



Article Intelligent Health Assessment of Aviation Bearing Based on Deep Transfer Graph Convolutional Networks under Large Speed Fluctuations

Xiaoli Zhao ^{1,2}, Xingjun Zhu ¹, Jianyong Yao ^{1,*}, Wenxiang Deng ¹, Yudong Cao ³, Peng Ding ³, Minping Jia ³ and Haidong Shao ⁴

- School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; xlzhao@njust.edu.cn (Xiaoli Zhao); zhuxj@njust.edu.cn (Xingjun Zhu)
- ² State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310027, China
- ³ School of Mechanical Engineering, Southeast University, Nanjing 211189, China; mpjia@seu.edu.cn (M.J.)
- ⁴ College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China
- Correspondence: jerryyao.buaa@gmail.com

Abstract: As a critical support and fixed component of aero engines, electro-hydrostatic actuators, and other equipment, the operation of aviation bearings is often subject to high speed, high-temperature rise, large load, and other continuous complex fluctuation conditions, which makes their health assessment tasks more difficult. To solve this problem, an intelligent health assessment method based on a new Deep Transfer Graph Convolutional Network (DTGCN) is proposed for aviation bearings under large speed fluctuation conditions. First, a new DTGCN algorithm is designed, which mainly uses the domain adaptation mechanism to enhance the performance of Graph Convolutional Network (GCN) and the generalization performance of transfer properties. Specifically, order spectrum analysis is employed to resample the vibration signals of aviation bearings and transform them into order spectral signals. Then, the trained 1dGCN is used as the feature extractor, and the designed Dynamic Multiple Kernel Maximum Mean Discrepancy (DMKMMD) is calculated to match the difference in edge distribution. Finally, the aligned features are fed into the softmax classifier for intelligent health assessment. The effectiveness of the proposed diagnostic algorithm and method are validated by using aviation bearing fault data set under large speed fluctuation conditions.

Keywords: aviation bearings; intelligent health assessment; large speed fluctuations; graph convolutional network (GCN); transfer learning

1. Introduction

Currently, high-speed rotating equipment represented by aero engines and aerohydrostatic actuators is widely used in industry, defense, and military industries [1–3]. Specifically, aviation bearings and others are core components of the rotating equipment drive system. Its working environment is exceptionally harsh and complex, with frequent high speeds, high-temperature rises, large loads, and other continuously fluctuating conditions. Once the aviation bearing failure or damage directly affects the safety of aviation equipment, the light will lead to an abnormal increase in vibration or noise signals. Heavy will lead to irreversible catastrophic accidents [4,5]. Accordingly, it is necessary to carry out effective condition monitoring and fault diagnosis for bearings and other core components of equipment under complex operating conditions (such as variable conditions, speed fluctuation conditions, etc.) [4–6].

Specifically, the operating speed and load of aviation bearings are determined by the cruising speed and maneuvering action. The working condition of speed fluctuation is almost everywhere. The problems associated with intelligent health assessment of aviation bearings under large speed fluctuation conditions are illustrated below [7,8]:



Citation: Zhao, X.; Zhu, X.; Yao, J.; Deng, W.; Cao, Y.; Ding, P.; Jia, M.; Shao, H. Intelligent Health Assessment of Aviation Bearing Based on Deep Transfer Graph Convolutional Networks under Large Speed Fluctuations. *Sensors* 2023, 23, 4379. https://doi.org/ 10.3390/s23094379

Academic Editor: Genda Chen

Received: 17 April 2023 Revised: 23 April 2023 Accepted: 27 April 2023 Published: 28 April 2023



Copyright: © 2023 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 (https:// creativecommons.org/licenses/by/ 4.0/). (1) The frequency domain signals of bearing faults under large speed fluctuation will generate frequency shift and amplitude change problems, which make the traditional signal processing methods unable to make accurate distinctions. (2) The time interval fluctuations between two adjacent shocks caused by speed changes. (3) Significant differences exist in the distribution between signals or samples. (4) The working environment is more complex, with many excitation sources, complex transmission paths, and more severe noise distribution. Namely, all these problems will make intelligent health assessment of aviation bearings more arduous.

For traditional aviation bearing fluctuating vibration signals, the main processing methods are based on two modes such as order-based analysis and time-frequency-based analysis [8–10]. Strictly speaking, the traditional spectral analysis method is employed to conduct discrete Fourier analysis on the vibration signal to obtain the main frequency components of the signal rather than the frequency change law with time. Therefore, the Fourier transform can only analyze the smooth signal if the spectral analysis interprets the vibration signal of the speed fluctuation. "Frequency blurring" and other phenomena are bound to occur. Specifically, short-time Fourier transforms (STFT), Gabor transforms, Wigner-Ville distribution (WVD), continuous wavelet transform (CWT), order tracking, and other methods can be employed to analyze non-stationary vibration signals [8–10]. The vibration of rotating machinery is often related to the rotational speed. Its operating state can be identified by the interrelationship between the order frequency components of the vibration signal proportional to the rotational speed. Therefore, order ratio analysis is more advantageous to represent the rotational speed-dependent vibrations well for monitoring the state characteristics of rotating machinery and practical health assessment [10].

Compared with the constant operating conditions, the forces on aviation bearings under fluctuating speed conditions are more variable and prone to failure. The mapping relationship between their signal signs and failure mechanisms becomes more complex. Therefore, studying bearing fault feature extraction and diagnosis under the large speed fluctuation condition is of practical theoretical and engineering value. Zhao, et al. [10], pointed out that the dynamic signals of mechanical equipment under variable speed will have non-smooth features such as frequency modulation, amplitude modulation, and modulation, which are often coupled with each other, making the fault feature extraction very difficult. Zhu, et al. [11], studied the transient shock response pattern of bearing faults under variable speed based on signal sparse representation theory, A sparse representationbased fault feature extraction method for variable speed bearings was proposed.

In summary, although the above-mentioned feature extraction methods have progressed and achieved the accuracy requirements for signal analysis under severe speed or load changes and random fluctuations, they still need to benefit from solving the highprecision time-frequency representation of arbitrary targets perfectly. Accordingly, the research on steady-state and relocatable fault feature extraction methods still needs to be explored and further decoded.

Deep learning (DL), one of the most advanced data and information processing tools, is used to enhance the accuracy of classification or prediction by multi-layer feature extraction [12,13]. Accordingly, scholars have researched intelligent fault diagnosis methods based on deep learning [13,14]. Miao, et al. [15], first used the AMESim simulation platform and experimental platform to simulate the fault of the aero-electro hydrostatic actuation system, then constructed an intelligent fault diagnosis model based on a deep convolutional neural network. Shao, et al. [13], designed a new fault diagnosis framework based on Dual-Threshold Attention-Guided Gan and limited infrared thermal images for rotating machinery. Zhao, et al. [1], developed an intelligent fault diagnosis method incorporating adaptive parametric rectifier linear units and deep residual networks for rotating machinery vibration signals that can vary significantly under different operating conditions. Due to the "high-dimensional and massive" state of the data measured by the aerospace equipment, the information from multiple sources from cross-coupling conditions, modes, and channels makes it difficult to guarantee the generalization of the diagnostic model. However, DL

only performs nonlinear mapping of Euclidean space data, which is easy to ignore the interdependent correlation between data [13–15].

With the continuous in-depth research of DL models, many deep learning algorithms are combined with graph theory, where deep neural networks are injected with fresh blood [16–19]. Currently, the hot graph neural networks (GNN) focus more on the connection relationship of the data, and its goal is to establish specific network connections for the data stored in the graph domain to deal with the structural information of non-Euclidean spatial data. Owing to traditional neural networks are not translation invariant on non-Euclidean data structures (they cannot employ kernels of the same size for convolution). Still, this type of data is widely available in the real world. Graph Convolutional Networks (GCN) [17] came into being to handle such data, making it possible to convolve on irregular graph structures. Recently, algorithms such as Graph Attention Network (GAT) [18] and Graph Convolutional Auto-encoder (GCAE) [19] have been proposed, which have stronger data sparsity and reconfigurability. In Ref. [20], an intelligent fault diagnosis framework for multi-receiver field graph convolutional networks was offered for the existing DL algorithms that cannot mine the correlation information between signals. Zhao, et al. [21], proposed an intelligent fault diagnosis method for electro-mechanical systems based on a novel semi-supervised graph convolutional deep belief networks (SSGCDBN) algorithm.

However, the fluctuations in real operating conditions can produce a series of disturbances to the training of traditional machine learning (e.g., shallow learning, DL, GNN, etc.) models, which can be summarized as follows [4,20,21]: (1) It is challenging to obtain fault data during the service life of an aviation bearing, as fault data and label data are relatively scarce, but traditional models need a sufficient amount for model training. (2) The distribution of bearing measurement data varies greatly across conditions and component levels, which is challenging to ensure that the trained model can be adapted to other distributed data, etc. Accordingly, these factors are the bottlenecks that hinder the development of aviation-bearing fault diagnosis and intelligent maintenance.

To this end, transfer learning (TL) [22,23] provides a novel solution to the aforementioned problem by applying models and knowledge learned in the old domain (source domain) to the new domain (target domain) by exploiting the similarity between data, tasks, or models. Liao, et al. [22], first proposed a deep semi-supervised domain generalization network for fault diagnosis under rotating machinery variable speed for the problem that testing and training samples are not independently and identically distributed. Li, et al. [23], also established a deep adversarial multi-classifier optimization model based on cross-domain fault diagnosis to solve the domain transfer problem. However, the large fluctuation of fault conditions and multi-source parameter coupling of aviation bearings can result in a geometric structure relationship. Moreover, the non-Euclidean space correlation between their graph domains needs to be addressed. To solve the above-mentioned problems, deep graph transfer learning (DGTL) [22–26] is born, mainly by fusing the geometric structure relationship extraction performance of graph neural network and the knowledge transfer ability and strong generalization, which opens up a new research idea to solve the existing problems of aerospace equipment fault diagnosis.

To solve the problem of fault diagnosis difficulties arising from fluctuating conditions of aviation bearings, a new deep graph transfer learning algorithm is designed in this paper. Specifically, an intelligent health assessment method is proposed based on a deep transfer graph convolutional network (DTGCN) for aviation bearings under large fluctuating working conditions. Firstly, the proposed DTGCN algorithm uses a feature transfer learning mechanism to enhance GCN to have both solid geometric feature extractions in the graph domain and stronger generalization performance with more robust transfer characteristics. The aviation-bearing fault dataset under large speed fluctuation has validated the proposed diagnostic method and algorithm. Namely, the main contributions of this paper are included in the following three aspects:

 A deep transfer graph convolutional network (named DTGCN) algorithm is first proposed in this paper;

- (2) Based on the proposed DTGCN algorithm, an intelligent health assessment method based on DTGCN algorithm is proposed for aviation bearing under large speed fluctuations;
- (3) The intelligent health assessment method based on DTGCN can be validated using experimental data from aero engine bearing failure simulation.

The rest of this paper is mainly described as follows. Section 2 reviews a brief theory of GCN and MKMMD. Section 3 describes the proposed approach in detail. The validation and analysis of case studies are presented in Section 4. Finally, some conclusions are summarized in Section 5.

2. Preliminaries

The basics of graph convolutional network (GCN) and multiple kernel maximum mean discrepancy (MKMMD) [24] can be briefly introduced in this section to provide a research basis for the subsequent design of intelligent health assessment algorithm and method.

2.1. Graph Convolutional Network (GCN)

To effectively handle non-Euclidean spatial data, GCN [17] was created to perform convolution on irregular graph structures. The schematic diagram of GCN based on frequency domain convolution can be illustrated in Figure 1. The specific GCN implementation process is described as follows.



Figure 1. Schematic diagram of Graph Convolutional Network (GCN).

First, the spectral graph relationship with the Laplacian matrix as nodes and edges must be established. The graph can be represented as G (V, E, W), where V denotes the node in the graph, E denotes the edge between two nodes, and W denotes the edge weight between two vertices. In addition, the graph can be represented by a Laplace matrix as L = D - A, where D and A represent the degree matrix and adjacency matrix, respectively. Ultimately, the characteristic decomposition of the Laplacian matrix is defined as

$$L = U \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} U^{-1} = U\Lambda U^{-1}$$
(1)

where $U = \begin{bmatrix} \vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \end{bmatrix}$ is the unit eigenvector and Λ is a diagonal matrix consisting of the eigenvalues of the Laplacian matrix. Since U is an orthogonal matrix (i.e., UUT = E), its eigendecomposition can also be expressed as,

$$L = U \begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{pmatrix} U^T = U \Lambda U^T$$
⁽²⁾

where is $U^{-1} = U^T$.

Based on the frequency domain convolution theory, the graph Fourier transform on the graph signal can be performed and convoluted in the spectral domain using the Fourier transform. Specifically, the graph Fourier transform operation is described as,

$$F(\lambda_l) = \hat{f}(\lambda_l) \sum_{i=1}^{N} f(i) u_l^*(i)$$
(3)

where *f* is the N-dimensional component of the graph, and f(i) corresponds to the graph nodes one by one. Denotes $u_l(i)$ is the *i*-th element of the *l*-th feature vector and $u_l^*(i)$ is the conjugate vector $u_l(i)$ of the representation.

The Fourier transform matrix form of the graph is $\hat{f} = U^T f$. Specifically, the Fourier transform has specific properties: $f \times g = F^{-1}{F\{f\} \cdot F\{g\}}$. Calculating the graph domain convolution is equivalent to Fourier domain multiplication, which first does the Fourier transform of the graph and convolution kernel after multiplication. Then, the Fourier inverse transforms back. The graph domain convolution can be obtained. The corresponding graph Fourier inverse transform is written as,

$$f(i) = \sum_{i=1}^{N} \hat{f}(\lambda_i) u_l(i).$$

$$\tag{4}$$

According to the convolution theorem (i.e., the convolution of both functions f(t) and g(t) is the inverse of the product of the Fourier transforms of their functions), the process of convolution of the graph is rewritten as,

$$(f * g)_G = U \begin{pmatrix} g(\lambda_1) \\ \dots \\ g(\lambda_n) \end{pmatrix} U^T f$$
(5)

where $(f * g)_G$ denotes the convolution of the functions f(t) and g(t), $U^T f$ denotes the Fourier transform, and g denotes the convolution kernel.

According to the properties of the Laplacian matrix, there is $(f \times h)_G = U(U^T(h) \cdot U^T(f))$, where \circ is the Hadamard product, which represents the element-by-element product of two vectors of the same dimension at the corresponding positions. The output of the graph convolution network can be simply expressed as,

$$y_{output} = \sigma(Ug_{\theta}(\lambda)U^T x) \tag{6}$$

More theoretical knowledge about GCN can be found in Ref. [17].

2.2. Multi Kernel-Maximum Mean Discrepancies (MKMMD)

MKMMD [24] is one of the most commonly used non-parametric methods to measure the difference in distribution between two domain datasets. Specifically, the feature representations in the source and target domains in transfer learning are mapped into the reproducing kernel Hilbert space (RKHS). Then the distance between the means of the two data types is calculated. The mean matching is calculated spatially to estimate the variability between edge distributions. Specifically, let h^s and h^t denote two types of deep features that fluctuate or cross-machine/cross-bearing, respectively, then MKMMD distance based on the two types of features is described as,

$$\ell_{MK-MMD}(h^{s}, h^{t}) = \left\| \frac{1}{M_{s}} \sum_{i=1}^{M_{s}} \phi(h_{i}^{s}) - \frac{1}{M_{t}} \sum_{i=1}^{M_{t}} \phi(h_{i}^{t}) \right\|_{H}^{2}$$
(7)

where $\|\cdot\|_H$ denotes the regenerated kernel Hilbert space and $\phi(\cdot)$ represents a series of characteristic mapping functions associated with the kernel mapping $k(h^s, h^t) = \langle \phi(h^s), \phi(h^t) \rangle$

for mapping the original variables to the regenerated kernel Hilbert space, $k_l(h^s, h^t)$ defined as a convex combination of the l underlying kernels.

3. The Proposed Method

3.1. Problem Description

Specifically, the domain adaptation problem and task for cross-bearing fault feature extraction can be described as follows [27,28]: First, assuming that a labeled source domain $D_s = \{x_s^i, y_s^i\}_{i=1}^n$ and an unlabeled target domain $D_t = \{x_s^i\}_{i=1}^m$ are given. However, the edge and the conditional distribution are different, i.e., $P(x_s) \neq P(x_t)$ and $Q(y_s, x_s) \neq Q(y_t, x_t)$. Accordingly, the model trained based on the source domain cannot be used directly in the target domain. Accordingly, the source and target domains can be mapped to a shared feature space by constructing a mapping relationship *f*. The mapped feature distribution satisfies the requirement that the source domain distribution is approximately equal to the target domain distribution, which can be used for target domain classification, and other problems described in Figure 2.



Figure 2. Problem description for domain adaptation based-intelligent health assessment.

3.2. The Constructed DTGCN Algorithm

To implement domain adaptation fault feature extraction and intelligent health assessment under fluctuating operating conditions, the proposed DTGCN-based-intelligent health assessment is designed in this section. This algorithm mainly consists of three components: feature extractor G, label predictor P, and domain discriminator D. Specifically, the structure of the proposed DTGCN algorithm is schematically displayed in Figure 3. In addition, the proposed DTGCN algorithm mainly involves a graph construction phase, a 1dGCN feature learning phase, and a domain adaptation and classification phase, which are illustrated below.

Figure 3. Schematic structure of the designed deep transfer graph convolutional network algorithm.

3.2.1. Graph Learning Construction Stage

Firstly, the multi-head attention mechanism and one-dimensional GCN are mainly used to construct the graph representation layer, graph attention layer, graph convolution, and graph pooling layer for multi-source signals and information fuse. The specific operation process is illustrated as follows.

Graph representation layer: An undirected connected graph can be defined as $G = \left\{ (V_i)_{i=1}^n, (E_j)_{j=1}^m, A \in \mathbb{R}^{N \times N} \right\}$. $\{ (V_i)_{i=1}^n \}$ represents all samples in the graph and the vertices that make up the graph. $\left\{ (E_j)_{j=1}^m \right\}$ denotes all edges connecting the vertices. Based on this, the algorithmic steps are refined, and the core steps of graph representation learning are described in Figure 4.

Figure 4. Graph representation learning core steps: (**a**) Sample relationship graph; (**b**) graph convolution process.

3.2.2. Feature Fusion and Extraction Stage

Graph attention layer: after realizing the signal-to-graph domain conversion, the graph attention layer inputs are both two sets of signals, which are randomly selected in the same domain with $V^b = \left\{ v_1^b, v_2^b, \dots, v_n^b \middle| v_i^b \in \mathbb{R}^L \right\}$ and $V^c = \left\{ v_1^c, v_2^c, \dots, v_n^c \middle| v_i^c \in \mathbb{R}^L \right\}$ corresponding to the adjacency matrix $A^{bc} \in \mathbb{R}^{n \times n}$, respectively. *L* is the signal length. A^{bc}

is extracted by $F(\cdot)$ for each signal, and the Euclidean distance between two features is calculated as

$$A^{bc} = \begin{bmatrix} \left\| F\left(v_{1}^{b}\right), F\left(v_{1}^{c}\right) \right\|_{2}^{2} \left\| F\left(v_{1}^{b}\right), F\left(v_{2}^{c}\right) \right\|_{2}^{2} \cdots \left\| F\left(v_{1}^{b}\right), F\left(v_{n}^{c}\right) \right\|_{2}^{2} \\ F\left(v_{2}^{b}\right), F\left(v_{1}^{c}\right) \right\|_{2}^{2} \left\| F\left(v_{2}^{b}\right), F\left(v_{2}^{c}\right) \right\|_{2}^{2} \cdots \left\| F\left(v_{2}^{b}\right), F\left(v_{n}^{c}\right) \right\|_{2}^{2} \\ \vdots & \cdots \\ \left\| F\left(v_{n}^{b}\right), F\left(v_{1}^{c}\right) \right\|_{2}^{2} \left\| F\left(v_{n}^{b}\right), F\left(v_{2}^{c}\right) \right\|_{2}^{2} \cdots \left\| F\left(v_{n}^{b}\right), F\left(v_{n}^{c}\right) \right\|_{2}^{2} \end{bmatrix}$$
(8)

Each element of A^{bc} represents the correlation between two signal samples. With the benefit of A^{bc} revealing the correlation between V^b and V^c vertices, the graph attention layer efficiently fuses the information contained in V^b and V^c , and the attention matrix E can be calculated as

$$E = Attention\left(V^{b}W^{b}, V^{c}W^{c}\right)$$
(9)

The multi-head-attention (MAE) mechanism is introduced to filter irrelevant information. Essentially, MAE is a combination of multiple self-attentive structures using Scaled Dot Product Attention (SDPA) to compute the input vector sequence query Q, key K, and value V attention output, which can be expressed as

$$SDPA(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V$$
 (10)

where $Q = W^q e_i$, $K = W^k e_i$, $V = W^v e_i$, *d* is the network hidden layer size, i.e., the projection size. Accordingly, the input features in the MAE are projected to different subspaces by multiple self-attention operations, and then the set of multiple attention vectors are calculated as

$$Mutilhead(Q, K, V) = Concat(head_1, head_2, \cdots, head_n)$$
(11)

$$head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right)$$
(12)

The multi-headed attention output is cascaded with a single-layer feedforward network, and the output is obtained. Then, each head attention output can be expressed as

$$head_i = SDPA\left(QW_i^Q, KW_i^K, VW_i^V\right)$$
(13)

Equation (9) to Equation (13) were employed to weigh the combined form to form a multi-head attention matrix.

First, the graph structure of the multi-source signal at a specific moment is normalized $L = I_N - D^{-1/2}AD^{-1/2} \in R^{N \times N}$, where I_N denotes the unit matrix, A denotes the adjacency matrix, and $D \in R^{N \times N}$ denotes the diagonal matrix ($D_{ii} \in \sum_j A_{ij}$). Since the Laplacian matrix L is a real symmetric matrix, an eigenvalue decomposition $L = U\Lambda U^T$ is performed on it, where U denotes the Fourier basis $L = U\Lambda U^T$ and the diagonal matrix $\Lambda = diag([\lambda_0, \ldots, \lambda_{N-1}]) \in R^{N \times N}$ is composed of eigenvalues. The graph signal x of Fourier transform can be expressed as $\hat{x} = U^T x$, where x is the graph signals from feature fusion. Since U is an orthogonal matrix, its Fourier inverse transform is $x = U\hat{x}$. Doing the convolution operator g and the graph Fourier transform of the graph signal x separately, multiplying the transform results, and then obtaining the graph convolution results by the graph Fourier inverse transform,

$$g * x = U\left(U^T g \cdot U^T x\right) \tag{14}$$

Here, * specifically refers to the graph convolution operation. Therefore, Chebyshev polynomials are proposed for the approximate solution with guaranteed accuracy. Ultimately, the propagation rule of the convolutional layer in GCN is expressed as

$$H^{(l+1)} = \sigma \left(\tilde{D}^{-1/2} \mathbf{A} \tilde{D}^{-1/2} H^{(l)} W^{(l)} \right)$$
(15)

where $\sigma(\cdot)$ is the activation function, *H* is the output, *W* is the weight matrix, and *l* is the number of layers. Finally, GCN is then formed.

3.2.3. Domain Adaptation and Classification Stages

Accordingly, the feature extractor consists of three 1dGCN layers and one fully connected layer, and the propagation rules of the 1dGCN layer are described above. The corresponding domain discriminator D determines whether the data comes from the source or target domain, takes the learned features as input, and outputs the predicted domain labels D(G(x)).

Therefore, the domain adversarial learning strategy is introduced to minimize the training loss function of the feature extractor *G*, and the training loss function of the domain discriminator *D* is maximized. Through such a dynamic "game" process, the feature extractor *G* and the domain discriminator *D* are optimally trained. Finally, the label predictor *P* is further used to optimize and predict *G*(*x*). Assuming that the parameters of the feature extractor, label predictor, and domain discriminator are θ_G , θ_P , and θ_D , respectively, the loss function in model optimization can be expressed as

$$L(\theta_G, \theta_D, \theta_P) = \frac{1}{ns} \sum_{\substack{x_i \in D_s \\ ns+nt}} L_p(P(G(x_i)), y_i) - \frac{\lambda}{ns+nt} \sum_{\substack{x_i \in (D_s \cup D_t)}} L_D(D(G(x_i)), d_i)$$
(16)

where L_p and L_D are label predictor and domain discriminator, respectively. λ are the adjustment factors, and d_i is the label of inputting samples. Finally, the label predictor parameters are updated by minimizing and maximizing the objective function to update the domain discriminator parameters.

For the distribution constraint of DMKMMD, when the edge distributions D_f of two domain features are different, the dynamic distribution alignment \overline{D}_f can be quantitatively evaluated (D_f and D_f^c). \overline{D}_f can be expressed as

$$\overline{D}_f = (1 - \gamma)D_f(P_s, P_t) + \gamma D_f^c(P_s, P_t)$$
(17)

where the adaptive factors $\gamma \in [0,1]$ are used for the alignment weights D_f^c and D_f . The dynamic alignment distribution for the domain invariant mapping f can be further expressed as

$$\overline{D}_{f} = (1 - \gamma) \left\| \frac{1}{n} \sum_{i=1}^{n} f(z_{i}) - \frac{1}{m} \sum_{j=n+1}^{n+m} f(z_{j}) \right\|_{H_{K}}^{2} + \gamma \left\| \frac{1}{n} \sum_{i=1}^{n} f(z_{i}^{c}) - \frac{1}{m} \sum_{j=n+1}^{n+m} f(z_{j}^{c}) \right\|_{H_{K}}^{2} = tr(\widetilde{G}M_{0}) + tr(\widetilde{G}M_{c})$$
(18)

where tr denotes the operation of finding the matrix trace, and the optimal combination of multiple kernels makes the feature extraction have a more accurate and reasonable expression by setting K to be a convex combination of U different kernel functions, which can be further expressed as

$$K: \left\{ k = \sum_{u=1}^{U} a_{u} k_{u} \middle| \sum_{u=1}^{U} a_{u} = 1, au \ge 0 \right\}$$
(19)

The above M_o and M_c are the MMD matrices for conditional distribution adaptation and edge distribution adaptation, respectively, denoted as

$$M_{0} = \begin{cases} \frac{1}{n^{2}} & (z_{i}, z_{j} \in D_{s}) \\ \frac{1}{m^{2}} & (z_{i}, z_{j} \in D_{t}) \\ -\frac{1}{mn} & (\text{otherwise}) \end{cases}$$
(20)

$$M_{c} = \begin{cases} \frac{1}{n_{c}^{2}} & \left(z_{i}, z_{j} \in D_{s}^{(c)}\right) \\ \frac{1}{m_{c}^{2}} & \left(z_{i}, z_{j} \in D_{t}^{(c)}\right) \\ -\frac{1}{m_{c}n_{c}} \left(\begin{bmatrix} z_{i} \in D_{s}^{(c)}, z_{j} \in D_{t}^{(c)} \\ z_{i} \in D_{t}^{(c)}, z_{j} \in D_{s}^{(c)} \end{bmatrix} \right) \\ 0 & (\text{otherwise}) \end{cases}$$
(21)

The output of its hidden layer is used as the depth features, and the deep feature variability is measured and characterized under fluctuating working conditions. DMKMMD is employed for the feature distributions of the *s* source domains. The data are projected into the Hilbert space, where mean matching is computed over the space, enabling estimation of the variability between edge distributions and avoiding the computation of intermediate probability densities. Let h^s and h^t denote the two types of depth features of work fluctuations or cross-bearing, respectively. The DMKMMD distance based on the two types of features can be expressed as

$$\ell_{DMK-MMD}(h^{s}, h^{t}) = \left\| \frac{1}{M_{s}} \sum_{i=1}^{M_{s}} \phi(h_{i}^{s}) - \frac{1}{M_{t}} \sum_{i=1}^{M_{t}} \phi(h_{i}^{t}) \right\|_{H}^{2}$$
(22)

where $\|\cdot\|_H$ denotes the regenerative kernel Hilbert space and $\phi(\cdot)$ represents a series of feature mapping functions $k(h^s, h^t) = \langle \phi(h^s), \phi(h^t) \rangle$ associated with the kernel mapping, $k_l(h^s, h^t)$ defined as a convex combination of l underlying kernels. Then the joint objective function with min-max optimization training strategy can be rewritten as

$$\min_{C,Q,P} \max_{D} \ell_{GCNs} + \lambda_1 \ell_{DMKMMD} + \lambda_2 \ell_{GAN}$$
(23)

where ℓ_{MTGAE} is the GCN loss, ℓ_{DMKMMD} is the DMKMMD loss, and ℓ_{GAN} is the adversarial loss.

Accordingly, the training phase focuses on the dynamic learning and optimization of the diagnosis model by using domain data input to the built algorithm framework, and then it is employed by generating adversarial learning strategies and the distribution variability between DMKMMD-aligned domains, the joint objective function and the parameter optimization can be displayed as

$$\hat{\theta}_{E} = \arg\left\{ \min_{\theta_{E}} L_{C}(\theta_{E}, \hat{\theta}_{C}), \max_{\theta_{E}} L_{D}(\theta_{E}, \hat{\theta}_{D}) \right\}
\hat{\theta}_{C} = \arg\min_{\theta_{C}} L_{C}(\hat{\theta}_{E}, \theta_{C})
\hat{\theta}_{D} = \arg\min_{\theta_{D}} L_{D}(\hat{\theta}_{E}, \theta_{D})$$
(24)

3.3. The Proposed Intelligent Health Assessment Method for Aviation Bearings

The proposed health assessment method can be divided into two main steps: signal processing based on order ratio analysis and intelligent health assessment based on DTGCN algorithm.

3.3.1. Signal Processing Based on Simultaneous Extraction Transform-Order Ratio Analysis

First, the time domain signal is low-pass filtered and downsampled to remove the high-frequency interference components. Then, the time-frequency spectrum of the downsampled signal is obtained by using the synchronous extraction transform. Next, the instantaneous frequency curve $f_{Cs}(k)$ of the rotational frequency is obtained from the time-frequency spectrum by defining the peak search (*k* is the sampling point number), and then the instantaneous frequency curve is fitted by reconstructing the least-squares fitting method. Then, the instantaneous frequency curve $f_i(t)$ is obtained from the instantaneous frequency curve. Finally, the original data are resampled by the phase identification time scale at equal angles to obtain the quasi-steady-state angular signal sequence $x(T_n)$.

$$x(T_n) = \sum_m x(m\Delta t_s)h_s(m\Delta t_s - T_n)$$
(25)

where Δt_s is the time domain sampling interval, $m\Delta t_s$ is the value around T_n , and $h_s(t)$ is the different filter. Finally, the order ratio spectrum analysis is obtained.

3.3.2. Feature Extraction and Intelligent Health Assessment Based on DTGCN Algorithm

The above transformed quasi-smooth spectral signals are firstly input into the DTGCN algorithm. The fault information under fluctuating operating conditions is represented through the graph learning layer. The shared high dimensions of the source and target domains are extracted using the graph attention, graph convolution, and graph pooling layers (i.e., the improved 1DGCN), respectively. Afterward, DMKMMD and domain adversarial learning mechanisms are employed. Finally, the training of the health assessment algorithm under fluctuating conditions is implemented by the label predictor and domain discriminator.

Specifically, the flow chart of the DTGCN-based intelligent health assessment method for aviation bearings under large speed fluctuation conditions is described in Figure 5.

Figure 5. Flow chart of DTGCN-based intelligent health assessment method for aviation bearings.

4. Validation and Analysis

4.1. Description for Aviation Bearing Fault Simulation Test Bench

The data used in this section are from the aero engine high-speed bearing fault simulation test bench at the Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Italy [29]. Specifically, the vibration acceleration data of aero bearings at different high speeds with other loads can be measured. A realistic view of the test rig is described in Figure 6. Accordingly, B1, B2, and B3 are three bearing supports. The triaxial vibration acceleration sensors are installed at A1 and A2, respectively. The vibration signal from the damaged bearing supports B1 and B2 can be measured and subjected to the maximum external load. The failure simulation process was roughly displayed as follows: First, the bearing was run briefly at 100 Hz rpm (6000 r/min) under no load, and after the correct installation, the magnitude of the external load was gradually changed, and the external load was slowly changed and increased in steps of 100 Hz to realize the modal simulation of speed fluctuation conditions. The rotational speed and external load of aviation bearing fault simulation are given in Table 1.

(a) Aero-engine high-speed bearing fault simulation tester

(b) Measurement and test systems

(c) Aero-engine high-speed bearing

Figure 6. Aero engine high-speed bearing fault simulation test bench.

Table 1. Test load and speed conditions.

Rated Load/N	0	1000	1400	1800
Rotating speed (r/min)	6000	6000	6000	6000
	12,000	12,000	12,000	12,000
	18,000	18,000	18,000	18,000
	24,000	24,000	24,000	/
	30,000	30,000	/	/

4.2. Construction of Aviation Bearing Dataset

To simulate the actual conditions of aero engine bearings operating at high speed and heavy load for a long time, vibration data under three loads (0 N, 1000 N, and 1400 N, respectively, A, B, and C data sets under fluctuating working conditions with speed fluctuation from 6000 rpm to 24,000 rpm) are selected to verify the effectiveness of the proposed method. Among them, the specific descriptions related to the typical experimental data set T1 are described in Table 2. Accordingly, the time domain and frequency domain waveform of specific speed fluctuations under the aviation bearing normal H1 and serious roller failure H5 are drawn in Figure 7.

Labels	Damaged Areas	Diameter (um)	Rotational Speed (r/min)	Number of Samples	
H1	No	No	6000~24,000	1000	
H2	Inner ring	450	6000~24,000	1000	
H3	Inner ring	250	6000~24,000	1000	
H4	Inner ring	150	6000~24,000	1000	
H5	Roller	450	6000~24,000	1000	
H6	Roller	250	6000~24,000	1000	
H7	Roller	150	6000~24,000	1000	

Table 2. Description of aviation bearing data set T1 (load = 0 N).

Figure 7. Time domain and frequency domain waveforms of H1 and H5 under rotational speed fluctuation.

4.3. Experimental Settings

To verify the effectiveness and generalization of the proposed method, cross-validation transfer tasks between different loads and rotational speeds are set. The specific transfer task numbers and rotational speed settings are described in Table 3. Accordingly, the number of samples in both source and target domains is set to 1000, and the sample length is 1024. FFT is the pre-processing data method to transform the original vibration signal into the frequency domain.

Table 3. Different transfer task settings.

Source Domain Data Set	Transfe	er Tasks
Training sample set A	$A \rightarrow B$ (T1)	A→C (T2)
Training sample set B	В→С (ТЗ)	B→A (T4)
Training sample set C	C→A (T5)	С→В (Т6)

First, the main parameters of the constructed DTGCN algorithm model and the operating environment need to be briefly introduced and analyzed. Generally speaking, the number of network layers and the main parameters of the designed DTGCN are described in Table 3. In the following study, 10 trials of each experiment were performed for averaging to reduce the effect of randomness. The software tool used to run the fault diagnosis program was the PyCharm framework. Python was chosen as the programming language, and the programming framework for the deep learning algorithm was TensorFlow 1.2.0. The computer platform used for fault diagnosis was an i7-6700 with 8 cores and 8 GB of RAM, running in a Windows 64-bit operating system. The parameters of the weight initialization network and the model weights were updated using the Adam optimizer. A Dropout of 50% was implemented to reduce over-fitting.

4.4. Experimental Results

To further illustrate and analyze the proposed method's effectiveness, the parametersharing mechanism selected 1dGCN as the benchmark comparison algorithm for DTGCN. Figure 8 shows the diagnostic confusion matrix of this diagnostic method based on DT-GCNs and migrated 1dGCNs algorithm under different migration task conditions. The proposed method effectively classifies aerospace bearing faults under various operating conditions. Additionally, Figure 9 shows the training error curves under fluctuating operating conditions with different algorithms.

Figure 8. Diagnostic confusion matrix of the different methods at different tasks: (**a**) DTGCN (0.938, T1); (**b**) DTGCN (0.912, T2); (**c**) 1dGCN (0.833, T1); (**d**) 1dGCN (0.800, T2).

Figure 9. Training error curves of the proposed method and the standard algorithm: (**a**) DTGCN; (**b**) 1dGCN.

To explore the characteristics of the feature maps in each layer of the deep neural network, the commonly used t-distributed stochastic neighbor embedding (t-SNE) [30–36] was used to reduce the high-dimensional features of the test samples. As shown in Figure 10a, compared with the ordinary 1DGCN model, the constructed DTGCN model reflects higher separability among testing samples under different health states, which verifies from the perspective of the distribution of test samples in the feature space the improvement of the diagnostic accuracy.

Figure 10. Visualization of low-dimensional embedded testing samples of different methods under fluctuating working conditions: (**a**) DTGCN; (**b**) 1dGCN.

4.5. Comparison with Other Relevant Methods

CNN [6], GCN [17], Transfer Component Analysis (TCA) [30], JDA (Joint Distribution Adaptation) [31], Deep Adaptation Networks (Deep Adaptation Network (DAN) [32], and other DTGCN comparison methods were used to verify the effectiveness of the suggested method. Specifically, a preliminary normalization of the aforementioned datasets can be performed to improve the models' diagnostic effectiveness, and we chose T1, T2, and T3 as specific training tasks multiple times as average results.

In addition, sensitivity and specificity were further introduced to evaluate the proposed model's ability to predict different fault types. A positive and negative two-level problem was used as a better explanation. True positives (TP) are the number of correctly identified positive samples. False negatives (FN) are the number of samples incorrectly classified as positive. False Positive (FP) is the number of samples incorrectly classified as positive. True negative (TN) is the number of corrected and classified as negative samples. Generally, sensitivity and specificity can be defined as follows.

erany, sensitivity and specificity can be defined as follows.

Sensitivity =
$$(TP/(TP + FN))$$
 (26)

Specificity =
$$TN/(TN + FP)$$
 (27)

where sensitivity measures the proportion of actual positives correctly identified, and specificity measures the ratio of real negatives correctly identified. Good models usually have high sensitivity and high specificity. Finally, the specific diagnostic results of the

above-mentioned diagnostic models for the above three tasks are displayed in Figure 11. Accordingly, DTGCN achieves better migration diagnosis results in different migration tasks with an average accuracy of 95.57%. Owing DTGCN utilizes the nearest neighbor extraction performance of the geometric structure of 1dGCNs, which has a powerful transfer learning capability.

Figure 11. Comparison of diagnostic results of different diagnostic methods on the aerospace bearing dataset.

To validate the superiority of the proposed method, the above-mentioned related algorithms were selected for comparison with the proposed DTGCN algorithm. In addition, it should be noted that the variable condition dataset utilized here is between 6000 and 24,000 rpm. The model parameters are mainly tuned several times according to the above experiments. T4, T5, and T6 are selected as migration tasks. The training and testing sample subsets are input to the above six types of fault diagnosis methods for diagnosis to obtain the diagnosis recognition accuracy after multiple averaging. To increase the difficulty of diagnosis and recognition, Gaussian white noise with different degrees of signal-to-noise ratio (SNR) is added to the original fluctuating working condition aviation-bearing fault data set to simulate the actual industrial fault diagnosis situation [30–33]. Meanwhile, 60% of the new data set after noise addition is randomly divided into training sample sets, and the remaining 40% is the testing sample set. The diagnostic recognition accuracy of the above six diagnostic methods with different degrees of random noise interference is displayed in Figure 12.

Figure 12. Average accuracy under different random noise.

From Figure 12, it can be seen that the recognition rate of all diagnostic methods decreases with increasing random noise interference. Still, the overall noise resistance performance of the DTGCN model is stable. Some slight random noise interference before SNR = 5 has little effect on DTGCN. The traditional variable condition fault diagnosis methods (such as CNN and GCN) are more influenced by random noise, and the multiscale fault diagnosis methods have a specific good effect on fluctuating conditions. In summary, the DTGCN algorithm based on the diagnosis method is more stable and noise resistant than other traditional diagnosis frameworks.

5. Conclusions

In future work, extensive research and experiments will be conducted to improve the CNN model further and fuse data from multiple sensors to improve the diagnostic accuracy and robustness of the method. An intelligent health assessment method based on a deep transfer graph convolutional network is proposed to solve the poor diagnostic effect of aviation bearings under large speed fluctuation conditions. First, a novel DTGCN algorithm is designed, which uses an adaptive domain mechanism to enhance the convolutional graph network to have strong geometric feature extraction performance in the graph domain, which is a more robust generalization performance in transfer characteristics. The aviation-bearing fault data set under fluctuating conditions validate and advance the proposed diagnostic algorithm and method. Accordingly, the experimental results illustrated that the proposed method has higher diagnostic accuracy and robustness, which can eliminate the variability of health state sample distribution under large speed fluctuation conditions. Future work will continue to conduct in-depth research on the diagnosis of cross-machine and cross-part generalization and transfer of key aero engine components under extreme working conditions.

Author Contributions: X.Z. (Xiaoli Zhao) proposed the main idea and finished the draft manuscript; X.Z. (Xingjun Zhu), J.Y. and W.D. participated in the discussion of diagnostic thinking and provided correction of English grammar; Y.C. and P.D. drew the figures and tables; M.J. conceived the experiments; and H.S. analyzed the data. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant No. 52205062), the Natural Science Foundation of Jiangsu Province (Grant No BK20220950), the Fundamental Research Funds for the Central Universities (Grant No.30922010701), the National Key R&D Program of China (Grant No.2021YFB2011300), Funded by Open Foundation of the State Key Laboratory of Fluid Power and Mechatronic Systems(No.GZKF-202218), and also Funded by China Postdoctoral Science Foundation (Grant No.2022M711624), (Corresponding author: Jianyong Yao).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Zhao, M.; Zhong, S.; Fu, X.; Tang, B.; Dong, S.; Pecht, M. Deep Residual Networks with Adaptively Parametric Rectifier Linear Units for Fault Diagnosis. *IEEE Trans. Ind. Electron.* 2021, 68, 2587–2597. [CrossRef]
- 2. Chao, Q.; Zhang, J.; Xu, B.; Huang, H.; Pan, M. A review of high-speed electro-hydrostatic actuator pumps in aerospace applications: Challenges and solutions. *J. Mech. Des.* **2019**, *141*, 050801. [CrossRef]
- Chen, Z.; Liao, Y.; Li, J.; Huang, R.; Xu, L.; Jin, G.; Li, W. A Multi-Source Weighted Deep Transfer Network for Open-Set Fault Diagnosis of Rotary Machinery. *IEEE Trans. Cybern.* 2023, 53, 1982–1993. [CrossRef] [PubMed]
- 4. Zhao, X.; Yao, J.; Deng, W.; Ding, P.; Zhuang, J.; Liu, Z. Multi-Scale Deep Graph Convolutional Networks for Intelligent Fault Diagnosis of Rotor-Bearing System under Fluctuating Working Conditions. *IEEE Trans. Ind. Inform.* **2022**, *19*, 166–176. [CrossRef]
- 5. Chen, M.; Shao, H.; Dou, H.; Li, W.; Liu, B. Data Augmentation and Intelligent Fault Diagnosis of Planetary Gearbox Using ILoFGAN Under Extremely Limited Samples. *IEEE Trans. Reliab.* **2022**, *early access*, 1–9. [CrossRef]

- 6. Cao, H.; Shao, H.; Zhong, X.; Deng, Q.; Yang, X.; Xuan, J. Unsupervised domain-share CNN for machine fault transfer diagnosis from steady speeds to time-varying speeds. *J. Manuf. Syst.* **2022**, *62*, 186–198. [CrossRef]
- Zhao, X.; Yao, J.; Deng, W.; Jia, M.; Liu, Z. Normalized Conditional Variational Auto-Encoder with adaptive Focal loss for imbalanced fault diagnosis of Bearing-Rotor system. *Mech. Syst. Signal Process.* 2022, 170, 108826. [CrossRef]
- 8. Liu, Z.; Tang, X.; Wang, X.; Mugica, J.E.; Zhang, L. Wind Turbine Blade Bearing Fault Diagnosis under Fluctuating Speed Operations via Bayesian Augmented Lagrangian Analysis. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4613–4623. [CrossRef]
- 9. Xiang, Z.; Zhang, X.; Zhang, W.; Xia, X. Fault diagnosis of rolling bearing under fluctuating speed and variable load based on TCO spectrum and stacking auto-encoder. *Measurement* **2019**, *138*, 162–174. [CrossRef]
- 10. Lu, S.; Yan, R.; Liu, Y.; Wang, Q. Tacholess Speed Estimation in Order Tracking: A Review with Application to Rotating Machine Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 2315–2332. [CrossRef]
- 11. Fan, W.; Cai, G.; Zhu, Z.; Shen, C.; Huang, W.; Shang, L. Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction. *Mech. Syst. Signal Process.* **2015**, *56*, 230–245. [CrossRef]
- 12. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436-444. [CrossRef] [PubMed]
- Shao, H.; Li, W.; Cai, B.; Wan, J.; Xiao, Y.; Yan, S. Dual-Threshold Attention-Guided Gan and Limited Infrared Thermal Images for Rotating Machinery Fault Diagnosis under Speed Fluctuation. *IEEE Trans. Ind. Inform.* 2022, *early access*, 1–10. [CrossRef]
- Zhao, X.; Yao, J.; Deng, W.; Ding, P.; Ding, Y.; Jia, M.; Liu, Z. Intelligent Fault Diagnosis of Gearbox under Variable Working Conditions with Adaptive Intraclass and Interclass Convolutional Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* 2022, *early access*, 1–15. [CrossRef]
- 15. Wang, J.; Miao, J.; Wang, J.; Yang, F.; Tsui, K.L.; Miao, Q. Fault diagnosis of electrohydraulic actuator based on multiple source signals: An experimental investigation. *Neurocomputing* **2020**, *417*, 224–238. [CrossRef]
- 16. Chen, Z.; Xu, J.; Alippi, C.; Ding, S.X.; Shardt, Y.; Peng, T.; Yang, C. Graph neural network-based fault diagnosis: A review. *arXiv* **2021**, arXiv:2111.08185.
- Scarselli, F.; Gori, M.; Tsoi, A.C.; Hagenbuchner, M.; Monfardini, G. The graph neural network model. *IEEE Trans. Neural Netw.* 2008, 20, 61–80. [CrossRef]
- 18. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. arXiv 2017, arXiv:1710.10903.
- 19. Kipf, T.N.; Welling, M. Variational graph auto-encoders. arXiv 2016, arXiv:1611.07308.
- Li, T.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Multi-receptive Field Graph Convolutional Networks for Machine Fault Diagnosis. *IEEE Trans. Ind. Electron.* 2020, 68, 12739–12749. [CrossRef]
- 21. Zhao, X.; Jia, M.; Liu, Z. Semisupervised Graph Convolution Deep Belief Network for Fault Diagnosis of Electormechanical System with Limited Labeled Data. *IEEE Trans. Ind. Inform.* **2021**, *17*, 5450–5460. [CrossRef]
- Liao, Y.; Huang, R.; Li, J.; Chen, Z.; Li, W. Deep Semisupervised Domain Generalization Network for Rotary Machinery Fault Diagnosis under Variable Speed. *IEEE Trans. Instrum. Meas.* 2020, 69, 8064–8075. [CrossRef]
- Li, X.; Zhang, W.; Ma, H.; Luo, Z.; Li, X. Deep learning-based adversarial multi-classifier optimization for cross-domain machinery fault diagnostics. J. Manuf. Syst. 2020, 55, 334–347. [CrossRef]
- Cheng, J.; Liang, R.; Liang, Z.; Zhao, L.; Huang, C.; Schuller, B. A Deep Adaptation Network for Speech Enhancement: Combining a Relativistic Discriminator with Multi-Kernel Maximum Mean Discrepancy. *IEEE/ACM Trans. Audio Speech Lang. Process.* 2021, 29, 41–53. [CrossRef]
- 25. Yang, C.; Liu, J.; Zhou, K.; Yuan, X.; Ge, M.F. Transfer Graph-Driven Rotating Machinery Diagnosis Considering Cross-Domain Relationship Construction. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 5351–5360. [CrossRef]
- Li, T.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Domain Adversarial Graph Convolutional Network for Fault Diagnosis under Variable Working Conditions. *IEEE Trans. Instrum. Meas.* 2021, 70, 3515010. [CrossRef]
- Zhu, H.; He, Z.; Xiao, Y.; Wang, J.; Zhou, H. Bearing Fault Diagnosis Method Based on Improved Singular Value Decomposition Package. Sensors 2023, 23, 3759. [CrossRef]
- Zhang, W.; Li, X.; Ma, H.; Luo, Z.; Li, X. Universal Domain Adaptation in Fault Diagnostics with Hybrid Weighted Deep Adversarial Learning. *IEEE Trans. Ind. Inform.* 2021, 17, 7957–7967. [CrossRef]
- Daga, A.P.; Fasana, A.; Marchesiello, S.; Garibaldi, L. The Politecnico di Torino rolling bearing test rig: Description and analysis of open access data. *Mech. Syst. Signal Process.* 2019, 120, 252–273. [CrossRef]
- Li, J.; Huang, R.; He, G.; Liao, Y.; Wang, Z.; Li, W. A Two-Stage Transfer Adversarial Network for Intelligent Fault Diagnosis of Rotating Machinery with Multiple New Faults. *IEEE/ASME Trans. Mechatron.* 2021, 26, 1591–1601. [CrossRef]
- 31. Han, T.; Liu, C.; Yang, W.; Jiang, D. Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application. *ISA Trans.* 2020, *97*, 269–281. [CrossRef] [PubMed]
- Long, M.; Cao, Y.; Wang, J.; Jordan, M. Learning transferable features with deep adaptation networks. *Int. Conf. Mach. Learn.* PMLR 2015, 37, 97–105.
- Toma, R.N.; Prosvirin, A.E.; Kim, J.M. Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers. *Sensors* 2020, 20, 1884. [CrossRef] [PubMed]
- 34. Liu, D.; Zhong, S.; Lin, L.; Zhao, M.; Fu, X.; Liu, X. Highly imbalanced fault diagnosis of gas turbines via clustering-based downsampling and deep siamese self-attention network. *Adv. Eng. Inform.* **2022**, *54*, 101725. [CrossRef]

- 35. Lu, S.; He, Q.; Wang, J. A review of stochastic resonance in rotating machine fault detection. *Mech. Syst. Signal Process.* **2019**, *116*, 230–260.
- Wang, J.; Feng, W.; Chen, Y.; Yu, H.; Huang, M.; Yu, P.S. Visual domain adaptation with manifold embedded distribution alignment. In Proceedings of the 26th ACM International Conference on Multimedia, Seoul, Republic of Korea, 22–26 October 2018; pp. 402–410.

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.