Fault Diagnosis of Rolling Bearings of Different Working Conditions Based on Multi-Feature Spatial Domain Adaptation

TAO WEN[ID], **RENXIANG CHEN**[ID], **AND LINLIN TANG**[ID]

Chongqing Engineering Laboratory for Transportation Engineering Application Robot, Chongqing Jiaotong University, Chongqing 400074, China

Corresponding author: Renxiang Chen (manlou.yue@126.com)

ABSTRACT The running state of rolling bearings is complex in operation, and the data are generally collected under different working conditions. However, when single-source domain adaptation is used to model the heterogeneously distributed data obtained under different working conditions, the domain-invariant representations can hardly be used for representation, which directly affects the fault diagnosis rate. To this end, a method for the fault diagnosis of rolling bearings under different working conditions based on multi-feature spatial domain adaptation is proposed. Firstly, all the data from source and target domains are mapped into a feature space to learn the common representations of all domains. Secondly, the data for each pair of source and target domains are mapped into different feature spaces to get the fault feature representations under various working conditions. And the multi-domain adaptation network is used for the domain-specific distribution alignment to learn multiple domain-invariant representations. Thirdly, these representations are used to train multiple domain-specific classifiers, thus obtaining the recognition result for each domain-invariant representation. Finally, the domain-specific decision boundaries predicted by multiple classifiers are employed to align the classifiers' output of target samples and thus to reduce the influence from different classifiers. The effectiveness and feasibility of this proposed method have been verified by diagnostic experiments conducted according to the rolling bearing data from Case Western Reserve University and Laboratory, respectively.

INDEX TERMS Rolling bearing, different working conditions, multi-feature spatial, domain adaptation, fault diagnosis.

I. INTRODUCTION

Rotating machinery is critical to the efficiency and safety of mechanical systems [1]. Rolling bearing is a major component of rotating machinery, and its running state significantly affects the health of such machinery. Therefore, it is necessary to diagnose this running state to avoid production loss or even casualty caused by such critical fault [2]–[4]. Thereinto, the state monitoring system is essential for normal equipment operation. As the manufacturing industry enters a new era of big data and intelligence, an increasing number of data

has been collected from this system under various working conditions [5], making it more necessary to study how to utilize these data to diagnose and analyze the bearing state.

Deep transfer learning has become a new research direction. According to its basic concept, an adaptive layer is added between the feature extraction and classification layers so that the data on the source and target domains can be distributed more similarly [6]–[10]. Besides, more expressive features can be extracted automatically by deep transfer learning and it utilizes the self-extraction of deep learning together with transfer learning to acquire "new knowledge", which assist to solve small sample data problem well in few-shot learning [11]. This learning method is commonly

The associate editor coordinating the review of this manuscript and approving it for publication was Yuan Zhuang[ID].

used for mechanical fault diagnosis [12]–[16]. For example, Zhao *et al.* [17] reported a transfer learning framework based on the deep multi-scale convolutional neural network. In this framework, the adaptive layer's weighting parameters are adjusted slightly to diagnose the rolling bearings intelligently. Wang *et al.* [18] diagnosed the fault of rolling bearings using the transfer learning strategy that combines the variable convolutional neural network with the deep long short-term memories. Tong *et al.* [19] put forward a new method to diagnose the bearing fault, which is known as the domain adaptive feature transfer learning under variable working conditions. That is, the pseudo test labels are refined on the basis of maximum mean discrepancy (MMD) and domain-invariant clustering in a public space to represent the transferable features of training and testing data. Zhang *et al.* [20] proposed a transfer learning method for fault diagnosis based on the neural network so as to improve the diagnostic performance through the data of different working conditions. As shown in the bearing data from Case Western Reserve University, the effectiveness of this method has been verified.

The idea of deep transfer has been used for the fault diagnosis of rotating machinery by the above methods, which achieves good results. However, there are some problems as follows. 1) The focus is completely on the single-source unsupervised domain adaptation. However, in actual machinery operation, the labeled data can be collected from many different sources, such as different rotational speeds or loads [21]–[24]. Therefore, the adapters from multiple source domains shall not be modeled in the same way. 2) The common domain-invariant representations of all domains are primarily extracted by aligning the source distribution as well as target domains in a common feature space. However, the representations can hardly be extracted for the data under all working conditions in the multi-source unsupervised domain adaptation (MUDA) [25]. Besides, when we try to align multiple source and target domains, the bigger mismatch might result in unsatisfying performance in Figure 1. 3) The domain-specific decision boundary between clusters is not considered when these methods match the distribution. Hence, it is necessary to improve the diagnostic accuracy of rolling bearings under variable working conditions through the heterogeneously distributed data and the decision boundary of different domains. Zhu *et al.* [26] proposed a migration learning method based on adaptive multi-source domain. A multiple adversarial learning strategy is used to obtain feature representations, but the method is an adversarial-based learning method and the training time can be too much. Besides, the method does not consider the effects of different classifiers. Zhu *et al.* [27] introduced a new MUDA framework that aligns the distribution of each pair of source and target domains in multiple specific feature spaces, respectively. Also, it aligns the classifiers' outputs through domain-specific decision boundaries.

To sum this, This paper proposes a method to diagnose rolling bearings' fault under different working conditions
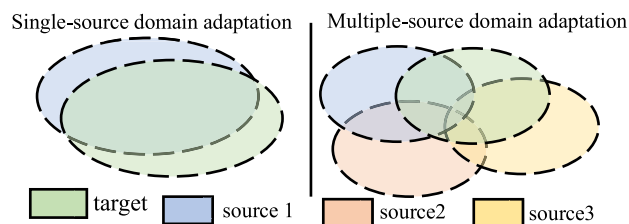

FIGURE 1. Main idea diagram.

based on multi-feature spatial domain adaptation to solve the above problems. First, the data of all source and target domains are mapped into a feature space to learn the common domain-invariant representations. Second, the data about each pair of source and target domains are mapped into multiple different feature spaces, and the domain-specific distribution is aligned with in order to learn multiple domain-invariant representations. Third, these representations are used to train multiple domain-specific classifiers. Finally, such classifiers are aligned by the domain-specific decision boundary to solve the problem that the target samples close to this boundary may get various labels, which is predicted by different classifiers. Numerous experiments have proved that this proposed method is more accurate than existing ones, showing prominent advantages.

The main contributions of this paper are summarized as follows:

1) Most previous fault diagnosis approaches based on multi-source domain adaptation have focused on extracting domain invariant representations of all domains without considering domain-specific decision boundaries between clusters. In this paper, we propose an approach based on a multi-feature spatial adaptation that aligns domain-specific distributions of each pair of source and target domains by learning multiple domain-invariant representations and the classifiers' output from multiple domains.

2) Traditional methods use only a set of source domain data for training, while our approach collected many sets of source domain data that has been monitored from different working conditions to identify the target data, which extracts various fault features from multiple source domains to achieve a more effective result.

This paper is organized as follows: Section 2 is the theory of MFSAN. Section 3 describes the entire processing of fault diagnosis procedure. In Section 4, the comparison methods and experimental result analysis are given. Section 5 introduces the generalization ability verification experiment. Finally, the conclusions are drawn in Section 6.

II. ALIGNMENT OF MULTI-FEATURE SPATIAL DOMAIN ADAPTATION

In the MUDA, there are N different source distributions in the bottom layer, which are expressed as $\{p_{sj}(x, y)\}_{j=1}^{N}$. The source domain data are labeled as $\{(X_{sj}, Y_{sj})\}_{j=1}^{N}$, where
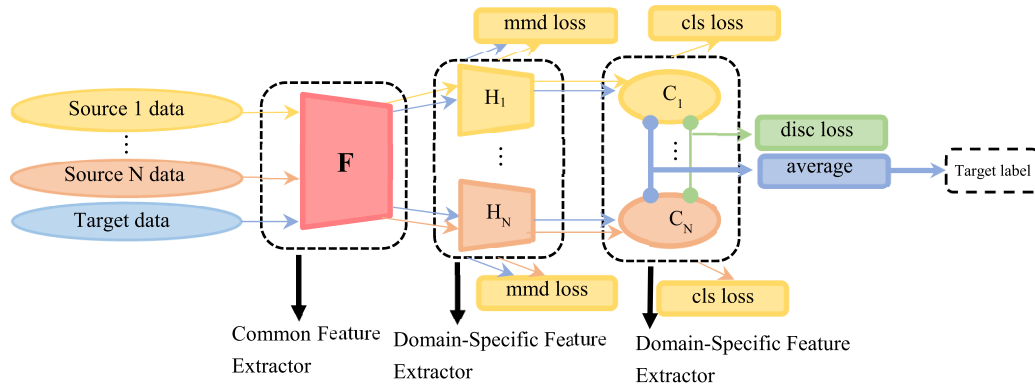
FIGURE 2. Network structure diagram.

$X_{sj} = \left\{x_i^{sj}\right\}_{i=1}^{|X_{sj}|}$ refers to the sample of the source domain j, and $Y_{sj} = \left\{y_i^{sj}\right\}_{i=1}^{|X_{sj}|}$ refers to the corresponding label. Besides, the target distribution is $p_t(x, y)$, from which the target source data $X_t = \left\{x_i^t\right\}_{i=1}^{|X_t|}$ are sampled, but Y_t is not labeled.

Figure 2 shows the major idea of Multi-feature spatial adaptive network (MFSAN), which includes two alignment phases, namely, learning the source-specific domain-invariant representations and aligning the classifier output of target samples. Our framework is composed of one common feature extractor, N domain-specific feature extractors, and N source-specific classifiers of target samples.

A. COMMON FEATURE EXTRACTOR

In this paper, a common sub-network $f(\cdot)$, which can map the images from the original feature space into a common one, is proposed to extract the common representations of all domains. In our study, the deep learning network ResNet50 [28] that has been frequently used is used as the common sub-network.

B. DOMAIN-SPECIFIC FEATURE EXTRACTOR

We expect that each pair of source and target domain data can be mapped into a specific feature space. Two batch images, x^{sj} and x^t, are given from the source domain (X_{sj}, Y_{sj}) and target domain X^t, respectively. These domain-specific feature extractors will receive the common features, $f(x^{sj})$ and $f(x^t)$, from the common feature extractor. Then, each pair of source and target domains will be mapped to a specific feature space by their corresponding unshared domain-specific subnetworks $h_j(\cdot)$, which exists in N source domain (X_{sj}, Y_{sj}).

C. DOMAIN-SPECIFIC CLASSIFIER

C is a multi-output network composed of N domain-specific predictors $(C_j)_{j=1}^N$. Each predictor C_j is a Softmax classifier that receives the specific domain-invariant feature through j-th domain-specific feature extractors of the source domain, i.e., $H(F(x))$. $F(\cdot)$ represents the common feature extractor

and $H(\cdot)$ is a domain-specific extractor. Then, the classification loss is added to each classifier using the cross-entropy, with the formula expressed as:

$$L_{cls} = \sum_{j=1}^{N} \mathbf{E}_{x \sim X_{sj}} J(C_j(H_j(F(x_i^{sj}))), y_i^{sj}) \qquad (1)$$

where $\mathbf{E}_x[\cdot]$ is the expected value operator with respect to the distribution x and $J(\cdot, \cdot)$ represents the cross-entropy loss function (classification loss).

D. DOMAIN-SPECIFIC DISTRIBUTION ALIGNMENT

To complete the first alignment phase (aligning the distribution of each pair of source and target domains), MMD [29] is used to estimate the interpolation between two domains. MMD refers to the kernel test that refuses or accepts the null hypothesis $p = q$ by observing samples, where p and q are two different distributions. Formatively, it defines the following measurement of the discrepancy:

$$d_{\mathcal{H}}(p, q) \triangleq \left\| \mathbf{E}_p\left[\varnothing\left(x^s\right)\right] - \mathbf{E}_q[\varnothing\left(x^t\right)] \right\|_{\mathcal{H}}^2 \qquad (2)$$

where $\mathcal{H}$ represents the reproducing kernel Hilbert space (RHKS) of the feature kernel k. Here $\varnothing(\cdot)$ means mapping the original samples to RKHS. In practice, the MMD is estimated to compare the average kernel embedding between squared distances.

$$\hat{d}_H(p, q) = \left\| \frac{1}{n_s} \sum_{x_i \in D_S} \varnothing(x_i) - \frac{1}{n_t} \sum_{x_j \in D_t} \varnothing(x_j) \right\|_H^2 \qquad (3)$$

We estimate the discrepancy between source and target domains through (3), where $\hat{d}_H(p, q)$ is the unbiased estimation of $d_H(p, q)$. The MMD loss is rewritten as:

$$L_{mmd} = \frac{1}{N} \sum_{j=1}^{N} \hat{d}\left(H_j(F(X_{sj})), H_j\left(F\left(X_t\right)\right)\right) \qquad (4)$$

Every specific feature extractor can be used to learn the domain-invariant representation for each pair of source and target domains by minimizing the (4).

E. DOMAIN-SPECIFIC CLASSIFIER ALIGNMENT

The target samples that are near the cluster boundary are more likely to be misclassified by classifiers from the source samples. Because classifiers are trained on various source domains, their prediction on target samples, especially those near the cluster boundary, may be different. Intuitively, the same target samples predicted by different classifiers shall produce an identical result. Therefore, the discrepancy among all classifiers will be minimized in the second alignment phase. The absolute value for the discrepancy of the classifier to the target data output is taken as the discrepancy loss:

$$L_{disc} = \frac{2}{N \times (N-1)} \sum_{j=1}^{N-1}$$

$$\times \sum_{j=j+1}^{N} \mathbf{E}_{x \sim X_t} \left[\left| C_i(H_i(F(x_k))) - C_j(H_j(F(x_k))) \right| \right] \quad (5)$$

Xu *et al.* [30] propose a target classification operator to combine various source classifiers. However, it will be complex to vote on the labels of target samples. A similar probability output among all classifiers can be achieved by minimizing (5). Finally, the mean value of all classifiers' outputs is calculated to forecast the labels of target samples.

F. NETWORK TRAINING

In general, our method contains classification loss, MMD loss, and classifier discrepancy loss. Concretely speaking, the network can classify the source domain data accurately by minimizing the classification loss; the domain-invariant representations can be learned by minimizing the MMD loss; the discrepancy between classifiers can be reduced by minimizing the classifier discrepancy loss. At last, the total loss is expressed as:

$$L_{total} = L_{cls} + \lambda L_{mmd} + \gamma L_{disc} \quad (6)$$

The training is mainly conducted based on the standard's small-batch stochastic gradient descent (SGD) algorithm.

III. METHOD FOR THE FAULT DIAGNOSIS OF ROLLING BEARINGS UNDER DIFFERENT ROTATIONAL SPEEDS BASED ON MULTI-FEATURE SPATIAL DOMAIN ADAPTATION

This paper introduces a method to detect the fault of rolling bearings under various working conditions based on multi-feature spatial domain adaptation. Firstly, we map the data of all source and target domains into a feature space to learn the common domain-invariant representations. Then, we map the data about each pair of source and target domains to several different feature spaces and align the domain-specific distribution to learn multiple domain-invariant representations. Thirdly, we use these representations to train multiple domain-specific classifiers. The

TABLE 1. The fault types.

Fault Types	Degree of faults/(mm)	Labels
Normal		N
Outer ring	0.18/0.36/0.53	O1/O2/O3
Inner ring	0.18/0.36/0.53	I1/I2/I3
Rolling element	0.18/0.36/0.53	R1/R2/R3

classifier's output of target samples will be aligned by the domain-specific decision boundary because the labels may be different for target samples near this boundary as predicted by different classifiers. Finally, we calculate the mean value of all classifier outputs to predict such samples' labels. The flow chart is shown in Figure 3:

1) Collect data samples. Calculate the time-frequency map through wavelet transform. Divide the training set (source domain) and testing set (target domain) in a certain proportion. Construct the data set under different working conditions.

2) Initialize the MFSAN's network parameters, and input multiple source and target domain data into the MFSAN in bulk.

3) Extract the common feature representations of all domains through the common sub-network ResNet50.

4) Construct the domain-specific feature extractor that receives the common features, $f\left(x^{sj}\right)$ and $f\left(x^{t}\right)$, from the common feature extractor.

5) Use the MMD to measure the discrepancy, and align all the domain-specific distributions to reduce their discrepancy.

6) Construct N domain-specific classifiers. Every predictor C_j is a Softmax classifier that receives the specific domain-invariant features extracted by j-th specific feature extractor of the source domain.

7) Take the absolute value for the interpolation of the classifier to the probability output of target data as the discrepancy loss to minimize the discrepancy among all classifiers.

8) At last, calculate the mean value of all classifiers' outputs to predict the labels of target samples.

IV. VERIFICATION EXPERIMENT ON THE DATA FROM CASE WESTERN RESERVE UNIVERSITY

A. DATA DESCRIPTION

In this experiment, the proposed method's effectiveness was verified by faults of the deep groove ball bearing 6205-2RS. The bearing data were assembled from the Case Western Reserve University Bearing Data Center Website. Single-point faults were set on the bearing's inner ring, outer ring, and rolling element through the electric discharging machining technique, with the fault diameter of 0.18mm, 0.36mm, and 0.53mm, respectively, and the depth of 0.28mm. Finally, there were 9 fault states, as shown in Table 1.
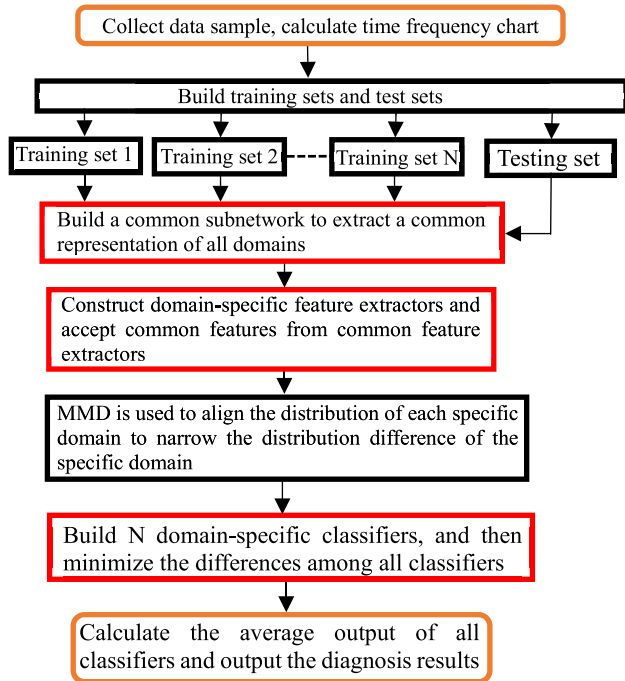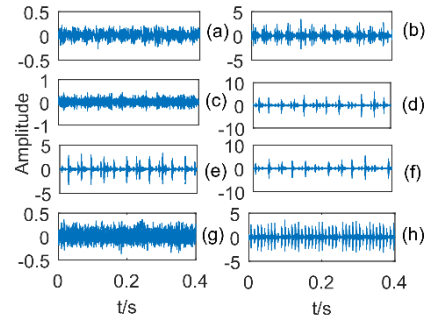
FIGURE 3. The flow chart of method.

FIGURE 4. The time domain signal under different fault types: (a) N/1797. (b) I3/1797. (c) B3/1797. (d) O3/1797. (e) O3/1772. (f) O3/1750. (g) O2/1797. (h) O1/1797.
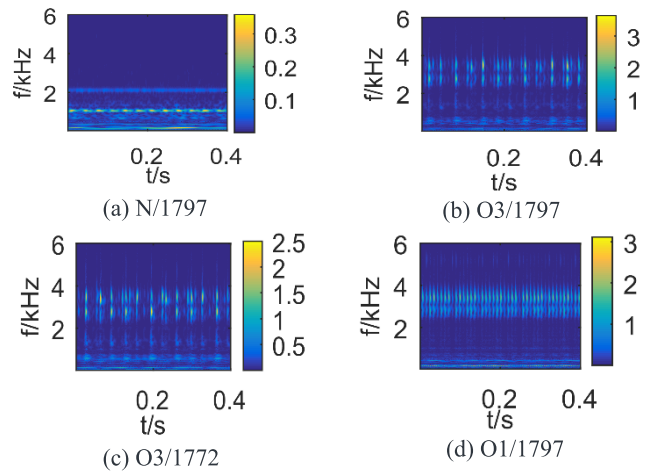
TABLE 2. Condition information.

Speeds(r/min)	Load/(HP)	Labels
1797	0	A
1772	1	B
1750	2	C

FIGURE 5. The wavelet transform time-frequency diagram.

When the data were collected, the bearing was working at a constant speed of 1797rpm, 1772rpm, and1750rpm under working conditions A, B, and C, respectively. Besides, the load was 0HP, 1HP, and 2HP, respectively, and the sampling frequency was 12KHz. Table 2 exhibits the working conditions of this experiment.

B. SET UP A DATA SET

Samples were intercepted from the collected vibration data at 1024 points, and the data were resampled every 500 points. Thus, there were 200 samples in each fault type. First, each sample's wavelet transform was conducted to obtain the time-frequency map samples, from which 150 ones were selected randomly. Thus, there were 1500 samples used as the training set among the 10 fault types. Second, the remaining 50 ones were selected for the test. Thus, there were 500 samples used as the testing set among the 10 fault types.

Figure 4 shows the time domain waveform of bearing components' fault signals in different fault states. The rolling element's damage degree and fault types can hardly be determined through this waveform.

Figure 5 is the time-frequency map about the normal state and the outer ring fault under different pitting degrees and rotational speeds. Due to limited space, the time-frequency maps of the remaining faults will not be described in this paper. By comparing Figures 4 and 5, the faults were much more diverse and changeable in the wavelet transform's time-frequency representation than in the time domain.

As shown in Figure 5, the bearing's signal energy was mainly concentrated in the low-frequency range in a normal state and kept steady and low during the whole period. However, the energy was distributed primarily in the high-frequency range with a wide frequency domain when the outer ring witnessed different damage degrees. During this period, the energy showed a higher value and fluctuated significantly, with an apparent non-stationary property. According to the comparison between Figures 5 (b) and (c), the energy was distributed in a similar structure when the outer ring had the same degree of damage under different working conditions. However, as the rotation accelerated, the vibration energy increased. The comparison results showed that the vibration energy increased with the increase of fault degree, but the elapsed time of fluctuation reduced accordingly. It can be seen from this that under different working conditions, the fault signals were much more diverse

TABLE 3. Performance comparison of classification accuracy(%).

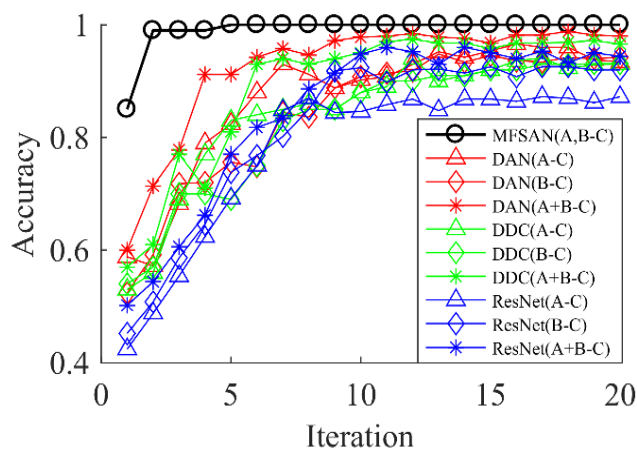| Standards	Method	A,B→C		A,C→B		B,C→A		Avg	Std	A→A
Multi-source	MFSAN	100		99		99		99.3	0.57	100
Source Combine	DAN	99		98.8		96.6		98.1	1.33	
	DDC	97.6		96		95.1		96.2	1.27	
	ResNet	96.4		94.8		94.8		95.3	0.92	
Single Source	DAN	A→C 95.4		A→B 98		B→A 92.6		95.2	2.11	98.2
		B→C 95.2		C→B 97		C→A 93.1				
	DDC	A→C 93		A→B 96.1		B→A 91.1		93.7	2.85	96.9
		B→C 94.1		C→B 97.6		C→A 90.2				
	ResNet	A→C 87.2		A→B 89.4		B→A 90.2		89.9	3.53	97.2
		B→C 93		C→B 94.6		C→A 85.1				

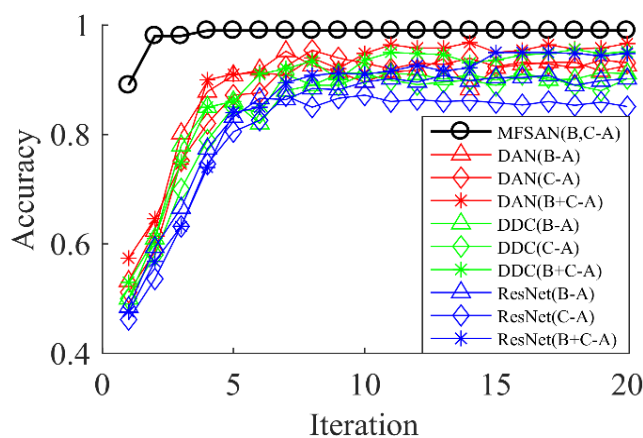**FIGURE 6.** A,B-C Transfer results.

FIGURE 7. B,C-A Transfer results.

in the wavelet time-frequency domain than in the time domain and changed more abundantly. This indicates that the wavelet transform can fully display the fault features.

C. EXPERIMENTAL RESULTS AND ANALYSIS

In the proposed method, the MMD distance that is frequently used for transfer learning was adopted to measure the distributional discrepancy between two domains. Therefore, this method was compared with those following the same measure criteria, such as deep adaptive methods DAN [31] and DDC [32]) and deep learning method (ResNet). Besides, the algorithms were compared and tested using source combine and single source, respectively. Thereinto, the former meant that various source domain data were mixed as the training set to test the single target data. As shown by the comparison results in Table 3, the proposed method's average diagnostic precision was highest (99.3%), and the standard deviation was 0.57%, indicating the effectiveness of this method.

According to Table 3, the single source's ResNet showed a recognition precision of 97.2% under the same working

condition A→A, but the average diagnostic precision was merely 89.9% under different conditions. This indicates that ResNet can effectively extract fault features in bearings, but it lacks the ability of adaptation and cannot reduce the difference in data distribution caused by various working conditions. This shows the need for migration adaptation of data from different working conditions. DAN and DDC of the single source domain approaches were more precise than ResNet because they have an adaptation layer, which can reduce the difference of data distribution in different working conditions. The average diagnosis precision of DAN is higher than DDC because the MMD used by the former is multi-kernel while by the latter is single-kernel. A single fixed kernel may not be the optimal one. The DAN, DDC, and ResNet of the source combined were more precise than those of the single source, which indicated that the recognition precision was improved by data diversity.

However, compared with the proposed method, the source combine's values were relatively low, which implied that the mixed transfer could affect the recognition result. In this method, multiple source domains were adapted,
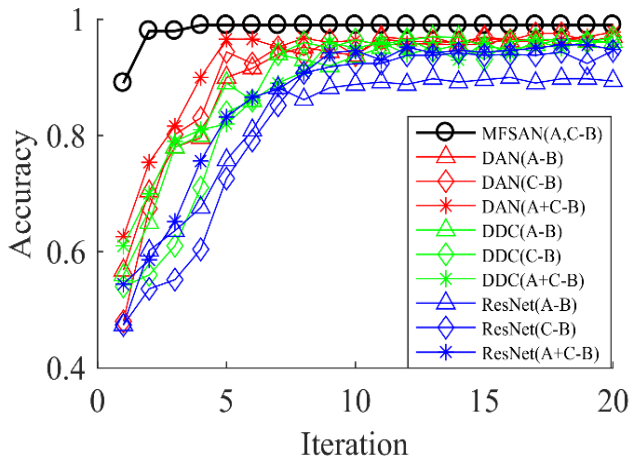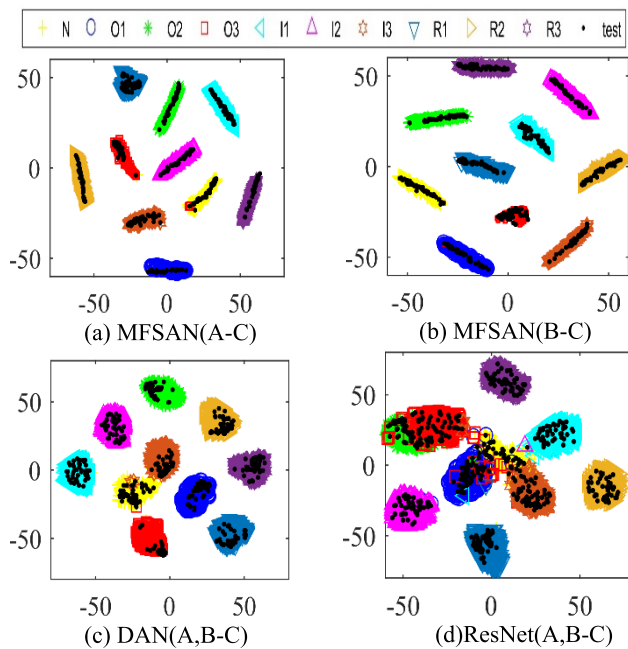
FIGURE 8. A,C-B Transfer results.

FIGURE 9. Characteristics of figure.

FIGURE 10. The experimental station.

TABLE 4. The fault types.

Fault Types	Degree of faults/(mm)	Labels
Normal		N
Outer ring	0.2	O
Inner ring	0.2	I
Rolling element	0.2	R

respectively, to learn more specific domain-invariant representations. Besides, the domain-specific distribution of each pair of source and target domains and the domain-specific classifier output of target samples were aligned in various feature spaces. Figures 6, 7, and 8 are iteration diagrams for the test precision of the proposed method and other ones. From this, we can see that the proposed method has converged at the fifth iteration, showing higher convergence speed and stability.

To verify the effectiveness of the proposed method further, we visualize the characteristics of DAN, ResNet, and this method under working conditions A, B, and C, as shown in Figure 9. The characteristics of MFSAN's two different feature spaces are presented in Figures 9 (a) and (b), respectively. This indicates that the types are separated from each other in these two spaces, and the training and testing sets
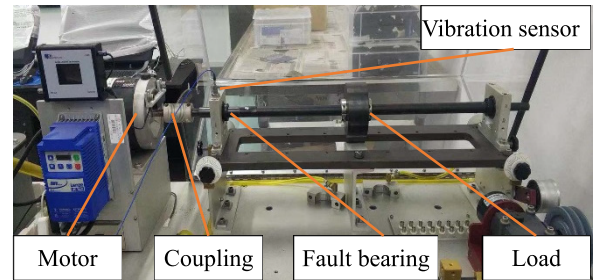
are located in similar positions. Figure 9 (c) describes the characteristics of DAN, which shows a better overall effect compared with ResNet. However, these characteristics are overlapped because the training sets A and B locate in different feature spaces. In this case, the characteristics may be overlapped when the two data sets are mixed to adapt the testing set C.

V. GENERALIZATION ABILITY VERIFICATION EXPERIMENT ON THE DATA FROM LABORATORY

A. DATA DESCRIPTION

The experimental data were collected from the synthetic test-bed for mechanical fault simulation (MFS), which is mainly composed of motor, coupler, rolling bearing, and vibration acceleration sensor. (Figure 10) Thereinto, the acceleration sensor was attached to the bearing support through a magnet base. In this experiment, we adopted the deep groove ball bearing er-16k with a nodal diameter of 38.51mm. There were 9 rolling elements in total, with the contact angle of 9.08°. Single-point faults were set on the bearing's inner ring, outer ring, and rolling elements using the electric discharging machining technique, with the fault diameter of 0.1mm and the depth of 0.2mm. Four fault states were simulated in this experiment, as shown in Table 4.

When the data were collected, the bearing was working at a constant speed of 1200rpm, 900rpm, and 600rpm under working conditions D, E, and F, respectively. Besides, the load was 5kg, the sampling frequency was 25.6KHz, and the sampling time was 10s. The data were collected 4 times for each fault.

B. SET UP A DATA SET

Samples were intercepted from the collected vibration data at 1024 points, and 1000 samples were collected from each fault type. First, the wavelet transform of each sample was conducted to obtain the time-frequency map samples, from

TABLE 5. Performance comparison of classification accuracy(%).

Standards	Method	D,E→F		D,F→E		E,F→D		Avg	Std	D→D
Multi-source	MFSAN	100		99		97		98.7	1.52	100
Source Combine	DAN	97		95.4		93		95.1	2.01	
	DDC	95.4		95.1		92.1		94.2	1.52	
	ResNet	85		86.7		85.2		85.6	0.92	
Single Source	DAN	D→F 94.6		D→E 93		E→D 92.6		93	1.60	96.9
		E→F 93.2		F→E 93.5		F→D 90.1				
	DDC	D→F 95.7		D→E 92.2		E→D 90.1		92.3	2.02	97.5
		E→F 91.4		F→E 93.6		F→D 91				
	ResNet	D→F 84.6		D→E 81.8		E→D 83.5		83.7	2.91	97.6
		E→F 89		F→E 82.5		F→D 80.0				
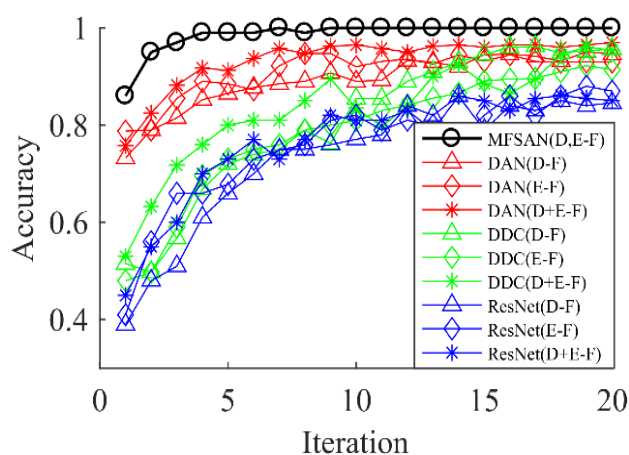
FIGURE 11. D,E-F transfer results.

FIGURE 12. D,F-E transfer results.

FIGURE 13. E,F-D transfer results.

C. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental results were identical to those in the previous section when the proposed method was compared with DAN, DDC, and ResNet. The algorithms were compared and tested by means of source combine and single source, respectively. As shown by the comparison results in Table 5, the proposed method's average diagnostic precision was highest, reaching 98.7%, and the standard deviation was 1.52%, indicating that this method was effective.

It can be found by further observing Table 5 that the overall diagnostic accuracies of DAN, DDC and ResNet of the sources combined domain was better than those in the single source domain, where the standard deviation of ResNet in the mixed source domain was the smallest at 0.92%, but the average diagnostic accuracy was only 85.6%. The average diagnostic accuracies of the mixed-source domain DAN and DDC methods were 95.1% and 94.2%, respectively, which were both lower than the proposed method. In the MFSAN, multiple sources and target domains were adapted, respectively, to learn various specific domain-invariant representations. The domain-specific distribution of each pair of source
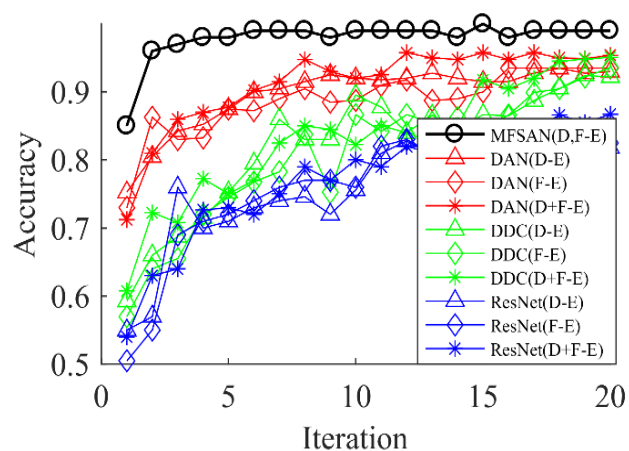
which 300 ones were selected randomly. Thus, 1200 samples were used as the training set among the 4 fault types. Second, 100 samples were selected randomly for the test. Thus, 400 samples were used as the testing set among the 4 fault types.

and target domains and the domain-specific classifier output of target samples were aligned in multiple feature spaces to improve the recognition precision of rolling bearings under different working conditions. Figures 10, 11, and 12 are iteration diagrams for the test precision of the proposed method and other ones. From this, we can see that the proposed method shows higher convergence speed and stability.

VI. CONCLUSION

This paper proposes a method to detect the fault of rolling bearings under different working conditions based on multi-feature spatial domain adaptation. Firstly, the data of all source and target domains are mapped into a feature space to learn the common domain-invariant representations. Then, the data for each pair of source and target domains are mapped into a number of different feature spaces, and the domain-specific distribution is aligned to learn multiple domain-invariant representations. Thirdly, multiple domain-specific classifiers are trained by these representations to obtain the recognition result for each of them. Finally, the classifiers' output of target samples is aligned through the domain-specific decision boundary predicted by various classifiers, thus reducing the influence of such classifiers. Diagnostic experiments have been carried out based on the rolling bearing data from Case Western Reserve University and Laboratory, respectively. Experiments have shown that compared with existing methods, our method can improve the recognition precision for the fault diagnosis of rolling bearings under different working conditions. The characteristics of this method are described as follows:

1) Align the domain-specific distribution of each pair of source and target domains; and align the domain-specific classifier output of target samples in multiple feature spaces.

2) Diagnose the bearing state based on the data collected under different working conditions so that the manufacturing industry can adapt to the new era of big data and intelligence.

REFERENCES

[1] B. Keshtegar, M. Bagheri, C.-W. Fei, C. Lu, O. Taylan, and D.-K. Thai, "Multi-extremum-modified response basis model for nonlinear response prediction of dynamic turbine blisk," *Eng. Comput.*, to be published, doi: 10.1007/s00366-020-01273-8.

[2] R. Chen, X. Huang, L. Yang, X. Xu, X. Zhang, and Y. Zhang, "Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform," *Comput. Ind.*, vol. 106, pp. 48–59, Apr. 2019.

[3] H. Zhao, J. Zheng, J. Xu, and W. Deng, "Fault diagnosis method based on principal component analysis and broad learning system," *IEEE Access*, vol. 7, pp. 99263–99272, 2019.

[4] W. Qian, S. Li, and J. Wang, "A new transfer learning method and its application on rotating machine fault diagnosis under variant working conditions," *IEEE Access*, vol. 6, pp. 69907–69917, 2018.

[5] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 1, pp. 136–144, Jan. 2019.

[6] Z. Chen, K. Gryllias, and W. Li, "Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 339–349, Jan. 2020.

[7] X. Wang, L. Li, W. Ye, M. Long, and J. Wang, "Transferable attention for domain adaptation," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, Jul. 2019, pp. 5345–5352.

[8] Z. Murez, S. Kolouri, D. Kriegman, R. Ramamoorthi, and K. Kim, "Image to image translation for domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4500–4509.

[9] A. Rozantsev, M. Salzmann, and P. Fua, "Residual parameter transfer for deep domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4339–4348.

[10] J. Wang and Y. Zhai, "Prototypical siamese networks for few-shot learning," in *Proc. IEEE 10th Int. Conf. Electron. Inf. Emergency Commun. (ICEIEC)*, Jul. 2020, pp. 1–11.

[11] K. Saito, K. Watanabe, Y. Ushiku, and T. Harada, "Maximum classifier discrepancy for unsupervised domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3723–3732.

[12] J. Li, X. Li, D. He, and Y. Qu, "A domain adaptation model for early gear pitting fault diagnosis based on deep transfer learning network," *Proc. Inst. Mech. Eng., O, J. Risk Rel.*, vol. 234, no. 1, pp. 168–182, Feb. 2020.

[13] J. Zhu, N. Chen, and C. Shen, "A new deep transfer learning method for bearing fault diagnosis under different working conditions," *IEEE Sensors J.*, vol. 20, no. 15, pp. 8394–8402, Aug. 2020.

[14] J. An, P. Ai, and D. Liu, "Deep domain adaptation model for bearing fault diagnosis with domain alignment and discriminative feature learning," *Shock Vibrat.*, vol. 2020, pp. 1–14, Mar. 2020.

[15] Z. He, H. Shao, X. Zhang, J. Cheng, and Y. Yang, "Improved deep transfer auto-encoder for fault diagnosis of gearbox under variable working conditions with small training samples," *IEEE Access*, vol. 7, pp. 115368–115377, 2019.

[16] T. Han, C. Liu, W. Yang, and D. Jiang, "Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application," *ISA Trans.*, vol. 97, pp. 269–281, Feb. 2020.

[17] B. Zhao, X. Zhang, Z. Zhan, and S. Pang, "Deep multi-scale convolutional transfer learning network: A novel method for intelligent fault diagnosis of rolling bearings under variable working conditions and domains," *Neurocomputing*, vol. 407, pp. 24–38, Sep. 2020.

[18] Z. Wang, Q. Liu, H. Chen, and X. Chu, "A deformable CNN-DLSTM based transfer learning method for fault diagnosis of rolling bearing under multiple working conditions," *Int. J. Prod. Res.*, pp. 1–15, Aug. 2020.

[19] Z. Tong, W. Li, B. Zhang, F. Jiang, and G. Zhou, "Bearing fault diagnosis under variable working conditions based on domain adaptation using feature transfer learning," *IEEE Access*, vol. 6, pp. 76187–76197, 2018.

[20] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions," *IEEE Access*, vol. 5, pp. 14347–14357, 2017.

[21] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, "Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data," *IEEE Trans. Ind. Electron.*, vol. 66, no. 9, pp. 7316–7325, Sep. 2019.

[22] C. Chen, F. Shen, J. Xu, and R. Yan, "Domain adaptation-based transfer learning for gear fault diagnosis under varying working conditions," *IEEE Trans. Instrum. Meas.*, vol. 70, no. 7, pp. 1–10, 2020.

[23] J. Singh, M. Azamfar, A. Ainapure, and J. Lee, "Deep learning-based cross-domain adaptation for gearbox fault diagnosis under variable speed conditions," *Meas. Sci. Technol.*, vol. 31, no. 5, May 2020, Art. no. 055601.

[24] S. Dong, K. He, and B. Tang, "The fault diagnosis method of rolling bearing under variable working conditions based on deep transfer learning," *J. Brazilian Soc. Mech. Sci. Eng.*, vol. 42, no. 11, pp. 585–601, Oct. 2020.

[25] 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.

[26] J. Zhu, N. Chen, and C. Shen, "A new multiple source domain adaptation fault diagnosis method between different rotating machines," *IEEE Trans. Ind. Informat.*, early access, Sep. 3, 2020, doi: 10.1109/TII.2020.3021406.

[27] Y. Zhu, F. Zhuang, and D. Wang, "Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, Jul. 2019, pp. 5989–5996.

[28] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4 inception-resnet and the impact of residual connections on learning," in *Proc. Int. Conf. Learn. Represent. (ICLR) Workshop*, Feb. 2016, p. 12.

[29] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2011.

[30] R. Xu, Z. Chen, W. Zuo, J. Yan, and L. Lin, "Deep cocktail network: Multi-source unsupervised domain adaptation with category shift," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3964–3973.

[31] M. Long, Y. Cao, M. I. Jordan, and J. Wang, "Learning transferable features with deep adaptation networks," in *Proc. ICML*, Jun. 2015, pp. 97–105.

[32] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," 2014, *arXiv:1412.3474*. [Online]. Available: https://arxiv.org/abs/1412.3474

RENXIANG CHEN was born in Sichuan, China, in 1983. He received the B.E. and Ph.D. degrees in mechanical engineering from Chongqing University, Chongqing, China, in 2007 and 2012, respectively. From 2015 to 2016, he was a Visiting Research Scholar with the University of Illinois at Chicago, IL, USA. He is currently a Professor and a Ph.D. Supervisor with Chongqing Jiaotong University, Chongqing. He has authored or coauthored more than 50 articles in his research career.

His main research interests include mechanical and electrical equipment security service, life prediction and health management, mechanical signal processing, intelligent manufacturing, machinery condition monitoring, and fault diagnosis.

TAO WEN was born in Yibin, Sichuan, China, in October 1999. He is currently pursuing the B.E. degree in mechanical Engineering with Chongqing Jiaotong University.

His current research interests include transfer learning, deep transfer learning, and bearing fault diagnosis. He received National Endeavor Scholarship in 2019.

LINLIN TANG was born in Chongqing, China, in October 1997. He received the B.E. degree in mechanical engineering from the City College of Science and Technology, Chongqing University, Chongqing, in 2018. He is currently pursuing the M.E. degree in mechanical engineering with the School of Mechanical and Electrical Engineering, Chongqing Jiaotong University. His research interests include machine learning, especially transfer learning, deep neural networks, and deep transfer learning, powered fault diagnosis, prognostics, and health management of mechanical and electrical systems.

· · ·