

Received January 31, 2021, accepted February 16, 2021, date of publication February 23, 2021, date of current version March 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3061530

Deep Convolutional Neural Network Using Transfer Learning for Fault Diagnosis

DONG ZHANG^{1,2} AND TAOTAO ZHOU^{1,2} 

¹MOE Key Laboratory of Marine Intelligent Equipment and System, Shanghai Jiao Tong University, Shanghai 200240, China

²China Ship Development and Design Center, Wuhan 430074, China

Corresponding author: Taotao Zhou (kttkt@163.com)


This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61703385, and in part by the National Defense Program under Grant 61400020401.

ABSTRACT Fault diagnosis is critical in industrial systems since early detection of problems can not only save valuable time but also reduce maintenance costs. The feature extraction process of traditional fault diagnosis is time-consuming and laborious work. Recently, with the rapid development of the deep learning (DL) method, it has shown its superiority with an end-to-end process and has been applied to classification and other fields. To a certain extent, it solves the disadvantages of manual feature extraction in the traditional fault diagnosis method. However, the available training data is often limited, and it will degrade the performance of DL methods. A new DL method that combines deep convolutional neural network (DCNN) and transfers learning (TL) for fault diagnosis is proposed in this paper to handle different fault types. A signal processing method that converts one-dimensional time-series signals into grayscale images is firstly applied, and it can eliminate the effect of handcrafted features. Secondly, an optimal DCNN is designed and trained with the ImageNet datasets, which can extract the high-level features of massive images. Finally, TL is further developed to apply the knowledge learned in the source data distribution to the target data distribution, which greatly reduces the dependence on training data and improves the generalization performance of DCNN. Three well-known datasets, including the bearing vibration dataset from the Case Western Reserve University (CWRU), self-priming centrifugal pump dataset (SPCP), and bearing force dataset from the University of Paderborn, are utilized to the performance of the proposed method. Some popular classification methods are also added to the comparison. Results show that the proposed method can precisely identify different fault types and have the highest classification accuracy among other methods.

INDEX TERMS Fault diagnosis, deep convolutional neural network, transfer learning, image classification, signal processing.

I. INTRODUCTION

Rotating machinery is a key component in industrial systems which is widely used in mechanical equipment. It usually works in a harsh environment for a long time. If a fault occurs in rotating machinery components such as bearing and gear, it is very likely to affect the normal operation of the entire mechanical equipment, or even threaten the safety of workers. Thus, to ensure the operational reliability of rotating machinery and avoid disastrous accidents in industrial systems, it is necessary to accurately detect defects and failures of the rotating machinery components as early as possible.

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino .

Due to the need for long-term continuous monitoring and the existence of massive data, machine learning (ML)-based fault diagnosis methods have attracted more and more attention. Because ML-based diagnosis methods can process and utilize useful information from adequate historical data, they are considered a very promising and powerful tool [1]–[3].

Fault diagnosis can be defined as a kind of pattern recognition in essence. It consists of three steps to detect the fault types: vibration signal collection, feature extraction, and fault classification. The latter two are the most critical steps which affect the performance of fault diagnosis methods. The existing fault diagnosis methods can be generally divided into two types: the traditional data-driven-based fault diagnosis method and the deep learning (DL)-based fault diagnosis

method. The traditional data-driven-based fault diagnosis methods need to extract and select features manually. In contrast, DL-based fault diagnosis methods can fulfill the whole process of feature extraction, feature selection, and classification automatically.

Vibration signal contains a wealth of information, hence different types of faults can be diagnosed accurately through the vibration signal. The traditional data-driven-based fault diagnosis method consists of four steps: collect the vibration signal, extract the feature from the vibration signal, filter out the feature which can reflect the fault state, and input the feature into the classification model to get the classification results. Liang *et al.* [4] input the characteristic parameters into a backpropagation (BP) neural network for training and testing, combined with the water-lubricated stern bearing test to verify that the method has high accuracy in identifying different types of faults. Liu *et al.* [5] presented an idea, which combines impact time-frequency dictionary, short-term matching, and SVM. The proposed method can extract good features under an extremely low signal-to-noise ratio, monitor the running status of bearing in real-time, and effectively detect early breakdown of bearing. Zheng *et al.* [6] came up with a composite multiscale fuzzy entropy method to extract the non-linear features hidden in the vibration signal and used the ensemble SVM to classify multiple types of faults. The experiment results showed that different types of bearing faults were effectively identified. To improve the classification accuracy of SVM, Hui *et al.* [7] suggested an SVM model based on Dempster-Shafer (DS) evidence theory. Compared with the original SVM, this model is more accurate and effective in dealing with common multiple fault diagnosis and classification problems. Yan and Jia [8] extracted features from the time domain, frequency domain, and time-frequency domain, and used the Laplace score feature selection algorithm to select useful sensitive feature information. SVM based on radial basis function and particle swarm optimization algorithm was used to fulfill the classification task. Two bearing data sets were applied to test the efficacy and superiority of the fault diagnosis method for bearing.

Although the traditional data-driven-based fault diagnosis methods have achieved high accuracy, the manual extraction and selection of features not only require signal processing methods but also rely on expert knowledge in the big data era, which consumes a large amount of manpower and time cost. What's more, extracting features manually sometimes may not contain fault information, which will lead to a decline in recognition accuracy and poor generalization ability. Therefore, the DL-based fault diagnose method is urgently needed to solve the existing problems.

DL-based fault diagnose method is popular with its advantages of automatically extracting features and performing recognition and classification. The process of the DL-based fault diagnose method is as follows: collect the vibration signal, and directly input the vibration signal into the DL model

to obtain the classification result. However, many researchers tend to extract feature at first, and then use the DL model to extract high-level features and classify automatically [9]. This method has higher accuracy in identifying different faults. Deng *et al.* [10] presented a method by inputting features of the time domain, frequency domain, and time-frequency domain into a deep Boltzmann machine (DBM) to detect faults of rolling bearing. DBM showed high reliability in rolling bearing fault diagnosis through the data sets of seven fault modes. Tao *et al.* [11] built the fault diagnosis system based on multi-signal fusion and proposed a deep belief network (DBN) based on multi-signal fusion. Firstly, the time domain features were extracted from three vibration sensors, and the features were directly used as the input of the DBN. The accuracy of the multi-sensor fusion method is about 10% higher than that of a single sensor. Janssens *et al.* [12] proposed a CNN model for bearing condition monitoring model, which can automatically learn the data's features. This method is better than the feature selection method based on manual feature extraction, and having an improved classification accuracy of about 6%. Guo *et al.* [13] proposed a novel CNN, which added the adaptive learning rate and momentum component to the traditional CNN. The experiment results showed that it has high accuracy in bearing fault diagnosis and severity judgment. Zhang *et al.* [14] presented a new CNN with training interference (TICNN), which directly took the original vibration signal as the input and used dropout interference training. Azamfar *et al.* [15] proposed a novel two-dimensional convolutional neural network structure, which fuses multiple sensor data and directly uses it for classification. Compared with the classical machine learning algorithm, the method showed the best classification performance in gearbox fault diagnosis. Kolar *et al.* [16] input the original signal of the three-axis accelerometer into a convolution neural network to automatically extract and select features. The method can complete the classification of different rotating machinery states. The experiment results showed that TICNN has high accuracy under noisy environments and changing workloads. However, when there is a distribution difference between the training dataset and the test dataset, the performance of the DL-based fault diagnosis method will degrade.

In this study, a novel fault diagnosis method that combines DCNN and TL is proposed to detect the fault timely and accurately. The proposed method not only reduces the computational cost but also improves classification accuracy. What's more, it also works well in the case of a small dataset. The summaries of the contributions of the present work are as follows:

- 1) A simple and effective method is applied to convert original time-series signals into two-dimensional grayscale image, which can quickly extract the two-dimensional features of the original data;
- 2) Instead of using a fully connected layer, global average pooling (GAP) is applied in DCNN; the optimal

pre-trained DCNN model is applied to obtain the high-level features of images;

- 3) A new fault diagnosis method based on DCNN and TL for the mechanical system is presented. The representative features are fully utilized and the classification accuracy is greatly improved;
- 4) Three experimental cases, including the CWRU dataset, SPCP dataset, and UPB dataset, are used to verify the efficacy of the suggested approach.

The remaining work is organized as follows. The background of DCNN and TL are introduced in section 2. The details of the proposed DCNN-TL model are introduced in Section 3. Section 4 verifies the effectiveness of the proposed approach through experiments, followed by the conclusion and future work in Section 5.

II. RELATED WORKS

A. DEEP CONVOLUTIONAL NEURAL NETWORK

CNN has fewer parameters than fully connected neural networks due to local connections and weight sharing. A standard CNN is mainly composed of three parts: the convolution layer, the pooling layer, and the fully connected layer.

The convolution layer is utilized to extract the features of the input data, and the convolution kernel in the convolution layer completes the function. The convolution kernel can be regarded as a scanner with specified window size. The scanner scans the input data again and again to extract features. The convolution operation needs to flip the convolution kernel. The cross-correlation operation can be called non-flipped convolution operation, which does not need to flip the convolution kernel. Although the convolution layer is named after the convolution operation, a more intuitive cross-correlation operation is often used in the convolution layer. Because convolution is used for feature extraction, and whether the convolution kernel is flipped or not has nothing to do with its ability to extract features. Especially when the convolution kernel is a learnable parameter, convolution and cross-correlation are equivalent. The mathematical model of convolution operation is as follows:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \right)$$

where the operator $*$ represents the convolution operation. M_j is the set of characteristic graphs; l is the l th network; k_{ij}^l is the convolution kernel of l th; b is the network bias; x_j^l is the output of l th; x_i^{l-1} is the input of l th; $f(\cdot)$ is the activation function.

The commonly used activation functions include Sigmoid function, Tanh function, and rectified linear unit (ReLU), where the ReLU function is expressed as:

$$f(x) = \max(0, x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$$

TABLE 1. The top-1 and top-5 accuracy refers to the DCNN's performance on the ImageNet validation dataset.

Model	Top-1 Accuracy(%)	Top-5 Accuracy(%)	Size (MB)	Parameters	Depth
VGG16[18]	0.713	0.901	528	138,357,544	23
VGG19[18]	0.713	0.900	549	143,667,240	26
ResNet50[19]	0.749	0.921	99	25,636,712	-
InceptionV3[20]	0.779	0.937	92	23,851,784	159
InceptionRes-NetV2[21]	0.803	0.953	215	55,873,736	572
MobileNet[22]	0.704	0.895	16	4,253,864	88
MobileNetV2[23]	0.713	0.901	14	3,538,984	88
Xception[24]	0.790	0.945	88	22,910,480	126
DenseNet121[25]	0.750	0.923	33	8,062,504	121
DenseNet169[25]	0.762	0.932	57	14,307,880	169
DenseNet201[25]	0.773	0.936	80	20,242,984	201
NASNetMobile[26]	0.744	0.919	23	5,326,716	-
NASNetLarge[26]	0.825	0.960	343	88,949,818	-

The pooling layer is utilized to select the core feature and reduce the number of features. It not only realizes the compression of the original data but also greatly reduces the parameters involved in the model calculation. Thus it improves the calculation efficiency. The two most common pooling ideas are average pooling and maximum pooling. The calculation method of neurons in the pooling layer can be expressed as:

$$x_j^l = f \left(\beta_j^l \text{down} \left(x_i^{l-1} \right) + b_j^l \right)$$

where $\text{down}(\cdot)$ is the pooling function.

Generally, the input data processed by the pooling layer is the feature map generated after the convolution operation. The fully connected layer is utilized to compress the feature extracted after convolution and pooling operations and complete the classification according to the compressed features.

The Back Propagation (BP) algorithm, a widely used gradient-based parameter learning algorithm [17], is usually used to train CNN. This process is trained by adjusting and optimizing the parameters of CNN. When the training is completed, the pre-trained CNN model is further applied to predict or classify.

Table 1 shows the image size, weights size, top-1 accuracy, top-5 accuracy, parameters, size, and depth of some commonly used DCNN architecture [18]–[26]. All the above architectures can be created using either Theano or

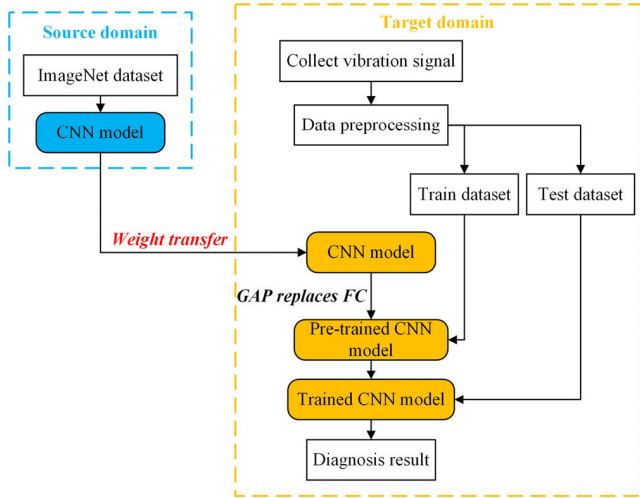


FIGURE 1. Framework of the DCNN-TL method.

TensorFlow except Xception and MobileNet (as they depend on Separable Convolutions and Depth wise Convolutions which are only available in TensorFlow).

B. TRANSFER LEARNING

TL is an essential machine learning method that applies knowledge learned from the source domain to different but related target domain. It can achieve an excellent result and promote the development of areas that are difficult to move forward due to a lack of training data [27].

The domain consists of two parts: feature space χ and edge distribution function $P(X)$. Suppose there are source domain $D_S = \{\chi_S, P(X_S)\}$ and source task $T_S = \{Y_S, f_S(\cdot)\}$, where Y_S is the label space and $f_S(\cdot)$ is the prediction function. Transfer learning can apply the knowledge information $\{(x_{S1}, y_{S1}), \dots, (x_{Sn}, y_{Sn})\}$ in the source domain to the target domain $D_T = \{\chi_T, P(X_T)\}$, making it more accurate to predict the category of output Y_T corresponding to the characteristics χ_T of the target domain.

The main research subjects of TL are about two parts: domain adaption and multi-source domain transfer. Among them, domain adaption is a popular TL method, which performs learning tasks on source data (called source domain) and performs the same task on different but related target data (called target domain). What’s more, Deep transfer learning applies deep neural networks to study how to apply the knowledge of the source domain to the target domain.

III. THE PROPOSED DCNN-TL METHOD

In this part, a new diagnosis method combining DCNN and TL (DCNN-TL) for mechanical faults is suggested, which can effectively solve the problem of less raw data further save time and cost. The proposed DCNN-TL method consists of two parts: preprocessing of signals and fault diagnosis based on suggested DCNN-TL. The framework of the method is shown in Figure 1. Firstly, the source domain DCNN is constructed and trained on the ImageNet dataset. Then, the weights of the source domain DCNN are transferred to the

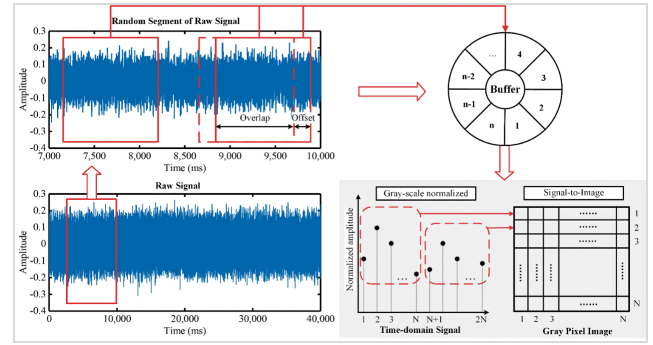


FIGURE 2. Signal-to-image conversion method.

target domain DCNN through TL. The fully connected layer of DCNN is replaced by GAP to construct the pre-trained DCNN. Finally, the collected vibration signals are converted into grayscale images and input into the pre-trained DCNN to obtain the diagnosis results.

A. PREPROCESSING OF VIBRATION SIGNALS

Data preprocessing is vital for traditional data-driven methods since it is difficult to use raw data directly. One of the preprocessing methods is the artificial extraction of features, which is usually time-consuming and needs expert knowledge. In this study, a simple and effective method is utilized to converted one-dimensional raw data into two-dimensional images [28]–[29].

As shown in Figure 2, the steps of the proposed method are as follows:

Step 1: Load and cut raw signal randomly

To obtain $N \times N$ size images, N segment signals of the length N^2 would be obtained randomly from the original signal. A random number is generated according to the uniform distribution or normal distribution, and it is taken as the starting point of time to intercept signals.

Step 2: Gray level normalization

The amplitude of each sample of vibration signal is first normalized ranging from 0 to 255, which is the significant pixel intensity range for a gray image, as shown in the following equation:

$$output = round \left(\frac{(255 - 0) * (input - \min(input))}{\max(input) - \min(input)} \right)$$

where round (\cdot) is the rounding function. After calculating by the above formula, all the values in the sample are normalized to the range of 0 to 255.

Step 3: Construct $N \times N$ matrix and convert it into an $N \times N$ single-channel grayscale map.

The signal has N^2 sample points, that is, the size of the image is $N \times N$ (N value represents the row and column of the image, respectively). The value of N is generally taken according to the capacity of the original signal. The results of Signal-to-Image can be seen in Figure 3. However, the computational complexity of this method is proportional to the value of N . Thus, a smaller value of N is preferred, but

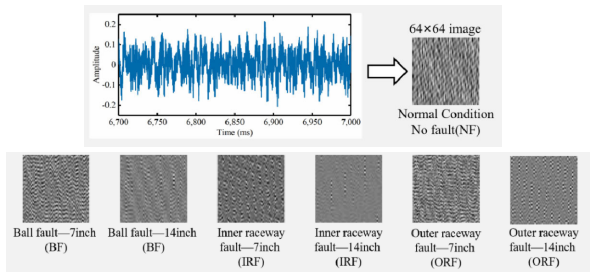


FIGURE 3. Results of signal-to-image conversion.

cannot be too small to prevent loss of the most significant features. The recommended value of N is 32, 64, or 128.

B. DIAGNOSIS OF MECHANICAL FAULTS BASED ON THE PROPOSED DCNN-TL

After preprocessing the vibration signal data, a large number of images with different pixels will be obtained. Then, the neural network is used to identify and classify the images to achieve mechanical fault diagnosis. Aiming at the problem of insufficient training data and time-consuming training of DCNN, this work suggests a new method DCNN-TL, which improves the training speed of the DCNN model and the classification accuracy of fault diagnosis. The shallow features of DCNN are universal for relevant tasks, while the deep features are specific for different tasks. Therefore, the shallow layer of DCNN is considered a universal feature extractor. The detailed process of the DCNN-TL method mentioned above is as follows:

Step 1: Acceleration sensors on different machines are used to collect vibration signals in time series.

Step 2: The preprocessing is applied to convert a one-dimensional time-series signal into a grayscale image.

Step 3: Construct DCNN and use GAP [3] to replace the fully connected layer in DCNN, and then train DCNN.

Step 4: The parameters of source DCNN are transferred to target DCNN through the TL, which can use general features for image classification.

Step 5: The training sets are input into the DCNN model to gain meaningful features in DCNN-TL. Moreover, the gained features and parameters are input into the DCNN-TL model, which can be trained by the generalized inverse operation.

Step 6: The test sets are input into the pre-trained DCNN-TL to gain the classification results.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this part, three cases, containing the CWRU dataset, SPCP dataset, and UPB dataset, are used to verify the effectiveness of the built DCNN-TL model. Three experiments were performed using the DL framework TensorFlow based on Windows 10. The basic information of the DL platform is as follows: the graphics card version is Nvidia RTX 2080 GPU, the CPU is i7-9700K@3.60GHz, the Python version is 3.7.4, and the Keras version is 2.2.4.

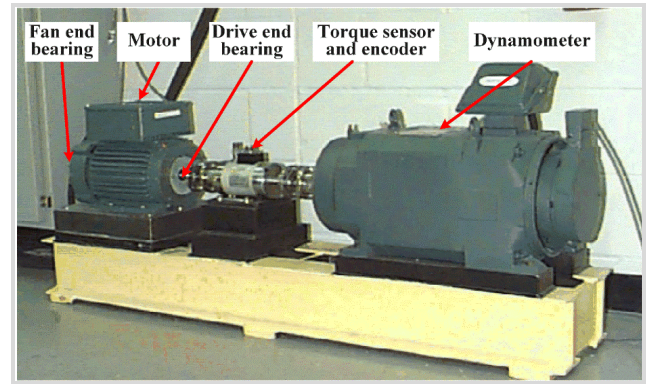


FIGURE 4. CWRU bearing dataset test rig.

A. PREPROCESSING OF THE VIBRATION SIGNALS

1) CWRU DATASET DESCRIPTION

As shown in Figure 4, the left side is a two hp induction motor, the middle is a torque sensor/encoder, and the right side is coupled to a dynamometer. The CWRU bearing dataset is obtained from the test standard.

The general process of collecting data is as follows. Firstly, electrical discharge machining technology (EDM) is used to manufacture faults with a diameter of 0.007 inches to 0.040 inches at three positions of the inner race, rolling elements, and outer race of the motor bearing. Then, the faulty bearing is installed into the motor. The vibration signal of the motor with a load of 0 to 3 hp is finally recorded. The structure of the experimental platform is shown in Figure 3, which is composed of equipment such as a test 2 hp motor, torque sensor/encoder, dynamometer. The main function of the test bearing is to brace motor shaft. The EDM method was used to make single-point failures with diameters of 7 mils, 14 mils, 21 mils, 28 mils, and 40 mils (1 mil = 0.001 inch) to the test bearing. Using SKF bearings for 7 mils, 14 mils, and 21 mils diameter faults, and NTN equivalent bearings for 28 mils and 40 mils diameter faults. The vibration signal was collected by accelerometers mounted on the housing with the magnetic base. The locations of the accelerometers were at 12 o'clock on the drive end and fan end. In some experiments, an accelerometer was mounted to the motor support base plate. The vibration signals under different faults were collected and stored in *.mat format to be read and processed by MATLAB. The sampling frequency of the drive end bearing is 12KHz and 48KHz. The speed and horsepower which were recorded manually were gained by a torque sensor/encoder.

The dataset contains three different faults (roller fault (RF), outer race fault (OF), and inner race fault (IF)), and each failure contains three different levels of damage (7, 14, and 21 milliseconds). Therefore, the dataset contains ten states, namely nine fault states and one normal state (NO).

2) CWRU IMAGE CONVERSION RESULT

Table 2 is the contents of the CWRU-12@6 dataset and the CWRU-48@6 dataset. The CWRU-12@6 dataset includes 10 fault types, each of which has 1000 pictures of 64*64*

TABLE 2. CWRU-12@6 and CWRU-48@6 dataset.

Label	Fault Type	Fault Description	Samples Number
0	IF07	Inter Race Fault 7 mils	1000
1	IF14	Inter Race Fault 14 mils	1000
2	IF21	Inter Race Fault 21 mils	1000
3	OF07	Outer Race Fault 7 mils	1000
4	OF14	Outer Race Fault 14 mils	1000
5	OF21	Outer Race Fault 21 mils	1000
6	RF07	Roller Fault 7 mils	1000
7	RF14	Roller Fault 14 mils	1000
8	RF21	Roller Fault 21 mils	1000
9	NF	No Fault	1000

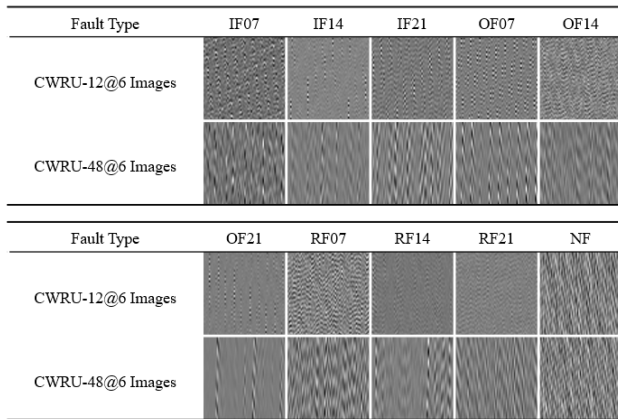


FIGURE 5. Converted images in CWRU (hp = 0).

TABLE 3. Parameter setting of the CWRU-12@6.

Types	Value
Epoch	30
Batch size	256
Optimizer	Adam
Learning rate	0.0003
Test size	0.2

1-pixel size (4 HP states, each HP state has 250 pictures), so the dataset contains 10000 samples in total. The difference between the CWRU-48@6 dataset and the CWRU-12@6 dataset is that the original signal sampling frequency and the values of random sampling points (uniformly distributed sampling) are different.

An example of converting the signals of ten states into images is shown in Figure 5. It can be seen from the converted image that the images of different states are different and can be effectively identified and classified.

3) CWRU RESULTS COMPARED WITH OTHER METHODS

a: CWRU-12@6 RESULTS

The specific settings of the CWRU-12@6 are shown in Table 3.

The performance of different DLNNs is compared through the test accuracy and efficiency. The comparison results are shown in Table 4. The parameter column in the table refers that the number of weights included in each DCNN, while

TABLE 4. Test results of CWRU-12@6.

Model	Accuracy(%)	Time(s)	Size(MB)	Parameters
CNN-2560-768	91.25	38.47	37.7	3,136,010
VGG16	100	161.09	176.9	14,735,178
ResNet50	99.95	201.30	284.4	23,669,642
Xception	99.95	232.18	251.5	20,943,410
DBMobileNetV1	99.85	86.73	39.1	3,239,114
DenseNet121	100	241.84	85.5	7,047,754
MobileNetV2	97.90	113.70	27.6	2,270,794

TABLE 5. Parameter setting of the CWRU-48@6.

Types	Value
Epoch	30
Batch size	256
Optimizer	Adam
Learning rate	0.0003
Test size	0.2

TABLE 6. Test results of CWRU-48@6.

Model	Accuracy(%)	Time(s)	Size(MB)	Parameters
CNN-2560-768	92.90	40.13	37.7	3,136,010
VGG16	98.70	164.56	176.9	14,735,178
ResNet50	94.80	200.87	284.4	23,669,642
Xception	95.65	231.22	251.5	20,943,410
DBMobileNetV1	98.50	85.88	39.1	3,239,114
DenseNet121	98.90	239.11	85.5	7,047,754
MobileNetV2	94.4	121.45	27.6	2,270,794

the size column denotes the required room to save these data. It can be seen from the results that the VGG16 model and the DenseNet121 model achieve the highest test accuracy compared with other models. The VGG16 model requires only 161.09s to train after TL. The running time of the CNN-2560-768 model is 38.47s, which is faster than all the other models. However, this model achieves the lowest classification accuracy. The DBMobileNetV1 model can achieve a very high classification accuracy of 99.85% with a relatively low training time of 86.73s.

b: CWRU-48@6 RESULTS

The specific settings of the CWRU-48@6 are given in Table 5.

The results of CWRU-48@6 are shown in Table 6. It can be seen from the results that the DenseNet121 model has better results compared with other models. The VGG16 model achieves almost the same accuracy of 98.70% but it saves about 31% of the training time compared with the DenseNet121 model. The CNN-2560-768 model still requires less time to train but has the lowest classification accuracy. It is worth mentioning that the parameters of MobileNetV2 are 2,270,794, which is fewer than all the other methods, thus the complexity of this model is the lowest.

The confusion matrix of the DenseNet121 test result is shown in Figure 6. The transverse axis represents the actual state, while the longitudinal axis represents the predicted state. The bar on the right represents the probability that an actual state will be diagnosed with a certain state. The darker

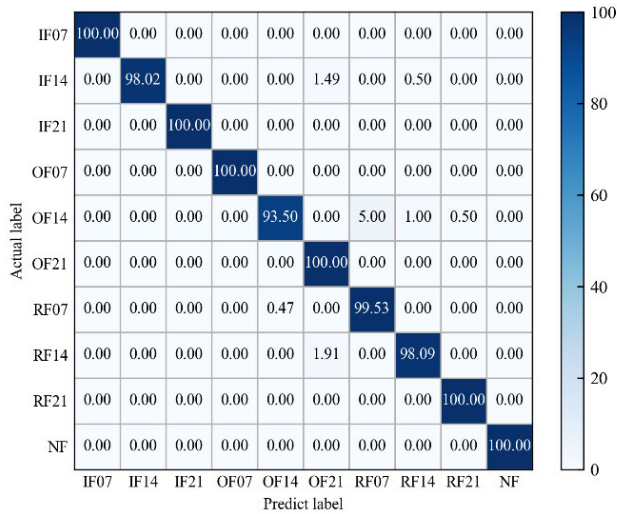


FIGURE 6. Confusion matrix of the result of DenseNet121.

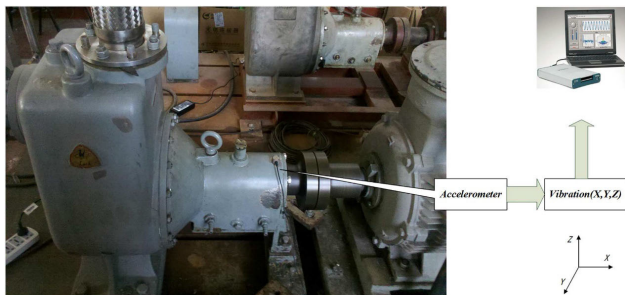


FIGURE 7. Self-priming centrifugal pump data acquisition system.

the color, the greater the probability. The results in Figure 5 shows that IF07, IF21, OF07, OF21, RF21, and NF have 100% accuracy. More importantly, no other conditions are misclassified into them, which means that they are completely separate from the other conditions. OF14 is the worst one which has an accuracy of 93.50%. OF14 receives the highest misclassification. 5.00% out of RF07, 1.00% out of RF14, and RF21 are misclassified to RF0.36.

B. FAULT DIAGNOSIS OF SPCP DATASET

1) SPCP DATASET DESCRIPTION

Figure 7 is the vibration signal acquisition system of the self-priming centrifugal pump (SPCP) [31]. The acceleration sensor is mounted to the top of the motor housing to collect data, and it is fixed on a specific base.

The rotation speed of the experiment is 2,900/min. The sensor is used to collect data at a sampling frequency of 10239 Hz.

The vibration data of five states were collected, including normal state and four different faults (bearing roller wear fault (RF), inner ring wear fault (IF), outer ring wear fault (OF), and impeller wear fault (IW)). The sampling time is set to 2s, while a group of data is collected every 5 seconds.

TABLE 7. Parameter setting of the SPCP.

Types	Value
Epoch	10
Batch size	256
Optimizer	Adam
Learning rate	0.0003
Test size	0.2
Early stopping	2

TABLE 8. Test results of SPCP.

Model	Accuracy(%)	Time(s)	Size(MB)	Parameters
CNN-1536-128	96.5	37.89	11.3	990,085
VGG16	non-convergence	-	-	-
ResNet50	non-convergence	-	-	-
Xception	99.72	229.07	239	20,902,445
DBMobileNetV1	100	63.75	37.1	3,233,989
DenseNet121	100	116.18	84.4	7,042,629
MobileNetV2	non-convergence	-	-	-

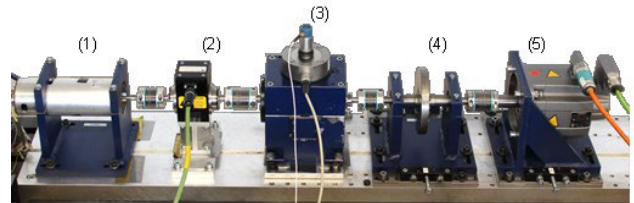


FIGURE 8. UPB bearing dataset test rig.

2) SPCP RESULTS COMPARED WITH OTHER METHODS

The specific settings of the SPCP are given in Table 7.

The results of the SPCP are shown in Table 8. It can be seen from the results that the DBMobileNetV1 model achieves excellent results with a test accuracy of 100%, which is better than all other models. The test accuracy of CNN-1536-128 [30] and Xception are 96.50% and 99.72%, respectively.

C. UPB DATASET FAULT DIAGNOSIS

1) UPB DATASET DESCRIPTION

A test platform was developed at the University of Paderborn in Germany to generate experimental data by using current signals by the motor [32]. The test platform adopts a modular system, which can flexibly utilize various faults in the electromechanical transmission system. By creating faults in mechanical equipment (such as faults in gearboxes or motors), fault data is produced in the platform. The measurement data is generated according to the recorded current signal. Besides, the vibration signal of the test bearing housing is also measured as a reference.

The test bench contains the following parts: an electric motor (1), a torque-measurement shaft (2), a rolling bearing test module (3), a flywheel (4), and a load motor (5), as shown in Figure 8. Experiments on rolling bearing with different faults are carried out and the experimental data are generated. The rolling bearing module can continuously adjust the radial load of the bearing to a fixed value within 10kN before each experiment. An adapter measures the vibration of the internal

TABLE 9. Parameter setting of the UPB.

Types	Value
Epoch	10
Batch size	256
Optimizer	Adam
Learning rate	0.0003
Test size	0.2
Early stopping	2

TABLE 10. Test results of UPB.

Model	Accuracy(%)	Time(s)	Size(MB)	Parameters
CNN-1536-128	95.1	8.39	11.9	985,987
VGG16	non-convergence	-	-	-
ResNet50	non-convergence	-	-	-
Xception	97.1	58.26	250.8	20,886,059
DBMobileNetV1	100	24.47	39.0	3,231,919
DenseNet121	100	74.90	85.4	7,040,579
MobileNetV2	100	33.13	27.5	2,261,827

shell by maintaining the load in the main direction of the bearing.

The experimental data consists of 32 different degrees of damage of the bearing. There are three main types of experimental bearings:

- Undamaged (healthy) bearings (6x);
- Artificially damaged bearings (12x);
- Bearings with real damages caused by accelerated life-time tests (14x);

2) UPB RESULTS COMPARED WITH OTHER METHODS

The specific settings of the UPB are given in Table 9.

The results of the UPB are shown in Table 10. It can be seen from the results that the test accuracy of the DBMobileNetV1, DenseNet121, and MobileNetV2 models are all 100%, which is better than other models. The prediction accuracy of the CNN-1536-128 model and the Xception model is 95.10% and 97.10%, respectively. On the other hand, the training of the VGG16 model and the ResNet50 model cannot converge under the limited data. For efficiency, the CNN-1536-128 model has the least number of parameters and requires the shortest training time, while the Xception model has the most parameters.

V. CONCLUSION AND FUTURE WORK

A new DCNN-TL diagnosis method combining DCNN and TL is proposed for mechanical faults in this work, which is capable of high-level feature extraction and classification. The optimal DCNN constructed by GAP and deep transfer learning is used as a feature extractor to enhance the feature learning ability. Three datasets, namely the CWRU dataset, SPCP dataset, and UPB dataset, are utilized for fault diagnosis experiments. It can be seen from the results that the suggested DCNN-TL method can achieve 100% accuracy on the three datasets. And the proposed method can increase the diagnostic accuracy of fault diagnosis and decrease the calculation cost. In future work, we will incorporate other transfer learning methods and classifiers into the proposed classification framework to improve its robustness.

Moreover, the vibration signal in real applications is usually non-standard, and we will verify it in our future work.

REFERENCES

- [1] Z. Li, R. Outbib, S. Giurgea, D. Hissel, S. Jemei, A. Giraud, and S. Rosini, "Online implementation of SVM based fault diagnosis strategy for PEMFC systems," *Appl. Energy*, vol. 164, pp. 284–293, Feb. 2016.
- [2] W. Du, M. Kang, and M. Pecht, "Fault diagnosis using adaptive multifractal detrended fluctuation analysis," *IEEE Trans. Ind. Electron.*, vol. 67, no. 3, pp. 2272–2282, Mar. 2020.
- [3] 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.
- [4] F. Liang, D. Zhang, and R. Zhang, "Fault diagnosis of water lubricated stern bearing based on BP neural network," *J. Mech. Transmiss.*, vol. 39, no. 7, pp. 118–121, 2015.
- [5] R. Liu, B. Yang, X. Zhang, S. Wang, and X. Chen, "Time-frequency atoms-driven support vector machine method for bearings incipient fault diagnosis," *Mech. Syst. Signal Process.*, vol. 75, pp. 345–370, Jun. 2016.
- [6] J. Zheng, H. Pan, and J. Cheng, "Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines," *Mech. Syst. Signal Process.*, vol. 85, pp. 746–759, Feb. 2017.
- [7] K. H. Hui, M. H. Lim, M. S. Leong, and S. M. Al-Obaidi, "Dempster-Shafer evidence theory for multi-bearing faults diagnosis," *Eng. Appl. Artif. Intell.*, vol. 57, pp. 160–170, Jan. 2017.
- [8] X. Yan and M. Jia, "A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing," *Neurocomputing*, vol. 313, pp. 47–64, Nov. 2018.
- [9] D.-T. Hoang and H.-J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, Mar. 2019.
- [10] S. Deng, Z. Cheng, C. Li, X. Yao, Z. Chen, and R.-V. Sanchez, "Rolling bearing fault diagnosis based on deep Boltzmann machines," in *Proc. IEEE Prognostics Syst. Health Manage. Conf.*, Oct. 2016, pp. 1–6.
- [11] J. Tao, Y. Liu, and D. Yang, "Bearing fault diagnosis based on deep belief network and multisensor information fusion," *Shock Vib.*, vol. 2016, pp. 1–9, Aug. 2016.
- [12] O. Janssens, V. Slavkovic, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. Van de Walle, and S. Van Hoecke, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016.
- [13] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement*, vol. 93, pp. 490–502, Nov. 2016.
- [14] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018.
- [15] M. Azamfar, J. Singh, I. Bravo-Imaz, and J. Lee, "Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis," *Mech. Syst. Signal Process.*, vol. 144, Oct. 2020, Art. no. 106861.
- [16] D. Kolar, D. Lisjak, M. Pajak, and D. Pavković, "Fault diagnosis of rotary machines using deep convolutional neural network with wide three axis vibration signal input," *Sensors*, vol. 20, no. 14, p. 4017, Jul. 2020.
- [17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015, pp. 1–14.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE-CVPR*, Jun. 2016, pp. 770–778.
- [20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE-CVPR*, Jun. 2016, pp. 2818–2826.
- [21] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. AAAI-AI*, 2017, pp. 1–7.
- [22] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: <https://arxiv.org/abs/1704.04861>

- [23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE-CVPR*, Jun. 2018, pp. 4510–4520.
- [24] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE-CVPR*, Jul. 2017, pp. 1251–1258.
- [25] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE-CVPR*, Jul. 2017, pp. 4700–4708.
- [26] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proc. IEEE-CVPR*, Jun. 2018, pp. 8697–8710.
- [27] P. Cao, S. Zhang, and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning," *IEEE Access*, vol. 6, pp. 26241–26253, 2018.
- [28] V. T. Do and U.-P. Chong, "Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain," *Strojnikski Vestnik*, vol. 57, no. 9, pp. 655–666, Sep. 2011.
- [29] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*. [Online]. Available: <https://arxiv.org/abs/1312.4400>
- [30] C. Lu, Y. Wang, M. Ragulskis, and Y. Cheng, "Fault diagnosis for rotating machinery: A method based on image processing," *PLoS ONE*, vol. 11, no. 10, Oct. 2016, Art. no. e0164111.
- [31] L. Christian, K. James, Z. Detmar, 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.*, 2016, pp. 5–8.
- [32] B. Hasan, "Model-free drive system current monitoring: Faults detection and diagnosis through statistical features extraction and support vector machines classification," School Eng. Des. Technol., Univ. Bradford, Bradford, U.K., Tech. Rep. 106927644, 2012.



DONG ZHANG was born in 1986. He received the B.S. degree from the Nanjing University of Aeronautics and Astronautics, in 2007, and the M.S. degree from Shanghai Jiao Tong University, in 2010. He is currently working with the China Ship Development and Design Center. His research interests include machine learning, fatigue diagnose, and structural design.



TAOTAO ZHOU was born in 1990. He received the B.S. degree from the Wuhan University of Technology, in 2010, and the M.S. degree from the China Ship Development and Design Center, in 2015. He is currently working with the China Ship Development and Design Center. His research interests include marine structural design and optimization, machine learning, and surrogate model.

...