Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 14, No. 3, 2025

Bearing Fault Diagnosis Method Based on SE-CNN

Taohuang Liu, Zhihua Hu*, Yanyang Zhang, Xiang Liu, Sunjie Liu, Weijian Shao

School of Intelligent Manufacturing and Control Engineering, Shanghai Polytechnic University, Shanghai, 201209, China
* Corresponding author: Zhihua Hu (Email: zhhu@sspu.edu.cn)

Abstract: As a key component in rotating machinery, the health condition of rolling bearings directly affects the operational efficiency and safety of the equipment. Traditional bearing fault diagnosis methods rely on signal processing techniques and empirical feature extraction, but in practical applications, with the change of working conditions, the signal features often appear to be mixed, which brings a greater challenge to the diagnosis. In this paper, a bearing fault diagnosis method based on Squeeze-and-Excitation Convolutional Neural Network is proposed. The method enhances the network's focus on key fault features by introducing the SE module to adaptively adjust the weights of each feature channel in the convolutional network, which optimizes the feature extraction ability of the model in complex vibration signals. The experimental results show that the SE-CNN performs superiorly in terms of precision rate, recall rate and F1-score, especially in the case of category imbalance, the SE-CNN has better robustness.

Keywords: Rolling Bearing, Fault Diagnosis, SE-CNN.

1. Introduction

In the process of industrial equipment operation, rolling bearings as the core components of rotating machinery, its health status directly affects the operational efficiency and safety of the equipment. Due to long-term exposure to complex mechanical loads, vibration and environmental factors, bearings are prone to fatigue damage, wear and even failure. Therefore, the accurate identification of bearing health condition, which provides the basis for maintenance decision-making, has become an important task to ensure the smooth operation of equipment.

Traditional intelligent diagnosis methods for bearing faults generally require the help of fast Fourier transform, wavelet transform and empirical modal decomposition[1-3],The acquired signals are processed, the features are artificially designed and extracted by combining expert experience, and finally the fault identification is realized by a back-end classifier, such as artificial neural network[4], SVM[5], and RF[6], etc. However, such methods have the defects of relying on expert experience, cumbersome procedure, and poor nonlinear fitting ability. However, such methods have the defects of relying on experts' experience, cumbersome procedures and poor nonlinear fitting ability of shallow classifiers[7-8].

The deep neural networks that have emerged in recent years have a strong nonlinear ability to automatically learn deep feature representations from data, making up for the shortcomings of traditional methods of intelligent diagnosis of bearing faults and bringing new solution ideas.Janssens et al[9] introduced convolutional neural networks into the field of fault diagnosis, where the raw signals are processed by discrete Fourier transform and then used as inputs to a CNN model for fault identification.Shebo et al[10] proposed a fault diagnosis method based on Deep Convolutional Variational Self-Coding Network(DCVAEN), which directly utilizes the spectral information of vibration signals as training data for fault diagnosis.

However, the theory has achieved great success, but in practical applications, bearings in the load, speed and other conditions of the working conditions change conditions, different changes in working conditions will cause non-linear changes in vibration signals, resulting in different fault location, different severity of the vibration signals appeared in the overlap, which in turn increases the difficulty of the bearing fault diagnosis.

To address the above problems, this paper proposes a bearing fault diagnosis method based on Squeeze-and-Excitation Convolutional Neural Network.The SE module enhances the expression of key features by adaptively adjusting the feature weights between channels, and the CNN targetedly adjusts the feature channel weights of the convolutional layer to increase the weights of special channels that contain fault information and cut down the useless feature channels.

2. Theoretical Foundations

2.1. Squeeze-And-Excitation Networks

SE is a lightweight attentional mechanism focusing on channel information, which is added to existing neural networks to improve the model performance with only a small increase in computational effort [11-12].SE networks are able to recalibrate the channel response to features through the interdependencies between channels, enhancing the network's representational capabilities.

The SE network consists of three parts, the Squeeze, the Excitation and the Scale and Multiply feature fusion operation. Where the global average pooling of Squeeze is a channel statistic generated by equation (1) that encodes the spatial features on a channel into a global feature that preserves the overall semantic information.

$$z_c = F_{\mathrm{sq}}(u_c) = \frac{1}{H \times \mathrm{W}} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j) \tag{1}$$

Where $u_c(i,j)$ is the data at position (i,j) in the cth channel， z_c is the response data of the compressed part in the cth channel; H, W are the length and width of the data, respectively.

The goal of The Excitation is to fully capture the dependencies on channel dimensions and parameterize the attention mechanism through two fully connected layers. By connecting the two fully-connected layers in series, not only is it possible to achieve a nonlinear mapping that enhances or suppresses the feature representation of a particular channel, but it is also possible to optimize the parameter scales of the model and improve the ability to fit inter-channel correlations. The realization of this excitation operation can be expressed by equation (2).

$$\begin{aligned} s &= F_{\text{ex}}(z, W) \\ s &= \sigma(g(z, W)) \\ s &= \sigma(W_2 \delta(W_1 z)) \end{aligned} \tag{2}$$

Among them. $W_1 \in R^{\frac{C}{r} \times C}, W_2 \in R^{C \times \frac{C}{r}}$, δ is the ReLU activation function for the FC layer, σ is the Sigmoid activation function for the FC layer, W_1, W_2 are FC layer weights.

The Scale realizes the adaptive adjustment of channel features by multiplying each channel weight obtained from the Excitation computation, point by point, with each element within the corresponding channel. After this process, the feature mapping U is reweighted based on the channel weights to generate the final output of the SE module[13].

2.2. Convolutional Neural Networks

Convolutional Neural Networks is a deep learning model designed to deal with the local correlation of data, compared with the traditional Fully Connected Neural Network (CNN), CNN effectively reduces the model complexity and improves the feature extraction capability through mechanisms such as local sensory wildness, weight sharing and pooling[14-16].

The structure mainly consists of Convolutional Layer, Activation Function, Pooling Layer, Fully Connected Layer and Output Layer, where the convolutional computation is represented by Equation (3), and the maximum pooling to extract the key features is computed by Equation (4).

$$Y_{i,j}^{k} = \sum_{m}\sum_{n} X_{i+m,j+n} \cdot W_{m,n}^{k} + b^{k} \tag{3}$$

$$Y_{i,j} = \max(X_{i+m,j+n}), \quad m, n \in P \tag{4}$$

Where X is the input feature, W is the convolution kernel weight, b is the bias term, Y is the convolution output, and P is the pooling layer window size.

2.3. Bearing Fault Diagnosis Method Based on SE-CNN

In bearing fault diagnosis, vibration signals usually have complex spatio-temporal characteristics, and the energy distribution in different frequency bands is closely related to the fault type. However, the traditional convolutional neural network often assigns the same weight to the features of all channels when processing vibration signals, failing to fully consider the importance of the information of different channels, which may lead to the weakening or loss of some of the key fault features. The convolutional neural network combined with the channel attention network shown in Figure 1 can effectively enhance the feature extraction ability of the model and improve the accuracy of bearing fault classification[17].

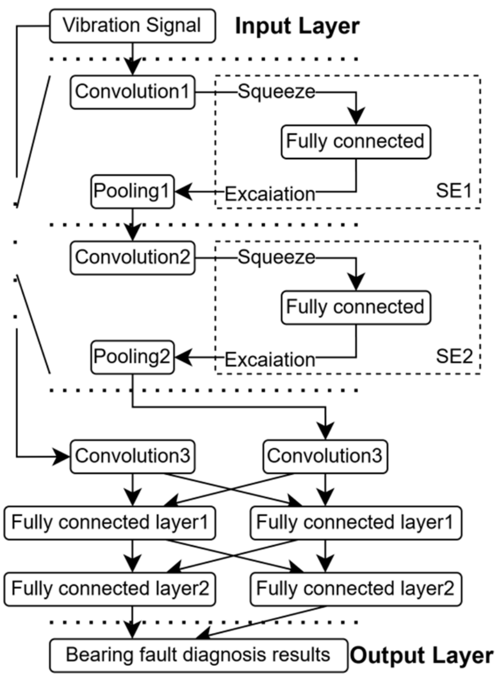

Figure 1. SE-CNN Structure

The whole algorithm model is architected and trained on python3.8.3 platform based on TensorFlow deep learning framework, and the training parameters of SE-CNN bearing fault diagnosis method are shown in Table 1.

Table 1. SE-CNN Training Parameter

Learning Rate	Epochs	Size	SE-FC1	CNN Activation Function
0.001	**50**	**512**	**1/16**	**ReLU**

3. Experimental Examples

3.1. Experimental Data

The experimental data are provided by the Bearing Experiment Center of Case Western Reserve University (CWRU), and the main components of the test bench include traction motor, torque sensor and dynamometer, and the motor shaft support bearing model is 6205-2RS JEM SKF.

The experimental data in this paper are collected from the motor drive end bearings in different health states, the sampling frequency are 12KHz, according to the different conditions of the construction of the data set, the data set training samples and test samples in the ratio of 0.8:0.2, and the use of the data amplification technology, the length of each sample is 2048, and normalized, as shown in Table 2.

Table 2. Specific Components Of The Data Set

Fault Location	Pitting Diameter(mil)	Category Tag	Number Of Training/Testing Samples
0.001	**50**	**512**	**1/16**
Normal		**Normal**	**400/100**
Inner Race	**7**	**IR_07**	**400/100**
	14	**IR_14**	**400/100**
	21	**IR_21**	**400/100**
Outer Race	**7**	**OR_07**	**400/100**
	14	**OR_14**	**400/100**
	21	**OR_21**	**400/100**
Ball	**7**	**Ball_07**	**400/100**
	14	**Ball_14**	**400/100**
	21	**Ball_21**	**400/100**

3.2. Evaluation Indicator

In order to comprehensively evaluate the performance of the improved SE-CNN model in the bearing fault diagnosis task, Macro-F1 , Recall, Accuracy and Precision are selected as the main evaluation indexes in this paper. These indexes can comprehensively reflect the classification effect of the model in different fault categories and ensure its generalization ability and stability.

1.Accuracy:Indicates the correct rate of overall classification of the model, i.e. the proportion of correctly classified samples to all samples.

2.Precision:Indicates the degree of accuracy of the model's prediction of a category, reflecting how many of the samples predicted to be in that category actually belong to that category.

3.Recall:Measures the model's ability to recognize a category.

4.Macro-F1:Calculates the arithmetic mean of F1-score for all categories, which is used to evaluate the classification effect of different categories in a balanced way.

With these indicators, this paper is able to comprehensively evaluate the performance of the SE-CNN model in the bearing fault diagnosis task and verify its improvement.

3.3. Experimental Results and Analysis

In this experiment, a fault diagnosis task was performed on the CWRU bearing dataset based on the SE-CNN model. The classification effect of the model is further analyzed by combining the Loss value iteration curve and confusion matrix during the model training process.

The Loss iteration curve is shown in Fig. 2, where the loss value gradually decreases as training proceeds, indicating that the model is continuously learning and optimizing its weights. The smooth decrease in the training process reflects the good convergence of the model and indicates that the training data is effectively fitted.

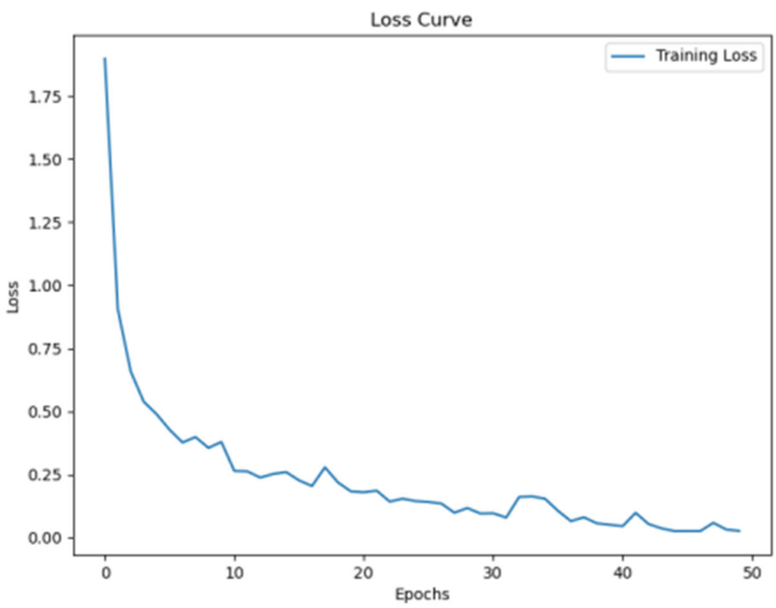

Figure 2. Loss Curve

Figure 3 shows the confusion matrix of the model on the test set. As can be seen from the confusion matrix, the vast majority of the test samples are correctly categorized and have large values on the diagonal line, indicating that the model makes accurate predictions on most of the categories. However, the model is prone to confusion in the failure type of 17 mil for the outer ring and 7 mil for the rolling body.

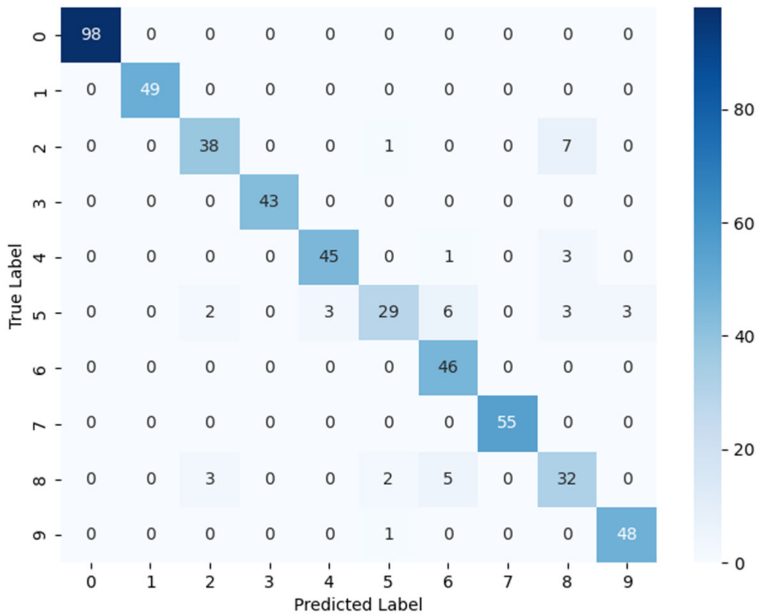

Figure 3. Confusion Matrix

In order to verify the effectiveness of SE-CNN model in this experimental task, network models with the same depth and different structures are built, which are ResNet model, LSTM model and original CNN model, and the comparison results are shown in Table 3.

Table 3. Comparison Of Training Results of Each Model On The Dataset

Average	precision	recall	f1-score
ResNet	**0.8773**	**0.8658**	**0.8634**
LSTM	**0.8926**	**0.8891**	**0.8825**
CNN	**0.9039**	**0.8948**	**0.8961**
SE-CNN	**0.9263**	**0.9235**	**0.9219**

It can be learned that ResNet performs relatively weakly in the bearing fault diagnosis task, its precision and recall are low, and its F1-score is also at a low level.LSTM performs well in capturing the long term dependencies of the time series data, and thus can effectively identify the fault features in the bearing vibration signals.The performance of LSTM is better compared to that of ResNet in this experiment, but it is still not as good as that of the CNN model and SE-CNN model. And although the original CNN model has good classification performance in this experimental task, it is still slightly inferior to the SE-CNN model, which enhances the selective expression of features by introducing the SE module, effectively reduces the rate of misclassification and omission, and shows better robustness, especially in the case of category imbalance.

4. Conclusion

In this paper, a bearing fault diagnosis method based on SE-CNN is proposed, which enhances the feature extraction capability of convolutional neural network in bearing vibration signal analysis by introducing the SE module to adaptively adjust the weights of channel features. Experimental results show that SE-CNN has higher precision, recall and F1-score in the bearing fault diagnosis task compared with traditional ResNet, LSTM and CNN models, and exhibits better robustness especially in the case of category imbalance.

Compared with the other models, ResNet performs less well in fault diagnosis, and LSTM captures timing information better but still not as good as CNN and SE-CNN.The primitive CNN has good classification performance but there is still room for improvement in feature selectivity.SE-CNN, by introducing the channel attention mechanism, significantly improves the ability of capturing key signal features and optimizes the diagnosis accuracy.

References

[1] ZHANG X L. Discussion on the acquisition and processing of motor bearing vibration signals[J]. Explosion-proof Electric Machines, 2024, 59(04): 87-90.

[2] ZHAO L, MENG Y, JIANG Z L, et al. Bearing fault diagnosis based on wavelet transform and attention mechanism[J/OL]. Vibration, Test & Diagnosis, 1-8 [2025-03-19]. http://kns.cnki.net/kcms/detail/32.1361.V.20241113.1120.002.html.

[3] PAN T H, CAI J H, GAO Y, et al. Fault feature extraction of hydropower unit vibration signals based on improved empirical mode decomposition[J/OL]. Hydropower Energy Science, 2025, (03): 182-185+162 [2025-03-19]. https://doi.org/10.20040/j.cnki.1000-7709.2025.20240783.

[4] KANKAR P K,SHARMA S C,HARSHA S P.Rolling Element Bearing Fault Diagnosis Using Wavelet Transform[J].Neurocomputing,2011(74):1638-1645.

[5] LIU C, WU X, LIU T, et al. Rolling bearing fault diagnosis based on approximate isometric projection and support vector machine[J]. Journal of Vibration and Shock, 2018, 37(5): 234-239.

[6] WEN J T, ZHOU X N. Rotating machinery fault diagnosis based on fuzzy granulation unsupervised learning combined with random forest fusion[J]. Mechanical Science and Technology, 2018, 37(11): 1722-1730.

[7] ZHAO R,YAN R Q,CHEN Z H,et al.Deep Learning and Its Applications to Machine Health Monitoring[J].Mechanical Systems and Signal Processing,2019,115:213-237.

[8] SAMIR K,TAKEHISA Y.A Review on The Application of Deep Learning in System Health Management[J].Mechanical Systems and Signal Processing,2018,107(1):241-26.

[9] JANSSENS O,SLAVKOVIKJ V,VAERVISCH B,et al.Convolutional Neural Network Based Fault Detection for Rotating Machinery [J]. Journal of Sound and Vibration, 2016, 377:331-345.

[10] SHE B, TIAN F Q, LIANG W G. Fault diagnosis method based on deep convolutional variational autoencoder network [J]. Journal of Instrumentation, 2018, 39(10): 27-35.

[11] MA Z. Remaining useful life prediction of rolling bearings based on EEMD and MCNN-SE-GRU [D]. Xinjiang University, 2021. DOI:10.27429/d.cnki.gxjdu.2021.000597.

[12] ZHAO G W, ZENG J. Rolling bearing fault diagnosis method based on GAF and SE-ResNeXt [J]. Mechanical Design and Manufacturing Engineering, 2025, 54(02): 63-67.

[13] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks [J]. IEEEXplore, 2017: 7132-7141.

[14] ZHOU F Y, JIN L P, DONG J. A review of convolutional neural networks [J]. Journal of Computer Research and Development, 2017(6): 1229-1251.

[15] ZHANG D, KONG K, ZHU C, et al. Bearing remaining useful life prediction based on CNN-Transformer encoder-BiLSTM model [J/OL]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 1-16 [2025-03-19]. https://doi.org/10.13245/j.hust.24091.

[16] CHEN Y, QU J L, WANG X F, et al. CNN-LSTM based rolling bearing fault diagnosis method with enhanced temporal memory [J]. Noise and Vibration Control, 2025, 45(01): 105-111.

[17] SUN Y, ZHANG L, ZHOU K. Research on rolling bearing fault diagnosis method based on genetic algorithm optimized convolutional neural network [J]. Manufacturing Automation, 2025, 47(01): 89-95.