*Advances in Artificial Intelligence, Big Data and Algorithms*
*G. Grigoras and P. Lorenz (Eds.)*

*© 2023 The Authors.*
*This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).*
*doi:10.3233/FAIA230855*

# Multi-Scale Segmentation Attention-Based Residual Network for Fault Diagnosis of Rolling Bearings

Duo CHEN[a], Yuehua LAI[b,c,1], Changjun YANG[d], Chunhua HUANG[e], Wei LI[d], Hong GAO[e], Lei NIU[b]

[a] *China Energy Group Ningxia Coal Industry Co., Ltd., Yinchuan 750000, China*
[b] *Beijing Tianma Intelligent Control Technology Co., Ltd., Beijing 101399, China*
[c] *China Coal Research Institute, Beijing 100013, China*
[d] *Hongliu Coal Mine, China Energy Group Ningxia Coal Industry Co., Ltd., Yinchuan 750408, China*
[e] *China Coal Technology and Engineering Group, Chongqing Research Institute, Chongqing 400039, China*

**Abstract.** Rolling bearings are widely used in coal mine equipment, real-time monitoring of bearing status and intelligent diagnosis of bearing fault is very important to the safe and efficient mining of coal mine. In order to solve the problem of insufficient representation ability about fault diagnosis model, a multi-scale segmentation attention-based residual network is proposed for the fault diagnosis of rolling bearings, which can fully and accurately extract vibration signal features to realize intelligent diagnosis. For time-frequency images of vibration signals, residual networks was used for feature extraction. Furthermore, the pyramid split attention mechanism was combined to optimize the feature selection to construct the intelligent diagnosis model for bearing. The proposed method has been applied to the task of fault diagnosis on SKF6205 bearing, and the experimental results show its superior diagnostic performance.

**Keywords.** Rolling bearing, time-frequency images, residual network, pyramid split attention

## 1. Introduction

Rolling bearing is a kind of precision rotating parts, widely used in various mechanical equipment. In order to ensure the safety and efficient mining of coal mines, it is of great significance to carry out health monitoring of rolling bearings and timely intelligent alarm of early failure [1, 2]. Deep learning model can automatically extract signal features and conduct modeling and analysis on complex problems, which has been widely concerned and studied. Importantly, fault diagnosis models built based on deep learning have also achieved important success and wide application [3-5].

Based on convolutional neural network (CNN) and multilayer perceptron (MLP), Sinitsin *et al.* proposed a hybrid fault diagnosis model for rolling bearings by combining

[1] Corresponding Author: Yuehua LAI, Beijing Tianma Intelligent Control Technology Co., Ltd.; China Coal Research Institute; E-mail: laiyh@tdmarco.com

mixed inputs [6]. Based on soft threshold and global context, Lyu *et al.* extracted signal features from vibration signals, and then proposed a fault diagnosis method by combining the residual building unit [7]. Ye used an adaptive signal processing method to process vibration signals of bearing and built the diagnosis model for rolling bearings based on an improved one-dimensional CNN to automatically identify different faults of rolling bearings [8]. Li *et al.* extracted the discriminant features from the original vibration signals of bearing from multiple dimensions such as sequence, channel and space, and built a fault diagnosis model for bearing with different operating conditions based on transfer learning [9]. However, the existing fault diagnosis models do not deeply and effectively integrate many model features. On the one hand, they do not integrate features from multiple dimensions, on the other hand, they do not optimize the selection of features, and then extract sensitive and effective features to form a feature combination with strong characterization ability. Therefore, the multi-scale segmentation attention-based residual network, in which residual networks was used for feature extraction and pyramid split attention (PSA) mechanism was combined to optimize the feature set, is proposed to improve the generalization ability of the intelligent fault diagnosis model for rolling bearings.

## 2. Structure of Residual Block

Residual network is one of the widely used CNN network architectures. It introduces residual learning in the process of learning potential mapping to mitigate gradient disappearance and degradation caused by the increase of network layers through identity shortcut connection. The residual block, which is shown in Figure 1, is the core basicblock of residual network.

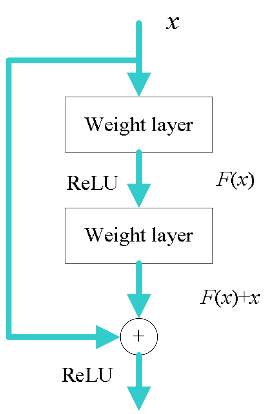


**Figure 1.** Basic structure of residual block.

Suppose the input of model is $x$, and the feature learned by neural network is assumed to be $H(x)$, the residual $F(x)$ refers to the $H(x)$-$x$ part. When the residuals are $F(x) = 0$, the residual block is equivalent to implementing the identity mapping. In this way, gradient disappearance and degradation caused by the increase of network layers can be greatly alleviated [10].

## 3. The Proposed Method

In order to fully and accurately extract vibration signal features, so as to solve the problem of insufficient representation ability of fault diagnosis model, this paper proposed a multi-scale segmentation attention-based residual network for fault diagnosis of rolling bearings. It mainly includes time-frequency image conversion of vibration data, data preprocessing, establishment of fault diagnosis model and model training and application.

### *3.1. Time-Frequency Image Conversion of Vibration Data*

The classical time-frequency analysis method in the field of signal processing, which named continuous wavelet transform (CWT), can convert sequential vibration signals into time-frequency images to realize time-frequency analysis of vibration signals. As shown in Figure 2, the time-frequency image can be transformed from one-dimensional vibration signal. In the process of conversion, Morlet wavelet was selected as the basis function.

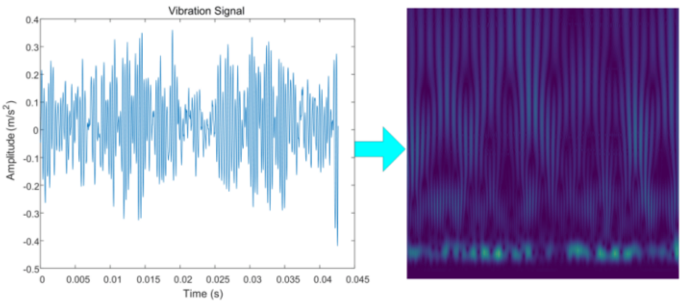


**Figure 2.** Time-frequency image conversion of original vibration signal.

When saving time-frequency images, the aspect ratio of the image needs to be constrained and the image has a certain resolution. In this study, all the time-frequency images are both transformed from the vibration signals with a length of 2048, and the images are saved with a dimension of 620x616x3.

### *3.2. Feature Optimization Selection Based on Pyramid Split Attention*

As is well known, stacked convolution layers can extract features of different levels, significantly improving the diversity of overall features. However, different levels of features have different contribution degrees to model performance, so it is necessary to further optimize the extracted features, increase the weight of favorable features and weaken the weight of invalid features. In contrast to the squeeze and excitation attention, which focuses on different channel characteristics, we explore the relationship between spaces, because local information is more important for bearing state characterization. In this paper, PSA, as an important part of feature extraction, carries out feature weighting for multi-layer convolution features, increases spatial attention on the basis of considering channel attention, and enriches feature space by capturing spatial information of different scales. Divide the model inputs into $S$($S$=3) groups from the channel as shown in the Figure 3. Each group was convolved with convolution kernel, and different convolution kernel sizes are set to obtain receptive fields of different scales, which represent the important information of different scales. Then, the weighting coefficient of each group of channels can be calculated through squeeze and excitation (SE) module, and the weighting coefficients of group $S$ are further normalized and weighted by Softmax.

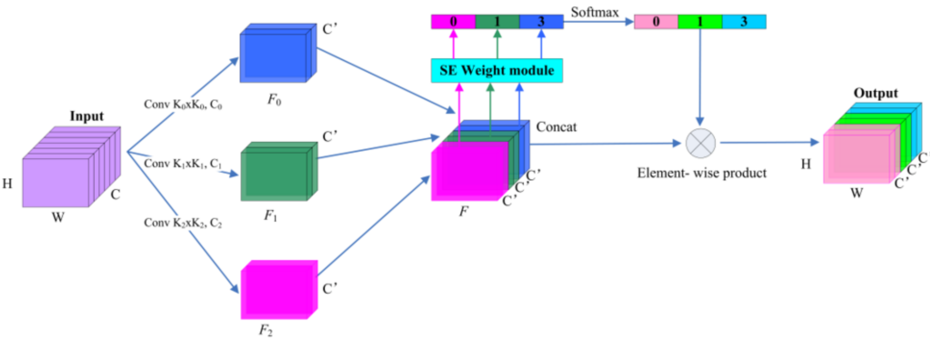


**Figure 3.** The structure of PSA module.

### *3.3. Multi-scale Segmentation Attention-Based Residual Network*

The vibration signal of bearings is a periodic random signal, and the damage of bearing mainly occurs in the following four parts: inner ring, outer ring and rolling body. The complex neural network can extract image features of different channels and levels to fully characterize the state information of bearings. Therefore, a lightweight fault diagnosis model is constructed by referring to the residual network, and the network structure is shown in Table 1.

**Table 1.** Network structure parameters of fault diagnosis model

| Layer Name | Output Size | RNPSA |
|---|---|---|
| conv1 | 112x112 | 7x7, 64, stride 2 |
| | | 3x3 max pool, stride 2 |
| conv2_x | 56x56 | $\begin{bmatrix} 1\times 1, & 32 \\ PSA & 32 \\ 1\times 1, & 64 \end{bmatrix}$ |
| conv3_x | 28x28 | $\begin{bmatrix} 1\times 1, & 32 \\ PSA & 32 \\ 1\times 1, & 64 \end{bmatrix}$ |
| conv4_x | 14x14 | $\begin{bmatrix} 1\times 1, & 32 \\ PSA & 32 \\ 1\times 1, & 64 \end{bmatrix}$ |
| conv5_x | 7x7 | $\begin{bmatrix} 1\times 1, & 32 \\ PSA & 32 \\ 1\times 1, & 64 \end{bmatrix}$ |
| | 1x1 | Average pool, fc(64, 4), softmax |

Based on the proposed fault diagnosis model, image features of different channels and levels can be extracted to fully characterize bearing state information. Since the dimension of model input is 224x224x3, the time-frequency images need to be preprocessed.

### *3.4. Intelligent fault diagnosis method of rolling bearing*

Aiming at the problem of insufficient representation ability of fault diagnosis model, a new intelligent bearing fault diagnosis method is proposed in this paper, which can fully and accurately extract vibration signal features to realize intelligent bearing fault diagnosis. Based on bearing vibration time-frequency images, the CNN with residual learning were used for image feature extraction, and feature optimization selection was carried out in combination with pyramid split attention mechanism to construct an intelligent bearing fault diagnosis model. The specific steps are as follows:

Step 1: Acceleration sensors are used to collect bearing vibration signals, and the time-frequency images can be obtained from vibration signals through CWT time-frequency analysis method.

Step 2: Image feature extraction was carried out based on CNN with residual learning, feature optimization selection was carried out in combination with PSA, and the proposed multi-scale segmentation attention-based residual network can be constructed for fault diagnosis.

Step 3: The multi-scale segmentation attention-based residual network is trained based on the time-frequency images corresponding to different states, and the fault diagnosis model can be achieved when the training is completed.

Step 4: The testing samples can be fed to the trained fault diagnosis model to identify the status of bearing.

The method flow is shown in Figure 4.

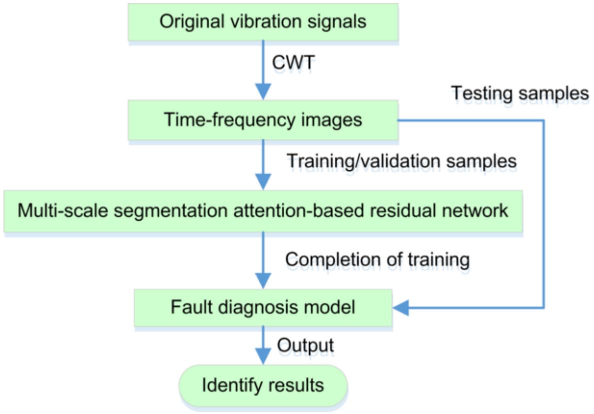


**Figure 4.** Flow of the proposed method.

## 4. Experimental Verification

In this study, the measured data disclosed by case western reserve university (CWRU) was used for experimental analysis, and the results verify the effectiveness of the proposed method.

### *4.1. Test Bench and Data Acquisition*

The test bench is composed of a 1.5kW motor, electronic controller, power tester, torque sensor, etc., as shown in Figure 5. The bearing data of CWRU is collected from SKF6205 rolling bearing. The driving end shaft of motor is supported by the experimental bearing.

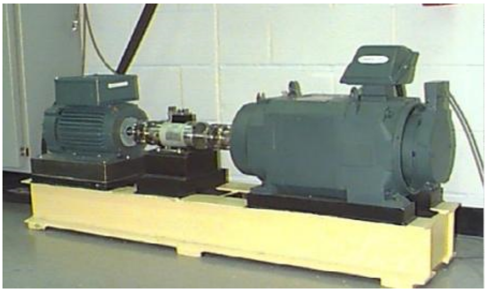

**Figure 5.** Test bench of rolling bearing.

The vibration status of bearing was monitored by acceleration sensors, and the sensor data were collected by the recorder with 16-channels, and the sampling frequency of the bearing at the driving end was 48Khz. At the same time, the power and speed were measured by the torque sensor. By replacing the normal bearing with the machined faulty bearing, the vibration data of the bearing under the fault state can be obtained on the test bench. The vibration data of bearing under different conditions were collected under 2 HP load condition of motor.

### *4.2. Data Processing and Model Training*

First, the original vibration signal of the bearing is intercepted, and the length of the intercepted signal is 2048. The time-frequency image of bearing vibration is obtained by analyzing the time-frequency signal through CWT. For each bearing state, 203 training samples, 87 test samples and 100 verification samples were randomly intercepted. Therefore, the training data set has 812 samples, the test data set has 348 samples, and the verification data set has 400 samples. Then, the image is preprocessed by size adjustment, center clipping, random flipping, normalization, etc., and the training and test data set is constructed. A deep learning classification model was built based on residual learning and pyramid split attention mechanism. The batch size was set as 64 and the learning rate was 1e-4. adam was selected as the optimizer, and the model parameters were trained using time-frequency image data set after model initialization. The early termination criterion of model training is set, and the training stops when the training loss of the model tends to be stable, which corresponds to the bearing fault diagnosis model.

### *4.3. Intelligent Diagnosis of Rolling Bearing Fault*

In order to visualize the feature extraction effect of multi-scale segmentation attention-based residual network, the t-distributed stochastic neighbor embedding algorithm was adopted to simplify the features of the average pool layer, and three-dimensional visualization was performed, as shown in Figure 6.

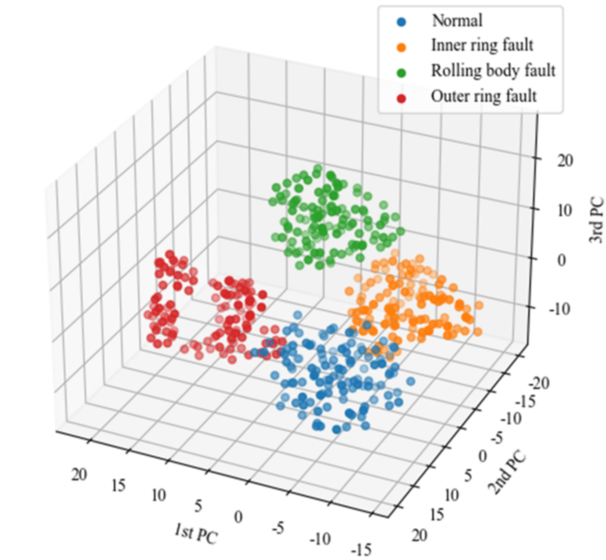


**Figure 6.** Feature visualization.

The trained fault diagnosis model was used to identify the sample of the test set. The result was shown in the confusion matrix in Figure 7, and the fault recognition accuracy reached 99.25%.

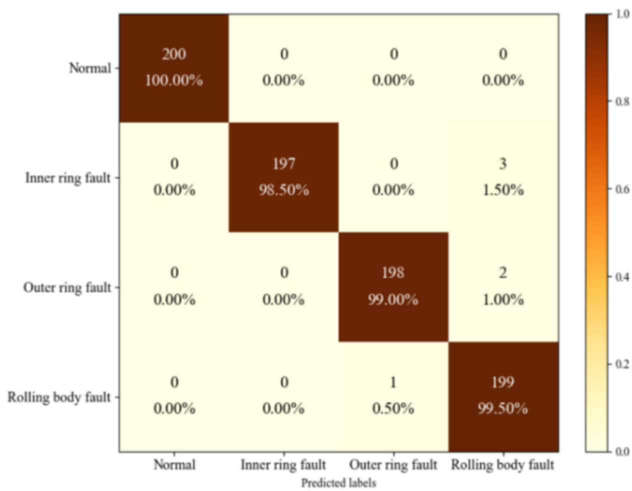


**Figure 7.** The confusion matrix of the results.

## 5. Conclusion

Aiming at the problem of insufficient representation ability of fault diagnosis model, a multi-scale segmentation attention-based residual network is proposed for the fault diagnosis of rolling bearings in this paper. In this method, CWT is used to pretreat the vibration signal of bearing, so as to facilitate feature extraction based on CNN. At the same time, rich features are extracted based on CNN with residual learning, and the pyramid split attention mechanism is used to enhance favorable features and weaken the weight of irrelevant features. Finally, based on CWRU data set, the effectiveness of the proposed method is verified.

## References

[1] A. S. Minhas, S. Singh. A new bearing fault diagnosis approach combining sensitive statistical features with improved multiscale permutation entropy method. *Knowledge-Based Systems*, **218**(2021), 106883.
[2] S. Afrasiabi, M. Mohammadi, M Afrasiabi, et al. Modulated Gabor filter based deep convolutional network for electrical motor bearing fault classification and diagnosis. *IET science, measurement & technology*, **15**(2021), 154-162.
[3] M. Ma, C. Sun, C. Zhang, et al. Subspace-based MVE for performance degradation assessment of aero-engine bearings with multimodal features, *Mech. Syst. Signal Process.* **124** (2019), 298–312.
[4] P. Ding, M. Jia, and X. Yan, Stationary subspaces-vector autoregressive with exogenous terms methodology for degradation trend estimation of rolling and slewing bearings, *Mech. Syst. Signal Process.* **150** (2021), 107293.
[5] M. Zhao, S. Zhong, X. Fu, et al. Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis, *IEEE Trans. Ind. Electron.* **68** (2021), 2587–2597.
[6] V. Sinitsin, O. Ibryaeva, V. Sakovskaya, et al. Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model, *Mech. Syst. Signal Process.* **180** (2022), 109454.
[7] P. Lyu, K. Zhang, W. Yu, et al. A novel RSG-based intelligent bearing fault diagnosis method for motors in high-noise industrial environment, *Advanced Engineering Informatics* **52** (2022), 101564.
[8] M. Ye, X. Yan, N. Chen, et al. Intelligent fault diagnosis of rolling bearing using variational mode extraction and improved one-dimensional convolutional neural network, *Applied Acoustics* **202** (2023), 109143.
[9] X. Li, X. Jiang, Q. Wang, et al. Multi-perspective deep transfer learning model: A promising tool for bearing intelligent fault diagnosis under varying working conditions, *Knowledge-Based Systems* **243** (2022), 108443.
[10] K. He, X. Zhang, S. Ren, et al. Deep residual learning for image recognition, *IEEE Conf. Comput. Vis. Pattern Recognit.* (2016), 770778.