

Research Article

Information Fusion of Infrared Images and Vibration Signals for Coupling Fault Diagnosis of Rotating Machinery

Tangbo Bai ^{1,2}, Jianwei Yang ^{1,2}, Dechen Yao ^{1,2} and Ying Wang ³

¹School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

²Beijing Key Laboratory of Performance Guarantee on Urban Rail Transit Vehicles, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

³Beijing Institute of Aerospace Control Device, Beijing 100094, China

Correspondence should be addressed to Tangbo Bai; baitangbo@bucea.edu.cn

Received 3 January 2021; Revised 15 March 2021; Accepted 26 March 2021; Published 5 April 2021

Academic Editor: Huaitao Shi

Copyright © 2021 Tangbo Bai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rotating machinery has a complicated structure and interaction of multiple components, which usually results in coupling faults with complex dynamic characteristics. Fault diagnosis methods based on vibration signals have been widely used, however, these methods are intricate when identifying coupling faults, especially in the situation where coupling faults share similar patterns. As a noncontact and nonintrusive temperature-measuring technique, methods by infrared images can recognize multiple faults through temperature variations; however, it is not effective if the faults are temperature-insensitive. In this paper, an improved machinery fault diagnosis technique based on information fusion of infrared images and vibration signals is studied, to have better utilization of multisource sensors and to solve the problems when one single type of data is separately used. Firstly, data enhancement for infrared images and data visualization for vibration are performed on the dataset by using the principle of graphics and Short-Term Fourier Transform, which increases the diversity of the dataset and enhances the generalization ability of the model. Then, a multichannel convolution neural network-based method is constructed to achieve data-level information fusion and improve the fault diagnosis accuracy. The effectiveness of the presented method is validated by the experimental studies on a rotor test stand, the results illustrate that the coupling faults can be effectively identified by the information fusion method, and the fault diagnosis accuracy is higher in comparison with the method by a signal from single-source sensors.

1. Introduction

As key equipment, rotating machinery has been widely used in modern industry. However, once accidental failure occurs, it often causes serious consequences. So, the diagnosis and prognosis of mechanical equipment play an important role in avoiding accidents. Vibration analysis has been widely investigated in current equipment condition monitoring [1, 2]. But when multiple failures occur simultaneously, vibration signals sometimes produce frequency aliasing, so it is difficult to extract the characteristic information of different faults in this case [3, 4]. To diagnose coupling faults of machinery, Gu et al. proposed a novel method by using multivariate empirical mode decomposition (MEMD) and band energy for fault diagnosis by four

channels of vibration signals in diesel engine diagnosis [5]. Zhou et al. presented a method based on unscented Kalman filter (UKF) and radial basis function for multiple faults of a pumping unit [6]. However, sensor installation location affects diagnosis accuracy, and it is difficult to determine the exact number of the sensors and the accurate location for installation to obtain an optimal diagnosis performance. Moreover, because the sensors cannot be installed directly on the rotating components, vibration signals cannot be measured directly on the defective part and consequently lead to low accuracy.

As a noncontact and nonintrusive temperature-measuring technique, infrared thermography is able to monitor all components of the whole equipment by intuitively observing temperature variations of abnormal parts. It has

wide applications in machinery defect diagnosis [7–9], building monitoring [10], medical inspection [11–13], and nondestructive testing [14–16]. Younus et al. proposed a two-dimensional discrete wavelet transformation-based diagnosis system, to diagnose machinery using infrared images [17]. To detect motor rotors, Eftekhari et al. presented a method using infrared images' features taken from the hottest region of the motor surface [18]. Olivier et al. proposed a novel automatic fault detection system using infrared images, to detect the bearing system by differencing the consecutive image frames which are subsequently summarized by their distribution along the image axes [19]. However, the temperature changes caused by some weak faults or early faults are small, so it is difficult to diagnose mechanical equipment fault effectively only through infrared images.

Through the above analysis, it can be seen that vibration signals and infrared images are complementary in fault diagnosis of mechanical equipment. So, the multisensor information fusion method is proposed in this paper, to realize rotor fault diagnosis using these two kinds of signals. In general, multiple sensors can provide more information and make equipment monitoring more accurate [20, 21]; thus, various sensors are used in the monitoring of large rotating equipment [22, 23]. To process data from multiple sensors, there are mainly three types of information fusion methods, including the data-level, the feature-level, and the decision-level fusion methods. In the application of feature-level fusion, it needs to extract the features from the data of each sensor [21, 23]; then, the feature information of all sensors should be integrated into a feature vector by certain methods and optimize the parameters [24–26], and the combine features are trained and recognized by pattern recognize method such as support vector machine (SVM) [27], artificial neural networks (ANN) [28], convolution neural network (CNN) [29], and long short-term memory (LSTM) [30]. In decision-level fusion, data from each sensor is analyzed independently, the corresponding results are obtained to represent the state of the equipment, and then the final decision is made by using all results acquired from each sensor with certain methods like Dempster–Shafer (D-S) evidence theory [31, 32] and fuzzy decision theory [33, 34]. Although the above two kinds of methods can realize information fusion with multiple sensors, there are often evidence conflicts between results of different sensors when using decision-level fusion, which lacks decision generality. The feature-level fusion is more suitable for sensors of the same type, but when it is used for different types of sensors, the fault information is easy to get lost due to the different distribution characteristics of the feature information. Data-level fusion does not need to extract features and make decisions from the data acquired by sensors but directly fuses the data, thus avoiding the above problems. However, similar data types are required in the traditional data-level fusion. Vibration signal is one-dimensional waveform data, and the infrared image is two-dimensional image data, so it is difficult to directly realize data-level fusion.

To solve these problems, a data-level information fusion method by fusing infrared images and vibration signals is investigated in this paper, to take advantage of these two types of sensors and acquire a better performance in rotating machinery fault diagnosis. In order to achieve data-level fusion, the vibration signals are visualized by the principle of graphics and time-frequency analysis method, and the infrared images are enhanced by operation in graphics to expand the dataset and improve the generalization of the model. Then a multichannel convolution neural network-based model is proposed, using the two-dimensional Time-Frequency Graph of vibration signals and enhanced infrared images as a multi-input dataset, to construct data-level information fusion. Finally, experiment tests on a rotor test stand are performed. The results show that the proposed method can effectively identify the coupling faults, and the classification accuracy is higher than the methods based on the vibration signals or infrared images only.

The intellectual merits of this study rest on three folds. (1) Data enhancement is performed on the infrared images dataset by using the principle of graphics, which improves the size of the dataset and generalizes the adaptability of the model to various infrared images acquisition situations. (2) A multichannel deep learning model is built by changing the input layer of the CNN network, which can process multiple input data simultaneously. (3) Data-level information fusion is introduced to fault diagnosis, taking the advantages of infrared images and vibration signals to improve the accuracy. The rest of the paper is organized as follows: the theoretical background of CNN and Short-Term Fourier Transform are first discussed in Section 2. Section 3 is the theoretical framework of fault diagnosis based on the fusion of infrared images and vibration signals. In Section 4, the experimental studies on a rotor test stand are performed to demonstrate the effectiveness of the information fusion approach in the paper. Finally, the concluding remarks are drawn in Section 5.

2. Theoretical Background

2.1. Convolution Neural Network. To realize the data-level fusion without feature extraction, vibration signals and infrared images need to be processed directly. Convolution neural network, as a deep learning algorithm, is widely used in the field of image processing and mechanical equipment fault diagnosis [35]. It can directly process with matrix and image information. In addition, it also has the advantages of strong adaptability, easy implementation, and less training parameters [36]. Therefore, CNN is chosen to process vibration signals and infrared images in this paper.

Compared with the traditional neural network, CNN constructs a deeper network structure in order to simulate the biological neural network more accurately. At the same time, CNN introduces the method of local connections and weight sharing, which avoids the problems in the traditional network such as net complexity, too many nodes and parameters, slow convergence speed, and difficult computation [37]. Local connection means that the neuron

nodes of a layer in the neural network are connected with part of adjacent neurons of the adjacent upper and lower layers by certain rules, instead of all the neurons, which greatly reduces the number of neuron nodes and simplifies the structure of the neural network. Weight sharing indicates that the parameters of the neurons are identical in CNN when different neurons are connected from upper and lower layers. Thus, it reduces the number of network parameters and improves the generalization ability of the network.

The basic CNN is usually composed of the input layer, multiple convolution layer, pooling layer, fully connected layer, and output layer, and the most important structure is convolution, activation, and pooling. The output of CNN is the specific feature space of each image. When dealing with the image classification task, the feature space output by CNN is taken as the input of the fully connected layer or fully connected neural network (FCN), and the fully connected layer is used to complete the mapping from the input image to the label set.

The concept of channels is important in the CNN input layer, which represents the channel composition of the input image. For a typical RGB image, the number of channels is 3 (red, green, and blue), while the number of channels for the monochrome image is 1. The structure of CNN with 3 input channels for RGB image is shown in Figure 1.

2.2. Short-Term Fourier Transform. Vibration signal is one-dimensional time-domain waveform data. If the signal is processed directly without feature extraction, the frequency domain information will be lost. Short-Term Fourier Transform (STFT) is a time-frequency domain analysis method. Its basic idea is to decompose a time-varying signal into several short-term stationary ones by a window function, so as to obtain locally stationary signals. Therefore, it is suitable for nonlinear and nonstationary signal processing, and it is widely used in vibration, acoustic, and other time-varying signal analysis. The calculation of STFT is shown in equation (1), after the signal is divided by the window function, Fourier transform is carried out to obtain the local spectrum within a range of short time t .

$$\text{STFT}_x(t, f) = \int_{-\infty}^{\infty} [x(t')g(t' - t)]e^{-j2\pi ft'} dt', \quad (1)$$

where t is the time period of the window function, f is the angular frequency, $x(t')$ is the signal to be analyzed, and $g(t' - t)$ is window function.

According to the equation of STFT, the most important factor affecting the result is the choice of window function and its related parameters. The type of window functions has a decisive influence on spectral leakage and interspectral interference. The width of the window function affects the resolution in the time domain and frequency domain. Meanwhile, the data length of signals determines the resolution in the time domain. The time-domain and frequency-domain resolutions are calculated by the following equations:

$$R_t = \left\lfloor \frac{N_x - N_o}{N_w - N_o} \right\rfloor, \quad (2)$$

$$R_f = \begin{cases} \frac{N_f}{2} + 1, & N_f \% 2 = 0, \\ \frac{N_f + 1}{2}, & N_f \% 2 \neq 0, \end{cases} \quad (3)$$

where R_t stands for time-domain resolution and R_f indicates frequency-domain resolution. N_x is the length of the signal to be processed, N_o is the overlap width of the window, N_w is the width of the window, N_f is the length of the short term in the signal to process.

3. The Proposed Methodology

3.1. The Proposed Approach for Coupling Fault Diagnosis. To realize fault diagnosis based on information fusion, infrared images and vibration signals are processed, respectively, to obtain the input dataset of the deep learning network. Then, the processed infrared images and vibration time-frequency image are taken as a multi-input dataset, to construct data-level information fusion. Additionally, a novel multichannel CNN model is built to realize the fusion and to obtain diagnosis results.

The proposed system for machinery fault diagnosis using infrared images and vibration signals is shown in Figure 2. This system consists of several modules: data acquisition, data preprocessing, dataset building, and information fusion. Data enhancement for infrared images and data visualization for vibration signals is performed first on the training set by using the principle of graphics and Short-Term Fourier Transform, which improves the size of the dataset and enhances the generalization ability of the model. Then, a multichannel convolution neural network-based classifier fusion is constructed for fault diagnosis, taking the advantages of infrared images and vibration signals to improve the accuracy.

3.2. Data Preprocessing. In the process of infrared images acquisition, when the distance and angle between the infrared camera and the object to be measured vary, it results in the changes of the infrared images states, and the acquisition environmental changes are also likely to cause the change of infrared images contrast. However, the experimental data acquisition is difficult to cover all conditions. Therefore, if the model is trained by infrared images of one single acquisition condition, the adaptability and diagnostic accuracy of the model under different acquisition conditions are limited.

To solve the above problems, this paper proposes to utilize the rotating, scaling, flipping, brightness adjustment, and mixed operation in graphics to enhance the data, expand the dataset, and improve the generalization of the model, as shown in Figure 3.

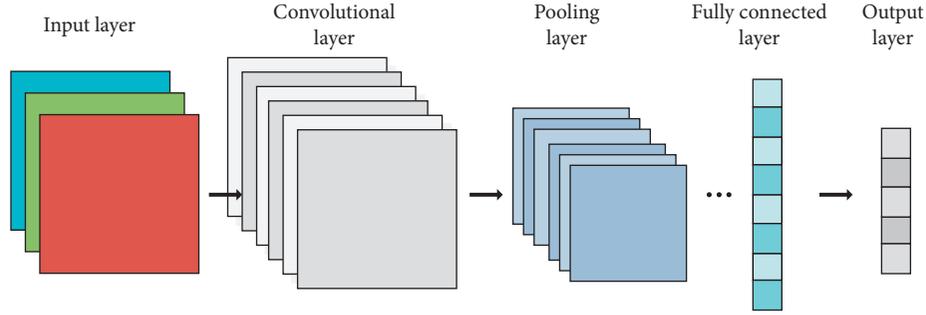


FIGURE 1: Structure of convolutional neural network for RGB image.

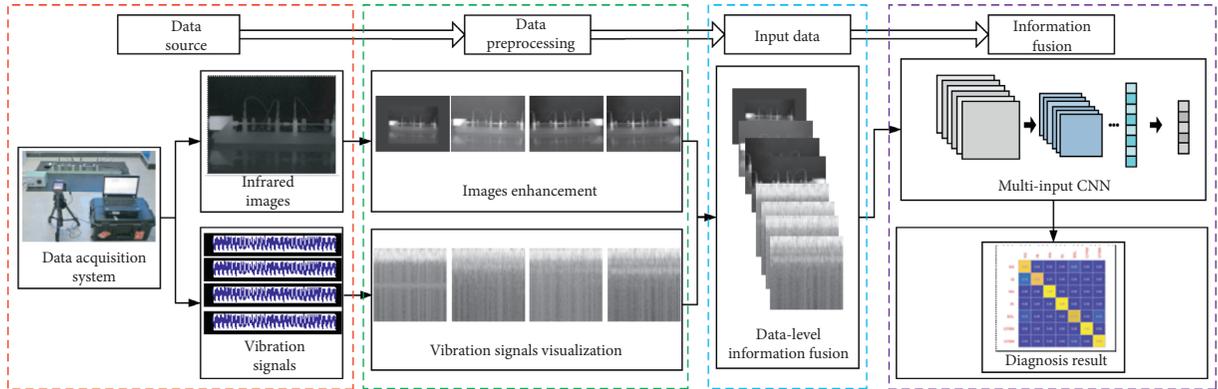


FIGURE 2: The diagram of the proposed method.

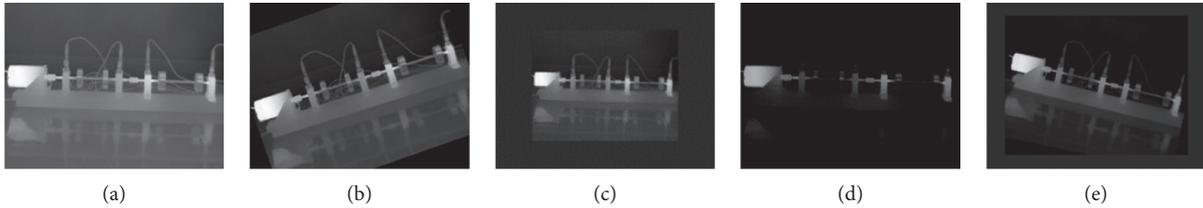


FIGURE 3: Infrared image enhancement. (a) Original. (b) Rotating. (c) Scaling. (d) Brightness adjustment. (e) Mixed operation

Generally, vibration signal is one-dimensional time-domain waveform data, however, the form of two-dimensional images is needed as the input of CNN. Therefore, the vibration data must be two-dimension visualized before application. So, the vibration signals are processed into a two-dimensional Time-Frequency Graph (TFG) using STFT in this paper, to realize the data-level fusion of vibration signals and infrared images. Hanning window for STFT is used as a window function in equation (1) in this paper, and the parameters in equations (2) and (3) are set as $N_x = 10288$, $N_o = 48$, $N_w = 128$, and $N_f = 128$. After calculation, the TFG with resolution of 128×128 is obtained, as shown in Figure 4.

3.3. Assessment Based on Fusion of Infrared Images and Vibration Signals. The input layer of the traditional CNN network is the image of single-channel gray or 3-channel

RGB type. In order to make use of multisource heterogeneous data and to realize data-level fusion, it is necessary to change the structure of the CNN network, construct multichannel input layer, and simultaneously input the infrared images and the TFGs of vibration signals to train the CNN model. Through this method, the infrared images and vibration data information are collaboratively extracted, and fusion diagnosis is realized in deep learning network.

The structure of the proposed multichannel CNN network is shown in Figure 5. The data preprocessing layer is used for infrared images enhancement and vibration signals visualization. After the data preprocessing layer, we get 4 infrared images and 4 vibration TFGs with the same resolution of 128×128 , and the 8 images are combined as 8-channel training data of the input layer, and then the CNN model with 6 convolutional layers, 6 pooling layers, and 3 fully connected layers is designed for fault diagnosis. ReLU

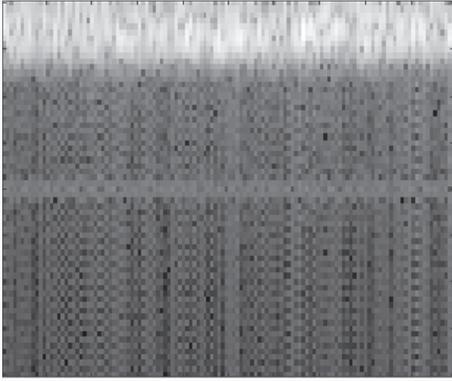


FIGURE 4: TFG of the vibration signal.

function is chosen as the activation function of each layer. Every pooling is connected to a Dropout layer with the shielding probability of 0.5. The initial learning rate is set to 0.001, and the exponential decay strategy is used to adjust the learning rate, decay rate and decay steps are set to 0.9 and 100, to achieve the best performance. Softmax classifier is chosen in the output layer, and the output is the fault diagnosis results.

4. Experimental Study

4.1. Experimental Setup. To verify the proposed method, a rotor fault simulation test stand is built to simulate different faults of the rotor system and collect vibration signals and infrared images. The schematic diagram of the experiment is shown in Figure 6(a). The hardware system includes the ZT-3 test stand, MDES vibration signal acquisition system, and FLIR E50 infrared image acquisition system.

The ZT-3 rotor test stand is composed of the base, rotor system, motor, and speed controller. The rotor system includes rotors, shafts, couplings, and bearings. During the whole experiment, the speed controller is set to 6000 rpm.

MDES vibration signal acquisition system is used to collect vibration signals in this experiment, which includes a computer, acquisition instrument, and acceleration sensors. The 4 measuring points for vibration signals are distributed as shown in Figure 6(b), which are marked as V1, V2, V3, and V4. The vibration signals are collected from each measurement point with a sampling frequency of 20000 Hz and a sampling length of 20000 points.

The FLIR E50 infrared thermal camera is used to collect infrared images with the image resolution 320×240. In order to verify the effectiveness of the proposed method under different collection conditions, the infrared images are acquired in two different cases. As shown in Figure 6(c), the camera is perpendicular to the test stand in Case 1, and the angle between the camera and the test stand is about 45 degrees in Case 2. Seven rotors system conditions are simulated in each case, including normal state (NS), imbalance (IB), misalignment (MA), rub-impact (RI), bearing set loose (BSL), and coupling faults of rub-impact and misalignment (CFRM) and coupling faults of bearing set loose and misalignment (CFBM). 100 datasets are measured

at each condition for each case, that is, the total of 700 images and 700 vibration signals at each measuring point for each case. Of the 100 images and 100 sets of vibration data at each state, 50 datasets are used for training and the remaining 50 for testing. The experiment is conducted under the TensorFlow deep learning framework and hardware with an I5-9400 CPU and a RTX2080 GPU.

The data samples collected in this experiment are shown in Figure 7, where 4 vibration signals from the four different measuring points are shown in Figure 7(a). The amplitudes of the four data are different, but it is difficult to determine the equipment state only by these time-domain waveforms. Figures 7(b) and 7(c) show the infrared images obtained in Cases 1 and 2, which can cover the entire test stand. In these two cases, the temperature distribution of the equipment at run time can be seen from different angles.

4.2. Analysis of the Infrared Images. Compared with the wide application of the rotor fault diagnosis method based on vibration signals, the method based on infrared images is less used. To test the effect of infrared images in the rotor state monitoring, infrared images under different states are labeled by Region of Interest (ROI), which contains bearings, coupling, and rotor, as shown in Figure 8.

Figure 8(a) shows the infrared image in the normal state and its ROI region, and Figure 8(b) shows the ROIs in six different fault states. Through Figure 8, it can be seen that the temperature of the rotor system under the normal state is lower and uniformly distributed, while under fault states, the fault in different positions of the rotor system will cause the corresponding temperature to rise. Taking MA and RI states for example, in both cases, the temperature increases because of the complex vibration caused by the relative displacement of the coupling and the friction between the rotor and the rubbing stick, which will increase the brightness of the corresponding areas in the infrared images. In particular, when multiple faults occur at the same time, the increased brightness areas in the infrared images are the superposition of that for one single fault. For example, in the CFRM state, the increased brightness areas are the superposition of the increased brightness area in the MA and RI state. This phenomenon indicates that the coupling faults of the rotor system can be directly reflected through the infrared images, making it suitable for fault diagnosis.

4.3. Fault Diagnosis Results

4.3.1. Fault Diagnosis Result Using Infrared Images. To verify the effectiveness of infrared images in rotor system fault diagnosis, the original and the enhanced infrared images in Case 1 are used for fault classification experiments respectively, and the CNN network in Section 3 is used for training and testing. The detailed results are shown in Figure 9.

The averaged accuracy of fault recognition by original and enhanced infrared images is 91.71% and 93.71%, respectively, as shown in Figure 9. Some IB samples are classified as NS, and, meanwhile, some NS samples are

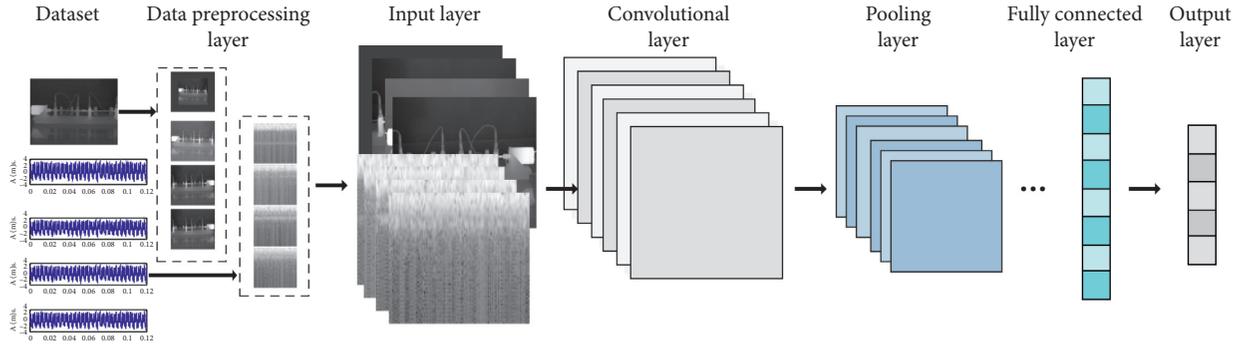


FIGURE 5: The structure of the multichannel CNN network.

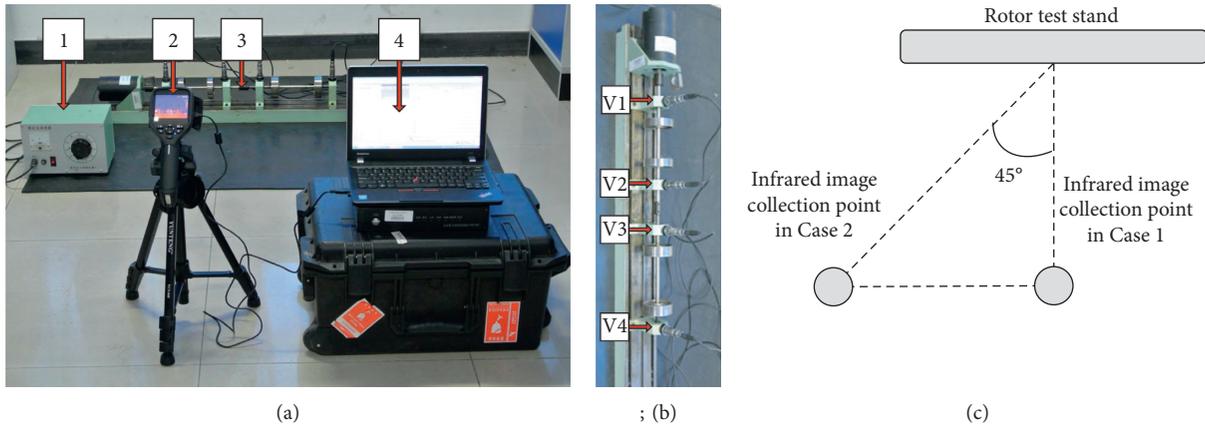


FIGURE 6: Illustration of the experimental setup: (a) acquisition systems of the experiment: (1) speed controller, (2) thermal camera, (3) rotor test stand, and (4) computer; (b) measuring points of vibration signals; (c) Positions of test stand and thermal camera in two cases.

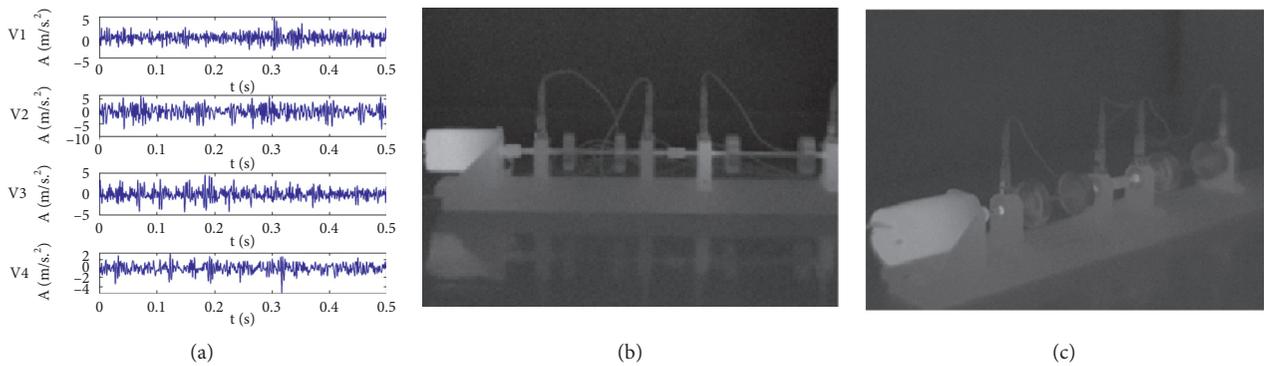


FIGURE 7: Sample data obtained in the experiments: (a) 4 vibration signals from different measuring points, (b) infrared image in Case 1, and (c) infrared image in Case 2.

classified as IB. Some BSL samples are recognized as NS and CFBM, and a few of CFBM faults are identified as BSL. More IB and BSL samples are wrongly classified because the minor imbalance causes slight temperature variation. The features in infrared images are not evident and similar to those at normal state. The other samples are correctly recognized.

4.3.2. Fault Diagnosis Result Using Vibration Signals. The vibration signals in Case 1 are processed using the method in

Section 3.2, and TFGs are obtained for the vibration signals collected from four vibration measuring points in the experiment. The traditional CNN method is used to process the data of each measurement point respectively for training and testing, and the test is performed consistently with that in Section 4.3.1. The classification results for the samples at different measuring points are illustrated in Figure 10.

The recognition accuracy rates at 4 measuring points are 74.00%, 81.43%, 84.57%, and 69.71%, respectively. It can be seen from Figure 10 that the accuracy of recognition for

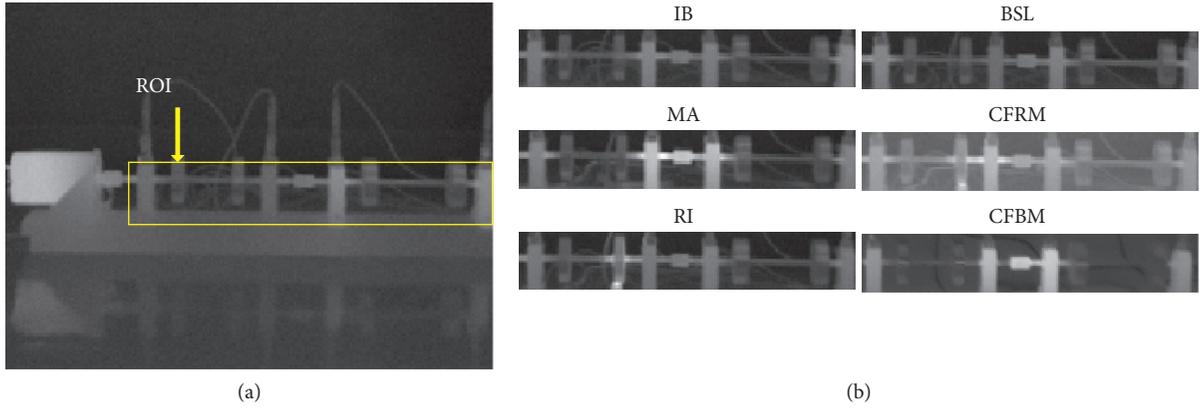


FIGURE 8: Comparison of infrared images in different states., (a) Infrared Image and ROI in the normal state. (b) ROIs at different conditions

	NS	IB	MA	RI	BSL	CFRM	CFBM
NS	0.88	0.04	0.00	0.00	0.08	0.00	0.00
IB	0.18	0.82	0.00	0.00	0.00	0.00	0.00
MA	0.00	0.00	1.00	0.00	0.00	0.00	0.00
RI	0.00	0.00	0.00	1.00	0.00	0.00	0.00
BSL	0.10	0.00	0.00	0.00	0.78	0.00	0.12
CFRM	0.00	0.00	0.00	0.00	0.00	1.00	0.00
CFBM	0.00	0.00	0.00	0.00	0.06	0.00	0.94

(a)

	NS	IB	MA	RI	BSL	CFRM	CFBM
NS	0.92	0.04	0.00	0.00	0.04	0.00	0.00
IB	0.16	0.84	0.00	0.00	0.00	0.00	0.00
MA	0.00	0.00	1.00	0.00	0.00	0.00	0.00
RI	0.00	0.00	0.00	1.00	0.00	0.00	0.00
BSL	0.08	0.00	0.00	0.00	0.84	0.00	0.08
CFRM	0.00	0.00	0.00	0.00	0.00	1.00	0.00
CFBM	0.00	0.00	0.00	0.00	0.04	0.00	0.96

(b)

FIGURE 9: Classification result by infrared images. (a) Original infrared images. (b) Enhanced infrared images.

coupling faults is not high at each point. And the precision of recognizing different faults is related to the positions of the four points.

4.3.3. Fault Diagnosis Result Using the Proposed Method.

Infrared images and vibration data in Case 1 are simultaneously processed following the method in Section 3, and the fault diagnosis result is shown in Figure 11.

Compared with the recognition results through infrared images or vibration signals individually, the accuracy of diagnosis by fusion is increased to 99.14%, with only 3 samples wrongly identified from the test dataset. It illustrates that the diagnosis based on fusion is of higher accuracy than that of a single information source.

Moreover, in order to compare with the traditional methods, feature-level and decision-level fusion methods are used for fault diagnosis.

In the feature-level fusion method, 6 histogram features are extracted from infrared images [38]; meanwhile, wavelet packet transformation is applied to the vibration

signals at 7 states, with a wavelet basis of db5 and decomposition level of 3. Energy entropy of 8 frequency bands is calculated as characteristic vectors. SVM and ANN are used for fault diagnosis, using the composite vector composed of infrared images features and vibration signals features as input data.

In the decision-level fusion method, infrared images and vibration signals are processed according to the methods in Section 3.2, with CNN as a classifier, and the TFGs of 4 vibration sensors and the infrared images are used for fault diagnosis, respectively, so that 5 diagnosis results are obtained; then, the D-S evidence theory and fuzzy decision theory are used to synthesize the fault diagnosis results. The results are shown in Table 1.

It can be seen from Table 1 that when infrared images and vibration signals are used for rotor fault diagnosis, using the information fusion methods, the effect is better than using one single type of data source. Besides, the proposed method avoids fault information losing due to the different distribution characteristics of the feature information and evidence conflicts between results of different sensors. Compared with other

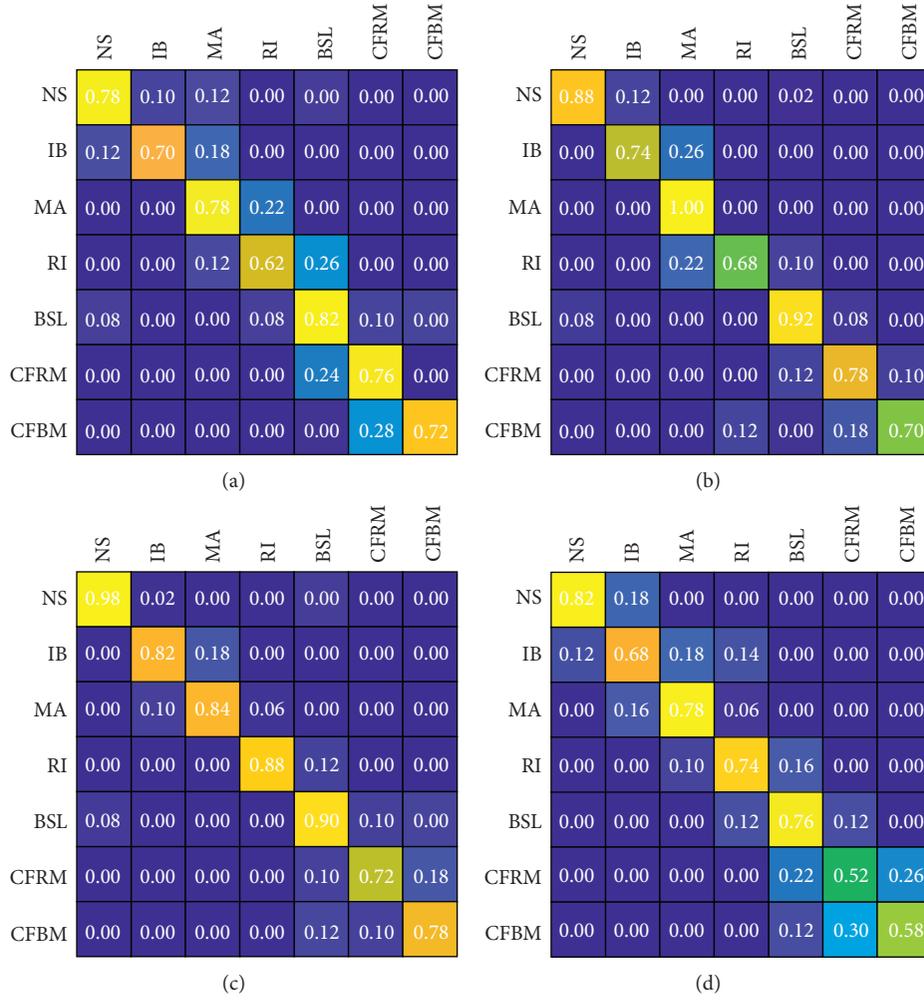


FIGURE 10: Classification results by vibration signals at different measuring Points: (a) V1, (b) V2, (c) V3, and (d) V4.

	NS	IB	MA	RI	BSL	CFRM	CFBM
NS	0.98	0.00	0.00	0.00	0.02	0.00	0.00
IB	0.00	1.00	0.00	0.00	0.00	0.00	0.00
MA	0.00	0.00	1.00	0.00	0.00	0.00	0.00
RI	0.00	0.00	0.00	1.00	0.00	0.00	0.00
BSL	0.00	0.00	0.00	0.00	0.96	0.00	0.04
CFRM	0.00	0.00	0.00	0.00	0.00	1.00	0.00
CFBM	0.00	0.00	0.00	0.00	0.00	0.00	1.00

FIGURE 11: Classification result by the proposed method.

TABLE 1: Comparison of test accuracy from different diagnostic models.

Methods	Accuracy (%)	
	Case 1	Case 2
SVM based feature-level fusion	93.14	92.29
ANN based feature-level fusion	93.43	92.57
D-S based decision-level fusion	96.00	94.29
Fuzzy decision based decision-level fusion	95.14	94.00
The proposed method	99.14	98.29

methods, the proposed method achieves the highest accuracy, which verifies the effectiveness of the method.

5. Conclusions

In this study, a new method for machinery fault diagnosis by fusing infrared images and vibration signals is proposed. Experiment tests on a rotor test stand are performed to validate the effectiveness and accuracy of the method. The conclusions can be drawn as follows:

- (1) The presented multichannel CNN model can realize the data-level fusion of infrared images and vibration signals; meanwhile, it avoids the information losing caused by feature extraction and information synthesis.
- (2) The accuracy of fault diagnosis based on vibration signals is sensitive to positions and the recognition for coupling faults is not effective. Coupling faults can be identified using infrared images. However, faults such as IB and BSL are not sensitive to temperature variation; thus, they cannot be correctly recognized.
- (3) The proposed method can effectively identify the coupling faults and improve the classification accuracy up to 99%, which is higher than the methods based on the vibration signals or infrared images only.

In addition to the above conclusions, with the rapid development of signal processing method, the idea of data-level fusion fault diagnosis method proposed in this paper can be extended to different deep learning networks, and fusing different types of signals, to test whether better performance can be obtained in complex fault diagnosis. At the same time, in order to verify the validity of the proposed method and avoid introducing more variables, the contents concerning the processing of images or vibration signals were not introduced. It can be inferred that if effective signal preprocessing is carried out, improved performance of fault diagnosis would be expected.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research acknowledges the financial support provided by the General Project of Scientific Research Program of Beijing Education Commission (KM202010016003), Post-doctoral Science Foundation of Beijing, China (ZZ2019-98), National Natural Science Foundation of China (51975038 and 51605023), and Nature Science Foundation of Beijing, China (19L00001).

References

- [1] H. Shi, Y. Li, X. Bai et al., "Investigation of the orbit-spinning behaviors of the outer ring in a full ceramic ball bearing-steel pedestal system in wide temperature ranges," *Mechanical Systems and Signal Processing*, vol. 149, Article ID 107317, 2021.
- [2] L. You, W. Fan, Z. Li, Y. Liang, M. Fang, and J. Wang, "A fault diagnosis model for rotating machinery using VWC and MSFLA-SVM based on vibration signal analysis," *Shock and Vibration*, vol. 2019, Article ID 1908485, 16 pages, 2019.
- [3] L. Xiang, Z. Deng, A. Hu, and X. Gao, "Multi-fault coupling study of a rotor system in experimental and numerical analyses," *Nonlinear Dynamics*, vol. 97, no. 4, pp. 2607–2625, 2019.
- [4] M. Hou and H. Shi, "Stator-winding incipient shorted-turn fault detection for motor system in motorized spindle using modified interval observers," *Institute of Electrical and Electronics Engineers Transactions on Instrumentation and Measurement*, vol. 70, p. 1, 2021.
- [5] C. Gu, X. Qiao, Y. Jin, and Y. Liu, "A novel fault diagnosis method for diesel engine based on MVMD and band energy," *Shock and Vibration*, vol. 2020, Article ID 8247194, 17 pages, 2020.
- [6] W. Zhou, X. Li, J. Yi, and H. He, "A Novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit," *Institute of Electrical and Electronics Engineers Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1415–1424, 2019.
- [7] B. Wang, M. Dong, M. Ren et al., "Automatic fault diagnosis of infrared insulator images based on image instance segmentation and temperature analysis," *Institute of Electrical and Electronics Engineers Transactions on Instrumentation and Measurement*, vol. 69, no. 8, pp. 5345–5355, 2020.
- [8] R. F. Deng, Y. B. Lin, W. J. Tang, F. S. Gu, and A. Ball, "Object-based thermal image segmentation for fault diagnosis of reciprocating compressors," *Sensors*, vol. 20, no. 12, 11 pages, Article ID 20123436, 2020.
- [9] S. Chellamuthu and E. C. Sekaran, "Fault detection in electrical equipment's images by using optimal features with deep learning classifier," *Multimedia Tools and Applications*, vol. 78, no. 19, pp. 27333–27350, 2019.
- [10] A. Kylili, P. A. Fokaides, P. Christou, and S. A. Kalogirou, "Infrared thermography (IRT) applications for building diagnostics: a review," *Applied Energy*, vol. 134, no. 1, pp. 531–549, 2014.
- [11] B. B. Lahiri, S. Bagavathiappan, T. Jayakumar, and J. Philip, "Medical applications of infrared thermography: a review," *Infrared Physics & Technology*, vol. 55, no. 4, pp. 221–235, 2012.
- [12] U. R. Acharya, J. H. Tan, J. E. W. Koh et al., "Automated diagnosis of dry eye using infrared thermography images," *Infrared Physics & Technology*, vol. 71, pp. 263–271, 2015.
- [13] B. F. Jones, "A reappraisal of the use of infrared thermal image analysis in medicine," *Institute of Electrical and Electronics Engineers Transactions on Medical Imaging*, vol. 17, no. 6, pp. 1019–1027, 1998.
- [14] B. B. Lahiri, S. Bagavathiappan, C. Soumya et al., "Infrared thermography based defect detection in ferromagnetic specimens using a low frequency alternating magnetic field," *Infrared Physics & Technology*, vol. 64, pp. 125–133, 2014.
- [15] C. Meola and G. M. Carlomagno, "Recent advances in the use of infrared thermography," *Measurement Science and Technology*, vol. 15, no. 9, pp. 27–58, 2004.
- [16] B. B. Lahiri, S. Bagavathiappan, R. Shunmugasundaram, J. Philip, and T. Jayakumar, "Measurement of annular air-gap using active infrared thermography," *Infrared Physics & Technology*, vol. 61, pp. 192–199, 2013.
- [17] A. M. D. Younus and B.-S. Yang, "Intelligent fault diagnosis of rotating machinery using infrared thermal image," *Expert Systems with Applications*, vol. 39, no. 2, pp. 2082–2091, 2012.
- [18] M. Eftekhari, M. Moallem, S. Sadri, and M.-F. Hsieh, "A novel indicator of stator winding inter-turn fault in induction motor

- using infrared thermal imaging,” *Infrared Physics & Technology*, vol. 61, pp. 330–336, 2013.
- [19] J. Olivier, S. Raiko, S. Viktor, S. Kurt, L. Mia et al., “Thermal image based fault diagnosis for rotating machinery,” *Infrared Physics and Technology*, vol. 73, pp. 78–87, 2015.
- [20] M.-K. Liu, M.-Q. Tran, and P.-Y. Weng, “Fusion of vibration and current signatures for the fault diagnosis of induction machines,” *Shock and Vibration*, vol. 2019, Article ID 7176482, 17 pages, 2019.
- [21] Z. H. Duan, T. H. Wu, S. W. Guo, T. Shao, M. Reza, and Z. X. Li, “Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: a review,” *International Journal of Advanced Manufacturing Technology*, vol. 96, no. 1-4, pp. 803–819, 2018.
- [22] B. Sergio, M. Mario, A. Alberto, C. Pablo, L. Alberto, and M. Raquel, “Dissolved gas analysis equipment for online monitoring of transformer oil: a review,” *Sensors*, vol. 19, 21 pages, 2019.
- [23] G. Qian, S. Lu, D. Pan, H. Tang, Y. Liu, and Q. Wang, “Edge Computing: a promising framework for real-time fault diagnosis and dynamic control of rotating machines using multi-sensor data,” *Institute of Electrical and Electronics Engineers Sensors Journal*, vol. 19, no. 11, pp. 4211–4220, 2019.
- [24] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, “An improved quantum-inspired differential evolution algorithm for deep belief network,” *Institute of Electrical and Electronics Engineers Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7319–7327, 2020.
- [25] Y. Song, D. Wu, W. Deng et al., “MPPCEDE: multi-population parallel co-evolutionary differential evolution for parameter optimization,” *Energy Conversion and Management*, vol. 228, Article ID 113661, 2021.
- [26] Y. Song, D. Wu, A. Wagdy Mohamed, X. Zhou, B. Zhang, and W. Deng, “Enhanced success history adaptive DE for parameter optimization of photovoltaic models,” *Complexity*, vol. 2021, Article ID 6660115, 22 pages, 2021.
- [27] L.-L. Jiang, H.-K. Yin, X.-J. Li, and S.-W. Tang, “fault diagnosis of rotating machinery based on multisensor information fusion using SVM and time-domain features,” *Shock and Vibration*, vol. 2014, Article ID 418178, 8 pages, 2014.
- [28] J. Kamal, M. Mohammadsadegh, J. M. Ruholla, and R. Elaheh, “Misfire and valve clearance faults detection in the combustion engines based on a multi-sensor vibration signal monitoring,” *Measurement*, vol. 128, pp. 527–536, 2018.
- [29] Y. Liu, X. S. Yan, C. A. Zhang, and W. Liu, “An ensemble convolutional neural networks for bearing fault diagnosis using multi-sensor data,” *Sensors*, vol. 19, no. 23, 20 pages, Article ID 19235300, 2019.
- [30] L. Gao, D. H. Li, D. Li, L. L. Yao, L. M. Liang, and Y. N. Gao, “A novel chiller sensors fault diagnosis method based on virtual sensors,” *Sensors*, vol. 19, no. 13, 15 pages, Article ID 19133013, 2019.
- [31] F. Xiao, “Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy,” *Information Fusion*, vol. 46, pp. 23–32, 2019.
- [32] F. Y. Xiao and B. W. Qin, “A weighted combination method for conflicting evidence in multi-sensor data fusion,” *Sensors*, vol. 18, no. 5, 20 pages, Article ID 18051487, 2018.
- [33] J. An, M. Hu, L. Fu, and J. Zhan, “A novel fuzzy approach for combining uncertain conflict evidences in the Dempster-Shafer theory,” *Institute of Electrical and Electronics Engineers Access*, vol. 7, pp. 7481–7501, 2019.
- [34] H. Hao, M. Wang, Y. Tang, and Q. Li, “Research on data fusion of multi-sensors based on fuzzy preference relations,” *Neural Computing and Applications*, vol. 31, no. S1, pp. 337–346, 2019.
- [35] S. Wang and J. Xiang, “A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps,” *Soft Computing*, vol. 24, no. 4, pp. 2983–2997, 2020.
- [36] L. Y. Jing, T. Y. Wang, M. Zhao, and P. Wang, “An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox,” *Sensors*, vol. 17, no. 2, 15 pages, Article ID 17020414, 2017.
- [37] N. Strom, “Sparse connection and pruning in large dynamic artificial neural networks,” in *Proceedings of the European Conference on Speech Communication and Technology*, pp. 2807–2810, Rhodes, Greece, September 1997.
- [38] T. Bai, L. Zhang, L. Duan, and J. Wang, “NSCT-based infrared image enhancement method for rotating machinery fault diagnosis,” *Institute of Electrical and Electronics Engineers Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2293–2301, 2016.