electronics

Article

WaveCORAL-DCCA: A Scalable Solution for Rotor Fault Diagnosis Across Operational Variabilities

Nima Rezazadeh [1,2,*], Mario De Oliveira [2], Giuseppe Lamanna [1], Donato Perfetto [1] and Alessandro De Luca [1]

[1] Department of Engineering, University of Campania "Luigi Vanvitelli", 81031 Aversa, Italy; giuseppe.lamanna@unicampania.it (G.L.); donato.perfetto@unicampania.it (D.P.); alessandro.deluca@unicampania.it (A.D.L.)
[2] School of Architecture, Built Environment, Computing and Engineering, Birmingham City University, Birmingham B4 7XG, UK; mario.deoliveira@bcu.ac.uk
* Correspondence: nima.rezazadeh@unicampania.it

Abstract

This paper presents WaveCORAL-DCCA, an unsupervised domain adaptation (UDA) framework specifically developed to address data distribution shifts and operational variabilities (OVs) in rotor fault diagnosis. The framework introduces the novel integration of discrete wavelet transformation for robust time–frequency feature extraction and an enhanced deep canonical correlation analysis (DCCA) network with correlation alignment (CORAL) loss for superior domain-invariant representation learning. This combination enables more effective alignment of source and target feature distributions without requiring any labelled data from the target domain. Comprehensive validation on both experimental and numerically simulated rotor datasets across three health conditions—i.e., normal, unbalanced, and misaligned—demonstrates that WaveCORAL-DCCA achieves an average diagnostic accuracy of 95%. Notably, it outperforms established UDA benchmarks by at least 5–17% in cross-domain scenarios. These results confirm that WaveCORAL-DCCA provides robust generalisation across machines, fault severities, and operational conditions, even with scarce target domain samples, offering a scalable and practical solution for industrial rotor fault diagnosis.

Keywords: rotor fault diagnosis; operational variabilities; unsupervised domain adaptation; wavelet transformation; deep canonical correlation analysis; correlation alignment

Academic Editor: Davide Astolfi

Received: 10 July 2025
Revised: 30 July 2025
Accepted: 4 August 2025
Published: 7 August 2025

Citation: Rezazadeh, N.; De Oliveira, M.; Lamanna, G.; Perfetto, D.; De Luca, A. WaveCORAL-DCCA: A Scalable Solution for Rotor Fault Diagnosis Across Operational Variabilities. *Electronics* **2025**, *14*, 3146. https://doi.org/10.3390/electronics14153146

1. Introduction

As a crucial component of health management and prognostics, fault diagnosis is cohesively and strategically integrated into industrial systems, making it a fundamental technology for their operations [1–4]. Intelligent fault diagnosis methods, driven by advancements in artificial intelligence, have gained prominence by enabling automatic fault detection; however, these methods often depend on extensive fault history and labelled data, typically scarce in industrial settings [5]. Researchers address this scarcity by generating synthetic data through numerical modelling, lab-scale experiments, and data from similar systems. Although these approaches help mitigate data scarcity, they introduce domain mismatches due to OVs such as variations in angular velocity and environmental conditions, as extensively discussed in recent rotor modelling studies, which can degrade diagnostic performance [6–12].

Recent advancements in intelligent fault diagnosis have focused on overcoming challenges related to class imbalance and domain shifts, specifically changes in data distribution caused by these OVs. Techniques such as the self-supervised approach by Miao et al. [13,14] integrate noise filtering and class balancing to improve diagnostics under significant class imbalance and OVs. However, extending this method to more complex scenarios, such as varying rotational speeds and multiple fault types, remains challenging. Rajagopalan et al. [15] address these issues using synthetic oversampling, genetic algorithms, and adversarial adaptation to diagnose rotor mass imbalance faults. Despite its robustness, this method suffers from minor class boundary overlaps and complexity, indicating a need for further refinement.

Evolving diagnostic approaches include transfer learning (TL) and adversarial domain adaptation (DA) to bridge the gap between diverse data sources and operational realities [16]. Wang et al. [17] introduced an adversarial transfer network for bearing fault diagnosis, achieving superior accuracy through a hybrid loss function, albeit at a high computational cost that may limit real-time applicability. Xu et al. [18] utilised a domain-adversarial neural network combined with short-time Fourier transform for reliable classification of rotor faults with minimal labelled data, performing well even in noisy environments. However, reliance on accurate simulations and high computational demands could restrict its implementation. Liu et al. [19] developed a method using adversarial discriminative DA to improve fault diagnosis in gas turbine rotors under OVs, effectively reducing domain offsets. The complexity of this method and the requirement for validation against new fault types remain as significant challenges. Autoencoder-based alignment strategies have also been proposed for intelligent fault diagnosis of rotating machines, such as the stacked autoencoder based partial adversarial domain adaptation (SPADA) model of Liu et al. [20]. While such methods leverage deep autoencoders to extract representative features from source and target domains and utilise adversarial training for domain adaptation, they may still encounter challenges in explicitly maximising cross-domain statistical dependence, especially under complex or partial domain adaptation scenarios, and may require careful weighting or adaptation to avoid negative transfer.

UDA, an extension of TL, has become critical, particularly in scenarios where acquiring labelled data is costly and time-consuming. UDA aims to transfer knowledge by learning domain-invariant features or by translating source data to resemble the target domain, enhancing fault diagnosis robustness in mechanical systems [21]. Recent techniques like adversarial training and self-training have proved effective in extracting these invariant features, improving model generalisation across different domains [7,22–24]. Challenges in TL include limitations imposed by the quality of source data. Simulations, which extend datasets when real data is limited, play a complementary role. However, their reliance on simplified models, as explored by Han et al. [25] and Khan et al. [26], raises concerns about their generalisability and precision in complex real-world settings.

Domain alignment techniques have also been developed to address distribution shifts. Ma et al. [27] introduced a domain distribution alignment network using an enhanced joint distribution adaptation mechanism with adaptive softmax, achieving high accuracy on experimental datasets. However, the reliance on simplified models and extended target datasets suggests a need for improvements in capturing real-world complexities. Xiang et al. [28] presented a classifier-constrained domain adaptation network for rotor fault diagnosis that combines simulated and experimental data under varied conditions. While it achieves high accuracy, its dependence on accurate rotor parameters poses practical challenges. Wang et al. [29] proposed a weight-based dual DA model focusing on noisy environments and domain shifts. This model, using local maximum mean discrepancy (MMD) and batch nuclear-norm maximisation for alignment, assumes sufficient target

domain observations and is primarily effective in intra-machine scenarios, with challenges in cross-machine generalisation and hyperparameter tuning. Similarly, Xiao et al. [30] developed a multi-label deep TL structure integrating multi-stage MMD and manifold regularisation for domain alignment, leveraging annotated data from varied conditions to improve accuracy with limited labelled target data. Nevertheless, this approach faces challenges with kernel parameter sensitivity and extreme domain shifts and sensor variations. Additionally, advanced graph-based frameworks for rotating machinery such as the edge-enhanced self-supervised attention graph convolutional network and the multi-sensor multi-head graph attention network, both recently validated on axial flow pumps, have also been proposed [31,32].

While various approaches have advanced the field of TL in rotating system fault diagnosis, significant challenges remain in achieving reliable performance with limited target domain observations, ensuring generalisation across disparate machines and fault types, and managing the scarcity of labelled data. Methods based on short-time Fourier transform or autoencoder-driven DA offer certain advantages; however, they often struggle to fully capture transient, highly localised features or to ensure robust alignment between source and target domains, particularly under complex or severe distribution shifts. Addressing these persistent limitations is essential for the development of truly generalisable diagnostic frameworks. In this work, the WaveCORAL-DCCA framework was specifically devised to tackle these issues, leveraging the multi-scale capabilities of wavelet transformation (WT) in combination with an enhanced DCCA technique. This integration provides a robust solution that markedly improves the model's ability to generalise across a wide range of operational scenarios, even in the presence of severely limited unlabelled target data. Importantly, industrial rotor fault diagnosis is frequently conducted using data acquired under varying conditions, involving differences in sensor types (such as acceleration and displacement), acquisition methods, and operational environments. The challenge posed by such real-world variability remains significant. To address this, the WaveCORAL-DCCA framework has been rigorously evaluated across highly heterogeneous datasets, thereby demonstrating its capacity for effective generalisation and for the transfer of diagnostic knowledge between fundamentally different data sources.

The main contribution of this paper includes the following:

- Introduced a novel framework that integrates wavelet-based feature extraction with a DCCA network enhanced by the CORAL loss function. This framework effectively mitigates domain shifts and OVs, offering a robust solution for rotor fault diagnosis;
- Applied label smoothing to the entropy loss function, improving classification performance by mitigating overconfident predictions, enhancing generalisation, and increasing robustness to noisy labels;
- Achieved high diagnostic accuracy even with limited unlabelled target domain data, addressing the significant challenge of data scarcity in real-world industrial applications;
- Demonstrated the framework's broad applicability through experiments on both experimental and simulated rotor system datasets, showing its effectiveness across different machines and fault types.

The remainder of this paper is structured as follows: Section 2 outlines the proposed framework and covers the necessary preliminaries. Section 3 details the case studies conducted. Section 4 presents and discusses the results, while Section 5 concludes by addressing limitations and offering suggestions for future research directions.

2. Materials and Methods

The WaveCORAL-DCCA framework comprises a combination of feature extraction, alignment, and classification phases for effective fault diagnosis in rotating systems, which are detailed in the subsequent sections. The framework operates by: (i) using WT to extract essential time–frequency features from vibration signals, (ii) employing an enhanced DCCA network with CORAL loss to align feature distributions across domains by matching their second-order statistics, and (iii) classifying faults based on the aligned features, diagnosing normal, unbalanced, and misaligned rotor conditions, even with limited target domain data.

2.1. Wavelet Transformation

WT is a powerful signal processing technique that breaks down signals into their frequency components, each tailored to its respective scale. By applying both high-pass ($h_i[n]$) and low-pass filters ($g_i[n]$), this method allows for precise analysis of both stationary and non-stationary signals [33]. To perform a discrete WT on a vibration signal $X[n]$, two essential functions are introduced: the scaling function $\phi(t)$ and the wavelet function $\psi(t)$, as well as the desired level of decomposition. The approximation coefficients at scale j, which capture the low-frequency components, are derived by passing the signal through the low-pass filter. The new approximation coefficients at scale $(j+1)$ are computed as follows:

$$cA_{j+1}[m] = \sum_n cA_j[n] \cdot g[n-2m] \tag{1}$$

In this process, m represents the index at the next scale $(j+1)$, where new approximation coefficients are computed. m indicates the position of the coefficients at the next level, while n is the summation index over the current scale's approximation coefficients, which capture low-frequency components. Similarly, the detail coefficients (cD), derived from the high-pass filters, are computed in the same way. Approximation coefficients move to the next level, while detail coefficients are useful for tasks like pattern recognition and anomaly detection. Figure 1 shows the decomposition of signal $X[n]$ up to level 3. The detail coefficients at scale $(j+1)$ are calculated as:

$$cD_{j+1}[m] = \sum_n cA_j[n] \cdot h[n-2m] \tag{2}$$

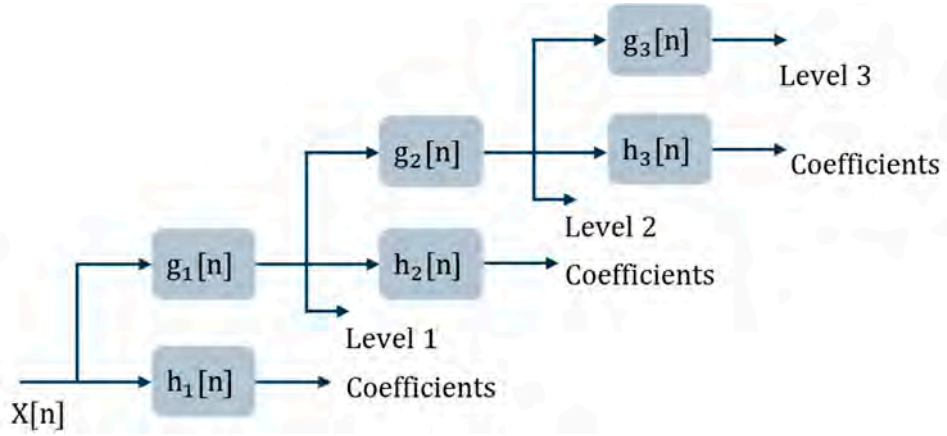


Figure 1. Illustration of signal decomposition utilising WT.

2.2. CORAL Loss

Correlation alignment (CORAL) loss serves as a metric to gauge alignment between two domains by measuring the difference in their feature covariance matrices. It is used in DA to align data distributions of source and target domains, reducing statistical variations

for better generalisation. By minimising this loss, the learned features become more consistent across domains, improving prediction accuracy for the target domain. The core concept behind CORAL loss is addressing domain shift by aligning the second-order statistics, specifically minimising the Frobenius norm of the difference between the source and target covariance matrices; the CORAL loss function can be represented as:

$$L_{CORAL} = \frac{1}{4d^2}\|C_S - C_T\|_F^2 \tag{3}$$

in which C_S and C_T are the covariance matrices of the source and target features, d is the feature dimensionality, and $\|\cdot\|_F$ denotes the Frobenius norm. The CORAL loss allows for quantitative comparison between DA methods by assessing which method minimises the discrepancy between domains. While a lower CORAL loss suggests better distribution alignment, it should be evaluated alongside other metrics, such as classification accuracy on the target domain, to ensure that alignment improves or maintains task-specific performance; CORAL loss can be calculated on a class-wise basis or as an aggregate. The aggregate CORAL loss is typically more significant as it reflects the overall alignment across all classes, ensuring that the adaptation generalises well across the domain rather than overfitting to specific classes.

2.3. Deep Canonical Correlation Analysis

Designed to project two sets of variables, or "views," into a shared subspace, deep canonical correlation analysis (DCCA) captures the key relationships between them. In vibration signal analysis, it is beneficial for reducing the dimensionality of data while preserving crucial information from both the source domain (view 1) and the target domain (view 2). The objective of DCCA is to maximise the correlation between these two sets of data by projecting them into a common space, improving the transfer of knowledge between the domains. Unlike traditional canonical correlation analysis (CCA), which operates on a single layer, DCCA uses multiple deep layers to capture more complex relationships between the views. In this context, the extracted features from the source and target domains are represented as vectors S (from the source) and T (from the target). DCCA projects these feature vectors through deep layers, aligning them in a way that preserves the most informative correlations, enhancing the model's ability to generalise across different conditions.

The deep structure of DCCA allows for a richer representation of the data, which is essential in applications like vibration analysis, where subtle signal variations carry significant meaning. Features from the source and target domains are transformed into canonical variates U and Y, which are linear combinations of their respective domains, using linear transformations W_S for the source and W_T for the target. This ensures that DCCA provides a robust and nuanced framework for analysing vibration signals.

$$S = \begin{pmatrix} F_{S1} \\ F_{S2} \\ \vdots \\ F_{Sp} \end{pmatrix},\ T = \begin{pmatrix} F_{T1} \\ F_{T2} \\ \vdots \\ F_{Tq} \end{pmatrix},\ W_S = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \dots & a_{pp} \end{bmatrix},\ W_T = \begin{bmatrix} b_{11} & \dots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{q1} & \dots & b_{qq} \end{bmatrix} \tag{4}$$
$$U = W_S S,\ Y = W_T T$$

For the purpose of computational efficiency, the values of p and q should be selected to satisfy $p \leq q$. As an illustration, (U_1, Y_1) is the initial canonical variate pair that can be shown as the following relationships:

$$U_1 = a_{11}F_{S1} + a_{12}F_{S2} + \ldots + a_{1p}F_{Sp}$$
$$Y_1 = b_{11}F_{T1} + b_{12}F_{T2} + \ldots + b_{1q}F_{Tq} \quad (5)$$

The aim of this process is to find linear combinations that maximise the correlations between the components of each canonical variate pair. This is achieved through an optimisation procedure, where the objective function can be expressed as $Max\|corr(U, Y)\|$. The correlation, $corr(U, Y)$, is introduced as:

$$L_{CORAL} = corr(U, Y) = \frac{cov(U, Y)}{\sqrt{var(U)var(Y)}} \frac{1}{4d^2}\|C_S - C_T\|_F^2 \quad (6)$$

In this context, *var* refers to the variance of each canonical variate and $cov(U, Y)$ represents the covariance of U and Y. In DCCA, the loss function typically minimised (as maximising correlation can be reformulated as a minimisation task) is often the negative of the canonical correlation or a modified version that incorporates regularisation terms.

$$L_{DCCA} = -corr\left(W_S^T S, W_T^T T\right) + regularization\ terms \quad (7)$$

The features processed from both views are combined to generate the final feature representation. This representation gains from redundancy analysis, which assesses the extent to which the variance in one set of variables is accounted for by the canonical variates from the other set, increasing the model's efficiency in handling new or related tasks. The architecture of DCCA is shown in schematic form in Figure 2. In the upcoming sections, this standard version of DCCA will be referred to as the original DCCA, and its performance will be compared with that of an improved DCCA-based framework.

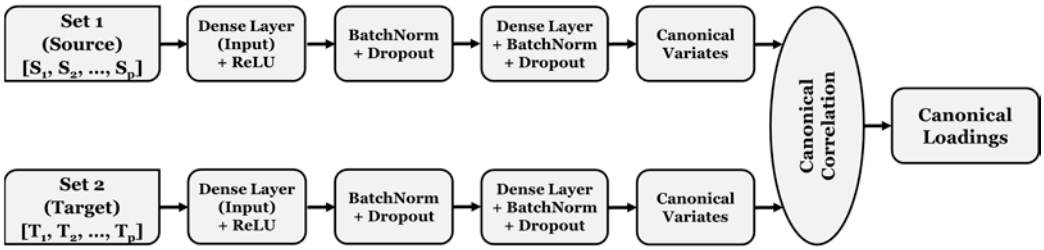


Figure 2. Process diagram of DCCA.

2.4. WaveCORAL-DCCA

WaveCORAL-DCCA leverages an enhanced DCCA model that integrates a modified loss function, incorporates overfitting prevention techniques, and includes a feature extraction phase prior to the DA stage. These enhancements significantly improve the model's performance and are detailed as follows: WaveCORAL-DCCA begins with a WT stage as the feature extraction stage. Since one of the released datasets ('S_E_D') was published in the form of WT coefficients, this stage was implemented separately for the remaining datasets using MATLAB® R2024a before feeding them into the DA and classification structure. The DA framework employs multiple standard techniques such as normalisation, batch normalisation, and dropout layers to enhance model stability and prevent overfitting. However, the primary focus of WaveCORAL-DCCA lies in the two distinct loss functions used during training, which govern the model's DA and classification tasks.

The first training process is the DA phase, where the model aligns the feature representations of the source and target domains using CORAL loss. This loss function ensures that the covariance matrices of the source and target domain features are aligned, which

minimises the domain shift and facilitates better generalisation to the target domain, so the DA loss can be introduced as:

$$L_{Domain\ adaptation} = L_{CORAL} \tag{8}$$

The second training process focuses on the classification phase, where the model classifies the DCCA-transformed features from the source domain. For this, a multi-layer perceptron (MLP) model was designed, consisting of three fully connected layers. The MLP model includes two hidden layers with ReLU activation functions, followed by batch normalisation and dropout layers to prevent overfitting. The final layer produces predictions for the target classes. The specific hyperparameters for this classifier include the sizes of the hidden layers and the dropout rates. In this phase, the total loss function is a combination of two components: the classification loss (implemented via label smoothing) and the entropy loss. Label smoothing modifies the cross-entropy loss by distributing a portion of probability mass to all classes, reducing the model's overconfidence in its predictions. The entropy loss penalises low-entropy (overconfident) predictions, encouraging the model to maintain a level of uncertainty and improving its generalisation capabilities.

$$L_{Classification} = L_{Label\ smoothing} + \epsilon \times L_{Entropy} \tag{9}$$

where ϵ is a coefficient that controls the assigned weight to the entropy loss. This approach not only ensures that the classifier attains both precision and generalisation, but also in the DA phase, CORAL loss governs the alignment of feature distributions across the two domains. By leveraging these two distinct training processes, each guided by its own loss function, the model is able to address both feature alignment and classification challenges effectively, resulting in improved performance on the target domain. It is important to highlight that the target domain data were included in the training phase of WaveCORAL-DCCA without labels, and these data played no role in training the classifier. As a result, the framework operates as a UDA structure. Figure 3 illustrates the overall structure of WaveCORAL-DCCA and the fault diagnosis framework.

In the grid search process for finding the most optimum hyperparameters based on the accuracy of the classification of the test set, as illustrated in Figure 3, various hyperparameters are involved, such as learning rates, dropout rate, batch size, hidden dimension, output dimension, number of epochs, label smoothing coefficients, and the entropy coefficient. A more detailed of the designed fault diagnosis procedure is displayed in Algorithm 1, where the WT process was not involved in this code.

Algorithm 1. Pseudocode of WaveCORAL-DCCA and the classification phase.

Input: Datasets from source dataset S (S_{train}, S_{val}) and target dataset T (T_{train}, T_{test}), wavelet level (k), learning rates (η), dropout rate (δ), batch size, hidden dimension (d_{hid}), output dimension (d_{out}), number of epochs (E), label smoothing (λ), entropy coefficient (ϵ);

Output: Trained domain adaptation and classifier parameters (θ_{DCCA} and $\theta_{classifier}$), evaluation metrics (accuracy Acc, confusion matrix CM);

Feature Extraction using WT

1 Load source S (S_{train}, S_{val}) and target T (T_{train}, T_{test}) datasets,

2 Apply WT on S and T to extract detail coefficients at level k (cD_k);

Algorithm 1. *Cont.*

Domain adaptation using DCCA with CORAL loss

3 Initialise DCCA model with input dimension (d_{in}), hidden dimension (d_{hid}), output dimension (d_{out}), dropout rate (δ), separate weights for source and target features,

4 For each mini-batch B_j in S_{train} and T_{train}:

- Perform a forward pass through the DCCA model for both source (cD_{Sk}) and target (cD_{Tk}) domains;

5 Initialise CORAL loss function;

6 For each epoch $i = 1, 2, \ldots, E$:

7 For each mini-batch B_j in S_{train} and T_{train}:

8 Pass source and target features through DCCA and extract feature representations,

9 Compute CORAL loss (L_{CORAL}) between source and target domains,

10 Update model parameters (θ_{DCCA}) using gradient descent:

- $\theta_{DCCA} \leftarrow \theta_{DCCA} - \eta \times \nabla_{\theta_{DCCA}} L_{\text{CORAL}}$;

Train Classifier on Extracted Features

11 Initialise classifier ($\theta_{\text{classifier}}$) with:

- Hidden layers,
- Output layer,
- Dropout rate (δ);

12 For each epoch $i = 1, 2, \ldots, E$:

- For each mini-batch B_j in S_{train}:

13 Extract DCCA features from the source domain,

14 Perform a forward pass through the classifier to obtain predictions,

15 Compute classification loss using label smoothing (λ):

- $L_{classification} = LabelSmoothingLoss\left(p\left(y\middle|\theta_{classifier}\right), y_{true}\right)$;

16 Compute entropy loss $(L_{entropy})$:

- $L_{entropy} = -\sum p\left(y\middle|\theta_{classifier}\right)\log p\left(y\middle|\theta_{classifier}\right)$,

17 Compute total loss (L_{total}) :

- $L_{total} = L_{classification} + \epsilon \times L_{entropy}$,

18 Update model parameters $\left(\theta_{classifier}\right)$ using gradient descent:

- $\theta_{classifier} \leftarrow \theta_{classifier} - \eta \times \nabla_{\theta_{classifier}}(L_{total})$,

Evaluation on Target Domain

19 Evaluate the classifier on T_{test},

- Extract DCCA features from T_{test} and compute test accuracy:

$$Acc = \sum\left(1\left(y_i = argmax\left(p\left(y \mid \theta_{classifier}\right)\right)\right)\right) / \left|T_{test}\right|;$$

20 Generate confusion matrix $CM = \left[cm_{ij}\right]$;

Result: Return trained model parameters θ_{DCCA} and $\theta_{classifier}$, test accuracy Acc, confusion matrix CM.

A sequential training approach, in which feature alignment and classification are performed in two distinct phases, was adopted in this work. This structure has been shown to offer greater robustness and interpretability in practical settings, particularly when target domain data are scarce, and has been widely employed in prior DA studies [34].

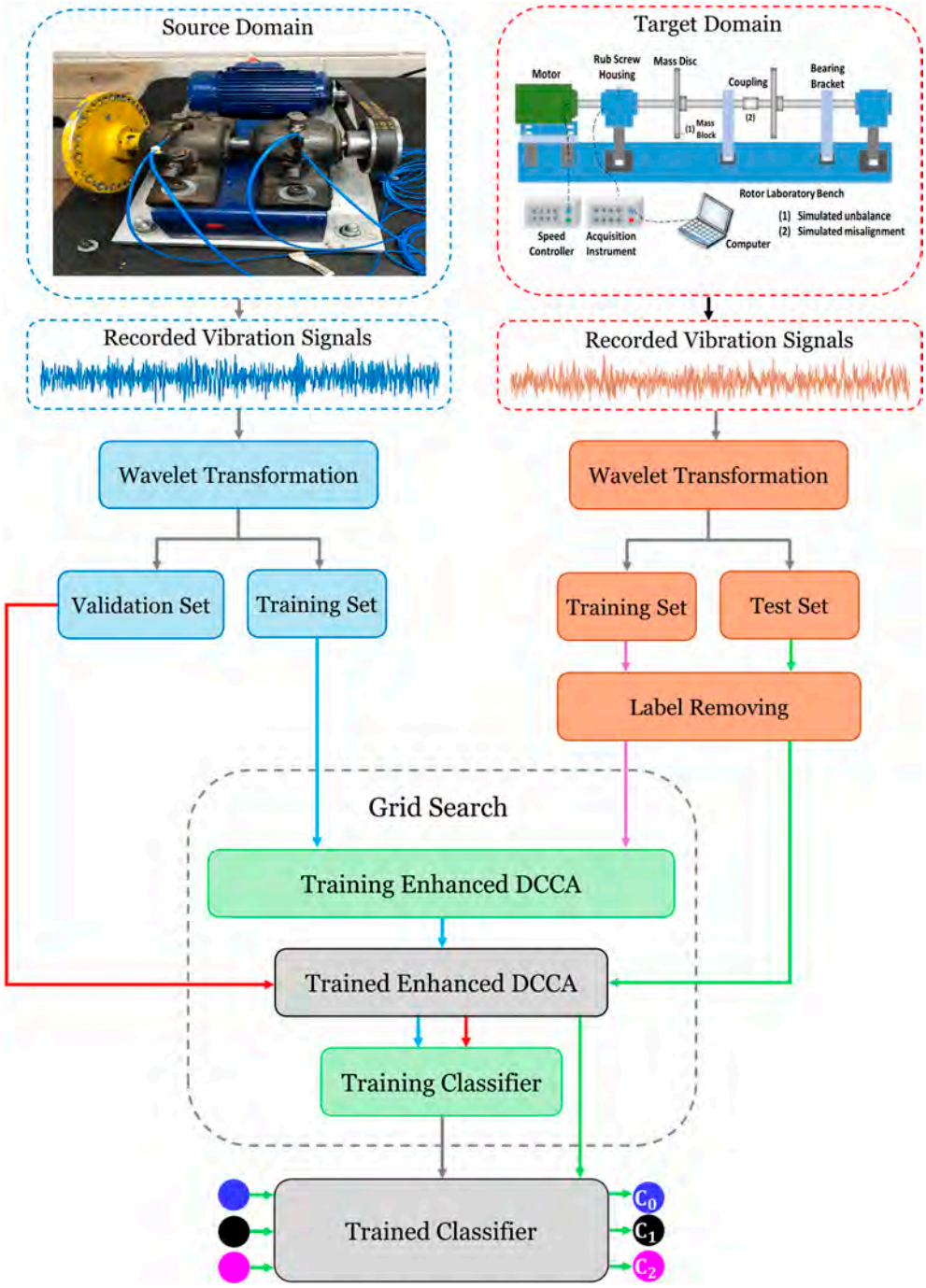


Figure 3. Workflow of WaveCORAL-DCCA and the classification phase.

3. Case Studies

In this work, to assess the effectiveness of WaveCORAL-DCCA in fault diagnosis, three datasets were employed, where two of them are the experimental sets released by two various research teams and the one remaining was generated by the authors through numerical simulation of a rotor-bearing-disc in MATLAB® R2024a utilising the finite element (FE) method; these datasets are elaborated in the following.

3.1. Large Experimental Dataset 'L_E_D'

A recent study [35] evaluated the performance of rotary machinery under various operational conditions, including imbalance, misalignment, mechanical looseness, and normal functioning. The generated dataset, referred to as the source domain data, is foun-

dational for ongoing research. The experimental setup included a motor with a frequency inverter, bearings, pulleys, a belt, and a rotor. Data was captured using accelerometers placed at strategic points around the machine, recording 420 signals per experiment, each containing 25,000 data points, sampled at 25 kHz. A total of 8400 signals were collected, 2100 per condition, with each observation producing four distinct vibration signals; the dataset, named 'L_E_D', reflects fives trials for each scenario. The data was collected in a controlled lab environment, with the test rig disassembled and reassembled before each trial to allow for fault introduction, simulating real-world conditions. Data acquisition was managed by a Python script, and accelerometers were calibrated accordingly. The rotational speed was maintained at 1772 rpm; Figure 4 reveals this test rig.

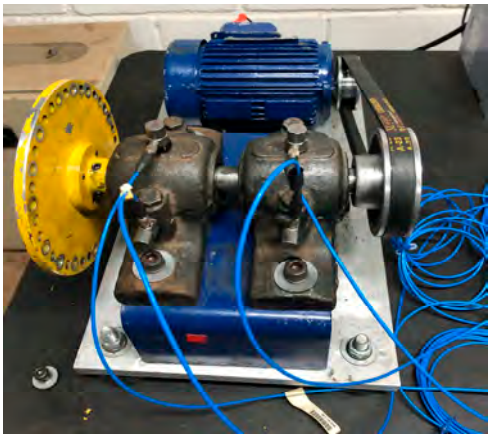


Figure 4. Setup of the source domain testing apparatus.

3.2. Small Experimental Dataset 'S_E_D'

The second experimental rotor dataset, referred to as 'S_E_D' in this study, was provided by Wuhan University [36]. It features vibration signals from a lab-based rotating machinery system consisting of a rotor test platform, speed controller, front-end processor, and computer (Figure 5). Vibration data were collected using two eddy current sensors positioned above the rotor to detect four rotor states: normal, contact-rubbing, unbalance, and misalignment. Each state includes 45 observations, totalling 180 instances, with each sample containing 2048 data points, captured at 1200 rpm and 2048 Hz. Pre-processing involved denoising with wavelet thresholding, and the data were converted into a 2D matrix format. The system used a DC motor with a 1.95A current, 148 W output, and rotor measuring 10 mm in diameter and 850 mm in length. Data acquisition was overseen by the GTS3-TG simulator, which forwarded signals to the front-end processor for amplification before computer analysis. This dataset has been widely used in condition monitoring research [37–41].

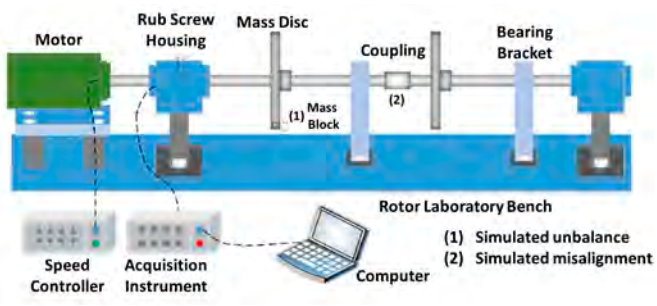
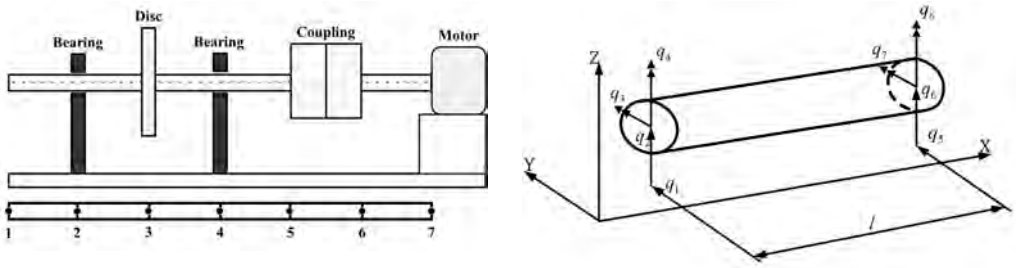


Figure 5. Setup of 'S_E_D' testing rig.

3.3. Numerical Dataset 'N_D'

To produce a numerical dataset, the FE method was implemented in MATLAB® R2024a to model a rotor-bearing-disc. The FEM model of the rotor includes the disc, coupling, bearing supports, two shafts, and electrical motor. Figure 6a shows a graphical depiction of the modelled system. In the FE formulation, each component is represented by a two-node element, as depicted in Figure 6b, with each nodal point having four degrees of freedom: rotations about the Y and Z axes and translations along these axes. In this model, the length of the smaller shaft as well as the length of the coupling were selected as the length of each element, and the motor was simulated as an alternative tortional torque at node 7. To simulate the unbalancing in the system, the centre of gravity of the disc was changed.



(**a**) (**b**)

Figure 6. (**a**) Diagram of the rotor-bearing-disc system with coupling; (**b**) FE model and coordinate framework of the shaft–beam component.

To model the desired damages in the system, i.e., unbalanced, and in the misaligned rotor, the right-hand side of the following equation of motion ought to be changed.

$$M\{\ddot{q}\} + C\{\dot{q}\} + K\{q\} = f(t) \tag{10}$$

The mass (M), damping (C), and stiffness (K) matrices, each 4 × 4 per node, are calculated based on [42]. The force vector ($f(t)$) primarily reflects the motor's external force but adjusts for system faults. For the shaft (six elements), disc, bearings, and coupling, a global 28 × 28 matrix is assembled using a connectivity table for the seven nodes. Journal bearings require additional damping and stiffness as per [42], and the coupling element has a distinct stiffness matrix compared with adjacent elements. An unbalanced rotor can be modelled through the introduction of the excess force, resulting from the centrifugal force as:

$$F_u = m_d e\omega^2 \tag{11}$$

To represent the harmonic essence of the implemented unbalancing force, it can be re-written in the Y and Z directions as:

$$\begin{aligned} f_u(t) &= [F_Y \; F_Z \; F_{\theta Y} \; F_{\theta Z}]^T \\ F_Y &= m_d e\omega^2 cos(\omega t - \varphi_u) \\ F_Z &= m_d e\omega^2 sin(\omega t - \varphi_u) \end{aligned} \tag{12}$$

in which m_d, e, ω, and φ_u show the mass of the disc, distance of the centre of gravity from the centre of geometry (eccentricity), angular velocity of the disc, and phase angle of the unbalancing force, respectively. In the system shown in Figure 6a, the unbalance force is applied at the disc node (e.g., node 4), while the misalignment force is applied at the coupling nodes (nodes 5 and 6). To model the misalignment force accurately, the forces on

these two nodes must be considered separately, as they differ; Figure 7 presents the nodal forces and moments, for which parallel misalignment can be expressed as follows:

$$\begin{gathered} MY_1 = T_q sin\theta_1 + K_b \varnothing_1,\ MY_2 = T_q sin\theta_2 - K_b \varnothing_2 \\ MZ_1 = T_q sin\varnothing_1 - K_b \theta_1,\ MZ_2 = T_q sin\varnothing_2 + K_b \theta_2 \\ FY_1 = \frac{(-MZ_1 - MZ_2)}{X_3},\ FY_2 = -FY_1 \\ FZ_1 = \frac{(MY_1 + MY_2)}{X_3},\ FZ_2 = FZ_1 \end{gathered} \tag{13}$$

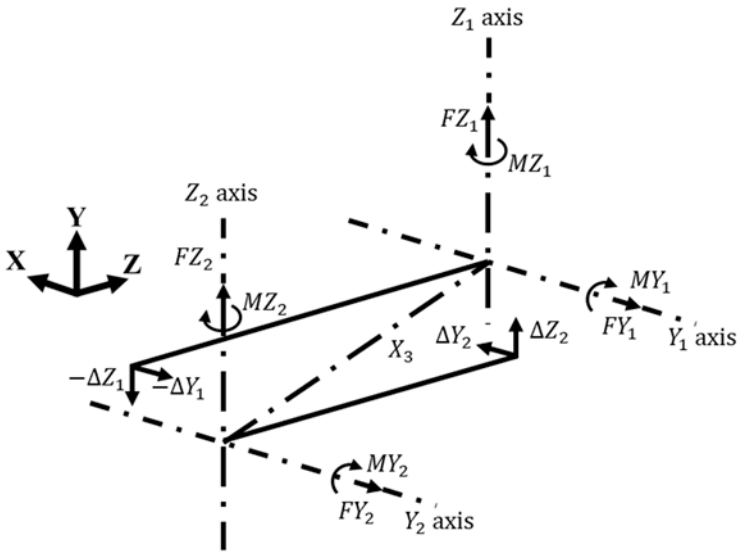


Figure 7. Coordinate system for coupling in the case of parallel misalignment.

Here, K_b represents the flexure coupling's bending spring rate per diaphragm or disc pack, and T_q denotes the external torsional torque; the values for θ_1, $\varnothing_1$, θ_2, and $\varnothing_2$ can be determined as:

$$\begin{gathered} \theta_1 = sin^{-1}\left(\frac{\Delta Y_1}{X_3}\right),\ \varnothing_1 = sin^{-1}\left(\frac{\Delta Z_1}{X_3}\right) \\ \theta_2 = sin^{-1}\left(\frac{\Delta Y_2}{X_3}\right),\ \varnothing_2 = sin^{-1}\left(\frac{\Delta Z_2}{X_3}\right) \end{gathered} \tag{14}$$

In these equations, ΔY_1 and ΔZ_1 represent the misalignment severities in the Y and Z directions at node 6, while ΔY_2 and ΔZ_2 correspond to the misalignment severities at node 5. X_3 is the centre of articulation. The nodal forces due to misalignment on the left (node 5) and right (node 6) sides of the element containing the coupling, denoted as Q_m^1 and Q_m^2, are expressed as:

$$\begin{gathered} Q_m^1 = \begin{Bmatrix} FY_1 sin\omega t + FY_1 sin2\omega t + FY_1 sin3\omega t + FY_1 sin4\omega t \\ FZ_1 cos\omega t + FZ_1 cos2\omega t + FZ_1 cos3\omega t + FZ_1 cos4\omega t \\ 0 \\ 0 \end{Bmatrix} \\ Q_m^2 = \begin{Bmatrix} FY_2 sin\omega t + FY_2 sin2\omega t + FY_2 sin3\omega t + FY_2 sin4\omega t \\ FZ_2 cos\omega t + FZ_2 cos2\omega t + FZ_2 cos3\omega t + FZ_2 cos4\omega t \\ 0 \\ 0 \end{Bmatrix} \end{gathered} \tag{15}$$

The dataset was constructed at one specific rotational speed, i.e., 100 rad/s. The experimental protocol involved executing 45 iterations; each iteration yielded a single observation relevant to the specific health condition being monitored, resulting in a total of 135 observations (45 for each of the three health scenarios); this dataset is referred as 'N_D'.

It is worth noting that only the stationary portions of the signals were saved, specifically data points ranging between 10,003 and 12,050. As such, each observation encompasses 2048 data points; this selection was informed by the data point consistency observed in the 'S_E_D' dataset. To solve the equation of motions, the Houbolt time marching technique was employed with a time increment of 0.9766 milliseconds; Table 1 outlines the system parameters employed to simulate the rotor system.

Table 1. Characteristics of the system being modelled.

Parameter	Magnitude	Parameter	Magnitude
Disc mass	2 kg	Bearing stiffness	5×10^7 N/m
Disc eccentricity	0.5 cm	Bearing damping	350 N·S/m
Diametral moment of inertia of the disc	6.5×10^{-4} kg·m^2	Motor phase angle	3.14 rad
Unbalancing phase angle	0 rad	Motor-side pulley's radius	1.9 cm
Shaft diameter	5 cm	Torsional torque	2.55 N·m
Larger shaft length	50 cm	Centre of articulation	2 cm
Smaller shaft length	10 cm	Flexure coupling bending	1600 N·m/rad
Shaft density	7800 kg/m^3	Misalignment in the Y direction at node 6	0.2 cm
Shaft Young's modulus	2.08×10^{11} N/m^2	Phase angle of misalignment	0.79 rad

The code was designed to ensure each loop produced unique observations despite constant system parameters. A random seed, initialised with the current time, generated different results in every iteration. Small random variations were added to the rotor system's initial conditions using normally distributed noise, creating slight differences at the start of each simulation. Random noise was also applied to the force resulting from the motor, mimicking real-world unpredictability and leading to distinct dynamic responses in each run.

3.4. Summary of the Case Studies

This study examined three health conditions across all datasets: normal operation, unbalanced disc, and misaligned rotor systems, which were assigned the labels 0, 1, and 2, respectively. These datasets, referred to as 'L_E_D', 'S_E_D', and 'N_D', are outlined in Table 2, which provides details on the number of observations and the labels used in the classification section.

Table 2. Overview of datasets with health conditions and labels.

Health Scenario	Number of Observations			Label
	'L_E_D'	'S_E_D'	'N_D'	
Normal	2100	45	45	0
Unbalanced	2100	45	45	1
Misaligned	2100	45	45	2

For each target domain dataset, 15% of the samples were reserved for testing and 11% for validation in a class-wise balanced manner.

It is noteworthy that the three datasets vary in both sensor type and physical quantity: the 'L_E_D' dataset comprises acceleration measurements from accelerometers, 'S_E_D'

provides displacement data from eddy current sensors, and the 'N_D' dataset is numerically simulated displacement. This deliberate selection ensures a challenging evaluation scenario, reflecting the diverse conditions encountered in practical industrial monitoring.

4. Results

The effectiveness of WaveCORAL-DCCA in detecting malfunctions in rotating machines was evaluated in two scenarios: (i) using 'L_E_D' as the source domain and 'S_E_D' as the target domain; and (ii) using 'L_E_D' as the source domain and 'N_D' as the target domain. Additionally, two other classification frameworks, one without a DA stage and another using the original DCCA, were assessed using the same source and target domain sets to compare their performance with WaveCORAL-DCCA.

In configuring the feature extraction stage, the Daubechies-4 (db4) mother wavelet function was selected for the discrete WT because of its effectiveness in capturing transient and non-stationary features that are crucial for rotor fault diagnosis [43]. The decomposition level was set to 3 since one of the experimental datasets ('S_E_D') was originally published in the form of level 3 WT coefficients; thus, for methodological consistency, the same decomposition level was applied to all datasets. This approach achieves a balance between capturing essential fault-related information and maintaining computational efficiency, in line with established practices in vibration-based machinery diagnostics.

Furthermore, the hidden layer sizes and output dimension in the DCCA network were determined by considering the complexity and dimensionality of the extracted wavelet features, as well as the expected diversity of fault signatures present in the datasets. These parameters were set to ensure the network possessed sufficient capacity to model the underlying data structures without excessive risk of overfitting, keeping in line with established engineering knowledge and prior studies on fault diagnosis network design.

4.1. Fault Diagnosis Without DA

Initially, the damage detection task was carried out without the designed DA structure. To achieve this, the two mentioned scenarios were implemented, and the extracted WT coefficients from the training phase of 'L_E_D' were used to train the classifier. Given that the designed classifier within the DA framework is a simple three-layer MLP, which may not be sufficient for complex tasks, a one-dimensional CNNs model was constructed with a more robust architecture. For the MLP, the classifier utilised two hidden layers with 512 and 128 neurons, respectively, each followed by batch normalisation and dropout with a rate of 0.3 to mitigate overfitting. The MLP was trained using the Adam optimiser with a learning rate of 0.001 and cross-entropy loss for 100 epochs. The architecture and hyperparameters of the MLP were tailored for multi-class classification using three output classes.

The applied CNN begins with a Conv1D layer utilising 64 filters and a kernel size of 5, followed by two additional Conv1D layers with 128 and 256 filters, respectively, to increase feature extraction capabilities. The output of the final convolutional layer was flattened and passed through a dense layer with 512 units. A softmax output layer was used for multi-class classification, and dropout with a rate of 0.5 was applied to prevent overfitting. The model was trained using a learning rate of 0.001 for a total of 100 epochs. The outcomes of the fault diagnosis for the two scenarios (CNNs and MLP) are displayed in the confusion matrices of Figure 8 and the bar chart in Figure 9. These results provide insights into the classification performance of both models across different conditions.

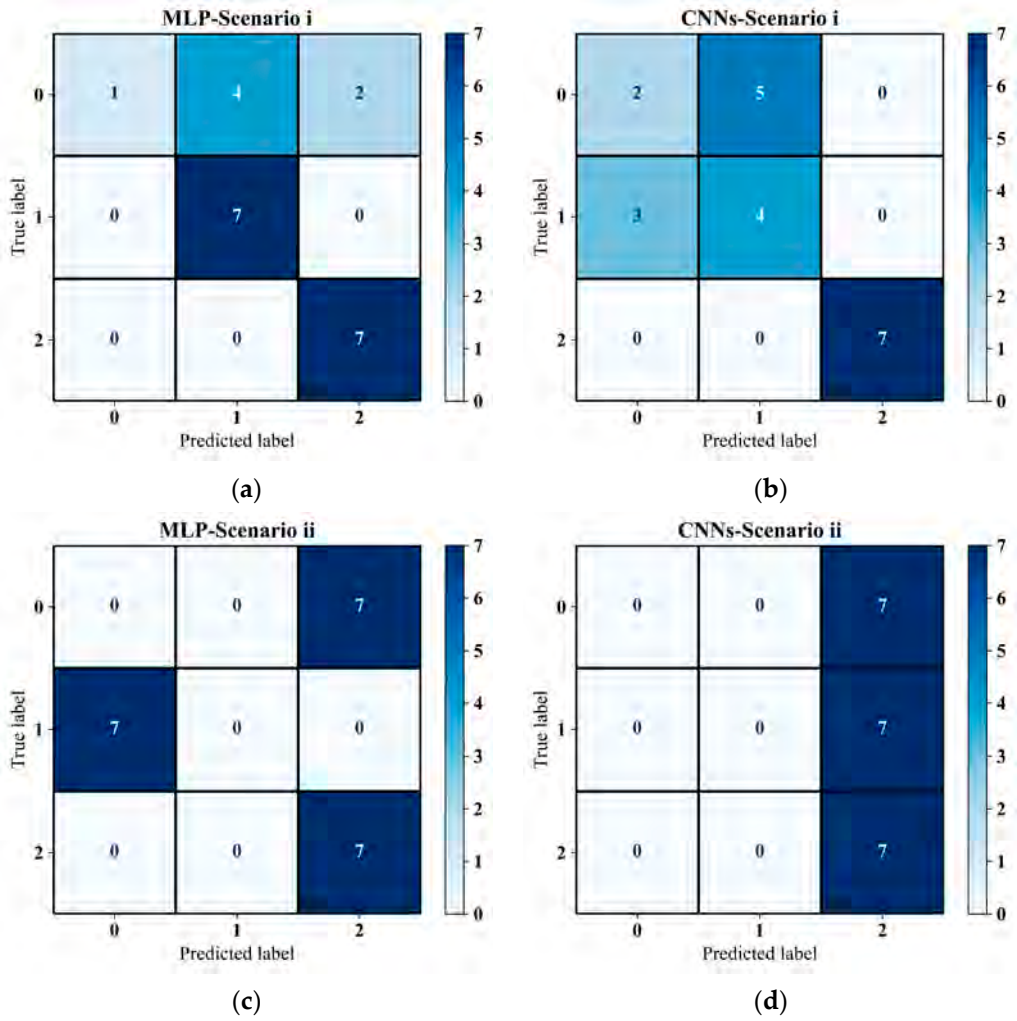


Figure 8. Confusion matrices for fault diagnosis in Scenario i: (**a**) MLP and (**b**) CNNs; and in Scenario ii: (**c**) MLP and (**d**) CNNs.

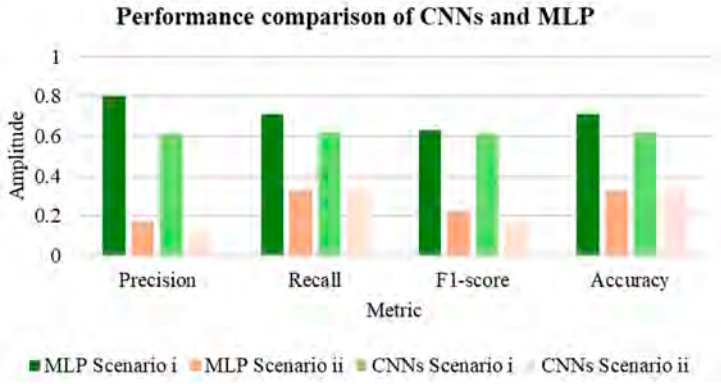
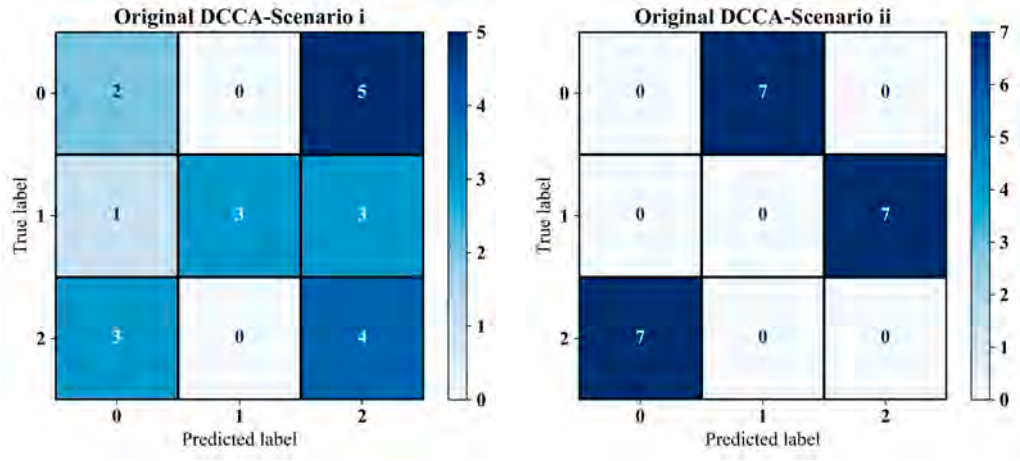


Figure 9. Performance comparison of CNNs and MLP models for fault diagnosis.

The confusion matrices in Figure 8 and the statistics in Figure 9 indicate that under Scenario i, the trained models, particularly the MLP using source domain data, performed well in diagnosing sample tests with labels 1 and 2; however, these models demonstrated lower accuracy when identifying observations with label 0. In contrast, under Scenario ii, both classifiers showed significantly deficient performance in diagnosing samples from all three classes, achieving only 33% accuracy. This considerable drop in accuracy highlights the need for a more generalised fault diagnosis framework.

4.2. Fault Diagnosis with the Original DCCA Framework

To further demonstrate the advantages of WaveCORAL-DCCA framework, the same training, validation, and testing subsets of Scenario i and Scenario ii were applied in the DCCA model referenced in Section 3.4. This implementation used MLP with hidden layers of size [1024, 512, output_dim], where the output dimension size was set to 10 for both data views. Key hyperparameters included a learning rate of 1×10^{-3} and a batch size of 64, and the model was trained for 1000 epochs. The DCCA model was optimised using the Adam optimiser, with a regularisation parameter of 1×10^{-4} being incorporated in the DCCA loss function to prevent overfitting. During training, model checkpoints were saved based on validation loss to ensure that the best-performing model was retained. Post-training, linear CCA was optionally applied to refine the learned features. The model's performance was evaluated using support vector machine classification, providing a comprehensive assessment of the model's effectiveness. The outcomes of this assessment are detailed in Figure 10 and Table 3.



(**a**) (**b**)

Figure 10. Confusion matrices of fault diagnosis using the original DCCA: (**a**) Scenario i; (**b**) Scenario ii.

Table 3. Effectiveness of the original DCCA framework under Scenario i and Scenario ii.

DA Model	Scenario	Precision	Recall	F1-Score	Accuracy
Original DCCA	i	0.56	0.43	0.44	0.43
	ii	0	0	0	0

The fault diagnosis results using the original DCCA baseline revealed its inability to handle distributional differences effectively, leading to inferior performance that was even worse than when no DA technique was applied. This poor generalisation, especially the complete failure in Scenario ii, stems from DCCA's limited feature expressiveness and the lack of explicit domain alignment. Without additional regularisation or alignment constraints, DCCA fails to reconcile substantial cross-domain discrepancies.

4.3. Fault Diagnosis with WaveCORAL-DCCA

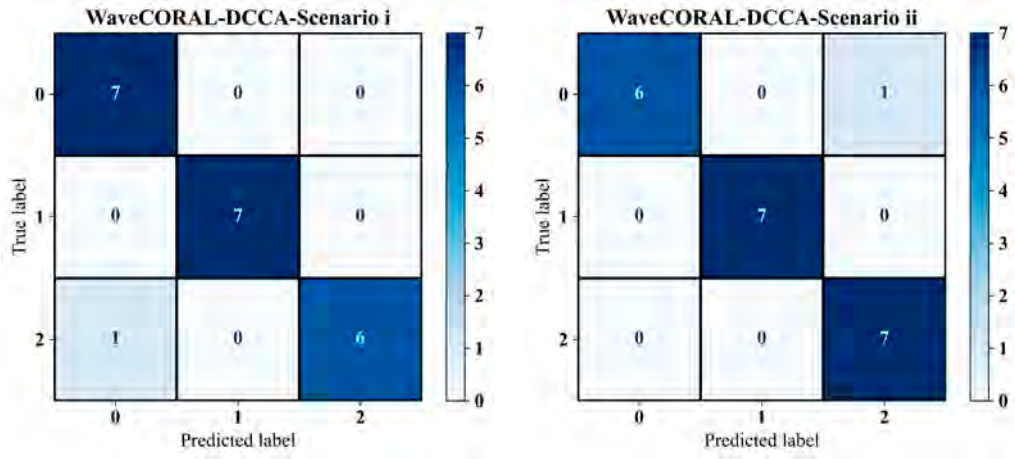
The previous sections showed the requirement for applying a robust DA more specifically on the assumption of Scenario ii, where the fault detection framework without applying a DA as well as employing the original DCCA structure presented very inferior accuracies. To this end, the designed framework (WaveCORAL-DCCA) was performed in combination with a grid search for discovering the most optimum hyperparameters

based on the test phase of the target domain; Table 4 presents the discrete values of the hyperparameters used in the grid search.

Table 4. Grid search hyperparameter values for the DA framework.

No.	Hyperparameter	Value(s)	No.	Hyperparameter	Value(s)
1	Output dimension of DCCA	16, 32	6	Number of epochs in DCCA	3, 10
2	Hidden dimension of neural networks	16, 32	7	Coefficient of label smoothing	0.01, 0.1, 1
3	Learning rate for DCCA	1×10^{-4}, 1×10^{-3}	8	Coefficient of entropy loss	0.01, 0.1, 1
4	Batch size	16, 32	9	Random seed	42
5	Dropout rate in DCCA	0.3, 0.4	10	Optimiser for DCCA	Adam

A total of 576 hyperparameter combinations, listed in Table 4, were assessed to analyse their impact on WaveCORAL-DCCA framework's performance; the malfunction detection results showed that the framework is overly sensitive to changes in these hyperparameters. As a result, it is necessary to use either a grid search or an advanced optimisation algorithm to find the best hyperparameter set. Figure 11 and Table 5 present the malfunction diagnosis results for Scenario i and Scenario ii, using the optimal hyperparameters identified for each scenario.



(**a**) (**b**)

Figure 11. Confusion matrices of fault diagnosis using WaveCORAL-DCCA: (**a**) Scenario i; (**b**) Scenario ii.

Table 5. Effectiveness of WaveCORAL-DCCA framework under Scenario i and Scenario ii.

DA Model	Scenario	Precision	Recall	F1-Score	Accuracy
WaveCORAL-DCCA	i	0.96	0.95	0.95	0.95
	ii	0.96	0.95	0.95	0.95

The results shown in Figure 11 and Table 5 demonstrate that WaveCORAL-DCCA structure significantly improved fault diagnosis performance, particularly in Scenario ii, where classification models without DA and using the original DCCA yielded lower accuracy; in both Scenario i and Scenario ii, the accuracy exceeded 95%.

As an illustration of how the WaveCORAL-DCCA framework positively influences the alignment of data distribution between the source and target domains, Figure 12 presents scatter plots for Scenario i, both before and after applying the designed DCCA structure.

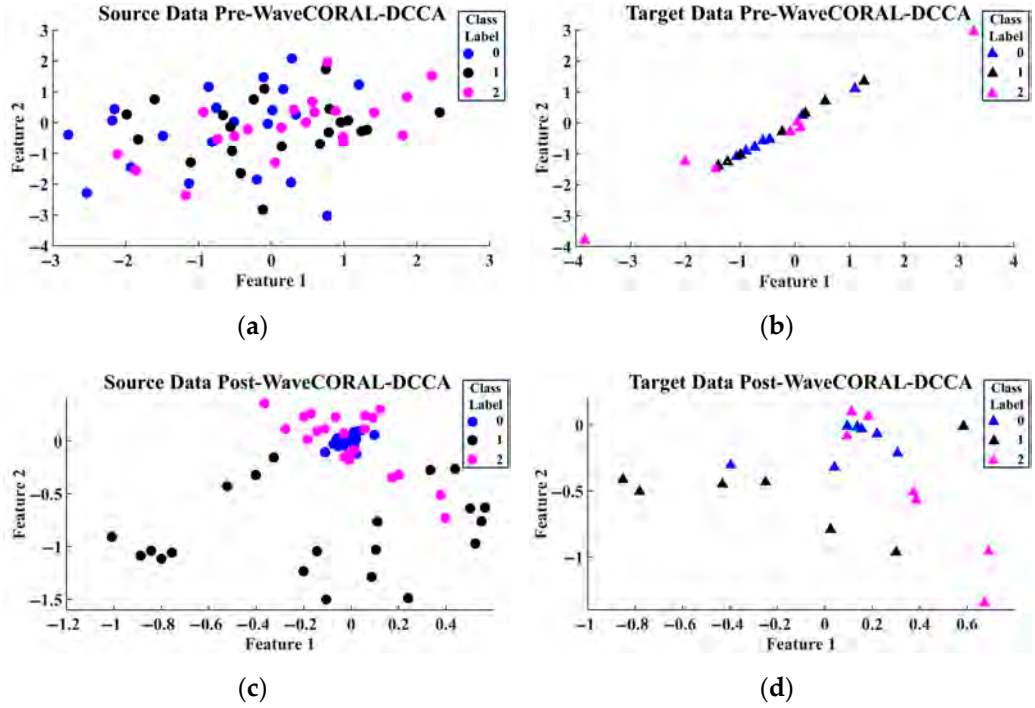


Figure 12. Scatter plots for Scenario i: (**a**) source domain before WaveCORAL-DCCA; (**b**) target domain before WaveCORAL-DCCA; (**c**) source domain after WaveCORAL-DCCA; (**d**) target domain after WaveCORAL-DCCA.

From Figure 12, a comparison of the feature distributions before (plots a and b) and after DCCA (plots c and d) clearly shows that the designed framework successfully aligns the source and target domains into a shared feature space, where observations from different classes are clustered in similar regions. Notably, 21 randomly selected samples from each class in the source domain were used as the test set for these scatter plots, and feature alignment was performed on these samples after DCCA training and validation. The same observations were also used for the pre-DCCA stage in the scatter plots.

5. Comparison Study

To further evaluate the effectiveness of the proposed framework, a comparative analysis was conducted against six well-known UDA approaches: maximum mean discrepancy (MMD) [44], joint distribution adaptation (JDA) [45], transfer component analysis (TCA) [34], balanced distribution adaptation (BDA) [46], adaptation regularisation-based transfer learning (ARTL) [47], and conditional domain adversarial network (CDAN) [48]. The performances of these methods were analysed in comparison with the WaveCORAL-DCCA framework. Figure 13 illustrates the comparative results, highlighting the diagnostic accuracies achieved by each method on the assumption of Scenario i and Scenario ii.

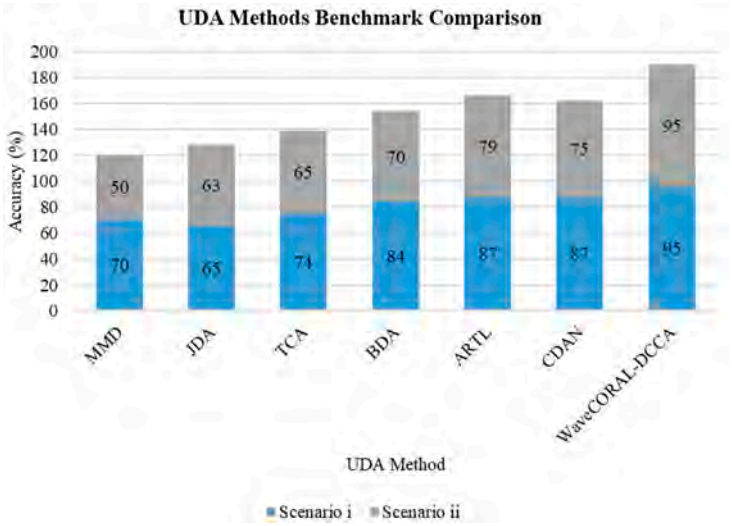


Figure 13. Comparative diagnostic performance of WaveCORAL-DCCA and the UDA benchmarks.

As shown in Figure 13, WaveCORAL-DCCA consistently outperforms the other UDA approaches across both scenarios. In particular, it achieves the highest diagnostic accuracies in both Scenario i and Scenario ii (95%), surpassing the performance of MMD, JDA, TCA, BDA, ARTL, and CDAN by a substantial margin. Notably, while ARTL and CDAN also demonstrate competitive performance, their accuracies remain below those of the proposed framework. These findings underscore the robustness of WaveCORAL-DCCA for DA tasks and highlight its efficacy in handling variations between different diagnostic scenarios.

6. Computational Efficiency and System Setup

The computational efficiency and practical implementation of the WaveCORAL-DCCA framework were prioritised in this study. All training, evaluation, and the 576-combination grid search were performed using Python and PyTorch on a standard workstation (CPU only), with MATLAB® R2024a being used for wavelet pre-processing as needed. Despite the large grid search, each scenario was completed in under five hours, and model inference required less than 0.3 s per batch. Scenario ii took slightly longer due to the increased complexity of the simulated dataset, but overall, the approach remains feasible for real-time industrial use. The system characteristics are listed in Table 6.

Table 6. System hardware and software specifications for WaveCORAL-DCCA experiments.

Component	Specification	Component	Specification
Operating System	Microsoft Windows 10 Enterprise (Build 19045)	Main libraries	PyTorch 2.0.1, Scikit-learn, NumPy, Matplotlib
Processor	Intel Core i7-14700 (20 cores, 28 threads, 2.10 GHz)	Wavelet pre-processing	MATLAB® R2024a (external, for selected datasets)
RAM	32 GB (31.6 GB usable)	Execution mode	CPU only (no GPU utilised)
Execution environment	Jupyter Notebook (Python 3.10)	Virtual memory	36.4 GB total, 19.9 GB available

The computation time for the two scenarios, i.e., Scenario i and Scenario ii, have been calculated and are listed in Table 7.

Table 7. Scenario-wise computational summary for model training and grid search.

Scenario	Source Domain	Target Domain	Total Samples	Grid Search Time	Maximum RAM Usage
i	6300	180	6480	3.2 h	12.3 GB
ii	6300	135	6435	4.1 h	13.1 GB

In summary, the WaveCORAL-DCCA framework was shown to be both computationally efficient and practically viable for real-time industrial applications. The full grid search and model training procedures are tractable with standard workstation resources, and the final diagnostic model is capable of rapid inference, meeting the operational demands of modern condition monitoring systems.

7. Discussion

The results demonstrate that WaveCORAL-DCCA achieves consistently high diagnostic accuracy for rotor fault diagnosis under significant domain shift. The integration of wavelet-based features, deep canonical correlation analysis, and CORAL loss enabled robust adaptation across all datasets, as shown in Table 5. Figure 11 further highlights the superiority of this approach in the comparison study, where it outperformed traditional models and other DA methods, particularly in challenging transfer scenarios. The confusion matrices in Figures 8–10 show that models without effective domain alignment struggled with class separation under severe distribution differences. The fault diagnosis results using the original DCCA baseline revealed its inability to handle distributional differences effectively, leading to inferior performance that was even worse than when no DA technique was applied. This poor generalisation, especially the complete failure in Scenario ii, stems from DCCA's limited feature expressiveness and the lack of explicit domain alignment. Without additional regularisation or alignment constraints, DCCA fails to reconcile substantial cross-domain discrepancies.

A key strength of the proposed framework is the efficiency of CORAL loss in aligning second-order statistics, which supported adaptation with limited target data. Nonetheless, reliance on covariance alignment may not fully address domain discrepancies in systems with strong non-Gaussian or nonlinear dynamics, as reflected in occasional misclassifications in some cases. While model performance exhibited some sensitivity to hyperparameter selection, the overall evidence supports the practical utility and reliability of WaveCORAL-DCCA for industrial rotor fault diagnosis.

A unique strength of WaveCORAL-DCCA lies in its ability to deliver robust diagnostic performance across datasets that differ not only in domain (experimental and simulated), but also in sensor modality and signal type. The consistent accuracy observed indicates that the combination of wavelet-based feature extraction and deep DA can bridge differences between acceleration and displacement data. This highlights the practical value of the framework for real-world deployments where data heterogeneity is unavoidable.

8. Conclusions

This paper introduces WaveCORAL-DCCA, an advanced UDA framework designed for rotor fault diagnosis under varying operational conditions and domain shifts. The framework leverages wavelet-based feature extraction combined with an enhanced DCCA network, where the CORAL loss function is employed to improve feature alignment between source and target domains without the need for labelled target data. Additionally, a numerical dataset of rotor systems was generated using the FE method in MATLAB® R2024a, supplementing experimental data and providing a comprehensive evaluation platform.

The proposed WaveCORAL-DCCA approach demonstrated a high diagnostic accuracy of 95% across three health scenarios—normal, unbalanced, and misaligned rotor systems—showcasing its ability to generalise effectively, even when limited unlabelled data from the target domain are available. The combination of wavelet-based features and enhanced DCCA ensures robust performance across different machines and fault types, making the framework a powerful tool for rotor fault diagnosis in industrial applications. The method's ability to maintain high accuracy while addressing domain shifts highlights its potential for broader applications in real-time fault diagnosis tasks. Furthermore, the comparative study unequivocally distinguished the effectiveness of the proposed method in handling scarce target domain data for adaptation.

Although noteworthy results were achieved with WaveCORAL-DCCA, sensitivity to hyperparameter values remains a limitation, with grid search optimisation being required to ensure peak performance. This may introduce complexity in practical implementation. In future work, this limitation will be addressed through the development of more automated or adaptive hyperparameter selection strategies, and the integration of higher-order or distributionally robust alignment techniques will also be investigated. Furthermore, the framework will be extended to encompass more complex fault scenarios, including composite and multi-faulted systems, and applied to a wider range of mechanical equipment to further assess its generalisability and industrial applicability.

Author Contributions: Conceptualisation, N.R.; investigation, N.R., M.D.O., D.P. and A.D.L.; data curation, N.R., D.P. and A.D.L.; writing—original draft preparation, review, and editing, N.R., M.D.O., D.P. and G.L.; supervision, A.D.L. and G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research study was carried out in the framework of the project "TU-LEARN—sTrUctural Life Extension enhAnced by aRtificial iNtelligence", funded by Unione Europea—Next Generation EU, as part of Prin 2022 PNRR—D.D. n. 1409 del 14-09-2022 program.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ARTL	Adaptation regularisation-based transfer learning
BDA	Balanced distribution adaptation
CCA	Canonical correlation analysis
CDANs	Conditional domain adversarial networks
CNNs	Convolutional neural networks
CORAL	Correlation alignment
DCCA	Deep canonical correlation analysis
DA	Domain adaptation
FE	Finite element
JDA	Joint distribution adaptation
MLP	Multi-layer perceptron
MMD	Maximum mean discrepancy
OVs	Operational variabilities
TCA	Transfer component analysis
TL	Transfer learning
UDA	Unsupervised domain adaptation
WT	Wavelet transformation

References

1. Wang, D.; Zhang, M.; Xu, Y.; Lu, W.; Yang, J.; Zhang, T. Metric-Based Meta-Learning Model for Few-Shot Fault Diagnosis Under Multiple Limited Data Conditions. *Mech. Syst. Signal Process.* **2021**, *155*, 107510. [CrossRef]
2. Yin, Z.; Yang, Y.; Shen, G.; Li, Y.; Huang, L.; Hu, N. Dynamic Modeling, Analysis, and Experimental Study of Ball Screw Pairs with Nut Spalling Faults in Electromechanical Actuators. *Mech. Syst. Signal Process.* **2023**, *184*, 109751. [CrossRef]
3. Li, Y.; Men, Z.; Bai, X.; Xia, Q.; Zhang, D. A Bearing Fault Diagnosis Method Based on M-SSCNN and M-LR Attention Mechanism. *Struct. Health Monit.* **2025**, *24*, 830–852. [CrossRef]
4. Rezazadeh, N.; De Luca, A.; Perfetto, D.; Salami, M.R.; Lamanna, G. Systematic Critical Review of Structural Health Monitoring Under Environmental and Operational Variability: Approaches for Baseline Compensation, Adaptation, and Reference-Free Techniques. *Smart Mater. Struct.* **2025**, *34*, 073001. [CrossRef]
5. Rezazadeh, N.; Annaz, F.; Jabbar, W.A.; Vieira Filho, J.; De Oliveira, M. A Transfer Learning Approach for Mitigating Temperature Effects on Wind Turbine Blades Damage Diagnosis. *Struct. Health Monit.* **2025**, 14759217241313350. [CrossRef]
6. Tang, S.; Ma, J.; Yan, Z.; Zhu, Y.; Khoo, B.C. Deep Transfer Learning Strategy in Intelligent Fault Diagnosis of Rotating Machinery. *Eng. Appl. Artif. Intell.* **2024**, *134*, 108678. [CrossRef]
7. Gong, X.; Feng, K.; Du, W.; Li, B.; Fei, H. An Imbalance Multi-Faults Data Transfer Learning Diagnosis Method Based on Finite Element Simulation Optimization Model of Rolling Bearing. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2024**, *238*, 8924–8940. [CrossRef]
8. Rezazadeh, N.; Perfetto, D.; Polverino, A.; De Luca, A.; Lamanna, G. Guided Wave-Driven Machine Learning for Damage Classification with Limited Dataset in Aluminum Panel. *Struct. Health Monit.* **2024**, 14759217241268394. [CrossRef]
9. Liang, P.; Yu, Z.; Wang, B.; Xu, X.; Tian, J. Fault Transfer Diagnosis of Rolling Bearings across Multiple Working Conditions via Subdomain Adaptation and Improved Vision Transformer Network. *Adv. Eng. Inform.* **2023**, *57*, 102075. [CrossRef]
10. Zhu, M.; Liu, J.; Hu, Z.; Liu, J.; Jiang, X.; Shi, T. Cloud-Edge Test-Time Adaptation for Cross-Domain Online Machinery Fault Diagnosis via Customized Contrastive Learning. *Adv. Eng. Inform.* **2024**, *61*, 102514. [CrossRef]
11. Rezazadeh, N.; Perfetto, D.; de Oliveira, M.; De Luca, A.; Lamanna, G. A Fine-Tuning Deep Learning Framework to Palliate Data Distribution Shift Effects in Rotary Machine Fault Detection. *Struct. Health Monit.* **2024**, 14759217241295951. [CrossRef]
12. Hou, Y.; Wang, H.; Wang, Y.; Wu, P.; Huang, W.; Wu, D. Structural Rotor Rub-Impact Diagnosis Under Intricate Noise Interferences Based on Targeted Component Extraction and Stochastic Resonance Enhancement. *Struct. Health Monit.* **2025**, *24*, 255–294. [CrossRef]
13. Miao, M.; Wang, Y.; Yu, J. Temporal Self-Supervised Domain Adaptation Network for Machinery Fault Diagnosis Under Multiple Non-Ideal Conditions. *Reliab. Eng. Syst. Saf.* **2024**, *251*, 110347. [CrossRef]
14. Rezazadeh, N.; Perfetto, D.; Caputo, F.; De Luca, A. Enhancing Air Compressor Fault Diagnosis: A Comparative Study of GPT-2 and Traditional Machine Learning Models. *Macromol. Symp.* **2025**, *414*, 7057. [CrossRef]
15. Rajagopalan, S.; Purohit, A.; Singh, J. Genetically Optimised SMOTE-Based Adversarial Discriminative Domain Adaptation for Rotor Fault Diagnosis at Variable Operating Conditions. *Meas. Sci. Technol.* **2024**, *35*, 106109. [CrossRef]
16. Espinoza-Sepulveda, N.; Sinha, J. Two-Step Vibration-Based Machine Learning Model for the Fault Detection and Diagnosis in Rotating Machine and Its Blind Application. *Struct. Health Monit.* **2025**, *24*, 1029–1042. [CrossRef]
17. Wang, J.; Ahmed, H.; Chen, X.; Yan, R.; Nandi, A.K. Improved Adversarial Transfer Network for Bearing Fault Diagnosis Under Variable Working Conditions. *Appl. Sci.* **2024**, *14*, 2253. [CrossRef]
18. Xu, Y.; Liu, J.; Wan, Z.; Zhang, D.; Jiang, D. Rotor Fault Diagnosis Using Domain-Adversarial Neural Network with Time-Frequency Analysis. *Machines* **2022**, *10*, 610. [CrossRef]
19. Liu, S.; Wang, H.; Tang, J.; Zhang, X. Research on Fault Diagnosis of Gas Turbine Rotor Based on Adversarial Discriminative Domain Adaption Transfer Learning. *Measurement* **2022**, *196*, 111174. [CrossRef]
20. Liu, Z.H.; Lu, B.L.; Wei, H.L.; Chen, L.; Li, X.H.; Wang, C.T. A Stacked Auto-Encoder Based Partial Adversarial Domain Adaptation Model for Intelligent Fault Diagnosis of Rotating Machines. *IEEE Trans. Ind. Inform.* **2021**, *17*, 6798–6809. [CrossRef]
21. Zou, Y.; Yu, Z.; Vijaya Kumar, B.V.K.; Wang, J. Unsupervised Domain Adaptation for Semantic Segmentation via Class-Balanced Self-Training. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer: Berlin/Heidelberg, Germany, 2018; Volume 11207.
22. Long, M.; Cao, Y.; Cao, Z.; Wang, J.; Jordan, M.I. Transferable Representation Learning with Deep Adaptation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 3071–3085. [CrossRef]
23. Cha, Y.-J.; Ali, R.; Lewis, J.; Büyüköztürk, O. Deep Learning-Based Structural Health Monitoring. *Autom. Constr.* **2024**, *161*, 105328. [CrossRef]
24. Sun, K.; Yin, A.; Lu, S. Domain Distribution Variation Learning via Adversarial Adaption for Helicopter Transmission System Fault Diagnosis. *Mech. Syst. Signal Process.* **2024**, *215*, 111419. [CrossRef]

25. Han, S.; Wang, Z.; Zhang, H.; Zhang, F.; Han, Q. Flexible Rotor Unbalance Fault Location Method Based on Transfer Learning from Simulation to Experiment Data. *Meas. Sci. Technol.* **2023**, *34*, 125053. [CrossRef]
26. Khan, A.; Kim, J.-S.; Kim, H.S. Damage Detection and Isolation from Limited Experimental Data Using Simple Simulations and Knowledge Transfer. *Mathematics* **2021**, *10*, 80. [CrossRef]
27. Ma, Z.; Fu, L.; Dun, G.; Tan, D.; Xu, F.; Zhang, L. A Robust Domain Distribution Alignment Discriminative Network Driven by Physical Samples for Rotor-Bearing Fault Diagnosis. *Knowl. Based Syst.* **2024**, *300*, 112216. [CrossRef]
28. Xiang, L.; Zhang, X.; Zhang, Y.; Hu, A.; Bing, H. A Novel Method for Rotor Fault Diagnosis Based on Deep Transfer Learning with Simulated Samples. *Measurement* **2023**, *207*, 112350. [CrossRef]
29. Wang, M.; Li, J.; Xue, Y. A New Weight-Based Dual Domain Adaptation Transfer Model for Bearing Fault Diagnosis Under Noisy and Cross-Domain Conditions. *IEEE Access* **2023**, *11*, 123766–123783. [CrossRef]
30. Xiao, Y.; Zhou, X.; Zhou, H.; Wang, J. Multi-Label Deep Transfer Learning Method for Coupling Fault Diagnosis. *Mech. Syst. Signal Process.* **2024**, *212*, 111327. [CrossRef]
31. Zhang, X.; Zhang, X.; Liu, J.; Wu, B.; Hu, Y. Graph Features Dynamic Fusion Learning Driven by Multi-Head Attention for Large Rotating Machinery Fault Diagnosis with Multi-Sensor Data. *Eng. Appl. Artif. Intell.* **2023**, *125*, 106601. [CrossRef]
32. Zhang, X.; Liu, J.; Zhang, X.; Lu, Y. Self-Supervised Graph Feature Enhancement and Scale Attention for Mechanical Signal Node-Level Representation and Diagnosis. *Adv. Eng. Inform.* **2025**, *65*, 103197. [CrossRef]
33. Rezazadeh, N.; de Oliveira, M.; Perfetto, D.; De Luca, A.; Caputo, F. Classification of Unbalanced and Bowed Rotors Under Uncertainty Using Wavelet Time Scattering, LSTM, and SVM. *Appl. Sci.* **2023**, *13*, 6861. [CrossRef]
34. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain Adaptation via Transfer Component Analysis. *IEEE Trans. Neural Netw.* **2011**, *22*, 199–210. [CrossRef]
35. Brito, L.C.; Susto, G.A.; Brito, J.N.; Duarte, M.A.V. Mechanical Faults in Rotating Machinery Dataset (Normal, Unbalance, Misalignment, Looseness). Mendeley Data, V3. 2023. Available online: https://data.mendeley.com/datasets/zx8pfhdtnb/3 (accessed on 3 August 2025).
36. Liu, D.; Xiao, Z.; Hu, X.; Zhang, C.; Malik, O.P. Feature Extraction of Rotor Fault Based on EEMD and Curve Code. *Measurement* **2019**, *135*, 712–724. [CrossRef]
37. Peng, B.; Wan, S.; Bi, Y.; Xue, B.; Zhang, M. Automatic Feature Extraction and Construction Using Genetic Programming for Rotating Machinery Fault Diagnosis. *IEEE Trans. Cybern.* **2021**, *51*, 4909–4923. [CrossRef] [PubMed]
38. Zhang, W.; Zhang, T.; Cui, G.; Pan, Y. Intelligent Machine Fault Diagnosis Using Convolutional Neural Networks and Transfer Learning. *IEEE Access* **2022**, *10*, 50959–50973. [CrossRef]
39. Wu, Q.; Zhang, X.; Zhao, B. A Novel Adaptive Kernel-Guided Multi-Condition Abnormal Data Detection Method. *Measurement* **2023**, *206*, 112257. [CrossRef]
40. Li, X.; Hu, H.; Zhang, S.; Tang, G. A Fault Diagnosis Method for Rotating Machinery with Semi-Supervised Graph Convolutional Network and Images Converted from Vibration Signals. *IEEE Sens. J.* **2023**, *23*, 11946–11955. [CrossRef]
41. Wang, J.; Ran, R.; Fang, B. GNPENet: A Novel Convolutional Neural Network with Local Structure for Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 3504316. [CrossRef]
42. Darpe, A.K.; Gupta, K.; Chawla, A. Coupled Bending, Longitudinal and Torsional Vibrations of a Cracked Rotor. *J. Sound Vib.* **2004**, *269*, 33–60. [CrossRef]
43. Bhaumik, D.; Bhaumik, D. Misalignment-Related Defect Detection Using Discrete Wavelet Transform. *Int. J. Recent. Technol. Eng.* **2023**, *12*, 97–101. [CrossRef]
44. Yan, H.; Ding, Y.; Li, P.; Wang, Q.; Xu, Y.; Zuo, W. Mind the Class Weight Bias: Weighted Maximum Mean Discrepancy for Unsupervised Domain Adaptation. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017.
45. Long, M.; Wang, J.; Ding, G.; Sun, J.; Yu, P.S. Transfer Feature Learning with Joint Distribution Adaptation. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013.
46. Wang, J.; Chen, Y.; Hao, S.; Feng, W.; Shen, Z. Balanced Distribution Adaptation for Transfer Learning. In Proceedings of the IEEE International Conference on Data Mining, ICDM, New Orleans, LA, USA, 18–21 November 2017.
47. Long, M.; Wang, J.; Ding, G.; Pan, S.J.; Yu, P.S. Adaptation Regularization: A General Framework for Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1076–1089. [CrossRef]
48. Long, M.; Cao, Z.; Wang, J.; Jordan, M.I. Conditional Adversarial Domain Adaptation. In Proceedings of the Advances in Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018.