IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 27, 2020, accepted September 29, 2020, date of publication October 2, 2020, date of current version October 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3028465

Input Feature Mappings-Based Deep Residual Networks for Fault Diagnosis of Rolling Element Bearing With Complicated Dataset

LIANGSHENG HOU, RUIZHENG JIANG, YANGHUI TAN, AND JUNDONG ZHANG
College of Marine Engineering, Dalian Maritime University, Dalian 116026, China
Corresponding author: Jundong Zhang (zhjundongpaper@163.com)

This work was supported in part by the Research on Intelligent Ship Testing and Verification under Grant [2018]473.

ABSTRACT Most rolling element bearing (REB) fault diagnosis algorithms are evaluated on the Case Western Reserve University (CWRU) bearing dataset for its popularity and simplicity. However, the diagnosis accuracy on CWRU bearing dataset is overly saturated; it is nearly up to 100%. In this study, an input feature mappings (IFMs)-based deep residual network (ResNet) is proposed to conduct detailed and comprehensive fault diagnosis on REB with complicated bearing dataset. Firstly, a new data preprocessing method named as a signal-to-IFMs method is proposed to automatically extract features from raw signals without predefined parameters. Then, a deep ResNet is used as the fault classifier to learn the discriminative features from IFMs and identify the faults of REB. Finally, the proposed model is evaluated on the artificial, real, and mixed damages of the Paderborn university bearing dataset. The proposed method yields the average testing accuracies of 99.7%, 99.7%, and 99.81% in artificial, real, and mixed bearing damages, which outperforms other methods.

INDEX TERMS Rolling element bearing, fault diagnosis, signal-to-input feature mappings, deep residual networks.

I. INTRODUCTION

Rotating machinery is widely used in modern industries. Rolling element bearing (REB) are the key component of rotating machinery to maintain the normal operation of rotating machinery. The REB faults account for 45%-55% of total failures in rotating machinery [1]. Once the REB fails, it will cause huge economic and irreversible damages to rotating machinery. Hence, practical REB fault diagnosis is essential to rotating machinery.

The model-based methods are popular in REB fault diagnosis; they use mathematical, experimental, and simulation models to detect the REB damages [2]. However, one problem exists with model-based methods. The model-based methods rely on the constructed models. That is to say, a low-quality model can yield a poor diagnosis result.

To overcome the drawback of model-based methods, machine learning (ML) methods were developed. ML methods can directly learn the fault discriminative features from the historical data without considering prior models [3]. Many ML algorithms have been used for REB fault diagnosis, and achieved considerable success [4]–[12]. Van *et al.* [13] proposed a support vector machine (SVM)-based model for REB fault diagnosis. Firstly, they used the nonlocal means method and empirical mode decomposition to extract the fault features from the raw signals. Then, they selected the optimal fault features using the maximum-relevance technology. Finally, their method yielded an ideal diagnostic result. Zhao *et al.* [14] proposed a new REB fault diagnosis model based on hidden Markov model (HMM). Firstly, they extracted the time- and frequency-domain components from the raw signals. Then, they used principal component analysis method to get low-dimension features. Finally, they achieved a satisfactory fault diagnosis result using the HMM technology. Li *et al.* [15] developed a new feature learning method for REB fault diagnosis. Firstly, they used wavelet multiscale transform to decompose the vibration signals. Then, they used back-propagation neural network to learn the discriminative features. Finally, they selected SVM as the fault classifier to detect the REB faults. Nevertheless, one limitation exists

The associate editor coordinating the review of this manuscript and approving it for publication was Sudipta Roy.

TABLE 1. Comparison of two bearing datasets.

Benchmark dataset	Damage type	Damage method	Collected signal
CWRU bearing dataset	Artificial damages	EDM	Vibration signal
Paderborn university bearing dataset	Artificial damages	EMD; drilling; electric engraver; fatigue: pitting; plastic deform: indentations	Current signals; vibration signal
	Real damages		

with conventional ML methods; they need to manually select fault features [15]. That is to say, the same fault classifier with different feature extraction methods can yield different diagnosis results.

Recently proposed deep learning (DL) methods have resolved the limitation of conventional ML methods. DL methods can automatically learn abstract representation features without manually selecting fault features, and yield the best-in-class performance [16]. Various DL algorithms have been used for REB fault diagnosis [17]–[26]. Chen and Li [27] proposed a fusion method to fuse the vibration signals from time and frequency domains. Then, they used a stacked autoencoder to extract the fault features. Finally, they used a deep belief network (DBN) as the fault classifier to classify the REB faults. Their model yielded a high diagnosis accuracy of 97.82%. Zhao *et al.* [28] proposed a gated recurrent unit-based sparse autoencoder to conduct REB fault diagnosis. They used grey wolf optimizer algorithm to select the optimal model parameters. The optimized model yielded the satisfactory results. Janssens *et al.* [29] used convolutional neural network (CNN) for REB fault diagnosis. Firstly, they used the discrete Fourier transform method to process the raw signals. Then, they used CNN to classify REB faults. Their method yielded a high testing accuracy.

Most current diagnostic algorithms are evaluated on Case Western Reserve University (CWRU) bearing dataset. The accuracy of REB fault diagnosis on CWRU bearing dataset is already overly saturated [30]. The highest accuracy of 99.99% has been yielded by Wen *et al.* [31]. Recently proposed Paderborn university bearing dataset has limited researches. The data from Paderborn university bearing dataset are more complicated than CWRU bearing dataset [30]. Table 1 represents the comparison of two bearing benchmark datasets. Only artificial damages using electro-discharge machining (EDM) exist in the CWRU bearing dataset. In Paderborn university bearing dataset, the REB damage contains artificial and real damages, and the damage method consists of EMD, drilling, electric engraver, pitting, and indentations. Only the vibration signals are collected in the CWRU bearing dataset, while both vibration and current signals are collected in the Paderborn university bearing dataset.

This study aims to conduct REB fault diagnosis with complicated bearing dataset. To improve the diagnosis accuracy of complicated bearing dataset, an input feature mappings-based deep residual network is proposed. Firstly, a new data reprocessing method called signal-to-input feature mappings (IFMs) is proposed to convert different signals to the IFMs without any predefined parameters and experts' experiences. Then, a deep residual network (ResNet) is designed as the fault classifier to automatically learns abstract discriminative features from the IFMs. Finally, the proposed model is evaluated on the Paderborn university bearing dataset. to detect the artificial, real, and mixed damages of REBs. The experimental results show that the proposed method outperforms other methods. To our best knowledge, it is the first attempt to conduct REB fault diagnosis on the Paderborn university bearing dataset in detail.

The remainder of this paper is organized as follows. The theoretical background is briefly introduced in Section II. In Section III, the proposed method for REB fault diagnosis is described. The experimental investigation is conduct in Section IV. Section V gives the experimental results and result analysis. Finally, conclusions are given in Section VI.

II. THEORETICAL BACKGROUND

In this section, the CNN and deep ResNet are briefly introduced.

A. CONVOLUTIONAL NEURAL NETWORK

CNN is inspired by the mechanism of the receptive biological field [32]. Different from the traditional neural network, CNN has fewer weight parameters and connections in the neural nodes. CNN is a feedforward neural network which consists of convolutional layer (CL), pooling layer (PL), and fully connected layer (FCL). The CL learns the feature representations from the input images; it extracts the features from a local region, and different convolution kernels have different feature extractors, as shown in Eq. (1).

$$y_k = \mathrm{f}(w_{k*}x + b_k) \tag{1}$$

where x is the input image, w_k stands for the convolutional filter, $\mathrm{f}(\cdot)$ denotes the nonlinear activation function, b_k represents the bias.

Although CL can reduce the number of connections in the network, the number of neurons in the feature mappings doesn't significantly decrease. To resolve this problem, the PL is added after the CL to reduce the feature dimensions and avoid overfitting. The PL is used for feature selection; it can reduce the number of feature dimensions and network parameters. The pooling operator converts the nearby feature values to one value. The pooling operator consists of maximum and mean pooling, as shown in Eq. (2) and (3):

$$y_{kmn} = maxx_i, i \in R_{kmn} \tag{2}$$

$$y_{kmn} = \frac{1}{|R_{kmn}|}\sum_{i \in R_{kmn}} x_i \tag{3}$$

where y_{kmn} denotes the pooling output in the kth feature mapping, x_i is the neuron activation value in the pooling region. R_{kmn} is the pooping region with the length m and width n.

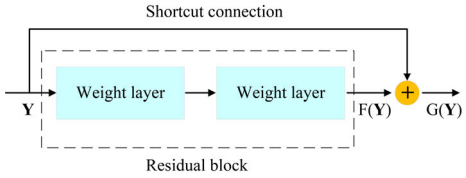

FIGURE 1. Residual block.

FCL is used for classification or regression; it is a traditional multilayer perceptron that all nodes are connected. Softmax function is generally used on the output layer.

B. DEEP RESIDUAL NETWORK

CNN suffers from gradient vanishing as the growing CLs, which results in the low accuracy and overfitting. In 2016, He *et al.* [33] proposed a deep ResNet by adding the residual blocks to CNN. The deep ResNet resolves the gradient vanishing problem. A shortcut connection connects the input and output of the residual block, as shown in Figure 1. The output of residual block can be calculated by Eq. (4).

$$\mathrm{G}(\mathbf{Y}) = \mathrm{F}(\mathbf{Y}) + \mathbf{Y} \tag{4}$$

where $\mathbf{Y}$ is the input of residual block, $\mathrm{F}(\mathbf{Y})$ is residual mapping.

The conventional deep ResNet consists of the input layer, CL, PL, FCL, and output layer [33]. The residual block contains the CLs with a shortcut connection. The gradient in deep ResNet never disappears due to the special architecture. Deep ResNet significantly improve the performance of DL.

III. PROPOSED METHOD

A. PROPOSED SIGNAL-TO-IFMs METHOD

This paper proposes a new data preprocessing method to convert different kinds of raw signals to the IFMs. The schematic illustration of signal-to-IFMs conversion is shown in Figure 2. Firstly, the raw signals $1, 2, \ldots,$ and n are normalized into the 0 - 255 range separately using Eq. (5). Secondly, the normalized signals $1, 2, \ldots,$ and n (left panel in Figure 2) are separately split into the segments with same length. Then, segments $1, 2, \ldots,$ and n (middle panel in Figure 2) are output sequentially as the pixels of the IFMs $1, 2, \ldots,$ and n (middle panel in Figure 2) by sequence. Finally, the IFMs $1, 2, \ldots,$ and n make up of IFMs (right panel in Figure 2). The signals $1, 2, \ldots,$ and n refer to different REB measurement signals. In this study, there are two kinds of REB measurement signals on the Paderborn university bearing dataset. The signals 1 and 2 are the current signals of the phase 1 and 2, and signal 3 is the vibration signal.

$$\mathrm{y} = \mathrm{round}\left\{\frac{x - x_{min}}{x_{max} - x_{min}} \times 255\right\} \tag{5}$$

where x is the raw signal; x_{min} and x_{max} represent the minimum and maximum value of the raw signals; round $(\cdot)$ denotes the rounding function.

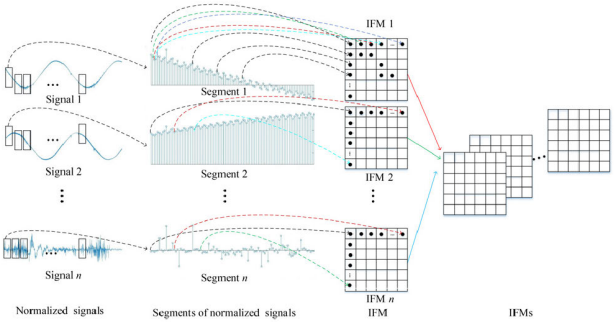

FIGURE 2. Converting the raw signals to IFMs.

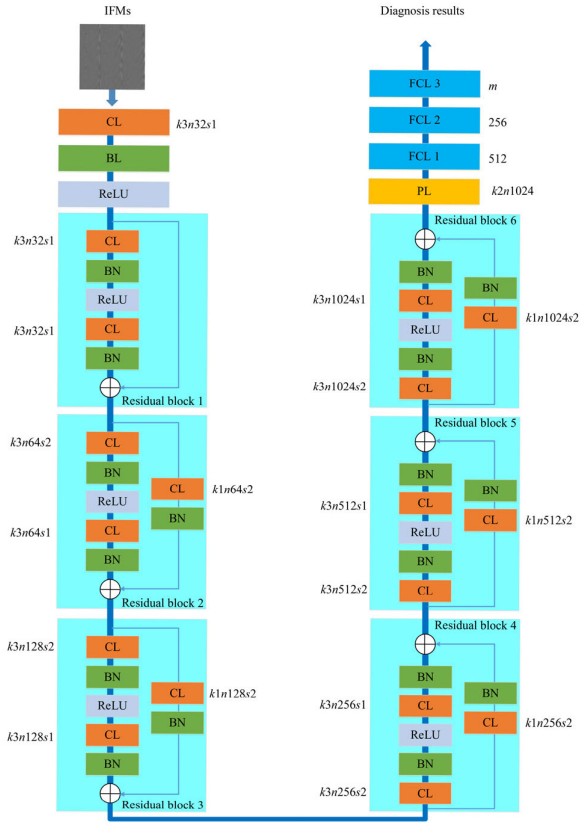

FIGURE 3. The architecture of proposed model. *k*: kernel size; *n*: number of filters; *s*: stride size; ReLU: rectified linear unit; BN: batch normalization layer.

B. PROPOSED FAULT DIAGNOSIS MODEL

In light of its strong generation ability and pattern recognition, a ResNet is used as the fault classifier after IFMs to identify the REB faults. The proposed ResNet consists of a CL, six residual blocks, a PL, and three FCLs, as shown in Figure 3.

The proposed model is evaluated on the Paderborn university bearing dataset. Two phases current signals and vibration signal are collected as measurement signals on the Paderborn university bearing dataset. Here, n is set to 3, as shown in Figure 2. The signal 1 and 2 denote current signals of phase 1 and 2. The signal 3 is vibration signal. Firstly, the raw

signals 1, 2, and 3 are converted to IFMs 1, 2, and 3 using the proposed signal-to-IFMs, respectively. The IFMs 1, 2, and 3 have the same mapping size of $p \times p$. Note that p is set to 64 in this study. Three IFMs with size 64×64 are obtained. Then, the obtained IFMs are fed to a CL ($k3n32s1$), BN, and ReLU. $k3n32s1$ means the CL consisting of 32 filters with a kernel size of 3, and a stride of 1. The 32 feature mappings with size 64×64 (denoted as $32 \times 64 \times 64$ feature mappings) are outputted. Afterwards, the $32 \times 64 \times 64$ feature mappings are fed to residual block 1. Residual block 1 consists of two CLs with $k3n64s1$. An identical connection is connected to the input and output of residual block 1. Residual block 1 output the $32 \times 64 \times 64$ feature mappings. Next, the $32 \times 64 \times 64$ feature mappings are fed to residual block 2. Residual block 2 contains two CL with $k3n64s2$ and $k3n64s1$, respectively. In addition, a CL with $k1n64s1$ is added to the residual connection so that the input of residual block 2 can be added to its output. Note that the number of zero padding in CLs with kernel size 3 is one. The CLs with a kernel size of 1 has no zero padding. Residual block 2 outputs the $64 \times 32 \times 32$ feature mappings. Afterwards, the $64 \times 32 \times 32$ feature mappings pass through residual block 3, 4, 5, and 6, respectively. The $1024 \times 2 \times 2$ feature mappings are obtained after residual block 6. The $1024 \times 2 \times 2$ feature mappings pass through a PL ($k2n1024$) and the $1024 \times 1 \times 1$ feature mappings is obtained. Finally, the $1024 \times 1 \times 1$ feature mappings pass through three FCLs with channels of 512, 256, m. Note that m is the number of the fault labels. The cross-entropy function L_c [34] is used to calculate the classification loss. Finally, the proposed model outputs the fault diagnosis results.

C. FAULT DIAGNOSIS PROCEDURE

The flowchart of REB fault diagnosis process is presented in Figure 4. The specific diagnosis procedure is as shown as follows:

Step 1: Collect the REB raw signals.

Step 2: Convert the raw signals to IFMs.

Step 3: Prepare the training, validation, and testing samples. The training samples are used to train the fault diagnosis model. The validation samples are applied to select the optimal hyper-parameters. The performance of the trained model is evaluated on the testing samples.

Step 4: Built the fault diagnosis model and initialize weights and bias.

Step 5: Train the fault diagnosis using the training samples and calculate the cross-entropy loss.

Step 6: Evaluate the trained model using the validation samples. If the condition is not meted, the fault diagnosis model will continue to update the model parameters, weights, and bias until an optimal trained model that meets the condition is generated.

Step 7: Conduct fault diagnosis on testing samples using the optimal fault diagnosis model.

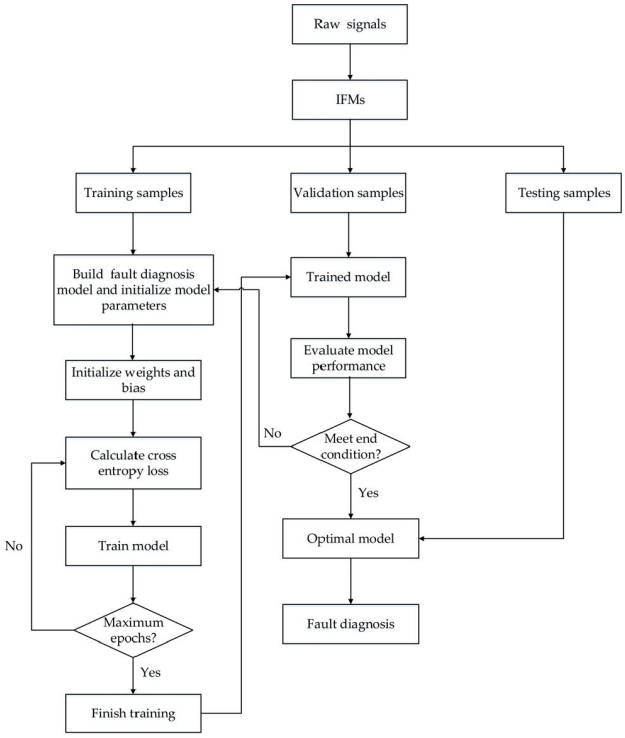

FIGURE 4. Flowchart of the REB fault diagnosis process.

IV. EXPERIMENTAL INVESTIGATION

A. DATA SOURCE DESCRIPTION

The Paderborn university bearing dataset was proposed in 2016 by Christian Lessmeier *et al.* in Paderborn university [35]. There is a wide distribution of bearing damages in this dataset. The data from artificial and real damages are collected. Not only vibration signal but also current signals are collected. The Paderborn university bearing dataset contains a massive amount of data and various fault categories. This subsection gives a brief introduction. A detailed description can be found in [35].

The test rig consists of an electric motor, a torque-measurement shaft, a rolling bearing test module, a flywheel, and a load motor, as shown in Figure 5. There are three categories of REBs used in the test rig. Some are healthy, and others sustain artificial and real damages. Each REB was run 20 times under each different load, as shown in Table 2. In each time, the current and vibration signals were collected lasting approximately 4 s at a rate of 64kHz and 256000 points are collected. There are artificial inner ring damage (AIRD), artificial outer ring damage (AORD), real inner ring damage (RIRD), and real outer ring damage (RORD) on the Paderborn university bearing dataset. Table 3 summarizes the different fault categories and their datasets.

B. DATA PROCESSING

The size of the signal segment should be carefully chosen. A too-long segment will increase the complexity of the

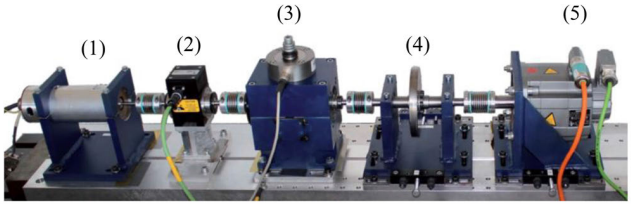

FIGURE 5. Experiment setup of the test rig. (1) Electric motor; (2) torque-measurement shaft; (3) rolling bearing test module; (4) flywheel; (5) load motor.

TABLE 2. Operating conditions of REBs.

load	Rotational speed (rpm)	Load torch (Nm)	Radial force (N)
1	900	0.7	1000
2	1500	0.1	1000
3	1500	0.7	400
4	1500	0.7	1000

TABLE 3. Categorization of the Paderborn university bearing dataset.

Healthy	AIRD	AORD	RIRD	RORD
K001	KI01	KA01	KI04	KA04
K002	KI03	KA03	KI14	KA15
K003	KI05	KA05	KI16	KA16
K004	KI07	KA06	KI17	KA22
K005	KI08	KA07	KI18	KA30
K006		KA08	KI21	
		KA09		

classifier model, while a too-short segment will fail to capture localized feature of REB signals. Besides, the size should be a square number for converting signal segments into the input feature mappings. In this study, the vibration signal and two-phase current signals are used. We divide each type of signal into 139200 segments. Each segment contains 4096 points. Then each segment is converted into a 64×64 IFM using the proposed signal-to-IFMs method We obtain 139200 groups of IFMs are converted from the raw data. They are randomly split into 60% for the training dataset, 20% for the validation dataset, and 20% for testing dataset. The proportion of healthy and faulty samples in each of these datasets is equal. Figure 6 shows the results of converted IFMs under the condition of rotational speed (900 rpm), load torch (0.7 Nm), and radial force (1000 N).

C. COMPARED METHODS

Various compared methods are used for REB fault diagnosis. They mainly contain conventional ML methods, and DL approaches.

(1) The traditional ML methods. We choose the SVM and multi-layer perceptron (MLP) as the compared methods. The 15 statistical features in the time and frequency domains are selected from vibration signal as the statistic features,

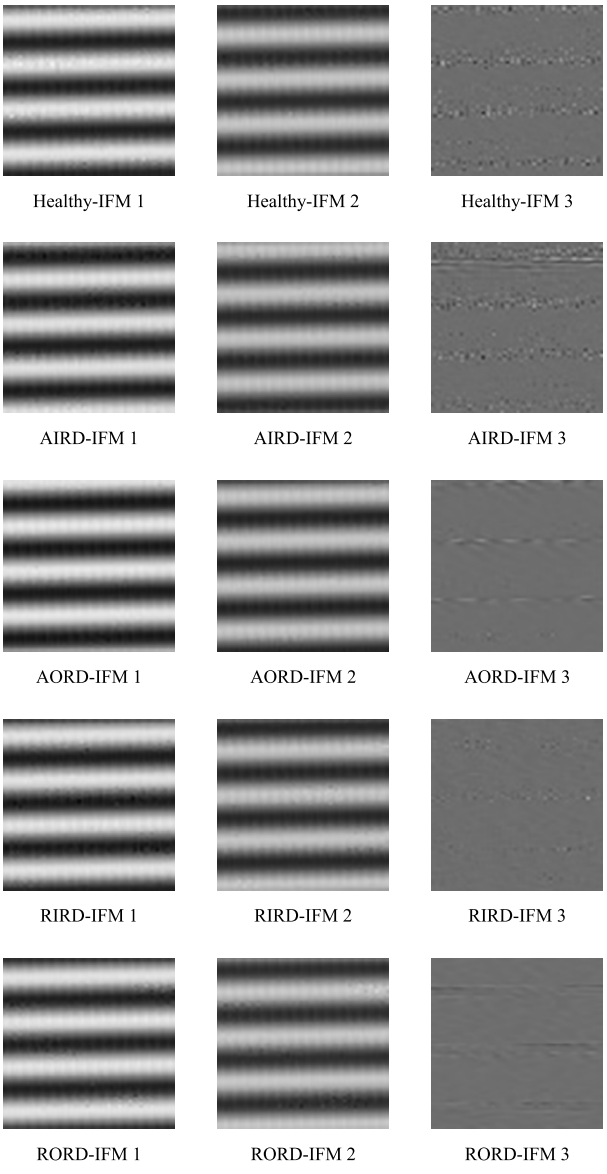

FIGURE 6. Converted IFMs under the condition of rotational speed (900 rpm), load torch (0.7 Nm), and radial force (1000 N).

as shown in Figure 7 [36]. In SVM, the penalty factor is set to 2, and the kernel function parameter is set to 2. In MLP, BP algorithm added momentum factor, and adaptive learning rate is used as learning algorithm. We use a three-layer MLP model. The number of neurons in the input layer is 15; the number of the nodes in the output layer equals to the fault categories; the number of the neurons in the hidden layer is 12. The other parameters of MLP is as follows: hyperbolic tangent transfer function for both hidden and output layer neurons; the learning rate: 0.01; momentum constant:0.9; ratio to increase learning rate:1.05; ratio to decrease learning rate:0.7; and maximum performance increase:1.04; performance goal: 0.01.

(2) The DL methods. DBN [37] and CNN [38] are selected as the compared method. DBN contains two hidden layers (25, 25). The number of the neurons in the input layer is 15

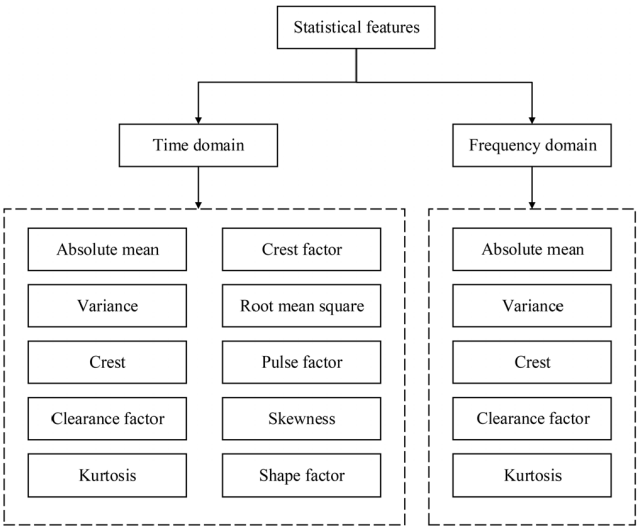

FIGURE 7. Statistical features in the time and frequency domains.

TABLE 4. CNN architecture.

Layer name	Layer configure
1	CL(k3n32s1)
2	PL (k2n32)
3	CL(k3n64s1)
4	PL (k2n64)
5	CL(k3n128s1)
6	PL (k2n128)
7	CL(k3n256s1)
8	PL (k2n256)
9	FCL (4096)
10	FCL (number of fault categories)

(15 statistic features that are same with statistic features of traditional ML methods), and the number of the nodes in the output layer equals to the fault category. In DBN, the learning rate and momentum are set to 0.01 and 0.9. The CNN is used as the fault classifier after IFMs; it contains four CLs, four PLs, and two FCL, as shown in Table 4. The input of CNN is the converted IFMs. The batch size is 32, and the dropout rate is 0.5. One zero is added to the image borders. The SGD optimizer is employed. The momentum factor is 0.9, and the weight decay is 1e-5. The learning rate is same with the learning rate of proposed method.

V. FAULT DIAGNOSIS RESULT AND ANALYSIS

This study focuses to conduct the REB fault diagnosis on the Paderborn university bearing dataset. The proposed model and CNN are written in Python 3.7 using PyTorch. MLP, SVM, and DBN are written in MATLAB 2018a. All methods are run on Windows 10.

A. MODEL PARAMETER SELECTION

The proposed model parameters of the CL, PL, and FCL have been given in Section III. The batch size is 32, and the dropout rate is 0.5. We also use the zero-padding technique to solve the problem of dimension loss. The SGD optimizer is employed. The momentum factor is 0.9, and the weight decay is 1e-5. In this study, we detect the artificial, real, and mixed damages. Learning rate is a hyper-parameter that affects performance of deep learning networks. A high learning rate can lead to divergent oscillations and miss the optimal final set of weights; a low learning rate can lead to very slow reduction in error and need more training epochs [39]. Different kinds of fault dataset, such as, artificial, real, or mixed damage dataset, has different optimal learning rate. The optimal learning rates are selected by experiments. We test five learning rates from 0.001 to 0.02 in each kind of damages. The testing experiments are repeated five times. Then, the optimal learning rate is selected based on the accuracy of the training and validation datasets. Table 5 shows the accuracies with different learning rates. For artificial damages, the proposed model has the optimal training and validation accuracy (99.76%, 99.63%) when the learning rate is 0.01. For real damages, 99.99% and 99.68% are the best training and validation accuracy when the learning rate is 0.001. For mixed damages, the best accuracies are achieved when the learning rate is 0.001.

TABLE 5. Model learning rate of the proposed model.

Case	Learning rate	Average training accuracy	Average validation accuracy
Artificial damages	0.001	98.75	98.72
	0.005	98.83	98.72
	0.01	**99.76**	**99.63**
	0.015	97.28	97.21
	0.02	96.30	96.15
Real damage	**0.001**	**99.99**	**99.68**
	0.005	99.81	99.25
	0.01	99.96	99.45
	0.015	99.42	98.78
	0.02	99.77	99.17
Mixed damages	**0.001**	**99.96**	**99.84**
	0.005	97.91	97.69
	0.01	98.51	98.29
	0.015	96.66	96.47
	0.02	95.54	95.32

B. REB FAULT DIAGNOSIS FOR ARTIFICIAL DAMAGES

In this subsection, we detect all the artificial damages in the REBs. We evaluate the proposed method on healthy dataset (K001, K002, K003, K004, K005, K006), AIRD dataset (KI01, KI03, KI05, KI07, KI08), and AORD dataset (KA01, KA03, KA05, KA06, KA07, KA08, KA09). In these datasets, there are two artificial damages (AIRD, AORD), and one healthy condition. The testing experiments are repeated ten times.

Firstly, we calculate the confusion matrices for the first trial. Figure 8 represents the confusion matrix for all methods. We can see that the proposed method achieves the best results. Then, we calculate training and testing accuracies of proposed method for ten trials, as listed in Table 6. Finally, we calculate the average training and testing accuracies for all methods, as listed in Table 7. The average training accuracy with the proposed method is 99.81%; it is higher than

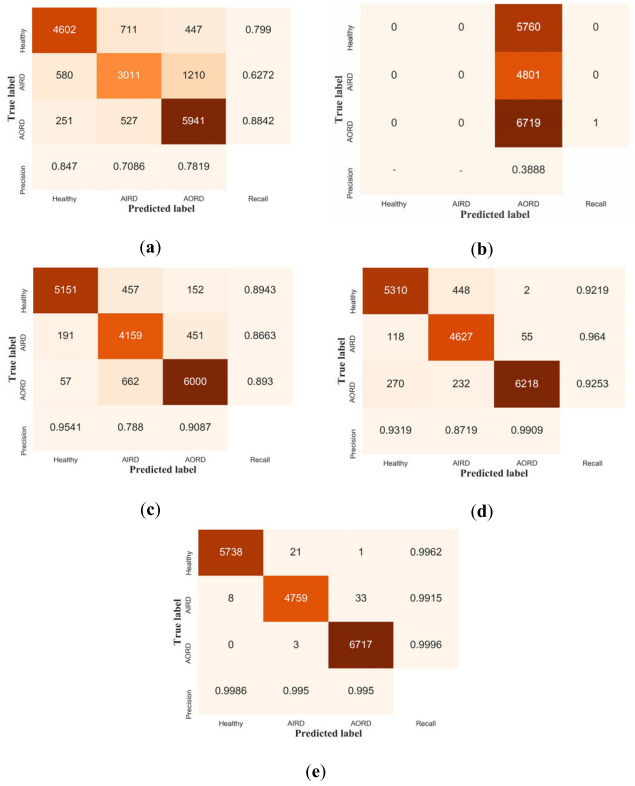

FIGURE 8. The confusion matrices for artificial damages. (a) MLP; (b)SVM; (c) DBN; (d) CNN; (e) proposed method.

TABLE 6. The training and testing accuracies for artificial damages using the proposed method.

Run No.	Training accuracy (%)	Testing accuracy (%)
1	99.71	99.62
2	99.67	99.54
3	99.96	99.92
4	99.57	99.38
5	99.83	99.76
6	99.88	99.76
7	99.81	99.66
8	99.92	99.8
9	98.86	98.76
10	99.92	99.79
Mean	99.81	99.7
Std	0.13	0.15

MLP, SVM, DBN, and CNN, which are 76.04%, 38.89%, 87.81%, and 95.89%, respectively. The average testing accuracy with the proposed method is 99.7%; it is higher than other methods, which are 75.72%, 38.88%, 87.61%, and 94.58%, respectively. We can see that the proposed method achieves the best results compared with other methods.

C. REB FAULT DIAGNOSIS FOR REAL DAMAGES

In this subsection, the proposed method is evaluated on all the real damages datasets. There are two real damages (RIRD, RORD) and one healthy condition in this case. Healthy dataset (K001, K002, K003, K004, K005, K006), RIRD dataset (KI04, KI14, KI16, KI17, KI18, KI21), RORD dataset

TABLE 7. The average accuracies for artificial damages.

Methods	Average accuracy (%)	
	Training	Testing
MLP	76.04	75.72
SVM	38.89	38.88
DBN	87.81	87.61
CNN	95.89	94.58
Proposed	**99.81**	**99.7**

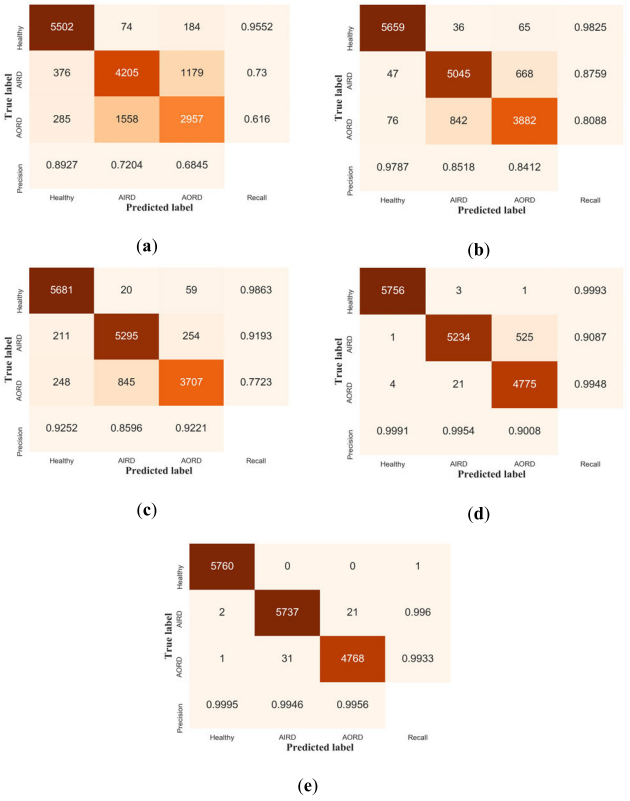

FIGURE 9. The confusion matrices for real damages. (a) MLP; (b) SVM; (c) DBN; (d) CNN; (e) proposed method.

(KA04, KA15, KA16, KA22, KA30) are used to conduct the REB fault diagnosis.

Firstly, the confusion matrices for the first trial is calculated, as shown in Figure 9. The proposed method outperforms the other methods. Ten trials of the proposed method are listed in Table 8. Table 9 gives the average training and testing accuracies for all methods. The average training and testing accuracies are 100% and 99.7% using the proposed model, which is higher than 80.43% and 80.01%, 89.61% and 89.38%, 90.53% and 90.52%, and 98.24% and 98.06% using MLP, SVM, DBN, and CNN, respectively. We can see that the proposed model achieves higher precision, recall, and average accuracies than the other models.

D. REB FAULT DIAGNOSIS FOR MIXED DAMAGES

In subsection C and D of this section, we evaluate the proposed method on the mixed damages. We further evaluate the proposed model on all REB damages. Here, we identify five

TABLE 8. The training and testing accuracies for real damages using the proposed method.

Run No.	Training accuracy (%)	Testing accuracy (%)
1	99.99	99.66
2	100	99.59
3	100	99.97
4	100	99.82
5	100	99.66
6	100	99.63
7	100	99.75
8	100	99.68
9	100	98.53
10	100	99.69
Mean	100	99.7
Std	0.003	0.12

TABLE 9. The average accuracies for real damages.

Methods	Average accuracy (%)	
	Training	Testing
MLP	80.43	80.01
SVM	89.61	89.38
DBN	90.53	90.52
CNN	98.24	98.06
Proposed	**100**	**99.7**

TABLE 10. The training and testing accuracies for mixed damages using the proposed method.

Run No.	Training accuracy (%)	Testing accuracy (%)
1	99.96	99.84
2	99.93	99.86
3	99.98	99.88
4	99.91	99.88
5	99.98	99.85
6	99.94	99.8
7	99.76	99.67
8	99.95	99.77
9	98.99	98.89
10	99.8	99.63
Mean	99.92	99.81
Std	0.08	0.09

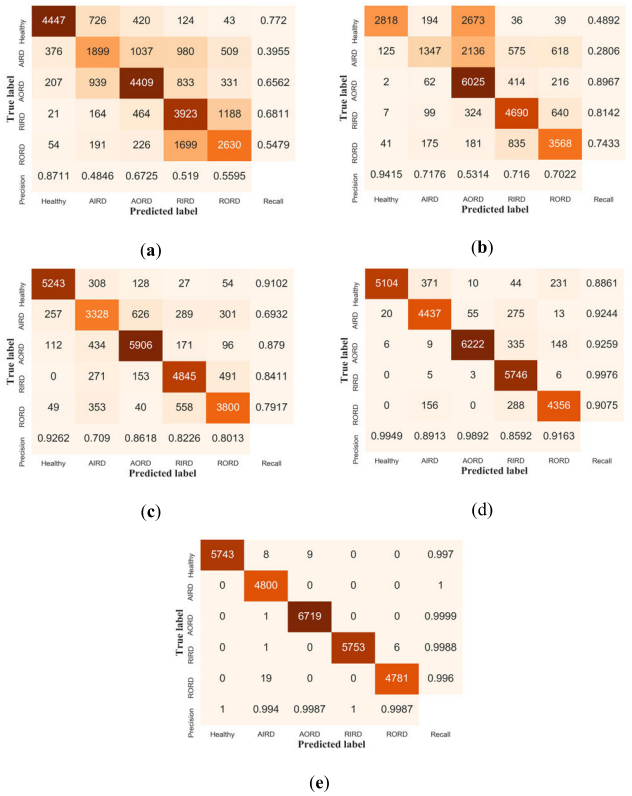

FIGURE 10. The confusion matrices for mixed damages with different methods. (a) MLP; (b)SVM; (c) DBN; (d) CNN; (e) proposed method.

TABLE 11. The average accuracies for mixed damages.

Methods	Average accuracy (%)	
	Training	Testing
MLP	61.87	61.43
SVM	66.59	66.26
DBN	82.78	82.26
CNN	97.15	96.67
Proposed	**99.92**	**99.81**

REB working conditions (healthy, AIRD, AORD, RIRD, and RORD). The testing experiments are repeated for ten times.

Figure 10 shows the confusion matrices of testing samples with different method for the first trial. It is found that the recalls and precisions of the proposed model are higher than other methods. Table 10 gives training, testing, average accuracies, and standard deviation (Std) of ten trails using the proposed method. Table 11 presents the average training and testing accuracies for all methods. The average training accuracy with the proposed method is 99.92%; it is higher than MLP, SVM, DBN, and CNN, which are 61.87%, 66.59%, 82.78%, and 97.15%, respectively. The testing accuracy with the proposed method is 99.81%; it is higher than other methods, which are 61.43%, 66.26%, 82.26%, and 96.67%, respectively. The results show that the proposed approach outperform other approaches.

Generally, the growing training sets can improve accuracy of deep learning methods under the same classification categories [40]. There are two reasons why the proposed method (deep ResNet) outperforms on mixed damage case than artificial and real damage cases. The main reason is that the powerful pattern recognition and generalization ability of deep ResNet. The deep ResNet can prevent gradient vanishing and overfitting when the training samples and fault categories growing; it significantly improves the classification accuracy [33], [41]–[43]. Another reason is that the mixed damages consist of artificial and real damages. The fault features extracted from artificial (or real) damage samples can improve the fault diagnosis performance on real (or artificial) damage samples. That is say, the generation ability of fault classifier using mixed damage samples are improved to some extent. These are the reasons why deep ResNet obtain better training and testing accuracy on mixed damage as compared to artificial and real damages. Other traditional deep learning methods, such as CNN, may has different results. CNN has lower pattern recognition and generalization ability compared to ResNet; it suffer gradient vanishing problem due to lacking of residual connections [33]. CNN may result in low accuracy

or overfitting when growing the training set and fault categories, such as using mixed fault samples. Although, the fault features extracted from artificial (or real) damage samples can improve the fault diagnosis performance of CNN on real (or artificial) damage samples to some extent, the total generation ability are reduced when conduct fault diagnosis on mixed damages. For traditional machine learning methods, such as MLP and SVM, their performance can significantly reduce compared with DL methods, especially with the increase of training dataset [30]. Hence, the traditional learning methods obtain the lower accuracy than the DL methods on REB fault diagnosis with the complicated dataset. All above analysis is assuming that the machine and deep learning have the optimal hyperparameters. Different results may be obtained if the sub-optimal hyperparameters are selected on machine and deep learning methods.

VI. CONCLUSION

In this study, we propose an input feature mappings-based deep residual network to conduct detailed and comprehensive fault diagnosis on REB with complicated bearing. Firstly, we propose a new data preprocessing method to automatically extract features from raw signals without predefined parameters. Secondly, we use a deep ResNet as the fault classifier to learn the discriminative features from IFMs. Finally, we evaluate the proposed method on the Paderborn university bearing dataset. The proposed model outperforms other models.

However, the proposed method has some limitations. Firstly, the proposed method is applied in supervised learning, which meaning humans must label the faults. If the fault type is unknown, the model may be misclassification. Secondly, the hyperparameters selection is still a challenge for the proposed model. Our future works can be conducted as follows. Firstly, we can introduce the unsupervised learning method to the proposed model to diagnose unknown faults. Secondly, other hyperparameters selection methods should be proposed to further improved fault accuracy.

REFERENCES

[1] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[2] A. Nabhan, N. M. Ghazaly, A. Samy, and M. O. Mousa, "Bearing fault detection techniques-a review," *Turkish J. Eng. Env. Sci.*, vol. 3, no. 2, pp. 1–19, 2015.

[3] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106587.

[4] Y. Wei, Y. Q. Li, M. Q. Xu, and W. H. Huang, "A review of early fault diagnosis approaches and their applications in rotating machinery," *Entropy*, vol. 21, no. 4, p. 409, Apr. 2019.

[5] L. F. de Almeida, J. W. Bizarria, F. C. Bizarria, and M. H. Mathias, "Condition-based monitoring system for rolling element bearing using a generic multi-layer perceptron," *J. Vibrat. Control*, vol. 21, no. 16, pp. 3456–3464, Dec. 2015.

[6] Y. Tian, J. Ma, C. Lu, and Z. Wang, "Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine," *Mechanism Mach. Theory*, vol. 90, pp. 175–186, Aug. 2015.

[7] X. Liu, H. Huang, and J. Xiang, "A personalized diagnosis method to detect faults in a bearing based on acceleration sensors and an FEM simulation driving support vector machine," *Sensors*, vol. 20, no. 2, p. 420, Jan. 2020.

[8] P. Baraldi, F. Cannarile, F. Di Maio, and E. Zio, "Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions," *Eng. Appl. Artif. Intell.*, vol. 56, pp. 1–13, Nov. 2016.

[9] S. G. Kumbhar and P. E. Sudhagar, "An integrated approach of adaptive neuro-fuzzy inference system and dimension theory for diagnosis of rolling element bearing," *Measurement*, vol. 166, Dec. 2020, Art. no. 108266.

[10] Y. Guo, J. Na, B. Li, and R.-F. Fung, "Envelope extraction based dimension reduction for independent component analysis in fault diagnosis of rolling element bearing," *J. Sound Vibrat.*, vol. 333, no. 13, pp. 2983–2994, Jun. 2014.

[11] J. Xiang, Y. Zhong, and H. Gao, "Rolling element bearing fault detection using PPCA and spectral kurtosis," *Measurement*, vol. 75, pp. 180–191, Nov. 2015.

[12] Z. Huo, Y. Zhang, L. Shu, and M. Gallimore, "A new bearing fault diagnosis method based on fine-to-coarse multiscale permutation entropy, Laplacian score and SVM," *IEEE Access*, vol. 7, pp. 17050–17066, 2019.

[13] M. Van, D. T. Hoang, and H. J. Kang, "Bearing fault diagnosis using a particle swarm optimization-least squares wavelet support vector machine classifier," *Sensors*, vol. 20, no. 12, p. 3422, Jun. 2020.

[14] W. Zhao, T. Shi, and L. Wang, "Fault diagnosis and prognosis of bearing based on hidden Markov model with multi-features," *Appl. Math. Nonlinear Sci.*, vol. 5, no. 1, pp. 71–84, Mar. 2020.

[15] J. Li, X. Yao, X. Wang, Q. Yu, and Y. Zhang, "Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis," *Measurement*, vol. 153, Mar. 2020, Art. no. 107419.

[16] D.-T. Hoang and H.-J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, Mar. 2019.

[17] H. Shao, H. Jiang, X. Li, and T. Liang, "Rolling bearing fault detection using continuous deep belief network with locally linear embedding," *Comput. Ind.*, vol. 96, pp. 27–39, Apr. 2018.

[18] S. Wang, J. Xiang, Y. Zhong, and Y. Zhou, "Convolutional neural network-based hidden Markov models for rolling element bearing fault identification," *Knowl.-Based Syst.*, vol. 144, pp. 65–76, Mar. 2018.

[19] H. Liu, J. Zhou, Y. Zheng, W. Jiang, and Y. Zhang, "Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders," *ISA Trans.*, vol. 77, pp. 167–178, Jun. 2018.

[20] F. Xu, W. T. P. Tse, and Y. L. Tse, "Roller bearing fault diagnosis using stacked denoising autoencoder in deep learning and Gath–Geva clustering algorithm without principal component analysis and data label," *Appl. Soft Comput.*, vol. 73, pp. 898–913, Dec. 2018.

[21] Z. Zhu, G. Peng, Y. Chen, and H. Gao, "A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis," *Neurocomputing*, vol. 323, pp. 62–75, Jan. 2019.

[22] Y. Han, N. Ding, Z. Geng, Z. Wang, and C. Chu, "An optimized long short-term memory network based fault diagnosis model for chemical processes," *J. Process Control*, vol. 92, pp. 161–168, Aug. 2020.

[23] S. Wang and J. Xiang, "A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps," *Soft Comput.*, vol. 24, no. 4, pp. 2983–2997, Feb. 2020.

[24] H. Liu, D. Yao, J. Yang, and X. Li, "Lightweight convolutional neural network and its application in rolling bearing fault diagnosis under variable working conditions," *Sensors*, vol. 19, no. 22, p. 4827, Nov. 2019.

[25] Y. Wang, J. Yan, Q. Sun, Q. Jiang, and Y. Zhou, "Bearing intelligent fault diagnosis in the industrial Internet of Things context: A lightweight convolutional neural network," *IEEE Access*, vol. 8, pp. 87329–87340, 2020.

[26] S. Wu, X.-Y. Jing, Q. Zhang, F. Wu, H. Zhao, and Y. Dong, "Prediction consistency guided convolutional neural networks for cross-domain bearing fault diagnosis," *IEEE Access*, vol. 8, pp. 120089–120103, 2020.

[27] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017.

[28] K. Zhao, H. Jiang, X. Li, and R. Wang, "An optimal deep sparse autoencoder with gated recurrent unit for rolling bearing fault diagnosis," *Meas. Sci. Technol.*, vol. 31, no. 1, Jan. 2020, Art. no. 015005.

[29] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. V. de Walle, and S. V. Hoecke, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016.

[30] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Machine learning and deep learning algorithms for bearing fault diagnostics-a comprehensive review," 2019, *arXiv:1901.08247*. [Online]. Available: http://arxiv.org/abs/1901.08247
[31] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 6111–6124, May 2020.
[32] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, no. 1, pp. 106–154, Jan. 1962.
[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[34] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," in *Proc. Workshop Deep Learn. Represent. Learn.*, 2015, pp. 1–9.
[35] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, "Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification," in *Proc. Eur. Conf. Prognostics Health Manage. Soc.*, Bilbao, Spain, 2016, pp. 1–18.
[36] M. Xia, T. Li, L. Xu, L. Liu, and C. W. de Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 101–110, Feb. 2018.
[37] H. Shao, H. Jiang, X. Zhang, and M. Niu, "Rolling bearing fault diagnosis using an optimization deep belief network," *Meas. Sci. Technol.*, vol. 26, no. 11, Nov. 2015, Art. no. 115002.
[38] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.
[39] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," 2012, *arXiv:1206.5533*. [Online]. Available: http://arxiv.org/abs/1206.5533
[40] J. Hestness, S. Narang, N. Ardalani, G. Diamos, H. Jun, H. Kianinejad, M. M. A. Patwary, Y. Yang, and Y. Zhou, "Deep learning scaling is predictable, empirically," 2017, *arXiv:1712.00409*. [Online]. Available: http://arxiv.org/abs/1712.00409
[41] F. He, T. Liu, and D. Tao, "Why ResNet works? Residuals generalize," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 5, 2020, doi: 10.1109/TNNLS.2020.2966319.
[42] S. Wu, S.-H. Zhong, and Y. Liu, "Steganalysis via deep residual network," in *Proc. IEEE 22nd Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Wuhan, China, Dec. 2016, pp. 1233–1236.
[43] S. Ozcan and A. F. Mustacoglu, "Transfer learning effects on image steganalysis with pre-trained deep residual neural network model," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Seattle, WA, USA, Dec. 2018, pp. 36–45.

LIANGSHENG HOU received the M.Sc. degree from Dalian Maritime University, Dalian, China, in 2017, where he is currently pursuing the D.Sc. degree with the College of Marine Engineering. His research interests include fault diagnosis, intelligent algorithms, machine learning, and deep learning.

RUIZHENG JIANG received the B.Sc. and M.Sc. degrees in marine engineering from Dalian Maritime University, Dalian, China, in 2007 and 2009, respectively, where he is currently pursuing the Ph.D. degree in marine engineering. He is currently an Associate Professor with the College of Marine Engineering, Dalian Maritime University. His current research interests include the simulation of marine power systems, optimization methods, artificial intelligence, and smart ships.

YANGHUI TAN received the B.Sc. and M.Sc. degrees in marine engineering from Dalian Maritime University, Dalian, China, in 2013 and 2016, respectively, where he is currently pursuing the D.Sc. degree in marine engineering. His research interests include the applications of machine learning on marine ships, smart ships, marine engine room simulator, and virtual reality.

JUNDONG ZHANG received the B.Sc., M.Sc., and D.Sc. degrees in marine engineering from Dalian Maritime University, Dalian, China, in 1989, 1992, and 1998, respectively. He is currently a Full Professor with Dalian Maritime University. His research interests include marine engineering automation and control, integrated supervision, the applications of computers and networks, marine engineering education, and electrical system design.