Hindawi
Shock and Vibration
Volume 2023, Article ID 6928871, 12 pages
https://doi.org/10.1155/2023/6928871

Research Article

An Intelligent Diagnostic Method for Multisource Coupling Faults of Complex Mechanical Systems

Xianbin Sun,[1,2] Mohan Wang,[1,3] Bo Zhan,[4] Yuanyuan Xiong,[5] and Wei Yu[3,6]

[1] *Qingdao University of Technology, 777 Jialingjiang East Street, Huangdao District, Qingdao 266520, Shandong, China*
[2] *Key Lab of Industrial Fluid Energy Conservation and Pollution Control (Qingdao University of Technology), Ministry of Education, 777 Jialingjiang East Street, Huangdao District, Qingdao 266520, Shandong, China*
[3] *Qingdao Mingyu Intelligent Equipment Technology Research Institute, 2 Heilong Jiangnan Road, Shibei District, Qingdao 266000, Shandong, China*
[4] *Qingdao Haier International Trade Co., Ltd., No. 1 Haier Road, Laoshan District, Qingdao 266000, Shandong, China*
[5] *Haier COSMO IoT Ecosystem Technology Co., Ltd., No. 1 Haier Road, Laoshan District, Qingdao 266000, Shandong, China*
[6] *Shanghai Municipal Engineering Design Institute (Group), The Seventh Design Institute Co., Ltd., 23 Dexian Road, Shinan District, Qingdao 266000, Shandong, China*

Correspondence should be addressed to Xianbin Sun; robin_sun@qut.edu.cn

Received 7 October 2022; Revised 18 November 2022; Accepted 6 December 2022; Published 7 January 2023

Academic Editor: Antonio Batista

In the actual environment, there are difficult points such as complicated mechanical system fault types, random fault locations, and inconspicuous minor fault signals, which make it difficult to accurately diagnose faults. This paper proposes a new method for fault diagnosis of an adaptive multisensor bearing-gear system based on GAF/MTF (Gramian angular fields and Markov transition fields) and ResNet (deep residual network). First, we establish a multisensor signal acquisition system to monitor the running signals of the bearing-gearbox composite test bench in real time. Faulty parts include multiple types of composite faults of different sizes, different fault types, and different transmission stages. Second, based on GAF/MTFs, the multichannel timing signal collected by using the acquisition system is converted into multichannel pictures, and pictures are fused and compressed into three-channel pictures. Finally, we input these pictures into ResNet for fault diagnosis. The experimental results show that the GAF/MTF-ResNet model has a recognition accuracy of 72.14% for a total of 520 classification label test sets under different motor speeds, different sampling times, and different types of faults. Among them, the accuracy of the motor speed and sampling time is close to 100%, and the accuracy of gearbox failure and bearing failure is 75.25% and 88.97%, respectively. This shows that the method provides new ideas for the composite fault diagnosis of mechanical systems under different working conditions and different types of faults and has theoretical guiding significance.

1. Introduction

With the progress of science and technology and the rapid development of the economy, mechanical equipment plays an increasingly irreplaceable role in people's daily life. This puts forward more stringent requirements for the safety and stability of the equipment. Prognostic health management (PHM), as an extremely critical part of equipment operation and maintenance, has attracted widespread attention [1, 2].

Bearings and gearboxes, as the basic components of mechanical equipment, have the characteristics of large usage, high precision requirements, and the complex operating environment. Their failure will directly affect the overall operating status of the equipment. They have become a research hotspot in the field of PHM [3, 4]. However, there are still insufficient research studies on early failures, composite failures, and multiple types of failures of these components [5, 6].

For the diagnosis of a single component failure, Sun and Jia proposed a new method based on experimental data-driven random fuzzy evidence acquisition and intuitionistic fuzzy set fusion [7]. Meng et al. calculate the time-varying mesh stiffness of gears teeth with different crack lengths. The results show that the impulse factor is sensitive to fault characteristics [8]. Shao et al. use transfer learning for bearing fault diagnosis and have achieved good results [9]. Although these studies have good performance in the fault diagnosis of a single component, they have not considered the complexity of the mechanical system and compound faults caused by the coupling of various components in actual operation.

Gearboxes have become the main platform for compound failure experiments due to their wide range of applications and simple structure [10]. Compared with the signal generated by the failure of a single component, the signal, vibration, sound, current, torque, and other different signals, generated by the compound failure of the gearbox is more diverse and complex. These signals have different performances and require different sensors to collect. Different sensors have higher sensitivity to specific fault types. Therefore, fusion of signals from multiple sensors is a key link in gearbox fault diagnosis [11]. Feature extraction and selection of different types of signals have become two major difficulties in multisensor information fusion [12, 13].

The neural network has the advantages of self-adaptation and high precision and has been widely used in various fields [14, 15]. As a typical structure of the neural network, the convolutional neural network (CNN) [16] has been used in image recognition [17], pose estimation [18], and other fields. Deep convolutional neural networks (DCNNs) [19] is to meet people's higher precision requirements for neural networks, which are obtained by further deepening the number of network layers based on the CNN network structure. Studies have found that DCNN has better recognition accuracy than CNN [20–22]. The deep residual network (ResNet), as a kind of DCNN, solves the degradation problem of deep networks [23]. Therefore, the network can be used to build a super deep network, such as 101-layer depth to solve complex problems [24, 25].

In the field of mechanical fault diagnosis, DCNN has also been well applied. Shao et al. use the wavelet transform to generate time-frequency distribution (TFD) images from multisensor signals and then use DCNN to learn discriminative representations from TFD images to perform fault diagnosis of asynchronous motors [26]. Based on multisensor data fusion, Jing proposed an adaptive multisensor data fusion method based on DCNN for fault diagnosis. The results demonstrate that the proposed method can detect the conditions of the planetary gearbox effectively [27].

Although these studies have made in-depth research on the fault diagnosis of mechanical components, there are the following problems [28]:

(1) At present, much research exists on the fault diagnosis of a single component of the mechanical system, and the research on coupling faults is not in depth. In particular, coupling failures of different transmission stage components have not been studied [26, 27].

(2) The operating conditions are single, fault types are few, and the diagnostic results are single [29], and the complex operating state under actual operation cannot be well simulated.

(3) The classical neural network model has the problem of gradient disappearance or gradient explosion when dealing with long-term series data, which reduces the accuracy of fault diagnosis.

In response to the abovementioned problems, we built a multisensor acquisition system, including three-way accelerometers, microphones, current sensors, torque sensors, and rotary encoders based on the bearing-gearbox composite test bench. The designed fault types include different sizes and types of faults, individual faults, composite faults in the same transmission stage, and composite faults in different transmission stages. After that, the GAF/MTF algorithm is used to fuse the multidimensional time series into two-dimensional RGB images. These images are used to train ResNet in order to find the optimal model. Through the GAF/MTF algorithm, the complex multidimensional time series signal task is transformed into a two-dimensional image classification task suitable for neural networks. The multilayer advantage of ResNet can better deal with the transformed images. Finally, the multisensor bearing-gear system for multitype coupling fault diagnosis is realized, and accuracy is high.

2. Time Series Imaging

2.1. Gramian Angular Field and Markov Transition Field. Deep learning has the problem that it performs well in computer vision and pattern recognition, but it performs poorly in processing time series. In response to this problem, Wang and Oates proposed a GAF/MTF method for encoding time series into pictures [30]. This method can encode any type of time series into Gramian angular summation fields (GASF), Gramian angular difference fields (GADF), and Markov transition field (MTF) pictures, respectively. The loss of information after encoding is very small.

Given a time series $X = \{x_1, x_2, \ldots, x_n,\}$ of the length n, the range of the X interval is scaled to $[-1, 1]$ through min-max scaler.

$$\tilde{x}_i^{[-1,1]} = \frac{(x_i - \max(X)) + (x_i - \min(X))}{\max(X) - \min(X)}. \tag{1}$$

The standardized time series $\tilde{X}$ is recalibrated by the polar coordinate encoding of Formula (2), where radian $\varnothing$ is the arc cosine of $\tilde{x}_i$, the interval range is $[0, \pi]$, the radius r is determined by the timestamp t_i corresponding to $\tilde{x}_i$ and the interval $[0, 1]$ phase mapping, and N is a constant factor to regularize the span of the polar coordinate system and is related to the number of time stamps included in the time series.

$$\begin{cases} \varnothing = \arccos(\tilde{x}_i), & -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X}, \\ \\ r = \dfrac{t_i}{N}, & t_i \in \mathbb{N}. \end{cases} \tag{2}$$

The coding of Formula (2) has the following advantages: the bijective mapping is realized from the time series to polar coordinates; that is, given a time series point $\tilde{x}_i$, there is only one point $(\varnothing, r)$ corresponding to it in the polar coordinate system. Contrary to the Cartesian coordinate system, the radius r maintains the time dependence in the polar coordinate system, which can prevent the loss of the time label.

Therefore, in polar coordinates, we can identify the time correlation in different time intervals by calculating the triangular sum or difference between each point. We define GASF and GADF as follows:

$$\mathrm{GASF} = \left[\cos\left(\varnothing_i + \varnothing_j\right)\right] = \tilde{X}'\bullet\tilde{X} - \sqrt{I - \tilde{X}'^2}\bullet\sqrt{I - \tilde{X}^2}, \tag{3}$$

$$\mathrm{GADF} = \left[\sin\left(\varnothing_i - \varnothing_j\right)\right] = \tilde{X}\bullet\sqrt{I - \tilde{X}'^2} - \tilde{X}'\bullet\sqrt{I - \tilde{X}^2}, \tag{4}$$

where I is the unit row vector $[1, 1, \ldots, 1]$ and subscripts i and j of $\varnothing$ represent the different rows and columns of the matrix GASF and GADF, respectively. After converting to the polar coordinate system, we regard the time series of each time step as a one-dimensional quantity space. By defining $\langle x, y\rangle = x\bullet y - \sqrt{1 - x^2}\bullet\sqrt{1 - y^2}$ and $\langle x, y\rangle = y\bullet\sqrt{1 - x^2} - x\bullet\sqrt{1 - y^2}$, the GAFs of the two inner products are actually quasi-Gramian matrices $G_{i,j}$.

The diagonal matrices of GASF and GADF are

$$\mathrm{diag}\,(\mathrm{GASF}) = \{\cos(2\varnothing_1), \ldots, \cos(2\varnothing_n)\} = \left\{2\cos^2(\varnothing_1) - 1, \ldots, 2\cos^2(\varnothing_n) - 1\right\} = \{\mathrm{GASF}_{ii}\}_{1\leq i\leq n}, \tag{5}$$

$$\mathrm{diag}\,(\mathrm{GADF}) = \{\sin(2\varnothing_1), \ldots, \sin(2\varnothing_n)\} = \{2\sin(\varnothing_1)\cos(\varnothing_1), \ldots, 2\sin(\varnothing_n)\cos(\varnothing_n)\} = \{\mathrm{GADF}_{ii}\}_{1\leq i\leq n}. \tag{6}$$

GAFs have several advantages. First, they provide a way to maintain time dependence, because when the matrix position moves from the upper left corner to the lower right corner, the corresponding time also increases. Second, they include time correlation because $G_{(i,j\|i-j|=k)}$ represents the relative correlation of superposition or difference with respect to the direction of the time interval k. The main diagonal $G_{i,i}$ is a special case when $k = 0$ and contains the original value or angle information. Finally, the time series can be reconstructed from the main diagonal. However, when the length of the original time series is n, the size obtained by the Gramian matrix formula is $n \times n$, which leads to larger GAFs.

Similarly, for time series $X = \{x_1, x_2, \ldots, x_n, \}$, we define MTF as follows:

$$M = \begin{bmatrix} \omega_{ij}\big|_{x_1\in q_i, x_1\in q_j} & \cdots & \omega_{ij}\big|_{x_1\in q_i, x_n\in q_j} \\ \omega_{ij}\big|_{x_2\in q_i, x_1\in q_j} & \cdots & \omega_{ij}\big|_{x_2\in q_i, x_n\in q_j} \\ \vdots & \ddots & \vdots \\ \omega_{ij}\big|_{x_n\in q_i, x_1\in q_j} & \cdots & \omega_{ij}\big|_{x_n\in q_i, x_n\in q_j} \end{bmatrix}. \tag{7}$$

By dividing the data into Q quantiles, the Markov transition matrix (W) of $Q \times Q$ can be established. Among them, q_i and q_j $(q \in [1, Q])$ are the data quantiles containing the time stamp i and j (temporal axis), respectively. $\omega_{i,j}$ is given by the frequency with which a point in the quantile q_j is followed by a point in the quantile q_i. By considering the time position, the matrix W containing the transition probability on the amplitude axis is expanded into the MTF matrix (M). The main diagonal M_{ii} is the probability from each quantile to itself when $k = 0$ ((the self-transition probability). The MTF matrix (M) is more sensitive to the distribution of X and the time dependence of the time step t_i than the Markov transition matrix (W), which is constructed by directly calculating the transitions between quantiles along the time axis in a first-order Markov chain [30].

2.2. Multidimensional Time Series Imaging. We introduce the GAF/MTF framework into multisensor data fusion. First, the GAF/MTF unit is established, and then, the multidimensional time series imaging framework is established based on this unit. The framework is used to synchronously compress and fuse the time series collected by different sensors into RGB images of a specified size according to the timestamp.

GASF, GADF, and MTF conversion are performed on the time series $X = \{x_1, x_2, \ldots, x_n, \}$ of the length n, and a three-dimensional matrix I of size $[n, n, 3]$ is obtained. Each layer in the third dimension of matrix I corresponding to GASF, GADF, and MTF conversion results from top to bottom. Because the size of the matrix I is related to the length n of the time series X, when n is large, the size of the matrix I will be too large, which is not conducive to subsequent calculations. Therefore, after downsampling I with a decimation filter, a matrix I' of size $[n', n', 3]$ is obtained. The GAF/MTF unit is established, as shown in Figure 1, where the input is a single-channel timing signal and the output is expanded into three two-dimensional single-channel pictures, represented by grayscale images.

Further more, for the multichannel timing signal $\{\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_m\}$, the number of channels is m and the signal length is n. After GAF/MTF conversion, a matrix set containing m pictures of size $[n', n', 3]$ is obtained. The set size is $[m, n', n', 3]$.

In Formulas (4)–(6), the larger the value of the calculation result, the more obvious the feature. We propose the concept of vertical maximum pooling; that is, the maximum

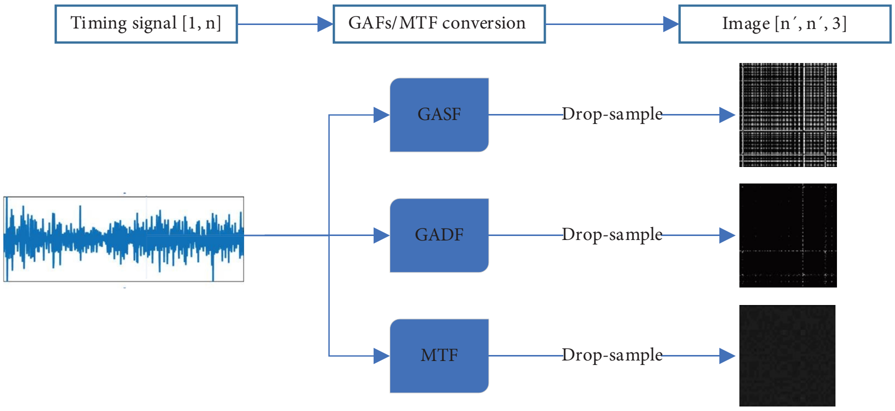


Figure 1: GAF/MTF unit.

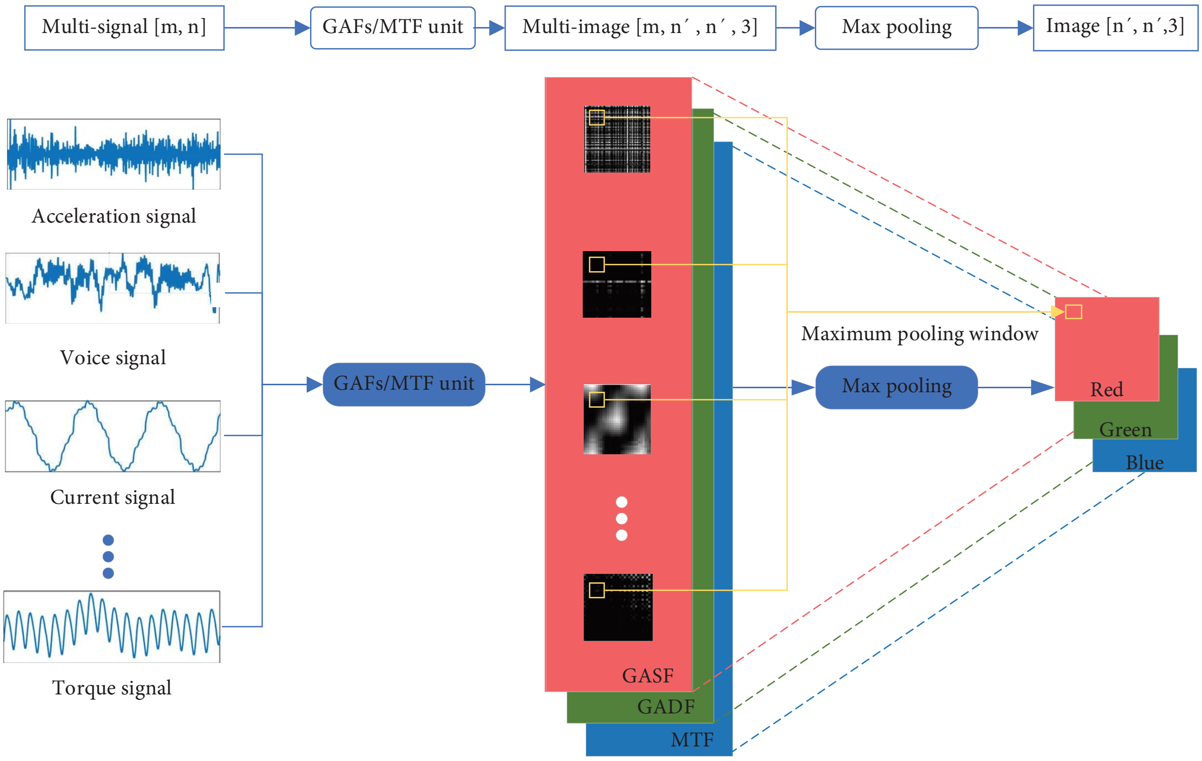


Figure 2: Multidimensional time series imaging.

pooling window with a depth of m is established at the corresponding position of the multidimensional picture, and the window size and sliding step are set to 1. By taking the maximum value in each window, the size of the matrix set is changed from $[m, n', n', 3] \longrightarrow [1, n', n', 3] \longrightarrow [n', n', 3]$. After that, the GASF layer, GADF layer, and MTF layer are placed on the red layer, green layer, and blue layer, respectively, and saved as RGB images. The flowchart is shown in Figure 2.

Through this method, the time series signal classification problem is converted to an image classification problem, which will bypass the shortcomings of deep learning that does not perform well in processing the time series signal. At the same time, it realizes the feature-level fusion of the signal, reduces data dimension, and is more conducive to the calculation of subsequent steps. The GAF/MTF conversion used does not need to set hyperparameters, which avoids the interference caused by manual feature selection.

3. Deep Residual Network

3.1. Architecture of the Deep Residual Network. For the two-dimensional image classification problem, there are already relatively mature neural networks available for use, including AlexNet [31], GoogleNet [32], and VGG [33]. These networks are deep convolutional networks. With the improvement of people's requirements for image classification accuracy and data volume, the depth of the network model continues to increase, and the learning ability is further enhanced. However, the degradation problem of the deeper

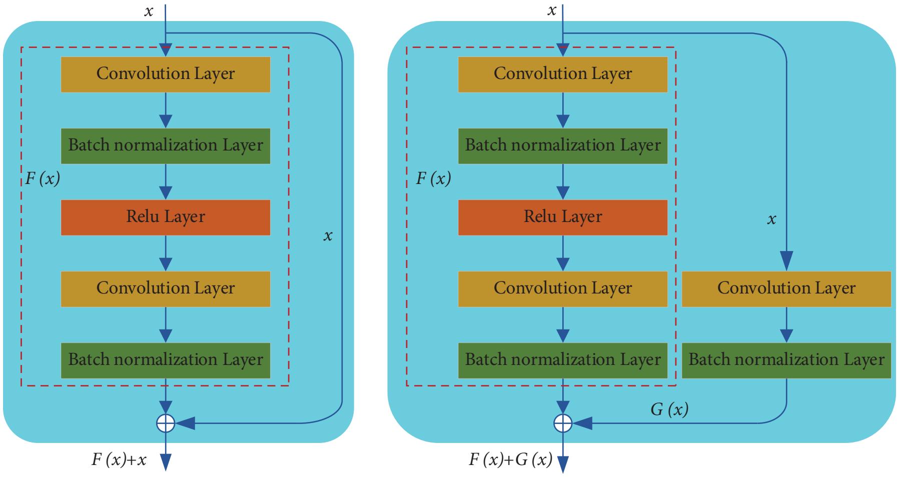


Figure 3: Shortcut connections.

network may cause a higher error rate than the shallower network [34]. In response to this problem, He et al. proposed the deep residual network (ResNet), which has a special "shortcut connection" method compared with the previous network [23]. Figure 3 shows the two types of shortcut connections.

Through this structure, the input x and output $F(x)$ of the block undergo element-wise superposition. This simple addition does not add additional parameters and calculations to the network, but it can greatly improve the speed and effect of model training. When the number of layers of the model increases, this simple structure can solve the degradation problem well.

3.2. Training Network Parameter. In the process of model building, the training time of the network is the most important part of the time cost. We introduce some special mechanisms to improve the training speed of the network, including stochastic gradient descent with momentum (SGDM) solver, L2 regularization, and minibatch. The SGDM solver is based on the stochastic gradient descent (SGD) solver combined with first-order momentum to simulate the concept of momentum in physics. It replaces the true gradient by accumulating previous momentum [35]. The solver can achieve good acceleration in the early stage of the descent; in the middle and late stages of the descent, it jumps out of the local minimum; when the gradient changes direction, it suppresses oscillation and accelerates convergence. L2 regularization will add an L2 norm after the original loss function. This constraint usually imposes a large penalty on sparse and peaked weight vectors and prefers uniform parameters [36]. This effect will encourage the neural unit to use all the input of the upper layer instead of part of the input. Therefore, it will avoid the overfitting phenomenon. The minibatch divides a large training set into several small training sets, and each time only the samples contained in the small training set are used for training [37]. Compared with the batch gradient descent trained with all samples, the parameters using mini batch gradient descent are updated faster. It is conducive to more robust convergence and avoids local optimal solutions and can avoid the amount of data imported into the network at a time and reduce the hardware demand.

3.3. Adaptive Diagnostic Model for Multicondition Composite Fault of the Bearing-Gear System. An adaptive multisensor bearing-gear system for the fault diagnostic model based on GAF/MTF and ResNet is proposed. This model can fuse different lengths and different types of time series collected by using sensors at different sampling frequencies into two-dimensional images through GAF/MTF adaptive feature-level data. All two-dimensional images have the same size and number of channels. In theory, this fusion method can fuse arbitrary sensor data. ResNet used later, through its ultradeep network structure and special residual mechanism, can mine deep fault features in a relatively short time and make decisions to obtain the final diagnostic result.

As the network depth increases, the model accuracy and training time will increase accordingly. Therefore, we have selected ResNet18 for the model, which contain 18 convolutional layers, representing the deep network and ultradeep network.

The flowchart of the model is shown in Figure 4. First, different types of sensors collect experimental data. Sensors include three-way accelerometers, microphones, current sensors, torque sensors, and rotary encoders. The sampling frequency of sensors is related to the motor speed. After that, the collected experimental data are preprocessed, including operations such as deleting invalid signals and randomly dividing signal segments. Through these operations, the n-channel sensor data can be preprocessed to obtain a matrix of size $[n, l, z]$. In the matrix, l indicates the length of the signal segment and z represents the number of signal segments divided. The preprocessed time series signal matrix uses GAF/MTF for feature-level data fusion to obtain an RGB image dataset of size $[m,\ m,\ z]$. Finally, ResNet is used to classify the image set to obtain the fault diagnostic result.

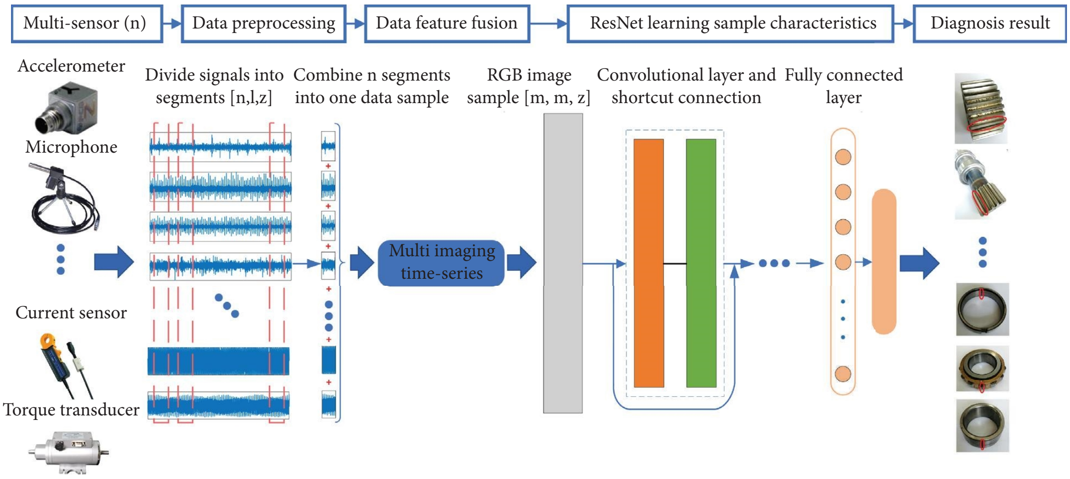

Multi-sensor (n)
Data preprocessing
Data feature fusion
ResNet learning sample characteristics
Diagnosis result
Accelerometer
Microphone
Current sensor
Torque transducer
Divide signals into segments [n,l,z]
Combine n segments into one data sample
RGB image sample [m, m, z]
Convolutional layer and shortcut connection
Fully connected layer
Multi imaging time-series

Figure 4: Flowchart of the GAF/MTF-ResNet model.

In Figure 4, the input of the model is the multichannel time series collected by using sensors, and the output is the diagnostic result, which realizes end-to-end diagnosis. In the operation of the model, the internal structure can be adjusted adaptively without manual participation. The adaptive step includes the automatic fusion of time series signals of different lengths into pictures of equal size in the GAF/MTF conversion; in the training of ResNet, the adaptive adjustment of hyperparameters facilitates the sharing of parameters in network operation.

4. Experiment and Analysis

4.1. Construction of a Test Bed. The mechanical part of the bearing-gear system fault diagnostic test bench includes a drive motor, a bearing seat with replaceable bearings, and a planetary gearbox, as shown in Figure 5. The bearing is fixed in the bearing seat by the bearing end cover and can be replaced. The planetary gearbox includes a sun gear at the driving end, three planet wheels surrounding the sun gear, and a ring gear fixed to housing. The motor drives the bearing and the sun gear to rotate synchronously through the transmission shaft, and the planetary gear is driven by the meshing of the sun gear. There is a long transmission distance between the bearing and the gearbox. By replacing the faulty parts of the bearing and the gearbox, it is possible to simulate compound failures under different transmission stages.

In this study, we built a multisensor acquisition system, as shown in Figure 6. The acquisition system includes two three-way accelerometers, two microphones, torque sensors, rotary encoders, and current sensors. The three-way accelerometer and the microphone are, respectively, arranged at the bearing seat and the planetary gearbox; the torque sensor is arranged at the output end of the motor; the rotary encoder is arranged at the input end of the planetary gearbox; the current sensor is arranged on an input line of the driving motor. Table 1 shows the types of sensors used in the experiment and the types of signals collected separately. Through these sensors, various data during the operation of the experimental platform can be monitored synchronously and saved by using the NI industrial computer.

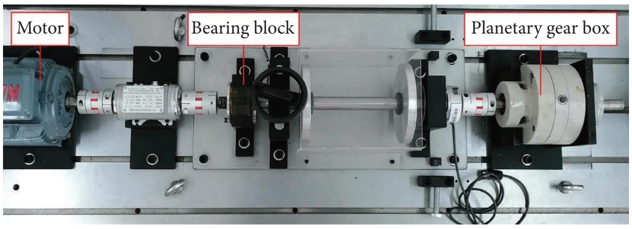


Figure 5: The mechanical system of the test bed.

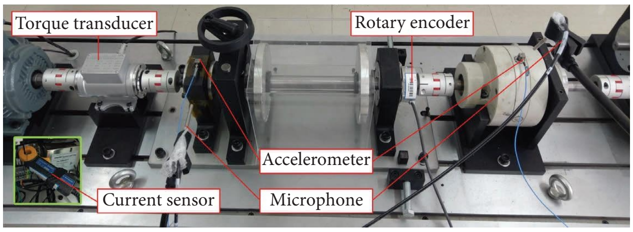


Figure 6: The acquisition system of the test bed.

Table 1: Types of sensors and signals.

Sensor type	Sensors number	Signal type
Three-way accelerometers	2	Vibration signal
Microphones	2	Sound signal
Torque sensors	1	Torque signal
Rotary encoders	1	Rotating speed signal
Current sensors	1	Current signal

4.2. Types of Faulty Parts. In the experiment, the artificial faults of bearings and gears were made to simulate the faults generated under the actual operation of the equipment. The bearing is a cylindrical roller bearing, which is divided into

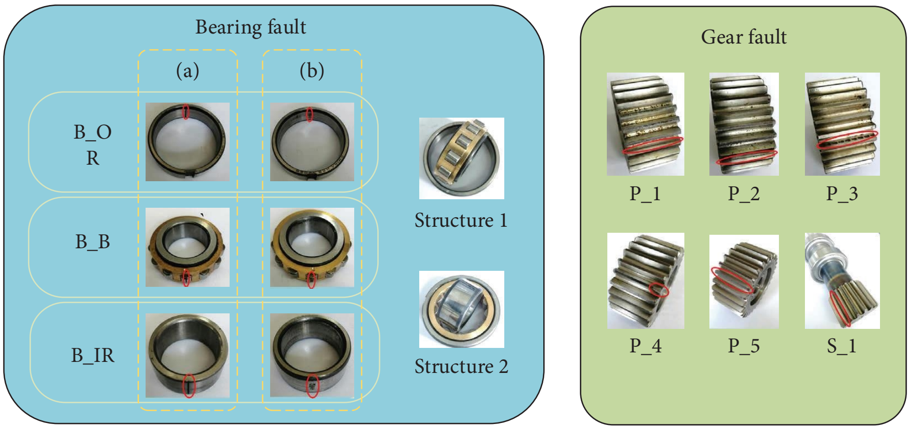


Figure 7: Failure part list (the red circle marks the fault location).

Table 2: Failure parts list.

Fault location		Fault type	Processing methods	Fault size (mm)	Coding
Bearing	Outer ring	Line erosion	Wire cutting	0.10	B_OR(a)
	Outer ring	Pitting	EDM	0.05	B_OR(b)
	Rolling element	Line erosion	Wire cutting	0.10	B_B(a)
	Rolling element	Pitting	EDM	0.05	B_B(b)
	Inner ring	Line erosion	Wire cutting	0.10	B_IR(a)
	Inner ring	Pitting	EDM	0.05	B_IR(b)
Gearbox	Planetary gear	Severe line erosion	Wire cutting	0.20	P_1
	Planetary gear	Slight line erosion	Wire cutting	0.05	P_2
	Planetary gear	Pitting on the tooth surface	EDM	0.05	P_3
	Planetary gear	Gear missing angle	Milling	0.10	P_4
	Planetary gear	Tooth surface microwear	Milling	0.02	P_5
	Sun gear	Severe line erosion	Wire cutting	0.20	S_1

two structures: the outer ring is detachable, which is used to make the outer ring and rolling element failure; the inner ring is detachable, which is used to make the inner ring failure. The two bearings have the same size. Regarding the gearbox, we made different types of failures of the planetary gear and the sun gear. Figure 7 shows the types of faulty parts of bearings and gearboxes. Through the combination of fault types of bearings and gearboxes, many mechanical faults in the actual environment can be simulated.

Table 2 shows a list of faulty parts, corresponding to the faulty part coding shown in Figure 7. Three processing methods are mainly used in the production of faulty parts, electrodischarge machining (EDM), wire cutting, and milling. Among them, wire cutting is used to produce larger faults; EDM and milling are used to produce smaller faults. There are 12 types of faulty parts and 2 types of normal parts in the experiment. The 12 types of faulty parts include 6 types of bearing failure parts and 6 types of gearbox failure parts.

Correspondingly, we designed the experiments, as shown in Table 3, including the no failure experiment (No. 1), single failure experiment (No. 2–No. 12), compound failure experiment in the same transmission stage (No. 13–No. 19), compound failure experiment in different transmission stages (No. 20–No. 25), and multiple compound failure experiment (No. 26). Each experiment in the table contains subexperiments with four motor speeds. The approximate speeds of the motors are 86.4 RPM, 288 RPM, 576 RPM, and 864 RPM, respectively. The precise speed can be calculated by using the rotary encoder on the test bench. A total of 104 experiments are performed.

4.3. Comparative Model. We divide the collected data into the training set, validation set, and test set at a ratio of 7 : 1 : 2. For the 26 types of faults at the four speeds, 5 sampling lengths of 0.2 s, 0.4 s, 0.6 s, and 1 s were used to intercept. Finally, a total of 520 categories are divided, and each category contains 200 signal samples.

In order to verify the superiority of the GAF/MTF-ResNet model, we established some appropriate comparison models based on the results of peer research. Figure 8 shows the structure of each comparative model. (a) In 2019, Shao et al. proposed the fault diagnostic model for multisensor [26], which uses the accelerometer and current sensor to collect the operating signal of the test bed, and then the authors perform wavelet transform on the signal collected by using each sensor to obtain the corresponding time-frequency diagrams and establish a multi-input single-output DCNN network for single fault diagnosis of the gearbox. (b) On the basis of (a) model, we increase the input sensor signal to 10 types, including the vibration signal, current signal, torque signal, and sound pressure signal. (c) Jing et al. proposed a multisensor diagnostic model for compound faults in 2017 [27]. The model preprocesses the signals

Table 3: Experiment list.

Number	Experiment type	Experiment name	Failure part coding
1	Normal	Normal gear-normal bearing	NAN
2	Single failure	Normal gear-bearing outer ring line erosion	B_OR(a)
3		Normal gear-bearing outer ring pitting	B_OR(b)
4		Normal gear-bearing rolling element line erosion	B_B(a)
5		Normal gear-bearing rolling element pitting	B_B(b)
6		Normal gear-bearing inner ring line erosion	B_IR(a)
7		Normal gear-bearing inner ring pitting	B_IR(b)
8		Planetary gear severe line erosion-normal bearing	P_1
9		Planetary gear slight line erosion-normal bearing	P_2
10		Pitting corrosion of planetary gear tooth surface-normal bearing	P_3
11		Planetary gear missing angle-normal bearing	P_4
12		Planetary gear tooth surface microwear-normal bearing	P_5
13		Sun gear severe line erosion-normal bearing	S_1
14	Compound failure in the same transmission stage	Normal gear-bearing outer ring and rolling element line erosion	B_OR(a), B_B(a)
15		Normal gear-bearing outer ring and rolling element pitting	B_OR(b), B_B(a)
16		Severe line erosion of planets and sun gear-normal bearing	P_1, S_1
17		Severe and slight line erosion of planetary gear-normal bearing	P_1, P_2
18		Pitting corrosion of the planetary gear tooth surface and severe line erosion of sun gear-normal bearing	P_3, S_1
19		Missing angle of the planetary gear and severe line erosion of sun gear-normal bearing	P_4, S_1
20	Compound failure in different transmission stages	Planetary gear severe line erosion-bearing outer ring pitting	P_1, B_OR(b)
21		Planetary gear severe line erosion-bearing rolling element pitting	P_1, B_B(b)
22		Planetary gear severe line erosion-bearing inner ring pitting	P_1, B_IR(b)
23		Planetary gear missing angle-bearing outer ring pitting	P_4, B_OR(b)
24		Planetary gear missing angle-bearing rolling element pitting	P_4, B_B(b)
25		Planetary gear missing angle-bearing inner ring pitting	P_4, B_IR(b)
26	Multiple compound failure	Severe line erosion and slight line erosion of the planetary gear and severe line erosion of sun gear-normal bearing	P_1, P_2, S_1

collected by using sensors and directly inputs them into the one-dimensional deep convolutional network (1D-DCNN) model. The output of the network is the type of fault. This model uses some convolutional layers to replace specialized feature extraction units. (d) Thus, we have built the GAF/MTF-ResNet model for multisensor complex fault diagnosis.

All comparative models are built using the Keras framework [38] based on TensorFlow [39], and Numba [40] is used to accelerate training. TensorFlow is an open source code developed and maintained by Google Brain. It is a symbolic mathematics system based on dataflow programming, which has been widely used in various machine learning. Keras is an open source neural network library written in Python and has a high-level application program interface with TensorFlow. The code structure is written by an object-oriented method that is fully modular and extensible, which simplifies the difficulty of model building. Numba is a JIT compiler that can compile Python functions into machine codes. The Python code is compiled by Numba (only for array operations), and its running speed can be close to those of the C code or Fortran code.

The model running environment is the Windows10 system equipped with the NVIDIA TITAN XP graphics card, and the GPU is used to accelerate training.

Table 4 shows the parameters of each model including the comparative model. Among them, the parameter setting of the batch size should be chosen as large as possible based on enough computer video memory. Keskar et al. [41] and Smith et al. [42] proved that when other conditions are the same, a larger batch size will be selected, and the model will show better performance.

4.4. Experimental Results and Analysis. Figure 9 shows the model using ResNet18. As the number of epochs increases, the accuracy of the training set and validation set changes. In order to ensure the reliability of the experimental results, for composite faults, if the model does not predict all faults, the accuracy rate is considered 0.

In Figure 9, for ResNet networks of different depths, the accuracy of the training set can eventually reach 100% accuracy. As the epoch increases, the accuracy of the verification set does not decrease, which shows that the "shortcut connection" mechanism can well avoid the overfitting phenomenon. The accuracy of the validation set of ResNet18 can reach 71.92%, and the accuracy of the test set is 72.14%. The results show that the model has the ability to identify complex fault types. In order to test this ability, we carried out the experiments, as shown in Tables 5 and 6.

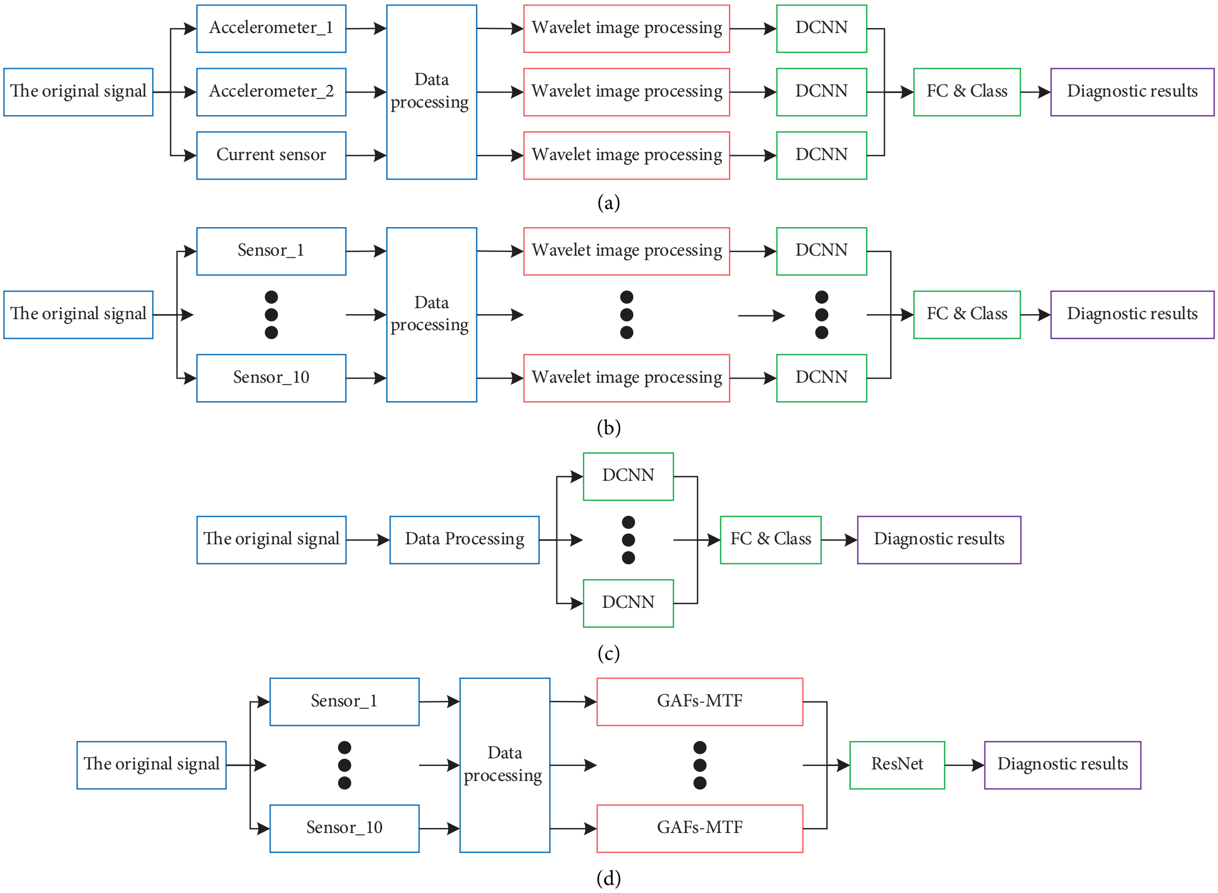


Figure 8: The structure of comparative models.

Table 4: Model parameter settings.

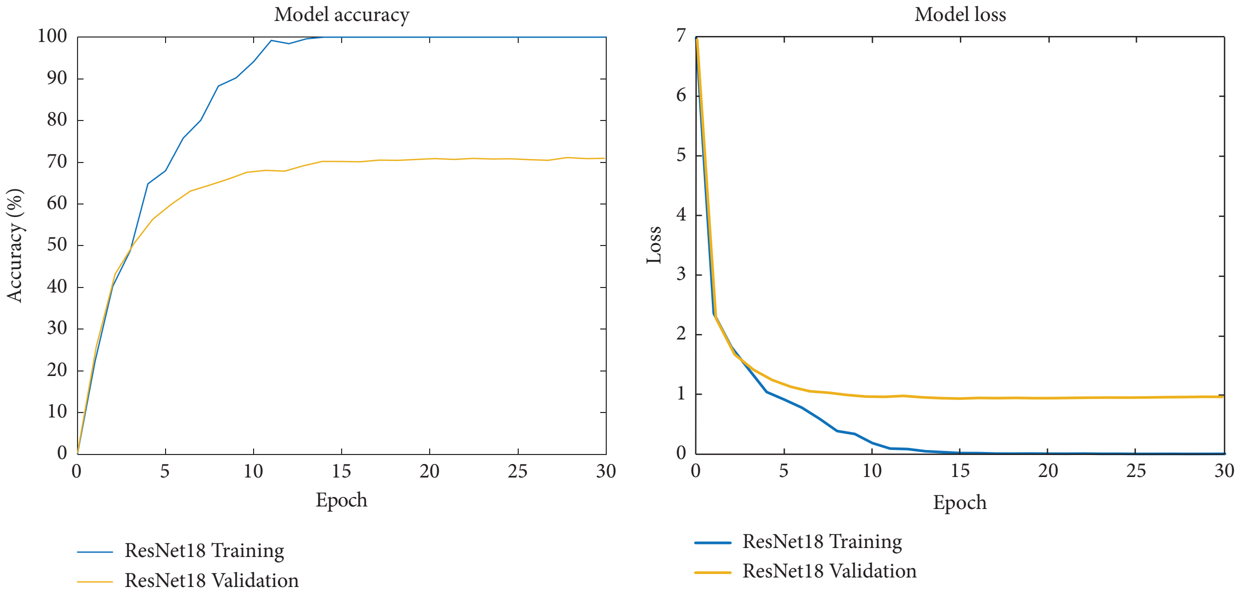
Parameter	Wavelet DCNN	Wavelet DCNN (max)	1D-DCNN	MAF/MTF-ResNet18
Sensor number	3	10	10	10
Input size	128 × 128	128 × 128	120	120 × 120 × 3
Deep network	8	8	7	69
Convolutional layer	3 × 3	10 × 3	10 × 3	18
Epoch	30	30	30	30
Batch size	512	256	512	256
Optimizer	Adam	Adam	SGDM	SGDM
Learning rate	0.0001	0.0001	0.01	0.01

We conducted a more detailed analysis of the results of the test set. All classification labels are divided into five types: the gearbox fault label, bearing fault label, gearbox and bearing fault label, speed label, and time label, among which compound faults at the same transmission stage are regarded as separate fault types. The five types of labels show the accuracy of the test set under each label.

In Table 5, our model has a particularly good classification effect for speed and time labels, and the recognition rate is close to 100%. Compared with gearbox faults, this model has better classification capabilities for bearing faults.

In order to reflect the advantages of the model built compared to those of other models, experiments were carried out on all comparative models. Using the model parameters, as shown in Table 5, all the experiments used the same training set, validation set, and test set and the same running environment. Table 6 shows the results of the comparative experiments. We recorded the accuracy of the training set, validation set, and test set, as well as the training time of the model.

In Table 6, compared to that of other models, the GAF/MTF-ResNet model has better performance for the training set, validation set, or test set. The training time of the model using ResNet18 is not much different from that of other models. The wavelet-DCNN model performs poorly compared to the wavelet-DCNN (max) model because the former uses less sensor data than the latter. The 1D-DCNN model with the one-dimensional time series signal as input has the worst performance, and it has almost no classification ability for different types of faults. The reason may be



(a)

(b)

Figure 9: (a) Model training accuracy. (b) Model training loss.

Table 5: Test set accuracy (%) after subdividing labels.

Model	Gearbox fault label	Bearing fault label	Gearbox and bearing fault label	Speed label	Time label
ResNet18	75.25	88.97	72.16	99.94	99.90

Table 6: All model experiment results.

Model	Training set accuracy (%)	Validation set accuracy (%)	Test set accuracy (%)	Training time (minute)
Wavelet DCNN	12.98	11.55	11.61	**89.5**
Wavelet DCNN (max)	61.87	59.10	59.71	333.90
1D-DCNN	0.23	0.21	0.10	673.38
MAF/MTF-ResNet18	**100**	**71.92**	**72.14**	261.03

In the table, the data with the best performance in each group are bold, including those with the highest accuracy and the least time, to highlight the advantages of the method used in this paper.

that in the process of building the model, the input of the neural network must be guaranteed to have the same size, so the strategy of adopting signals of different lengths adaptively according to the motor speed is used in this experiment. These signals are scaled in data preprocessing. In the process of scaling, the loss of information is inevitable, and especially, the loss of information from the one-dimensional time series signal may be more obvious, which ultimately leads to poor performance of the model.

5. Conclusion

In this paper, a GAF/MTF-ResNet model based on multisensor detection is proposed for coupled fault diagnosis of bearing-gear mechanical systems. The model first uses the GAF/MTF algorithm, which bypasses the gradient disappearance or gradient explosion problem existing in the classic neural network model, to extract multichannel time series signals of different types and lengths into two-dimensional RGB pictures through features. Then, we use the ResNet network to train and classify picture sets, and finally, we realize the diagnosis of different fault types.

In the process of building the model, the artificial selection of key parameters such as feature extraction was avoided, and the end-to-end operation of the model was realized. The experiment we designed has complex failure types, small failure sizes, and low motor speed, which requires the model to have high resolution capabilities. More importantly, we have designed mechanical faults in different transmission stages so that the experimental content can more realistically simulate the coupling faults in the actual complex working conditions of the mechanical system.

The GAF/MTF algorithm is used to realize the fusion and dimensionality reduction of multidimensional information based on the time-frequency domain information from the original signal. Moreover, this algorithm transforms the original time series classification problem into a two-dimensional image classification problem, which can avoid the shortcomings of the neural network's poor performance in processing time series. ResNet introduced afterwards has a deeper number of layers and better performance than other networks, and the training time has not increased excessively.

The experimental results show that the GAF/MTF-ResNet model has a training set accuracy of 100% and a test set accuracy of 72.14% for 520 types of experimental tags with different speeds, different sampling times, and different types of failures. Among them, the resolution of the speed and sampling time is close to 100%, and the resolution of gearbox failure and bearing failure is 75.25% and 88.97%, respectively. It has high application performance and further research value.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of Shandong Province, China (No. ZR2021ME026), the Natural Science Foundation of Shandong Province, China (No. ZR2020QE158), the Shandong Province Science and Technology Small and Medium Enterprise Innovation Capability Improvement Project, China (No. 2021TSGC1045), and the Key Research and Development Project of Qingdao, China (No. 21-38-04-0002).

References

[1] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, “Intelligent manufacturing in the context of industry 4.0: a review,” *Engineering*, vol. 3, no. 5, pp. 616–630, 2017.

[2] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, “Prognostics and health management design for rotary machinery systems-reviews, methodology and applications,” *Mechanical Systems and Signal Processing*, vol. 42, no. 1-2, pp. 314–334, 2014.

[3] D. S. Shah and V. N. Patel, “A review of dynamic modeling and fault identifications methods for rolling element bearing,” *Procedia Technology*, vol. 14, pp. 447–456, 2014.

[4] Y. Lei, J. Lin, M. J. Zuo, and Z. He, “Condition monitoring and fault diagnosis of planetary gearboxes: a review,” *Measurement*, vol. 48, pp. 292–305, 2014.

[5] Y. Wei, Y. Li, M. Xu, and W. Huang, “A review of early fault diagnosis approaches and their applications in rotating machinery,” *Entropy*, vol. 21, no. 4, p. 409, 2019.

[6] D. Cordoneanu and C. Nițu, “A review of fault diagnosis in mechatronics systems,” in *Proceedings of the International Conference of Mechatronics and Cyber-MixMechatronics – 2018*, pp. 173–184, Springer, Cham, Switzerland, 2019.

[7] X. Sun and X. Jia, “A fault diagnosis method of industrial robot rolling bearing based on data driven and random intuitive fuzzy decision,” *IEEE Access*, vol. 7, pp. 148764–148770, 2019.

[8] Z. Meng, G. Shi, and F. Wang, “Vibration response and fault characteristics analysis of gear based on time-varying mesh stiffness,” *Mechanism and Machine Theory*, vol. 148, Article ID 103786, 2020.

[9] S. Shao, S. McAleer, R. Yan, and P. Baldi, “Highly accurate machine fault diagnosis using deep transfer learning,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2446–2455, 2019.

[10] T. Wang, Q. Han, F. Chu, and Z. Feng, “Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: a review,” *Mechanical Systems and Signal Processing*, vol. 126, pp. 662–685, 2019.

[11] J. He, S. Yang, E. Papatheou, X. Xiong, H. Wan, and X. Gu, “Investigation of a multi-sensor data fusion technique for the fault diagnosis of gearboxes,” *Proceedings of the Institution of Mechanical Engineers - Part C: Journal of Mechanical Engineering Science*, vol. 233, no. 13, pp. 4764–4775, 2019.

[12] L. Kong, X. Peng, Y. Chen, P. Wang, and M. Xu, “Multi-sensor measurement and data fusion technology for manufacturing process monitoring: a literature review,” *Int. J. Extrem. Manuf*, vol. 2, no. 2, Article ID 022001, 2020.

[13] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, M. Pichler, and H. Efendic, “Fault detection in multi-sensor networks based on multivariate time-series models and orthogonal transformations,” *Information Fusion*, vol. 20, pp. 272–291, 2014.

[14] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: a review and new perspectives,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[15] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[16] J. Gu, Z. Wang, J. Kuen et al., “Recent advances in convolutional neural networks,” *Pattern Recognition*, vol. 77, pp. 354–377, 2018.

[17] W. Rawat and Z. Wang, “Deep convolutional neural networks for image classification: a comprehensive review,” *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, 2017.

[18] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, “A convolutional neural network for modelling sentences,” in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 655–665, Baltimore, MD, USA, June 2014.

[19] H.-C. Shin, H. R. Roth, M. Gao et al., “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[20] N. Aloysius and M. Geetha, “A review on deep convolutional neural networks,” in *Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP)*, pp. 0588–0592, Chennai, India, April 2017.

[21] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, “Efficient deep CNN-based fire detection and localization in video surveillance applications,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 7, pp. 1419–1434, 2019.

[22] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, “A comprehensive study on cross-view gait based human identification with deep CNNs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 2, pp. 209–226, 2017.

[23] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[24] Z. Wu, C. Shen, and A. van den Hengel, “Wider or deeper: revisiting the ResNet model for visual recognition,” *Pattern Recognition*, vol. 90, pp. 119–133, 2019.

[25] H. Jung, M.-K. Choi, J. Jung, J.-H. Lee, S. Kwon, and W. Y. Jung, “ResNet-based vehicle classification and localization in traffic surveillance systems,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 934–940, Honolulu, HI, USA, July 2017.

[26] S. Shao, R. Yan, Y. Lu, P. Wang, and R. X. Gao, “DCNN-based multi-signal induction motor fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 2658–2669, 2020.

[27] L. Jing, T. Wang, M. Zhao, and P. Wang, “An adaptive multisensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox,” *Sensors*, vol. 17, no. 2, p. 414, 2017.

[28] R. Liu, B. Yang, E. Zio, and X. Chen, “Artificial intelligence for fault diagnosis of rotating machinery: a review,” *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.

[29] D. Ho and R. B. Randall, “Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals,” *Mechanical Systems and Signal Processing*, vol. 14, no. 5, pp. 763–788, 2000.

[30] Z. Wang and T. Oates, “Imaging Time-Series to Improve Classification and Imputation,” in *Proceedings of the Presented at the 24th International Joint Conference on Artificial Intelligence*, Buenos Aires, Argentina, June 2015.

[31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[32] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Boston, MA, USA, June 2015.

[33] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015, https://arxiv.org/abs/1409.1556.

[34] K. Zhang, W. Zuo, and L. Zhang, “learning a single convolutional super-resolution network for multiple degradations,” pp. 3262–3271, 2018, https://arxiv.org/abs/1712.06116.

[35] S. Ruder, “An overview of gradient descent optimization algorithms,” 2017, https://arxiv.org/abs/1609.04747.

[36] Y. Zhang and W. Ye, “L2/3 regularization: convergence of iterative thresholding algorithm,” *Journal of Visual Communication and Image Representation*, vol. 33, pp. 350–357, 2015.

[37] C. Peng, T. Xiao, Z. Li et al., “MegDet: A Large Mini-Batch Object Detector,” in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6181–6189, Salt Lake City, UT, USA, June 2018.

[38] F. Chollet, *Keras: The Python Deep Learning library*, Astrophysics Source Code Library, 2018.

[39] M. Abadi, A. Agarwal, P. Barham, and E. Brevdo, “Tensorflow: large-scale machine learning on heterogeneous distributed systems,” 2016, https://arxiv.org/abs/1603.04467.

[40] Numba, “A High Performance Python Compiler,” 2022, https://numba.pydata.org.

[41] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, “On large-batch training for deep learning: generalization gap and sharp minima,” 2017, https://arxiv.org/abs/1609.04836.

[42] S. L. Smith, P.-J. Kindermans, C. Ying, and Q. V. Le, “Don’t decay the learning rate, increase the batch size,” 2018, https://arxiv.org/abs/1711.00489.