

Article

Research on Intelligent Fault Diagnosis of Rolling Bearing Based on Improved Deep Residual Network

Xinyu Hao ^{1,2,*}, Yuan Zheng ^{1,*}, Li Lu ^{3,*} and Hong Pan ^{4,*}¹ College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China² School of Mechanical Engineering, Yancheng Institute of Technology, Yancheng 224051, China³ China Institute of Water Resources and Hydropower Research, Beijing 100038, China⁴ College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China

* Correspondence: haoxinyu@ycit.edu.cn (X.H.); zhengyuan@hhu.edu.cn (Y.Z.); luli@iwhr.com (L.L.); hongpan@hhu.edu.cn (H.P.)

Abstract: Rolling bearings are the most fault-prone parts in rotating machinery. In order to find faults in time and reduce losses, this paper presents an intelligent diagnosis method for rolling bearings. At present, the deep residual network (RESNET) is the most widely used convolutional neural network (CNN) and has become one of the hotspots in fault diagnosis. However, the fully connected layer of the deep residual network has the disadvantage of too many training parameters, which makes the model training and testing time longer. So, we proposed a new network structure which the global average pooling (GAP) technology replaces the fully connected layer part of the traditional RESNET. It effectively solves the problem of too many parameters of the traditional RESNET model, and uses data enhancement, dropout, and other deep learning training techniques to prevent the model from overfitting. Experiments show that the accuracy of fault diagnosis of the improved algorithm reaches 99.83%, training time has been shortened. Also, the whole process of rolling bearing fault detection does not need any manually extract features, and this “end-to-end” algorithm has good versatility and operability.

Keywords: fault diagnosis; improved deep residual network; deep learning; GAP



Citation: Hao, X.; Zheng, Y.; Lu, L.; Pan, H. Research on Intelligent Fault Diagnosis of Rolling Bearing Based on Improved Deep Residual Network. *Appl. Sci.* **2021**, *11*, 10889. <https://doi.org/10.3390/app112210889>

Academic Editor: Alessandro Di Nuovo

Received: 27 September 2021
Accepted: 11 November 2021
Published: 18 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hydropower Units, Large wind power equipment, and other rotating machinery are developing towards high precision machinery field. A reliable health detecting system is key for the steady operation of mechanical equipment [1]. Rolling bearing affects the overall performance of Rotating machinery [2]. Fault diagnosis of rolling bearings has attracted more and more attention. It can minimize maintenance costs and increase system reliability [3].

Rolling bearing fault diagnosis is mainly to extract features from vibration signals that was collected, and to identify and classify faults. The fault diagnosis methods based on signal processing have been widely studied. A.H. Zamanian, et al. developed a fault diagnosis method based on the Gaussian correlation of vibration signals and wavelet coefficients for gear [4]. Wei Fan a, Gaigai CAI put forward fault diagnosis modus based on sparse representation in wavelet basis for gearbox and it is superior to empirical mode decomposition (EMD) especially in transient feature extraction [5]. Kang, M., Kim, J., Kim, J.-M. proposed a fault diagnosis scheme based on a binary bat algorithm (BBA), that is superior to other dimensionality [6]. Su, Z., et al. uses supervised extended local tangent space alignment (SE-LTSA) for dimensionality reduction to make better the effectiveness of fault diagnosis in rotating machinery [7]. Yu, D., et al. put forward a new morphological component analysis (MCA) method used for the complex fault diagnosis of gearboxes [8]. Dolenc B examines a method based on vibration analysis for diagnosis of distributed bearing faults [9]. An, X., et al. proposed a new vibration analysis method based on the adaptive local iterative filtering used for a hydropower unit [10]. Hu, A., et al. presented a novel method based on linear transformation of intrinsic time-scale decomposition (ITD)

and the cubic spline interpolation. Also, this method was used for diagnosing wind turbine faults with decomposing nonstationary vibration signal, that it did identify wind turbine gearbox fault. The method can solve the problem that is recognizing fault conditions when two or more fractal dimensions are close to each other [11]. From these features, fault detection and classification can be done through various machine learning techniques [12]. A support vector machine (SVM) is to detect bearing faults using harmonics of fault-related frequencies from vibration signals [13]. It also involved an ANN for bearing faults using a genetic algorithm [14]. Support vector machine (SVM) and Artificial neural network (ANN) are the two most popular intelligent diagnosis methods.

Although traditional intelligent diagnosis methods have made a great contribution to rolling element-bearing fault diagnosis, they still have some shortcomings. For example, feature extraction or selection relies heavily on expert knowledge and extensive human labor [15]. Furthermore, Artificial intelligence methods, such as SVM and ANN are shallow learning models. It is difficult for the shallow learning model to learn complex nonlinear relationships effectively [16–18]. Hence, deep learning has been investigated for automatic and effective fault feature learning of rolling-element bearing in recent years.

Deep learning is no longer a new concept [19]. At first, deep learning was used in image processing, audio processing, and other related fields, and has achieved great success [20,21]. Therefore, researchers have introduced deep learning into fault diagnosis. Wang, X., et al. proposed a new data-driven remaining useful life (RUL) estimation approaches of bearings based on Deep Spatiotemporal Convolutional-Neural-Network [22]. Feng, J., et al. applied deep neural networks to conquer the flaw of the before-mentioned intelligent diagnosis methods [23].

Tra, V., et al. presented a novel method using Convolutional Neural Networks and the Stochastic for detecting early bearing defects under changeable operating speeds [24]. Xia, M., et al. developed a new approach based on convolutional neural network (CNN) and multiple sensors for fault diagnosis of rotating machinery [25]. Khan MA et al. proposed a new DL model according to a dilated convolutional neural network (D-CNN) used for bearing faults detection in induction motors (IMs) [26]. Kumar, A., et al. put forward method that was applied to recognize faults of the centrifugal pump using advanced convolution neural network (ACNN) and acoustic images [27]. Shao, Y., et al. combine support vector machine (SVM) and convolutional neural network to propose a new hybrid intelligent fault diagnosis frame which is better than some traditional fault diagnosis methods and has high precision for rolling bearings [28]. Qin, Y., et al. applied the Optimized Deep Belief Networks with Improved Logistic Sigmoid Units for Planetary Gearboxes of Wind Turbines [29]. Jun, P., et al. developed a LiftingNet which achieved layer wise feature learning and effectively classify mechanical failure data even with different speeds and under the effects of random noise [30]. Zhou Q et al. combine convolutional neural networks and nonlinear auto-regression neural networks to put forward a new method for imbalanced fault diagnosis of the rotating machinery [31]. Azamfar, M., et al. put forward a method based on motor current signature analysis and 2-D convolutional neural network used for gearbox fault diagnosis [32]. Xie, S., et al. presented a new convolutional neural network with a one-dimensional structure (ODCNN) for the automatical fault diagnosis of rolling bearings [33]. However, these methods have too many parameters, and the network convergence speed is slow and cannot be used to practical projects.

To solve this issue, this paper developed an improved deep residual network for developing an intelligent fault diagnosis system. By changing the deep residual network structure, the network training time can be shortened, and the fault recognition accuracy can be improved. The novel method can extract features from the raw vibration signal and avoid manual feature extraction. The method is more reliable and effective than traditional fault diagnosis methods. Based on the excellent performance of the improved deep residual network, the rolling element bearing fault detection technology in a one-dimensional vibration signal is studied. Also, from the experimental results, it is validated that our proposed ResNet model is effective. To summarize, our contributions as following:

In this paper, the improved deep residual network structure is proposed, and GAP is introduced to replace the full connection layer. It is an effective method to solve the problem of too many parameters in traditional deep residual network model. The method is validated on rolling bearing fault data of different types. Compared with existing models, the training time of this model is shortened, and the classification accuracy is higher.

The rest of arrangement of this article is as follows. Section 2 presents the background of ResNet. In Section 3, the particulars of the proposed fault diagnosis method are presented. Section 4 presents one comprehensive case study of fault diagnosis for rolling element bearings to illustrate the effectiveness of the presented method. Section 5 is the summary of paper.

2. Deep Residual Network (ResNet)

ResNet was proposed by Dr. He in 2015 [34]. The ResNet model is an updated version of the ConvNet model. However, it is different from traditional deep learning. The ResNet adds identity mappings, it is convenient for the backpropagation of errors and the optimization of model parameters, and further reduces the training difficulty of deep neural networks. It has generated great outcomes in computer vision-related tasks such as image segmentation, image recognition, and target positioning. Therefore, ResNet is used in this study. ResNet principally consist of a certain amount of residual building blocks, several convolutional layers, a global average pooling, and a fully connected output layer. This section introduces the theory of ResNet in detail.

2.1. Convolutional Layer

In the convolutional layer, the inputs and the convolutional kernels are convolved to get the feature maps. Meanwhile, the weights of the convolutional kernels are allocated over the input. So, this significantly reduces the number of parameters required to train. The mathematical form of the convolutional operation can be expressed by:

$$X_{r'}^{(l)} = \sum_{r=1}^K w_{rr'}^{(l)} * x_r^{(l-1)} \tag{1}$$

where l indicates layer number of the network; $w_{kk'}^{(l)}$ is the convolutional kernel, $x_r^{(l-1)}$ is the input of feature map, $X_{r'}^{(l)}$ is the output of feature map, r is the index of the input feature maps, r' is the index of the output feature maps. The convolution operation can also be understood from Figure 1. Input is 4 by 4, $r = 1$, There are two kernels, each kernel is 2 by 2, $r' = 2$.

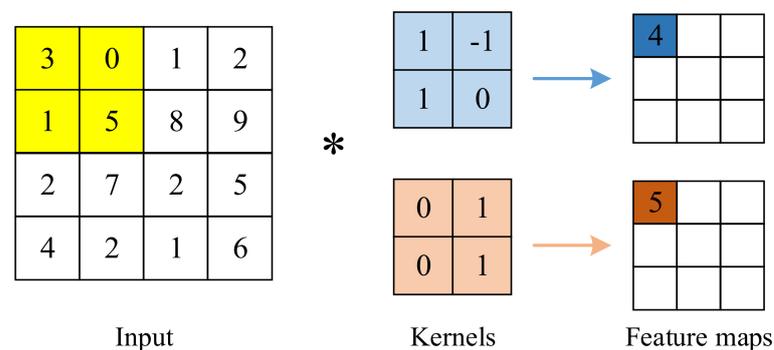


Figure 1. Convolution operation.

2.2. Max Pooling Layer

The pooling layer primarily performs a down-sampling operation. In this study, the input signal is the time domain signal. So, we use the maximum pooling function. The advantage is that you can obtain location-independent features. It's important for time-domain signals. The pooling operation can be understood from Figure 2. We are taking 2×2 region, and taking a stride of 2. Since we start from this pool kernel is like a 2×2

region, which gives you the 9. Also, then, you step it over two steps to look at this region to give you the 2. The mathematical form of the max pooling operation is expressed by:

$$P_{ij}^l = \max(a_{mn}^{l-1}) \tag{2}$$

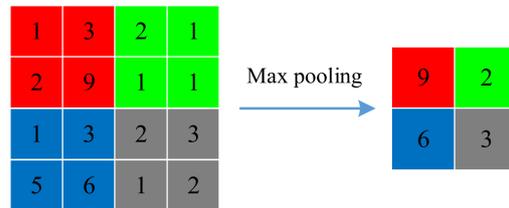


Figure 2. Maximum Pooling operation.

2.3. Residual Building Block

ResNet is often composed of several residual building blocks, and it is the core component of the model. Two common residual building blocks are shown in Figure 3 [35]. Figure 3a is the original residual building block, Figure 3b is the proposed residual building block. The old residual building block and the presented residual building block are consisted of two convolutional layers, two batch normalizations (BN), and two ReLU activation functions, but the location of the ReLU activation function is different. Enter the residual path and the identity mapping passed and add them before the next ReLU activation function of the original building block. BN and ReLU before each convolutional layer in the proposed residual building block. So, the proposed residual building block has a path directly connecting the input and output. It is more conducive to the backpropagation of errors in the neural network, and thus easier to train and improve generalization. Therefore, the proposed residual building block is used in this paper.

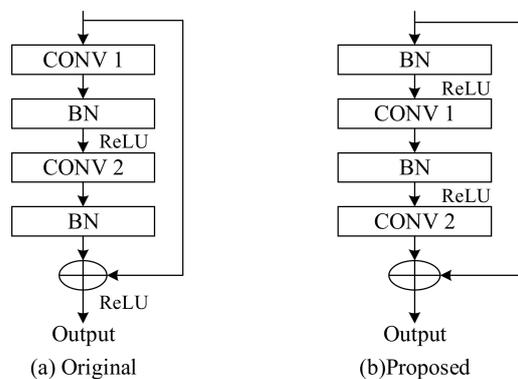


Figure 3. Residual building block. (a) original. (b) Proposed.

The frequently used activation function is the rectified linear unit (ReLU) which can accelerate convergence. ReLU activation function rarely encounter the gradient vanishing problems in that its derivative is either 1 or 0. The activation function can be expressed as:

$$Y(x) = \max \{0, x\} \tag{3}$$

where x is the input of the ReLU and Y(x) is the output of the ReLU, accordingly.

2.4. Global Average Pooling

To avoid overfitting, Global Average Pooling (GAP) layer is adopted in the ResNet. A brief introduction of GAP is presented below. The fully connected layer is usually located in the last two layers of traditional CNN. It can be connected with traditional neural network and convolutional structure. The full-connection layers make predictions, such

as classification while the convolutional layers extract features and output feature maps. However, due to large number of parameters in the full connection layers, overfitting is easy to occur. To address this problem, GAP is introduced into the deep residual network. In CNNs, GAP instead of full connection layer was firstly proposed in their work by Lin et al. [36] GAP averages the feature maps and outputs a single value and to obtain a vector, which can be interpreted as the category of the classification confidence map. GAP layer has no parameters to optimize. So, this greatly decreases the number of parameters to avoid overfitting.

2.5. The Objective Function

First, the output layer uses the softmax activation function. It achieves an event probabilities distribution over different event. The objective function calculates the probability of each target category in all possible target category. Softmax layer operation can be specified as:

$$P(y_i) = \frac{\exp(y_i)}{\sum_{j=1}^k \exp(y_j)} \quad (4)$$

k is the classes number (health status), y_i are input of the softmax function. $P(y_i)$ stand for output feature maps of the softmax function. $P(y_i)$ can be regarded as the reckoned possibility of an observation belonging to the i th class. Then, it calculates loss when training the layer of ResNet. The objective function of ResNet must be reduced to the least for precise data prediction. In multi-class classification problems, Cross-entropy error is usually used as the target to be minimized [37]. A cross-entropy loss function can be presented as:

$$CE = - \sum_{i=1}^k t_i \ln P(y_i) \quad (5)$$

t_i and $P(y_i)$ are the target value and the forecasted value separately.

3. The Proposed Method

In this research, a deep ResNet framework for feature learning and fault diagnosis of rolling bearing is proposed. The structure of the network composed of an input layer, a convolution layer, a max-pooling layer, eight residual blocks, then the following are a GAP and a softmax output layer. The RESNET is the most widely used Convolutional Neural Network (CNN) and has become one of the hotspots in terms of fault diagnosis. However, due to large number of parameters in the full connection layers, the convergence speed is slow during network training. This paper proposed an improved RESNET algorithm for intelligent fault diagnosis of rolling bearings. The complete network structure of the presented ResNet is demonstrated in Figure 4. This method improved the RESNET structure and introduced GAP technology [36] to replace the connection layer part, reducing the amount of training parameters and testing time of the model. The proposed method does not need to perform any manual feature extraction and feature transformation operations for the original data during the entire fault diagnosis process. It only needs to input the original fault data of the rolling bearing into the improved RESNET model, and the fault diagnosis results are automatically output. The “end-to-end” algorithm structure has better operability and versatility.

As demonstrated in Figure 4, the input of this network is the one-dimensional time-domain signal of the rolling bearing fault signal, and the probability distribution of each failure type is the output of the network.

Residual block is consisted of two convolutional operations, ReLU activation functions, batch normalizations (BNs), and one identity shortcut, as shown in Figure 3b. The parameters of the convolution layer in the residual block are shown in Table 1.

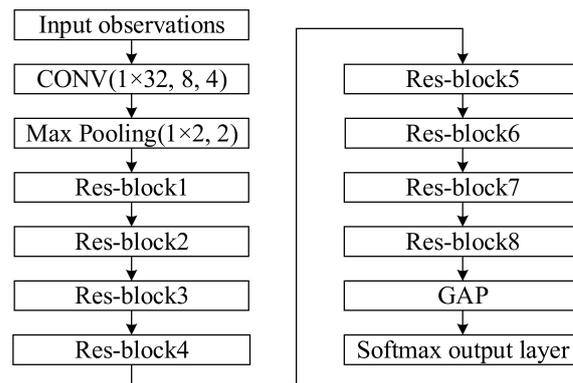


Figure 4. The complete structure of the designed ResNet.

Table 1. The parameters of the convolution layer in the residual block.

Residual Block	Parameters of the Convolution Layer	
	Conv1	Conv2
Res-block1	1×3, 8, 1	1×3, 8, 1
Res-block2	1×3, 8, 1	1×3, 8, 1
Res-block3	1×3, 16, 1	1×3, 16, 1
Res-block4	1×3, 16, 1	1×3, 16, 1
Res-block5	1×3, 32, 1	1×3, 32, 1
Res-block6	1×3, 32, 1	1×3, 32, 1
Res-block7	1×3, 64, 1	1×3, 64, 1
Res-block8	1×3, 64, 1	1×3, 64, 1

The size of convolution kernel in the residual block is all 1×3 , the quantity of convolution kernels is 1, and the main difference lies in the stride.

4. Experimental Verifications

Programming with open source Python language (version 3.5) and TensorFlow (version 2.0) toolkit from Google to realize the ResNet model. TensorFlow, developed by Google, is an open-source machine learning library based on TensorFlow graphs. It has the function of automatically solving the reverse gradient to optimize the model parameters (weights and bias), and is suitable for the rapid development of deep learning algorithms. At the same time, the TensorFlow toolkit supports large-scale and fast matrix computations based on image processing units, greatly reducing the training time required for the ResNet algorithm.

4.1. Experimental Data Collection

The Case Western Reserve University Bearing Data Center provides experimental data. As shown in Figure 5, the test stands mainly composed of a torque transducer/encoder (center), a 2 hp motor (left), and a dynamometer (right). Motor bearings were used to inoculate faults by electro-discharge machining (EDM). The fault 0.007 inches, 0.014 inches, and 0.021 inches in diameter were introduced separately at the inner raceway, rolling element (i.e., ball), and outer raceway. Therefore, totally we have one normal condition (no-fault) and nine different faults. Bearings are tested under different conditions consisting of normal condition, ball fault (BF), outer race fault (OR), and inner race fault (IR).

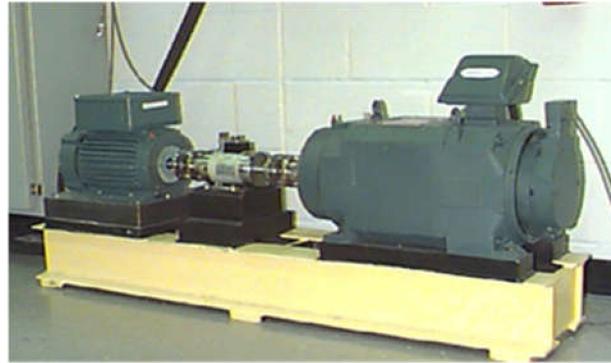


Figure 5. Test-stand.

Vibration data was collected for motor loads of 3 horsepower (motor speeds of 1730 RPM) at 12,000 samples per second.

For training our proposed ResNet model, we employ enough training samples.

4.2. Data Preprocessing

4.2.1. Signal Normalization

In order to increase the reliability of the model, the input signal needs to be normalized. As shown in the following equation.

$$\lambda = \frac{x - \bar{x}}{\max(x) - \min(x)} \quad (6)$$

4.2.2. Data Augmentation

Data augmentation is overlapping sampling, that is, for training samples, when training samples are recorded from the original signal, it has overlap between each segment of signal and the next segment of the signal. When the step is smaller than the signal length of a single sample, there is overlap between samples, and more samples can be extracted with the fixed-length signal. This is shown in Figure 6. When the step is the same as the data length, there is no data augmentation. In our case, the stride is 28. For each type of vibration signal collected, data segmentation is performed. A point is inserted randomly first, and then 1024 points are taken. In this way, 600 samples can be obtained after repeated operation 600 times. Since there are 10 kinds of signals, a total of 6000 samples are obtained, and then training set 4200, verification set 1200 and test set 600 are divided according to the proportion of 7:2:1. The composition of experimental sample data is shown in Table 2. The vibration data waveform of the bearing in 10 states is shown in Figure 7. Figure 7a1 is the waveform of the normal state. Figure 7a2–a10 are waveforms of fault state.

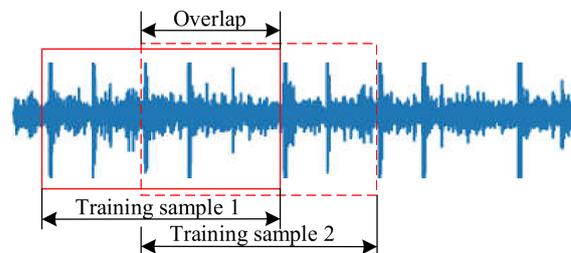


Figure 6. Schematic diagram of sample augmentation.

Table 2. Composition of experimental samples.

Condition Type	Size of Samples	Sample Number	Label
Ball (0.007")	1024	600	F1
Ball (0.014")	1024	600	F2
Ball (0.021")	1024	600	F3
Inner Race (0.007")	1024	600	F4
Inner Race (0.014")	1024	600	F5
Inner Race (0.021")	1024	600	F6
Outer Race (0.007")@6	1024	600	F7
Outer Race (0.014")@6	1024	600	F8
Outer Race (0.021")@6	1024	600	F9
Normal condition	1024	600	H0

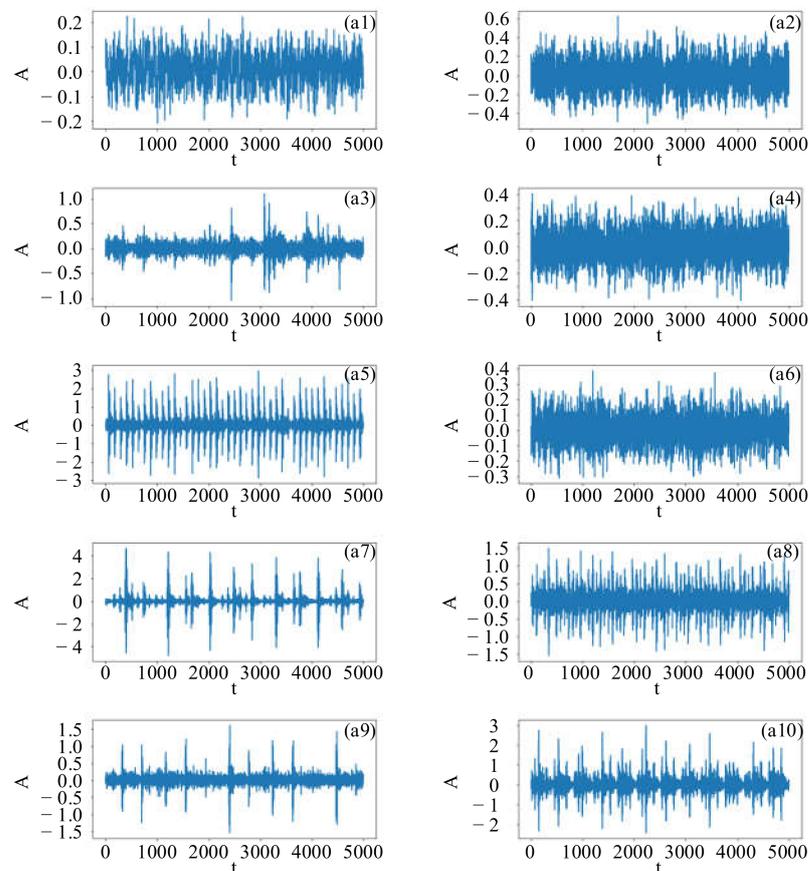


Figure 7. Vibration waveforms of rolling bearing in 10 states (a1) Normal condition. (a2) BF 0.007 inch. (a3) BF 0.014 inch. (a4) BF 0.021 inch. (a5) OR 0.007 inch. (a6) OR 0.014 inch. (a7) OR 0.021 inch. (a8) IR 0.007 inch. (a9) IR 0.014 inch. (a10) IR 0.021 inch.

4.2.3. Dropout

Among the deep learning algorithms in recent years, Dropout is a commonly used method to reduce overfitting [38]. During each training iteration, some neurons are dropped randomly by Dropout, so that the neural network only propagates forward and updates backward the parameters of retained neurons. In this way, Dropout can weaken the “cooperative relationship” between neurons and make each neuron function more independently, thus achieving the effect of model regularization. In this study, the dropout is used during training and not during testing. The dropout rate is set to 0.5 [38]. Dropout technologies were used to solve the overfitting problem.

4.3. Hyperparameter Setup

Hyperparameters have a great influence on the fault diagnosis accuracy of residual neural networks. According to the paper [39], the more important hyperparameters are optimizer, learning rate, activation function, convolution kernel, and pooling kernel. The hyperparameters of the experiments were made on the basis of the empirical recommendation. In our case, ReLU activation functions are selected. The convolution kernel and pooling kernel are shown in Figure 2. Adam optimizer was used in the experiment. The learning rate is too large to convergence, too small training is too slow. In this study, the exponential decay learning rate is used to optimize this problem. Set the initial learning rate to 0.01, then gradually reduce the learning rate through iteration, and the attenuation coefficient is 0.99. The best classification effect can be achieved when the learning rate is 0.001.

4.4. Outcome of Experiment

The established ResNet method was adopted in fault diagnosis based on vibration using a dataset collected from bearings. In the training of the improved ResNet model, the Adam optimization algorithm was also adopted to improve the overfitting problem. The Adam (Adaptive Moment Estimation) algorithm is an algorithm that combines Momentum algorithm and RMSProp algorithm. Also, data augmentation and Dropout technology were used to improve the overfitting problem. The mini-batch was 16, the ReLU activation function was used, and the number of cycle iterations was 100 rounds. The final improved ResNet has the highest accuracy of 99.83% on the test set, and its diagnostic results are shown in Table 3.

Table 3. Comparison of fault diagnosis results.

Method	Testing Accuracy	Training Time per Model	Testing Time per Observation
ResNet + GAP	99.83%	228.53 s	0.193 s
ResNet + FC	98.48%	245.96 s	0.270 s

By comparing with Table 3, it can be seen that the performance of the improved RESNET algorithm is significantly improved compared with the traditional fully connected RESNET algorithm. In terms of time, the number of model parameters is greatly reduced and the training time is significantly reduced in the improved RESNET algorithm because the full connection layer is removed, which is of great significance for the model to be applied to the online rapid diagnosis and monitoring of faults. In terms of accuracy, the accuracy of the improved RESNET algorithm has reached 99.83%, while the accuracy of the traditional RESNET algorithm is 98.48%; Figure 8 shows the training and testing result of one trial. Figure 8a shows the relationship between the epoch and the accuracy of the model. Figure 8b shows the relationship between the epoch and the cross entropy.

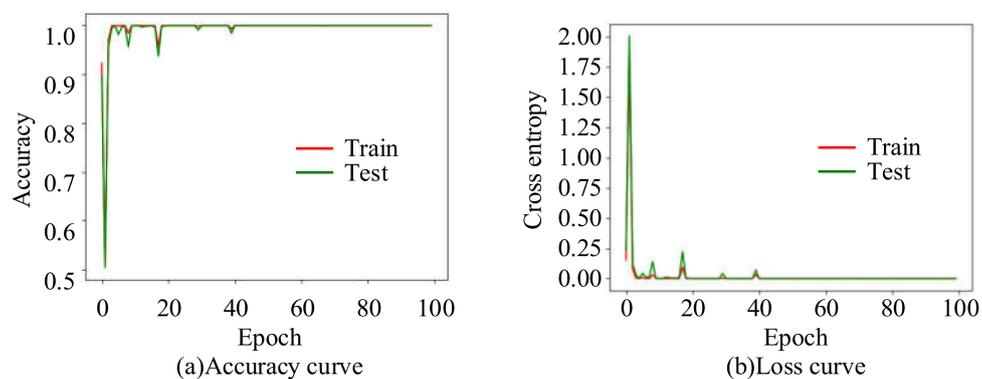


Figure 8. Experimental result of bearing dataset. (a) Accuracy curve. (b) Loss curve.

4.5. Discussion

In order to evaluate the accuracy of the developed ResNet more effectively, we quote the precision and recall to evaluate the algorithm [40–42]. Precision is the probability of actually positive samples out of all the predicted positive samples. The recall is the probability of being predicted to be positive samples out of actually positive samples. Precision and recall are shown as follows.

$$P = \frac{TP}{(TP + FP)} \quad (7)$$

$$R = \frac{TP}{(TP + FN)} \quad (8)$$

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (9)$$

where P is precision, R is recall. TP stands for actual 1, predicted 1, predicted correctly. FP stands for actual 0, predicted 1, forecast wrong. FN stands for actual 1, predicted 0, forecast wrong. $F1$ is the harmonic mean of precision and recall. In this study, According to Equation (7) and (8), the precision and recall are calculated only based on the experimental results of improved ResNet in Table 4. As shown in Table 4:

Table 4. The diagnosis result evaluation of improved ResNet.

Label	Precision (%)	Recall (%)	Observations
H0	98.36	100	600
F1	100	100	600
F2	100	100	600
F3	100	100	600
F4	100	100	600
F5	100	100	600
F6	100	100	600
F7	100	98	600
F8	100	100	600
F9	100	100	600
Average	99.83	99.8	600

In order to further show the ability of the improved ResNet algorithm to identify minor faults and the details of the fault misjudgment, we introduce the multi-classification confusion matrices [37] to conduct a detailed quantitative analysis of the fault results. The confusion matrices comprehensively reflect the diagnosis precision and the number of misjudgments of bearings under different fault grades, as well as the information of the real fault types being misjudged. The confusion matrices quantization diagram of bearing corresponding to Table 4 is shown in Figure 9.

The confusion matrix of the improved ResNet is shown in Figure 9. The X-axis represents the predicted category of the fault, and the Y-axis represents the true label of the fault. The numbers on the main diagonal represent the accuracy of the improved ResNet algorithm for the correct diagnosis of each type of fault state.

Depending on the confusion matrix, two among the ten health states present misclassifications. F1 was misjudged as F2, F7 misjudged as F4. By analyzing the types of fault misclassifications, it can be seen that the above misclassifications are all errors between different fault categories, which basically belong to small faults misjudged as larger faults, which is meaningful for risk prediction. The recognition accuracy of this algorithm is 100% between the normal state and the fault state. It can be seen that the comprehensive fault identification rate can reach 99.8%. Experimental show that the improved ResNet algorithm has superior identification ability and higher diagnostic accuracy for the micro-faults of rolling bearings.

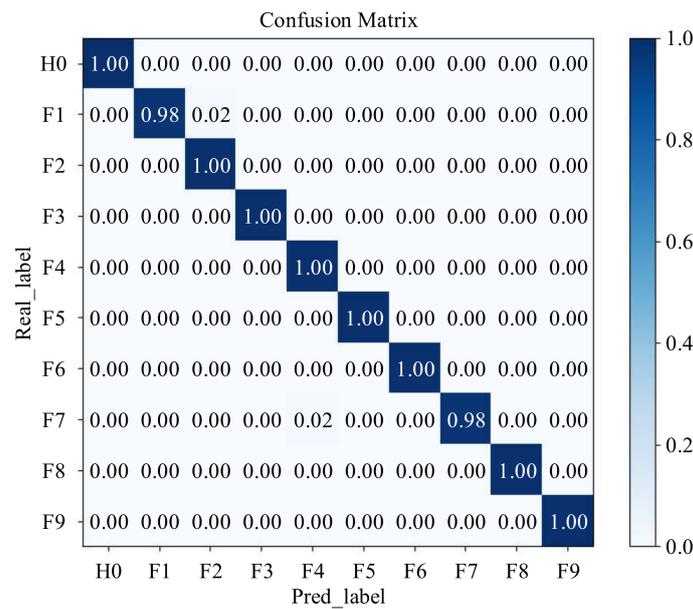


Figure 9. Confusion matrix demonstrating classification performance using testing data.

4.6. Performance Comparison

We did a comparison for the designed method with the modified CNN, SVM, KNN and DPBN. The results are shown in Table 5. The modified CNN gets an accuracy of 98.2%. The accuracy achieved using KNN is 91.9%. The accuracy of Support vector machine (SVM) is 94.1%. The accuracy of the DPBN is 92.3%. The comparison confirms that the improved ResNet method designed in this paper conducts better performance than the methods such as DPBN, SVM, and KNN and the existing CNN.

Table 5. A comparison of the proposed method Vs existing methods.

Diagnosis	The Improved ResNet	The Improved CNN	SVM	KNN	DPBN
Training time	228.53 s	229.53 s	2.264 s	0.049 s	327.09 s
Testing time	0.193 s	0.198 s	0.829 s	1.322 s	0.141 s
Fault category	Accuracy %				
H0	100	100	100	96.25	100
F1	100	100	96.72	78.00	92.86
F2	98	92.01	92.01	86.31	93.55
F3	100	98.36	98.31	90.15	98.36
F4	98	98.36	100	97.00	100
F5	100	98.33	93.25	92.85	66.59
F6	100	96.77	95.24	96.59	96.08
F7	100	100	87.32	100	100
F8	100	100	100	100	100
F9	100	98.31	78.59	81.69	75.56
Averages	99.6	98.2	94.1	91.9	92.3

This has been possible for the reason of the GAP involved in the improved ResNet. The GAP assures fewer amount of training parameters are needed and keep away overfitting troubles to training data. This GAP realizes deep learning and guarantees excellent defect identification result even if the data is unseen.

5. Conclusions

This paper developed an improved deep residual network for defects identification in the bearing. Modeling of modified ResNet is produced by using vibration signals attained by celerometers. Conclusions of the study are as following:

RESNET acquired better performance by modifying its FC layer. GAP replaced the full connection layer part of the traditional RESNET model. Also, accordingly, the amount of training parameters is reduced and over-fitting of ResNet is avoided. This GAP realizes deep learning and guarantees high accuracy of defect identification even if the data is unseen. A contrast has also been done for the proposed method with the present machine learning methods and deep learning methods. Result states that the reliability of the designed method is up to 99.8%, which is much higher than the present machine learning methods and the existing deep learning methods when to ascertain defects of the rolling bearing.

The improved ResNet algorithm needn't to do any manual feature extraction of the original fault data but inputs the original fault data directly as the model, then automatically outputs the fault classification results. The "end-to-end" model has much better versatility and operability. Furthermore, Dropout, adaptive variable learning rate and data enhancement can also be used to effectively decrease training parameters and calculation time of the model while preventing model overfitting

This method is verified based on experimental data, and the actual engineering data needs to be verified.

Author Contributions: Formal analysis, X.H., L.L., and; funding acquisition, Y.Z., H.P.; writing—original draft, X.H.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the "National Key Research and Development Program of China (2019YFE0105200)"; "National Natural Science Foundation of China (51809082)".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this paper are acquired from the Bearing Data Center of Case Western Reserve University (CWRU) and web page: <http://csegroups.case.edu/bearingdatacenter/home> (accessed on 1 August 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lei, Y.; Jia, F.; Zhou, X.; Lin, J. A Deep Learning-based Method for Machinery Health Monitoring with Big Data. *J. Mech. Eng.* **2015**, *51*, 49–56. [[CrossRef](#)]
2. El-Thalji, I.; Jantunen, E. A summary of fault modeling and predictive health monitoring of rolling element bearings. *Mech. Syst. Signal Process.* **2015**, *60–61*, 252–272. [[CrossRef](#)]
3. Hazra, B.; Narasimhan, S. Gearbox Fault Detection Using Synchro-squeezing Transform. *Procedia Eng.* **2016**, *144*, 187–194. [[CrossRef](#)]
4. Zamanian, A.H.; Ohadi, A. Gear fault diagnosis based on Gaussian correlation of vibrations signals and wavelet coefficients. *Appl. Soft Comput. J.* **2011**, *11*, 4807–4819. [[CrossRef](#)]
5. Wei, F.; Cai, G.; Zhu, Z.K.; Shen, C.; Huang, W.; Shang, L. Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction. *Mech. Syst. Signal Process.* **2015**, *5*, 230–245.
6. Kang, M.; Kim, J.; Kim, J.M. Reliable fault diagnosis for incipient low-speed bearings using fault feature analysis based on a binary bat algorithm. *Inf. Sci.* **2015**, *294*, 423–438. [[CrossRef](#)]
7. Su, Z.; Tang, B.; Deng, L.; Liu, Z. Fault diagnosis method using supervised extended local tangent space alignment for dimension reduction. *Measurement* **2015**, *62*, 1–14. [[CrossRef](#)]
8. Yu, D.; Wang, M.; Cheng, X. A method for the compound fault diagnosis of gearboxes based on morphological component analysis. *Measurement* **2016**, *91*, 519–531. [[CrossRef](#)]
9. Dolenc, B.; Bošković, P.; Juričić, Đ. Distributed bearing fault diagnosis based on vibration analysis. *Mech. Syst. Signal Process.* **2016**, *66/67*, 521–532. [[CrossRef](#)]
10. An, X.; Yang, W.; An, X. Vibration signal analysis of a hydropower unit based on adaptive local iterative filtering. *ARCHIVE Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2016**, *231*, 1339–1353. [[CrossRef](#)]

11. Hu, A.; Yan, X.; Xiang, L. A new wind turbine fault diagnosis method based on ensemble intrinsic time-scale decomposition and WPT-fractal dimension. *Renew. Energy* **2015**, *83*, 767–778. [CrossRef]
12. Zhang, X. Introduction to Statistical Learning Theory and Support Vector Machines. *Acta Autom. Sin.* **2000**, *26*, 32–42.
13. Ben Salem, S.; Bacha, K.; Chaari, A. Support vector machine-based decision for mechanical fault condition monitoring in induction motor using an advanced Hilbert-Park transform. *Isa Trans.* **2012**, *51*, 566–572. [CrossRef]
14. Unal, M.; Onat, M.; Demetgul, M.; Kucuk, H. Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network. *Measurement* **2014**, *58*, 187–196. [CrossRef]
15. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
16. Kang, M.; Kim, J.; Kim, J.-M.; Tan, A.C.C.; Kim, E.Y.; Choi, B.-K. Reliable Fault Diagnosis for Low-Speed Bearings Using Individually Trained Support Vector Machines with Kernel Discriminative Feature Analysis. *IEEE Trans. Power Electron* **2015**, *30*, 2786–2797. [CrossRef]
17. Jegadeeshwaran, R.; Sugumaran, V. Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines. *Mech. Syst. Sig. Process.* **2015**, *52*, 436–446. [CrossRef]
18. Pandya, D.; Upadhyay, S.; Harsha, S. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN. *Expert Syst. Appl.* **2013**, *40*, 4137–4145. [CrossRef]
19. Khan, S.; Yairi, T. A review on the application of deep learning in system health management. *Mech. Syst. Signal Process.* **2018**, *107*, 241–265. [CrossRef]
20. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
21. Li, J.; Feng, J.; Kuo, C.C.J. Deep convolutional neural network for latent fingerprint enhancement. *Signal Processing. Image Commun.* **2018**, *60*, 52–63. [CrossRef]
22. Wang, X.; Wang, T.; Ming, A.; Han, Q.; Li, A. Deep Spatiotemporal Convolutional-Neural-Network-Based Remaining Useful Life Estimation of Bearings. *Chin. J. Mech. Eng.* **2021**, *34*, 62. [CrossRef]
23. Jia, F.; Lei, Y.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **2016**, *72–73*, 303–315. [CrossRef]
24. Tra, V.; Kim, J.; Khan, S.A.; Kim, J.M. Bearing Fault Diagnosis under Variable Speed Using Convolutional Neural Networks and the Stochastic Diagonal Levenberg-Marquardt Algorithm. *Sensors* **2017**, *17*, 2834. [CrossRef]
25. Xia, M.; Li, T.; Xu, L.; Liu, L.; De Silva, C.W. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. *IEEE/ASME Trans. Mechatron.* **2017**, *23*, 101–110. [CrossRef]
26. Khan, M.A.; Kim, Y.H.; Choo, J. Intelligent fault detection using raw vibration signals via dilated convolutional neural networks. *J. Supercomput.* **2018**, *76*, 8086–8100. [CrossRef]
27. Kumar, A.; Gandhi, C.P.; Zhou, Y.; Kumar, R.; Xiang, J. Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images. *Appl. Acoust.* **2020**, *167*, 107399. [CrossRef]
28. Shao, Y.; Yuan, X.; Zhang, C.; Song, Y.; Xu, Q. A Novel Fault Diagnosis Algorithm for Rolling Bearings Based on One-Dimensional Convolutional Neural Network and INPSO-SVM. *Appl. Sci.* **2020**, *10*, 4303. [CrossRef]
29. Qin, Y.; Wang, X.; Zou, J. The Optimized Deep Belief Networks with Improved Logistic Sigmoid Units and Their Application in Fault Diagnosis for Planetary Gearboxes of Wind Turbines. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3814–3824. [CrossRef]
30. Pan, J.; Zi, Y.; Chen, J.; Zhou, Z.; Wang, B. LiftingNet: A Novel Deep Learning Network with Layerwise Feature Learning from Noisy Mechanical Data for Fault Classification. *IEEE Trans. Ind. Electron.* **2018**, *65*, 4973–4982. [CrossRef]
31. Zhou, Q.; Li, Y.; Tian, Y.; Jiang, L. A novel method based on nonlinear auto-regression neural network and convolutional neural network for imbalanced fault diagnosis of rotating machinery. *Measurement* **2020**, *161*, 107880. [CrossRef]
32. Azamfar, M.; Singh, J.; Bravo-Imaz, I.; Lee, J. Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis. *Mech. Syst. Signal Process.* **2020**, *144*, 106861. [CrossRef]
33. Xie, S.; Ren, G.; Zhu, J. Application of a new one-dimensional deep convolutional neural network for intelligent fault diagnosis of rolling bearings. *Sci. Prog.* **2020**, *103*, 36850420951394. [CrossRef]
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
35. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, the Netherlands, 8–16 October 2016; pp. 630–645.
36. Lin, M.; Chen, Q.; Yan, S. Network In-Network. 2013. Available online: <https://arxiv.org/pdf/1312.4400.pdf> (accessed on 1 August 2021).
37. Zhao, M.; Tang, B.; Deng, L.; Pecht, M. Multiple Wavelet Regularized Deep Residual Networks for Fault Diagnosis. *Measurement* **2019**, *152*, 107331. [CrossRef]
38. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
39. Zhao, M.; Kang, M.; Tang, B.; Pecht, M. Deep Residual Networks with Dynamically Weighted Wavelet Coefficients for Fault Diagnosis of Planetary Gearboxes. *IEEE Trans. Ind. Electron.* **2018**, *65*, 4290–4300. [CrossRef]
40. Flach, P. *Machine Learning: The Art and Science of Algorithms That Make Sense of Data*; Cambridge University Press: Cambridge, UK, 2012.

-
41. Ouadine, A.Y.; Mjahed, M.; Ayad, H.; El Kari, A. Aircraft Air Compressor Bearing Diagnosis Using Discriminant Analysis and Cooperative Genetic Algorithm and Neural Network Approaches. *Appl. Sci.* **2018**, *8*, 2243. [[CrossRef](#)]
 42. Zhuang, Z.; Lv, H.; Xu, J.; Huang, Z.; Qin, W. A Deep Learning Method for Bearing Fault Diagnosis through Stacked Residual Dilated Convolutions. *Appl. Sci.* **2019**, *9*, 1823. [[CrossRef](#)]