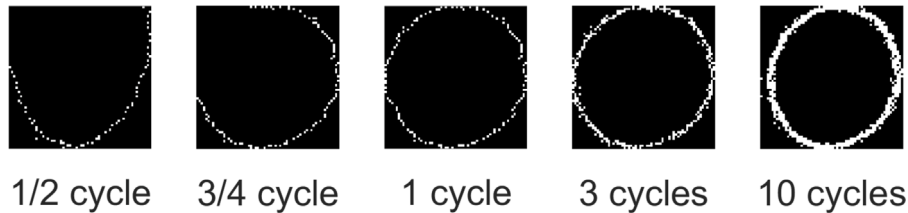**Fig. 5.** Illustration of SVIs at various speeds.

Fig. 6. The SVIs constructed by samples with variable lengths.

Finally, the coordinates are discretized, and the orbit area is divided into L^2 grids. The grids where the vibration points fall are set as active areas, vice versa, are set as inactive areas. Specifically, the active area is assigned with a value 1, and the inactive area is represented by 0. The orbit area can be regarded as a black-and-white image with L^2 pixels called shaft vibration image, which is essentially a 1–0 matrix.

Fig. 4 shows the formation process of the SVI.

2.3. Discussion

The SVI is based on the concept of the GSO and derives the fusion of two-sensor data, reflecting the coupling relationship between the signals captured by two arbitrary fixed sensors in AMBs. The SVI discretizes the two signals into a black-and-white image and compresses raw data, which transforms the problem of fault diagnosis into a shape recognition task. From the formation process of the SVI, the following characteristics can be summarized:

First, the SVI is independent of the rotating speed of the rotor. As the AMB-R operates normally, the SVI stably presents a standard shape regardless of the speed. Fig. 5 gives some SVIs collected at different rotating speeds during the speed increasing process of a high-speed motor and the shapes of SVIs are almost consistent. Hence, the algorithms based on SVIs can be utilized under the conditions of varying speeds.

Second, as long as the length of sample signals is greater than one cycle, the SVI presents a whole shape, as exhibited in Fig. 6. In a normal operating state, the shape of the SVI does not change significantly as the number of sampling points increases. Hence, when the signal length is greater than one cycle, the SVI can be constructed from signals with variable lengths or variable sampling rates, and still maintain a relatively fixed shape. It provides the opportunity to detect faults in the shortest time while making it unnecessary to collect enough data points.

Third, SVI represents the shape of the GSO with a clear physical meaning. The distinguishability of the image shapes makes it possible to achieve a high accuracy of fault classification.

Strictly speaking, the SVI is a reorganization of the raw data, which is a kind of data-level fusion. The description ability of SVI is still limited; thus, it requires further information mining in machine learning models.
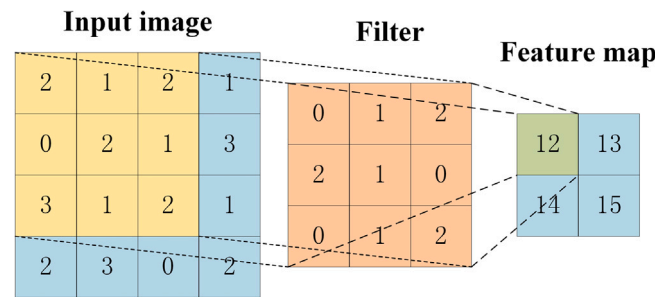

Fig. 7. Convolutional layer.

3. Multi-branch CNN

As mentioned previously, SVI is a discrete form of GSO that converts two signals into a black-and-white image to provide an input for machine learning models. It is necessary to utilize deep learning models to further extract high-level semantic information to carry out effective fault classification. CNN, as a successful deep learning model, is chosen as the basic model. We combine SVI with improved CNN to achieve end-to-end learning.

3.1. SVI-CNN

LeCun [29] first applied the backpropagation algorithm to CNN and designed LeNet to succeed in handwritten digit recognition. With the development of new deep learning technologies such as ReLu and dropout, AlexNet became the first modern deep convolutional network and achieved a major breakthrough in image classification [30]. Compared to fully connected deep neural networks, CNN is structurally composed of convolutional layers, activation functions, pooling layers, and fully connected layers, and its input can be 1D signals or 2D images.

Convolutional layer—Convolution is an important operation for extracting features in image understanding. A plurality of convolution kernels is introduced in the convolutional layer, and the convolution kernel is convolved with the input images to extract feature information such as edges and textures in the images. At the same time, the mechanism of weight sharing is adopted, which computes much less than the fully connected layer. The working principle of the convolutional layer is shown in Fig. 7

Activation functions are adhered after convolutional layers or fully connected layers to handle nonlinear problems. Since deep neural
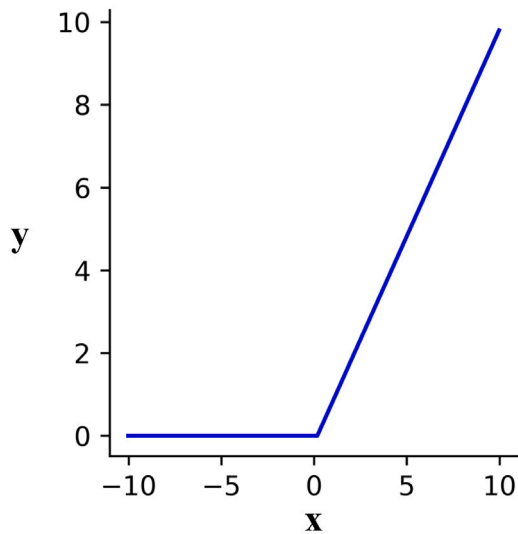
Fig. 8. Activation function: ReLu.
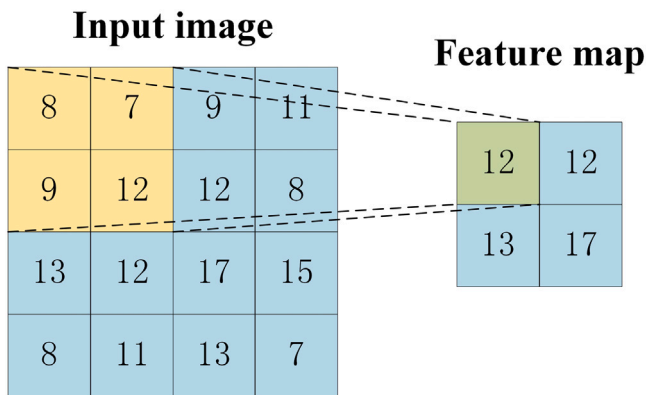
Input image   Feature map

Fig. 9. Maximum pooling.

networks require that the interior of the network is everywhere, it demands activation functions that are differentiable, such as tanh, ReLu, and sigmoid. ReLu is used in this paper (see Fig. 8).

The pooling operation is essentially a downsampling technique to reduce the dimensionality of the output feature vector of the convolutional layer, reducing the possibility of overfitting and improving the robustness of the model. Typical operations include average pooling and maximum pooling. The principle of maximum pooling used in the proposed algorithm is shown in Fig. 9.

Through a series stacked combinations of convolutional layers and pooling layers, the features of the input data are obtained. At the end of CNN, one or more fully connected layers are added to enable the network achieve high-order reasoning ability and act as a "classifier". The last fully connected layer is always activated by softmax in the classification task, and the loss function is generally based on cross entropy.

The proposed algorithm, named SVI-CNN, consists of a series of alternately stacked convolutional and pooling layers, and a number of fully connected layers. The input and output of SVI-CNN are SVIs and fault types separately corresponding to the SVIs. The framework of SVI-CNN is presented in Fig. 10.

The procedure for fault diagnosis using SVI-CNN is summarized as follows. First, the vibration signals are obtained under different fault states during the operation of AMB-R, and are cut into a series of samples, which are further divided into a training set and a testing set. Second, the SVI is calculated according to the signals in each sample.

Third, SVIs and the fault category of each samples in the training are set as input and output of the CNN respectively to train the network parameters of the CNN. Finally, the SVI of the testing samples are input to the trained CNN and fault categories are determined.

As an end-to-end learning mechanism, SVI-CNN effectively mines the fault feature information in SVIs. However, for AMB-R, there are five sensors, and 10 SVIs can be obtained at the same time. Consequently, the information of these 10 SVIs can be integrated to fully reflect the fault information; however, this situation create an impossible scenario for SVI-CNN to address. Therefore, it is necessary to design a new network structure and merge the information of SVIs composed of more than 2 sensors.

3.2. MCNN

As only two vibration sensors are installed in the rotating machinery, one SVI can be constructed and the diagnosis result is gained by SVI-CNN. However, when the system is equipped with N sensors satisfied $N > 2$, more SVIs are generated so that SVI-CNN cannot handle this circumstance. For AMB-R, $N = 5$. Accordingly, a multi-branch CNN is designed to merge multiple SVIs. Unlike the conventional CNN, MCNN consists of a multi-branch part and a combination part.

The multi-branch part of MCNN comprises $N(N - 1)/2$ branches. Every branch is independent of each other and is made up of several convolution layers and pooling layers alternately. The input of each branch is one SVI $\mathbf{s}_i, i = 1, 2, \ldots, N(N - 1)/2$, and there are

$$\mathbf{z}_i = F_{B_i}(\mathbf{s}_i) \tag{7}$$

where F_{B_i} and $\mathbf{z}_i$, $i = 1, 2, \ldots, N(N - 1)/2$ represent the ith branch network and its output respectively.

The combination part of MCNN is composed of a series of fully connected layers, and the input of the first fully connected layer is

$$\mathbf{z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{N(N-1)/2}] \tag{8}$$

where $\mathbf{z}$ is concatenated by the output of multiple branches. The fault type c corresponding to the input sample is acquired as Eq. (9):

$$c = F_C(\mathbf{z}) \tag{9}$$

where F_C denotes the combination part of MCNN.

The framework of SVI-MCNN is depicted in Fig. 11 and the procedure of SVI-MCNN is as follows:

Step 1: Capture samples containing N channels of vibration signals from sensors installed at diverse positions of the rotating machinery system.

Step 2: Divide the samples into two parts, one part for training and the other for testing.

Step 3: In the training set, combine the signals in pairs to construct $N(N - 1)/2$ SVIs for each sample.

Step 4: Train the MCNN model with $N(N - 1)/2$ branches.

Step 5: Construct $N(N - 1)/2$ SVIs for each sample in the testing set and input these SVIs into the trained MCNN model to obtain the fault category.

Compared with CNN, MCNN is employed to situations when the number of sensors is larger than 2. When $N = 2$, SVI-MCNN degenerates into SVI-CNN, thus SVI-CNN is a special case of SVI-MCNN.

3.3. Discussion

Compared to other fusion algorithms, there exist three advantages of SVI-MCNN.

First, there exist relations as follows:

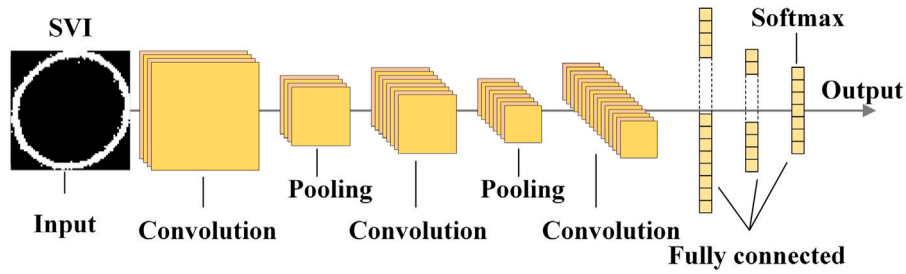$$S_{input} = \frac{N(N - 1)}{2} \cdot L^2 \tag{10}$$
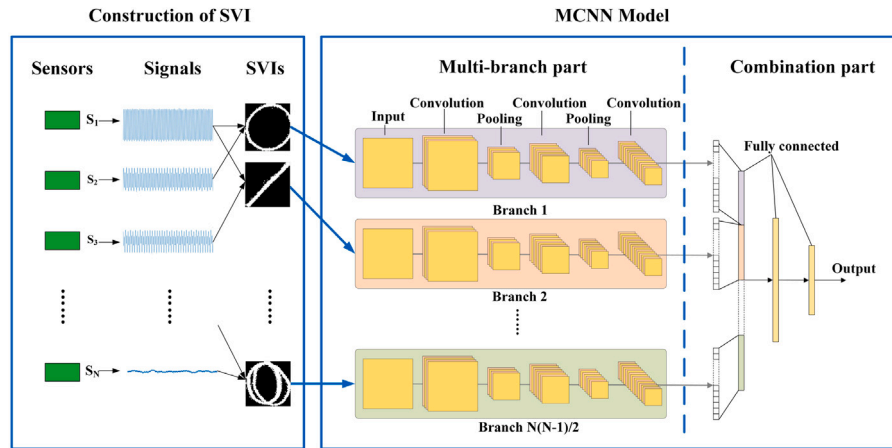
Fig. 10. The framework of SVI-CNN.

Fig. 11. The framework of SVI-MCNN.

where S_{input} is the size of the input parameters of MCNN. Set the length of samples as S_L which should satisfy:

$$S_{input} < S_L \cdot N \tag{11}$$

where $S_L \cdot N$ characterizes the number of points in one raw sample containing multi-sensor data. Eq. (11) indicates that S_{input} can be smaller than the dimension of the raw data by selecting proper N and L, which may bring the benefits of less training time, high recognition accuracy and robust generalization ability.

Second, the multi-branch structure of MCNN is a secondary fusion of sensor information while SVI is the first fusion of two sensor information. Each branch of MCNN extracts the high-level semantics of SVI through the stacking of multiple convolution and pooling layers. The combination part of MCNN concatenates the high-level semantics of multiple SVIs and then further merges the information of all sensors through a series of fully connected layers. The fault representation obtained includes richer information and are more discriminative .

Third, compared with the traditional algorithms based on shaft orbit, SVI-MCNN is an end-to-end learning framework that does not require the manual extraction of features based on SVIs. Through the stacked layers of CNN, the algorithm can automatically obtain more discriminative features than those that are manually extracted.

4. Experiments

To verify the performance of the proposed algorithm, a series of experiments are carried on the AMB-R1 dataset, which contains samples collected from a vertical rig with the rotor supported by AMBs (see Fig. 12).

The dataset consists of four types of samples, including normal, unbalanced, misalignment and rub-impact, which are all common health states in the AMB-R. Each sample contains five channels of signals and each signal consists of 15,000 points under a sampling rate of 25 kHz. The number of each type of samples is presented in Table 1. The

Table 1
AMB-R1 dataset settings.

Fault type	Normal	Unbalance	Misalignment	Rub-impact
Training number	150	150	150	150
Testing number	260	358	172	55

Table 2
CNN structure.

Layer name	Model
L1	Conv(40*3*3), Activation(Relu)
L2	Maxpool(2*2)
L3	Conv(60*3*3), Activation(Relu)
L4	Maxpool(2*2)
L5	Conv(100*3*3), Activation(Relu)
L6	Maxpool(2*2)
L7	Fulllayer(36), Activation(Relu)
L8	Fulllayer(4), Activation(Softmax)

samples are randomly divided into training samples and test samples for each trial, and all experimental results are the average of 20 trials.

To facilitate the experimental comparison, we specify the network structure of the CNN and MCNN as listed in Tables 2 and 3, where the batch size is set to 50, and the cross-entropy is selected as the loss function of the neural networks.

4.1. Comparison with other multi-sensor fusion algorithms

We have implemented four multi-sensor fusion algorithms in other literatures for comparison. Table 2 provides the results of these algorithms and SVI-MCNN outperforms the others.

Both CNN-A and CNN-B use CNN's end-to-end learning mechanism, the structure of CNN are consistent with that shown in Table 4, and the size of the input parameter scale is 75 000, which is equal to the
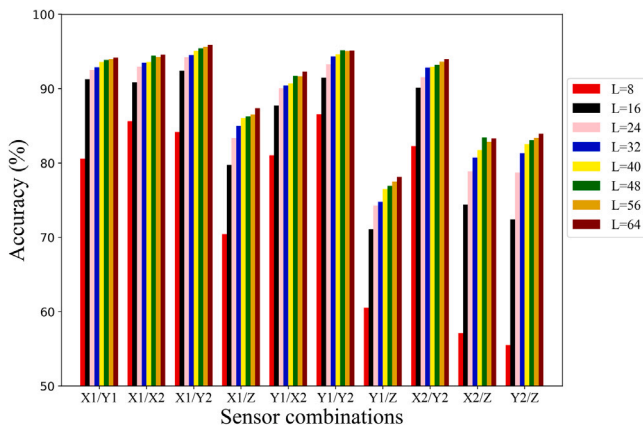
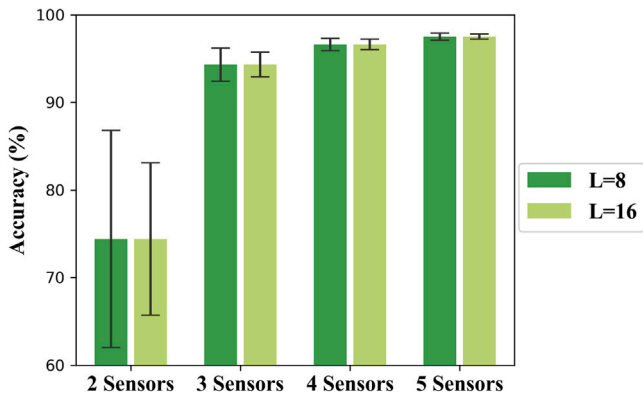Fig. 12. Test rig.

Fig. 13. Performance under a series of L.

Fig. 14. Effect of the number of multiple sensors on algorithm performance.

Table 3
MCNN structure.

Layer name	Branches	Model
L1	N(N-1)/2	Conv(40*3*3), Activation(Relu)
L2	N(N-1)/2	Maxpool(2*2)
L3	N(N-1)/2	Conv(60*3*3), Activation(Relu)
L4	N(N-1)/2	Maxpool(2*2)
L5	N(N-1)/2	Conv(100*3*3), Activation(Relu)
L6	N(N-1)/2	Maxpool(2*2)
L7	1	Fulllayer(100), Activation(Relu)
L8	1	Fulllayer(36), Activation(Relu)
L8	1	Fulllayer(4), Activation(Softmax)

Table 4
Comparison of SVI-MCNN and other fusion algorithms.

Algorithms	Accuracy
SVI+MCNN	**98.6%**
CNN-A [18]	92.9%
CNN-B [20]	96.3%
Orbit-A [17]	97.3%
Orbit-B [26]	93.6%

number of sampling points in raw sample. Yet, the input size of the SVI-MCNN is only 25 601 when the length $L = 16$ and $N = 5$, which is only 3.41% of the raw data points. Of course, the comparison of deep learning algorithms depends on multiple factors including initial values and parameter commission, and the comparison of the algorithms with different network structures is not rigorous. However, it shows that the SVI-MCNN can at least achieve a similar performance to the algorithms with the same number of layers. Furthermore, the parameter size is much smaller, which inevitably leads to a smaller probability of overfitting and a shorter training time.

The algorithms Orbit-A and Orbit-B also achieve good performance; however, they both need to manually design and extract features based on the shaft orbit, while the SVI-MCNN can automatically learn discriminative features from GSO without denoising, and incorporate the feature extraction and decision fusion process into one framework.

4.2. The influence of parameter L

Fig. 13 shows the results of SVI-CNN under various two-sensor combinations and different L values. When the sensor combination is fixed, the recognition accuracy gradually increases as the value of L increases. When L increases further, the recognition accuracy increases slightly.

From the above phenomenon, it can be concluded that the increase of L leads to a larger resolution of the SVI and the more information extracted from the samples. However, as L increases to a certain extent, the recognition accuracy is not further improved and it may cause overfitting. Meanwhile, the input parameter size L^2 grows fast to boost the training time. Therefore, choosing a proper L can ensure the performance, avoiding overfitting and long training time.

4.3. The influence of the number of sensors N

From Fig. 13, it can also be concluded that different sensor combinations have various abilities to discriminate faults. The X1/Y2 combination has the strongest discriminative power and the highest recognition accuracy exceeding 95%. The performance of the Z-related sensor combinations is not higher than 90%, even lower than 80%. Hence, the performance of different sensors can be complemented and better results can be achieved through the fusion algorithm.

To verify the influence of the number of sensors N on the proposed algorithm, the SVI-MCNN is tested under variable N. There are C_5^N kinds of sensor groups and the results are averaged and presented in Fig. 14. As N increases, the number of branches increases, and the
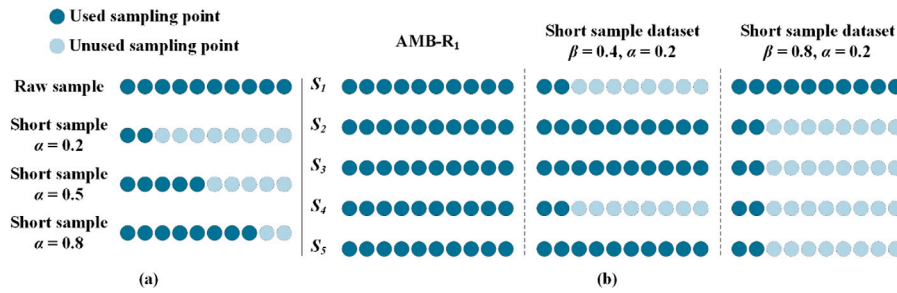
Fig. 15. The illustration of parameters α and β.

amount of information contained in $N(N-1)/2$ SVIs as the input of the MCNN also increases. In particular, from $N = 2$ to $N = 3$, the accuracy increased from 74.4% to 94.3% with an obvious improvement, which shows that the SVI-MCNN achieves a deeper information integration of multiple sensors. When N ranges from 3 to 5, the performance of the algorithm is improved slightly, reaching 98.6%, and the standard deviation of the results is further reduced under diverse sensor groups and iterations. Hence, as N increases, the provided information becomes richer, and the SVI-MCNN can effectively mine and fuse more information, leading to better results.

Hence, we can conclude that different sensor combinations have obvious differences in diagnosing faults. As more sensor combinations are added into SVI-MCNN, more sensor information are fused and the risk brought by some sensor combinations with poor performance is reduced.

4.4. The influence of samples with variable lengths

From Fig. 6, it can be concluded that when the sample length is greater than one cycle, the shape of SVI does not change substantially with the length of the sample. To verify the above characteristic of the SVI, some samples in AMB-R1 are instead short samples that are cut from the original samples. Two parameters are set as α and β that both range from 0 to 1. α indicates the length of the short sample relative to the original sample (as shown in Fig. 15(a)). When $\alpha = 1$, the sample length is consistent with the original sample length. In the experiment, α is set to 0.2, 0.5, and 0.8, satisfying that the length of short samples is greater than 1 cycle. β denotes the proportion of the short samples to the whole dataset of samples (as shown in Fig. 15(b)). When β is 0, all samples in the dataset are original samples. When β is equals to 1, the transformed dataset is full of short samples. The original and short samples are randomly distributed in the training and testing samples. The SVI-CNN is utilized after setting $L = 40$ and selecting two sensor combinations X1/Y2 and X2/Z to construct SVIs.

From Fig. 16, the following conclusions can be drawn:

(1) Overall, when short samples are mixed into the dataset, although the performance of the algorithm is reduced, the proposed algorithm still achieves good results.

(2) Essentially, α indicates the amount of information contained in the short sample. Under the same conditions, as α decreases, the recognition rate of the algorithm decreases.

(3) Generally, when the dataset is full of original samples, the performance of the algorithm is the best. As β increases, the performance of the algorithm decreases, and while β is close to 1, the performance of the algorithm improves again. Therefore, the better the sample length consistency in the dataset is, the better the performance of the algorithm becomes.
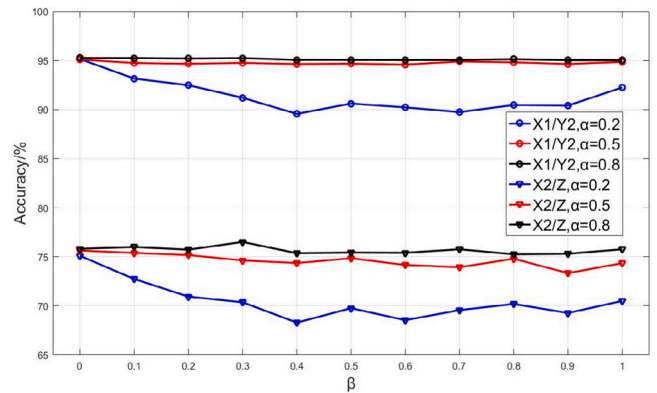

Fig. 16. Performance of the proposed algorithm under the condition of the dataset with variable-length samples.

4.5. The influence of samples with variable sampling rates

In practice, the sampling rate of the samples is often required to be fixed. To evaluate the influence of samples with variable sampling rates, AMB-R1 is converted into three different datasets named mode 1, mode 2 and mode 3. Each sample in the mode 1 dataset takes the first 5000 points of each original sample in AMB-R1. In mode 2, the sampling rate is reduced to 1/3 of AMB-R1. Mode 3 dataset is mixed with half of mode 1 and half of mode 2. The organization of the three datasets is illustrated in Fig. 17. Mode 1 and mode 2 both consist of samples at a constant sampling rate, while mode 3 is a mixture dataset consisting of samples with various sampling rates.

Two sensor combinations, X1/Y2 and X2/Z, are selected for testing, and the results of SVI-CNN are shown in Fig. 18. In most cases, the recognition accuracy of mode 3 is higher than that of mode 1 and slightly lower than that of mode 2, which suggests that the proposed algorithm can be applied to the datasets with different sampling rates.

According to the principle of FFT, the frequency analysis results of signals depend on the sampling rate and sample length. In Fig. 19, (a) and (b) show two signals with the same sampling rate but different rotating speeds. The signal in (b) is cut and stretched to form the signal in (c) which has the same period as the signal in (a), while the signal in (c) keeps the same amplitude and phase characteristics as the signal in (b). The signals in (a) and (c) have approximate frequency analysis results including the main frequency component, i.e., rotating speed. Therefore, the problem at different speeds can be converted into the problem at various sampling rates. The previous experiment indirectly proves that the algorithm can be applied to datasets with samples at different rotating speeds, which is consistent with what Fig. 5 shows. Of course, these experiments are not completely equivalent to the test at different speeds, but it provides a new idea to study the characteristics of the algorithms under various speeds.
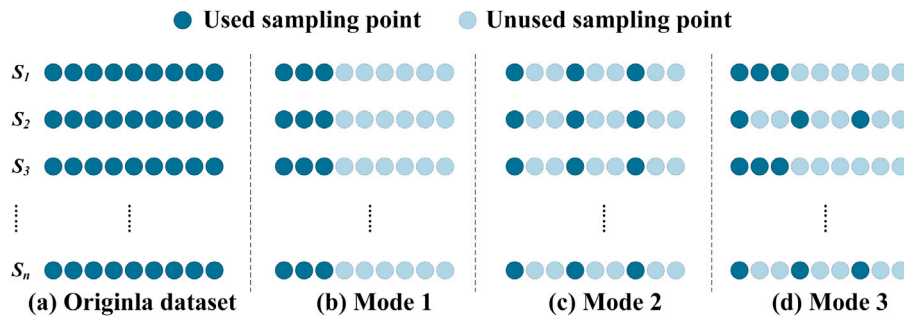
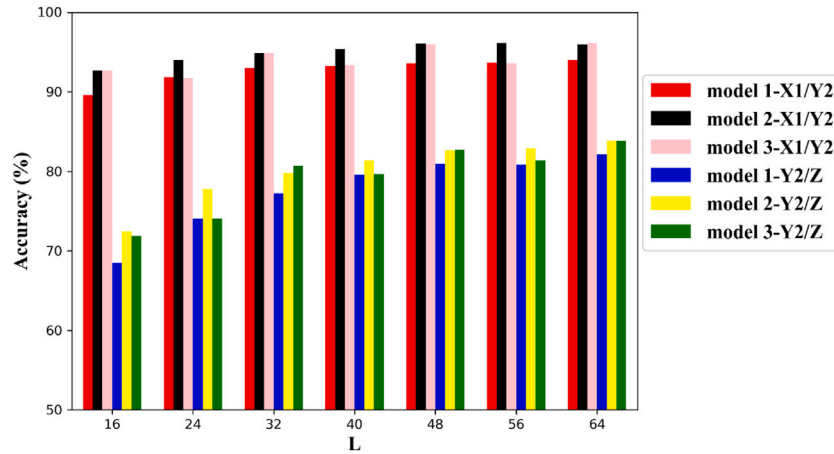**Fig. 17.** The organization of mode 1, mode 2 and mode 3 datasets.

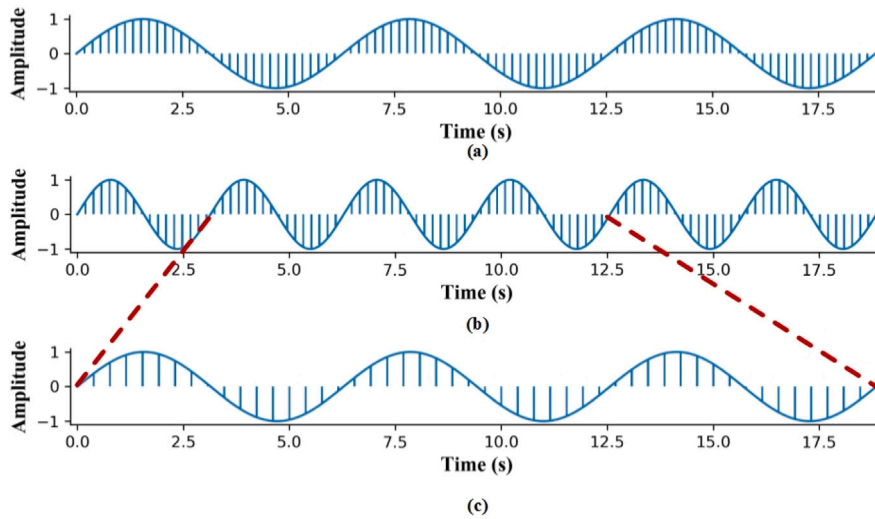**Fig. 18.** Experimental results under various sampling rates.

Fig. 19. Illustration of the relationship between the sampling rate and rotating speed. The sampling rate is set as fs, and rotation speed as ω (a) fs = k, ω = v (b) fs = k, ω = 2v (c) fs = k/2, ω = v.

4.6. Discussion

Through the above experiments, the effectiveness and efficiency of the proposed algorithm is verified, and the following conclusions can be drawn.

First, compared to other multi-sensor fusion algorithms, SVI-MCNN can better integrate multi-sensor information and achieve a higher accuracy and a smaller network parameter size.

Second, SVI-MCNN mainly includes two parameters: the resolution of the SVI L and the number of sensors N. By selecting the proper L

and N, it can obtain a high accuracy, a short training time and a robust generalization ability.

Third, SVI-MCNN is suitable for samples with different sampling rates, variable speeds and changeable lengths, which provides the possibility to make full use of operating data and provide sufficient samples for training better models.

In summary, SVI-MCNN yields excellent performance and can handle under variable operating conditions, which is an effective tool for fault diagnosis of AMB-R.

5. Conclusions

To address the issue that are encountered in AMB-R, such as multiple sources of faults, the obvious non-linearity and non-stationarity of vibration signals and the resulting difficult representation of fault characteristics, this paper proposes an end-to-end learning algorithm based on GSO for fault diagnosis of AMB-R. According to the results obtained, we can conclude our paper as follows:

(1) The proposed algorithm fully incorporates the characteristics of AMB-R without using additional sensors, and uses the information captured by the five displacement sensors in AMB. The concept of traditional shaft orbit is expanded to GSO, which is discretized into SVI as the data representation of multi-sensor signals;

(2) With the designed MCNN implemented as an end-to-end deep learning method, there is no need to manually extract features and perform denoising of the SVI. Through multiple branches of stacked layers and combination layers, the features are uncovered deeply and multi-sensor information is fused from both the data level and decision level. Compared with that of algorithms that use raw data as input, the parameter size of the SVI-MCNN are much smaller, leading to resistance to overfitting and shortening the training time.

(3) Due to the characteristics of SVI and the good generalization of MCNN, the proposed algorithm is applicable to a variety of working conditions. Through experimental verification, it has a good performance at different speeds, sampling rates, and sample lengths.

In the future, we will further study characteristic of the GSO and combine it with the deep learning algorithms to apply in the early fault detection of AMB-R.

CRediT authorship contribution statement

Yan Xunshi: Conceptualization, Methodology, Software, Writing - original draft. **Zhang Chen-an:** Writing - review & editing, Resources. **Liu Yang:** Visualization, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] G. Yang, Z. Shi, N. Mo, Technical design and engineering prototype experiment of active magnetic bearing for helium blower of HTR-PM, Ann. Nucl. Energy 71 (9) (2014) 103–110.

[2] J. Zhao, Z. Sun, X. Yan, G. Yang, Y. Zhou, X. Liu, Z. Shi, T. Fan, X. Zhang, Helium blower test based on aerodynamic force simulation, Ann. Nucl. Energy 118 (2018) 283–290.

[3] G. Schweitzer, E.H. Maslen, Magnetic Bearings: Theory, Design, and Application to Rotating Machinery, Springer, Berlin Heidelberg, 2009.

[4] International Organization for Standardization (ISO), ISO14839 - mechanical vibration - vibration of rotating machinery equipped with active magnetic bearings, 2003.

[5] M. Zhao, M. Kang, B. Tang, M. Pecht, Multiple wavelet coefficients fusion in deep residual networks for fault diagnosis, IEEE Trans. Ind. Electron. 66 (6) (2019) 4696–4706.

[6] L. Ou, D. Yu, Fault diagnosis of roller bearings based on laplacian energy feature extraction of path graphs, Measurement 91 (2016) 168–176.

[7] D. Zhao, T. Wang, F. Chu, Deep convolutional neural network based planet bearing fault classification, Comput. Ind. 107 (2019) 59–66.

[8] Y. Lei, Z. He, Y. Zi, A new approach to intelligent fault diagnosis of rotating machinery, Expert Syst. Appl. 35 (4) (2008) 1593–1600.

[9] H. Shao, H. Jiang, Y. Lin, X. Li, A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders, Mech. Syst. Signal Process. 102 (Supplement C) (2018) 278–297.

[10] L. Wen, X. Li, L. Gao, Y. Zhang, A new convolutional neural network-based data-driven fault diagnosis method, IEEE Trans. Ind. Electron. 65 (7) (2018) 5990–5998.

[11] G. Jiang, H. He, J. Yan, P. Xie, Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox, IEEE Trans. Ind. Electron. 66 (4) (2019) 3196–3207.

[12] J. Wang, J. Yan, C. Li, R.X. Gao, R. Zhao, Deep heterogeneous gru model for predictive analytics in smart manufacturing: Application to tool wear prediction, Comput. Ind. 111 (2019) 1–14.

[13] Z.-X. Hu, Y. Wang, M.-F. Ge, J. Liu, Data-driven fault diagnosis method based on compressed sensing and improved multiscale network, IEEE Trans. Ind. Electron. 67 (4) (2020) 3216–3225.

[14] D. Zhang, E. Stewart, M. Entezami, C. Roberts, D. Yu, Intelligent acoustic-based fault diagnosis of roller bearings using a deep graph convolutional network, Measurement 156 (2020) 107585.

[15] Z. Duan, T. Wu, S. Guo, T. Shao, R. Malekian, Z. Li, Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: a review, Int. J. Adv. Manuf. Technol. 96 (1) (2018) 803–819.

[16] M. Sadoughi, C. Hu, Physics-based convolutional neural network for fault diagnosis of rolling element bearings, IEEE Sens. J. 19 (11) (2019) 4181–4192.

[17] X. Yan, Z. Sun, J. Zhao, Z. Shi, C.-a. Zhang, Fault diagnosis of rotating machinery equipped with multiple sensors using space–time fragments, J. Sound Vib. 456 (2019) 49–64.

[18] H. Wang, S. Li, L. Song, L. Cui, A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals, Comput. Ind. 105 (2019) 182–190.

[19] T. Han, C. Liu, L. Wu, S. Sarkar, D. Jiang, An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems, Mech. Syst. Signal Process. 117 (2019) 170–187.

[20] M. Xia, T. Li, L. Xu, L. Liu, C.W. De Silva, Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks, IEEE/ASME Trans. Mechatronics 23 (1) (2018) 101–110.

[21] L. Jing, T. Wang, M. Zhao, P. Wang, An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox, Sensors 17 (2) (2017) 414.

[22] J. Liu, Y. Hu, Y. Wang, B. Wu, J. Fan, Z. Hu, An integrated multi-sensor fusion-based deep feature learning approach for rotating machinery diagnosis, Meas. Sci. Technol. 29 (5) (2018) 055103.

[23] R. Zhao, D.Z. Wang, R.Q. Yan, K.Z. Mao, F. Shen, J.J. Wang, Machine health monitoring using local feature-based gated recurrent unit networks, IEEE Trans. Ind. Electron. 65 (2) (2018) 1539–1548.

[24] J. Wang, P. Fu, L. Zhang, R.X. Gao, R. Zhao, Multilevel information fusion for induction motor fault diagnosis, IEEE/ASME Trans. Mechatronics 24 (5) (2019) 2139–2150.

[25] D. Shi, W. Wang, P. Unsworth, L. Qu, Purification and feature extraction of shaft orbits for diagnosing large rotating machinery, J. Sound Vib. 279 (3–5) (2005) 581–600.

[26] X. Yan, Z. Sun, J. Zhao, Z. Shi, C.-A. Zhang, Fault diagnosis of active magnetic bearing-rotor system via vibration images, Sensors 19 (2) (2019) 244.

[27] G. Wen, Z. Chen, Z. Zhang, Rotor state diagnosis based on fuzzy c-mean value clustering and its axial center orbit features, Zhendong yu Chongji/J. Vib. Shock 38 (15) (2019) 27–35.

[28] B. Wu, S. Feng, G. Sun, L. Xu, C. Ai, Fine-grained fault recognition method for shaft orbit of rotary machine based on convolutional neural network, J. Vibroeng. 21 (8) (2019) 2106–2120.

[29] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.

[30] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: 26th Annual Conference on Neural Information Processing Systems, Vol. 2, 2012, pp. 1097–1105.