

## Research Article

# A Novel Sparse Enhancement Neural Network for Rolling Bearing Fault Diagnosis

Yong Zhang,<sup>1</sup> Junjie Ye ,<sup>1,2</sup> Wenhui Yang,<sup>1</sup> Jinwang Shi,<sup>1</sup> Wangpeng He,<sup>2</sup> and Gaigai Cai<sup>1</sup>

<sup>1</sup>Research Center for Applied Mechanics, Key Laboratory of Ministry of Education for Electronic Equipment Structure Design, Xidian University, Xi'an 710071, China

<sup>2</sup>Shaanxi Key Laboratory of Space Extreme Detection, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Junjie Ye; ronkey6000@sina.com

Received 30 June 2022; Revised 9 August 2022; Accepted 19 September 2022; Published 6 October 2022

Academic Editor: Yi Qin

Copyright © 2022 Yong Zhang 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.

To ensure the operational reliability of machinery, rolling bearings exposed to complex and poor conditions should be monitored in real-time. Traditional bearing fault diagnosis methods are always dependent on signal analysis and feature extraction, which are complex and time-consuming. Deep learning method exhibits a good ability in extracting the fault feature, while it is limited to noise pollution and insufficient sample data during the training procedure. In this study, a new sparse enhancement neural network based on generalized minimax-concave penalty and convolutional neural network is proposed to capture fault features automatically. To this end, the generalized minimax-concave penalty is first employed to expand the dataset by pollution data denoise and sparse enhancement of the insufficient samples. Second, the amplified dataset is employed to train the fault classification. By employing the datasets of drive end and fan end derived from the Case Western Reserve University (CWRU), a good prediction accuracy can be found in fault diagnosis for rolling bearings.

## 1. Introduction

With the rapid development of modern industry, rotating machinery has been widely used in rail transit, wind power generation, aerospace, etc. Bearing, one of the key components in mechanical equipment, is always stated at a high-intensity working state [1, 2]. Due to the complex and poor working conditions, the bearing is prone to wear, crack, and even fracture, which may result in a huge economic loss [3, 4]. In order to improve the reliability and economic benefits of the mechanical equipment, it is critical to identify bearing faults in time.

Fault diagnosis technology, which is a benefit to identifying health status, shows a good ability in preventing catastrophic accidents caused as a result of bearing fault. Traditional fault diagnosis techniques, such as wavelet analysis [5], variational modal decomposition [6], and singular value decomposition [7], have been widely used in fault identification which are time-consuming and laborious and have a low recognition rate. With the advent of the era of

big data, bearing fault diagnosis algorithms driven by massive data have become a research hotspot in recent years, especially from a perspective transfer of deep learning from image processing to fault identification [8, 9]. The procedure can be summarized in the following two steps: (1) a two-dimensional time-frequency diagram was first constructed according to the acquired one-dimensional vibration signal by employing signal processing methods, such as a wavelet packet [10] or discrete wavelet transform [11], and (2) the processed vibration signal was reconstructed and further divided into the train set and the test set, which were then imported into the deep neural network after normalization. Moreover, bearing fault classification was performed by using a convolutional neural network (CNN) model [12]. Guo et al. [13] obtained time-frequency feature maps from vibration signals by using a continuous wavelet transform scalogram. The feature maps were then used as the model input, which was a LeNet-5 model for bearing fault diagnosis. Xu et al. [14] used the continuous wavelet transform to transform the time-domain vibration signals into two-

dimensional grayscale images with rich fault information. On this basis, a CNN model based on LeNet-5 was constructed to automatically diagnose bearing faults in the images. These studies represent a good ability in bearing fault, but restrict to the length of training time and the complex of network structure.

Bearing fault signals represent obvious nonlinear and nonstationary properties due to the mixture interference with various noises. Some scholars have noted that it is hard to directly extract fault features from the acquired original signal and the model training time is too long. In order to improve the diagnosis effectiveness, Che et al. [15] proposed an intelligent diagnosis model based on the stack denoising autoencoder (SDAE) and CNN according to vibration characteristics of the rolling bearing with a strong noise interference, wherein the SDAE was used to process time series data coupled with the multidimensional noisy interference. Qiao et al. [16] used a one-dimensional convolution layer to suppress the background noise of the rotating machine before extracting the multiscale features. In order to eliminate the interference of bearing fault noise, Jiang et al. [17] extracted fault features by using the sensitivity to the impulse of spectrum kurtosis. It should be noted that the constructed one-dimensional CNN network has many complex structural layers and naturally the network parameters increase sharply. However, the designed complex network did not convert into a huge improvement in the computational efficiency.

At present, the convolutional layer in the deep learning method always acts as a filter (denoising) in processing the data sequence signals. In order to fundamentally solve fault signals disturbed by a strong interference or noise, feature extraction based on the sparse representation has been used in mechanical fault diagnosis due to its good ability in extracting fault features [18, 19]. Selesnick [20] designed a generalized minimum concave (GMC) penalty to overcome the underestimation of the amplitude components in traditional algorithms. Meanwhile, the GMC penalty can be effectively used to induce sparsity and to keep the convexity of the cost function, which exhibits a good feature in reducing noise interference during the mechanical fault diagnosis [21–23]. Cai et al. [23] proposed a new reweight generalized minimax-concave sparse regulation to extract the repetitive transients according to the characteristic of the fault rotating machinery. However, it is hard for the GMC to automatically identify the weak faults, and a long post-processing procedure is required.

In this study, a new sparse enhancement neural network combining the GMC penalty and the deep CNN is proposed to identify rolling-bearing faults. The main work of this study is summarized in the next two steps: (1) the GMC

algorithm is used to filter and denoise the interference or noise signals in advance. The denoising procedure is equivalent to a convolutional layer, which is a facility to identify the fault feature, as well as to sparsely enhance the insufficient samples to expand the dataset, and (2) to improve diagnosis accuracy and reduce network depth and training time, the deep learning network is redesigned. The remainder of this study is organized as follows: Section 2 is dedicated to the theory scheme of the sparse algorithm (GMC) and CNN. In Section 3, the theoretical scheme of the deep learning model is artfully designed to identify bearing fault diagnosis. In Section 4, the proposed method exhibits a good ability in bearing fault diagnosis by comparing it with other theoretical methods. Conclusions are drawn in Section 5.

## 2. Theoretical Background

Here, an intelligent diagnosis method for rolling bearing fault is proposed. First, the sparse GMC algorithm is used to denoise the original vibration signal. Second, feature vectors of the sequence signals are extracted by one-dimensional convolutional neural networks (1DCNN) to classify and identify bearing faults.

*2.1. The GMC Algorithm.* During the operation of rolling bearings, the measured data will be seriously interfered by external noise. According to the transient characteristic components and noise characteristics of the vibration signals  $y \in \mathbb{R}^M$ , the GMC model can be established as follows:

$$y = \psi x + n, \quad (1)$$

where  $\psi x \in \mathbb{R}^M$  and  $n \in \mathbb{R}^M$  indicate the pure data and the noise interference, respectively. The matrix  $\psi \in \mathbb{R}^{M \times N}$  represents the dictionary library. The GMC is one of the nonconvex penalties for the sparse-regularized linear least squares, the value of which maintains the convexity of the cost function to be minimized and promotes stronger sparsity than the  $L1$  norm [24]. The function  $F(x, v)$  can be formulated as

$$F(x, v) = \frac{1}{2} \|y - \psi x\|_2^2 + \lambda \varphi(x), \quad (2)$$

where  $\lambda$  ( $\lambda > 0$ ) is a regularization parameter.  $\varphi(x)$  is the GMC penalty, which can be defined as

$$\varphi(x) = \|x\|_1 - \min_{v \in \mathbb{R}^N} \left\{ \|v\|_1 + \frac{1}{2} \|B(x - v)\|_2^2 \right\}. \quad (3)$$

When  $B$  satisfied  $B^T B \leq 1/\lambda \psi^T \psi$ , the optimization problem can be rewritten as a saddle-point problem,

$$(w^{\text{opt}}, u^{\text{opt}}) = \arg \min_{x \in \mathbb{R}^N} \max_{v \in \mathbb{R}^N} \left\{ \frac{1}{2} \|y - \psi x\|_2^2 + \lambda \|x\|_1 - \lambda \|v\|_1 - \frac{\lambda}{2} \|B(x - v)\|_2^2 \right\}. \quad (4)$$

Here, the GMC algorithm is used to denoise the acquired sequence signal in advance and it is also used to highlight the sample features. Sample points of the randomly intercepted signal data are equal to 4096 as shown in Figure 1. To demonstrate the sparse enhancement effects of the GMC algorithm, the range of intercepted data overlaps and the spacing are both different. It can be easily discerned that the GMC exhibit a huge contribution to noise reduction and fault feature enhancement.

**2.2. Deep Learning Model Framework.** At present, 1DCNN is widely used for processing time series data, which mainly contains the convolution, pooling, and full connection layers [25]. The network structure is shown in Figure 2. The convolution operation is first executed on the input sample data by employing the weight sharing of the convolution kernel. Moreover, the corresponding feature vectors are extracted and its mathematical model can be written as

$$x_j^l = \sum_{i=1}^M x_i^{l-1} \otimes H_{ij}^l + b_j^l, \quad (5)$$

where  $x_j^l$  and  $x_i^{l-1}$  indicate the corresponding eigenvectors of the output and the input sample data, respectively.  $H_{ij}^l$  and  $b_j^l$  are the convolution kernel and the offset term, respectively.  $\otimes$  is the convolution operation sign.

By reducing the output dimension of the previous layer and by retaining the important information, the pooling layer is introduced to reduce the required GPU occupancy and improve its calculation time. The pooling layer function can be expressed as

$$p_j^l = f(u(z_1, z_2) \cap x_j^{l-1}), \quad (6)$$

where  $f(\bullet)$  is the max pooling layer or the average pooling layer.  $u(z_1, z_2)$  is the pooled region.  $z_1$  and  $z_2$  are the widths of the corresponding pooled regions.  $x_j^{l-1}$  indicates the  $j$ th eigenvalue in the  $(l-1)$ th convolution layer.  $\cap$  is an overlap between  $u(z_1, z_2)$  and  $x_j^{l-1}$ .

Feature extraction is executed on a series of convolutional layers and pooling layers, and the acquired feature vectors are input into the full connection layer through the straightening process. The formula is expressed as follows:

$$y_j = \sigma\left((w_f)_{p_i}^T + b_f\right), \quad (7)$$

where  $\sigma(\bullet)$  and  $w_f$ , respectively, are the activation function and the weight matrix between adjacent layers.  $b_f$  is the offset term.

### 3. The Proposed Network Framework of the GMC-CNN Model

To accurately extract the fault features of rolling bearing with a high efficiency, a new network framework combining the GMC and the CNN is proposed, known as the GMC-CNN model. This chapter mainly includes how to construct the GMC-CNN model, how to perform data sparse enhancement for bearing fault signals, and how to solve the problem of insufficient samples.

**3.1. Design of the Deep Learning Model.** The GMC algorithm has exhibited its excellent performance in fault feature extraction [21–23]; based on this, a new network framework of the GMC-CNN model is constructed as shown in Figure 3, wherein the GMC is equivalent to a convolutional layer to extract fault features in advance. At the same time, the depth of the neural network and also the training time are greatly reduced. Herein, we use the GMC algorithm to denoise, and then the extracted feature is further imported into the designed CNN for training. Moreover, bearing faults are classified by the Softmax classifier. In order to easily distinguish the fault feature, the rectified linear unit is designed as an activation function. On this basis, the output of the convolution layer is designed as an input of the activation function to carry out nonlinear mapping, which facilitates fault feature learning. To prevent overfitting of the model, we added the batch normalization layer and the dropout layer. In addition, the max pooling layer is added after the convolution layer, and the output feature vector of the convolutional layer is reduced through the pooling layer operation to reduce the number of parameters.

In order to retain maximum data texture features and to accelerate the mode convergence speed, the max pooling layer is added after the convolution layer, and the output feature vector of the convolutional layer is reduced through the pooling layer operation to reduce the number of parameters. Before the full-connection layer, a mean pooling layer is used to ensure the integrity of data information and to avoid overfitting. Meanwhile, the dropout layer is also used to prevent overfitting. In order to optimize the model, the cross entropy is selected as a loss function, and the formula is as follows:

$$J = -\frac{1}{N} \sum_{i=1}^N \sum_{i=1}^k y_i \cdot \log(p_i), \quad (8)$$

where  $N$  and  $k$  are the total number of samples and categories, respectively.  $y_i$  indicates the real labels.  $p_i$  indicates the probability of calculating category  $i$ .

**3.2. The Principle of Bearing Fault Diagnosis.** Aiming at a series of urgent problems, including few meaningful engineering data and a long feature extraction time by traditional diagnosis methods, a new diagnosis method based on the GMC-CNN is proposed. Figure 4 indicates the flow chart of the proposed method, which can be summarized in the following four steps:

Step 1: an experimental platform to collect vibration signal data of fault bearing is set up;

Step 2: the GMC is employed to denoise the acquired sequence data;

Step 3: random interception of the denoising data is performed, whose length is equal to 4096. This procedure facilitates for data enhancement and makes up for the shortcoming of a few meaningful failure data and extends it to the sample set. The intercepted data is further divided into a training set and a test set;

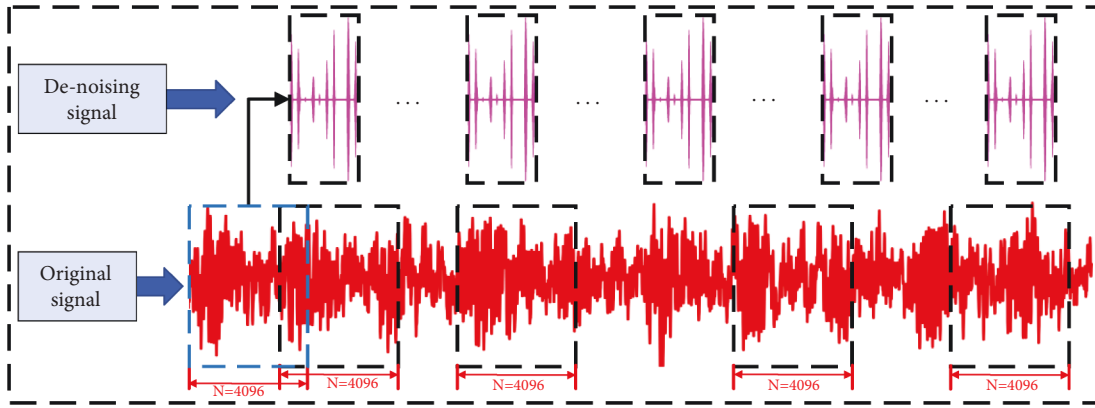


FIGURE 1: A schematic diagram of the random intercept signal processing.

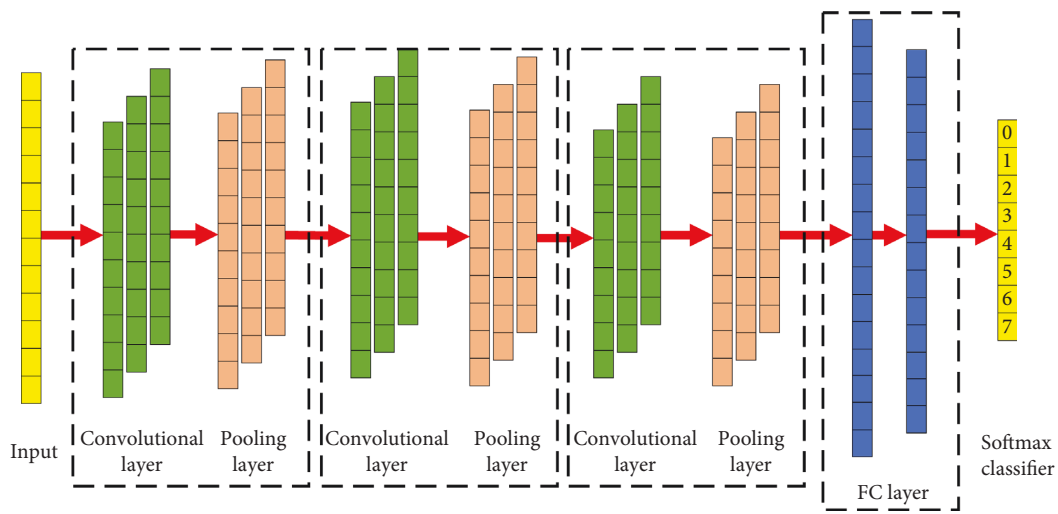


FIGURE 2: 1DCNN.

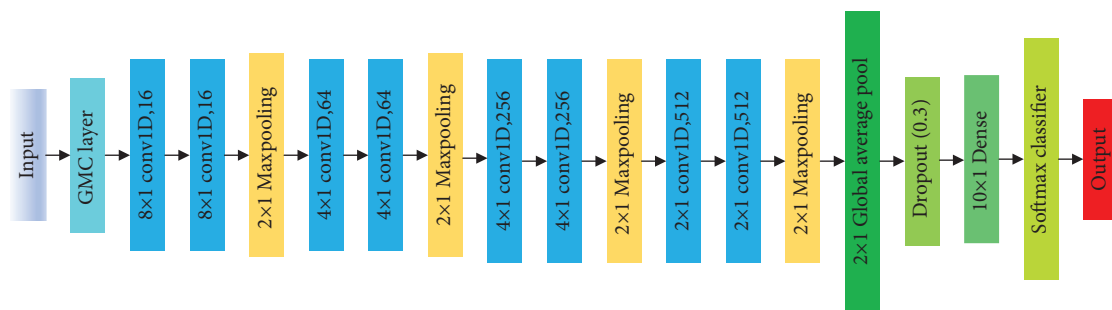


FIGURE 3: The framework of the proposed GMC-CNN model.

Step 4: importing the processed dataset into the GMC-CNN model for training. Moreover, the test set is also input into the trained model to validate its accuracy and efficiency in the fault identification;

Step 5: numerical result acquired by theoretical methods in the literature and the GMC-CNN model are both indicated for a comparison. Naturally, the training time and prediction accuracy can be easily observed to evaluate the proposed method.

**3.3. The Description of GMC-CNN Model Parameters.** The model parameters in this study are divided into two parts, the first part is the GMC algorithm, and the second part is the 1DCNN model. Firstly, the tunable  $Q$ -factor wavelet transform is used as the dictionary term for sparse noise reduction of the fault signal. The matrix  $B = \sqrt{\gamma/\lambda}A$  with  $0 \leq \gamma \leq 1$  is set to satisfy the convexity of  $F(x)$ . In general, the parameter is set between 0.5 and 0.8. Herein the parameter  $\gamma = 0.63$  is determined. The  $\lambda$  in the GMC algorithm was

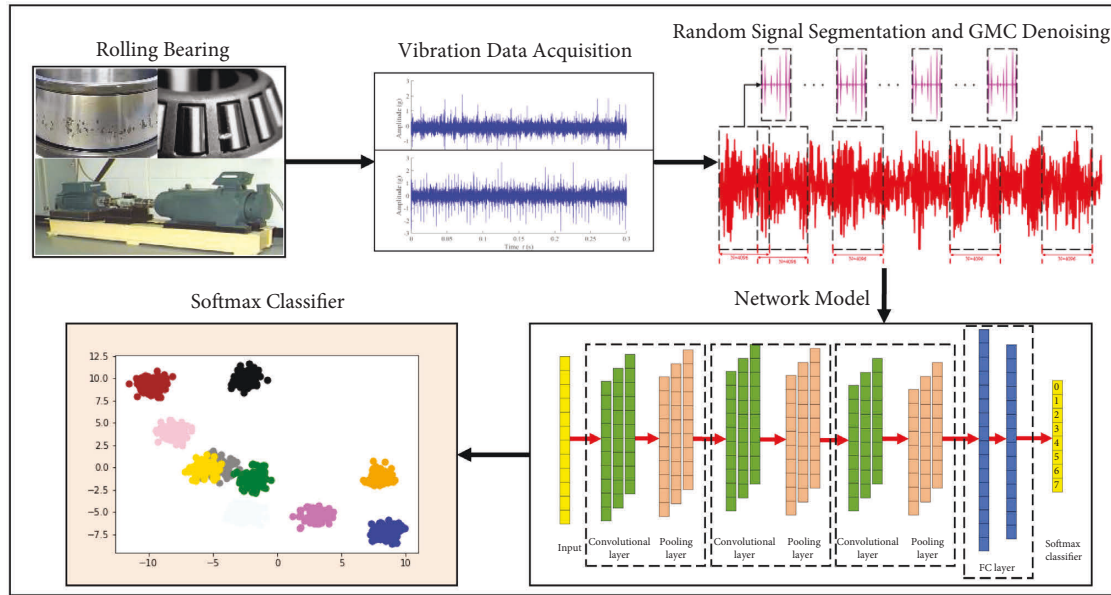


FIGURE 4: The flowchart of the bearing fault diagnosis.

adjusted by an adaptive strategy. The root mean square error and signal-to-noise ratio were both selected as the evaluation indexes, and the optimal  $\lambda = 0.53$  is determined.

In the CNN model, the parameter adjustment mainly contains the size of the convolution kernel, selection of the optimization algorithm, selection of the loss function, size of the pooling layer, the learning rate, the number of batch samples, the number of iteration rounds, and the proportion of dropout method. The sizes of the convolution kernel and the pooling layer, as well as the proportion of dropout layers, are shown in Figure 3. The optimization algorithm selects the Adam algorithm, in which the learning rate and other parameters select the default value. The numbers of batch samples and iteration rounds are 128 and 30, respectively.

#### 4. Validation of the Proposed Deep Learning Network Model

To verify the flexibility and utility of the constructed deep learning network model in bearing fault diagnosis, the public bearing dataset provided by the CWRU Bearing Data Center [26] is used. In this dataset, vibration signals are collected from a 2 hp electric motor, in which the accelerators are installed at the drive end and the fan end, respectively. Herein two sets of signal data are synchronously used and analyzed.

*4.1. Dataset A: The Drive End Data.* Three types of fault features, including inner race fault, ball fault, and outer race fault, always occurred in the drive end of the bearing. With full consideration of damage degree influences, damage sizes of 0.18 mm, 0.36 mm, and 0.54 mm for each fault type are considered. Therefore, ten health conditions, including one normal condition and nine fault conditions, are investigated in the dataset. Specific conditions are shown in Table 1.

TABLE 1: The description of ten health conditions and labels.

Name	Ten conditions and descriptions	Labels
Normal	The normal bearing	0
Ball_18	The ball fault with diameters of 0.18 mm	1
Ball_36	The ball fault with diameters of 0.36 mm	2
Ball_54	The ball fault with diameters of 0.54 mm	3
Inner_18	The inner race fault with diameters of 0.18 mm	4
Inner_36	The inner race fault with diameters of 0.36 mm	5
Inner_54	The inner race fault with diameters of 0.54 mm	6
Outer_18	The outer race fault with diameters of 0.18 mm	7
Outer_36	The outer race fault with diameters of 0.36 mm	8
Outer_54	The outer race fault with diameters of 0.54 mm	9

In order to obtain enough experimental samples and to ensure segmentation data with a high randomness and universality, the sample signal length of 96000 is randomly intercepted by 500 data groups, wherein the length of each data group is equal to 4096. This procedure is a facility to ensure that the vibration sample contains enough feature information. Figure 5 indicates a comparison between original data and denoising data by the GMC method, which are represented by a blue line and red line, respectively. It can be seen that data essential information and fault characteristics are both retained to a maximum during the denoising procedure. In general, a higher fault prediction accuracy always needs more sample data, which undoubtedly increases the model training time, while the model training time can also be effectively guaranteed for the sparse enhancement data, even if the dataset increases to 20000.

The randomly segmented training set and test set are separately imported into the GMC-CNN model, and the number of iterations is both set to 30. Figure 6(a) indicates a comparison of the prediction accuracy, wherein the blue line and the red line are, respectively, the accuracy of the train set and the test set. The prediction accuracies in the training set and the test set are 100% and 99.98%, respectively, which

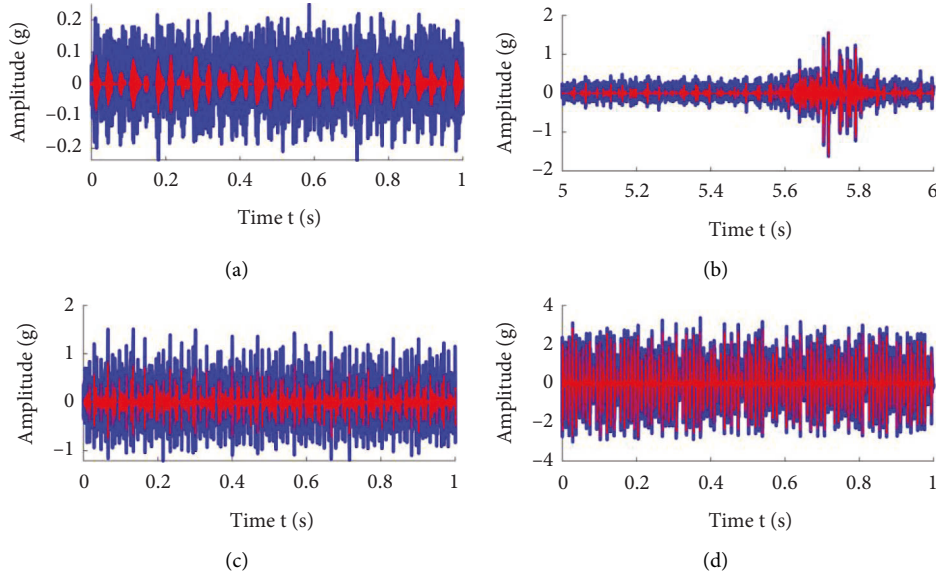


FIGURE 5: The drive end signal with and without noise: (a) normal data, (b) ball data, (c) inner race data, and (d) outer race data.

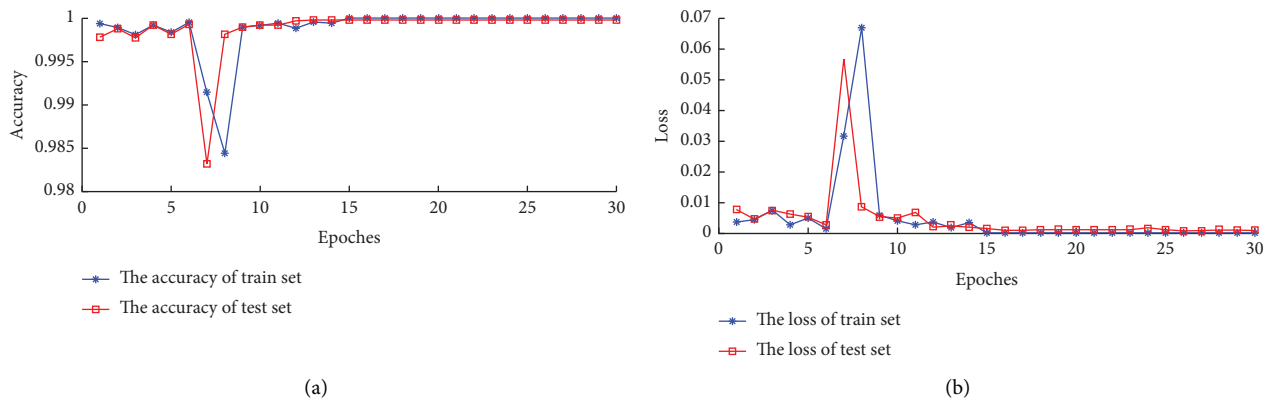


FIGURE 6: Results of the drive end data investigated by the proposed GMC-CNN model: (a) calculation accuracy and (b) the loss.

may be caused by the effect of GMC sparsity enhancement combined with the amplified sample set. It can be determined that the problem of insufficient sample is solved when the low amplitude noise is considered. A further comparison of the loss rate is shown in Figure 6(b). The loss rate in the training set shows a high consistency with that in the test set. In addition, the proposed method has little sign of overfitting and accuracy degradation.

The confusion matrix is indicated in Figure 7(a). It is revealed that only one judgment error can be found. Figure 7(b) is a two-dimensional graph of ten health conditions in the dataset. The clustering phenomenon of each damage type in bearing faults can be easily discerned, and each clustering group of the data represents a fault feature. In other words, the fault type can be easily determined with a high precision by the GMC-CNN model if the data are clean and clear.

**4.2. Dataset B: The Fan End Data.** Herein, ten identical damage types in the driver end are also considered in the fan end. In other words, three damage types in the fan end

combined with three damage sizes are considered in the example. Similar data processing procedure is indicated in Section 4.1 which is executed on the fan end data. The waveforms of the sequence signals before and after denoising are shown in Figure 8, wherein the red and blue lines represent vibration data and denoising data by the GMC method, respectively. It can be seen that the noise or interference components have been removed from the sampled signal.

As one of the most important evaluation indexes for the designed model, overfitting of the training samples has been paid special attention. Therefore, the average pooling layer and dropout layer are set up to prevent data overfitting during the study. It can be easily seen from Figure 9(a) that the overfitting is effectively controlled, and the training accuracy is maximum approximate to the test accuracy. So, when the iteration step is equal to 30, the prediction accuracy of the training set and test set are 100% and 99.97%, respectively. In addition, the loss rate of the fan end as shown in Figure 9(b) is small enough which can be ignored.

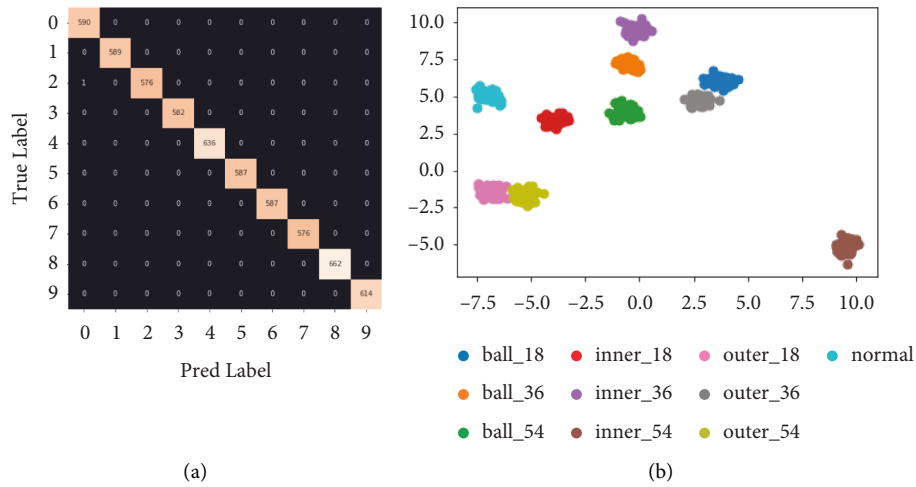


FIGURE 7: The representation of fault classification: (a) confusion matrix and (b) diagram of ten classifications.

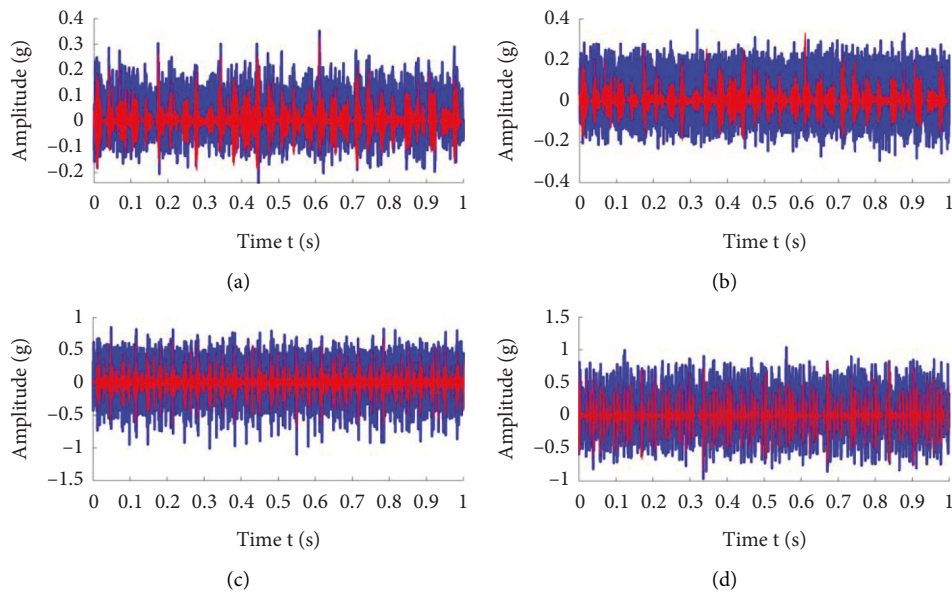


FIGURE 8: The fan end signal with and without noise: (a) normal data, (b) ball data, (c) inner race data, and (d) outer race data.

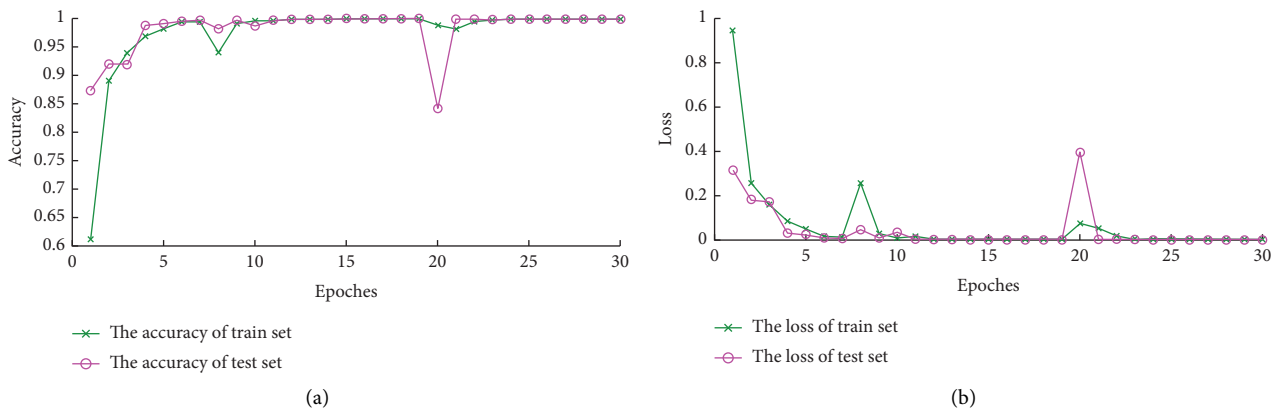


FIGURE 9: Results of the fan end data investigated by the proposed GMC-CNN model: (a) calculation accuracy and (b) the loss.

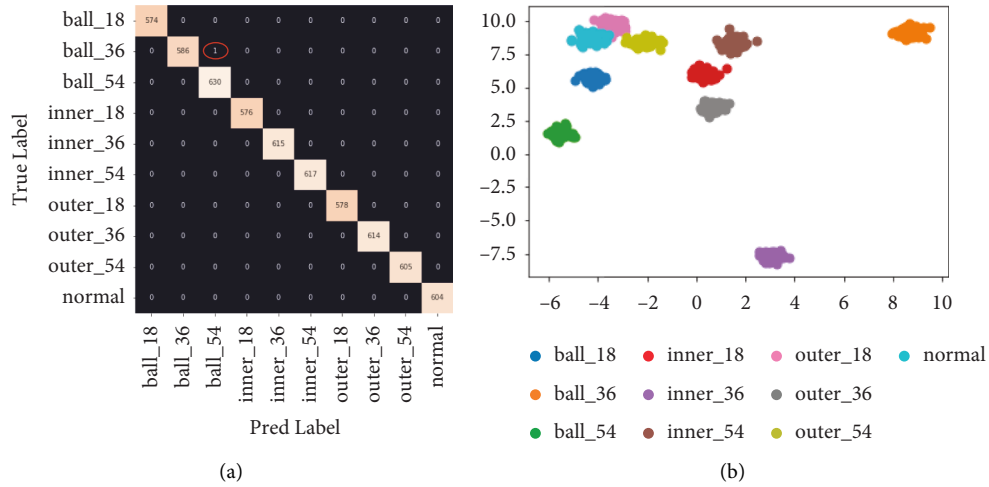


FIGURE 10: The representation of fault classification: (a) confusion matrix and (b) the diagram of ten classifications.

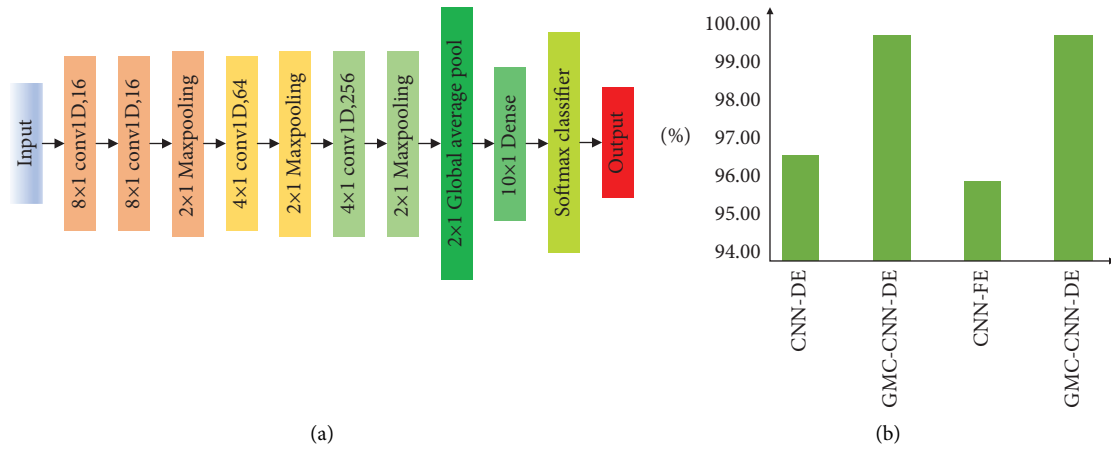


FIGURE 11: The structure and comparison of the model: (a) the network structure of the 1DCNN model and (b) a comparison between the 1DCNN model and the GMC-CNN model.

The confusion matrix is shown in Figure 10(a). A high fault prediction accuracy by the proposed method can be found, and only one error in the test set can be discerned. Figure 10(b) shows the distributions of ten damage types, and each color represents one type of damage. From the figure, each fault feature can be well clustered, which can be used for fault classification identification by utilizing the proposed method.

**4.3. A Comparison of the Theoretical Methods.** The traditional CNN model is shown in Figure 11(a), whose network structure is relatively simple. By employing a one-dimensional CNN model, it can be seen in Figure 11(b) that the fault recognition accuracies in drive end and fan end datasets are 96.79% and 96.11%, respectively. For a comparison, fault recognition by the proposed the GMC-CNN method is also shown in the figure. It is revealed that the GMC-CNN model shows a higher accuracy in bearing fault diagnosis, which effectively solved the problems of small samples and confusing data. In detail, the classification accuracy of the

proposed method increases by nearly 3% compared with the traditional CNN method. In detail, prediction accuracies by the GMC-CNN method are increased to 99.97% and 99.98%, respectively. At the same time, the GMC penalty function not only retains the convexity of the cost function but also represents the sparsity characteristic. In addition, sparse enhancement of the data can liberate the depth of the CNN model and can save the training time.

To verify the data enhancement by the GMC layer, a 2-dimension CNN model is also used for data comparison. Figure 12 shows the structural framework of each layer. Two-dimensional sliced data, whose dimension is  $64 \times 64$ , are imported into an identical CNN structure. During the investigation, the performance of AMD Ryzen 5 3600 6-Core Processor with 3.60 GHz CPU and NVIDIA GeForce GTX 1050 Ti is employed in running python 3.8.

The GMC-2DCNN model is further studied for a comparison, where the GMC is also used for data-sparse enhancement and data noise-cleaning. Numerical results as shown in Figures 13(a) and 13(c) indicate that the prediction accuracies in solving drive end data and fan end data are,



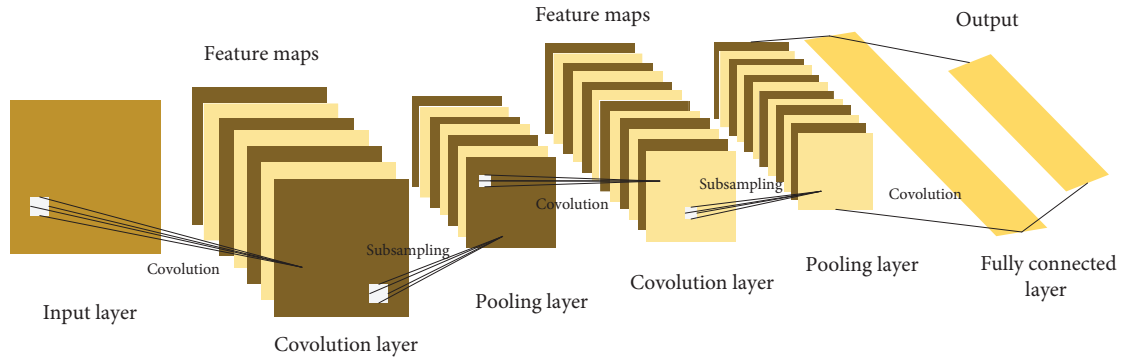


FIGURE 12: The structure of the 2DCNN model.

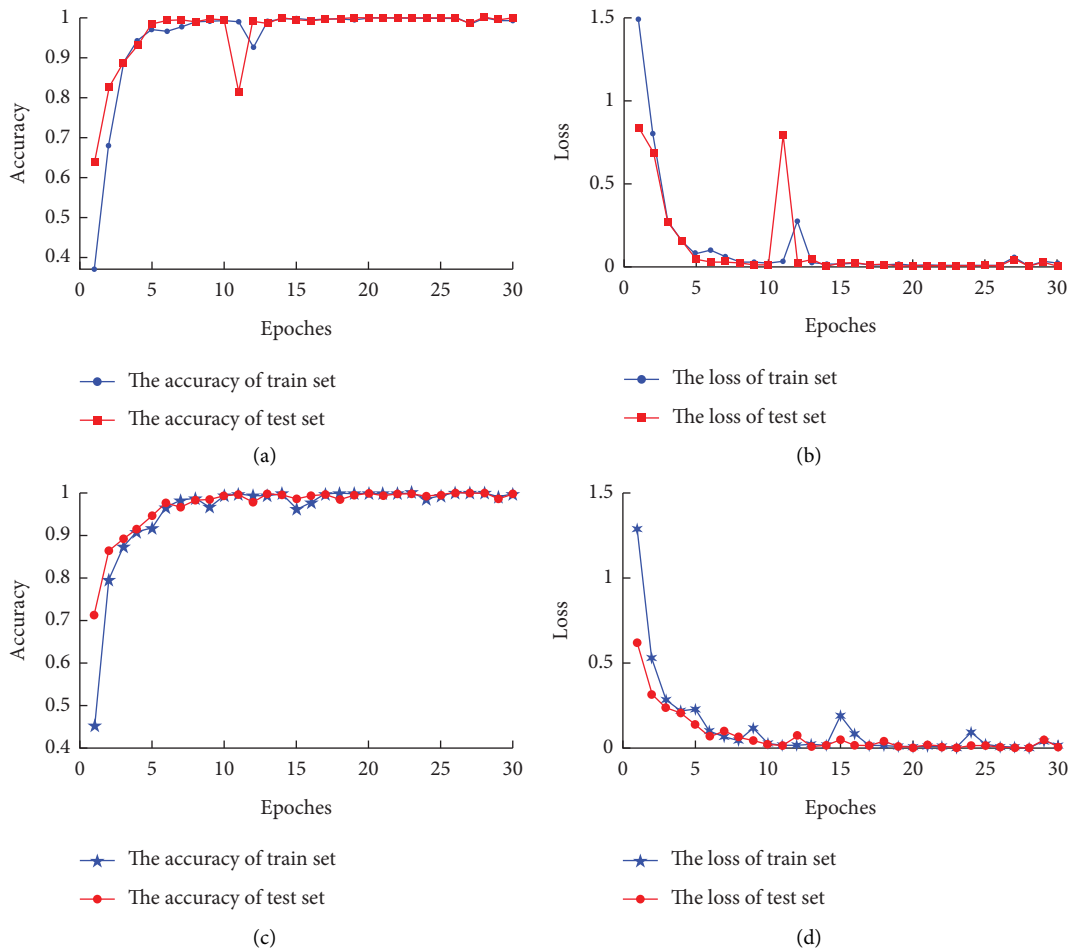


FIGURE 13: The representation of the drive end data and the fan end data on the GMC-2DCNN model: (a) the accuracy of the drive end data, (b) the loss of the drive end data, (c) the accuracy of the fan end data, and (d) the loss of the fan end data.

respectively, 99.97% and 99.85%. In addition, the GMC-2DCNN model has little overfitting phenomenon and identical accuracy can be found in the training set and the test set when the iteration step is equal to 15. Compared with the traditional one-dimensional model as shown in Figure 11(b), the accuracy of the model after data enhancement and data cleaning is significantly improved.

Figure 14 shows a comparison of precision and efficiency between the GMC-2DCNN model and the proposed GMC-

CNN model. In Figure 14(a), the two methods represent little difference in accuracy. Herein the advantage of the proposed method is illustrated from the perspective of time cost, in which the suffixes DE and FE denote drive end data and fan end data, respectively. It can be seen from Figure 14(b) that the training time required by the GMC-2DCNN model in drive end data and fan end data is equal to 782 s and 785 s, respectively. While the training time required by the proposed GMC-CNN model in drive end data

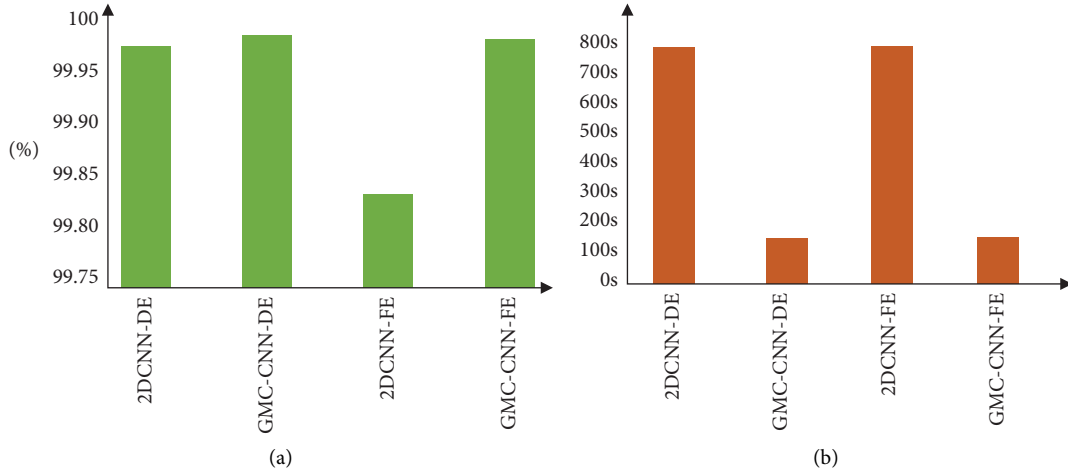


FIGURE 14: A comparison of accuracy and efficiency between the 2DCNN model and the proposed GMC-CNN model: (a) accuracy comparison and (b) time comparison.

TABLE 2: A comparison of the accuracy among different methods.

Model	SIRCNN [27]	1D-CLN [28]	D-CNN [29]	CNN-SVM [30]	ODCNN [31]	GMC-CNN-FE	GMC-CNN-DE
Accuracy	99.16%	99.93%	98.83% (s)	98.75% (s)	99.58%	99.97% (s)	99.98% (s)
Time	/	/	216.32	35.82	/	151	152

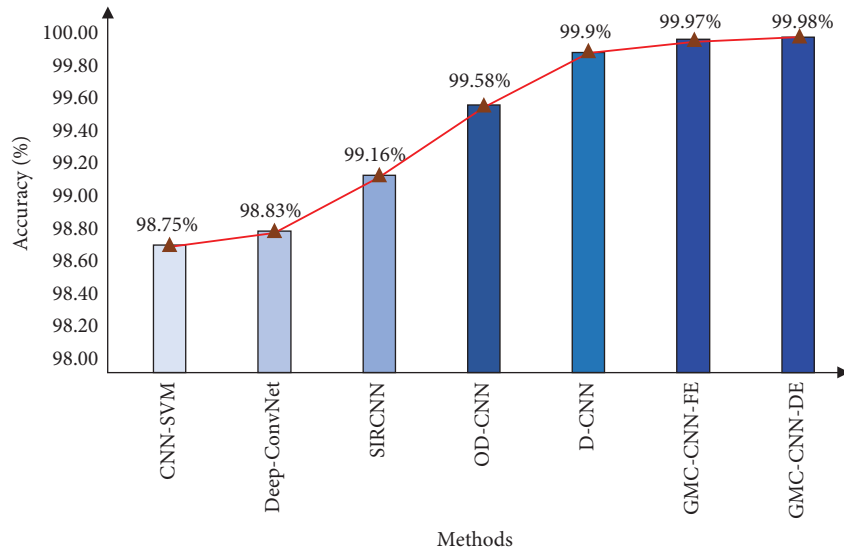


FIGURE 15: A comparison of prediction accuracy by numerical methods.

and fan end data is equal to 152 s and 151 s, respectively. It can be concluded that the proposed GMC-CNN model greatly improves the training time and reduce the operation cost. Moreover, the GMC used for data enhancement also improves the accuracy of fault classification to a certain extent.

For a further comparison of the solving accuracy, some other theoretical methods [27–31], such as the D-CNN model [29] and the CNN-SVM model [30] are considered. Table 2 indicates the results of classification accuracy and calculation time. In the studies, the structural dimension of the CNN-SVM model and the D-CNN model are 20 layers and 11

layers, respectively. Training time and classification accuracy of the CNN-SVM in bearing fault diagnosis are 35.82 s and 98.75%, respectively. Fault classification accuracy of the D-CNN model has improved to a certain extent, which decreases the calculation efficiency to a great extent. In detail, fault classification accuracy is equal to 98.83%. However, training time increases to 216.32 s, which is more than six times of the CNN-SVM model. This may be attributed to the fact that the analysis data transformed by wavelet transform into a two-dimensional time-frequency graph undesired increases the train time. Therefore, the dataset processed into one-dimensional data, will improve the accuracy and reduce

the time cost. To sum up, the proposed method has three advantages, first, it can sparsely enhance the data and expand the sample set; second, it has relatively low operation cost, reduces the training time, and can quickly respond to the fault type; third, it has high accuracy under the same condition of predicting the fault type.

In order to represent the accuracy more vividly, a bar chart is used as shown in Figure 15. It can be seen that the proposed GMC-CNN method improves the diagnosis speed and accuracy by taking advantage of the noise reduction ability in the GMC, which retains the convexity of the cost function and represents the sparsity characteristic, as well as the excellent feature extraction in the CNN. In detail, the accuracy of the proposed model during the training process is 99.98%, and the time loss is nearly 5 s in each iteration step. In other words, iteration steps 30 need about 150 s, which greatly improves the solving efficiency. In addition, the effect with CNN, as well as without consideration of the GMC, is also considered. In conclusion, the proposed deep learning method represents a higher accuracy than other models in fault identification and a low time cost.

## 5. Conclusions

In this study, a new sparse enhancement neural network is proposed to diagnose the bearing fault with a high accuracy and efficiency. First, the GMC is used to denoise the sequence signals and sparsely enhance the data derived from random signal slices. Second, the enhanced data are imported into the proposed model for training. Finally, the fault location and severity of rolling bearing are timely identified with a high accuracy. It is revealed that the diagnostic accuracy of the method reaches to 99.98% in the drive end and 99.97% in the fan end, respectively. By comparing with traditional methods (standard CNN model and GMC-2DCNN), the results indicate that the proposed model represents a better ability in fault feature extraction and training time reduction. Meanwhile, it also represents a good ability in identifying the fault with a high precision.

## Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

## Disclosure

This article was previously published as a preprint [32] at [https://studys.ssrn.com/sol3/studys.cfm?abstract\\_id=4092632](https://studys.ssrn.com/sol3/studys.cfm?abstract_id=4092632).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Yong Zhang wrote of the original draft, developed methodology, and visualized the study. Junjie Ye developed

methodology and supervised the study and reviewed and edited the manuscript. Wenhui developed methodology and curated the data. Jinwang Shi visualized the study and curated the data. Wangpeng He conceptualised and supervised the study. Gaigai Cai validated the study and curated the data.

## Acknowledgments

The authors would like to thank the National Natural Science Foundation of China, China (Nos. 52175112, 52075406, and 51805398), the National Natural Science Foundation of Shaanxi Province, China (No. 2018JZ5005), Fundamental Research Funds for the Central Universities (No. JB210421), and the 111 Project, China (No. B14042), for their support and cooperation.

## References

- [1] J. Jiao, M. Zhao, J. Lin, and J. Zhao, "A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes," *Knowledge-Based Systems*, vol. 160, pp. 237–250, 2018.
- [2] S. Haidong, J. Hongkai, L. Xingqiu, and W. Shuaipeng, "Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine," *Knowledge-Based Systems*, vol. 140, pp. 1–14, 2018.
- [3] J. Liu, "A dynamic modelling method of a rotor-roller bearing-housing system with a localized fault including the additional excitation zone," *Journal of Sound and Vibration*, vol. 469, Article ID 115144, 2019.
- [4] X. Wang, Y. Qin, Y. Wang, S. Xiang, and H. Chen, "ReLU-Tanh: an activation function with vanishing gradient resistance for SAE-based DNNs and its application to rotating machinery fault diagnosis," *Neurocomputing*, vol. 363, pp. 88–98, 2019.
- [5] H. Hong and M. Liang, "Fault severity assessment for rolling element bearings using the Lempel–Ziv complexity and continuous wavelet transform," *Journal of Sound and Vibration*, vol. 320, no. 1–2, pp. 452–468, 2009.
- [6] Z. Wang, N. Yang, N. Li, W. Du, and J. Wang, "A new fault diagnosis method based on adaptive spectrum mode extraction," *Structural Health Monitoring*, vol. 20, no. 6, pp. 3354–3370, 2021.
- [7] Z. Li, W. Li, and X. Zhao, "Feature frequency extraction based on singular value decomposition and its application on rotor faults diagnosis," *Journal of Vibration and Control*, vol. 25, no. 6, pp. 1246–1262, 2019.
- [8] G. San Martin, E. López Droguett, V. Meruane, and M. das Chagas Moura, "Deep variational auto-encoders: a promising tool for dimensionality reduction and ball bearing elements fault diagnosis," *Structural Health Monitoring*, vol. 18, no. 4, pp. 1092–1128, 2018.
- [9] H.-b. Yang, J.-a. Zhang, L.-l. Chen, H.-l. Zhang, and S.-l. Liu, "fault diagnosis of reciprocating compressor based on convolutional neural networks with multisource raw vibration signals," *Mathematical Problems in Engineering*, vol. 2019, Article ID 6921975, 7 pages, 2019.
- [10] G. Li, C. Deng, J. Wu, Z. Chen, and X. Xu, "Rolling bearing fault diagnosis based on wavelet packet transform and convolutional neural network," *Applied Sciences*, vol. 10, no. 3, p. 770, 2020.
- [11] D. Gao, Y. Zhu, X. Wang, K. Yan, and J. Hong, "a fault diagnosis method of rolling bearing based on complex morlet

- CWT and CNN,” in *Proceedings of the the 2018 Prognostics and System Health Management Conference (PHM-Chongqing)*, Chongqing, China, October 2018.
- [12] 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.
- [13] S. Guo, T. Yang, W. Gao, and C. Zhang, “A novel fault diagnosis method for rotating machinery based on a convolutional neural network,” *Sensors*, vol. 18, no. 5, p. 1429, 2018.
- [14] G. Xu, M. Liu, Z. Jiang, D. Soffker, and W. Shen, “bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning,” *Sensors*, vol. 19, no. 5, p. 1088, 2019.
- [15] C. Che, H. Wang, X. Ni, and Q. Fu, “Intelligent fault diagnosis method of rolling bearing based on stacked denoising autoencoder and convolutional neural network,” *Industrial Lubrication & Tribology*, vol. 72, no. 7, pp. 947–953, 2020.
- [16] H. Qiao, T. Wang, P. Wang, L. Zhang, and M. Xu, “An adaptive weighted multiscale convolutional neural network for rotating machinery fault diagnosis under variable operating conditions,” *IEEE Access*, vol. 7, Article ID 118954, 2019.
- [17] Q. Jiang, F. Chang, and B. Sheng, “bearing fault classification based on convolutional neural network in noise environment,” *IEEE Access*, vol. 7, Article ID 69795, 2019.
- [18] G. He, K. Ding, and H. Lin, “Fault feature extraction of rolling element bearings using sparse representation,” *Journal of Sound and Vibration*, vol. 366, pp. 514–527, 2016.
- [19] H. Yang, J. Mathew, and L. Ma, “Fault diagnosis of rolling element bearings using basis pursuit,” *Mechanical Systems and Signal Processing*, vol. 19, no. 2, pp. 341–356, 2005.
- [20] I. Selesnick, “Sparse regularization via convex analysis,” *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4481–4494, 2017.
- [21] S. Wang, I. Selesnick, G. Cai, Y. Feng, X. Sui, and X. Chen, “Nonconvex sparse regularization and convex optimization for bearing fault diagnosis,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, pp. 7332–7342, 2018.
- [22] G. Cai, I. W. Selesnick, S. Wang, W. Dai, and Z. Zhu, “Sparsity-enhanced signal decomposition via generalized minimax-concave penalty for gearbox fault diagnosis,” *Journal of Sound and Vibration*, vol. 432, pp. 213–234, 2018.
- [23] G. Cai, S. Wang, X. Chen, J. Ye, and I. W. Selesnick, “Reweighted generalized minimax-concave sparse regularization and application in machinery fault diagnosis,” *ISA Transactions*, vol. 105, pp. 320–334, 2020.
- [24] S. Wang, I. W. Selesnick, G. Cai, B. Ding, and X. Chen, “Synthesis versus analysis priors via generalized minimax-concave penalty for sparsity-assisted machinery fault diagnosis,” *Mechanical Systems and Signal Processing*, vol. 127, pp. 202–233, 2019.
- [25] S. Hao, F.-X. Ge, Y. Li, and J. Jiang, “Multisensor bearing fault diagnosis based on one-dimensional convolutional long short-term memory networks,” *Measurement*, vol. 159, Article ID 107802, 2020.
- [26] Csegroup, “CWRU dataset. Case western Reserve university bearing data center,” 2022, <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>.
- [27] D. Yao, H. Liu, J. Yang, and X. Li, “A lightweight neural network with strong robustness for bearing fault diagnosis,” *Measurement*, vol. 159, Article ID 107756, 2020.
- [28] X. Ding and Q. He, “Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 1926–1935, 2017.
- [29] J. Guo, X. Liu, S. Li, Z. Wang, and W. Caesarendra, “Bearing intelligent fault diagnosis based on wavelet transform and convolutional neural network,” *Shock and Vibration*, vol. 2020, pp. 1–14, Article ID 6380486, 2020.
- [30] L. Yuan, D. Lian, X. Kang, Y. Chen, and K. Zhai, “Rolling bearing fault diagnosis based on convolutional neural network and support vector machine,” *IEEE Access*, vol. 8, Article ID 137395, 2020.
- [31] S. Xie, G. Ren, and J. Zhu, “Application of a new one-dimensional deep convolutional neural network for intelligent fault diagnosis of rolling bearings,” *Science Progress*, vol. 103, no. 3, Article ID 003685042095139, 2020.
- [32] Y. Zhang, J. Ye, W. Yang, J. Shi, W. He, and G. Cai, “A novel sparse enhancement neural network for rolling bearing fault diagnosis,” 2022, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4092632](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4092632).