

Research Article

Intelligent Diagnosis of Rolling Bearing Fault Based on Improved Convolutional Neural Network and LightGBM

Yanwei Xu ^{1,2}, Weiwei Cai,¹ Liuyang Wang,¹ and Tancheng Xie^{1,2}

¹School of Mechatronics Engineering, Henan University of Science and Technology, Luoyang 471003, China

²Intelligent Numerical Control Equipment Engineering Laboratory of Henan Province, Luoyang 471003, China

Correspondence should be addressed to Yanwei Xu; xuyanweiluoyang@163.com

Received 24 July 2021; Accepted 25 November 2021; Published 9 December 2021

Academic Editor: Claudio Sbarufatti

Copyright © 2021 Yanwei Xu 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.

Aiming at the problems of weak generalization ability and long training time in most fault diagnosis models based on deep learning, such as support vector machines and random forest algorithms, one intelligent diagnosis method of rolling bearing fault based on the improved convolution neural network and light gradient boosting machine is proposed. At first, the convolution layer is used to extract the features of the original signal. Second, the generalization ability of the model is improved by replacing the full connection layer with the global average pooling layer. Then, the extracted features are classified by a light gradient boosting machine. Finally, the verification experiment is carried out, and the experimental result shows that the average training and diagnosis time of the model is only 39.73 s and 0.09 s, respectively, and the average classification accuracy of the model is 99.72% and 95.62%, respectively, on the same and variable load test sets, which indicates that the diagnostic efficiency and classification accuracy of the proposed model are better than those of other comparison models.

1. Introduction

Rolling bearing is one of the most critical components widely used in a modern machine, and it is easy to appear cracks, pitting corrosion, and other local damages or defects on the inner and outer ring raceways and rolling elements of the rolling bearings under the harsh working conditions of high temperature, alternating load, and long-time fatigue. As one key component, once the rolling bearing fails, it will affect the safe operation of mechanical equipment, or even damage the equipment and cause casualties. It is of great significance for the safe operation of the mechanical equipment to avoid the occurrence of catastrophic accidents, if we can accurately, timely, and intelligently identify the faults of the rolling bearing and carry out maintenance as soon as possible.

In recent years, various fields made great achievements in the research of algorithms. In order to overcome the slow convergence speed, poor global search ability, and difficult designing rotation angle of quantum-inspired evolutionary algorithm (QEA), Xing et al. [1] proposed an improved quantum-inspired cooperative coevolutionary algorithm,

named MSQCCEA, which is based on combining the strategies of cooperative coevolution, random rotation direction, and Hamming adaptive rotation angle, and the results demonstrate that the proposed MSQCCEA has faster convergence speed and higher convergence accuracy. In order to overcome the low solution efficiency, insufficient diversity in the later search stage, slow convergence speed, and a high search stagnation possibility of differential evolution (DE) algorithm, Deng et al. [2] studied the quantum computing characteristics of quantum evolutionary algorithm (QEA) and, combined with the divide and conquer idea of cooperative evolutionary algorithm (CCEA), proposed an improved differential evolutionary algorithm (HMCQDE), and the results proved that the proposed HMCQDE has higher convergence accuracy and stronger stability and a strong ability to optimize high-dimensional complex functions. Deep learning theory has made great progress in the field of fault diagnosis, and as one of the important models of deep learning theory, continuous neural network (CNN) has shown its own value and great potential in the field of bearing fault diagnosis. For example, Verstraete et al. [3] proposed a deep learning-enabled

featureless methodology to automatically learn the features of the data, and the proposed CNN architecture achieved better results. He [4] presented a deep learning-based approach for bearing fault diagnosis and built an optimized deep learning structure LAMSTAR neural network to diagnose the bearing faults, and the approach shows the accurate classification performance on various bearing faults under different working conditions. Hoang and Kang [5] provided a systematic review of deep learning-based bearing fault diagnosis, introduced the three popular deep learning algorithms for bearing fault diagnosis including autoencoder, restricted Boltzmann machine, and convolutional neural network, and reviewed their applications in the area of bearing fault diagnosis. Zhenghong et al. [6] proposed an adaptive deep transfer learning method for bearing fault diagnosis and verified the method with two kinds of datasets, and the results demonstrate the effectiveness and robustness of the proposed method. Sun et al. [7] proposed a novel intelligent diagnosis method for fault identification of rotating machines, which can not only reduce the amount of measured data that contained all the information of faults but also realize the automatic feature extraction in the transform domain, and the proposed method can reduce the need of human labor and expertise and provide a new strategy to more easily handle the massive data. Ding and He [8] proposed a novel energy-fluctuated multiscale feature mining approach based on wavelet packet energy (WPE) image and deep convolutional network (ConvNet) for spindle bearing fault diagnosis, which is quite suitable for spindle bearing fault diagnosis with multiclass classification regardless of the load fluctuation. Chen et al. [9] proposed a rolling bearing fault diagnosis method based on discrete wavelet transform and the convolution neural network so as to achieve the adaptive feature extraction and intelligent diagnosis of rolling bearing faults, and the experimental results showed that the proposed method has the better generalization ability and robustness. Haidong et al. [10] proposed a novel method for intelligent fault diagnosis of rolling bearing based on deep wavelet autoencoder and extreme learning machine, the method is applied to analyze the experimental bearing vibration signals, and the results showed that the method is superior to the traditional methods and standard deep learning methods. Ding and Jia [11] proposed a one-dimensional multiscale convolutional autoencoder fault diagnosis model of rolling bearings based on the standard convolutional autoencoder, and the test results show that the proposed model has a better recognition effect for rolling bearing fault data. These studies have achieved good diagnosis results. Although the convolutional neural network has achieved good results in the field of fault diagnosis, it cannot well separate the feature extraction and classification functions of the model using the softmax layer to classify the features extracted from the convolution layer, and it may lead to poor classification and generalization ability of the model.

Machine learning algorithms play an important role in the field of fault diagnosis. The single machine learning algorithm, such as support vector machine (SVM) [12, 13] and K-nearest neighbor (KNN) [14], and the ensemble

learning algorithm, such as random forest algorithm [15] and extreme gradient boosting [16, 17], all have made great achievements in the field of mechanical fault diagnosis. However, it is difficult for these classification algorithms to meet the requirements in terms of efficiency and accuracy in big data and high-dimensional environments. Light gradient boosting machine (LightGBM) [18, 19] is a gradient lifting algorithm based on the decision tree, it optimizes the classification accuracy and computational efficiency based on the boosting algorithm, and it is more suitable for classification in a large sample environment, while there will be a lot of unprocessed redundant signals if the original signals are directly input in LightGBM, and it will consume too much memory space in model training and easily cause overfitting of LightGBM classifier.

In order to solve the above problems, in this study, a bearing fault diagnosis model combined with LightGBM algorithm and the improved convolutional neural network that is optimized by replacing the full connection layer to the global average pooling (GAP) layer is proposed (hereinafter referred to as GCNN). The two kinds of data sets under the same load and variable load conditions are constructed. The improvement effect of the global average pooling layer on the model generalization ability and the effectiveness of the proposed model are proved through the comparative analysis with other models.

2. Improved Convolution Neural Network and LightGBM

The convolution neural network is one kind of feed-forward neural network, which adopts unsupervised or semi-supervised learning mode. It contains convolution calculation and deep structure and can classify the input information according to its hierarchical structure [20]. Figure 1 shows the structure of the convolution neural network [21, 22], and it includes the input layer, convolution layer, pooling layer, full connection layer, and output layer. Convolution layer, pooling layer, and full connection layer constitute the hidden layer.

2.1. Convolution Layer and Pooling Layer. Convolution layer is the most basic structure of the convolution neural network, it is the feature extraction layer, its main function is to extract features from the input data, and it uses the local link, weight sharing, and multiple convolution kernels to extract features from data. The most significant features of convolution layer are local sensing and parameter sharing compared with the general deep learning network structure, which can greatly reduce the model parameters and ensure the sparsity of the network. The convolution layer formula is

$$y^{l(i,j)} = \sum_{j'=0}^{m-1} k_i^{l(j')} x^{l(j+j')} m, \quad (1)$$

where $y^{l(i,j)}$ is convoluted output, $k_i^{l(j')}$ is the j' weight value of the convolution kernel i in the l layer, $x^{l(j+j')}$ is the

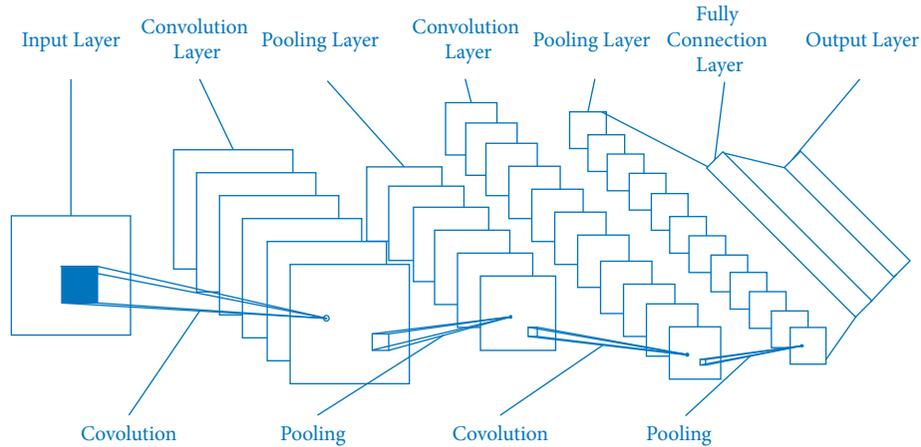


FIGURE 1: The model of the convolution neural network.

convoluted local region j in the l layer, and m is the width of convolution kernel.

The pooling layer is mainly used to select and filter the feature graph extracted from the convolution layer and replace the results of a single point in the feature graph with the statistics of its adjacent regions so as to reduce the number of nodes in the final fully connected layer. It can reduce the overfitting and improve the fault tolerance of the model. The common pooling methods are maximum pooling and average pooling. Compared with the average pooling method, the maximum pooling can select the most significant features in the region; therefore, in this study, the maximum pooling method is selected to select the maximum value in the region as the pooled value of the region. The expression of the maximum pooling method is

$$p^{l(i,j)} = \max\{a^{l(i,j)}\}; (j-1)n+1 \leq t \leq jn, \quad (2)$$

where $p^{l(i,j)}$ is pooled output and n is the width of the pooling area.

The comparison between the full connection layer and the global average pooling layer is shown in Figure 2. It can be seen from Figure 1 that it needs to expand all the features of each feature graph before using the full connection layer, while the global average pooling layer only needs to calculate the average value of each feature graph. It is easy to see from this simple comparison structure diagram that it can greatly reduce the parameter calculation of the classical convolutional neural network using the global average pooling layer to replace the full connection layer.

2.2. Light Gradient Boosting Machine. Gradient boosting decision tree (GBDT) is a long-standing model in machine learning. Its main idea is to use a weak classifier, decision tree, and iterative training to get the optimal model, which has the advantages of good training effect and is not easy to overfit. LightGBM is a framework to realize the GBDT algorithm, it supports efficient parallel training, and it has the advantages of faster training speed, lower memory consumption, better accuracy, support for distribution, and can quickly process massive data. LightGBM uses the negative

gradient of the loss function as the residual approximation of the current decision tree to fit the new decision tree. It uses the histogram algorithm, which takes up less memory and reduces the complexity of data separation. It adopts the leaf-wise strategy with depth restriction, and it will find the leaf with the largest splitting gain, the largest amount of data, from all the current leaves every time, and then splits it. In this way, if the splitting times are the same, the leaf-wise strategy can reduce more errors and get better accuracy. LightGBM can skillfully solve the problem that traditional boosting algorithm is very time-consuming in the large sample environment, and the key of LightGBM is to combine two new methods of gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB). GOSS is a balancing algorithm in reducing the amount of data and ensuring accuracy. GOSS is to reduce the amount of calculation by distinguishing the instances of different gradients, retaining the larger gradient instances, and randomly sampling the smaller gradients, so as to improve the efficiency. EFB is a way to reduce the feature dimension by binding features to improve computing efficiency. Usually, the bundled features are mutually exclusive so that the two features will not lose information.

3. Model of the Improved Convolution Neural Network and LightGBM

3.1. Structure Diagram of the Model. The structure diagram of the improved convolution neural network and LightGBM (GCNN-LightGBM) model is shown in Figure 3, and it is mainly composed of convolution layer, pooling layer, global average pooling layer, and LightGBM classifier. Before the original one-dimensional vibration signal was input into the convolution layer, the random deactivation with a probability of 0.2 was carried out on it, so as to improve the generalization ability of the training model and the stability of fault diagnosis under variable load conditions. There are two convolution layers and two pooling layers. In the first layer, a large convolution kernel is used to obtain more effective information in the low- and medium-frequency bands of the original signal. The feature maps obtained by

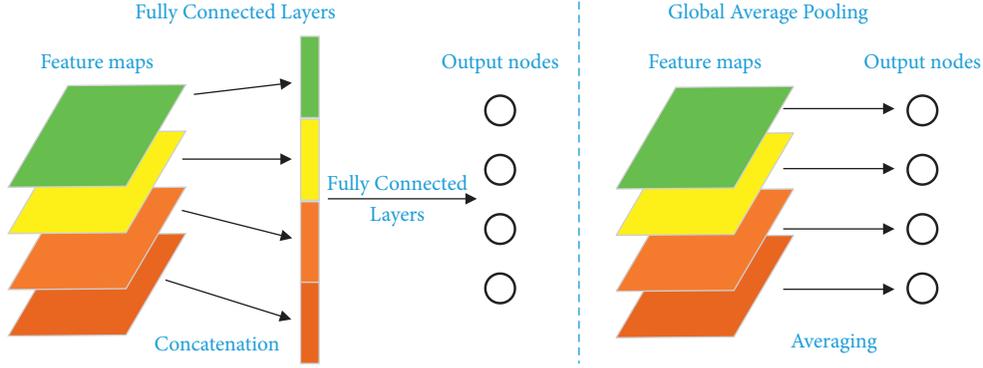


FIGURE 2: Comparison between the full connection layer and global average pooling layer.

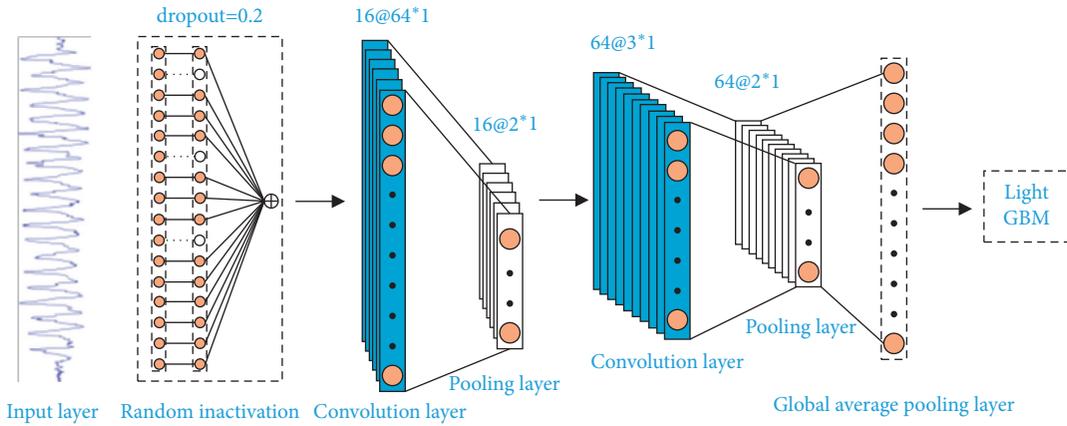


FIGURE 3: Structure diagram of the model of the improved convolution neural network and LightGBM.

two convolution layers and two pooling layers are input into the global average pooling layer, and the secondary feature extraction and data dimension reduction are realized by averaging each feature map. Finally, the extracted low-dimensional features are input into the LightGBM classifier for classification.

3.2. Parameter Setting of the GCNN-LightGBM Model. The GCNN-LightGBM model uses the improved convolution neural network to extract the adaptive features of the bearing vibration signals. It is very important for the feature extraction effect to select the superparameters of the convolution neural network. Therefore, the parameters of the convolution neural network are trained by softmax classifier (shown as Table 1). After the convolution part is trained, LightGBM is used to replace the softmax layer. The selection of LightGBM parameters is processed by the Bayesian parameter adjustment algorithm. The meanings and values of some important parameters are shown in Table 2.

4. Test and Performance Analysis

4.1. Test Data Set. The test data are selected from the Case Western Reserve University Bearing Data Center. The normal bearing data under different load conditions and the fault bearing data of the inner ring, steel ball, and outer ring

TABLE 1: Hyperparameter setting of the improved convolution neural network.

Names of the hyper parameters	Parameter value
Input random deactivation rate	0.2
Loss function	Categorical_crossentropy
Optimizer	Adam
Initial learning rate	0.001
Attenuation coefficient	0.5
Attenuation interval	5
Minimum learning rate	0.00001
Number of batches	32
Iterations	20

TABLE 2: Value of important parameters of LightGBM.

Names of the parameters	Parameter value
n_estimators	500
max_depth	6
num_leaves	34
learning_rate	0.07
bagging_fraction	0.68
bagging_freq	5
feature_fraction	0.6
lambda_l1 , lambda_l2	3.4, 2.1
min_data_in_leaf	33
min_split_gain	0.96
min_sum_hessian_in_leaf	0.003

with the damage diameters of 0.178, 0.356, and 0.534 mm are selected, and the fault bearing is at the drive end. A total of 10 kinds of bearing operation state data are selected. The sampling frequency is set to 12 kHz, and 1,024 data points are collected as a sample each time. The z -score normalization method is used to preprocess the data before feature extraction so as to accelerate the convergence speed of the convolution neural network. The expression of the z -score method is

$$x' = \frac{x - u}{\sigma}, \quad (3)$$

where x is the original sample value, u is the mean value of all sample data, σ is the standard deviation of all samples, and x' is the normalized value.

The selected data are divided into three data sets corresponding to the load of 1 HP, 2 HP, and 3 HP. Each data set contains 10,000 samples, there are 10 kinds of bearing operation state, and each bearing state includes 1,000 samples. About 70% of samples are selected as the training set, 20% samples are selected as the verification set, and 10% samples are randomly selected as the test set. The specific data sets are shown in Table 3.

Generally, the distribution of the data set is different due to the amplitude, fluctuation period, and phase inconsistency of the vibration signals under different working conditions. Therefore, it needs the classifier designed has strong generalization ability and robustness. However, it is not realistic to collect and mark enough training samples to make the classifier robust to all the working conditions. In this study, one method using the single load to train the fault diagnosis model and using the test set of the other loads to carry out the fault diagnosis is adopted [23]. For example, it requires the model trained under 1 HP load to not only have high classification accuracy in the 1 HP test set but also in the 2 HP or 3 HP test set, and the variable load adaptive data set constructed to achieve this goal is shown in Table 4.

4.2. Model Validation. In the experiment, the GCNN-LightGBM model uses the deep learning framework Keras in Python language, the classification module directly calls the LightGBM software package, and the established network is used to train and test using different data sets. Because the initialization of input data and neural network weights is random, the average value was calculated after each data set was trained 10 times so as to ensure the reliability of the test results.

In order to verify that the improved convolutional neural network has stronger generalization ability, the contrast model of the classical convolutional neural network and LightGBM is constructed, and the network structure and training parameters of the contrast model are consistent with the GCNN-LightGBM model except for the full connection layer. At the same time, in order to verify that LightGBM has a stronger classification ability than the softmax layer, one contrast model of the improved convolutional neural network and softmax also is constructed, and the feature extraction part of the contrast model is

TABLE 3: Bearing test data set.

Bearing status	Fault diameter (mm)	Number of samples
Normal	—	700/200/100
	0.178	700/200/100
	0.356	700/200/100
Ball failure	0.356	700/200/100
	0.534	700/200/100
	0.178	700/200/100
Inner ring fault	0.356	700/200/100
	0.534	700/200/100
	0.178	700/200/100
Outer ring fault	0.356	700/200/100
	0.534	700/200/100
	0.178	700/200/100

TABLE 4: Variable load adaptive data set.

Training data set	Training sample	Target test set	Test sample
1 HP	7000	2 HP, 3HP	1000, 1000
2 HP	7000	1 HP, 3HP	1000, 1000
3 HP	7000	1 HP, 2HP	1000, 1000

consistent with the model in this study except the softmax classifier.

The recognition accuracy of each model under different load conditions is shown in Figures 4 and 5.

It can be seen from Figures 4 and 5 that the average classification accuracy of the GCNN-LightGBM model is slightly higher than that of the CNN-LightGBM model under the same load condition, while the average classification accuracy of the GCNN-LightGBM model is 2.72% higher than that of the CNN-LightGBM model under variable load condition, and it is verified that the improved convolutional neural network has a better anti-overfitting effect and can improve the generalization ability of the model. The average classification accuracy of the GCNN-LightGBM model is 1.05% and 0.77% higher than that of the GCNN-softmax model, respectively, under the same load and variable load conditions, and it indicates that LightGBM has a stronger classification ability than softmax. LightGBM classifier can achieve good classification results under the same load condition, but the average classification accuracy is less than 68% under the variable load condition, which indicates that it is easy to overfit when LightGBM is directly used to train the original data although it is a very powerful classifier, and it is necessary to extract features from the original data. The classification accuracy between adjacent conditions is high under variable load conditions, which indirectly reflects that the distribution difference of adjacent load data sets is small but the distribution difference of nonadjacent load data sets is large.

4.3. Contrast Test

4.3.1. Accuracy Rate of Fault Diagnosis. Since the classification accuracy of the GCNN-LightGBM model is close to 100% under the same load test set, several deep learning models [24, 25] that have achieved good classification results under the same load conditions are selected to carry out the

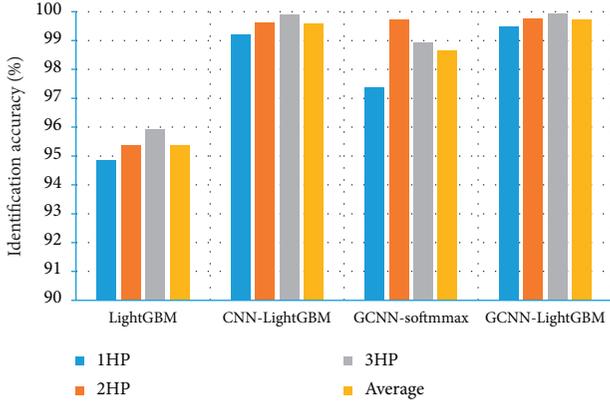


FIGURE 4: The classification accuracy of each model under the same load condition.

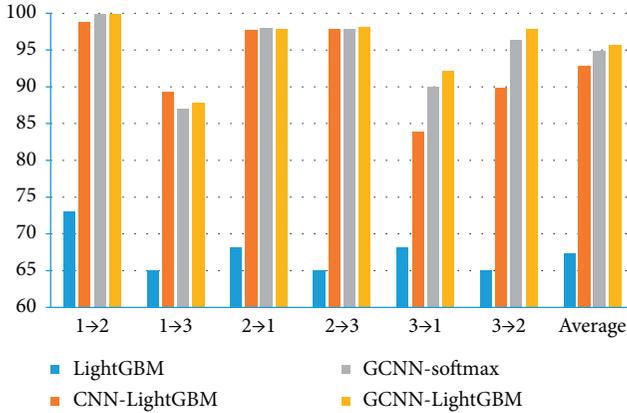


FIGURE 5: The classification accuracy of each model under different load conditions.

comparative analysis under the variable load test sets, so as to highlight the advantages of the generalization ability and load migration ability of GCNN-LightGBM model. The classification accuracy of different deep learning models under the variable load adaptive data sets is shown in Figure 6.

It can be seen from Figure 6 that the models of CNN-LSTM and WDCNN have strong adaptability when they are trained under the load conditions of 1 HP and 2 HP, and the classification accuracy can reach more than 90% in other variable load test sets. But the load migration ability of the models is not strong when they are trained under the load conditions of 3 HP, and the classification accuracy can only reach about 80% under the load conditions of 1 HP and 2 HP. The classification accuracy of the CNN-SVM model trained under the load condition of 3 HP in other test sets is close to 100%, but the classification accuracy of the CNN-SVM model trained under the load condition of 1 HP in other test sets is even less than 80%. It shows that the overall robustness and load migration ability of the three comparison models are not very strong, although they can achieve good classification results on a variable load test set.

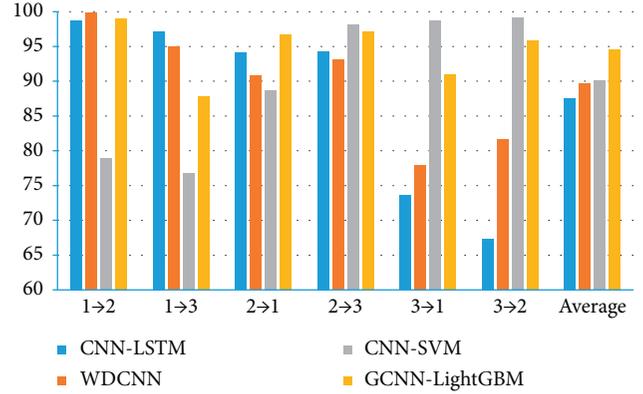


FIGURE 6: The classification accuracy of each model under the variable load condition.

The lowest classification accuracy of the GCNN-LightGBM model is about 87.89% under the variable load conditions, and it is increased by 20.57%, 9.93%, and 11.10%, respectively, compared with the worst cases of the models of CNN-LSTM, WDCNN, and CNN-SVM. The average classification accuracy of the GCNN-LightGBM model is 94.64%, which is significantly higher than that of other models. It can be seen that the GCNN-LightGBM model has the better overall classification effect under the variable load conditions and also has better generalization ability and load migration ability.

4.3.2. Efficiency of Fault Diagnosis. In order to further highlight the advantages of the GCNN-LightGBM model on the efficiency of rolling bearing fault diagnosis, the training time and diagnosis time of each model, the amount of training parameters of the deep learning model, and the number of required layers of training parameters (excluding pooling layer) are recorded in the process of comparative test, and these parameter values are shown as Table 5. Among them, the average duration is the average value of each model trained for 10 times under different load conditions, and the average value of different loads is calculated again.

It can be seen from Table 5 that the required training parameters and network layers of the GCNN-LightGBM model are the least, especially the amount of the training parameters are several orders of magnitude different from the other three networks, and the required average training and diagnosis time is the smallest among the four models. The training time is shortened by 90.81%, and the fault diagnosis efficiency is improved by 10.44 times compared with the CNN-LSTM model.

These data show that the less the parameters and layers of the model need to be trained, the shorter the time of model training and fault diagnosis. The advantage of the short fault diagnosis time of the GCNN-LightGBM model will be more obvious when there are hundreds or even millions of samples need to be trained or diagnosed, and it can save a lot of time cost.

TABLE 5: Analysis table of training time and diagnosis time of each model.

Models	Training parameters		Average duration (s)	
	Number of layers	Number	Training	Diagnosis
CNN-LSTM	8	507333	432.14	1.03
WDCNN	7	20755	107.26	0.12
CNN-SVM	4	63518	65.81	0.15
GCNN-LightGBM	2	2413	39.73	0.09

5. Conclusions

In order to solve the problem that weak generalization ability and long training time in most fault diagnosis models based on deep learning, one intelligent diagnosis method of rolling bearing fault based on the improved convolution neural network and light gradient boosting machine is proposed.

- (1) First, the random deactivation with a probability of 0.2 was carried out on the original one-dimensional vibration signal, so as to improve the generalization ability of the training model and the stability of fault diagnosis under variable load conditions. Second, the signal was input into the GCNN. In the first layer, a large convolution kernel is used to obtain more effective information in the low- and medium-frequency bands of the original signal. The feature maps obtained by two-layer convolution pooling operation are input into the global average pooling layer, and the secondary feature extraction and data dimension reduction are realized by averaging each feature map. Finally, the extracted low-dimensional features are input into the LightGBM classifier for classification.
- (2) The results show that (1) the average classification accuracy of the GCNN-LightGBM is 99.72% for the same load test set and 95.62% for the variable load test set; (2) the GCNN-LightGBM model has the higher average classification accuracy on the variable load test set compared with the models of CNN-LSTM, WDCNN, and CNN-SVM, and it has the stronger generalization ability and load migration ability; (3) the GCNN-LightGBM model only needs two training layers, and the amount of parameter calculation is less than 3,000, the training and fault diagnosis durations are 39.73 s and 0.09 s, respectively, and these data are far lower than other comparison models, which shows that the GCNN-Light GBM model has the advantages of simple structure, less parameter calculation, and high efficiency of training and fault diagnosis.
- (3) In this study, the generalization ability of the model is improved according to the change of load. In the future, the robustness of the model will be further improved by adding noise interference to the samples.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful for the financial support provided by the National Natural Science Foundation of China under grant no. 51805151 and the Key Scientific Research Project of the University of Henan Province of China under grant no. 21B460004.

References

- [1] C. Xing, Z. Huimin, S. Shifan et al., "An improved quantum-inspired cooperative co-evolution algorithm with multi-strategy and its application," *Expert Systems with Applications*, vol. 171, no. 1, Article ID 114629, 2021.
- [2] W. Deng, S. Shang, C. Xing et al., "Quantum differential evolution with cooperative coevolution framework and hybrid mutation strategy for large scale optimization," *Knowledge-Based Systems*, vol. 224, Article ID 107080, 2021.
- [3] D. Verstraete, A. Ferrada, E. Lopez 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, 2017.
- [4] M. He and D. He, "Deep learning based approach for bearing fault diagnosis," *IEEE Transactions on Industry Applications*, vol. 53, no. 3, pp. 3057–3065, 2017.
- [5] D. T. Hoang and H. J. Kang, "A survey on Deep Learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, 2019.
- [6] W. Zhenghong, J. Hongkai, Z. Ke, and X. Li, "An adaptive deep transfer learning method for bearing fault diagnosis," *Measurement*, vol. 151, Article ID 107227, 2020.
- [7] J. Sun, C. Yan, and J. Wen, "Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 185–195, 2018.
- [8] 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.
- [9] R. X. Chen, X. Huang, L. X. Yang, and B. P. Tang, "Rolling bearing fault identification based on convolution neural network and discrete wavelet transform," *Journal of Vibration Engineering*, vol. 31, no. 5, pp. 883–891, 2018, in Chinese.
- [10] 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.
- [11] Y. Ding and M. Jia, "A multi-scale convolutional auto-encoder and its application in fault diagnosis of rolling

- bearings,” *Journal of Southeast University*, vol. 35, no. 4, pp. 417–423, 2019.
- [12] Z. Wang, L. Yao, and Y. Cai, “Yongwu. Rolling bearing fault diagnosis using generalized refined composite multiscale sample entropy and optimized support vector machine,” *Measurement*, vol. 156, Article ID 107574, 2020.
- [13] D. Goyal, A. Choudhary, B. S. Pabla, and S. S. Dhimi, “Support vector machines based non-contact fault diagnosis system for bearings,” *Journal of Intelligent Manufacturing*, vol. 31, no. 5, pp. 1275–1289, 2020.
- [14] S. Dong, T. Luo, L. Zhong, L. Chen, and X. Xu, “Fault diagnosis of bearing based on the kernel principal component analysis and optimized k-nearest neighbour model,” *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 36, no. 4, pp. 354–365, 2017.
- [15] M. Cerrada, G. Zurita, D. Cabrera, R. V. Sánchez, M. Artés, and C. Li, “Fault diagnosis in spur gears based on genetic algorithm and random forest,” *Mechanical Systems and Signal Processing*, vol. 70–71, pp. 87–103, 2016.
- [16] Z. Wu, X. Wang, and B. Jiang, “fault diagnosis for wind turbines based on ReliefF and eXtreme gradient boosting,” *Applied Sciences*, vol. 10, no. 9, p. 3258, 2020.
- [17] J. Tao, C. Qin, W. Li, and C. Liu, “intelligent fault diagnosis of diesel engines via extreme gradient boosting and high-accuracy time-frequency information of vibration signals,” *Sensors*, vol. 19, no. 15, p. 3280, 2019.
- [18] A. A. Taha and S. J. Malebary, “An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine,” *IEEE Access*, vol. 825579 pages, 2020.
- [19] P. j. Chun, S. Izumi, and T. Yamane, “Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 1, pp. 61–72, 2020.
- [20] C. Zhang, X. Yao, J. Zhang, and E. Liu, “Tool wear monitoring based on deep learning,” *Computer Integrated Manufacturing Systems*, vol. 23, no. 10, pp. 2146–2155, 2017, (in Chinese).
- [21] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, “Deep learning and its applications to machine health monitoring,” *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019.
- [22] S. Khan and T. Yairi, “A review on the application of deep learning in system health management,” *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.
- [23] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, “A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals,” *Sensors*, vol. 17, no. 2, p. 425, 2017.
- [24] L. Zhang, W. Xu, and L. Jing, “Study on fault diagnosis methods of rolling bearing based on deep convolutional neural network and support vector machine,” *Modular Machine Tool & Automatic Manufacturing Technique*, vol. 2019, no. 7, pp. 69–71, 2019, in Chinese.
- [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.