Automation, Control and Intelligent Systems
2022; 10(2): 17-26
http://www.sciencepublishinggroup.com/j/acis
doi: 10.11648/j.acis.20221002.12
ISSN: 2328-5583 (Print); ISSN: 2328-5591 (Online)

Classification Fusion for Bearing Fault Diagnosis with Multi-source Domain Shift

Huang Mu-sheng, Wu Song-song, Zheng Shi-yuan, Yu Xi, Zhuang Jia-yang, Jing Xiao-yuan

School of Computer Science, Guangdong University of Petrochemical Technology, Maoming, China

Email address:
ms_huang2021@126.com (Huang Mu-sheng), sswuai@126.com (Wu Song-song)

To cite this article:
Huang Mu-sheng, Wu Song-song, Zheng Shi-yuan, Yu Xi, Zhuang Jia-yang, Jing Xiao-yuan. Classification Fusion for Bearing Fault Diagnosis with Multi-source Domain Shift. *Automation, Control and Intelligent Systems*. Vol. 10, No. 2, 2022, pp. 17-26.
doi: 10.11648/j.acis.20221002.12

Received: March 20, 2022; **Accepted**: May 6, 2022; **Published**: May 12, 2022

Abstract: Bearings are one of the most widely used components of rotating machinery, whose failure can cause serious injuries and economic losses, therefore bearing fault diagnosis is an essential step in maintaining the safe and stable operation of industrial processes. Bearing fault diagnosis aims to detect the bearing fault condition and accurately classify it into a fault category based on sensing signals, such as vibration data. In practical applications, bearings always work in different types of equipment and under various working conditions, leading to performance degradation of diagnosis models due to the domain gap between the training data and the test data. Domain adaptation has been developed to address the domain shift problem in bearing fault diagnosis with demonstrated efficacy. Current domain adaptation models focus on the single-source scenario, by ignoring that sensing data may be collected from multiple sources in practical applications, and then be annotated for mode training. In this situation, it is non-trivial to use the single-source domain adaptation model to address the multi-source domain shift problem, because the domain gap exists among the source domains and the target domain. To solve this problem, we propose a novel bearing fault diagnosis model based on classification decision fusion to address the problem of multi-source domain conversion. Firstly, we train a source-aware fault diagnosis model in each source domain and then use it to predict the fault labels of the target samples. Second, a similarity score between each source domain and target domain is computed based on their feature distributions using local discriminant analysis and Maximum Mean Discrepancy. Finally, the similarity scores are used as domain weights in a proposed classification decision fusion strategy that uses a weighted linear combination process of predicted fault labels to provide the final predicted labels for the target samples. The benefits of the adaptive weighting fusion based on the classification result level, which makes full use of the available data from multiple source domains, measures the differences in distribution between the source and target domains and automatically adjust the weights to improve the diagnostic capability of the target domain. The effectiveness of the proposed method for diagnosing bearing faults under different operating and measurement conditions was verified using a bearing data set provided by Case Western Reserve University.

Keywords: Multi-source Domain Transfer, Bearing Fault Diagnosis, Classification Fusion, Knowledge Transfer, Cross-domain Fault Diagnosis

1. Introduction

Rolling bearings are essential to rotating machines, whose health condition has a significant impact on machine operation performance. As a kind of consumable component, bearings always work in varying environments so their conditions determine the sensing data in a complex manner [1]. Currently, it is still a challenging task to accurately detect and classify the bearing faults based on sensing data such as vibration signals.

The vibration signal contains the running state information of the rolling bearing, which is widely used to diagnose the health state of the bearing. The bearing fault diagnosis based on machine learning extracts the identification features reflecting the bearing fault from the vibration signal data in a data-driven way, and obtains the prediction label of the bearing fault type through the classifier. Fault diagnosis can be supported by learning algorithms such as shallow learning support vector machine (SVM) [2], k-nearest neighbor (KNN) [3], logical regression (LR) [4], and can also be implemented with the help of deep confidence network [5] and convolution network [5] in deep learning. This learning based fault diagnosis idea uses the supervised learning paradigm, trains the fault diagnosis model with the vibration signal data with the fault type label, and then sends the vibration signal to be identified to the trained model for fault type prediction.

Most of the existing fault diagnosis methods assume that the vibration signal data in the training set and the test set come from the same distribution. However, due to factors such as cross unit, variable working conditions, and multiple measuring points, the vibration signals in the source domain and the vibration signals in the target domain have different data distribution characteristics, which is difficult to be established in the actual application scenario. This distribution difference between training and test samples will significantly reduce the diagnostic accuracy of traditional methods [6].

To solve the above problems, domain adaptive learning is used to migrate the fault diagnosis knowledge of the source domain to the target domain to obtain good cross domain generalization performance. The work in this field includes the method of TrAdaBoost [7] Based on case-based learning, the method of SSTCA [8] Based on kernel function transfer component analysis, the method of SSTCA-SVM based on feature transfer [9], and the improved LSSVM method using recursive quantitative analysis [10]. The basic principle of these methods is to seek the consistent data distribution between the source domain and the target domain in the feature stage, reduce the interference of domain differences on fault feature identification, and thus enhance the robustness of fault diagnosis algorithms to changes in data domains with different distributions.

In addition to the single source domain migration task, there are also migration tasks from multiple source domains to target domains in practical applications, such as predicting the bearing health state represented by the vibration signal in the new equipment according to the bearing vibration signal of multiple existing types of equipment. The difficulty of this task is that the source domain and the target domain have different data distributions, and there are also data distribution differences among multiple source domains. If the domain adaptive method of a single source domain is simply extended, it will be difficult to obtain satisfactory diagnosis results [6]. Recently, literature [11] proposed a multi-source domain migration method based on the mean subspace of the Glassman manifold. Its main idea is to extract the common features of multiple source domains by feature fusion and migrate them to the target domain.

From the perspective of fault diagnosis result fusion, this paper designs a weighted fusion strategy based on domain distribution similarity and proposes a multi-source domain migration bearing fault diagnosis method for complex working conditions. The method is divided into two steps: firstly, the local Fisher discriminant analysis is used to obtain the discriminant feature subspace of each source domain, and in this feature space, the label samples are used to train the support vector machine classifier to realize the bearing fault classification of the specific source domain. Then, calculate the maximum mean discrepancy (MMD) [12] between the target domain and all source domains as a measure of the distribution difference between the source domain and the target domain, and convert it into the weight of the source domain diagnosis result in the final diagnosis result. This method deals with the problem of multi-source domain migration from the level of fault classification and reduces the risk of negative migration by considering the difference in domain distribution. The simulation results on CWRU bearing vibration database show that the proposed method can effectively improve the performance of bearing fault diagnosis under the condition of multi-source domain migration.

2. Related Work

In order to overcome the influence of data distribution differences on the generalization performance of the model, single source domain adaptation methods have been continuously proposed to improve the domain generalization ability of the fault diagnosis model. Pan et al. [8] proposed a kernel function-based transfer component analysis method SSTCA, which reduces the data distribution difference between the source domain and the target domain samples in the feature space. Shen et al. [14] applied the representative case-based learning method TrAdaBoost to improve the diagnostic performance of rolling bearing under different working conditions. After the feature-based transfer strategy, the combination of deep learning and transfer learning is used for mechanical fault diagnosis. On this basis, Yan et al. [13] proposed a multi-sensor data fault diagnosis algorithm based on the K-nearest neighbor classification method based on the temporal and spatial relationship between signals. Sun et al. [15] proposed a depth migration method based on a sparse self-encoder. Zhang et al. [16] proposed a neural network bearing fault diagnosis method based on model parameter migration.

In practical application, the labeled training samples may come from multiple source domains. In this case, the diagnostic knowledge of multiple source domains needs to be migrated to the target domain. R. Gopalan et al. [17] proposed a subspace based domain adaptation method and B. Gong et al. [18] proposed a domain adaptation method based on domain interpolation, which can deal with the migration of single source domain and multi-source domain at the same time. Zheng et al. [11] proposed using multiple source domains to construct the mean subspace on the Grassmann manifold and then using the manifold learning model to obtain the mean subspace to diagnose the target domain. Yang et al. [19] proposed a convolution neural network model for multi-source vibration signals to deal with the problem of off-condition fault diagnosis.

Compared with the existing work that performs multi-source domain fusion at the feature level, this paper proposes a new multi-source domain fusion strategy from the level of diagnosis results, that is, the weight is calculated according to the distribution similarity between each source domain and the target domain, and the weighted average is calculated. The diagnostic results of multiple source domain models on target domain samples are fused into the final fault type identification result.

3. Multi-source Domain Transfer Bearing Fault Diagnosis

3.1. Problem Formalization

Given a set of K source domain vibration dataset $\left\{x_i^k, y_i^k\right\}_{i=1}^{n_k}$, $k=1,\ldots,K$, $x_i^k \in \mathrm{R}^d$ is the i th vibration signal in the source domain K, $y_i^k \in \{1,2,\ldots,C\}$ is the fault type label for x_i^k, and C is the number of fault categories. Given a target domain vibration dataset $\left\{x_i^t\right\}_{i=1}^{n_t}$, where $x_i^{\mathrm{t}} \in \mathrm{R}^d$ is the vibration signal from a distribution that is not consistent with the source domain data.

The purpose of this paper is to accurately predict the health state $\hat{y}_i^t$ of the bearing associated with the target domain vibration sample $\left\{x_i^t\right\}_{i=1}^{n_t}$ by learning a migration model f from the source domain to the target domain using a set of K source domain labeled vibration signal samples $\left\{x_i^k, y_i^k\right\}_{i=1}^{n_k}$ and a set of target domain unlabelled vibration signal samples $\left\{x_i^t\right\}_{i=1}^{n_t}$.

3.2. Algorithm

The key to multi-source domain migration lies in how to effectively extract the discriminative information from each source domain and effectively transfer it to the target domain. Based on this, this paper proposes a multi-source domain bearing fault diagnosis algorithm as shown in Figure 1. The algorithm is divided into two main steps: (1) learning a source domain-dependent fault diagnosis model (2) fusion of diagnosis results based on domain similarity.

3.2.1. Supervised Learning Based Source Domain Fault Classification

As shown by Figure 1, we use local Fisher discriminant analysis for each source domain to obtain a feature subspace that portrays the fault identification information of that source domain. An SVM fault diagnosis model is then trained in this feature space using labeled samples to achieve bearing fault classification.

LFDA mines the class structure by maximizing the inter-class scatter and minimizing the intra-class scatter within the local neighborhood of the sample, and seeks a feature space that preserves the local discriminative properties of the training samples.

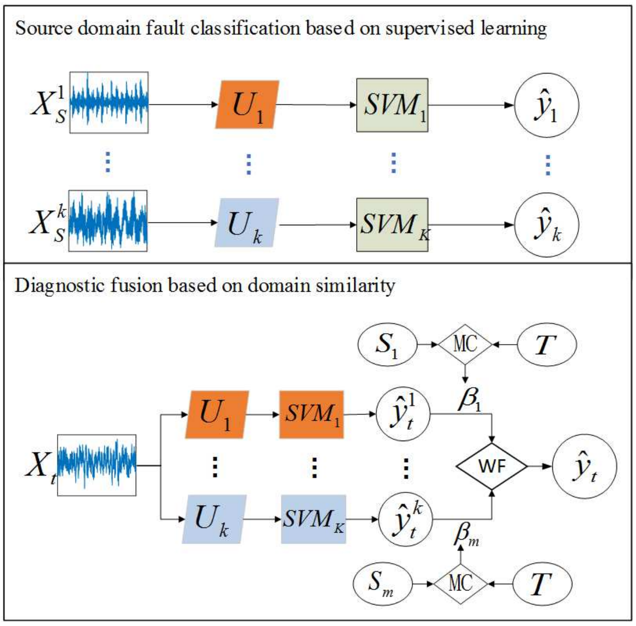


Figure 1. *Framework of the proposed diagnostic method.*

The optimal feature subspace U of LFDA is provided by the optimal solution to the following optimization problem

$$U^* \equiv \underset{U\in R^{D\times d}}{\arg} \max\ tr((U^T S^{(w)} U)^{-1} U^T S^b U) \tag{1}$$

where the intra-class discretization matrix S^w and the inter-class discretization matrix S^b are defined as

$$\begin{aligned} \hat{S}^w &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \hat{W}_{ij}^w (x_i - x_j)^T \\ \hat{S}^b &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \hat{W}_{ij}^b (x_i - x_j)^T \end{aligned} \tag{2}$$

where local intra-class weighting coefficients $\hat{W}_{ij}^w$ and local inter-class weighting coefficients $\hat{W}_{ij}^b$ are used to describe the local discriminative properties of the vibration data, defined as

$$\hat{W}_{i,j}^b = \begin{cases} A_{i,j} \times \left(\dfrac{1}{n} - \dfrac{1}{n_k}\right) & \text{if } x_i, x_j \in C_k \\ \dfrac{1}{n} & \text{else} \end{cases} \tag{3}$$

$$y_i^k \in \hat{W}_{i,j}^w = \begin{cases} \dfrac{A_{i,j}}{n_k} & \text{if } x_i, x_j \in C_k \\ 0 & \text{else} \end{cases} \tag{4}$$

where n_k is the number of kth class of training samples and A_{ij} denotes the degree of association between samples, defined as

$$A_{i,j} = e^{\frac{\|x_i - x_j\|^2}{\sigma_i \sigma_j}} \tag{5}$$

The $\sigma_i = \| x_i - x_{i,k} \|$ in Equation (5) is the distance between x_i and its k th local nearest neighbour sample. LFDA to obtain K optimal locally discriminated subspaces $U_1, \ldots, U_k$ as the feature subspace for the corresponding K source domains, the vibration signal samples from each of the K source domains will be projected into the corresponding feature subspace to obtain the corresponding source domain features, i.e.

$$z = U^T x \tag{6}$$

In each source domain, we learned a support vector machine (SVM) as a classifier based on the vibration features. To achieve a non-linear classification, we first transform the vibration features into a renewable Hilbert space (RKHS) and then learn a linear support vector machine to obtain the following kernel function-based SVM classification model.

$$f(x) = \mathrm{sgn}\left(\sum_{m=1}^{l} a_m^* y_i K(z, z_m) + b^* \right) \tag{7}$$

where $K(xx_i)$ is the kernel function, and the RBF kernel function is used in this experiment. a_m^*, b^* in Equation (7) are obtained by solving the following optimal problem

$$\begin{aligned} &\max_a \sum_{m=1}^{i} a_m^* - \frac{1}{2} \sum_{m=1}^{i} \sum_{j=1}^{i} a_m^* a_j^* y_m y_j K(z_m, z_j) \\ &s.t. \sum_{m=1}^{i} a_m^* y_m = 0 \text{ , } a_m^* \geq 0 \text{ , m=1,2,...,i} \end{aligned} \tag{8}$$

$$b^* = \frac{1}{|S|} \sum_{s \in S} \left(\frac{1}{y_s} - \sum_{m \in S} a_m^* y_m K(z, z_m) \right) \tag{9}$$

For the K source domain vibration datasets, we can train K LFDA feature subspaces $U_1, \ldots, U_K$ and the corresponding K SVM classifiers $\mathrm{SVM}_1, \mathrm{SVM}_2, \ldots, \mathrm{SVM}_k$. Since these K classification models will give K fault classification results for the target domain vibration signal x_i^t, we need to predict the fault type $\hat{y}_i^t$ of the target vibration signal based on these K fault diagnosis results.

3.2.2. Diagnostic Fusion Based on Domain Similarity

Due to the K source domains and the target domain are inconsistent in data distribution and have a certain correlation, the classification model of each source domain can predict the bearing fault type of the target domain vibration signal to a certain extent.

In this paper, we use the idea of integrated learning to achieve fault type identification for multi-source domain migration by fusing the identification results of the target vibration signals derived from multiple fault diagnosis models at the resulting level. If the source domain is more similar to the target domain distribution, the source domain classification model is more suitable for the target domain fault diagnosis task.

We utilize MMD to measure the distribution similarity between the target domain and the source domain,, defined as follows. $\mathcal{F}$ is a given set of functions, X the distribution of the source domain is, the distribution of the target domain is Y then we have

$$\begin{aligned} d_{st}^k &= \mathrm{MMD}[\mathcal{F}, X, Y] \\ &= \sup_{f \in \mathcal{F}} \left(\frac{1}{n_k} \sum_{i=1}^{n_k} f\left(x_i^k\right) - \frac{1}{n_t} \sum_{i=1}^{n_t} f\left(x_i^t\right) \right) \end{aligned} \tag{10}$$

Based on the similarity of the domain data distribution, we define the importance weight of the source domain for the target domain as

$$\beta_k = \frac{e^{\frac{-d_{st}^k}{\rho}}}{\sum_{i=1}^{N} e^{\frac{-d_{st}^k}{\rho}}}, \qquad k = 1, 2, \ldots, K \tag{11}$$

The parameter ρ is used to adjust the sensitivity of the weight β to the MMD measure, and the function of the denominator is to make the β value between [0, 1], and the sum is 1. It can be seen from formula (11) that if the MMD value of a source domain and the target domain is smaller, it means that the data distribution of the source domain is closer to the target domain distribution, and the learned classification model may have a relatively better diagnosis effect. The weight of the classification model obtained from the source domain in the fusion should be larger.

Let $\widehat{y_t^1}, \ldots, \widehat{y_t^k}$ be the fault type prediction label obtained by the target domain vibration signal x^t through K source domain classification model, then the final fault type of x^t is given by equation (12)

$$\widehat{y_t} = \max_{r \in \{1,2,\ldots,C\}} \sum_{\widehat{y_t^k} = r} \beta_k \widehat{y_t^k} \tag{12}$$

The method proposed in this paper is described in detail as shown in Algorithm.

Algorithm: *Multi-Source Domain Shift-based Algorithm for Fusion of Classification Results.*

Input: K source domain vibration Datasets $\{x_i^k, y_i^k\}_{i=1}^{n_k}, k=1,\dots,K$, target domain sample set $\{x_i^t\}_{i=1}^{n_t}$.

Output: Target domain sample set of predicted labels $\{\widehat{y_i^t}\}_{i=1}^{n_t}$

Start:

For $k=1:K$:

Learning the $\{x_i^k, y_i^k\}_{i=1}^{n_k}$ locally optimal discriminative feature space of the kth source domain based on $\{x_i^k, y_i^k\}_{i=1}^{n_k}$ using equation (1) U_k ;

Using equation (7) to train the support vector machine SVM_k in U_k ;

End

For $i=1:n$:

For $k=1:K$:

Use equation (6) to obtain the features of the target domain sample $\{x_i^t\}_{i=1}^{n_t}$ in the kth source domain feature space;

Using equation (7) to obtain the fault type prediction labels $\widehat{y_k^i}$ for the target domain in the k source domains;

End

For $h=1:K$:

Use equation (10) to calculate the maximum mean discrepancy d_{st}^h between the h th source domain $\{x_i^h\}_{i=1}^{n_h}$ and the target domain $\{x_i^h\}_{i=1}^{n_h}$;

The input d_{st}^h is calculated by equation (11) to obtain the weight β_h of the h th SVM diagnostic model;

End

The obtained weights $\beta_1, \beta_2, \dots, \beta_K$, $\widehat{y_1^i}, \widehat{y_2^i}, \dots, \widehat{y_K^i}$ into equation (12) to obtain the result for the current sample $\widehat{y_i^t}$;

End

4. Experiment

4.1. Database

In this paper, a multi-source domain migration bearing fault diagnosis experiment based on the Case Western Reserve University rolling bearing database CWRU [20] is conducted to test the effectiveness of the proposed method on cross-domain bearing fault diagnosis.

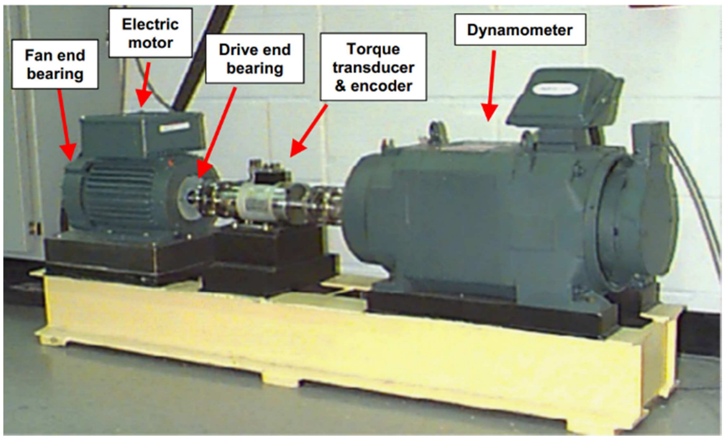


Figure 2. *Bearing test bench and its working schematic.*

As shown in Figure 2 above, the test stand consists of a 2 hp motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown). The test bearings support the motor shaft. Single point faults were introduced to the test bearings using electro-discharge machining. The test bearing was used to support the motor shaft and three single points of fault were introduced into the test bearing using EDM, namely outer ring fault, inner ring fault and ball fault. Vibration data was collected using accelerometers, which were placed at twelve o'clock on the drive and fan ends of the motor housing.

The sampling frequency at the drive end was 12KHz and the fan end used a frequency of 48KHz. vibration signals were collected at four motor load conditions of 0, 1, 2, and 3 hp.

Vibration data was collected using accelerometers, which were attached to the housing with magnetic bases. Accelerometers were placed at the 12 o'clock position at both the drive end and the fan end of the motor housing.

As shown in Table 1, according to the vibration signal sampling point and motor load, the vibration data are divided into eight data fields, which are noted as A, B, C, D, E, F, G, and H. Among them, A, B, C, and D come from the data of the drive end, and E, F, G, and H come from the data of the fan end.

Table 1. *Bearing parameters for the eight vibration data fields.*

Source Domain	bearing specifications	Operating environment	Location of Accelerometer
A/B/C/D	SKF 6205-2RS	0/1/2/3 hp	drive end
E/F/G/H	SKF 6203-2RS	0/1/2/3hp	fan end

Table 2. *Multi-source domain migration troubleshooting tasks.*

Task	domain	target	Task	domain	target
MT-A	E/F/G/H	A	MT-E	A/B/C/D	E
MT-B	E/F/G/H	B	MT-F	A/B/C/D	F
MT-C	E/F/G/H	C	MT-G	A/B/C/D	G
MT-D	E/F/G/H	D	MT-H	A/B/C/D	H

By selecting one of the eight data domains as the target domain and the remaining data domains at different ends as the source domains, we constructed eight multi-source domain migration fault diagnosis tasks, and the specific information of the data domains and fault diagnosis tasks are shown in Table 2. We use the data processing method of the literature [11], where each vibration data sample is a 4000-dimensional vector from four bearing condition categories, i.e., Normal, outer ring fault, inner ring fault, and ball fault. The extraction process of this first-level feature is referred to in the literature [11], and the algorithm in this paper takes this first-level feature as input and obtains the fault condition prediction label of the target bearing after a multi-source domain adaptive model.

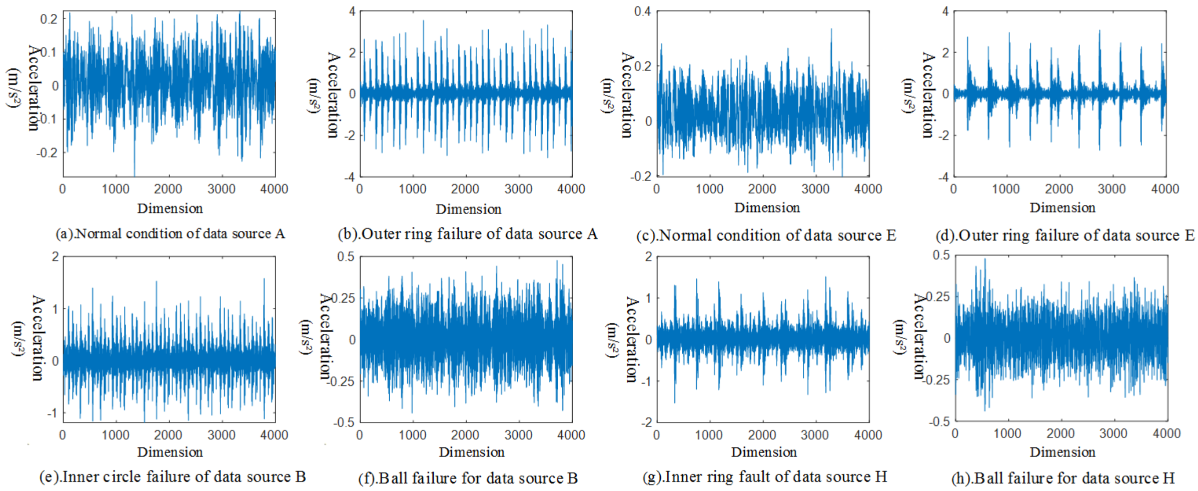


Figure 3. *Example of the vibration signal used in this paper.*

The number of samples for each bearing condition is 270, totaling 1080 vibration data samples, thus forming the experimental setup I. To test the fault diagnosis performance under the low-dimensional vibration data samples, we randomly intercepted the original vibration data with continuous 1200-dimensional signals and repeated 5 times to calculate the average accuracy and standard deviation, thus forming the experimental setup II. Figure 3 shows some of the vibration signal data samples of the experimental setup. For a fair comparison, we extracted 23-dimensional dimensionless features and 10-dimensional features about the vibration spectrum and envelope spectrum from the D-dimensional (D=4000 in the setup I and D=1200 in setup II) original vibration signals according to the data processing scheme of the literature [11], and combined them into 33-dimensional features as the input data of the algorithm.

4.2. Comparison Methods

To verify the effectiveness of the proposed method in this paper, we conduct comparative fault diagnosis experiments under setting I and setting II. The reference methods include.

1) SVM (Support Vector Machine) [2]: all source domain samples are used in this experiment to train the SVM model based on Gaussian kernel function, and then predict the state labels of the target domain samples.
2) KNN (k-nearest neighbor classifier) [3]: the state labels of the target domain samples are predicted based on the nearest neighbors of the target domain samples among all source domain samples, where the similarity is based on the Euclidean distance and the optimal k value is used with the setting of the literature [11].
3) LR (logistic regression): the potential relationship between the vibration signal and the bearing state is fitted using all labeled source domain samples.
4) SSTCA (semi-supervised migration principal component analysis) [8]: the data are mapped to a high-dimensional regenerative kernel Hilbert space, and then the distribution distance between the source and target domains is reduced with the help of kernel functions to reduce the distribution differences in the data domains. In this paper, SSTCA aggregates all source domain samples into a single source domain for the prediction of the target domain.
5) TrAdaBoost [14]: the weak classifier set is optimized iteratively to become a strong classifier that is robust to domain migration. In this paper, TrAdaBoost aggregates all source domains into a single source domain for single

source domain migration.

6) SGF [17]: solves the linear feature subspace of the source and target domains and performs feature space interpolation to achieve asymptotic alignment of the distribution from the source domain to the target domain.
7) GFK [18]: using kernel functions on Grassmannian manifolds to achieve infinite subspace interpolation from the source domain eigensubspace to the target domain eigensubspace. In this paper, GFK solves the multi-source domain migration problem by selecting the source domain with the most similar distribution to the target domain from multiple source domains with a single source domain migration method.
8) MSDGIF [11]: the local discriminative subspaces in each source domain are modeled as Grassmannian manifolds, and the multi-source domain migration problem is transformed into a single-source domain migration problem by finding the optimal average subspace to obtain a uniform portrayal of multiple source domains.

4.3. Comparison Methods

The multi-source domain migration fault diagnosis results for the experimental setup I am shown in Table 3. We observe that traditional classification models that do not consider domain migration, such as SVM, KNN, and LR have difficulty in maintaining the validity of the classification models on the target domain data because they do not consider the distribution differences between the source and target domains.

The accuracy of single-source domain migration methods, such as TrAdaBoost (combine) and SSTCA (combine), is below 70% when dealing with large spans migration tasks, such as MT-E, MT-F, MT-G, and MT-H, because these two methods ignore the distribution differences between source domains and treat all source domain training samples as the same distribution. It is worth noting that among the traditional classification models, KNN diagnostic performance outperforms other non-migration methods and some domain migration methods, probably because KNN relies on the local structure of the data for classification, which is less affected by the overall distribution changes of the data.

SGF does not consider labeling information in the subspace acquisition and interpolation process, which makes it difficult to retain diagnostic discrimination information during the transition from source to target domains. GFK selects the optimal source domain from multiple source domains to achieve domain migration, which inevitably loses diagnostic discrimination information from other source domains.

When the source and target domains have large differences in distribution, such as MT -F and MT-G, the strategy of using only a single source domain makes the fault diagnosis accuracy of GFK low. MSDGIFI considers the diagnostic discriminative information of each source domain and transforms the multi-source domain migration into a single-source domain migration problem using feature fusion. Therefore, MSDGIFI achieves higher diagnostic accuracy than the above methods. As for the method proposed in this paper, Table 3 shows that the highest accuracy is achieved in 7 out of 8 multi-source domain migration diagnosis tasks, while the average diagnosis accuracy is improved by 4.76% compared with MSDGIFI.

This is attributed to the fact that the adopted automatic source domain weight determination strategy can effectively use the diagnostic knowledge from multiple source domains and migrate it to the target domain based on the similarity of the distribution between the target and source domains.

The results of multi-source migration fault diagnosis under Experimental Setup II are shown in Table 4, and we observe that the diagnosis results of each method in Table 4 have similar characteristics and trends as those in Table 3.

Table 3. *Results of the multi-source domain migration troubleshooting task under experimental setup I.*

method	MT-A	MT-B	MT-C	MT-D	MT-E	MT-F	MT-G	MT-H	Accuracy
SVM	83.15	82.22	78.70	82.13	57.41	52.22	51.11	60.28	68.40
KNN	78.94	82.41	79.26	82.41	79.44	64.91	74.63	77.04	77.38
LR	83.33	82.04	76.57	82.13	58.06	50.09	50.46	50.56	66.66
SSTCA	86.57	83.06	80.19	83.33	69.63	44.44	48.80	50.28	68.29
TrAdaboost	76.94	79.72	77.13	74.44	52.96	50.00	50.28	50.74	64.03
SGF	69.35	69.91	71.11	67.31	65.65	55.09	56.48	55.09	63.75
GFK	85.19	79.44	71.76	72.13	65.37	36.48	42.50	58.43	63.91
MSDGIFI	83.33	83.33	83.33	83.33	89.81	81.57	77.41	73.89	82.00
Ours	89.17	85.65	88.33	85.93	84.81	86.02	86.48	87.69	86.76

Table 4. *Results of the multi-source domain migration troubleshooting task under experimental setup II.*

method	MT-A	MT-B	MT-C	MT-D	MT-E	MT-F	MT-G	MT-H	Accuracy
SVM	73.73±1.28	74.96±1.25	72.45±1.21	70.62±0.47	62.71±1.08	56.20±1.32	56.37±0.66	60.26±0.50	65.92±7.49
KNN	73.98±1.47	77.47±0.77	73.98±1.13	72.54±0.43	79.75±1.21	59.65±1.35	68.75±1.12	68.37±2.42	71.81±6.05
LR	81.44±0.50	81.44±1.56	80.76±2.47	77.52±2.11	65.94±1.18	60.66±3.91	57.69±3.17	56.09±0.52	55.73±10.8
SSTCA	79.62±1.18	81.64±1.47	75.24±4.53	80.94±1.41	68.25±1.72	46.79±6.50	51.45±1.21	53.01±1.91	66.56±13.8
TrAda*	81.79±0.09	82.42±0.01	81.90±2.38	73.59±1.77	64.30±3.26	54.26±3.76	50.47±5.14	55.11±2.23	67.98±13.3
SGF	71.53±2.04	72.05±2.11	70.88±1.78	67.57±1.19	67.48±2.19	54.01±6.36	56.53±1.01	57.25±1.37	64.66±7.57
GFK	81.03±3.56	75.40±2.64	72.58±2.85	68.37±2.74	76.75±10.7	47.50±9.44	49.32±7.49	54.87±6.31	65.73±14.0
MSDGIFI	82.15±1.09	82.81±1.03	81.87±1.24	81.39±1.86	85.76±2.12	77.12±5.31	74.09±4.54	77.05±4.89	80.28±4.86
Ours	86.14±2.10	87.13±2.53	85.92±1.88	83.89±1.88	83.85±2.67	79.66±2.34	76.66±1.46	79.72±3.64	82.87±4.15

*TrAda stands for TrAdaBoost.

Specifically, the multi-source migration method is more accurate and stable than the non-migration and single-source migration methods in terms of diagnostic accuracy and diagnostic stability.

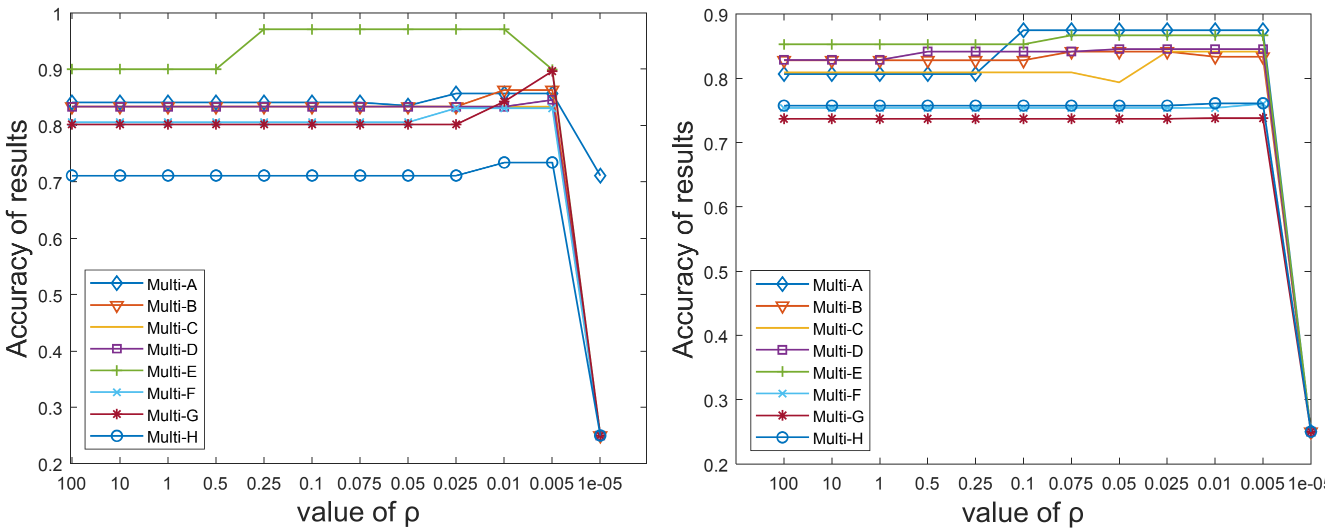
It is worth noting that the proposed method in this paper still achieves the highest diagnostic accuracy in most tasks under the experimental setup II, and its average diagnostic accuracy improves by 2.59% compared with MSDGIFI, indicating that this method has stronger fault diagnosis capability and diagnostic stability than MSDGIFI.

The results in Table 1 and Table 2 show that it is difficult to achieve effective migration from multiple source domains to the target domain by simply combining all source domains into one source domain or by directly averaging the diagnosis results of multiple source domains. Meanwhile, comparing the method in this paper and MSDGIFI, it can be found that diagnostic decision-level fusion and feature-level fusion are comparable or even better in multi-source domain migration, which proves that multi-source domain migration is based on diagnostic result fusion has a greater potential for handling complex scenario bearing fault diagnosis.

4.4. Analysis of Algorithm Parameters

The value of the parameter ρ in Equation (11) of the method in this paper has a direct impact on the weight of the fusion of multi-source domain results. In order to study the dependence of the method in this paper on the parameter ρ, we set different values of ρ under experimental settings I and II and observe the changes of the corresponding diagnostic accuracy, and the results are shown in Figure 4(a) and Figure 4(b), respectively.

As can be seen from Figure 4, the method achieves high diagnostic accuracy and stable recognition performance over a wide range of values of ρ. In particular, when the value is between 0.1 and 0.05, the method achieves the highest diagnostic accuracy in several tasks.



(a) Experimental setup I (b) Experimental setup II

Figure 4. *Effect of different ρ values on the diagnostic accuracy under two settings.*

When the value of ρ is less than 0.005, the recognition effect of this method drops abruptly because the value of ρ is too small to make the value of the negative exponential function in Eq. (11) tend to 0. The calculation of Eq. (11) will face numerical problems, making the calculated source domain weight values appear data anomalies and leading to the fault of the fusion strategy. The above results also prove the rationality of the fusion strategy of diagnostic results proposed in this paper.

4.5. Diagnostic Fusion Mechanism Validation

To verify that the fusion of diagnostic results strategy proposed in this paper can extract the relative importance of each source domain to the target domain fault diagnosis, we use the MMD metric to examine the similarity of each source domain to the target domain and its corresponding single source domain migration fault diagnosis accuracy. At the same time, we calculate the corresponding fusion weights of each source domain and the corresponding fault diagnosis accuracy of our fusion strategy. The experimental results are shown in Tables 5 and 6.

The results in Table 5 show that the diagnostic accuracy of single source domain migration is approximately proportional to the MMD values of the source and target domains, demonstrating that the closer the source domain is to the target domain, the better the migration effect is. It is worth noting that according to the fusion strategy of diagnostic results in this paper, the weights taken for each source domain are also proportional to the corresponding MMD values, indicating that the method in this paper does automatically mine the relative importance of the source domain for the diagnostic task of the target domain.

The results in Table 5 also show that the multi-source domain fusion approach in this paper achieves higher diagnostic accuracy than all single-source domain migrations, demonstrating that the approach in this paper is indeed effective in migrating diagnostic knowledge from multiple source domains to the target domain. Similar conclusions can be drawn from the results in Table 6.

Table 5. *MMD of MT-B with ρ=0.010 and its accuracy relationship in the case of experimental setup I.*

Setting	E→B	F→B	G→B	H→B	MT-B
MMD	0.1550	0.1544	0.1547	0.2688	/
weight	0.2680	0.3513	0.3806	0.0000	/
accuracy	81.1111	81.6667	84.1667	74.6296	85.6557

Table 6. *MMD and its accuracy relationship for MT-B with ρ=100 in case of experimental setup I.*

setting	E→B	F→B	G→B	H→B	MT-B
MMD	0.1550	0.1544	0.1547	0.2688	/
weight	0.2544	0.2545	0.2545	0.2366	/
accuracy	81.1111	81.6667	84.1667	74.6296	81.6667

The results in Table 5 also show that the multi-source domain fusion method in this paper achieves higher diagnostic accuracy than all single-source domain migration, proving that the method in this paper is indeed effective in migrating diagnostic knowledge from multiple source domains to the target domain. Similar conclusions can be drawn from the results in Table 6.

Comparing Table 5 and Table 6, it can be found that when the value of ρ is taken too large, although the numerical calculation problem is avoided, the important difference of each source domain to the target domain is ignored. At this time, the adaptive diagnosis result fusion strategy proposed in this paper degenerates to a direct average of the diagnosis results of each source domain.

Table 7. *MMD and its accuracy relationship for MT-F with ρ=0.01 in the case of experimental setup II.*

setting	A→F	B→F	C→F	D→F	MT-F
MMD	0.4966	0.5112	0.4720	0.5028	/
weight	0.2329	0.1742	0.3990	0.1938	/
accuracy	80.5556	78.7037	81.0185	80.3704	82.3154

Table 8. *MMD of MT-F with ρ = 100 and its accuracy relationship in the experimental setup II case.*

setting	A→F	B→F	C→F	D→F	MT-F
MMD	0.4966	0.5112	0.4720	0.5028	/
weight	0.2499	0.2500	0.2500	0.2500	/
accuracy	80.5556	78.7037	81.0185	80.3704	79.722

The experimental results under Experimental Setup II are shown in Tables 7 and 8, from which we observe that (1) determining the weights of each source domain in the fusion process of diagnostic results based on the distribution similarity between the source and target domains is beneficial for multi-source migration, and the proposed method in this paper outperforms the best performance of single-source migration; (2) taking too large a value of the parameter ρ will degrade the adaptive fusion strategy in this paper to direct averaging and weaken the multi-source migration capability of the method in this paper.

5. Conclusion

In conclusion, the main contributions of this paper can be summarized in two terms. 1) This paper proposes an adaptive multi-source domain fault diagnosis method based on the fusion of classification results at the level of classification, which effectively utilizes multiple source domains and improves the correct rate of fault diagnosis; and 2) In order to reduce the influence of negative migration on the results of multi-source domain fault diagnosis we introduce a model determined by similarity score and sensitivity parameters Weights. By using local discriminant analysis and Maximum Mean Discrepancy to calculate the similarity score between the source domain and the target domain, the similarity score is used to calculate the weight of the source domain model in multi-source domain fault diagnosis, which reduces the risk of negative migration in the multi-source domain context; in addition, the selection of sensitivity parameters for the method is investigated in depth and an effective interval on how to set the parameters is proposed. Finally, the validity of the proposed method is verified by a bearing data set provided by Case Western Reserve University for bearing fault diagnosis. We will further explore the fusion of feature levels through neural networks using similarity scores in the future work.

Acknowledgements

This paper is partially supported by the Key Laboratory of Data Science and Intelligence Application, Minnan Normal University under grant No. D202002 and the Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds) under grant No. pdjh2021b0338.

References

[1] ZHENG Jinde, WANG Xinglong, PAN Haiyang, TONG Jinyu, LIU Qingyun. Rolling Bearing Fault Diagnosis Method Based on Adaptive Autogram. China Mechanical Engineering, 2020, 39 (18): 118-124+131.

[2] ZHENG Jinde, PAN Haiyang, CHENG Junsheng. Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines. Mechanical Systems and Signal Processing. 2016.09.010:85.10.1016/j.ymssp.

[3] WANG Dong. K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: Revisited. Mechanical Systems and Signal Processing. 2015.10.007:70.10.1016/j.ymssp.

[4] Zhou JM, Lai HW, Zhang L, Li P. Evaluation of bearing performance degradation based on EMD and logistic regression. Mechanical Design and Research, 2016, 32 (05): 72-75+79.

[5] Jinde, Zheng, Pan, Haiyang, Junsheng, Cheng. (2017). Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines. Mechanical Systems and Signal Processing. 85. 10.1016/j.ymssp.2016.09.010.

[6] W. Li, F. Li, Y. Luo, P. Wang, and J. Sun, "Deep Domain Adaptive Object Detection: a Survey," 2020 IEEE Symposium Series on Computational Intelligence (SSCI), 2020, pp. 1808-1813, DOI: 10.1109/SSCI47803.2020.9308604.

[7] Yao, Yi, Doretto, Gianfranco. (2010). Boosting for transfer learning with multiple sources. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1855-1862. 10.1109/CVPR.2010.5539857.

[8] Sinno Jialin Pan, Ivor W TSANG, James T KWOK, Qiang YANG. Domain adaptation via transfer component analysis. IEEE Trans Neural Netw. 2011; 22 (2): 199-210.

[9] Kang, S., Hu, M., Wang, Y., Xie, J., Mikulovich, V. I. (2019). Fault Diagnosis Method of a Rolling Bearing Under Variable Working Conditions Based on Feature Transfer Learning. Zhongguo Dianji Gongcheng Xuebao/Proceedings of the Chinese Society of Electrical Engineering. 39. 764-772. 10.13334/j.0258-8013.pcsee.180130.

[10] Yunguang Gao, Shunbo Liu, Feng Li, Zhiguo Liu. Fault detection and diagnosis method for cooling dehumidifier based on LS-SVM NARX model. International Journal of Refrigeration, 2016, 61.

[11] ZHENG Huailiang, WANG Rixin, YANG Yuantao, LI Yuqing, XU Minqiang. Intelligent Fault Identification Based on Multisource Domain Generalization Towards Actual Diagnosis Scenario. IEEE Transactions on Industrial Electronics, 2020 (67): 1293-1304.

[12] A Kernel Two-Sample Test. Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, Alexander Smola; 13 (25): 723−773, 2012.

[13] YAN Xunshi, SUN Zhe, ZHAO Jingjing, SHI Zhengang, ZHANG Chen-an. Fault diagnosis of rotating machinery equipped with multiple sensors using space-time fragments. Journal of Sound and Vibration, 2019, 456: 49-64.

[14] Shen Fei, CHEN Chao, Ruqiang YAN, Robert X. GAO. Bearing fault diagnosis based on SVD feature extraction and transfer learning classification. 2015 Prognostics and System Health Management Conference (PHM), 2015: 1-6.

[15] CHUANG Sun, MA Meng, ZHAO Zhibin, TIAN Shaohua, YAN Ruqiang, CHEN Xuefeng. Deep Transfer Learning Based on Sparse Autoencoder for Remaining Useful Life Prediction of Tool in Manufacturing. IEEE Transactions on Industrial Informatics, 2019 (15): 2416-2425.

[16] ZHANG Ran, TAO Hongyang, WU Lifeng, GUAN Yong R. Transfer Learning With Neural Networks for Bearing Fault Diagnosis in Changing Working Conditions. IEEE Access, vol. 2017: 14347-14357.

[17] Raghuraman GOPALAN; Ruonan LI; Rama CHELLAPPA. Domain adaptation for object recognition: An unsupervised approach. 2011 International Conference on Computer Vision, 2011: 999-1006.

[18] Boqing GONG, Yuan SHI, Fei SHA, Kristen GRAUMAN. Geodesic flow kernel for unsupervised domain adaptation. 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012: 2066-2073.

[19] YANG, Hong-bai, ZHANG, Jiang-an, CHEN, Lei-lei, ZHANG, Hong-li, LIU, Shu-lin. Fault Diagnosis of Reciprocating Compressor Based on Convolutional Neural Networks with Multisource Raw Vibration Signals. Mathematical Problems in Engineering, 2019.

[20] SMITH, Wade, ROBERT B. Randall. Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study. Mechanical Systems and Signal Processing. 2015: 64-65.