machines

MDPI

Article

Experimental Cross-Domain Bearing Fault Diagnosis Method Based on Local Mean Decomposition and Improved Transfer Component Analysis

Jia-Peng Liu [1], Zi-Hang Lv [1], Jia-Li Wang [1], Xin-Cheng Yang [1], Zhen-Kun He [2] and Run-Sen Zhang [1,*]

[1] College of Mechatronkcal & Electrical Engineering, Hebei Agricultural University, Baoding 071001, China; 2023214020206@stu.hebau.edu.cn (J.-P.L.); 2023214211417@stu.hebau.edu.cn (Z.-H.L.)
[2] Magdeburg College, Hebei Agricultural University, Baoding 071001, China; 2025154030129@stu.hebau.edu.cn
* Correspondence: zhangrunsen@hebau.edu.cn

Abstract

To address the issue of reduced fault diagnosis accuracy caused by insufficient samples in laboratory datasets, this study proposes an improved Transfer Component Analysis (TCA) algorithm with dynamic kernel parameter adjustment, combined with Local Mean Decomposition (LMD). Firstly, the original signals are decomposed using LMD, and representative signal components are reconstructed based on the Pearson's correlation coefficient to enhance feature representativeness. Then, multidimensional features, including Root Mean Square (RMS), kurtosis, and main frequency (MF), are extracted from the reconstructed signals to comprehensively reflect signal characteristics in terms of energy distribution, impact properties, and frequency structure. Subsequently, a dynamic kernel parameter adjustment strategy is incorporated into TCA to adaptively optimize the kernel parameters, effectively reducing the distribution discrepancy between the source and target domains and enhancing the generalization capability of cross-domain feature transfer. Finally, a Least Squares Support Vector Machine (LSSVM) classifier is employed to perform fault diagnosis on the reconstructed features. The experimental results demonstrate that the proposed method achieves significantly higher diagnostic accuracy than traditional approaches under various operating conditions, especially when signals are complex and distribution differences are large, showing strong robustness and adaptability.

Keywords: local mean decomposition; transfer component analysis; Pearson correlation coefficient; adaptive optimization; LSSVM

Check for updates

Academic Editor: Ahmed Abu-Siada

Received: 22 October 2025
Revised: 8 December 2025
Accepted: 12 December 2025
Published: 12 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Bearings, as critical components of rotating machinery, play an essential role in ensuring the operational safety and reliability of mechanical systems [1]. However, in laboratory research environments, the limited number of samples and the imbalance in class distribution often hinder traditional models from adequately learning representative fault features during training, thereby reducing diagnostic accuracy [2]. In practical industrial applications, bearing vibration signals typically exhibit strong nonlinearity and non-stationarity [3], further increasing the difficulty of extracting effective features. Particularly under limited target-domain samples, distributional discrepancies between source and target domains—caused by variations in operating conditions or equipment—can significantly degrade model generalization performance [4]. These factors collectively restrict the

effectiveness of existing fault diagnosis methods under complex and variable conditions. Therefore, extracting discriminative features from small-sample data and enhancing the adaptability and robustness of diagnostic models through cross-domain transfer have become critical research challenges [5].

Vibration signals constitute the primary data source for bearing fault diagnosis [6]. However, due to their complex time–frequency characteristics, direct analysis of raw signals often fails to yield satisfactory diagnostic results. Time-domain analysis captures dynamic variations through statistical indicators such as mean, variance, and kurtosis, whereas frequency-domain analysis employs Fast Fourier Transform (FFT) to reveal fault-related frequency components [7]. Nevertheless, high-dimensional feature sets frequently contain redundant information, which increases computational complexity and may lead to model overfitting. To address this issue, this study employs Local Mean Decomposition (LMD) [8] to decompose raw vibration signals and reconstruct representative components using the Pearson's correlation coefficient. LMD effectively decomposes nonlinear and non-stationary signals, extracting physically meaningful instantaneous frequency features and thereby enhancing the separability of fault-related characteristics.

To ensure that the extracted features remain both discriminative and robust for cross-domain transfer, this study selects three representative indicators—RMS, kurtosis (K), and main frequency (MF). RMS reflects the signal energy and is closely related to fault severity; kurtosis captures impulsive components produced by localized defects; and MF highlights dominant frequency structures associated with bearing faults. Together, these features provide a compact and physically interpretable description of vibration characteristics. Other potential features, such as crest factor, skewness, entropy-based indices, and frequency-energy ratios, were initially considered. However, comparative analysis showed that many of them were either redundant with RMS and kurtosis or exhibited poor stability under varying loads, reducing their transferability in the TCA subspace. Entropy features, although informative, also increased computational cost and tended to introduce noise into the cross-domain mapping. Therefore, RMS, kurtosis, and MF were ultimately chosen because they offer the best balance of diagnostic sensitivity, robustness, and computational efficiency while avoiding redundancy and ensuring reliable cross-domain feature transfer.

Under limited sample conditions, improving cross-domain fault diagnosis performance remains a major challenge. Transfer learning leverages knowledge from the source domain to assist target-domain learning, effectively mitigating the performance degradation caused by insufficient target samples. Transfer Component Analysis (TCA) [9,10] is a widely adopted approach that aligns source and target distributions within a common subspace to reduce domain discrepancy [11]. However, traditional TCA employs fixed kernel parameters, which cannot adapt to diverse data distributions, thereby limiting the effectiveness of feature transfer. To overcome this limitation, this study proposes an improved TCA method with dynamic kernel parameter adjustment. The method adaptively optimizes kernel parameters according to the statistical characteristics of sample distributions, enabling more consistent feature mapping between source and target domains and enhancing the generalization capability of cross-domain transfer. Recently, adaptive and physics-informed diagnostic methods—such as surrogate modeling for pantograph–catenary interactions [12] and the CFFsBD frequency-guided blind deconvolution [13]—have shown strong interpretability and robustness by incorporating mechanical priors. However, these approaches typically rely on explicit physical models or predefined fault-frequency information, which may reduce their applicability under varying operating conditions or incomplete prior knowledge. In contrast, the proposed ITCA method adaptively adjusts kernel parameters based on data-driven distribution characteristics, without requiring additional physical

assumptions. Combined with LMD's physically meaningful signal decomposition, the proposed framework provides a more flexible and generalizable solution for small-sample cross-domain fault diagnosis.

Furthermore, traditional classifiers such as Support Vector Machine (SVM) [14] and K-Nearest Neighbors (KNN) [15] have demonstrated good performance in handling high-dimensional nonlinear features. However, they often incur high computational costs and exhibit reduced efficiency when dealing with complex data distributions or large sample sizes. The Least Squares Support Vector Machine (LSSVM) [16,17] reformulates the quadratic programming problem into a system of linear equations, thereby reducing computational complexity and improving training efficiency. In addition, LSSVM offers flexibility in kernel function selection and parameter tuning, maintaining high diagnostic accuracy while ensuring strong generalization capability. This makes it particularly suitable for small-sample cross-domain fault diagnosis tasks.

In summary, to address the decline in diagnostic accuracy caused by insufficient laboratory samples, this study proposes an improved TCA-based cross-domain fault diagnosis method that integrates LMD and an LSSVM classifier. The proposed approach not only enhances the effectiveness of feature extraction but also improves the adaptability of cross-domain feature transfer, thereby alleviating performance degradation under limited sample conditions. The experimental results demonstrate that the proposed method achieves high diagnostic accuracy and robustness under various operating conditions, confirming its potential applicability to small-sample cross-domain fault diagnosis and providing a feasible technical framework for translating laboratory research outcomes into practical engineering applications.

2. Signal Decomposition, Reconstruction, and Feature Extraction

2.1. Bearing Weak Signal Extraction Based on LMD

LMD is a nonlinear and non-stationary signal processing method based on signal decomposition [18]. It aims to effectively extract intrinsic modes from signals and overcome the limitations of traditional methods when handling complex signals. LMD decomposes a signal into several Intrinsic Mode Functions, each possessing good time-frequency characteristics. This enables LMD to adapt to various nonlinear and non-stationary signal features, making it suitable for bearing fault diagnosis under low-speed and high-noise conditions. The core idea of LMD is to adaptively decompose the signal, progressively extracting its inherent characteristics. Compared with conventional signal processing methods, LMD adaptively decomposes the signal according to its own characteristics [19].

For the original signal $x(t)$, all local maxima and minima points are first identified, and the local mean $m_k(t)$ and local envelope amplitude $a_k(t)$ between adjacent extrema points are calculated as follows:

$$m_k(t) = \frac{x_{k+1}(t) + x_k(t)}{2} \tag{1}$$

$$a_k(t) = \frac{|x_{k+1}(t) - x_k(t)|}{2} \tag{2}$$

where $x_k(t)$ and $x_{k+1}(t)$ represent the values at the k-th and $(k+1)$-th extrema points, respectively.

The normalized signal $s_k(t)$ is obtained by subtracting the local mean from the original signal and normalizing [20]:

$$s_k(t) = \frac{x(t) - m_k(t)}{a_k(t)} \tag{3}$$

The instantaneous phase $s_k(t)$ is then acquired by applying the Hilbert transform to the normalized signal, followed by calculating the modulation signal $h_k(t)$:

$$h_k(t) = \cos(\theta_k(t)) \tag{4}$$

where In Equation (4), $\theta_k(t)$ represents the instantaneous phase of the normalized signal, obtained from its Hilbert transform. This phase describes the local oscillation behavior of the decomposed component and is used to construct the modulation signal.

Finally, the Product Function component is derived by multiplying the envelope amplitude and modulation signal:

$$PF_k(t) = a_k(t) \cdot h_k(t) \tag{5}$$

2.2. Bearing Signal Reconstruction and Feature Extraction Based on Pearson's Correlation Coefficient

While LMD effectively decomposes the signal into multiple PF components containing different frequency information, some decomposed components may still consist of irrelevant interference that affects the accuracy of feature extraction. To further enhance signal validity and reduce noise impact, the Pearson's Correlation Coefficient (PCC) is used to quantify the correlation between each component and the original signal. This facilitates selection of components significantly contributing to fault features and removal of irrelevant or redundant parts, enabling effective signal reconstruction and improving the accuracy of subsequent fault diagnosis steps.

The Pearson's Correlation Coefficient measures the linear correlation between two variables and ranges from -1 to 1. Suppose the original signal $x(t)$ is decomposed by LMD into K intrinsic mode components $PF_k(t)$ [21]:

$$x(t) = \sum_{k=1}^{K} \mathrm{PF}_k(t) + r(t) \tag{6}$$

The PCC between each $PF_k(t)$ and $x(t)$ is calculated as follows:

$$r_k = \frac{Cov(\mathrm{PF}_k(t), x(t))}{\sigma_{\mathrm{PF}_k} \cdot \sigma_x} \tag{7}$$

where $\mathrm{Cov}(PF_k(t), x(t))$ is the covariance between the component and the original signal, and $\sigma_{\mathrm{PF}k}$ and σ_x are their respective standard deviations. The coefficient r_k indicates the degree of contribution of the k-th component to the original signal, with positive values denoting positive correlation, zero indicating no correlation, and negative values indicating negative correlation.

A threshold τ is set to select effective components such that those with $r_k \geq \tau$ are regarded as valid components, while others are considered noise or irrelevant. The set of effective components is defined as follows:

$$S = \{\mathrm{PF}_k(t) | r_k \geq \tau\} \tag{8}$$

The reconstructed signal $x_r(t)$ is obtained by summing the selected components:

$$x_r(t) = \sum_{\mathrm{PF}_k \in S} \mathrm{PF}_k(t) \tag{9}$$

2.3. Feature Extraction

After signal reconstruction based on LMD and Pearson's correlation coefficient screening, the reconstructed signal $x_r(t)$ contains reduced noise and enhanced fault characteristics. To effectively represent key diagnostic information in both time and frequency domains and provide inputs for subsequent cross-domain feature transfer and classification, three features are extracted from the reconstructed signal: Root Mean Square (RMS), kurtosis (K), and main frequency (MF).

The RMS reflects the overall energy level of the signal and is defined as follows:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_r^2(i)} \tag{10}$$

where N is the number of samples and $x_r(i)$ is the amplitude of the reconstructed signal at the i-th sample. In rotating machinery, faults often cause significant changes in vibration energy, making RMS effective for fault detection and severity assessment.

Kurtosis measures the sharpness of the signal waveform and is given by:

$$K = \frac{\frac{1}{N}\sum_{i=1}^{N} (x_r(i) - \mu)^4}{\left[\frac{1}{N}\sum_{i=1}^{N} (x_r(i) - \mu)^2\right]^2} \tag{11}$$

where μ is the mean of the signal. Kurtosis is sensitive to impulsive components; local damages in bearing rolling elements or raceways often produce pulse-like impacts; so, high kurtosis values are closely related to the presence of faults.

The main frequency corresponds to the frequency component with the highest amplitude in the signal's energy distribution. It is computed by performing the Fast Fourier Transform (FFT) on the reconstructed signal to obtain its spectrum $X(f)$, then identifying the frequency corresponding to the maximum amplitude:

$$MF = \underset{f}{\text{argmax}}|X(f)| \tag{12}$$

The main frequency helps reveal changes in fault characteristic frequencies and assists in distinguishing different fault types. Together, these features capture the signal's energy, impulsive behavior, and frequency content, characterizing the bearing's operational state and fault characteristics from multiple perspectives. The combined feature vector serves as the input for subsequent Transfer Component Analysis and Kernel Extreme Learning Machine classification, contributing to improved accuracy and robustness in cross-domain fault diagnosis.

3. Bearing Fault Diagnosis Based on ITCA-WELM

3.1. Transfer Component Analysis

Transfer Component Analysis is a transfer learning method based on feature space transformation, designed to enhance cross-domain learning performance by reducing the data distribution discrepancy between the source domain and the target domain. TCA aligns shared components between the source and target domains, effectively alleviating the problem of distribution differences between samples from these domains. Unlike traditional supervised learning methods, TCA does not require labeled data from the target domain; instead, it achieves cross-domain knowledge transfer by minimizing the distribution discrepancy between the source and target domains [22].

The objective of TCA is to learn a transformation matrix that maps the source domain and target domain data into a common feature space where they are as close as possible, thereby improving learning performance in the target domain. Given the source domain data X_s and target domain data X_t, TCA aims to find a mapping function ϕ that projects both datasets into a shared feature space. Through this mapping, the source and target data exhibit similar statistical properties in the new space. To reduce the distribution discrepancy between the source and target domains, TCA optimizes the feature mapping by minimizing the Maximum Mean Discrepancy (MMD). MMD measures the difference in means between two distributions and is typically defined as follows:

$$MMD^2(D_s, D_t) = \|\mu_s - \mu_t\|_H^2 \tag{13}$$

where μ_s and μ_t are the means of the source and target domain data, respectively, and $\|\cdot\|_H^2$ denotes the norm in a reproducing kernel Hilbert space (RKHS). By minimizing MMD, TCA effectively reduces the distribution discrepancy between source and target data.

To perform feature alignment in high-dimensional space, this work adopts the Gaussian kernel to compute similarities between data points:

$$K(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2}) \tag{14}$$

where σ is the kernel bandwidth parameter.

Ultimately, TCA optimizes the mapping function by minimizing the MMD between the source and target domains. The optimization objective can be formulated as follows:

$$\min_w MMD^2(D_s, D_t) = \min_w (\frac{1}{n^2}\sum_{i,j} K(x_i^s, x_j^s) - \frac{2}{nm}\sum_{i,j} K(x_i^s, x_j^t) + \frac{1}{m^2}\sum_{i,j} K(x_i^t, x_j^t)) \tag{15}$$

where W is the transformation matrix mapping data from the original space to the new feature space.

3.2. Improved TCA Algorithm Based on Adaptive Kernel Parameter Optimization

A limitation of traditional TCA is the fixed selection of kernel function parameters, which reduces its flexibility and adaptation capability when facing large distribution discrepancies between the source and target domains. This restriction may adversely affect the transfer learning performance. To overcome this, the present study proposes an improved TCA algorithm that adaptively optimizes the kernel function parameter to enhance TCA's adaptability in cross-domain fault diagnosis.

In conventional TCA, a fixed kernel parameter σ is used throughout optimization. Such a fixed parameter cannot accommodate variations in data distribution between domains, especially when distribution differences are substantial. Therefore, a dynamic kernel parameter adjustment strategy is proposed, enabling automatic optimization of σ during each iteration based on the current distributions of the source and target data. This approach achieves better cross-domain feature alignment.

To allow TCA to dynamically adjust the kernel parameter according to actual data distributions, a gradient descent-based adaptive optimization method is introduced. At each optimization iteration, the kernel parameter σ_k is updated as follows:

$$\sigma_k = \sigma_{k-1} - \eta \frac{\partial MMD(\sigma)}{\partial \sigma} \tag{16}$$

where σ_k is the kernel parameter at the k-th iteration, η is the learning rate, and $\frac{\partial MMD(\sigma)}{\partial \sigma}$ denotes the gradient of the MMD loss function with respect to the kernel parameter. This adaptive optimization mechanism enables σ to automatically adjust in response to changes in data distributions, resulting in more precise alignment of source and target features in the transformed space.

During training, to effectively address differences between the source and target domains, a staged optimization strategy is also introduced. In the initial stage, the focus is on optimizing global feature alignment by minimizing the overall distribution discrepancy between the two domains. As training progresses, the optimization gradually shifts towards local feature alignment to better cope with complex fault patterns and feature distributions in the target domain. This staged approach allows the algorithm to flexibly adapt to feature differences between domains and ultimately improves cross-domain learning performance.

3.3. Least Squares Support Vector Machine

In bearing fault diagnosis tasks, constructing a classifier with high generalization capability and computational efficiency is crucial, as the data often exhibit high dimensionality, nonlinearity, and noise interference. The Least Squares Support Vector Machine (LSSVM), an improved version of the standard Support Vector Machine (SVM), retains the superior classification properties of SVM while reducing computational complexity and solving time by transforming the original quadratic programming problem into a linear system of equations through the introduction of a squared loss function.

Let the training dataset be:

$$D = \{(x_i, y_i) | i = 1, 2, \ldots, N\} \tag{17}$$

where $x_i \in \mathbb{R}^d$ represents a d-dimensional feature vector, and $y_i \in \{-1, +1\}$ is the corresponding class label. The objective of LSSVM is to construct a decision function in the form:

$$f(x) = \mathrm{w}^{\mathrm{T}} \phi(x) + b \tag{18}$$

where $\phi(\cdot)$ denotes a nonlinear mapping function that projects the raw data into a high-dimensional feature space, w is the weight vector, and b is the bias term.

The working principle of SVM is illustrated in Figure 1.

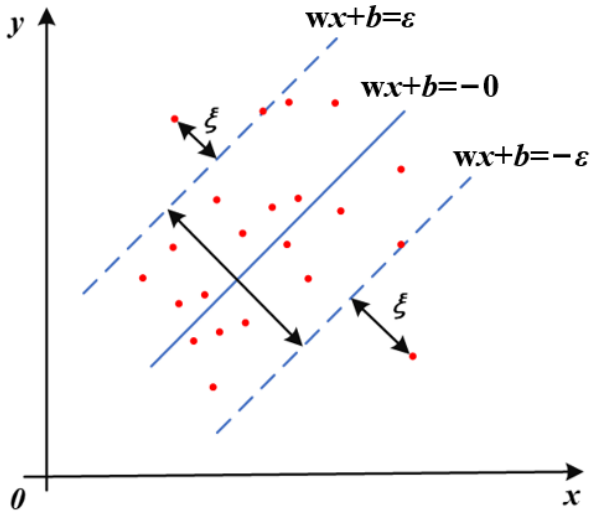


Figure 1. Schematic diagram of SVM.

Traditional SVM employs a hinge loss function and solves a quadratic programming (QP) problem to obtain the optimal solution. In contrast, LSSVM introduces a squared loss function, converting inequality constraints into equality constraints, thereby reformulating the problem into solving a linear system. This transformation simplifies the solving process and enhances computational efficiency for large-scale datasets [23]. The optimization objective of LSSVM is expressed as follows:

$$\min_{\mathrm{w},b,\xi} \frac{1}{2} \|\mathrm{w}\|^2 + \frac{\gamma}{2} \sum_{i=1}^{N} \xi_i^2 \tag{19}$$

where ξ_i represents the slack variable measuring the classification error of sample x_i, and γ is the regularization parameter that balances classifier complexity and training errors. The equality constraint is given by:

$$y_i(\mathrm{w}^{\mathrm{T}} \phi(x_i) + b) = 1 - \xi_i, i = 1, 2, \ldots, N. \tag{20}$$

To solve this constrained optimization problem, the Lagrangian function is constructed by introducing Lagrange multipliers α_i:

$$L(w, b, \xi, \alpha) = \frac{1}{2}\|\mathrm{w}\|^2 + \frac{\gamma}{2}\sum_{i=1}^{N}\xi_i^2 - \sum_{i=1}^{N}\alpha_i\left[y_i(\mathrm{w}^{\mathrm{T}}\phi(x_i) + b) - 1 + \xi_i\right] \tag{21}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)^{\mathrm{T}}$ is the vector of Lagrange multipliers.

By taking partial derivatives of $L(\mathrm{w}, b, \xi)$ with respect to w, b, ξ, the following optimality conditions are derived:

$$\frac{\partial L}{\partial \mathrm{w}} = \mathrm{w} - \sum_{i=1}^{N}\alpha_i y_i \phi(x_i) = 0 \Rightarrow \mathrm{w} = \sum_{i=1}^{N}\alpha_i y_i \phi(x_i) \tag{22}$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{N}\alpha_i y_i = 0 \Rightarrow \sum_{i=1}^{N}\alpha_i y_i = 0 \tag{23}$$

$$\frac{\partial L}{\partial \xi_i} = \gamma\xi_i - \alpha_i = 0 \Rightarrow \alpha_i = \gamma\xi_i, \;\; i = 1, 2, \ldots, N. \tag{24}$$

Equation (22) indicates that the final weight vector is a linear combination of all training samples in the feature space, with coefficients determined by the Lagrange multipliers and class labels. The constraint in Equation (23) ensures the rationality of the bias term b, maintaining balance between different classes. Equation (24) shows that each Lagrange multiplier α_i is proportional to its corresponding error ξ_i, governed by the regularization parameter γ.

Substituting these results into Equation (20) and applying the kernel trick with the kernel function definition [24]:

$$k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \tag{25}$$

The problem is transformed into solving the following linear system:

$$\begin{bmatrix} 0 & y^{\mathrm{T}} \\ y & K + \frac{1}{\gamma}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{26}$$

where $y = [y_1, y_2, \ldots, y_N]^{\mathrm{T}}$ is the class label vector, K is an $N \times N$ kernel matrix with elements $K_{ij} = K(x_i, x_j)$, I is the $N \times N$ identity matrix, and 1 is a vector of ones.

Solving this linear system yields the optimal Lagrange multipliers α and bias b. For a new sample x, its mapping in the high-dimensional feature space is computed via the kernel function. This kernel-based prediction enables LSSVM to effectively capture nonlinear data characteristics, enhancing the classifier's discriminative capability for complex fault patterns.

4. Experimental Cross-Domain Bearing Fault Diagnosis Model Based on Improved Transfer Component Analysis

The proposed method consists of three main stages: signal preprocessing and feature extraction, cross-domain feature transfer, and model training and diagnosis. First, the raw vibration signals from the source domain and target domain are decomposed using Local Mean Decomposition (LMD) to obtain multiple product function components. The Pearson's correlation coefficient is then used to select components that best preserve the characteristics of the original signal for reconstruction. Subsequently, RMS, kurtosis, and main frequency (MF) features are extracted from the reconstructed signals to form the initial feature set. The improved Transfer Component Analysis (TCA) is then applied, where kernel parameters are dynamically adjusted to reduce distribution discrepancy

between domains and enhance cross-domain adaptability. Finally, the aligned source-domain features train an LSSVM classifier, and the target-domain features are input to achieve fault identification. The overall workflow is illustrated in Figure 2.

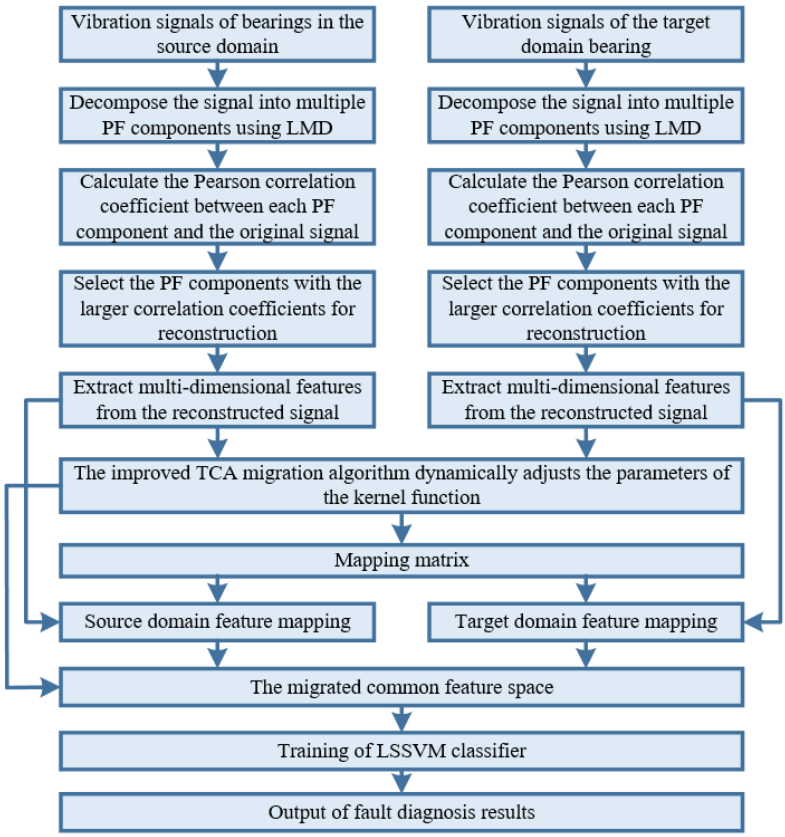


Figure 2. Overall flowchart of the experiment.

5. Experimental Results and Analysis

Figure 3 shows the experimental platform set up in the laboratory. The bearing fault data collected by this platform serves as the target domain, the setup consists, from left to right, of a driving motor, frequency converter, dynamic balancing rotor disk, and a belt-pulley transmission system. This configuration provides stable power transmission and controlled operating conditions for simulating various bearing fault states. Due to the difficulty and cost of reproducing real fault conditions in a laboratory environment, the number of available target-domain samples is relatively limited.

To address the issue of insufficient target-domain samples, the bearing dataset provided by Xi'an Jiaotong University was adopted as the source domain, as illustrated in Figure 4. This dataset, widely used in bearing fault diagnosis research, originates from an experimental platform composed of an AC motor, motor speed controller, rotating shaft, support bearings, hydraulic loading system, and test bearings [25]. The hydraulic loading system applies precise radial loads, while the motor speed controller maintains stable rotation across diverse working conditions. The dataset source is appropriately cited in this study.

For both domains, the data acquisition (DAQ) systems were also configured and described in detail. Specifically, for the source-domain dataset, vibration signals were sampled at 12 kHz, with each sample containing 5000 points under various fault conditions. A total of 100 samples were used for model training and transfer learning. In the target domain, vibration signals were collected through a laboratory DAQ module at a sampling frequency of 20 kHz, with each record containing 5000 points, resulting in 100 samples

for testing and validation. These specifications ensure consistency in signal length while reflecting the natural differences in acquisition environments between the two domains.

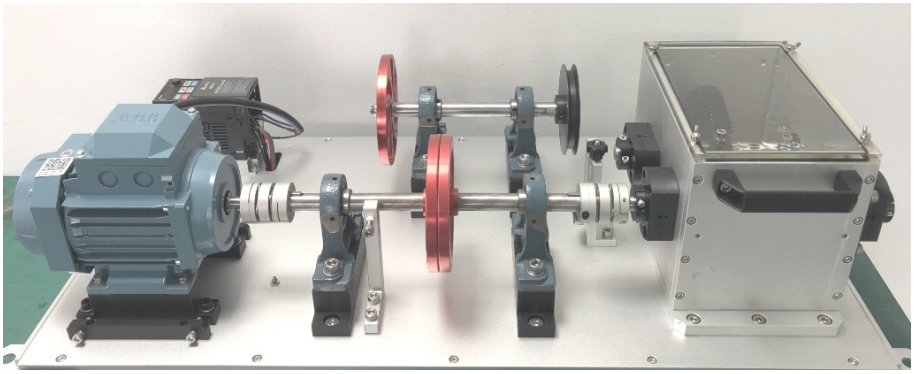


Figure 3. The laboratory signal acquisition device is used for collecting samples from the target domain.

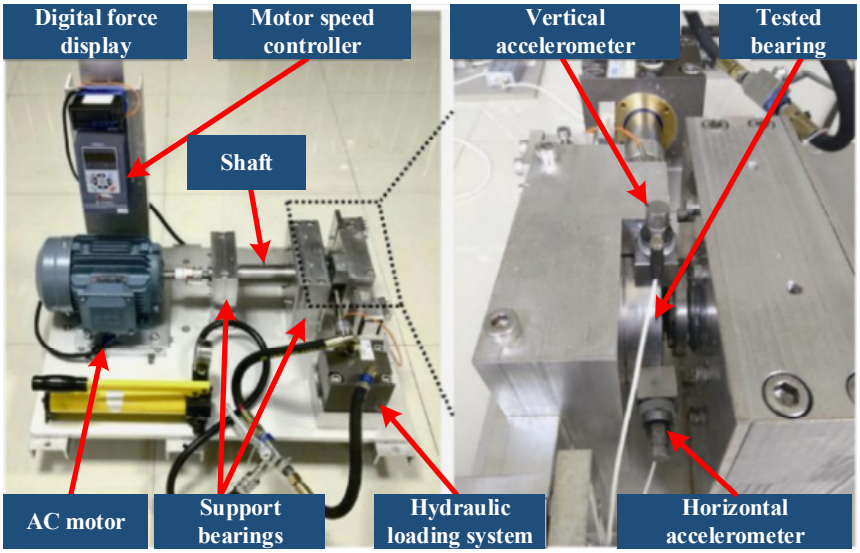


Figure 4. The signal acquisition device of Xi'an Jiaotong University [25] is used for collecting samples from the source domain.

To verify the effectiveness of the proposed method in bearing fault feature extraction and cross-domain diagnosis, this study takes the vibration signals collected from the laboratory test bench as the primary research object. The raw signals acquired under different operating conditions are analyzed to evaluate their overall variation trends in the time domain, the influence of noise interference, and the presence of characteristic components that may directly indicate fault states. The time-domain waveforms of the collected raw signals are presented in Figure 5.

From the waveform distribution in Figure 3, the signals contain obvious periodic components superimposed with random fluctuations, where the random part mainly originates from the mechanical system itself and background noise during sensor measurement. In some operating conditions, impact pulses can be observed in the waveform, but their amplitudes are small and submerged by noise, making it difficult to extract fault features directly from the raw signals. Additionally, due to differences in vibration energy distribution under varying speeds and load conditions, the frequency structure and amplitude changes of the signals exhibit strong non-stationarity. These characteristics indicate that the raw signals require further processing by adaptive time-frequency decomposition methods to separate frequency components and highlight effective fault-related information.

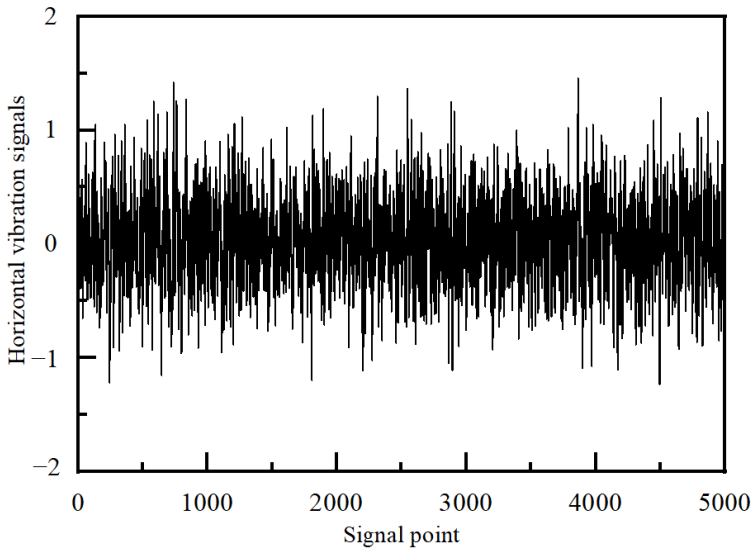


Figure 5. Raw Vibration Signals.

Since the raw vibration signals contain multiple frequency components and noise interference simultaneously, direct extraction of fault features is clearly limited. To effectively separate vibration information at different frequency scales, this paper applies the Local Mean Decomposition method to decompose the raw signals into several physically meaningful product functions. Each PF reflects the signal's variation characteristics within a specific frequency range, facilitating subsequent identification of fault-related specific patterns. Figure 6 shows the waveforms of the decomposed components after LMD.

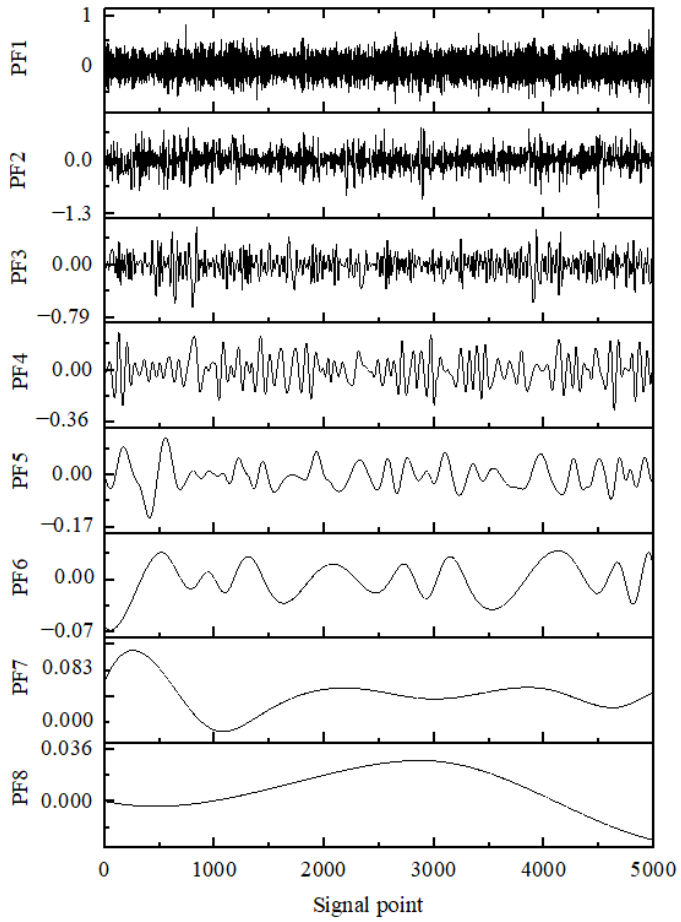


Figure 6. Components after LMD Decomposition.

As shown in Figure 6, LMD decomposes the original signal into multiple PF components, which differ significantly in frequency and amplitude characteristics. The lower-order components (e.g., PF1 and PF2) correspond to the highest frequency bands, typically above 1 kHz, and exhibit large amplitude fluctuations, containing most of the high-frequency impact information related to transient contacts between rolling elements and defects. The mid-order components (e.g., PF3–PF5) have moderate frequencies, approximately 500–1000 Hz, and relatively smaller amplitude variations, reflecting the system's intermediate vibration modes. The higher-order components (e.g., PF6 and above) correspond to low-frequency bands, generally below 500 Hz, and exhibit small amplitude fluctuations; some of these components may mainly consist of noise or irrelevant signals and thus do not directly represent fault features.

This threshold was empirically determined through preliminary experiments, which showed that it provides a reasonable balance between signal fidelity and feature purity: lower thresholds include more components but introduce redundant information, whereas higher thresholds may discard useful fault-related components. To further justify this choice, a sensitivity analysis was conducted by varying the PCC threshold from 0.2 to 0.5 and evaluating the corresponding fault classification accuracy using the LSSVM classifier. The results are summarized in Table 1, which shows that the diagnostic accuracy remains relatively stable for thresholds between 0.3 and 0.4, while performance slightly decreases when thresholds are set below 0.3 or above 0.4. This analysis confirms that 0.3 is an effective choice that balances reconstruction quality and noise suppression.

Table 1. Sensitivity of fault diagnosis accuracy to PCC threshold for signal reconstruction.

PCC Threshold	Number of PFs Selected	Accuracy (%)
0.2	5	93
0.25	4	94
0.3	3	97
0.35	2	93
0.4	2	93
0.45	1	89
0.5	1	89

As illustrated in Table 1, selecting PFs with PCC greater than 0.3 achieves high accuracy while maintaining a compact feature set. Additionally, this LMD–PCC selection strategy inherently suppresses noise: fault-related oscillatory patterns are concentrated in the dominant PFs, whereas noise is dispersed across low-correlation components. Therefore, reconstructing the signal using only highly correlated PFs effectively preserves major fault signatures and reduces noise interference, enhancing the reliability of subsequent feature extraction.

Although LMD can decompose the raw signal into different frequency components, not all components are closely related to fault features; some may mainly consist of noise or irrelevant vibrations, which interfere with feature extraction. To address this, Pearson's Correlation Coefficient (PCC) is introduced to measure the correlation between each PF component and the original signal. Components with correlation above a set threshold are selected as effective components, and their superposition reconstructs an optimized signal. Figure 7 shows the waveform of the reconstructed signal after correlation-based selection. The signal-to-noise ratios (SNRs) of each fault state before and after reconstruction are listed in Table 2, illustrating the enhancement in fault-related information through the reconstruction process.

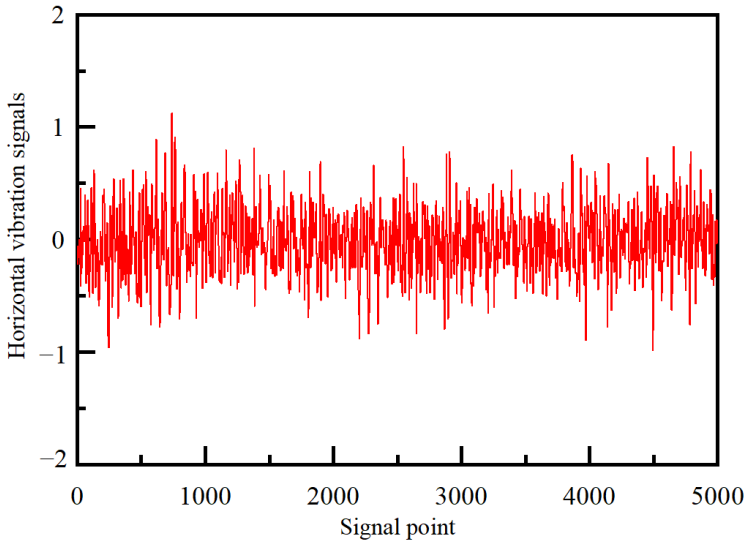


Figure 7. Reconstructed Signal Waveform.

Table 2. Signal-to-Noise Ratios of Different Fault States Before and After Reconstruction.

Fault State	SNR Before Reconstruction (dB)	SNR After Reconstruction (dB)
Normal Bearing	12.5	18.7
Inner Race Fault	14.2	22.3
Outer Race Fault	12.6	26.9
Rolling Element Fault	11.0	21.6

As shown in Figure 7, based on the LMD decomposition results, the Pearson's correlation coefficients between each PF component and the original signal were calculated and ranked by correlation magnitude. Components with high correlation clearly retain the main time-domain characteristics of the raw signal, while those with low correlation mostly contain random noise or irrelevant vibration. To preserve meaningful fault information while avoiding ineffective components, the top three PFs with correlation coefficients greater than 0.3 were selected for reconstruction. This threshold ensures a balance between signal fidelity and feature purity. Moreover, this LMD–PCC selection strategy inherently suppresses noise: fault-related oscillatory patterns are typically concentrated in a few dominant PFs, whereas noise is dispersed across components with weak structure and thus exhibits low correlation. As illustrated in Figure 7, these low-correlation PFs show irregular fluctuations and limited resemblance to the original waveform. Therefore, reconstructing the signal using only the highly correlated PFs effectively retains major fault signatures while reducing noise interference, improving the reliability of subsequent feature extraction.

The signal-to-noise ratios (SNRs) of the different fault states before and after reconstruction are summarized in Table 1. It can be observed that the reconstruction process significantly improves the SNR for all fault conditions: for the normal bearing, the SNR increased from 12.5 dB to 18.7 dB; for the inner race fault, from 14.2 dB to 22.3 dB; for the outer race fault, from 12.6 dB to 26.9 dB; and for the rolling element fault, from 11.0 dB to 21.6 dB. These results demonstrate that the correlation-based selection and reconstruction method effectively enhances dominant fault-related components while suppressing noise, thereby improving signal quality for subsequent feature extraction and fault diagnosis. Moreover, the reconstructed signal exhibits clearer periodic impact features in the waveform, with a significantly reduced noise background and more pronounced amplitude fluctuation patterns. Irrelevant frequency components are effectively removed, resulting in a more concentrated spectral distribution and providing a stable and reliable basis for feature extraction. Overall, selecting high-contribution components based on Pearson correlation

coefficients and reconstructing the signal not only significantly improves SNR but also enhances fault feature separability in the time domain, highlighting the method's effectiveness in improving feature representativeness and reducing noise interference, which positively impacts the accuracy of cross-domain fault diagnosis.

In summary, the LMD–PCC reconstruction process effectively streamlines the raw vibration signals by retaining only the components that contribute meaningful fault-related information. By combining adaptive decomposition with correlation-based selection, this step provides cleaner and more discriminative signals for subsequent feature extraction. The improved SNR and clearer impact characteristics confirm the necessity of this preprocessing stage.

For clearer comparison of cross-domain features and analysis of transfer effects, three types of features—root mean square (RMS), kurtosis (K), and main frequency (MF)—were extracted from the reconstructed signals and distinctly labeled for source and target domain data. Specifically, features from source domain samples are labeled as S-RMS, S-K, and S-MF, while those from target domain samples are labeled as T-RMS, T-K, and T-MF. This labeling facilitates intuitive distinction of feature origins during subsequent visualization and distribution difference assessment, aiding the analysis of feature distribution changes before and after cross-domain transfer. Prior to transfer, significant differences typically exist between S-RMS and T-RMS, S-K and T-K, and S-MF and T-MF in their numerical distributions. These differences mainly stem from changes in vibration energy levels, impact characteristics, and frequency structures under different operating conditions, which often cause conventional classifiers to perform poorly on the target domain.

For clarity and reproducibility, the main implementation parameters of the ITCA-LSSVM framework are summarized as follows. The learning rate η for adaptive kernel optimization was set to 0.05, and the number of kernel-update iterations was limited to 10, as the MMD value typically converged within 6–8 iterations. The TCA regularization parameter λ was set to 1.0, and the subspace dimensionality was fixed at 20. For LSSVM, the penalty factor γ was set to 100, and the RBF kernel width σ was selected using the median heuristic and refined through grid search. For baseline methods, SVM used C = 100 and σ = 0.5; ELM adopted 300 hidden nodes; WKNN used k = 5; and deep learning baselines were trained with a learning rate of 0.001 and batch size of 64. All experiments were conducted in MATLAB R2022b on an Intel i7-11700 workstation. These settings ensure fair comparison and enable reproduction of the proposed method.

To validate the effectiveness of the proposed method, the improved TCA method with dynamic kernel parameter optimization was compared with the commonly used Joint Distribution Adaptation (JDA) algorithm and the traditional TCA method without improvement. The transfer visualization results of each method are shown in Figure 8.

From the transfer visualization results in Figure 6, it can be seen that there are significant differences among the various transfer algorithms in their ability to align features between the source and target domains. For the JDA algorithm, although the RMS features can be clearly separated from other features in the common subspace, the K and MF features from the source domain show a high degree of overlap in distribution, making it difficult to form distinct decision boundaries, which results in insufficient feature discrimination. In contrast, the traditional TCA algorithm improves feature separation compared to JDA, with RMS, K, and MF forming relatively independent clusters in the common subspace. However, some overlap between K and MF features remains, affecting the purity of classification boundaries. After applying the proposed improved TCA method with dynamic kernel parameter optimization, the three features are clearly distinguished in the mapped space, cluster separation is enhanced, and the distributions of the same class features across

domains highly overlap. This effectively improves cross-domain feature consistency and discriminative ability.

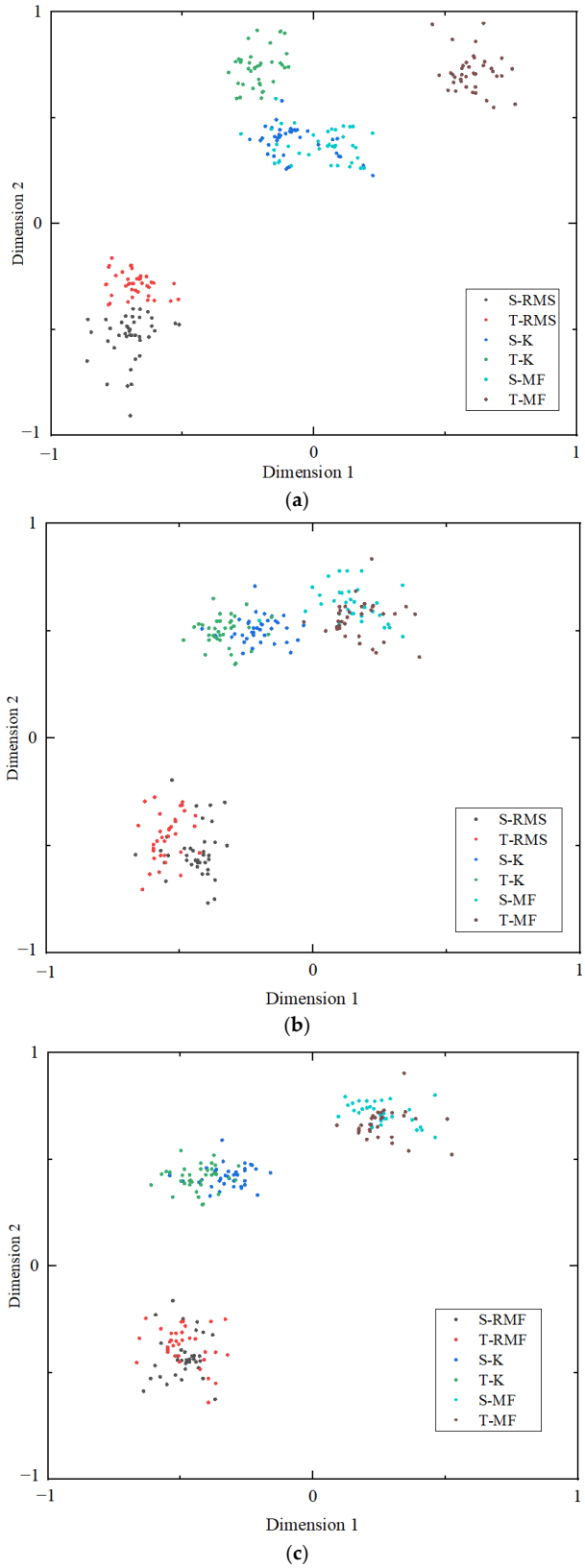


Figure 8. Visualization of transfer effects using different algorithms. (**a**) Visualization results of the JDA algorithm. (**b**) Visualization results of the TCA algorithm. (**c**) Visualization results of the improved TCA algorithm.

The underlying reason for this phenomenon is that while JDA can simultaneously align marginal and conditional distributions of source and target domains, its ability to model nonlinear relationships between feature dimensions is limited. Especially when impact characteristics (K) and frequency characteristics (MF) are coupled, feature mixing is prone to occur. On the other hand, TCA constructs a common subspace and minimizes domain differences, which inherently provides better feature separation than JDA for such problems. However, traditional TCA uses fixed kernel parameters, making it difficult to optimally map different feature distributions. The proposed method introduces a dynamic kernel parameter adaptive optimization mechanism within the TCA framework, allowing the mapped space to better fit the distribution characteristics of RMS, K, and MF features, thus achieving optimal domain alignment and inter-class separation.

Overall, the visualization results clearly show a progressive improvement in inter-domain feature alignment from JDA to traditional TCA and finally to the proposed ITCA. By enhancing cluster compactness and class separability, the improved method provides a more consistent shared feature space, which directly contributes to improved transferability and diagnostic performance in later experiments.

To further validate the effectiveness of the proposed fault diagnosis algorithm combining dynamic kernel parameter optimized Transfer Component Analysis with Least Squares Support Vector Machine (ITCA-LSSVM), a series of comparative experiments were conducted. The source domain dataset consists of 100 samples from Xi'an Jiaotong University, while the target domain dataset, collected in the laboratory, contains 100 samples. Vibration signals from both domains were preprocessed using LMD–PCC reconstruction. Diagnostic accuracy is calculated as follows:

$$\text{Accuracy } (\%) = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}} \times 100 \tag{27}$$

Traditional algorithms selected as baselines include Joint Distribution Adaptation with LSSVM (JDA-LSSVM), conventional Transfer Component Analysis with LSSVM (TCA-LSSVM), Long Short-Term Memory networks (LSTM), Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Extreme Learning Machines (ELM), Weighted K-Nearest Neighbors (WKNN) and domain-adaptive CNN (da-CNN). Additionally, ITCA transfer was combined with other classifiers (ITCA-SVM, ITCA-WKNN, ITCA-ELM) to evaluate the effect of dynamic kernel optimization across different machine learning methods.

The results, shown in Figure 9 and summarized in Table 3, indicate that ITCA-LSSVM achieves the highest diagnostic accuracy of 97%, surpassing traditional transfer learning methods (JDA-LSSVM 90%, TCA-LSSVM 92%) and deep learning models without transfer learning (LSTM 85%, CNN 81%, da-CNN 91%). The superior performance of ITCA-LSSVM is attributed to its dynamic kernel parameter optimization, which adaptively aligns source and target feature distributions, improving feature transfer and generalization. Compared with traditional transfer learning methods, the additional computation for dynamic kernel adjustment slightly increases the training time (ITCA-LSSVM: 11.7 s vs. JDA-LSSVM: 10.2 s, TCA-LSSVM: 10.8 s), but this increase is relatively modest and acceptable given the substantially improved diagnostic accuracy.

Table 3. Comparative performance and training time of different models.

Method	Accuracy (%)	Training Time (s)
ITCA-LSSVM	97	11.7
JDA-LSSVM	90	10.2
TCA-LSSVM	92	10.8
ITCA-SVM	90	11.2
ITCA-WKNN	93	11.0
ITCA-ELM	92	11.5
LSTM	85	50.3
CNN	81	45.7
da-CNN	91	58.4

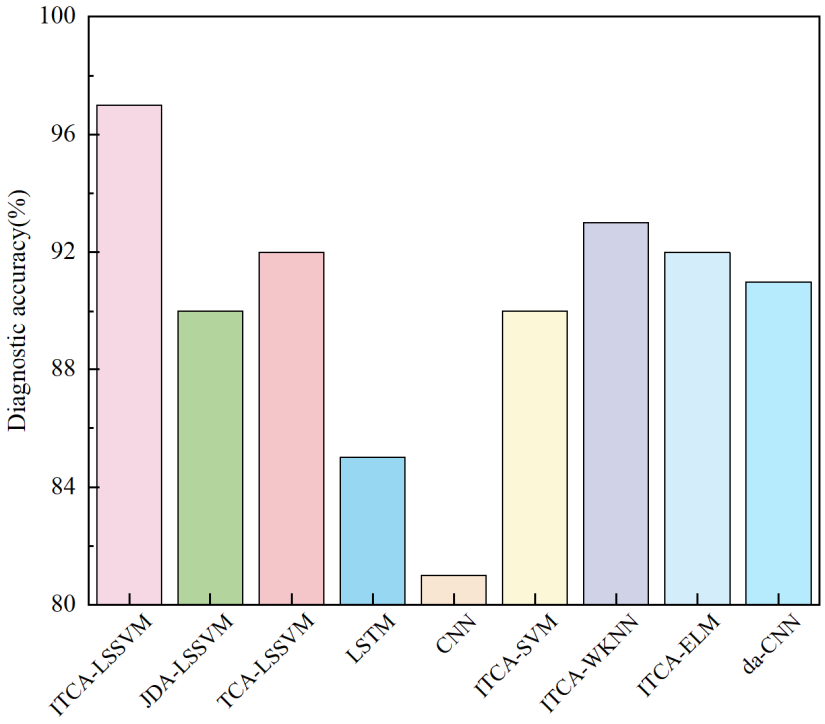


Figure 9. Diagnostic accuracy results of different models.

Moreover, combining ITCA with LSSVM leverages the classifier's high efficiency and robustness, allowing ITCA-LSSVM to achieve high accuracy while maintaining shorter training times than deep learning models (LSTM: 50.3 s, CNN: 45.7 s, da-CNN: 58.4 s). Other ITCA-based classifiers (ITCA-SVM, ITCA-WKNN, ITCA-ELM) also achieve high accuracy (90–93%), further demonstrating that the proposed transfer strategy effectively improves cross-domain feature alignment and recognition performance across various traditional machine learning models.

To further examine the convergence behavior of the proposed dynamic kernel optimization, the evolution of the maximum mean discrepancy (MMD) across iterations was monitored during training. The results show that MMD decreases rapidly within the first 5–8 iterations and stabilizes afterward, indicating that the adaptive kernel parameter update converges quickly toward a stable transfer mapping. This fast convergence is attributed to the monotonic update rule, in which kernel width adjustments are guided by the gradient of the domain divergence. No oscillations or divergence were observed across repeated runs, demonstrating stable convergence properties.

In terms of computational cost, the adaptive kernel optimization introduces only a lightweight overhead. For a dataset with n samples and feature dimension d, the complexity

of standard TCA is approximately $O(n^2d + d^3)$, dominated by kernel matrix construction and eigenvalue decomposition.

The proposed ITCA 0adds only an iterative kernel adjustment step with complexity $O(n^2)$ per iteration. Given that convergence occurs within fewer than 10 iterations, the overall complexity remains on the same order as traditional TCA. This is consistent with the experimental training time results in Table 3, where ITCA-LSSVM increases computation by only 0.9–1.5 s compared with TCA-LSSVM.

A sensitivity analysis was further conducted to assess the influence of the adaptive kernel parameter update. Specifically, three variants were compared:

(1) Fixed kernel parameter (i.e., traditional TCA);
(2) Coarse-step adaptive update;
(3) Fine-step adaptive update (proposed ITCA).

The results in Table 4 show that coarse updates weaken domain alignment performance, while fine-grained adaptive updates (ITCA) consistently achieve the best accuracy.

Table 4. Ablation study on kernel parameter update strategy.

Method	Update Strategy	Accuracy (%)	Training Time (s)
TCA-LSSVM	Fixed kernel	92	10.8
ITCA (coarse update)	Step size = 0.2	94	11.3
ITCA (fine update, proposed)	Step size = 0.05	97	11.7

These results demonstrate that the proposed adaptive kernel update is not only effective but also robust to parameter variations, and the full-version ITCA provides the optimal balance between computational cost and diagnostic accuracy.

6. Conclusions

This study addresses the challenges of signal non-stationarity and sample distribution differences in cross-domain bearing fault diagnosis by proposing a fault diagnosis method that combines LMD and improved TCA with dynamic kernel parameter optimization (ITCA). Firstly, LMD is used to perform multi-scale decomposition of vibration signals, and the effective components are selected based on the Pearson's correlation coefficient, significantly improving the signal-to-noise ratio and enhancing the representativeness of fault features. Secondly, the dynamic kernel parameter-optimized TCA effectively reduces the distribution differences between the source and target domain features, improving the accuracy and robustness of cross-domain feature alignment. Finally, a Least Squares Support Vector Machine (LSSVM) is employed for fault classification, constructing an efficient and stable diagnostic model.

Experiments are conducted using laboratory-collected bearing fault data as the target domain samples, combined with the source domain dataset for cross-domain diagnosis tasks. Transfer visualization results indicate that the dynamic kernel parameter-optimized TCA significantly outperforms traditional JDA and fixed kernel parameter TCA in feature distribution alignment and discriminability. In terms of diagnostic performance, the ITCA-LSSVM model achieves the highest accuracy of 97% among various comparison methods, demonstrating the advantages of the proposed approach in cross-domain fault identification. Moreover, applying the ITCA transfer strategy to traditional machine learning models such as SVM, WKNN, and ELM also significantly improves diagnostic accuracy, further validating the generality and applicability of the dynamic kernel parameter-optimized transfer method.

In summary, the proposed method combining LMD decomposition and dynamic kernel parameter-optimized TCA provides an effective solution for cross-domain adaptability

in bearing fault diagnosis under complex operating conditions. The approach not only enhances the accuracy and robustness of fault feature extraction but also significantly improves the generalization capability of the diagnostic model, offering important application value and promotion potential for intelligent monitoring and maintenance of industrial equipment. Future work may further integrate deep transfer learning and online adaptive mechanisms to enhance real-time diagnostic capabilities under more complex practical operating conditions. Moreover, combining the ITCA framework with meta-learning strategies or real-time adaptive transfer techniques could enable rapid model adaptation to new machinery and operational environments, facilitating online deployment in industrial settings.

Author Contributions: Conceptualization, J.-P.L. and R.-S.Z.; methodology, J.-P.L.; software, Z.-H.L.; validation, J.-L.W., X.-C.Y. and Z.-K.H.; formal analysis, J.-P.L.; investigation, J.-L.W.; resources, R.-S.Z.; data curation, Z.-H.L.; writing—original draft preparation, J.-P.L.; writing—review and editing, R.-S.Z.; visualization, X.-C.Y.; supervision, R.-S.Z.; project administration, R.-S.Z.; funding acquisition, R.-S.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grant number 702/3118123, under the project titled "Influence of Relative Twist Angle on the Structural, Mechanical, and Thermal Properties of Two-Dimensional van der Waals Materials and Structural Optimization Design". The APC was funded by the same foundation.

Data Availability Statement: The data presented in this study are not publicly available at this time because they will be used for further experimental validation in subsequent research. However, the data can be made available upon reasonable request from the corresponding author.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. Hoang, D.T.; Kang, H.J. A survey on deep learning based bearing fault diagnosis. *Neurocomputing* **2019**, *335*, 327–335. [CrossRef]
2. Chen, X.; Yang, R.; Xue, Y.; Huang, M.; Ferrero, R.; Wang, Z. Deep transfer learning for bearing fault diagnosis: A systematic review since 2016. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3508221. [CrossRef]
3. Zimroz, R.; Bartelmus, W.; Barszcz, T.; Urbanek, J. Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings. *Mech. Syst. Signal Process.* **2014**, *46*, 16–27. [CrossRef]
4. Mao, W.; Liu, Y.; Ding, L.; Li, Y. Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: A comparative study. *IEEE Access* **2019**, *7*, 9515–9530. [CrossRef]
5. Su, H.; Xiang, L.; Hu, A.; Xu, Y.; Yang, X. A novel method based on meta-learning for bearing fault diagnosis with small sample learning under different working conditions. *Mech. Syst. Signal Process.* **2022**, *169*, 108765. [CrossRef]
6. McInerny, S.A.; Dai, Y. Basic vibration signal processing for bearing fault detection. *IEEE Trans. Educ.* **2003**, *46*, 149–156. [CrossRef]
7. Fang, X.; Zheng, J.; Jiang, B. A rolling bearing fault diagnosis method based on vibro-acoustic data fusion and fast Fourier transform (FFT). *Int. J. Data Sci. Anal.* **2024**, *20*, 2377–2386. [CrossRef]
8. Goyal, D.; Choudhary, A.; Sandhu, J.K.; Srivastava, P.; Saxena, K.K. An intelligent self-adaptive bearing fault diagnosis approach based on improved local mean decomposition. *Int. J. Interact. Des. Manuf. (IJIDeM)* **2022**, 1–11. [CrossRef]
9. Ma, P.; Zhang, H.; Wang, C. Improved transfer component analysis and it application for bearing fault diagnosis across diverse domains. In Proceedings of the 2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), Xiamen, China, 5–7 July 2019; IEEE: Piscataway, NJ, USA, 2020; pp. 501–506.
10. Li, D.; Ma, M. A bearing fault diagnosis method based on improved transfer component analysis and deep belief network. *Appl. Sci.* **2024**, *14*, 1973. [CrossRef]
11. Li, X.; Zhang, Z.; Gao, L.; Wen, L. A new semi-supervised fault diagnosis method via deep CORAL and transfer component analysis. *IEEE Trans. Emerg. Top. Comput. Intell.* **2021**, *6*, 690–699. [CrossRef]
12. Cheng, Y.; Yan, J.; Zhang, F.; Li, M.; Zhou, N.; Shi, C.; Jin, B.; Zhang, W. Surrogate modeling of pantograph-catenary system interactions. *Mech. Syst. Signal Process.* **2025**, *224*, 112134. [CrossRef]

13. Cheng, Y.; Zhou, N.; Wang, Z.; Chen, B.; Zhang, W. CFFsBD: A candidate fault frequencies-based blind deconvolution for rolling element bearings fault feature enhancement. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–12. [CrossRef]
14. Han, T.; Zhang, L.; Yin, Z.; Tan, A.C. Rolling bearing fault diagnosis with combined convolutional neural networks and support vector machine. *Measurement* **2021**, *177*, 109022. [CrossRef]
15. Tian, J.; Morillo, C.; Azarian, M.H.; Pecht, M. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis. *IEEE Trans. Ind. Electron.* **2015**, *63*, 1793–1803. [CrossRef]
16. Li, X.; Yang, Y.; Pan, H.; Cheng, J.; Cheng, J. A novel deep stacking least squares support vector machine for rolling bearing fault diagnosis. *Comput. Ind.* **2019**, *110*, 36–47. [CrossRef]
17. Han, H.; Cui, X.; Fan, Y.; Qing, H. Least squares support vector machine (LS-SVM)-based chiller fault diagnosis using fault indicative features. *Appl. Therm. Eng.* **2019**, *154*, 540–547. [CrossRef]
18. Cheng, J.; Yang, Y.; Yang, Y. A rotating machinery fault diagnosis method based on local mean decomposition. *Digit. Signal Process.* **2012**, *22*, 356–366. [CrossRef]
19. Li, Y.; Xu, M.; Wang, R.; Huang, W. A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy. *J. Sound Vib.* **2016**, *360*, 277–299. [CrossRef]
20. Li, Y.; Si, S.; Liu, Z.; Liang, X. Review of local mean decomposition and its application in fault diagnosis of rotating machinery. *J. Syst. Eng. Electron.* **2019**, *30*, 799–814. [CrossRef]
21. Yang, J.; Bai, Y.; Cheng, Y.; Cheng, R.; Zhang, W.; Zhang, G. A new model for bearing fault diagnosis based on optimized variational mode decomposition correlation coefficient weight threshold denoising and entropy feature fusion. *Nonlinear Dyn.* **2023**, *111*, 17337–17367. [CrossRef]
22. Xie, J.; Zhang, L.; Duan, L.; Wang, J. On cross-domain feature fusion in gearbox fault diagnosis under various operating conditions based on transfer component analysis. In Proceedings of the 2016 IEEE International Conference on Prognostics and Health Management (ICPHM), Ottawa, ON, Canada, 20–22 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–6.
23. Gao, Q.W.; Liu, W.Y.; Tang, B.P.; Li, G.J. A novel wind turbine fault diagnosis method based on intergral extension load mean decomposition multiscale entropy and least squares support vector machine. *Renew. Energy* **2018**, *116*, 169–175. [CrossRef]
24. Liu, X.; Bo, L.; Luo, H. Bearing faults diagnostics based on hybrid LS-SVM and EMD method. *Measurement* **2015**, *59*, 145–166. [CrossRef]
25. Wang, B.; Lei, Y.; Li, N.; Li, N. A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.