

Research Article

A Multifault Diagnosis Method of Gear Box Running on Edge Equipment

Xiao-ping Zhao,^{1,2} Yong-hong Zhang ,³ and Fan Shao³

¹School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China

²Network Monitoring Centre of Jiangsu Province, Nanjing University of Information Science and Technology, Nanjing, China

³School of Automation, Nanjing University of Information Science and Technology, Nanjing, China

Correspondence should be addressed to Yong-hong Zhang; zyh@nuist.edu.cn

Received 4 June 2020; Revised 5 July 2020; Accepted 13 July 2020; Published 3 August 2020

Academic Editor: Dou Wanchun

Copyright © 2020 Xiao-ping Zhao 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.

In recent years, a large number of edge computing devices have been used to monitor the operating state of industrial equipment and perform fault diagnosis analysis. Therefore, the fault diagnosis algorithm in the edge computing device is particularly important. With the increase in the number of device detection points and the sampling frequency, mechanical health monitoring has entered the era of big data. Edge computing can process and analyze data in real time or faster, making data processing closer to the source, rather than the external data center or cloud, which can shorten the delay time. After using 8 bits and 16 bits to quantify the deep measurement learning model, there is no obvious loss of accuracy compared with the original floating-point model, which shows that the model can be deployed and reasoned on the edge device, while ensuring real time. Compared with using servers for deployment, using edge devices not only reduces costs but also makes deployment more flexible.

1. Introduction

Gearboxes play an important role in modern machinery and equipment, which are gradually developing toward complexity, precision, and intelligence. A gearbox is composed of gears, bearings, a shaft and box body, and other parts. It has the characteristics of a compact structure, high transmission efficiency, long service life, and reliable operation. It is an indispensable general component in modern industry, including aviation, power systems, automobiles, and industrial machine tools. But because of its complex structure and high running speed in a harsh environment, it can easily break down, so the gearbox is an important factor in machine failure. Gears and bearings are two important parts of gearboxes, and they are prone to local faults due to fatigue, wear, and tear, leading to abnormal operation of gearboxes, which may cause economic losses, including damage to machines. But the performance and life of some bearings and gears are higher than expected. To repair or replace them regularly will waste manpower, material, and production resources. Using edge computing devices for diagnosis can

generate faster network service response, meeting the industry's basic needs in real-time business, application intelligence, security, and privacy protection. So, using edge computing equipment [1] to monitor and diagnose mechanical equipment can effectively avoid the above situation. Consequently, the research of efficient gearbox condition monitoring and fault identification technology is of great significance for ensuring production safety, preventing and avoiding major accidents.

2. Materials and Methods

2.1. Model Compression on Edge Computing. Fault diagnosis has three main steps: feature extraction, feature dimension reduction, and classification. Traditional feature extraction generally adopts artificial methods such as wavelet transforms, statistical features, and empirical mode decomposition. PCA, ICA, and self-encoders are used to reduce the dimension of features; Bayesian and nearest neighbor classifiers are most commonly used for classification. The process is shown in Figure 1.

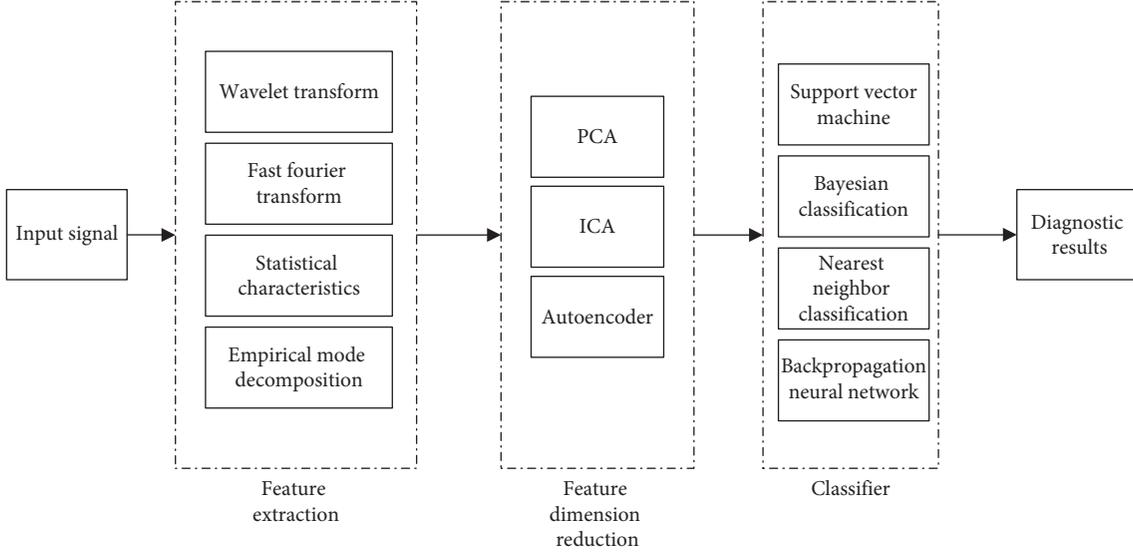


FIGURE 1: Fault diagnosis algorithm flow.

An increasing number of researchers are using neural networks for automatic extraction of fault features and feature dimensionality reduction and softmax for fault classification. Softmax is a generalization of a logistic classifier that mainly solves multiclassification problems. Assuming that the input sample in training data are x and the corresponding label is y , the probability of determining the sample as a class j is $p(y = j | x)$. The output of a K -class classifier will be a K -dimensional vector $h_{\theta}(x^{(i)})$. The elements of a vector sum to 1, and the category with the largest median of its elements is the prediction class, as shown in

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}, \quad (1)$$

where $\theta_1; \theta_2; \dots; \theta_k \in \mathfrak{R}^{n+1}$ is the model parameter and $1/\sum_{j=1}^k e^{\theta_j^T x^{(i)}}$ is the normalization function. The probability distribution is normalized so that the sum of all probabilities is 1.

Deep learning technology is developing rapidly, especially in the fields of image classification, target recognition, scene semantic analysis, and natural language processing. The deep neural network is demonstrably superior at processing complex data and predicting complex systems. Many experts and scholars in the field of mechanical failure have achieved good results when applying deep learning techniques to mechanical fault diagnosis.

Shao et al. [2], Wang et al. [3], and Chen et al. [4] used deep belief networks (DBNs) to diagnose the faults of rolling bearings and gearboxes, and the robustness and accuracy of

DBNs compared to some mainstream fault diagnosis methods was verified. Using DBN, Li et al. [5] studied information extraction and fusion under high background noise and achieved superior results compared to traditional methods. Convolutional neural network (CNN) has been applied to fault diagnosis to reduce the number of model parameters and improve calculation speed. CNN can be considered a neural network model for image processing. In the field of fault diagnosis, CNN can extract features and can be used to predict the classification of artificial features [6, 7]. Lu et al. [8] processed bearing running-health data based on a shallow CNN and extracted characteristic parameters and classifications of fault states. Zhang et al. [9] constructed a multilayer one-dimensional CNN and used the time-domain signals of bearing data to carry out fault diagnosis research, with good results. Wang et al. [10] and others used short-time Fourier transforms to convert collected motor-vibration signals to spectra and constructed a two-dimensional CNN for fault diagnosis, which achieved high diagnostic accuracy. Verstraete et al. [11] transformed the time domain signals of rolling bearings to time-frequency spectrograms by short-time Fourier transform, wavelet packet transform, and Hilbert-Huang transform, trained them by CNN, and studied the performance of the network by changing the size of the input time-frequency spectrograms and the denoising method. A one-dimensional signal can be transformed to a two-dimensional time-frequency diagram by time-frequency conversion, and then fault diagnosis by CNN can achieve good results. Zhang et al. [12] took the time-frequency spectrum of a rolling bearing vibration signal after Fourier transform as input. Using deep fully convolutional neural network (DFCNN), the vibration signal data of the rolling bearing rolling 2-3 turns is modeled by a large number of convolution layers. All effects reached 100%.

The above research shows that in-depth learning has strong adaptive feature extraction and classification ability in the face of large mechanical data tasks. These studies have

played a good role in the diagnosis of single target faults. However, in practice, gearboxes often have many kinds of faults simultaneously, and there are hundreds of combinations of complex faults. To solve this problem, a deep measurement learning model based on triplet loss is proposed in this paper. The multifault signals of gearbox bearings and gears are processed, and a variety of complex faults are simulated by using different bearings, gears, loads, and rotational speeds. Triplet loss is used as a loss function to optimize the model, thus efficiently completing the task of classifying complex faults [13–15].

Each triple [16, 17] is constructed by randomly selecting a sample from the training dataset as an anchor (x^a) and then randomly selecting a sample of the same type as the anchor called positive (x^p) and different classes of samples called negative (x^n). The anchor, positive, and negative constitute a complete triple. A neural network is trained for each sample in the triple, and the feature expressions of the three samples are denoted as $f(x_i^a)$, $f(x_i^p)$, $f(x_i^n)$.

The purpose of triplet loss is to make the distance between the characteristic expressions of x^a and x^p as small as possible, while making the distance between the characteristic expressions of x^a and x^n as large as possible. There is a minimum interval α between the distances (α is a hyperparameter, which can be set manually). As shown in Figure 2, the triplet learns to calculate the triplet loss multiple times to reduce the distance between similar samples and increase the distance between heterogeneous samples. In Euclidean space, a closer distance between two-fault data indicates greater similarity. The formula is

$$f(x_i^a) - f(x_i^p)_2 + \alpha < f(x_i^a) - f(x_i^n)_2, \quad (2)$$

where the subscript 2 represents the L2 paradigm and normalizes the data. The corresponding objective function is

$$\sum_i^n \left[f(x_i^a) - f(x_i^p)_2 - f(x_i^a) - f(x_i^n)_2 + \alpha \right]_+. \quad (3)$$

In the previous equation, the subscript + indicates that when the value in brackets is greater than zero, the loss is the value, and when it is less than zero, the loss is zero. It can be seen from the objective function that when the distance between the characteristic expressions of x^a and x^p is greater than that between the expressions of x^a and x^n minus α . If the value in brackets is greater than zero, the loss will occur. And, conversely, the loss will be zero. When the loss is not zero, all network parameters are adjusted by a backpropagation algorithm to optimize the features.

The choice of triplets is essentially a resampling process. There are many eligible triplets in the entire dataset. Assume a total of B fault data points of P types and K data points of each type. Then $B = P \times K$; that is, there are B original sample points, $K - 1$ similar sample points, and $B - K$ heterogeneous sample points. Therefore, the number of qualified triplets is $B \times (K - 1) \times (B - K)$. Some triplets satisfy the optimization goal; that is, the distance between x^a and x^p is much smaller than between x^a and x^n . Calculating these triplets is not helpful for optimizing the target. On the contrary, it will

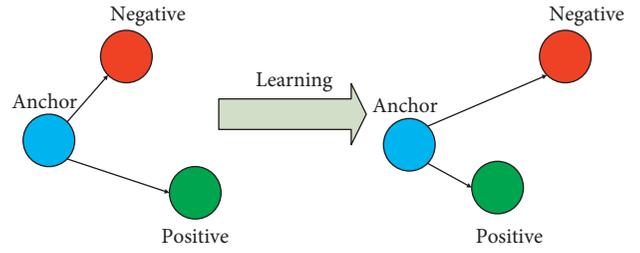


FIGURE 2: Triplet loss optimization goal.

reduce the training efficiency. Schroff et al. [18] mentioned that we should choose the combination which violates the optimization goal most seriously, where the distance between x^a and x^p is much larger than that between x^a and x^n . Whether the optimization goal is violated is found by calculating the Euclidean distance of the new features of embedding between the fault data, but the embedding updates the transformation each time; that is, the triplets that violate the optimization target may be different each time. If the triplet is reselected for each update, the training efficiency of the algorithm will be greatly reduced. There are currently two solutions to this issue.

- (1) For each n iterations, traverse the triplets to calculate the triplet loss on the latest training result until the network converges or iterates to a stop, instead of updating the triplets after each iteration.
- (2) Update the triplet loss online. A small number of triplets are selected to form a minibatch, in which all positive fault pairs are selected according to embedding at that time. Then, a negative fault that satisfies the condition that the distance between x^a and x^n is less than that between x^a and x^p is selected in the negative fault group. After calculating the triplet loss based on the above triplets, the embedding is updated and repeated until the network converges or the iteration stops.

The composite fault of the gear and bearing faults in the gearbox is taken as the research object of this paper, and the deep metric learning model is established. The model uses triplet loss as a loss function to construct a network. It maps features to Euclidean space and calculates the feature distance of similar samples and heterogeneous samples on Euclidean space. The closer the distance, the higher the similarity. By continually optimizing the triplet loss, the neural network continues to learn new features and bring the distances of similar samples ever closer, while the distance of heterogeneous samples increases.

Figure 3 shows the deep metric learning network model designed in this paper. The model consists of four layers (input layer, deep network, triple layer, and loss function calculation layer).

The task of the deep network layer is to extract the characteristics of the composite fault signal. There are many network structures to choose from, such as the convolutional neural network (CNN), long short-term memory neural network (LSTM), and fully connected neural networks.

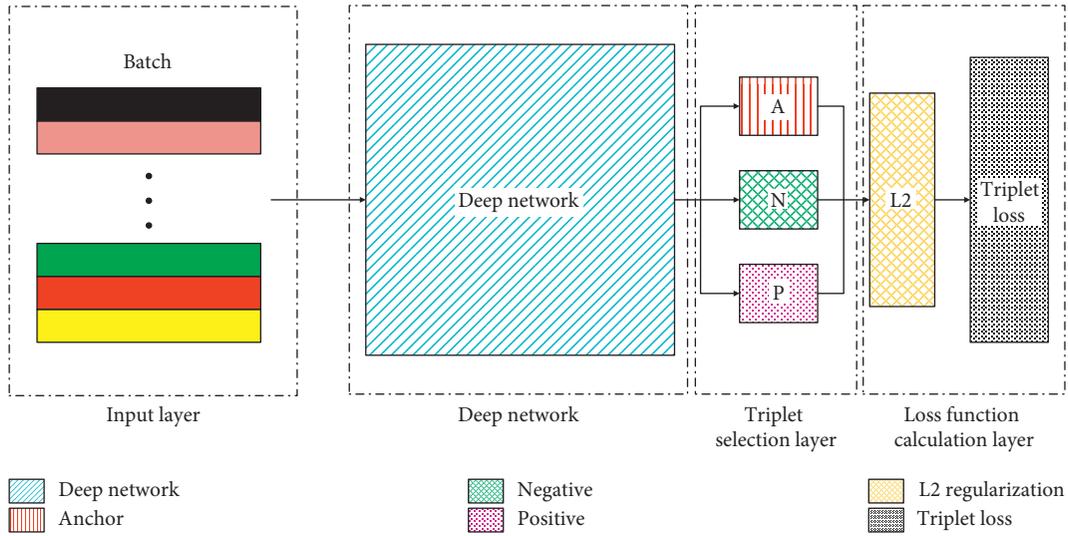


FIGURE 3: Structure of the deep metric learning model.

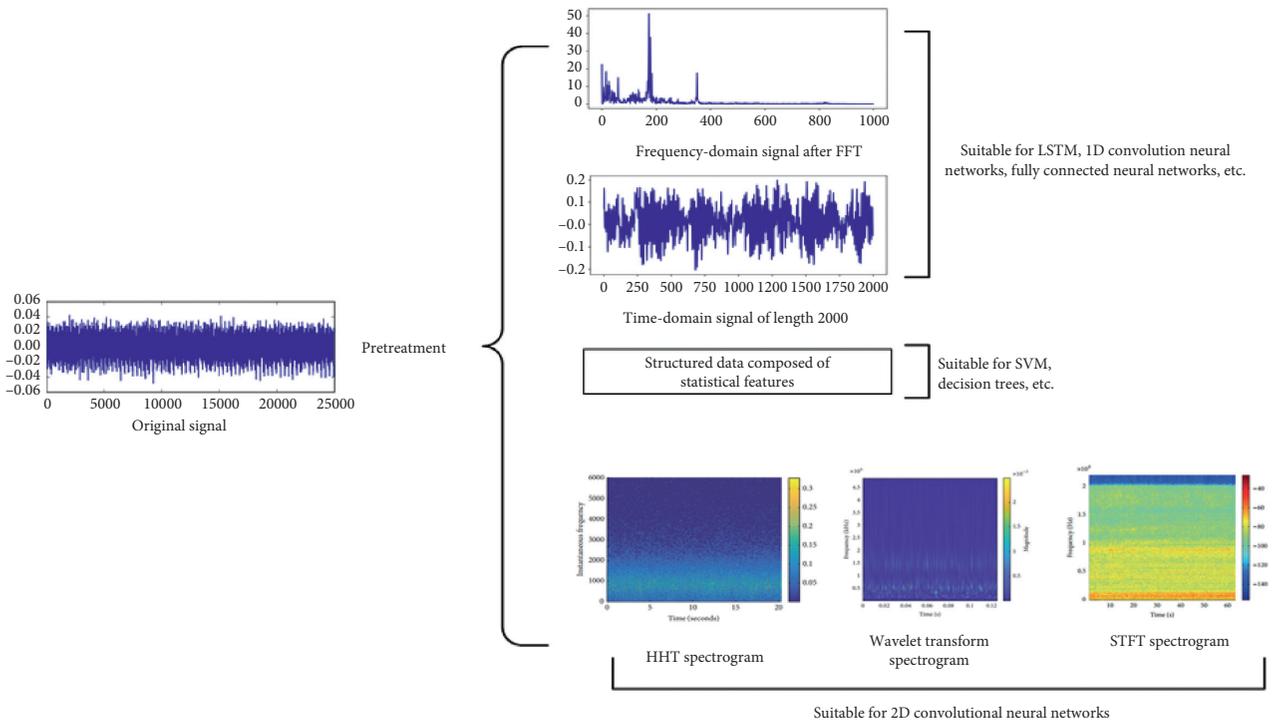


FIGURE 4: Pretreatment method.

The input forms corresponding to various networks are shown in Figure 4.

The literature [10, 19] shows that the effect of using the time domain signal directly as input for network training is not good. At the same time, the loss of the network cannot converge when the time domain signal is used as input, and the accuracy is only 30%. When diagnosing single faults of a gearbox [20], network training using the frequency domain signal as input data after fast Fourier transform achieved good results.

The composite fault signal is a nonstationary signal whose frequency varies with time. It is more complex than a

single fault signal. It is difficult to accurately diagnose the composite fault information in a gearbox using only the frequency domain signal, which only extracts the components of each frequency in the signal, and loses the time information of each frequency. Therefore, two signals with very different time domains may be the same as the spectrum.

Consider a nonstationary signal as a superposition of a series of short-term signals. In this paper, STFT [21] is used to divide the signal into several time intervals. On the basis of a traditional Fourier transform, the frequency spectrum is calculated by a sliding time window, and the frequency in a

certain time interval is determined. The time-frequency description of the signal is carried out so that the time information will not be lost. Assuming a nonstationary signal S_t , the short-time Fourier transform of S_t is defined as

$$\text{STFT}_s(t, w) = \int_{-\infty}^{+\infty} S(\tau)h(\tau - t)e^{-jw\tau} d\tau, \quad (4)$$

where t is the time translation parameter and $h(t)$ represents a window function centered on t , truncates the signal through the window function, and divides the signal into multiple segments.

The intercepted signal can be expressed as

$$S_t = S(\tau)h(\tau - t), \quad (5)$$

where S_t is a signal corresponding to the original signal for a fixed time t , and $S(T)$ is a signal whose execution time corresponds to T . A Fourier transform of S_t is used to obtain the spectrum of S_t :

$$S_t(w) = \frac{1}{\sqrt{2\pi}} \int e^{-jw\tau} S_t(\tau) d\tau = \frac{1}{\sqrt{2\pi}} \int e^{-jw\tau} S_t(\tau)h(\tau - t) d\tau. \quad (6)$$

By changing the size of the translation parameter t , the center position of the window function can be changed to obtain Fourier transforms at different times.

A different spectrum is obtained at each time interval, and the total of these spectra constitutes a time-frequency distribution, that is, a spectrogram.

After the short-time Fourier transform of the signal, the spectral energy relation of time t is

$$P_{\text{sp}}(t, w) = |S_t(w)|^2 = \left| \frac{1}{\sqrt{2\pi}} \int e^{-jw\tau} S_t(\tau)h(\tau - t) d\tau \right|^2. \quad (7)$$

As shown in Figure 5, the composite fault signal is converted into a time-frequency diagram through STFT and finally compressed to generate an 80×80 image for input to the network.

The structure of a convolutional network layer directly affects the effect of the network model, so the selection of network structure parameters is particularly important.

Figure 6 shows the structure of the convolutional network layer.

Table 1 shows the structural parameters adopted by the convolutional neural network layer. The network uses ReLU as the activation function and a uniform distribution when initializing network parameters. The range is $[-0.1, 0.1]$. The network uses the Adam optimizer with the learning rate set to 0.06. The dropout is set to 0.5 to avoid overfitting on the network. This structure ensures that the network can learn as many features as possible, and it prevents overfitting.

The convolutional neural network layer is followed by a triplet selection layer that shares the 32-dimensional features of the convolutional neural network output and generates a triplet for optimization. This article takes the online update of triplet loss mentioned in Section 3 to solve the triplet selection problem. The last layer is the loss function calculation layer, which normalizes the characteristics of the

output through the L2 paradigm and finally calculates the triplet loss.

Triplet loss is used as the loss function of the network (hyperparameter margin = 1). With the minimized triplet loss as the optimization goal of the network, the back-propagation (BP) algorithm is used to continuously update the weight of the neural network to train the optimal features.

In the new trained feature space, the distance between the data of different fault types is great, and the distance between data of the same fault type is small [5, 22, 23].

3. Operation Process on Edge Equipment

Figure 7 shows a flowchart of training and fault diagnosis of the deep measurement learning model based on triplet loss.

The steps are as follows.

The first step is sample collection. Through the short-time Fourier transform, the comprehensive fault data of gearbox is converted into time-frequency diagram.

The next is network training. The frequency domain signal is input to the network, and the deep network extracts a feature of each point of fault data. With the Euclidean distance between these features and the label of the fault data, the model can select the triplet according to the second scheme in Section 3 to calculate the triplet loss function. The network weights are updated by backpropagation and the above steps are repeated until the network converges or the iteration ends, saving the model parameters.

The third step is diagnosis. At the end of model training, forward propagation is used to obtain a feature of each fault category, which is the template needed to diagnose the unknown fault data. After deep network processing, the unknown fault data can also get a feature. The Euclidean distance between the feature of unknown data and the feature of the template can be calculated, and the minimum value of the Euclidean distance can be selected. The diagnosis result can be obtained by comparing the minimum value with the preset threshold.

The fourth step is to obtain the diagnosis results. The threshold is set because the unknown fault data probably indicate a completely new type of fault. Its characteristics are far from the Euclidean distance of each feature in the template, but the model selects a recent fault type as the output. If the minimum value is greater than the threshold, then the feature vector of the fault is stored in the template library, the label is recorded as unknown fault 1, and the diagnosis result outputs "unknown fault 1." When the unknown data of the fault type are encountered again, the deep metric learning model can accurately diagnose it. If the minimum value is less than the threshold, then the diagnosis result is output as the fault category of the template closest to the fault data (the smaller the Euclidean distance between fault data features is, the more similar they are).

4. Experiment and Analysis

4.1. Data Preprocessing Based on Edge Devices. Training deep learning networks requires a large amount of data support,

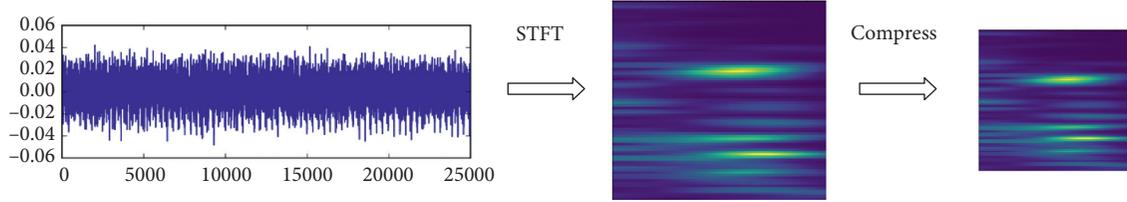


FIGURE 5: Pretreatment results.

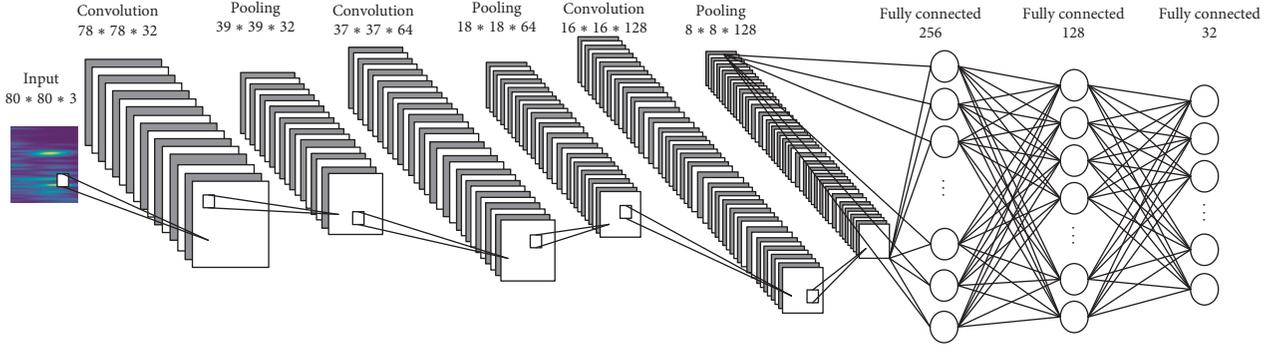


FIGURE 6: Structure of convolutional neural network.

TABLE 1: Network structure parameters.

Layers	Output size	Parameter
Conv2d	(None, 78, 78, 32)	320
Activation (ReLU)	(None, 78, 78, 32)	0
max_pooling2d	(None, 39, 39, 32)	0
Conv2d_1	(None, 37, 37, 64)	18, 496
Activation_1 (ReLU)	(None, 37, 37, 64)	0
max_pooling2d_1	(None, 18, 18, 64)	0
Conv2d_2	(None, 16, 16, 128)	73, 856
Activation_2 (ReLU)	(None, 16, 16, 128)	0
max_pooling2d_2	(None, 8, 8, 128)	0
Flatten	(None, 8192)	0
Dense	(None, 256)	2, 097, 408
Activation_3 (ReLU)	(None, 256)	0
Dropout	(None, 256) dropout = 0.5	0
Dense_1	(None, 32)	8, 224

and the quality of training data directly influences the output of the model. This paper used the power transmission fault diagnosis test bench (DDS) produced by Spectra Quest as the research object (as shown in Figure 8). When installing the acceleration sensor (SQI608A11-3F), this study refers to the acquisition method of the bearing data of Case Western Reserve University (<http://csegroups.case.edu/bearingdatacenter/home>). The acceleration sensor was mounted on the left and right sides of the fixed shaft of the gearbox by bolts (as shown in Figure 8 for sensors 1 and 2). The sampling frequency was 20 kHz, and the sampling time was 20 s.

The research object of this paper is gearbox compound fault. At present, there is no open dataset of compound fault. Therefore, we refer to the collection process of bearing data of Case Western Reserve University and carry out data collection on the Drivetrain Dynamics Simulator (DDS) by

ourselves. If there is an appropriate open dataset, we will conduct further research.

By replacing gears (e.g., those with missing teeth, broken teeth, eccentricity, excessive wear, and cracks) and bearings (inner ring fault, outer ring fault, rolling element fault, and compound fault) in the gearbox, 30 kinds of faults that may occur in the gearbox were simulated, as shown in Figure 9. To simulate a more realistic production environment, artificial noise pollution was carried out by tapping the gearbox or table with metal at random time, and the pollution signal accounted for about 5% of the total signal.

The picture of specific fault location and damage degree is shown in Figure 10.

At the same time, to increase the diversity of the sample, the speed was changed by controlling the driving motor of the front end when collecting data; the load was changed by controlling the load regulator, so as to simulate the type of working conditions that could occur in actual production. Each fault sample was collected at four motor speeds (1700, 1800, 3400, and 3800 rpm) and four loads (A, B, C, and D; see Table 2 for the load voltage and current of each load).

The time domain signals of the left and right channels were collected under each working condition to obtain 960 vibration signal files (30 multifault combination types \times 4 speeds \times 4 loads \times 2 channels). Each signal file contained 409,600 signal points.

The vibration signal file was divided randomly, and the 409,600 points in each signal file were evenly divided into 200 segments of 2,048 points. In order to fully study the recognition ability of the model, nine different data division methods are developed in this paper; the division is shown in Table 3.

In order to compare the diagnostic effect of convolution neural network + softmax classifier and the depth measurement learning model based on triplet loss, two kinds of

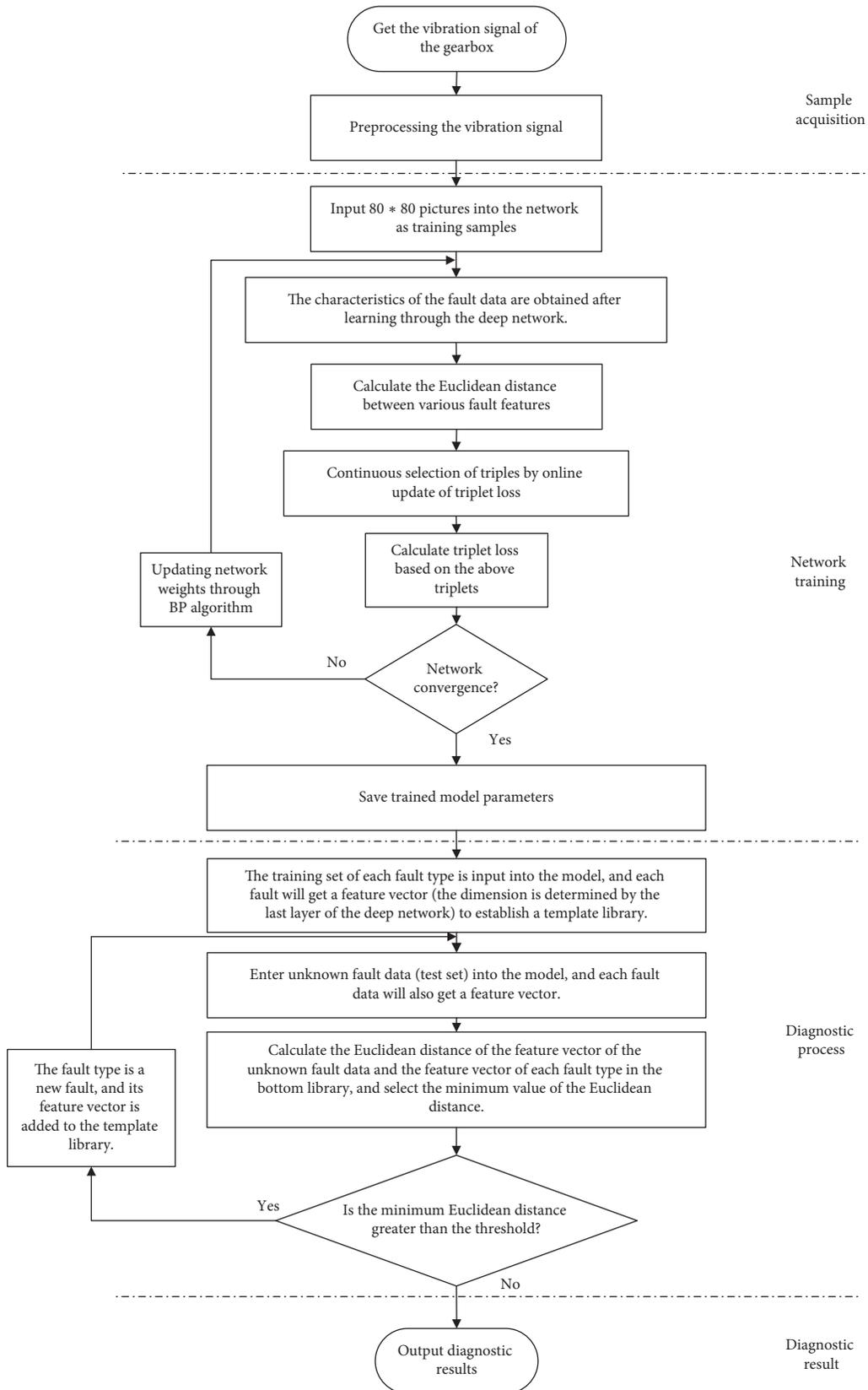


FIGURE 7: Model training and diagnostic process.

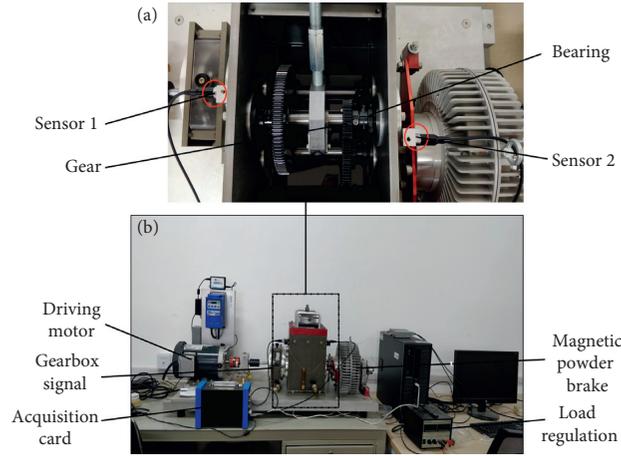


FIGURE 8: Power transmission fault diagnosis test bench.

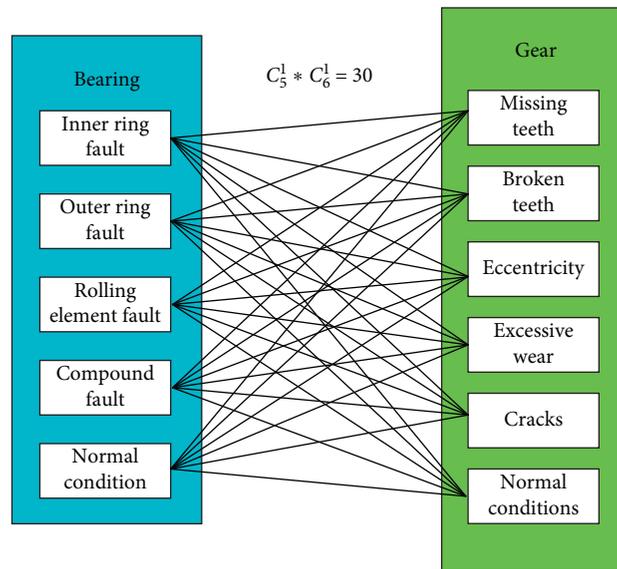


FIGURE 9: Combination types of bearing and gear faults.

labels are made (the structure and parameters of convolution network in the two models are identical).

- (1) The fault types were divided into 30 categories (five bearing faults and six gear faults). We created labels for training and test sets separately. The label corresponding to softmax must encode the label of the data. We used one hot encoding.
- (2) The fault types were divided into 30 categories (five bearing faults, six gear faults, four loads, and four speeds). We created separate labels for training and test sets. The label for entering the triplet loss was the ones matrix.

5. Results and Discussion

5.1. Experimental Verification. The data in Table 3 are input into the diagnostic model of convolution neural network+softmax classifier and the deep measurement

learning model based on triplet loss, respectively, for model training (the structure and parameters of convolution network are identical in the two models).

From the experimental results in Table 4, we can see that the accuracy of the diagnostic model of convolutional neural network+softmax classifier can reach 94.66% when the percentage cut data (i.e., sufficient training data, including various rotational speeds and loads) are normalized (i.e., Table 4, experiment 10), while the accuracy of the depth measurement learning model based on triplet loss can reach 97.73%.

When using the data missing from a certain load to train the network and using this load data for network testing (experiments 2–5, 11–14), the two models can obtain higher accuracy on the test set, but the deep measurement learning model is still better [24–26].

When using the data missing from a certain speed to train the network and using this speed data for network testing (experiments 6–9, 15–18), we can see from Figure 11

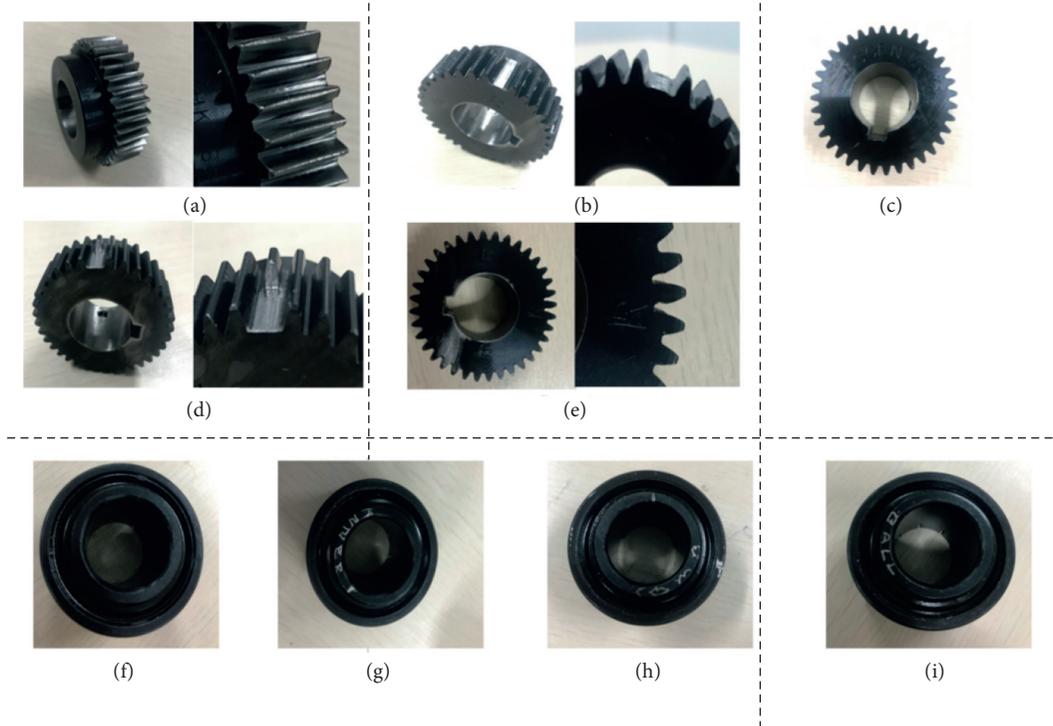


FIGURE 10: Bearing faulty bearings and gears. (a) Excessive wear. (b) Missing teeth. (c) Eccentricity. (d) Broken teeth. (e) Cracks. (f) Outer ring fault. (g) Inner ring fault. (h) Compound fault. (i) Rolling element fault.

TABLE 2: Type of load.

Load	Current (A)	Voltage (V)
A	0	0
B	0.37	4
C	0.56	6
D	0.75	8

TABLE 3: Segmentation of experimental data.

Dataset	Segmentation method	Training set		Test set		Total number of samples
		Types of signals	Number of samples	Types of signals	Number of samples	
A	Segmentation by percentage	Random 75%	144,000	Remaining 25%	48,000	192,000
B	Segmentation by load	B, C, D	144,000	A	48,000	192,000
C	Segmentation by load	A, C, D	144,000	B	48,000	192,000
D	Segmentation by load	A, B, D	144,000	C	48,000	192,000
E	Segmentation by load	A, B, C	144,000	D	48,000	192,000
F	Segmentation by speed	1800, 3400, 3800	144,000	1700	48,000	192,000
G	Segmentation by speed	1700, 3400, 3800	144,000	1800	48,000	192,000
H	Segmentation by speed	1700, 1800, 3800	144,000	3400	48,000	192,000
I	Segmentation by speed	1700, 1800, 3400	144,000	3800	48,000	192,000

TABLE 4: List of the highest test set accuracy achieved by each experiment.

Experiment	Data	Normalization	Accuracy of test set	
			Using softmax	Using triplet loss
1	A	No	0.93942	0.9611
2	B	No	0.89214	0.9437
3	C	No	0.90307	0.9489
4	D	No	0.91742	0.9509
5	E	No	0.92573	0.9593
6	F	No	0.68717	0.8334
7	G	No	0.67384	0.8376
8	H	No	0.48658	0.8197
9	I	No	0.38966	0.8168
10	A	Yes	<u>0.94668</u>	0.9773
11	B	Yes	0.91268	0.9615
12	C	Yes	0.92335	0.9703
13	D	Yes	0.93611	0.9756
14	E	Yes	0.93282	0.9632
15	F	Yes	<u>0.68371</u>	0.9124
16	G	Yes	<u>0.64178</u>	0.9133
17	H	Yes	<u>0.50438</u>	0.9098
18	I	Yes	<u>0.49335</u>	0.9045

Underlined italic values indicate extreme cases in the experimental results, and they are explained in detail below.

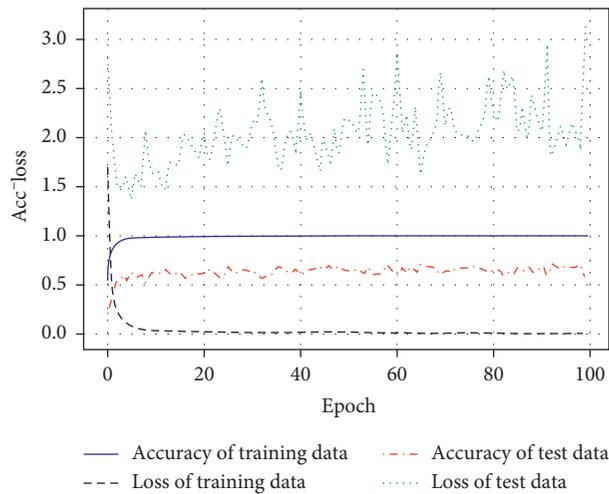


FIGURE 11: Accuracy of traditional network in experiment 10.

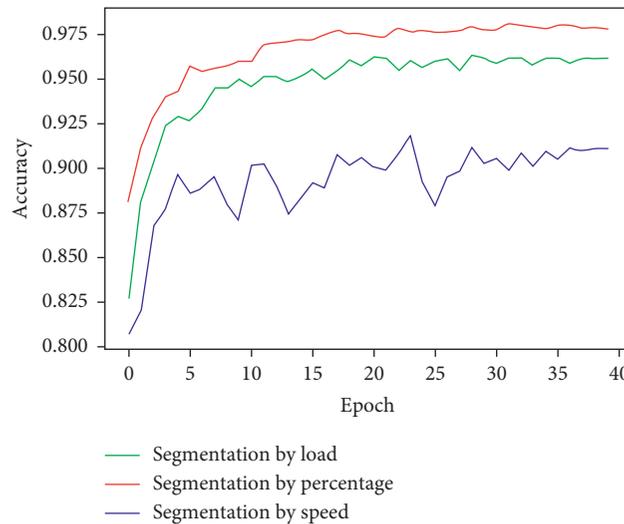


FIGURE 12: Accuracy of deep metric learning model with different data.

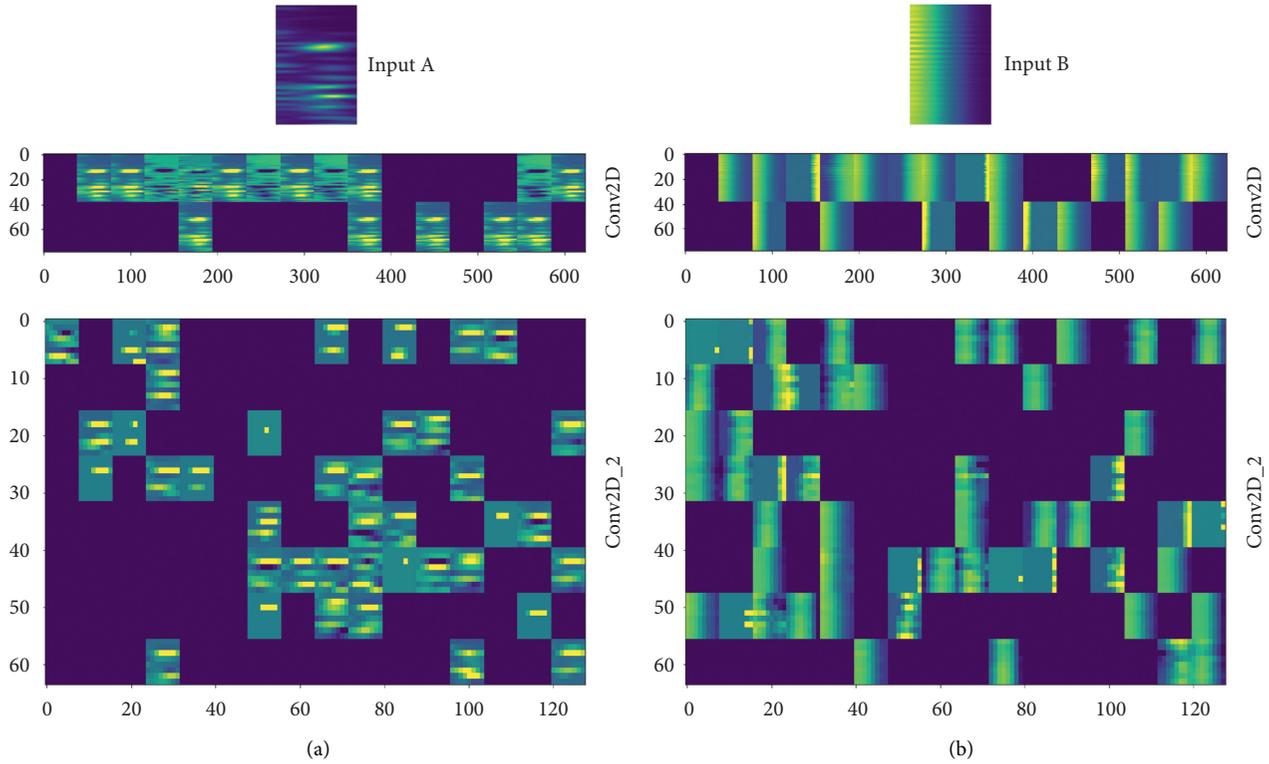


FIGURE 13: Feature visualization of faults A and B. (a) Input A. (b) Input B.

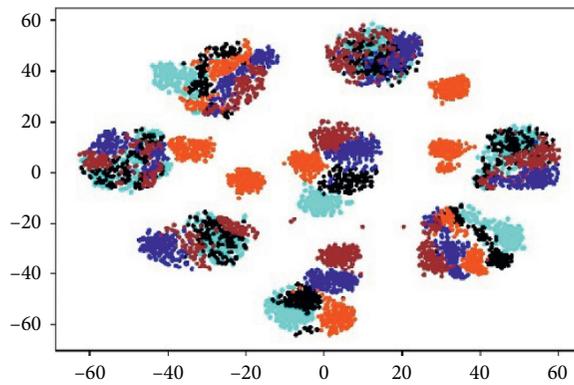


FIGURE 14: T-SNE visualization before processing.

that the diagnostic model of convolutional neural network + softmax classifier produces serious overfitting phenomenon, and the accuracy of the test set is very low.

But the deep metric learning model has strong generalization ability. As shown in Figure 12, the model can still obtain a high diagnostic accuracy even when some working condition data may be missing in practical application. Even in the experiment of missing motor speed data (experiments 15–18), it can achieve more than 90% accuracy.

The basic reasons why the method proposed in this paper can achieve higher diagnostic accuracy are as follows:

- (1) Vibration signals are transformed into time-frequency diagrams by STFT, and the features of time-frequency diagrams are extracted by convolutional

neural network; thus the frequency and time information of fault signals are effectively utilized. In order to verify the capability of the proposed method in feature extraction more intuitively, two different kinds of fault signals (A and B) are randomly selected and input into the network model with the highest accuracy obtained in Table 4, experiment 10. The output features of convolution layer Conv2d and convolution layer Conv2d_2 in Figure 6 are visualized [27]. The visualization results of the convolution kernel of signals A and B are shown in Figure 13.

- (2) Using triplet loss to measure the distance between different kinds of faults makes the distance between

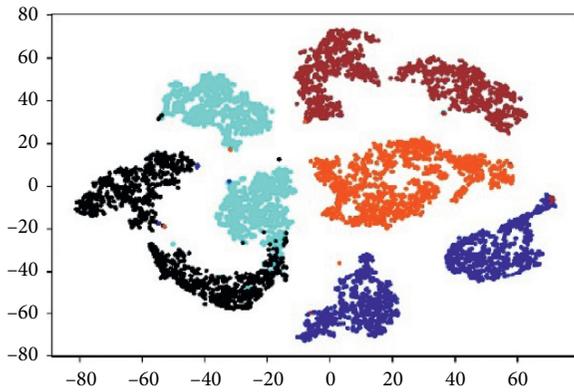


FIGURE 15: T-SNE visualization after processing.

similar fault features very close and that between different fault features very far, which makes the diagnosis more accurate and easier. However, the traditional convolutional neural network + softmax classifier model does not measure the distance between fault features. In order to prove that the model can make the fault data feature meet the distance between similar samples more and more close, at the same time, the distance between different samples is farther and farther [28]. Five fault types (A, B, C, D, and E) are selected randomly in this paper. 1600 data points ($5 * 1600 = 8000$ data) are selected randomly for each fault type. Figure 14 shows the original distribution of 8000 data points visualized by T-SNE.

Figure 15 shows the distribution of 8000 data points in t-SNE visualization after processing by deep measurement learning model.

6. Conclusions

In this paper, for the first time, the depth metric learning model is used to diagnose the faults of inner bearings and gears in gearbox at the same time, and the collected data are segmented in different ways to simulate the situation of missing some working condition data in practical application, so as to verify the performance of the network model. At the same time, the model of convolution neural network + softmax classifier is constructed to compare and classify. The conclusions are as follows:

- (1) When there is no missing data type, the depth metric learning model can extract features adaptively when dealing with complex faults of gearbox. The diagnostic accuracy of complex faults can reach 97.73%, which is higher than that of using convolution network + softmax classifier model.
- (2) When using the missing load data to train the network and use the missing load data to test the network, the accuracy of the convolution network + softmax classifier model is still $97 + 0.6\%$, while the accuracy of the convolution network + softmax classifier model is only $92 + 1\%$.

- (3) When using the data training network with missing speed and using the missing speed data for network testing, the convolution network + softmax classifier model produces serious overfitting phenomenon, and the accuracy of the test set is very low, only about 60%. However, the depth measurement learning model has not been fitted, and the accuracy of the test set is still higher than 90%.
- (4) The multifault dataset of gearbox collected in this paper has certain research value and can be used to evaluate the model for this kind of problem.

Data Availability

This article refers to the collection method of CWRU bearing data, sensor placement, etc., using Spectra Quest's powertrain fault diagnosis comprehensive test bench (Drivetrain Dynamics Simulator, DDS) as the test object to collect gearbox compound fault data. The storage address of the compound fault data in the gearbox is <https://pan.baidu.com/s/1zBJLV-O5v6nS9rfj16a5Kg> and the extraction code is 286w. These data include vibration signal data without any processing.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors thank LetPub (<http://www.letpub.com>) for its linguistic assistance during the preparation of this manuscript. The authors acknowledge the financial support of the National Science Foundation of China (Grant nos. 51505234, 51575283, and 51405241).

References

- [1] S. Li, S. Zhao, Y. Yuan, Q. Sun, and K. Zhang, "Dynamic security risk evaluation via hybrid bayesian risk graph in cyber-physical social systems," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 1133–1141, 2018.
- [2] H.-d. Shao, H.-k. Jiang, X. Zhang et al., "Rolling bearing fault diagnosis using an optimization deep belief network," *Measurement Science & Technology*, vol. 26, no. 11, 2015.
- [3] X.-q. Wang, L. Yan-feng, T. Rui et al., "Bearing fault diagnosis method based on Hilbert envelope spectrum and deep belief network," *Journal of Vibroengineering*, vol. 17, no. 3, pp. 1295–1308, 2015.
- [4] Z.-q. Chen, C. Li, and R. V. Sánchez, "Multi-layer neural network with deep belief network for gearbox fault diagnosis," *Journal of Vibroengineering*, vol. 17, no. 5, pp. 2379–2392, 2015.
- [5] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomputing*, vol. 168, no. C, pp. 119–127, 2015.
- [6] M. Bhadane and K. I. Ramachandran, "Bearing fault identification and classification with convolutional neural network," in *Proceedings of the International Conference on Circuit ,Power*

- and Computing Technologies (ICCPCT)*, pp. 1–5, IEEE, Kollam, India, April 2017.
- [7] H. Jeong, S. Park, S. Woo, and S. Lee, “Rotating machinery diagnostics using deep learning on orbit plot images,” *Procedia Manufacturing*, vol. 5, pp. 1107–1118, 2016.
 - [8] C. Lu, Z. Wang, and B. Zhou, “Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification,” *Advanced Engineering Informatics*, vol. 32, pp. 139–151, 2017.
 - [9] W. Zhang, G.-l. Peng, and C.-h. Li, *Advances in Intelligent Information Hiding and Multimedia Signal Processing*, pp. 77–84, Springer, Taiwan, 2016.
 - [10] L.-H. Wang, X.-P. Zhao, J.-X. Wu, Y.-Y. Xie, and Y.-H. Zhang, “Motor fault diagnosis based on short-time fourier transform and convolutional neural network,” *Chinese Journal of Mechanical Engineering*, vol. 30, no. 6, pp. 1357–1368, 2017.
 - [11] D. Verstraete, A. ferrada, E. L. Droguett, V. Meruane, and M. Modarres, “Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings,” *Shock and Vibration*, vol. 2017, Article ID 5067651, 17 pages, 2017.
 - [12] W. Zhang, F. Zhang, W. Chen et al., “Fault state recognition of rolling bearing based fully convolutional network,” *Computing in Science & Engineering*, vol. 21, no. 5, pp. 55–63, 2018.
 - [13] X. Xu, B. Shen, X. Yin et al., “Edge server quantification and placement for offloading social media services in industrial cognitive IoV,” *IEEE Transactions on Industrial Informatics*, 2020.
 - [14] X. Xu, X. Zhang, X. Liu, J. Jiang, L. Qi, and B. Zakirul Alam, “Adaptive computation offloading with edge for 5G-envisioned internet of connected vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2020.
 - [15] X. Xu, X. Liu, Z. Xu, F. Dai, X. Zhang, and L. Qi, “Trust-oriented IoT service placement for smart cities in edge computing,” *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4084–4091, 2020.
 - [16] R. Salakhutdinov and G. E. Hinton, “Replicated softmax: an undirected topic model,” in *Proceedings of the International Conference on Neural Information Processing Systems*, Curran Associates Inc., Vancouver, Canada, pp. 1607–1614, 2010.
 - [17] E. Hoffer and N. Ailon, “Deep metric learning using triplet network,” *Similarity-Based Pattern Recognition*, Springer International Publishing, Cham, Switzerland, 2015.
 - [18] F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: a unified embedding for face recognition and clustering,” pp. 815–823, 2015.
 - [19] L. I. U. Hong-mei, L. I. Lian-feng, and M. A. Jian, “Rolling bearing fault diagnosis based on STFT-deep learning and sound signals,” *Shock and Vibration*, vol. 2016, Article ID 6127479, 12 pages, 2016.
 - [20] X. Zhao, J. Wu, Y. Zhang, Y. Shi, and L. Wang, “Fault diagnosis of motor in frequency domain signal by stacked denoising auto-encoder,” *Computers, Materials & Continua*, vol. 57, no. 2, pp. 223–242, 2018.
 - [21] Y. Wang, J. Wang, and H. Huang, “Fault diagnosis of variable speed planetary gearbox based on nonlinear short time fourier transform order tracking,” *China Mechanical Engineering*, vol. 29, no. 14, pp. 54–61, 2018.
 - [22] D. Yu and L. Deng, “Deep learning and its applications to signal and information processing,” *IEEE Signal Processing Magazine*, vol. 28, no. 1, pp. 145–154, 2010.
 - [23] P. Tamilselvan and P. Wang, “Failure diagnosis using deep belief learning based health state classification,” *Reliability Engineering & System Safety*, vol. 115, no. 7, pp. 124–135, 2013.
 - [24] X. Xu, X. Zhang, H. Gao, Y. Xue, L. Qi, and W. Dou, “Be-Come: blockchain-enabled computation offloading for IoT in mobile edge computing,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4187–4195, 2020.
 - [25] L. Qi, X. Wang, X. Xu, W. Dou, and S. Li, “Privacy-aware cross-platform service recommendation based on enhanced locality-sensitive hashing,” *IEEE Transactions on Network Science and Engineering*, 2020.
 - [26] X. Xu, R. Mo, F. Dai, W. Lin, S. Wan, and W. Dou, “Dynamic resource provisioning with fault tolerance for data-intensive meteorological workflows in cloud,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, 2019.
 - [27] L. Qi, Q. He, F. Chen, X. Zhang, W. Dou, and Q. Ni, “Data-driven web APIs recommendation for building web applications,” *IEEE Transactions on Big Data*, 2020.
 - [28] C. P. Mbo’O and K. Hameyer, “Bearing damage diagnosis by means of the linear discriminant analysis of stator current feature,” in *Proceedings of the 2015 IEEE 10th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, IEEE, Guarda, Portugal, September 2015.