An Intelligent Fault Diagnosis Method for Aviation Precision Bearings based on Transfer Learning

Xiaoyue Peng, Lu Dong*, Zehua Hua, Tingyu Zhu

Civil Aviation Flight University of China, Guanghan 618300, Sichuan, China

*Corresponding author

Abstract

To meet the requirements of efficient, accurate, and cross-operating-condition fault diagnosis for aviation bearings, this paper proposes an intelligent fault diagnosis method based on transfer learning. Addressing the data difference between the source domain and target domain, preprocessing methods such as resampling (unifying sampling frequency), filtering denoising, and signal segmentation (improving sample set stability) are adopted to unify data characteristics, laying a foundation for subsequent transfer learning. Class weights are introduced to solve the sample imbalance problem. After multiple model parameter optimizations and comparisons of the recognition effects of various classifiers, it is concluded that Random Forest achieves the best recognition performance. To tackle the decline in diagnostic performance caused by inter-domain distribution differences, a three-layer transfer learning framework featuring feature-model-sample collaboration is proposed. To enhance the credibility and interpretability of the diagnostic process, a full-process interpretability analysis system covering the three-layer transfer learning strategy is constructed, and a visual analysis flow chart showing the joint completion of transfer by the three-layer transfer learning strategy and the original model is presented, thereby improving the transparency and acceptability of diagnostic results. Relevant experiments demonstrate that the proposed intelligent analysis method exhibits excellent performance in practicality, robustness, and interpretability. The research results of this paper can provide important reference for studies in related fields.

Keywords

Aviation Bearings; Vibration Signals; Intelligent Fault Diagnosis; Transfer Learning.

---

1. Introduction

Aviation precision bearings are key components in aircraft, significantly affecting the service life and flight safety of aircraft [1-2]. However, aviation bearings operate under harsh conditions such as high speed, high temperature, and strong vibration for a long time, resulting in a high failure rate [3-4]. Traditional early fault diagnosis methods rely on maintenance personnel to conduct manual inspections with handheld measuring equipment, which not only have low accuracy but also tend to cause unnecessary damage to bearing components. Therefore, intelligent, automated, and non-destructive fault diagnosis for aviation precision bearings is extremely necessary [5].

The intelligence of aviation bearing fault diagnosis is of great significance for flight safety and operation and maintenance efficiency [6-7]. It enables early accurate warning of faults, transforming traditional scheduled maintenance into predictive maintenance and effectively avoiding catastrophic accidents. Through automated diagnosis, reliance on expert experience can be greatly reduced, detection efficiency and consistency can be improved, and operation and maintenance costs as well as downtime losses can be significantly decreased [8]. Moreover, it can promote the upgrading of condition monitoring technology from "perception" to "cognition", serving as a key technical support for ensuring high reliability of aviation equipment and realizing intelligent aviation [9].

In practical applications, limited by complex noise environments and variable operating conditions, the original signals collected by sensors are often mixed with interference components, which weaken the significance of fault features and further lead to a decline in model performance. In contrast, bearing data collected under test bench conditions not only have abundant samples and clear labels but also share certain similarities in fault mechanisms with actual operating scenarios [10]. Therefore, how to transfer the knowledge learned from test bench data to actual scenarios has become a crucial issue to be solved urgently. Based on the basic theory of transfer learning, this paper constructs a diagnostic model in the source domain by combining fault mechanism analysis and multi-dimensional feature extraction of aviation bearings, and realizes knowledge transfer with the help of transfer learning technology, which can effectively improve the fault diagnosis performance in the target domain.

2. Analysis Process and Related Principles

2.1 Construction of Multi-Dimensional Features

To maximize the accuracy of intelligent diagnosis of aviation bearings, more than 40 dimensions of features are selected in this paper, covering time domain, frequency domain, and time-frequency domain. These features not only consider physical interpretability but also are suitable for deep learning models. Specifically including:

(1) Time-domain features: Mean value, standard deviation, root mean square (RMS), peak-to-peak value, skewness, kurtosis, etc. The mean value reflects the average level of the signal; the peak-to-peak value can reflect the maximum variation range of the signal; skewness and kurtosis describe the asymmetry and sharpness of the signal distribution respectively, and are sensitive to impact signals.

(2) Frequency-domain features: Spectral centroid, spectral spread, spectral rolloff, spectral flux, etc. The spectral centroid represents the center of gravity of the spectrum and reflects the dominant frequency of the signal; the spectral spread describes the degree of dispersion of the spectrum.

(3) Time-frequency domain features: Extracted through envelope analysis, variational mode decomposition (VMD), and wavelet transform. Envelope analysis can highlight the impact components of the signal, and the kurtosis and skewness of the envelope are sensitive to faults; wavelet transform analyzes the signal at multiple scales and extracts energy and standard deviation at different scales, thereby obtaining both time-domain and frequency-domain information.

2.2 Visualization of Feature Data

After extracting feature values from the collected data, it is necessary to visualize the feature information of fault signals to facilitate understanding, human-computer interaction, and daily maintenance work. As shown in Fig. 1, this paper compares the feature distribution map sets of normal bearing data and fault bearing data, showing the differences between the two types of data in different features from multiple dimensions (time domain, frequency domain, time-frequency domain, etc.), which can assist in understanding the principle of bearing fault diagnosis. Among them, red represents fault bearing signals and green represents normal bearing signals. This paper compares the differences of 12 types of typical feature values mentioned above between normal signals and fault signals. It can be seen from Fig. 1 that the discriminative accuracy represented by different feature value information varies:

(1) High-discriminative features: Represent feature values with extremely strong discriminative power, the highest accuracy, and the most obvious visualization information. For example: In fault signals, the values of kurtosis, crest factor, impulse factor, and envelope kurtosis are significantly increased; the values of impact property, peak ratio, impulse characteristic, and impact energy ratio are much larger than the corresponding values in normal signals.

(2) Medium-discriminative features: The distribution of these feature values in the two types of data overlaps partially, but their mean values are significantly different, which can assist in discrimination. For example: In fault signals, the RMS is more inclined to high frequencies; the spectral spread is larger; the frequency distribution range is wider; the basic vibration is abnormal; the cyclostationarity is more obvious, including periodic impacts.

(3) Low-discriminative/scenario-dependent features: The distribution of these feature values overlaps greatly between normal signals and fault signals, meaning they cannot be used for discrimination alone and need to be combined with other feature value information for discrimination, such as first-order amplitude.

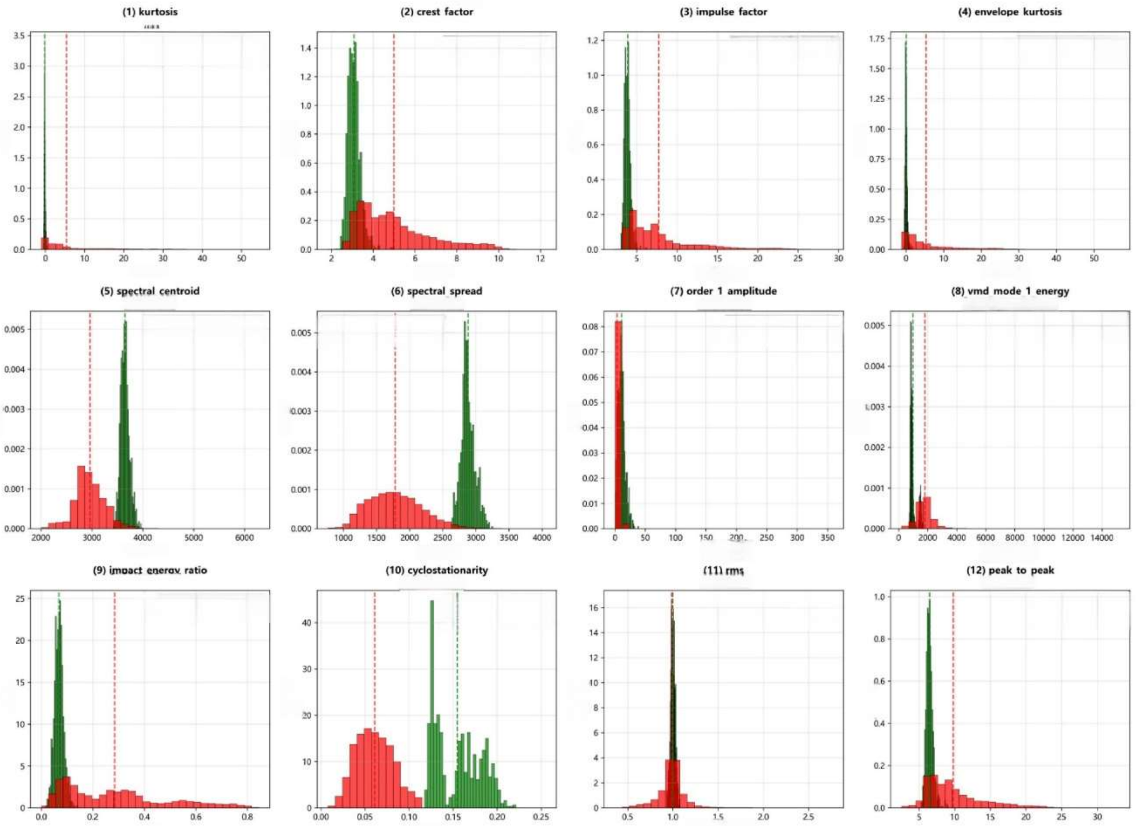

Fig. 1 Comparison of twelve typical features between fault signals and normal signals of aviation bearings

(1) Kurtosis; (2) Crest factor; (3) Impulse factor; (4) Envelope kurtosis; (5) Spectral centroid; (6) Spectral spread; (7) First-order amplitude; (8) VMD mode 1 energy; (9) Impact energy ratio; (10) Cyclostationarity; (11) RMS; (12) Peak-to-peak value

2.3 Model Selection and Analysis

After feature extraction, it is necessary to construct a high-precision source domain bearing fault diagnosis model for analysis and use it as the initial model for subsequent transfer learning. Obviously, the quality of the initial model will directly affect the final transfer learning effect, which is also a key step restricting the final fault diagnosis accuracy. Currently, commonly used learning models include the following:

(1) Support Vector Machine (SVM)

SVM is a machine learning method based on kernel trick and maximum margin classification theory. Studies have shown that it is particularly suitable for pattern recognition tasks in high-dimensional feature spaces and can be used as an effective tool for cross-domain fault feature modeling.

(2) Random Forest

As a representative model of ensemble learning, its core mechanism of "multi-decision tree collaborative voting" has both excellent feature mining ability and model robustness in dealing with the nonlinear modeling task of 49-dimensional fault features, and is an important part of nonlinear modeling in this study. Its principle is shown in Fig. 2.

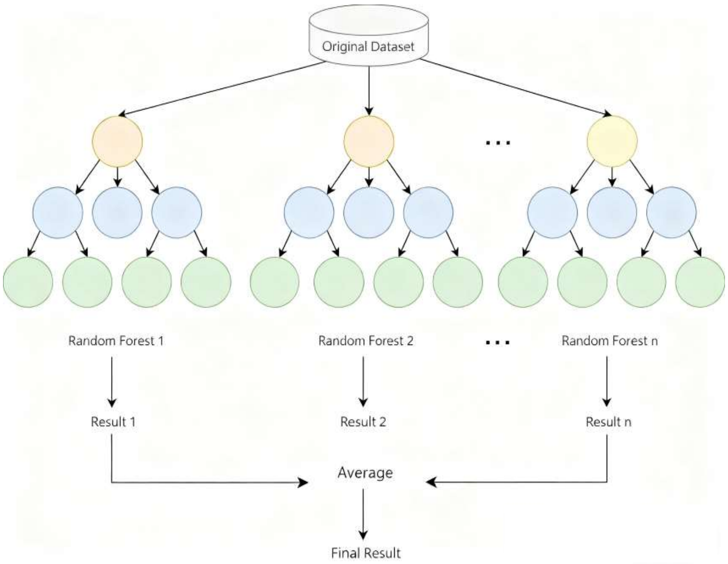

Fig. 2 Random Forest model diagram

(3) XG Boosting)

XG Boosting is a high-performance gradient boosting framework, which is in principle suitable for the complex nonlinear pattern recognition challenges in the 49-dimensional fault features in this study. By integrating multiple decision trees and iteratively correcting residuals round by round, the model significantly enhances the ability to express subtle fault signs in high-dimensional features.

(4) CNN

CNN is a type of deep learning model specially designed for processing grid-structured data (such as images and time-series signals). Through local connection, weight sharing, and hierarchical feature extraction mechanisms, it significantly reduces the number of model parameters and effectively captures spatial local patterns.

(5) Naive Bayes

Naive Bayes is a probabilistic classification model based on Bayes' theorem and the assumption of feature conditional independence. Its core idea is to calculate the posterior probability through prior probability and likelihood probability, and select the class with the maximum posterior probability as the prediction result.

3. Transfer Learning

The core task of this paper is to solve the distribution shift problem caused by differences in operating conditions. Although the fault mechanisms are consistent, there are significant differences in operating conditions such as load, speed, background noise, and installation methods between the test bench (source domain) and the actual aircraft (target domain), resulting in a mismatch between the feature distribution extracted from the test bench data and the actual data. Directly using the source domain model to diagnose the target domain data will lead to a sharp decline in performance. The solution proposed in this paper is to adopt a "domain-adaptive transfer learning framework". Its core idea is: during the training process, the model is not only required to accurately classify fault classes (diagnostic task) but also forced to learn features that can "confuse" the domain discriminator, that is, to make the model accurately classify so that the fault type is infinitely close to the actual data from the aircraft. Thus, the model can separate the differences in operating conditions, focus on the common features of the faults themselves, and complete the actual aviation bearing fault diagnosis and analysis. The algorithm framework proposed in this paper is shown in Fig. 3.

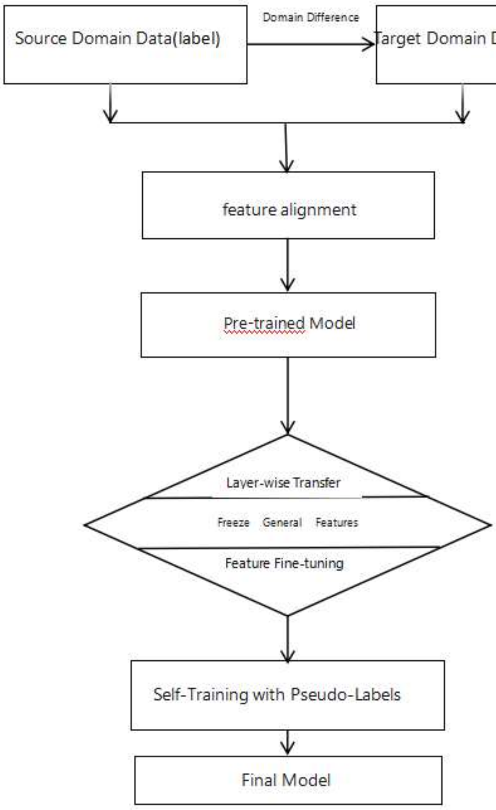

Fig. 3 Overall framework of the proposed transfer learning algorithm

As can be seen from the above figure, the two most critical steps in the transfer learning process proposed in this paper are "feature alignment" and "pseudo-label self-training".

3.1 Feature Engineering Alignment

One of the keys to the success of the transfer learning process in this paper is to use features that are sensitive to faults and not affected by operating conditions. Therefore, among the 49-dimensional feature values proposed earlier, dimensionless indicators such as kurtosis and crest factor should be preferred because they can effectively describe the fault impact characteristics and avoid the interference of changes in speed and load. More importantly, envelope spectrum analysis is adopted, and order analysis (dividing frequency by rotational frequency) is used to eliminate speed differences and realize direct comparison of fault features. At the same time, methods such as VMD can be used to extract the resonance frequency band energy ratio, reducing the impact of absolute energy values. All features need to be standardized based on test bench data, and then the aircraft data are processed according to the same standards to ensure consistent feature scales.

3.2 Pseudo-Label Self-Training

In cross-domain fault diagnosis, the scarcity of labeled data in the target domain is one of the core challenges. To overcome this problem, this paper designs and implements an "iterative pseudo-label self-training mechanism". The core idea of this mechanism is to use the initial transfer model pre-trained on the source domain and fine-tuned on a small amount of labeled data in the target domain to predict a large number of unlabeled samples in the target domain. High-confidence prediction results are selected and assigned "pseudo-labels". These newly labeled samples are merged with the original labeled samples to form an enhanced dataset for the next round of training, thereby gradually expanding high-quality training samples and continuously improving the model's adaptability and generalization ability to target domain features. The core logic of the entire process lies in the dual guarantee of "confidence screening" and "iterative update", aiming to effectively use unlabeled data while strictly avoiding noise interference caused by the introduction of low-confidence pseudo-labels, ensuring the steady improvement of model performance. The algorithm flow chart of this mechanism designed in this paper is shown in Fig. 4.

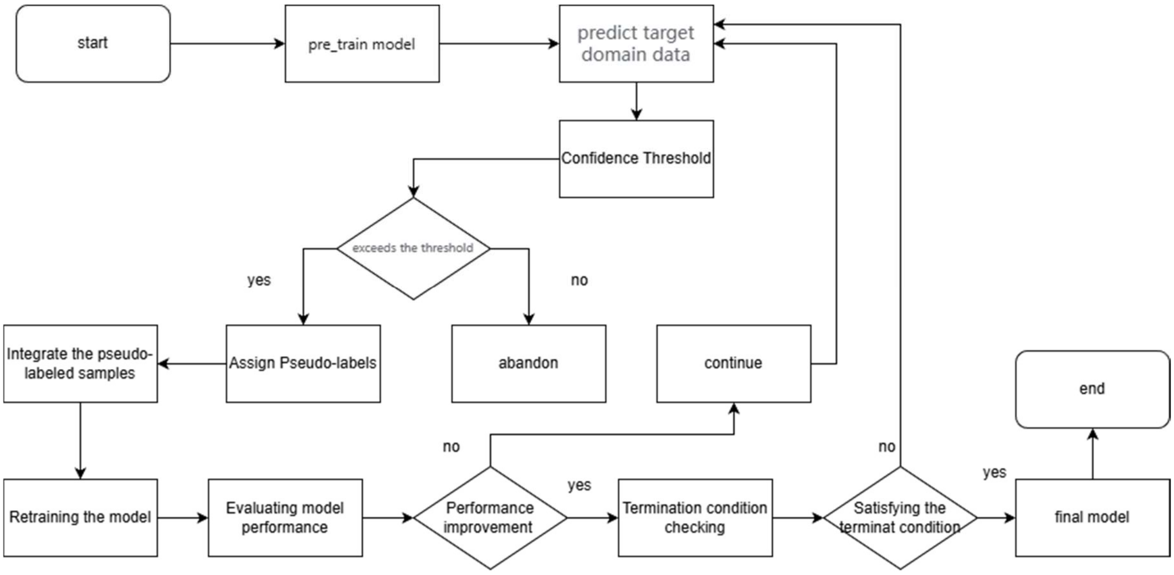

Fig. 4 Flow chart of pseudo-label self-training

3.3 Three-Layer Transfer Learning Strategy

In the process of transfer learning, the credibility and practical application of the model are often restricted, so the interpretability of transfer learning is very important. The core goal of interpretability is to enhance the transparency of the transfer learning process and the results after fault diagnosis. Therefore, based on a clear and progressive interpretability framework, this paper proposes a "three-layer transfer learning strategy". As shown in Fig. 5, the first layer of this strategy is feature space alignment. By using tools such as Maximum Mean Discrepancy (MMD) and CORrelation ALignment (CORAL), the difference in feature distribution between the two domains can be minimized, providing a fair and comparable feature platform for subsequent model transfer. Secondly, the second layer is model layering. First, a pre-trained model is used, then its bottom layer is reused through feature parameter extraction, and then the top task layer is replaced. Then, a small amount of labeled data in the target domain is used to fine-tune the top layer with a very small learning rate (optionally unfreezing a small number of bottom layers). This method can quickly adapt to the target task and solve the problem of small-sample training. Finally, the third layer is sample pseudo-label self-training, whose core is to select reliable samples using the model's own confidence as a clear indicator.

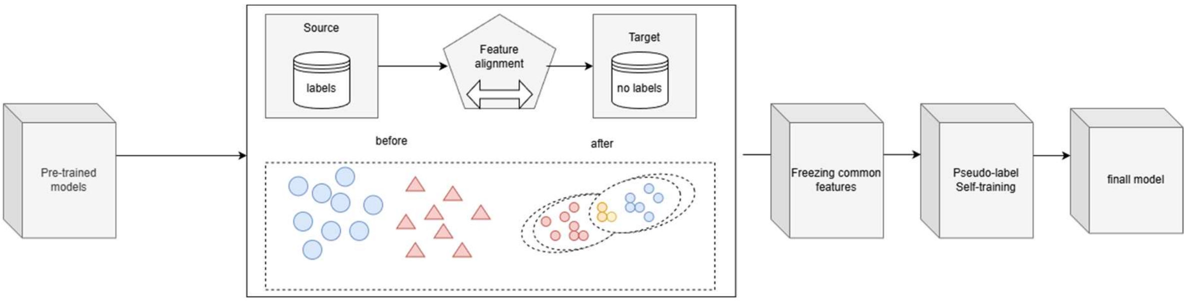

Fig. 5 Three-layer transfer learning strategy of "feature-model-sample collaboration"

As mentioned above, each layer of the "three-layer transfer learning strategy" proposed in this paper has a high degree of prior interpretability, which can ensure the transparency and predictability of the entire transfer learning process.

4. Simulation Analysis and Experimental Verification

To verify the effectiveness of the proposed method in aviation bearing fault diagnosis, the relevant algorithms are applied to the bearing fault dataset of Civil Aviation Flight University of China. The dataset (source domain and target domain) is divided into training set and test set according to the ratio of 8:2, and the verification is completed based on Python and libraries such as NumPy, Keras, SciPy, Pandas, and Seaborn, as shown in Fig. 6.

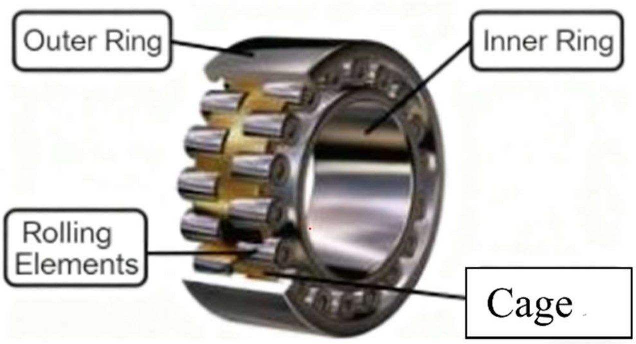

Fig. 6 Test bench bearing

4.1 Data Preprocessing and Analysis

(1) Resampling

The sampling rates of the source domain data include 12 kHz and 48 kHz, while the sampling rate of the target domain data is 32 kHz. To maintain consistency with the target domain data, all source domain data are resampled to 32 kHz to ensure the consistency of data sampling frequency.

(2) Signal Segmentation

This paper adopts a fixed-length sliding window to segment all vibration signals. The window length covers the key feature period to ensure that each segment contains sufficient information. A reasonable overlap rate is set to balance signal continuity and computational efficiency, and the original signal is divided into multiple overlapping short-time samples. This method not only ensures sufficient sample quantity but also enhances the statistical correlation between samples, improves the statistical stability of the sample set, and provides a high-quality data foundation for subsequent analysis. Finally, more than 20,000 valid samples are obtained.

(3) Low-Pass Filter and Normalization

To better denoise the sample signals, a 4th-order Butterworth low-pass filter with a passband width of [12800 Hz] is designed in this paper to complete signal denoising. The standard deviation normalization method is adopted to adjust the signals collected under different operating conditions and different sensor positions to a unified dimension, which can eliminate the adverse impact of amplitude differences on model training.

4.2 Fault Feature Analysis

In the collected vibration signals, the periodic impact components caused by defects present a series of quasi-periodic vibration impacts and attenuation responses. To more intuitively understand the vibration characteristics of bearings under different states, this paper takes a 1.75-inch outer diameter angular contact ball bearing as an example. When the sampling frequency is 48 kHz, the vibration signals of the bearing in ball fault (B), inner race fault (IR), outer race fault (OR), and normal (Normal) states are extracted on the test bench, as shown in Fig. 7. It is obvious in the figure that the vibration stability of fault bearings is poor, and it can be seen that the vibration signals collected when the bearing is in different fault types have different vibration characteristics.

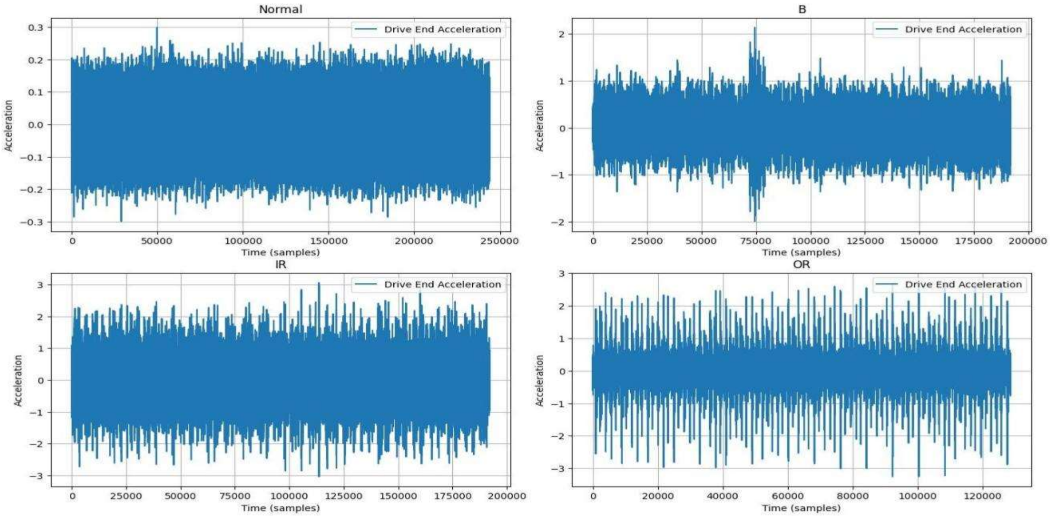

Fig. 7 Vibration characteristic diagrams of the experimental bearing in three typical fault states (B, IR, OR) and normal state (Normal)

4.3 Fault Feature Extraction

In the feature extraction process, clear bearing state labels are designed for each signal segment sample in the source domain and target domain. The experiment divides the bearing health status into 4 typical state types, including 1 normal state (Normal) and 3 typical fault states: ball fault (B), inner race fault (IR), and outer race fault (OR). A 49-dimensional feature vector is extracted from each preprocessed vibration signal segment sample, covering multiple dimensions such as time domain, frequency domain, and time-frequency domain, forming a data structure where feature vectors correspond to state labels one by one, as shown in Table 1. This "49-dimensional feature-single label" sample structure provides a standardized data foundation for the subsequent training of the transfer learning model, ensuring that the model can accurately learn the feature patterns of different fault states and realize effective knowledge transfer from the source domain to the target domain.

Table 1. Storage structure in fault and feature files

Feature 1	Feature 2	...	Feature n	Bearing state	Sample number
Feature value	Feature value	...	Feature value	IR (Inner race fault)	1
Feature value	Feature value	...	Feature value	OR (Outer race fault)	2
Feature value	Feature value	...	Feature value	B (Ball fault)	...
Feature value	Feature value	...	Feature value	N (Normal bearing)	n

4.4 Classification Model Analysis and Result Comparison

(1) Comprehensive Model Comparison

Different classification models have different effects on feature extraction. To find the most suitable model for this project, this paper tests the two different levels of fault features using the commonly used training models mentioned above. The test data comparison is shown in Table 2.

Table 2. Comprehensive performance comparison of different models

	accuracy	CV Mean	Recall (Abnormal)
Random Forest	0.989	0.990	0.99
XGBoost	0.978	0.984	0.99
Naive Bayes	0.788	0.792	0.78
CNN	0.987	0.976	0.98
Svm	0.956	0.967	0.95

It can be seen from the above table that Random Forest is the optimal classification model, so this paper selects Random Forest as the initial model for the next step of transfer learning.

(2) Detailed Evaluation of Random Forest Classification

Recall rate evaluation and analysis: As shown in the confusion matrix in Fig. 8, it presents the performance of the Random Forest model in bearing fault classification. The rows represent the actual classes, and the columns represent the predicted classes, including four categories: B (Ball fault), IR (Inner race fault), Normal, and OR (Outer race fault).

The classification report based on the Random Forest confusion matrix shows that the recall rate of type B is about 99.5% with 5 misdetected samples; the recall rate of type IR is about 99.1% with 9 misdetected samples; the recall rate of Normal type is about 100% with no misdetected samples; the recall rate of type OR is about 98.3% with 9 misdetected samples. The problem of class imbalance is effectively alleviated.

In summary, the model has good recognition effects on the four types of faults, meeting the expectations.

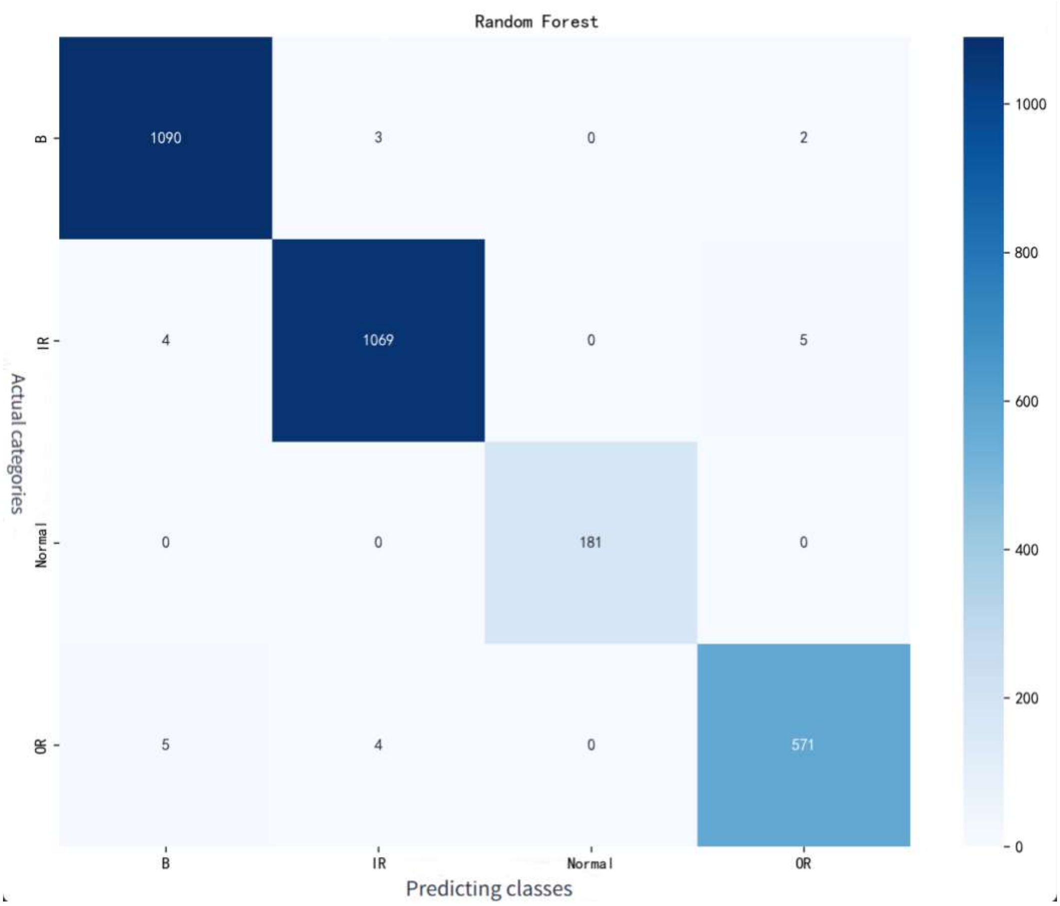

Fig. 8 Confusion matrix of bearing fault recognition by Random Forest model

The higher the importance of a feature, the more the impurity reduction it brings when splitting the tree in the Random Forest, that is, the greater the contribution of this feature to classification. Fig. 9 shows the importance analysis of all features in the model.

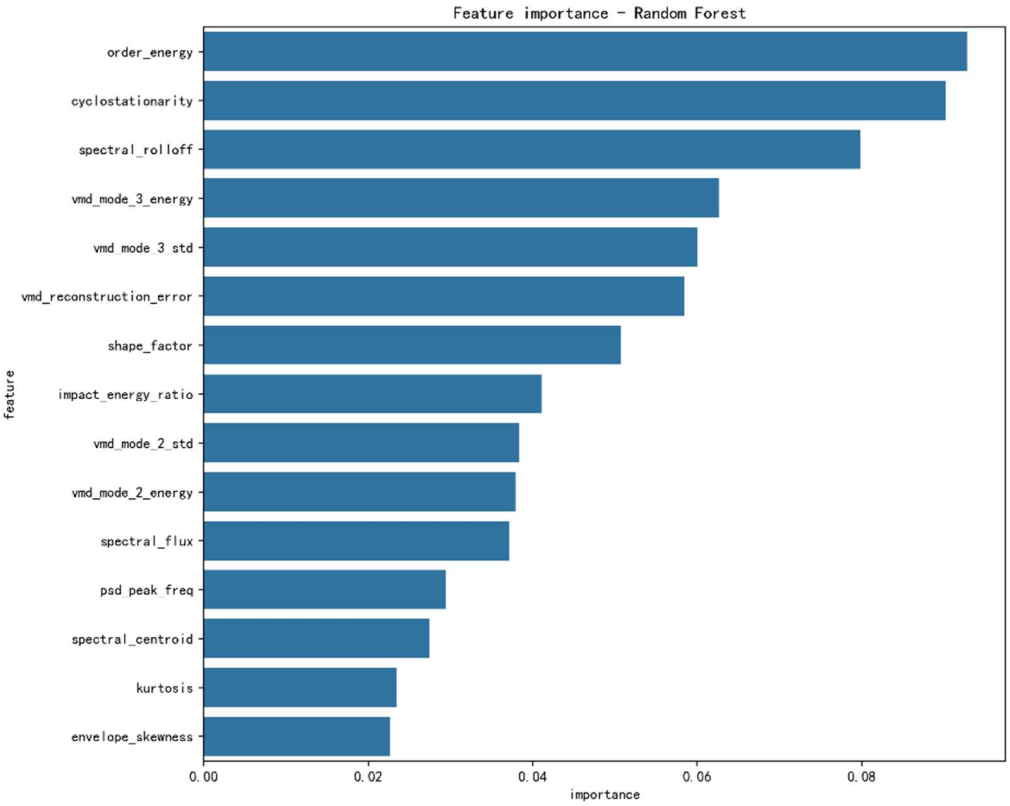

Fig.9 Feature importance in the Random Forest model (Top 15)

4.5 Transfer Learning

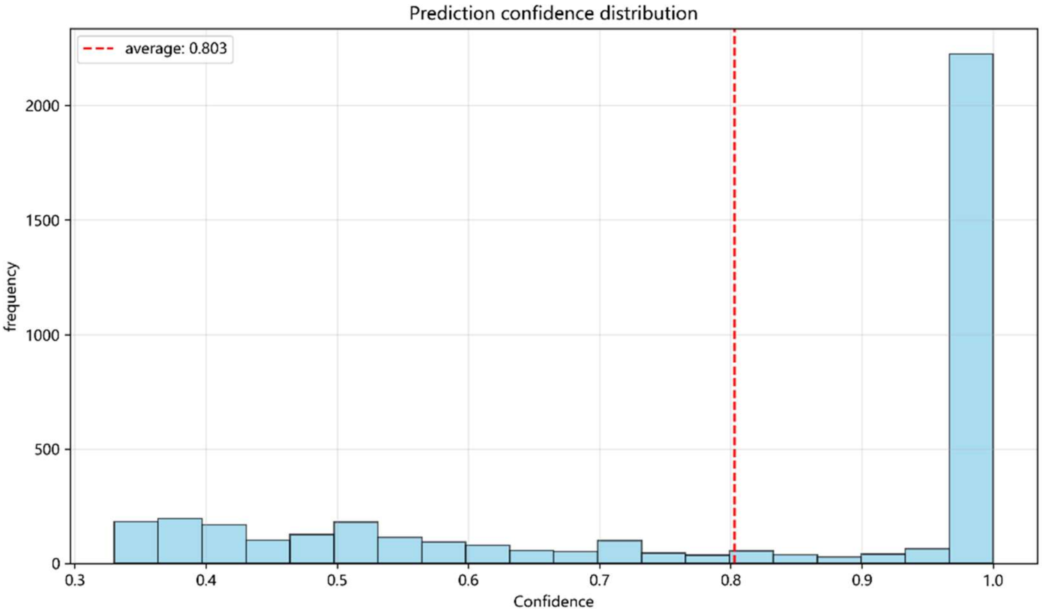

Fig.10 Confidence analysis after transfer

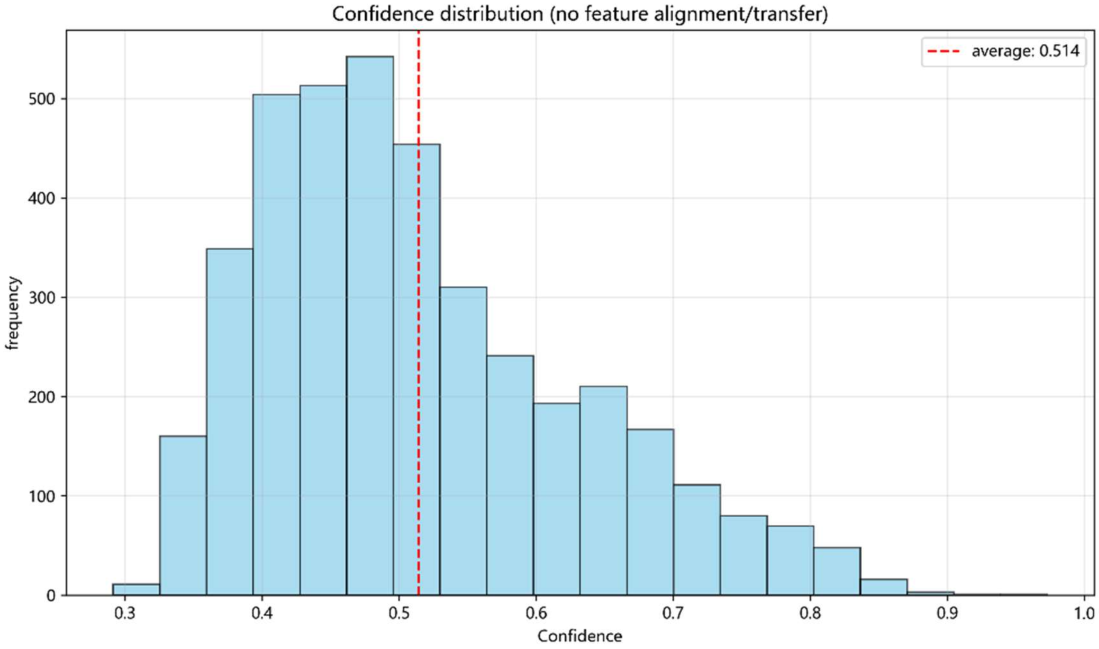

Fig. 11 Confidence analysis without transfer

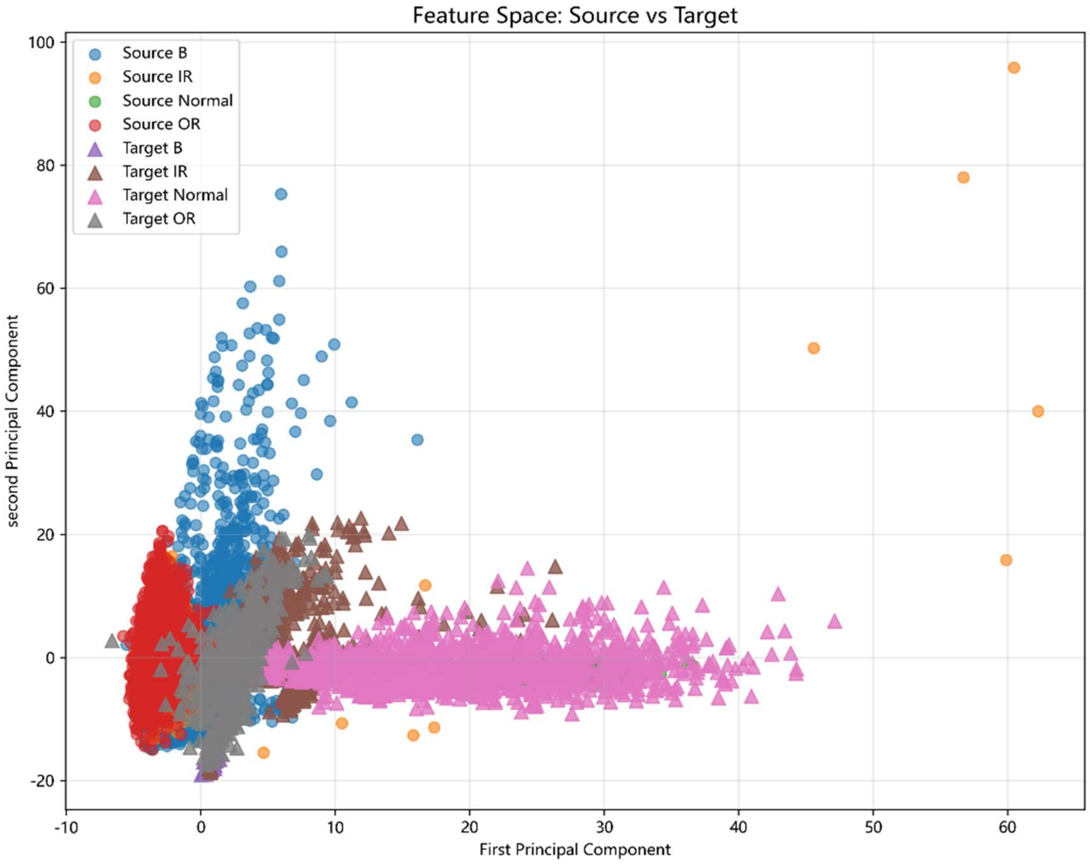

Fig. 12 Feature space distribution of fault classification.

This paper proposes a "domain-adaptive transfer learning" method and conducts experimental verification. The learning results are displayed in the form of confidence distribution, as shown in Figs 10-12. It can be seen from Fig. 10 that the confidence distribution shows that the average prediction confidence of the model is 0.803, and the overall prediction effect is good. Most samples have a confidence level concentrated above 0.9. As shown in Fig. 11, if transfer learning is not performed, the average confidence of directly using the model trained on the source domain (initial model) to classify and predict the target domain data is only 0.514, and most samples have a confidence level concentrated around 0.5. The model's prediction is completely uncertain, indicating that after transfer training, the model is more confident in judgment in most cases. After transfer learning, it has good prediction stability and reliability. As shown in the comprehensive feature space analysis in Fig. 12, the transfer learning model has the ability to diagnose actual aircraft bearing faults.

5. Conclusion

This paper proposes an intelligent diagnosis method for the fault diagnosis of aviation precision bearings. The innovations of this paper are as follows:

(1) Data preprocessing is performed before transfer learning. Resampling can unify the sampling frequency, data low-pass filtering can filter out high-frequency noise, and signal segmentation can improve the stability of the sample set. Such preprocessing methods can improve the accuracy of fault diagnosis and provide a higher-quality data foundation for subsequent transfer learning.

(2) Class weight processing is introduced to solve the problem of sample imbalance. At the same time, the recognition effects of various commonly used classifier models on the dataset of this paper are compared to select the most suitable classifier for this paper, and model parameter optimization is performed to further improve the accuracy of fault diagnosis.

(3) A three-layer transfer learning strategy of "feature-model-sample collaboration" is proposed. Each layer of the strategy is designed with clear intentions and expected effects, and has strong interpretability, enabling the strategy to be finally applied in practice.

(4) An iterative pseudo-label self-training mechanism is proposed, which can solve the problem of data scarcity through a self-iterative enhancement process, and can avoid noise interference caused by the introduction of low-confidence pseudo-labels through a confidence screening process. This dual guarantee can ensure the steady improvement of model performance and further guarantee the effect of final fault diagnosis.

Relevant experimental tests have been carried out on the test bench dataset. The test results show that the fault diagnosis method designed in this paper has good practicality, robustness, and reliability, and has certain popularization and application value.

Acknowledgments

This work was supported by the project "Standards for Emerging Engineering Majors and Construction of Characteristic Teaching Experiment Platforms" (Civil Aviation Education Talent Programs in China), the open project "Research on the Color Recognition and Appearance Defect Detection System for Solar Cells" [Key Laboratory of Solar Energy Technology Integration and Application Promotion in Sichuan Provincial Universities, Grant No. TYNSYS-2018-Y-06], the project "Research on SLAM for Intelligent Civil Aviation Transportation Robots Based on Stereo Vision" [Civil Aviation Flight University of China, Grant No. 25CAFUC04024], and the project "Design of a Single-Star Simulation System for Calibration of Aviation Attitude Determination Star Sensors" [Chengdu Aeronautic Industry and Cultural Development Research Center, Grant No. CAIACDRCXM2024-04].

References

[1] Qiu Z, Yuan X, Wang D, et al. Physical model driven fault diagnosis method for shie ld Machine hydraulic system[J]. Measurement, 2023, 220: 113436.

[2] Zhang X, Gao Y, Guo Z, et al. Physical model-based tool wear and breakage monito ring in milling process[J]. Mechanical Systems and Signal Processing, 2023, 184: 10 9641.

[3] Attouri K, Mansouri M, Hajji M, et al. Improved fault detection based on kernel PCA for monitoring industrial applications[J]. Journal of Process Control, 2024, 133: 103 143.

[4] Lin X, Sun R, Wang Y. Improved key performance indicator-partial least squares me thod for nonlinear process fault detection based on just-in-time learning[J]. Journal of the Franklin Institute, 2023, 360(1): 1-17.

[5] Kini K R, Madakyaru M. Monitoring multivariate process using improved Independe nt component analysis-generalized likelihood ratio strategy[J]. IFAC-PapersOnLine, 2020, 53(1): 392-397.

[6] Bengio Y, Courville A, Vincent P. Representation learning: A review and new persp ectives[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35 (8): 1798-1828.

[7] Song Q, Ma M, Guo W, et al. Health state diagnosis of air duct for photovoltaic inve rter based on PSO-SVM classifier[J]. Microelectronics Reliability, 2023, 150: 11508 4.

[8] Chen Z, Han F, Wu L, et al. Random forest based intelligent fault diagnosis for PV a rrays using array voltage and string currents[J]. Energy conversion and management, 2018, 178: 250-264.

[9] Li P, Liu Z, Anduv B, et al. Diagnosis for multiple faults of chiller using ELM-KNN model enhanced by multi-label learning and specific feature combinations[J]. Buildin g and Environment, 2022, 214: 108904.

[10] Kumar H S, Upadhyaya G. Fault diagnosis of rolling element bearing using continuo us wavelet transform and K-nearest neighbour[J]. Materials today: proceedings, 2023, 92: 56-60.