

# Physics-aware dynamic spectral modeling integrated with weakly supervised few-shot learning for planetary gearbox fault diagnosis

Peng Chen<sup>1,2</sup> , Jia Gao<sup>1</sup> , Yaqiang Jin<sup>3,4</sup> , Chengning Zhou<sup>5</sup>, Junyu Qi<sup>6</sup>   
and Changbo He<sup>7,\*</sup> 

<sup>1</sup> College of Engineering, Shantou University, Shantou 515063, Guangdong, People's Republic of China

<sup>2</sup> Key Laboratory of Intelligent Manufacturing Technology, Shantou University, Shantou 515063, Guangdong, People's Republic of China

<sup>3</sup> School of Qilu Transportation, Shandong University, Jinan 250061, Shandong, People's Republic of China

<sup>4</sup> Qingdao Mingserve Tech, Qingdao 266041, Shandong, People's Republic of China

<sup>5</sup> Nuclear Power Institute of China, Chengdu 610213, Sichuan, People's Republic of China

<sup>6</sup> Electronics & Drives, Reutlingen University, Reutlingen 72762, Germany

<sup>7</sup> College of Electrical Engineering and Automation, Anhui University, Hefei 230601, Anhui, People's Republic of China

E-mail: [changbh@ahu.edu.cn](mailto:changbh@ahu.edu.cn), [pengchen@alu.uestc.edu.cn](mailto:pengchen@alu.uestc.edu.cn) and [dr.pengchen@foxmail.com](mailto:dr.pengchen@foxmail.com)

Received 9 March 2025, revised 26 June 2025

Accepted for publication 2 July 2025

Published 9 September 2025



## Abstract

Intelligent fault diagnosis based on deep learning has shown promising results in industrial applications, yet the requirement for large labeled datasets remains a significant limitation in real-world deployments. This paper proposes a novel physics-aware dynamic spectral modeling integrated with weakly supervised few-shot learning (PADSM-WSFL) framework for fault diagnosis in planetary gearboxes. The key innovations include: (1) integration of physics-based modeling with deep learning to enhance feature extraction, (2) a unique combination of weakly supervised and few-shot learning that effectively utilizes abundant unlabeled data while requiring only extremely limited labeled samples, and (3) a graph-based feature extraction module that captures complex fault patterns. The framework consists of three main components: a physics-aware dynamic spectral modeling approach, a graph construct module for feature extraction, and the integration of weakly supervised learning with few-shot learning models. Experimental validation on two machinery fault diagnosis datasets demonstrates that PADSM-WSFL achieves superior robustness and generalization capabilities compared to state-of-the-art methods, providing an effective solution to the critical challenge of limited labeled data in industrial fault diagnosis.

Keywords: planetary gearbox, fault diagnosis, dynamic, weakly supervised learning, few-shot learning

\* Author to whom any correspondence should be addressed.

## 1. Introduction

Planetary gearboxes are crucial components in a wide range of industrial applications, including wind turbines, helicopters, electric motors, and hybrid vehicles, where they ensure efficient performance [1–5]. However, these gearboxes are prone to damage such as wear, cracking, and tooth breakage due to sustained high-speed operation and fluctuating load conditions. If these failures are not promptly detected and addressed, they can pose significant safety hazards. Therefore, developing advanced and effective fault diagnosis methods is essential for the early detection of potential failures, ensuring safe and reliable operation, and enhancing overall operational reliability.

In recent years, data-driven fault diagnosis methodologies, particularly those harnessing the power of deep learning techniques, have undergone rapid and significant advancements. This remarkable progress can be largely attributed to deep learning's exceptional capacity to automatically extract high-level representations from raw signals and subsequently achieve highly accurate diagnostic predictions in an end-to-end manner. Over the past decade, a plethora of impressive algorithms based on deep learning have emerged, encompassing, but not limited to, convolutional neural networks (CNNs) [6, 7], generative adversarial networks (GANs) [8], recurrent neural networks (RNNs) [9] and the Transformer model [10–12]. To illustrate the efficacy of these approaches, several studies have made contributions to the field. For instance, Wang *et al* [13] introduced a novel adaptive normalized convolutional neural network (ANCNN), which demonstrated remarkable performance by achieving over 99.8% diagnostic accuracy and excellent stability in experimental settings. In another significant development, Shi *et al* [14] proposed an innovative bidirectional-convolutional long short-term memory (BiConvLSTM) deep neural network. This advanced model simultaneously extracts spatial and temporal features from vibration and rotational speed measurements, thereby achieving superior accuracy in identifying gearbox faults compared to traditional methods. Furthermore, Chen *et al* [8] made a substantial contribution by introducing a novel threshold self-setting health condition monitoring (HCM) scheme for wind turbine generator bearings. This innovative approach utilizes deep convolutional generative adversarial networks (DCGANs) to automate the threshold-setting process, thereby enhancing monitoring accuracy and mitigating the risk of misdiagnosis. Additionally, Chen *et al* [15] proposed a physics-informed hyperparameters selection strategy for LSTM models, which enhances fault detectability by maximizing the discrepancy between healthy and fault states. Snyder *et al* [16] proposed a novel dual-head ensemble transformer (DHET) algorithm that integrates a 1D Transformer model and a 2D Vision Transformer model to enhance the classification of signals with time-frequency features.

Despite the excellent performance of deep learning-based approaches, they require substantial labeled data for training, which poses challenges for fault diagnosis methods in real-world industrial applications:

1. Analyzing and identifying fault types in mechanical systems is time-consuming and labor-intensive. Consequently, producing large amounts of manually labeled data is impractical and often infeasible.
2. Deep learning-based fault diagnosis heavily relies on substantial labeled fault samples. This requirement poses a significant challenge in real-world applications, as labeling often requires domain experts and equipment disassembly.
3. In real-world industrial equipment, machines rarely operate in faulty states for long periods, resulting in imbalanced data where normal operations dominate and fault signals are scarce.
4. Limited labeled samples frequently lead to overfitting in deep neural networks, resulting in suboptimal parameter optimization and reduced performance on new data.

In light of the persistent challenges associated with fault diagnosis in the context of limited dataset availability, researchers have diligently explored and implemented a range of sophisticated methodologies to mitigate the reliance on extensive supervised data. Among these approaches, traditional techniques such as data augmentation [8] and transfer learning [17] have been widely adopted and refined. Building upon these foundational methods, Luo *et al* [18] have proposed an innovative imbalanced fault diagnosis method that leverages a conditional-deep convolutional generative adversarial network (C-DCGAN). This advanced approach not only enhances feature extraction capabilities but also utilizes generated samples to significantly improve both the accuracy and stability of fault diagnosis processes. Furthermore, Ha and Fink [19] have introduced the novel concepts of scaled CutPaste and FaultPaste, which ingeniously incorporate domain knowledge to create highly realistic faulty samples and effectively scale fault severity. These techniques have demonstrated particular efficacy in the domain of planetary gearbox diagnostics, especially when confronted with domain shifts. Complementing these advancements, Cheng *et al* [20] have developed the Q networks calibrated ensemble (QCE) method, which aims to enhance both the generalization and robustness of cross-domain fault diagnosis. The effectiveness of this approach has been rigorously validated through experiments conducted on a nuclear circulating water pump test bench.

Furthermore, in order to address the aforementioned challenges, such as limited labeled samples, few-shot learning (FSL) has emerged as a promising and viable solution for mitigating these constraints in the field of intelligent fault diagnosis. This approach has garnered significant attention from researchers due to its potential to overcome data scarcity issues that have long hindered the development of robust diagnostic systems. To elucidate the growing importance of FSL in this domain, several noteworthy studies have been conducted, each contributing unique methodologies and insights to the field. For instance, Jin *et al* [21] propose an innovative ordinal classification prototypical networks (OCPN) model, which ingeniously integrates ordinal regression into prototypical networks. This integration aims to enhance fault diagnosis of offshore

wind turbines by simultaneously addressing the issues associated with limited high-quality labeled data and varying fault severity. Consequently, the OCPN model represents a significant advancement in tackling these dual challenges that have historically impeded accurate fault diagnosis in complex industrial systems. In a parallel effort, Yue *et al* [22] introduce a novel multiscale wavelet prototypical network (MWPN) that specifically targets the problem of few-shot fault diagnosis across different rotating components. By leveraging a multiscale wavelet convolution module for feature extraction and a metric meta-learner for distance measurement, their model exhibits superior performance in experimental trials. This approach demonstrates the potential of combining advanced signal processing techniques with meta-learning strategies to improve diagnostic accuracy in scenarios with limited training data. Moreover, Liu *et al* [23] have made a substantial contribution to this domain by proposing the attentional contrastive calibrated transformer (ACCT) for intelligent fault diagnosis. This sophisticated model addresses the challenges of sharp speed variations and limited data through a synergistic combination of convolutional layers and transformers for global dependency modeling. Additionally, it incorporates an innovative unsupervised contrastive learning regularization to enhance feature representation, thereby improving the model's ability to generalize from limited examples.

Inspired by the success of semi-supervised learning (SSL) in other domains, researchers have developed a hybrid approach known as few-shot semi-supervised learning, which merges SSL techniques with few-shot learning paradigms. This approach typically involves pretraining a prototypical network (PN) with limited labeled data, subsequently using SSL methods to select pseudolabels, and then fine-tuning the PN with these pseudolabels. As a result, this innovative combination has shown promise in enhancing generalization and overall performance in fault diagnosis tasks. In alignment with this trend, Lao *et al* [24] present a semi-supervised weighted prototypical network (SSWPN) specifically designed for fault classification in switch machines. This model enhances diagnostic performance in scenarios with limited labeled data by incorporating an effective dual-scale neural network architecture and a novel prototype updating strategy. Thus, the SSWPN represents a significant step forward in adapting few-shot learning techniques to the specific challenges of railway infrastructure maintenance. Furthermore, Su *et al* [25] propose an innovative semi-supervised temporal meta-learning method (SSTML) and a comprehensive deep learning framework (SeMeF) that effectively utilize scarce labeled data in conjunction with abundant unlabeled vibration data for wind turbine bearing fault diagnosis. Their approach demonstrates superior diagnostic accuracy compared to existing models, thereby highlighting the potential of combining temporal meta-learning with semi-supervised techniques in industrial applications.

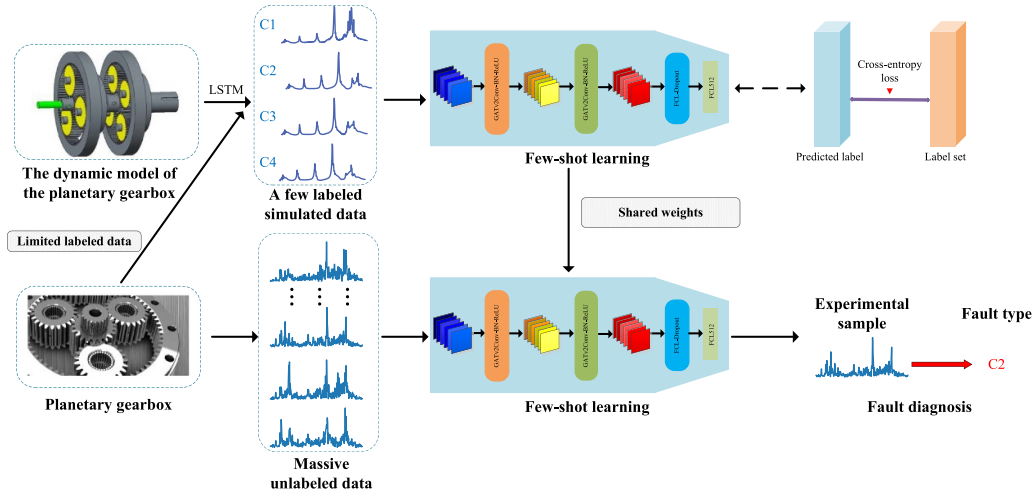
While the aforementioned models have undoubtedly demonstrated considerable efficacy when applied to limited training datasets or scenarios with sparse labeled samples,

there remain significant challenges in the field of few-shot learning. One of the primary obstacles is the difficulty in determining a precise and reliable decision boundary for knowledge transfer. Furthermore, the inherent variability and potential fluctuations in data distribution can potentially lead to a substantial degradation in the performance of few-shot learning methodologies.

For planetary gearboxes operating in real-world applications, substantial amounts of unlabeled data can be readily acquired during the operational process. This abundance of data presents a unique and promising opportunity to address the current challenges associated with few-shot learning methods and, consequently, enhance their robustness and applicability in the field of fault diagnosis and prognosis. The judicious and effective utilization of this unlabeled data, in conjunction with the extremely limited labeled samples typically available, has the potential to significantly improve the performance, accuracy, and reliability of diagnostic and prognostic models. Motivated by these considerations and the pressing need for more efficient and adaptive learning paradigms in the field of mechanical engineering, we propose a novel approach: physics-aware dynamic spectral modeling integrated with weakly supervised few-shot learning (PADSM-WSFL). This innovative methodology, as illustrated in figure 1, is specifically designed and dedicated to resolving the pervasive issue of limited labeled samples by leveraging easily accessible unlabeled data alongside advanced few-shot learning techniques. In the proposed framework, the two adopted learning paradigms—physics-aware dynamic spectral modeling and weakly supervised few-shot learning—are not only complementary but also meticulously designed to capitalize on their respective strengths, thereby creating a synergistic approach to fault diagnosis in planetary gearboxes.

The primary contributions of this paper can be summarized as follows:

1. Development of a novel graph construct module: An innovative graph construct module is designed specifically for extracting intrinsic features under extremely small sample sizes. This module addresses the critical challenge of feature extraction in scenarios where labeled data is scarce, thereby enhancing the overall efficacy of the learning process.
2. Proposition of a physics-aware dynamic spectral modeling approach: The study introduces a sophisticated pure torsional lumped parametric model of a one-stage planetary gear mechanism. In this model, only the torsional motion of each component is incorporated into the modeling process, allowing for a more focused and precise analysis of the system's dynamic properties. This approach enables a comprehensive capture of the dynamic behavior of planetary gearboxes, facilitating a more accurate and nuanced understanding of their operational characteristics.
3. Integration of weakly supervised learning and few-shot learning models: The research creatively combines weakly supervised learning techniques with few-shot learning models for planetary gearbox diagnosis. This innovative



**Figure 1.** The proposed fault diagnosis framework: Integrating physical dynamic models and graph neural networks for extreme limited labeled sample scenarios.

integration fully exploits the information contained in unlabeled data while effectively performing fault diagnosis with extremely limited labeled samples.

4. Leveraging the proposed physics-aware dynamic spectral model and synthesizing simulated vibration data across diverse fault conditions for model training. These physically consistent simulations significantly augment limited labeled samples, mitigating deep neural network overfitting while enhancing generalization capability and diagnostic performance on novel data.

The paper is structured as follows: section 2 reviews pertinent literature on weakly supervised learning and graph neural networks, which form the foundation of this research. Section 3 elaborates on the proposed physics-aware dynamic spectral modeling integrated with weakly supervised few-shot learning. Section 4 presents and analyzes experimental results from two fault diagnosis datasets. Finally, section 5 summarizes the key findings and concludes the paper.

## 2. Related work

In this section, the relevant literature of weakly supervised learning and graph neural networks, which are critical components of the research, will be revisited. Specifically, subsection 2.1 will focus on weakly supervised learning, exploring its methodologies and advancements, while subsection 2.2 will examine the developments and theoretical frameworks surrounding graph neural networks.

### 2.1. Weakly supervised learning for fault diagnosis

Weakly supervised learning for fault diagnosis has garnered considerable attention as a promising approach to address the limitations inherent in scenarios where obtaining fully labeled data is challenging or impractical. Within industrial scenarios, the acquisition of comprehensively labeled fault data can be not only costly and time-intensive but also, in some

instances, impracticable due to safety considerations or the infrequent occurrence of fault events. Consequently, researchers have been exploring innovative methods to overcome these obstacles and enhance the efficacy of fault diagnosis systems.

Compared with weakly supervised learning, semi supervised learning combines a small amount of labeled data and a large amount of unlabeled data, relying on structural information in unlabeled data and supervised signals in labeled data to improve performance. It is suitable for scenarios with rich unlabeled data but high cost of high-quality labeling. In summary, although both weakly supervised learning and semi supervised learning aim to reduce reliance on large-scale annotated data, the former focuses more on learning from imperfect labels, while the latter focuses on effectively utilizing rich unlabeled data resources.

In light of these challenges, several noteworthy studies have emerged, each contributing to the advancement of weakly supervised learning in fault diagnosis. For instance, Ruan *et al* [26] introduce an enhanced non-local weakly supervised fault diagnosis method that ingeniously incorporates few-shot learning to augment training tasks and refine feature extraction in convolutional neural networks (CNNs). This innovative approach has demonstrated significantly improved fault diagnosis accuracy when compared to conventional methods, particularly when applied to rolling bearing and bevel gear datasets. Similarly, Guo *et al* [27] propose a cost-effective, deep learning-based solution that leverages weakly supervised learning to directly estimate iron ore feed load from ore pellet images. This method effectively addresses challenges associated with data annotation and image size, while simultaneously exhibiting competitive performance for real-time optimization of the grinding process. Further expanding the field, Yan *et al* [28] present a novel digital twin-assisted framework designed to tackle imbalanced fault diagnosis. This approach successfully overcomes the limitations of existing methods by generating high-fidelity simulated fault data and employing a subdomain adaptive mechanism in conjunction with margin-aware regularization. Consequently, the

model's diagnostic performance is substantially enhanced, particularly in scenarios involving highly imbalanced data. Qian and Li [29] introduce an innovative weakly supervised oversampling method for fault diagnosis in industrial systems. This approach enhances the selection of high-quality synthetic samples through the application of graph semi-supervised learning and a cost-sensitive neighborhood component analysis, resulting in improved performance and robustness when applied to highly imbalanced datasets, outperforming existing approaches. Liu *et al* [30] propose an optimal sample selection strategy for weakly supervised visual tracking that enhances model performance by meticulously assessing the reliability of samples through score maps and judiciously replacing unreliable pseudo-labels with reliable ground truth.

Despite the notable domain-specific successes achieved by weakly supervised learning methods in fault diagnosis, it is important to acknowledge their limitations. These approaches often necessitate the incorporation of *a priori* information, which can lead to increased computational demands and potential constraints on model generalization. Furthermore, the diagnostic performance of weakly supervised approaches generally falls short of the benchmarks set by fully supervised models. Notwithstanding these challenges, the field continues to evolve rapidly, and ongoing research is expected to yield more robust and broadly applicable methodologies. As advancements are made in addressing current limitations, it is anticipated that weakly supervised learning will play an increasingly significant role in fault diagnosis applications, particularly in scenarios where the acquisition of fully labeled data remains challenging or impractical.

## 2.2. Graph neural networks

The application of graph neural networks (GNNs) in fault diagnosis is receiving increasing attention, as they are capable of handling non-Euclidean data and intricate relationships, particularly when complete data labeling is challenging. In industry, building detailed fault databases is both expensive and time-consuming. GNNs provide new avenues for fault prediction and learning from limited annotations by effectively spreading and aggregating node features, boosting the performance of diagnostic systems. Consequently, both academia and industry are actively engaged in GNNs model innovations, propelling fault diagnosis technology forward.

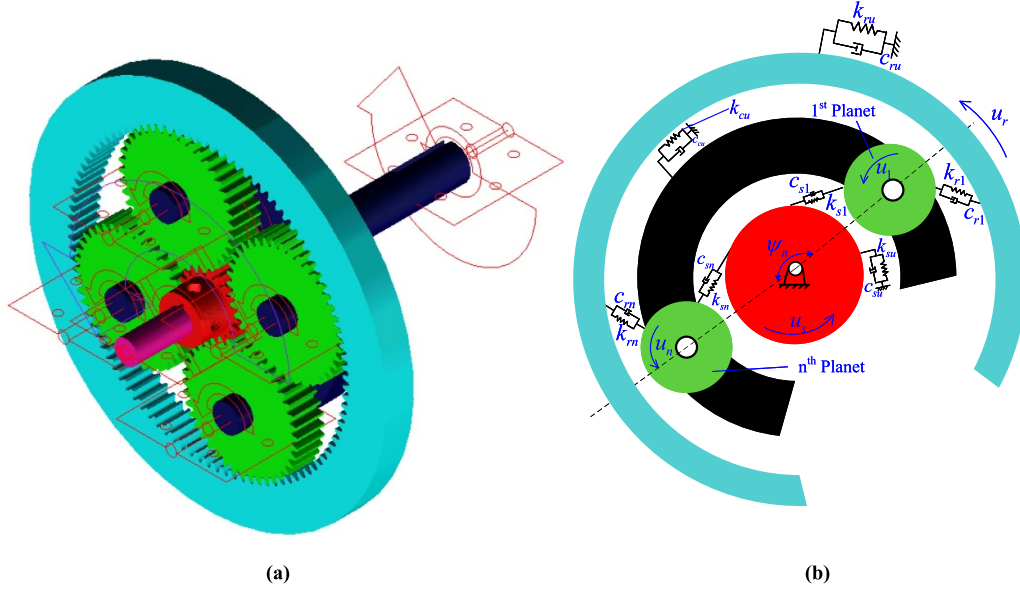
In view of the aforementioned obstacles, a number of prominent research efforts have surfaced, collectively driving forward the progress of GNNs in the domain of fault diagnosis. Specifically, Jiao *et al* [31] introduce an ensemble of simplified graph wavelet neural networks (SGWNN) for enhanced fault diagnosis in planetary gearboxes, addressing the challenge of scarcity of faulty samples in real-world scenarios. By utilizing diverse wavelet bases for feature extraction and a learnable weighting ensemble strategy, their method significantly outperforms conventional CNNs and other GNNs variants in terms of both accuracy and robustness. In a similar vein, Xu *et al* [32] present a novel graph-guided collaborative

convolutional neural network (GGCN) designed for enhanced fault diagnosis in electromechanical systems. This approach features a graph reasoning fusion module that effectively explores both modality-specific and shared features across multi-source signals. Consequently, it surpasses seven state-of-the-art (SOTA) methods, especially in noisy environments. Furthermore, Chen *et al* [33] introduce a Semi-supervised self-correcting graph neural network (SSGNN) for fault diagnosis in rotating machinery. Their method constructs a graph-structured representation from vibrational signals, utilizes an improved state transform algorithm for state propagation, and applies an alternative learning method based on the expectation-maximization (EM) algorithm to optimize feature extraction and graph structure. This approach demonstrates superior accuracy and convergence speed over existing methods. Additionally, Yu *et al* [34] propose a graph-weighted reinforcement network (GWRNet) that accurately diagnoses faults in rotating machines under conditions of small sample sizes and strong noise. By constructing an adjacency matrix for pre-classification and dynamically enhancing the node feature aggregation strategy to suppress noise, the effectiveness of their method is verified using datasets from the drivetrain diagnostics simulator (DDS) test rig and wind turbine gearboxes. Yin *et al* [35] develop a multiscale graph convolutional network (MS-GCN) for enhanced fault diagnosis in rolling bearings. This approach features a multiscale feature extraction module for comprehensive discovery of signal regularities, a multiscale graph iteration module for retaining local features while mining global information, and a mutual fusion module based on Bayesian methods. Collectively, these innovations significantly improve diagnostic accuracy over current SOTA methods.

GNNs have demonstrated remarkable domain-specific successes in various applications, yet they also present prominent limitations. One significant challenge is the high computational demand associated with GNNs, as they often require substantial processing power and memory to handle large-scale graph data. Additionally, the performance of GNNs can be highly sensitive to the quality and structure of the input graph, making them susceptible to issues arising from noisy or incomplete data. Despite these challenges, the field of GNNs is rapidly advancing, with ongoing research focused on developing more efficient algorithms and robust architectures. As improvements are made in these areas, GNNs are expected to become increasingly integral to applications requiring complex relational data analysis, particularly in scenarios where traditional neural network models struggle to capture the underlying graph structures.

## 3. Semi-supervised consistency models

In this section, the research present a comprehensive illustration of the proposed PADSM-WSFL framework. The PADSM-WSFL architecture is composed of two principal sub-architectures, each playing a crucial role in the



**Figure 2.** (a) The dynamic model of the planetary gearbox transmission (b) Schematic of planetary gear sets under pure torsion, with mesh interactions represented by linear spring-damper systems.

overall diagnostic process: Physics-Aware Dynamic Spectral Modeling (PADSM), see discussed in section 3.1, and Few-Shot Signal Representation Learning (FSSRL), detailed in section 3.2. Initially, the PADSM framework is engineered to generate highly accurate synthesized vibration signals by meticulously modeling the dynamic behavior of a Planetary Gearbox (PGB). This physics-based approach ensures that the generated signals closely mimic real-world vibration patterns, including various fault conditions. Ultimately, this approach enables the effective and efficient fault diagnosis of the PGB. By leveraging both physical modeling principles and advanced weakly supervised learning techniques, the PADSM-WSFL framework achieves robust diagnostic performance, even in scenarios with limited real-world fault data.

### 3.1. Physics-aware dynamic spectral modeling

In order to achieve highly accurate synthesized vibration of various types of failures, particularly the dynamic spectral characteristics of Planetary Gearboxes (PGBs), this research employs physics-aware dynamic spectral modeling to comprehensively capture the dynamic behavior of PGBs. The model utilized in this study is a sophisticated pure torsional lumped parametric model of a one-stage planetary gear mechanism, wherein only the torsional motion of each component is incorporated into the modeling process. This approach allows for a more focused and precise analysis of the system's dynamic properties.

The system under investigation comprises one Sun gear (s), one ring gear (r), one carrier (c), and N planet gears (p), collectively resulting in (3+N) degrees of freedom, as illustrated in figure 2. To accurately represent the interactions between components, the meshes of Sun-planet and ring-planet pairs

are modeled as linear springs and dampers, denoted by  $k_{hn}$  and  $c_{hn}$  respectively, where  $h = s$  or  $r$ , and  $n = 1, 2, \dots, N$ . It is important to note that, for the purposes of this study, non-linear factors of the gear mesh, such as backlash and transfer errors, have been intentionally omitted from the dynamic modeling procedure to maintain model simplicity and focus on the primary dynamics. Furthermore, the torsional stiffness and damping of individual parts are represented by  $k_{hu}$  and  $c_{hu}$  respectively, where  $u = s, r, c$ . These parameters play a crucial role in determining the system's overall dynamic response and are essential for accurate modeling. By meticulously formulating the differential equations of rigid body motion for each component, the comprehensive governing equation of the system can be derived. This equation encapsulates the complex interactions and dynamics of the entire planetary gearbox system, providing a solid foundation for further analysis and simulation of various failure modes and operating conditions.

$$\begin{aligned} \text{Carrier : } & \left( \frac{I_c}{r_c^2} + \sum_{n=1}^N m_n \right) \ddot{u}_c + c_{cu} \dot{u}_c \\ & + k_{cu} u_c - \cos \alpha \left[ \sum_{n=1}^N c_{sn} \dot{\delta}_{sn} + \sum_{n=1}^N c_{rn} \dot{\delta}_{rn} \right. \\ & \left. + \sum_{n=1}^N k_{sn} \delta_{sn} + \sum_{n=1}^N k_{rn} \delta_{rn} \right] = \frac{T_c}{r_c} \end{aligned} \quad (1)$$

$$\begin{aligned} \text{Ring : } & \left( \frac{I_r}{r_r^2} \right) \ddot{u}_r + c_{ru} \dot{u}_r + k_{ru} u_r \\ & + \sum_{n=1}^N c_{rn} \dot{\delta}_{rn} + \sum_{n=1}^N k_{rn} \delta_{rn} = \frac{T_r}{r_r} \end{aligned} \quad (2)$$

$$\text{Sun} : \left( \frac{I_s}{r_s^2} \right) \ddot{u}_s + c_{su} \dot{u}_s + k_{su} u_s + \sum_{n=1}^N c_{sn} \dot{\delta}_{sn} + \sum_{n=1}^N k_{sn} \delta_{sn} = \frac{T_s}{r_s} \quad (3)$$

$$n\text{-th Planet} : \left( \frac{I_n}{r_n^2} \right) \ddot{u}_n - c_{rn} \dot{\delta}_{rn} - k_{rn} \delta_{rn} + c_{sn} \dot{\delta}_{sn} + k_{sn} \delta_{sn} = 0 \quad (4)$$

where  $u_h$  represents the torsional displacement of the corresponding components, serving as a crucial variable in describing the rotational motion of each gear element. The mass and moment of inertia, denoted by  $m_h$  and  $I_h$  respectively, are fundamental properties that influence the dynamic response of the system. Furthermore,  $r_h$  signifies the radius of the base circle, which is a critical geometric parameter in gear design and analysis. The gear pressure angle, represented by  $\alpha$ , plays a significant role in determining the force transmission characteristics between mating gears. Additionally,  $T_h$  denotes the applied torque, which is the primary driving force in the system. The relative displacements,  $\delta_{sn} = u_s + u_n - u_c$  and  $\delta_{rn} = u_r - u_n - u_c \cos \alpha$ , are of particular importance as they project onto the action line of the Sun-planet and ring-planet pairs, respectively, thereby capturing the intricate interactions between these components.

The dynamic behavior of the gear mesh stiffness is a complex phenomenon that warrants careful consideration. Due to the cyclical engagement pattern alternating between single and double tooth contacts, coupled with the continuous alteration of the contact location, the mesh stiffness exhibits temporal fluctuations. This time-varying nature of the mesh stiffness significantly influences the overall system dynamics and is crucial for accurate modeling. In analyzing the flexure of a straight-toothed gear tooth, it is instructive to conceptualize it as analogous to a non-homogeneous cantilever beam. The stiffness characteristics of such a system can be effectively gauged through the application of the potential energy approach. This method allows for a comprehensive consideration of various deformation modes, including bending, shear, compression, and contact-induced deformations. By incorporating these effects, the aggregate effective stiffness for the gear mesh can be determined with greater precision. This enhanced accuracy is achieved by utilizing the following equation for computation

$$k_t = \sum_{i=1}^2 \left( \frac{1}{1/k_{h,i} + 1/k_{b1,i} + 1/k_{s1,i} + 1/k_{a1,i} + 1/k_{b2,i} + 1/k_{s2,i} + 1/k_{a2,i}} \right) \quad (5)$$

where the indices  $i = 1, 2$  represent the specific gear pair involved in the meshing process. The parameters  $k_h$ ,  $k_b$ ,  $k_s$  and  $k_a$  are used to denote, respectively, the Hertzian contact stiffness, the bending stiffness, the shearing stiffness, and the axial compressive stiffness of a pair of mating teeth. These stiffness

values are determined as follows:

$$\frac{1}{k_h} = \frac{4(1-\nu^2)}{\pi EL} \quad (6)$$

$$\frac{1}{k_b} = \int_0^d \frac{(x \cos \alpha_1 - h \sin \alpha_1)^2}{EI_x} dx \quad (7)$$

$$\frac{1}{k_s} = \int_0^d \frac{1.2 \cos^2 \alpha_1}{GA_x} dx \quad (8)$$

$$\frac{1}{k_a} = \int_0^d \frac{\sin^2 \alpha_1}{EA_x} dx. \quad (9)$$

where the elastic modulus ( $E$ ), shear modulus ( $G$ ), and Poisson's ratio ( $\nu$ ) are fundamental material properties that govern the tooth's response to applied loads. The tooth width ( $L$ ) is a critical geometric parameter that influences the load distribution along the tooth face. The area moment of inertia ( $I_x$ ) and the cross-sectional area ( $A_x$ ) are essential for calculating the tooth's resistance to bending and axial deformation, respectively. The angle  $\alpha_1$ , defined as the angle between the action line and the line perpendicular to the central line of the tooth, along with the distance  $x$  from the point of applied force to the tooth root, are crucial geometric parameters that affect the stress distribution within the tooth.

### 3.2. Few-shot signal representation learning

In the field of real-world industrial applications, the acquisition of labeled samples often presents a significant financial burden. Moreover, due to various constraints inherent in data collection processes, such as operating conditions or equipment limitations, researchers and practitioners are frequently compelled to rely on simulated samples generated by physics-based dynamic models as proxies for real-world labeled data. These circumstances give rise to a challenging scenario wherein one must contend with two distinct datasets: firstly, a dataset  $D_u = \{X_i\}_{i=1}^N$ , which encompasses a substantial quantity of unannotated data samples, and secondly, a dataset  $D_l = \{\hat{X}_i, Y_i\}_{i=1}^M$ , which contains a limited number of labeled simulated samples. Within this framework,  $X_i$  and  $\hat{X}_i \in \mathfrak{R}^{W \times 1}$  represent data samples with dimensions of  $W \times 1$ , where  $W$  denotes the number of features. The variable  $Y_i$  corresponds to the label associated with the simulated data sample  $\hat{X}_i$ . It is important to note that  $N$  and  $M$  represent the number of data samples in  $D_u$  and  $D_l$  respectively, with the critical distinction that  $M$  is significantly smaller than  $N$ , reflecting the scarcity of labeled data in comparison to the abundance of unlabeled samples.

To address the challenges posed by this data imbalance, researchers have developed advanced machine learning techniques, among which the dynamic graph attention network (DGAT) [36] stands out as a particularly promising approach. The DGAT is a specialized variant of graph neural networks

that incorporates a sophisticated attention mechanism. This network introduces a dynamic attention mechanism through judicious modifications to the original graph attention network (GAT) architecture. The operational principle of the DGAT can be elucidated as follows: Initially, a shared linear transformation, parameterized by the weight matrix  $W$ , is applied uniformly to all nodes in the graph. Subsequently, a shared attention mechanism, denoted as  $F$ , is executed on these transformed nodes. This process culminates in the calculation of dynamic attention scores, which can be formally expressed through the mathematical formula below.

$$e_{ij} = F \cdot \text{LeakyReLU}(W \cdot [h_i || h_j]) \quad (10)$$

where  $h_i$  and  $h_j$  denotes the node feature of node  $i$  and node  $j$ , respectively,  $||$  indicates a concatenation operation, and  $W$  and  $F$  are updated during iterations of the DGAT layer.

Based on the attention scores, the normalized scores  $\alpha_{ij}$  can be calculated, and the final updated node set  $H'$  can be obtained. This process can be mathematically expressed as:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik})} \quad (11)$$

$$h'_i = \sigma \left( \sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j \right) \quad (12)$$

where  $\mathcal{N}(i)$  represents the set of neighboring nodes of node  $i$ ,  $\sigma$  denotes a non-linear activation function.

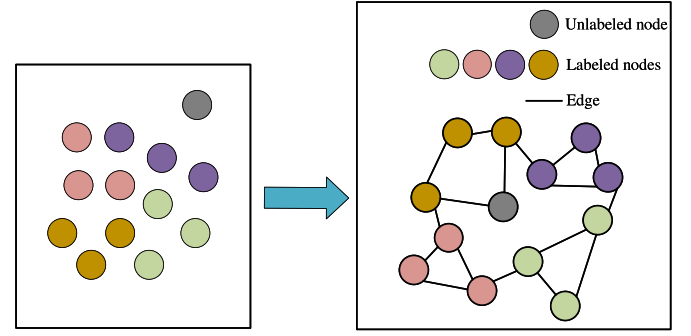
For few-shot signal representation learning, two tasks are considered: a target task  $T_t$  and a pretext task  $P_t$ . The target task  $T_t$  involves fault diagnosis, while the pretext task  $P_t$  focuses on signal transformation classification to learn inherent features from a limited amount of labeled data. To establish the pretext task, it is necessary to modify the input of the original graph neural network, transforming it from a traditional semi-supervised model to a weakly supervised model suitable for few-shot learning. Consequently, the process of graph construction transitions from equations (13) to (14):

$$G = g_{\text{base}}(H_l, H_u) \quad (13)$$

$$G' = g_{\text{improved}}(H_l) \quad (14)$$

where  $g$  represents the graph construction process,  $H_l$  and  $H_u$  denote labeled and unlabeled samples, respectively. However, it is important to note that traditional training processes of graph neural networks often struggle to effectively differentiate faults when the sample size is extremely small. Therefore, it becomes necessary to improve graph construction strategies by incorporating the concept of few-shot matching pairs. Through effective structural design, GNN can effectively utilize all samples appearing in the graph and integrate information between similar nodes through edge connections, thereby extracting more representative node features.

In this specific approach, when the number of samples is extremely limited, each training sample is considered as a



**Figure 3.** Graph construction process for extremely small sample sizes.

target node, and a graph is constructed with all other training samples. This process is then cyclically implemented. The constructed graphs are subsequently inputted into our DGAT network for feature extraction and classification. Similarly, during the prediction phase, the same strategy is employed, wherein a supporting group containing a small number of labeled samples, along with the test samples, is used to construct graphs and make predictions. In a K-way classification task, where K distinct types exist for classification, the graph construction methodology remains consistent during both training and testing phases. The initial procedure entails the creation of a support set  $S$

$$S = \{T_1, T_2, T_3, \dots, T_k\} \quad (15)$$

$$T_i = \{s_{i1}, s_{i2}, s_{i3}, \dots, s_{in}\} \quad (16)$$

where  $T_k$  represents that the support signal sample belongs to the  $k$ th type, and  $s_{in}$  denotes the  $n$ th sample in type  $T_i$ . Given an input test signal sample  $\tilde{X}_t$ , the dataset containing  $K$  different types samples is defined as a few-shot testing set, as shown below

$$S_t = \{T_1, T_2, T_3, \dots, T_k, \tilde{X}_t\} \quad (17)$$

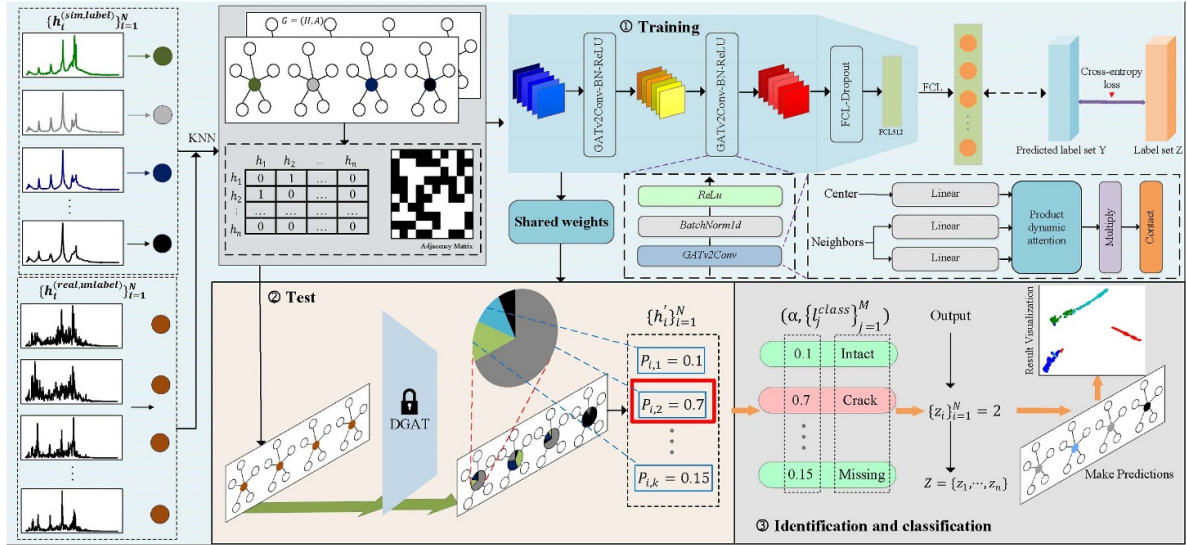
the original loss function has been adjusted, and a function tailored to the aforementioned method has been proposed. This function utilizes the predicted label set  $Z$  and the label set  $Y$ , and is expressed as equation (18)

$$\text{Loss} = - \sum_t^T y_n^{(t)} \ln(z_n^{(t)}) \quad (18)$$

where  $n$  is the specified location of the unknown sample that needs to be predicted,  $y_n^{(t)}$  is the  $t$ -th value of the label  $y_n$ , and  $z_n^{(t)}$  is the  $t$ th value of the predicted label  $z_n$ . The constructing process is illustrated in figure 3.

As the number of available samples increases, it becomes feasible to employ the traditional graph construction method. This approach involves constructing a single graph with all labeled samples. Consequently, the loss function is updated as below

$$\text{Loss} = - \frac{1}{I} \sum_i^I \sum_t^T y_i^{(t)} \ln(z_i^{(t)}) \quad (19)$$



**Figure 4.** The proposed PADSM-WSFL architecture for planetary gearbox diagnosis.

where  $I$  is the label number,  $T$  is the number of types,  $y_i^{(t)}$  is the  $t$ th value of the label  $y_i$ , and  $z_i^{(t)}$  is the  $t$ th value of the predicted label  $z_i$ .

### 3.3. The proposed architecture of PADSM-WSFL

As discussed above, the proposed architecture PADSM-WSFL can be designed based on physics-aware dynamic spectral modeling and few-shot signal representation learning, thereby presenting a new approach to fault diagnosis in planetary gearbox systems. The overarching workflow of this model is meticulously illustrated in figure 4, while the comprehensive algorithm is elaborated in Algorithm 1.

The framework operates in two distinct, yet interconnected stages. In the initial stage, the neural network is subjected to training on a pretext task to extract and learn inherent features from a limited set of labeled simulation signals. Subsequently, the acquired parameters are transferred to an identical network architecture for fault diagnosis using unlabeled real samples, thus bridging the gap between simulated and real-world data. To initiate the first stage, graphs are constructed for all available labeled simulated fault signals, albeit limited in number, by employing the K-nearest neighbors (KNN) algorithm [37]. The specific criteria for graph construction are meticulously described in section 3.2, ensuring a robust and meaningful representation of the signal relationships. These carefully crafted graph structures serve as the bedrock for training the neural network in the fault diagnosis task.

The training process adheres to the tenets of supervised learning, wherein a signal transformation classification neural network, represented as  $f_p(\theta_p)$ , is trained on the pretext task to distinguish various fault types within the simulation signals. Here,  $\theta_p$  symbolizes the trainable parameters of the neural network. These parameters are iteratively refined by

minimizing the loss function  $L_s$ , which is formally defined in equations (18) and (19).

In this research, a sophisticated GNN is developed to perform feature extraction and classification tasks. GNNs are renowned for their ability to aggregate global feature information and exhibit heightened sensitivity to inter-sample similarities based on the underlying graph structure. The proposed neural network architecture incorporates two DGAT layers, each followed by two Batch-Norm (BN) layers. While the DGAT layers are responsible for learning node feature representations, the BN layers serve to accelerate the training process and bolster the model's generalization capabilities. To further enhance the stability and efficacy of the attention mechanism, a multi-headed attention approach is implemented. This enhancement augments the model's feature extraction capabilities, with  $M$  denoting the number of attention heads employed for multiple attention aggregation operations. The final updated output of each layer is mathematically expressed as follows

$$h_i = \sigma \left( \frac{1}{M} \sum_{m=1}^M \sum_{j \in N(i)} \alpha_{ij} e_j \right) \quad (20)$$

where  $\sigma$  represents the Sigmoid activation function,  $M$  is set to 4 in this study,  $N(i)$  denotes the set of all nodes adjacent to node  $i$ ,  $\alpha_{ij}$  quantifies the attention weight of node  $j$  to node  $i$ ,  $e_j$  is the linear transformation of the lower layer embedding, and  $h_i$  represents the node vector updated by node  $i$  after information aggregation. To mitigate overfitting, a fully connected (FC) layer followed by a dropout layer is incorporated into the architecture. The network culminates with an additional FC layer to yield the ultimate output.

The second stage of the framework involves transferring the learned parameters to another identical network, denoted as  $f_i(\theta_i)$ , for initialization through weight sharing. Subsequently,

**Algorithm 1.** PADSM-WSFL Framework.

---

```

1: procedure LEARNINHERENTFEATURES( $D, f_p(\cdot), g_{improved}(\cdot)$ )
2:   Input: Synthesized dataset:  $D = \{\hat{X}_i\}_{i=1}^N$ 
3:   Create dataset for  $P_t$ :  $D_p = \{\hat{X}_i, Y_i\}_{i=1}^N$  via transformation
4:   Build graph:  $G' = g_{improved}(D_p)$  based on the number of  $i$ 
5:   Initialize the parameter  $\theta_p \leftarrow f_p(\theta)$ 
6:   while not converged do
7:     Compute output:  $z'_i = f_p(\hat{X}_i | \theta_p)$ 
8:     Calculate loss:  $L = -\sum_i^T y_i^{(i)} \ln(z_i^{(i)})$  or
        $L = -\frac{1}{I} \sum_i^I \sum_t^T y_i^{(i)} \ln(z_i^{(i)})$ 
9:     Update  $\theta_p$  by back propagation
10:  end while
11:  Output: Optimized parameter  $\theta_p$  for  $P_t$ 
12: end procedure
13: procedure TESTONREALWORLDDATA( $D_r, \theta_p, FC(\cdot)$ )
14:  Input: Real dataset:  $D_r = \{X_i\}_{i=1}^N$ 
15:  Initialize twin network:  $\theta_t \leftarrow \theta_p$  (weight-sharing)
16:  Compute label value:  $\{(P_{i,k})_{k=1}^n\}_{i=1}^N = FC(\{h_i'\}_{i=1}^N)$ 
17:  Make predictions for  $D_r$ :  $P_i = \max(\{P_{i,k}\}_{k=1}^n)$ 
18:  Calculate accuracy:  $\text{acc} = \frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C (\sum_{j=1}^C M_{ij}) + \sum_{i=1}^C TP_i}$ 
19:  return Accuracy and Predictions
20: end procedure

```

---

the performance of this twin neural network is rigorously evaluated using real experimental data, adhering to the fundamental training principles of graph neural networks. During this phase, the graph construction method for real data is dynamically adjusted based on previously determined training strategies, ensuring optimal adaptability to real-world scenarios.

To quantitatively assess the model's performance, accuracy is employed as the primary metric, defined as:

$$\text{Accuracy} = \frac{\sum_{i=1}^C TP_i}{\sum_{i=1}^C \left( \sum_{j=1}^C M_{ij} \right) + \sum_{i=1}^C TP_i} \quad (21)$$

where  $C$  represents the total number of categories,  $TP_i$  denotes the number of samples from the true class  $i$  correctly predicted as class  $i$ , and  $M_{ij}$  signifies the number of samples from the true class  $i$  erroneously predicted as class  $j$ . Thus, accuracy is computed as the ratio of correctly predicted samples to the total number of predicted samples, providing a comprehensive measure of the model's diagnostic capabilities.

#### 4. Experimental results and comparative analysis

This section systematically evaluates the PADSM-WSFL framework through two case studies employing drivetrain prognostics simulator (DPS) and DDS platforms, detailed in sections 4.1 and 4.2. And demonstrated the experimental details and process of synthesizing samples using a physical perception dynamic spectral model, as detailed in section 4.1.2. And display the final experimental results through accuracy, precision, F1 score, and recall. In terms of visualization, T-SNE diagrams, confusion matrices, and Roc

curves were used to express the stability and reliability of the PADSM-WSFL framework.

##### 4.1. Case study I

**4.1.1. Experimental apparatus and data acquisition.** The experimental apparatus employed in this investigation is the DPS, manufactured by SpectraQuest Inc. as illustrated in figure 5. This sophisticated system comprises several interconnected components, which are described in detail as follows: a variable speed drive motor, a planetary gearbox system, a two-stage parallel gearbox system, resistance-load gear boxes coupled to a resistance-load inducing electric load motor, and an electric control unit that governs the entire set-up. The physical parameters of the planetary set are listed in table 1. The primary focus of this research is directed towards the planetary gearbox system, which operates using spur gears, and the two-stage parallel gearbox system, which utilizes helical gears. It is worth noting that the dataset for case study I encompasses fault samples from four distinct categories: (1) Broken, (2) Chipped, (3) Crack, and (4) Normal.

In this experimental protocol, the signal sampled from the planetary gearbox transmission system was carefully selected for analysis. Horizontal position signal data is acquired at a sampling frequency of 24 kHz, under controlled conditions where the input speed was maintained at 20 Hz. These parameters are chosen to ensure consistency and reproducibility of the results. Furthermore, based on the physical parameters of the test rig, a comprehensive model of the gearbox is developed, effectively creating a dynamic model of this system, as shown in figure 2. This model is subsequently utilized to simulate the vibration response of each component under various input speeds and external loads, thereby providing a more holistic understanding of the system's behavior. The simulation signal is sampled at a frequency of 25 kHz, slightly higher than the experimental sampling rate to ensure adequate resolution. To facilitate analysis, the original signal segmented into discrete samples, each with a length of 2048 data points. This segmentation process allows for more manageable data handling and enables the application of various signal processing techniques.

The proposed algorithms are implemented in Python 3.9 and executed on a GeForce GTX 3090 GPU. The framework consists of two stages: first, training on a small set of labeled simulation data to learn vibration signal characteristics, and second, testing on real data. To avoid experimental coincidence, 8 replicate experiments are executed for all methods. The network structure and hyperparameters are set to the same for all GNNs. The network is trained for 200 epochs using an Adam optimizer with a weight decay of 0.0005, an initial learning rate of 0.01 and decays to 0.1 in 100, 150 iterations respectively. The  $k$  value is set to 3. For testing, 120 samples are randomly selected from each of the 3000 samples per category, totaling 480 test samples. This approach aims to effectively diagnose faults with limited training data while ensuring robust performance on real-world samples.

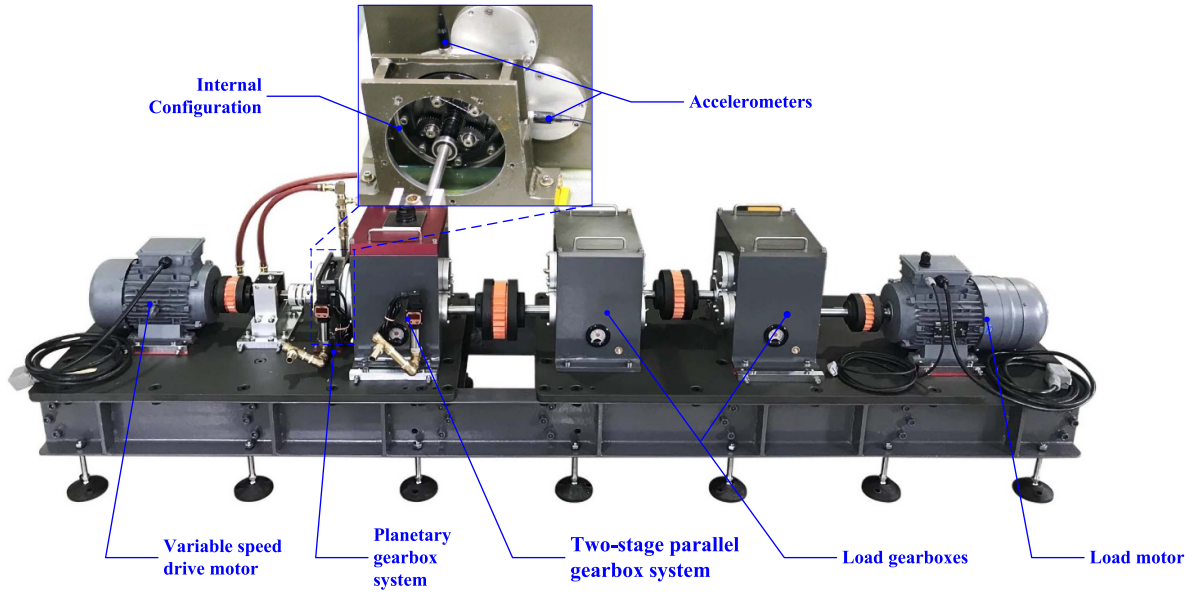


Figure 5. Illustration of the drivetrain prognostics simulation (DPS).

Table 1. Physical parameters of the planetary gear set in DPS.

| Parameters                                                          | Sun                   | Planet (4)            | Ring                  | Carrier            |
|---------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|--------------------|
| Number of teeth                                                     | 28                    | 36                    | 100                   | —                  |
| Module [mm]                                                         | 1                     | 1                     | 1                     | —                  |
| Pressure angle [°]                                                  | 20                    | 20                    | 20                    | —                  |
| Face width [mm]                                                     | 10                    | 10                    | 10                    | —                  |
| Young's modulus [Pa]                                                | $2.1 \times 10^{11}$  | $2.1 \times 10^{11}$  | $2.1 \times 10^{11}$  | —                  |
| Poisson's ratio                                                     | 0.3                   | 0.3                   | 0.3                   | —                  |
| Mass [kg]                                                           | —                     | —                     | $9.86 \times 10^{-2}$ | —                  |
| Moment of inertia [ $\text{kg} \cdot \text{m}^2$ ]                  | $2.41 \times 10^{-6}$ | $1.60 \times 10^{-5}$ | $9.20 \times 10^{-3}$ | $4.99 \times 10^4$ |
| Base circle [mm]                                                    | 13.2                  | 16.9                  | 47.0                  | —                  |
| Torsional stiffness [ $\text{N} \cdot \text{m}^{-1}$ ]              | 0                     | —                     | $1 \times 10^9$       | 0                  |
| Torsional damping [ $\text{N} \cdot \text{s} \cdot \text{m}^{-1}$ ] | 0                     | —                     | $1 \times 10^3$       | 0                  |

**4.1.2. Synthesized sample using physics-aware dynamic spectral model.** When considering the impact of local faults on gear performance, it is important to note that such defects typically result in a decrease in the stiffness of the affected tooth. The nature and extent of this stiffness reduction vary depending on the specific failure mode. For instance, a root crack fundamentally alters the tooth's cross-sectional area, leading to a significant reduction in both bending and shearing stiffness. In the case of a chipped tooth, the change can be modeled as a reduction in the tooth surface area, which primarily affects the Hertzian contact stiffness. Pitting, another common form of gear damage, can be represented in the model by a shortened contact line between mating teeth. The most severe failure mode, a broken tooth, results in a complete loss of stiffness in the affected region.

To improve the accuracy of predicting planetary gearbox behavior, an LSTM network [14] and the generalized Welch method [38] are used to estimate power spectral density (PSD). The LSTM network is initially trained to correct errors in the dynamic model. The simulated vibrations of various components serve as input data, while the measured signals from the

test rig are used as output sample labels. Notably, the training is conducted using the discrete Fourier transform (DFT) sequences of the samples rather than the original time-domain waveforms, which allows for more efficient processing of frequency-domain characteristics. Following the successful training of the LSTM network, it becomes possible to estimate the spectral density of unknown conditions using Welch's method. This approach involves averaging the modified periodograms of windowed overlapping segments. The mathematical expression for this process is provided in the subsequent equations, offering a rigorous framework for spectral analysis of planetary gearbox vibrations

$$\hat{P}_{\text{PER}}^i(e^{j\omega}) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_N^i(n) w(n) e^{-j\omega n} \right|^2 \quad (22)$$

$$\hat{P}_{\text{PER}}^i(e^{j\omega}) = \frac{1}{U} \left| X_N^i(e^{j\omega}) * W(e^{j\omega}) \right|^2 \quad (23)$$

$$\tilde{P}_{\text{PER}}(e^{j\omega}) = \frac{1}{L} \sum_{i=1}^L \hat{P}_{\text{PER}}^i(e^{j\omega}). \quad (24)$$

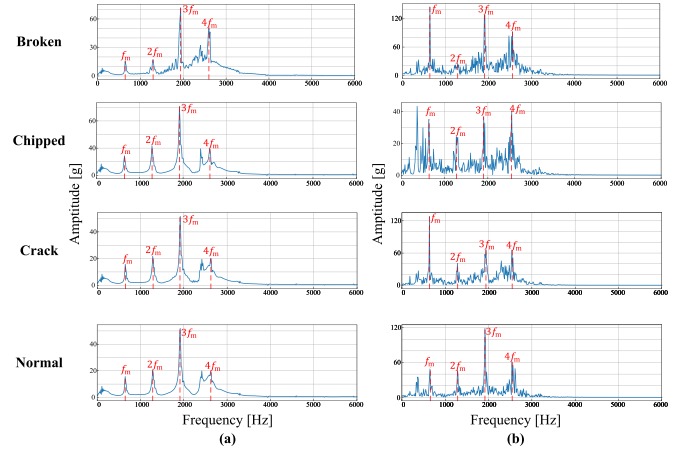
The term  $\hat{P}_{\text{PER}}^i(e^{j\omega})$  represents the discrete time Fourier transform (DTFT) of the windowed  $i$  – th segment of length  $M$ . This transformation is fundamental in converting time-domain data into the frequency domain, thereby enabling detailed spectral analysis. It is worth noting that this DTFT can also be computed through the convolution of the DTFT sequences of the signal with the window function, providing an alternative computational approach that may be advantageous in certain scenarios. The energy of the window function, denoted as  $U = (1/M) \cdot \sum_{n=1}^{M-1} w^2(n)$ , plays a critical role in the analysis process. This parameter is essential for normalizing the spectrum, thereby ensuring an unbiased estimation of the PSD. The normalization process is crucial in maintaining the integrity of the spectral analysis, as it compensates for the energy introduced by the windowing process. A key aspect of this methodology is the averaging of multiple segment estimations, represented by  $\bar{P}_{\text{PER}}$ . This average, taken over  $L$  segments, serves a vital purpose in reducing the variance of the estimation. By combining multiple estimates, the method effectively mitigates the impact of random fluctuations and noise, thereby enhancing the overall reliability and accuracy of the spectral analysis. This improvement in estimation performance is particularly valuable when dealing with complex systems such as planetary gearboxes, where subtle spectral features may hold significant diagnostic information.

The robustness and generalization capability of the trained network are ensured through a comprehensive data acquisition strategy. Both simulated and measured signals are collected under a diverse range of operating conditions and failure modes. This approach is crucial in capturing the full spectrum of potential gear behaviors and fault characteristics. By exposing the network to this wide array of scenarios during training, its ability to generalize to new, unseen conditions is significantly enhanced. In the training process, labels corresponding to the simulated fault types are assigned to the simulated signals. This labeling is a critical step in supervised learning, providing the network with the necessary context to associate specific signal characteristics with particular fault conditions. The labeled, simulated data, represented as  $\{h_i^{(\text{sim}, \text{label})}\}_{i=1}^N$ , serves as the primary input to the network during the training phase. Following the training process, the network's generalization capabilities are put to the test. The trained network parameters are applied to the same network architecture, but this time using real, unlabeled data  $\{h_i^{(\text{real}, \text{unlabel})}\}_{i=1}^N$  as input. This step is crucial in assessing the network's ability to transfer its learned knowledge to real-world scenarios

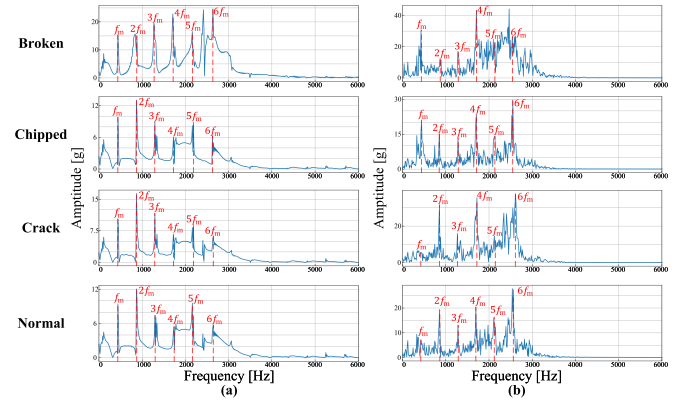
$$\{(P_{i,k})_{k=1}^n\}_{i=1}^N = \text{FC}(\{h_i'\}_{i=1}^N) \quad (25)$$

$$P_i = \max(\{P_{i,k}\}_{k=1}^n) \quad (26)$$

where  $n$  represents the total number of fault types under consideration, while  $k$  denotes a specific fault number within this set. The output  $P_{i,k}$  is a critical parameter, indicating the degree of similarity between the input signal and the  $k$ th fault type. This similarity measure is normalized to fall within the range



**Figure 6.** Signal comparison analysis samples for Case Study I: (a) four types of synthetic vibration signals with their corresponding meshing frequencies, and (b) experimental fault vibration signals with their corresponding meshing frequencies.



**Figure 7.** Signal comparison analysis samples for Case Study II: (a) four types of synthetic vibration signals with their corresponding meshing frequencies, and (b) experimental fault vibration signals with their corresponding meshing frequencies.

of 0 to 1, providing an intuitive scale for interpretation. The interpretation of these similarity values is straightforward yet powerful: a larger value of  $P_{i,k}$  suggests a higher likelihood that the input signal corresponds to the fault category represented by  $k$ . The fault category with the highest similarity value, denoted by  $P_i$ , is then considered the most probable classification for the input signal.

Based on the comprehensive analysis presented above, the signal comparison analysis samples from Case Study I and Case Study II are illustrated in figures 6 and 7, respectively. These figures provide valuable insights into the effectiveness of our proposed methodology. In figure 6, it is evident that the principal frequencies, such as the meshing frequency with harmonics  $f_m$ ,  $2f_m$ ,  $3f_m$ , and  $4f_m$ , can be clearly identified in the synthesized signals generated through the physics-aware spectral model. Notably, these meshing frequencies in the experimental signals closely align with those observed in the synthesized signals. This strong correlation demonstrates

**Table 2.** Details of the used DGAT

| Layer      | Input Channels | Output Channels | Params                       |
|------------|----------------|-----------------|------------------------------|
| GATv2Conv1 | feature        | 1024            | heads = 4                    |
| Linear1    | 1024           | 1024            | —                            |
| BatchNorm1 | 1024           | 1024            | —                            |
| GATv2Conv2 | 1024           | 1024            | heads = 4                    |
| Linear2    | 1024           | 1024            | —                            |
| BatchNorm2 | 1024           | 1024            | —                            |
| FCL1       | 1024           | 512             | Act: ReLU,<br>inplace = True |
| Dropout1   | 512            | 512             | $p = 0.2$                    |
| FCL2       | 512            | out_channel     | —                            |

the high efficacy of the dynamic model in creating synthesized signals that accurately represent real-world conditions. Consequently, this validates the model's suitability for further training the proposed few-shot learning method for planetary gearbox diagnosis.

Similarly, figure 7 exhibits comparable results, wherein the meshing frequency and its harmonics, specifically  $f_m$ ,  $2f_m$ ,  $3f_m$ ,  $4f_m$ ,  $5f_m$ , and  $6f_m$ , can also be distinctly identified in the synthesized signals. This consistency across both case studies further reinforces the robustness and reliability of our proposed approach. These findings collectively underscore the significant potential of the proposed methodology in accurately simulating and analyzing complex planetary gearbox systems, thus paving the way for more efficient and effective diagnostic techniques in industrial applications.

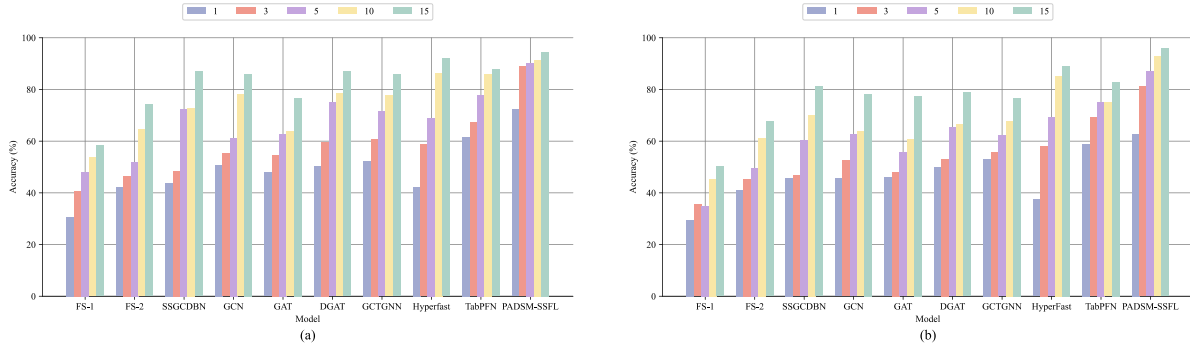
**4.1.3. Planetary gearbox diagnosis results and comparative analysis.** To rigorously evaluate and demonstrate the efficacy and superiority of the proposed PADSM-WSFL framework, a comprehensive validation process is conducted using measured datasets under controlled conditions. For a thorough comparative analysis, several current SOTA methods are employed as benchmarks. These methods include Few-shot learning based on WDCNN [39], referred to as FS-1, Few-shot learning based on IDCNN [40], denoted as FS-2, SSGCDBN [41], GCN [42], GAT [43], GNN method based on the Granger causality test (GCTGNN) [44], TabPFN [45], and HyperFast [46]. These methodologies are selected due to their relevance and established performance in related domains, thereby providing a robust foundation for comparative analysis. Furthermore, it is noteworthy that the proposed DGAT component. TabPFN and HyperFast directly load existing pre trained model weights to generate adapted network structures and parameters without the need for additional training and tuning. The specific architectural details and hyperparameters of the employed DGAT are meticulously documented in table 2, facilitating reproducibility and further analysis by the scientific community.

The diagnostic results obtained from the DPS datasets are comprehensively presented in table 3 and illustrated in figure 8(a). In this study, the proposed method, PADSM-WSFL, is compared with established supervised learning

models, such as FS-1 and FS-2. Upon careful analysis, it becomes evident that PADSM-WSFL demonstrates a marked improvement in performance compared to these above models. Specifically, when the models are trained using a single sample for each faulty class, PADSM-WSFL achieves a noteworthy accuracy of 72.23%. This performance significantly surpasses that of FS-1 and FS-2 by 30.75% and 42.13%, respectively. Furthermore, as the number of training samples increases to 15, PADSM-WSFL exhibits a remarkable diagnostic accuracy of 94.58%, whereas the supervised methods struggle to exceed 75% accuracy. A comprehensive comparison results in other metrics is also recorded in table 4. The values highlighted in bold represent outstanding performance indicators obtained through different methods under experimental conditions.

In addition to supervised learning models, this study also incorporates a comparative analysis with semi-supervised learning methods, including SSGCDBN, GCN, GAT, DGAT and GCTGNN. These models are characterized by their ability to be trained with a limited amount of labeled data supplemented by a substantial volume of unlabeled data, thereby enabling them to extract inherent features from the unlabeled dataset. Leveraging initialization weights derived from auxiliary tasks, these semi-supervised models demonstrate superior performance compared to traditional supervised networks when confronted with limited training samples. Nevertheless, the proposed PADSM-WSFL method consistently outperforms these aforementioned models. Notably, when the number of training samples per class is increased to three, PADSM-WSFL continues to exhibit superior performance, surpassing SSGCDBN, GCN, GAT, DGAT and GCTGNN by 48.48%, 55.21%, 54.64%, 59.63% and 60.83% in terms of accuracy, respectively. These experimental results provide compelling evidence that the implementation of PADSM-WSFL networks can substantially enhance the accuracy of fault diagnosis in scenarios where only a limited number of training samples are available. When utilizing 10 training samples, PADSM-WSFL demonstrates a remarkable improvement, outperforming DGAT and GCTGNN by 12.8% and 13.62%, respectively. This performance disparity reaches its zenith at 15 training samples, where PADSM-WSFL achieves an impressive accuracy of 94.58%, while GCN and DGAT attain 86.04% and 86.93%, respectively.

A comprehensive comparative analysis is conducted to validate the effectiveness of the proposed approach against well-established few-shot learning methodologies, particularly TabPFN and HyperFast, whereby the performance metrics of all approaches are meticulously documented in table 3 and systematically visualized in figure 8(a). Through rigorous empirical evaluation, it has been demonstrated that the proposed PADSM-WSFL method not only matches but consistently and significantly outperforms these SOTA benchmark models across diverse testing scenarios. Of particular significance is the observation that when the training dataset is expanded to incorporate three samples per class, PADSM-WSFL maintained its superior diagnostic capabilities, exceeding both TabPFN and HyperFast by substantial margins of 67.22% and 58.75% in classification accuracy, respectively.



**Figure 8.** Accuracy comparison: (a) Case study I, (b) Case study II. The height of each bar represents the number of training samples (1–15) available for each distinct fault category, differentiated by color.

**Table 3.** The comparison results in accuracy (%) on case study I.

| Method     | 1                 | 3                 | 5                 | 10                | 15                |
|------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| FS-1       | 30.75±0.70        | 40.79±0.64        | 48.09±4.13        | 53.70±0.52        | 58.42±0.64        |
| FS-2       | 42.13±3.09        | 46.47±3.02        | 51.87±4.31        | 64.47±4.32        | 74.27±8.42        |
| SSGCDBN    | 43.74±3.85        | 48.48±4.16        | 72.37±4.33        | 72.91±1.89        | 86.91±4.56        |
| GCN        | 50.76±7.12        | 55.21±4.44        | 61.25±0.31        | 78.12±0.48        | 86.04±2.20        |
| GAT        | 47.98±0.64        | 54.64±3.65        | 62.77±2.10        | 63.75±2.50        | 76.64±4.00        |
| DGAT       | 50.41±5.53        | 59.63±3.33        | 74.94±1.86        | 78.52±1.42        | 86.93±4.25        |
| GCTGNN     | 52.29±3.61        | 60.83±2.19        | 71.48±6.14        | 77.70±1.16        | 85.83±4.66        |
| HyperFast  | 42.29±5.35        | 58.75±5.70        | 68.96±5.77        | 86.32±2.78        | 92.02±0.37        |
| TabPFN     | 61.67±2.73        | 67.22±4.82        | 77.92±5.16        | 86.04±1.82        | 87.64±1.21        |
| PADSM-WSFL | <b>72.23±4.52</b> | <b>88.82±4.89</b> | <b>90.16±3.12</b> | <b>91.32±4.67</b> | <b>94.58±1.25</b> |

**Table 4.** The comparison results in other metrics (%) on case study I.

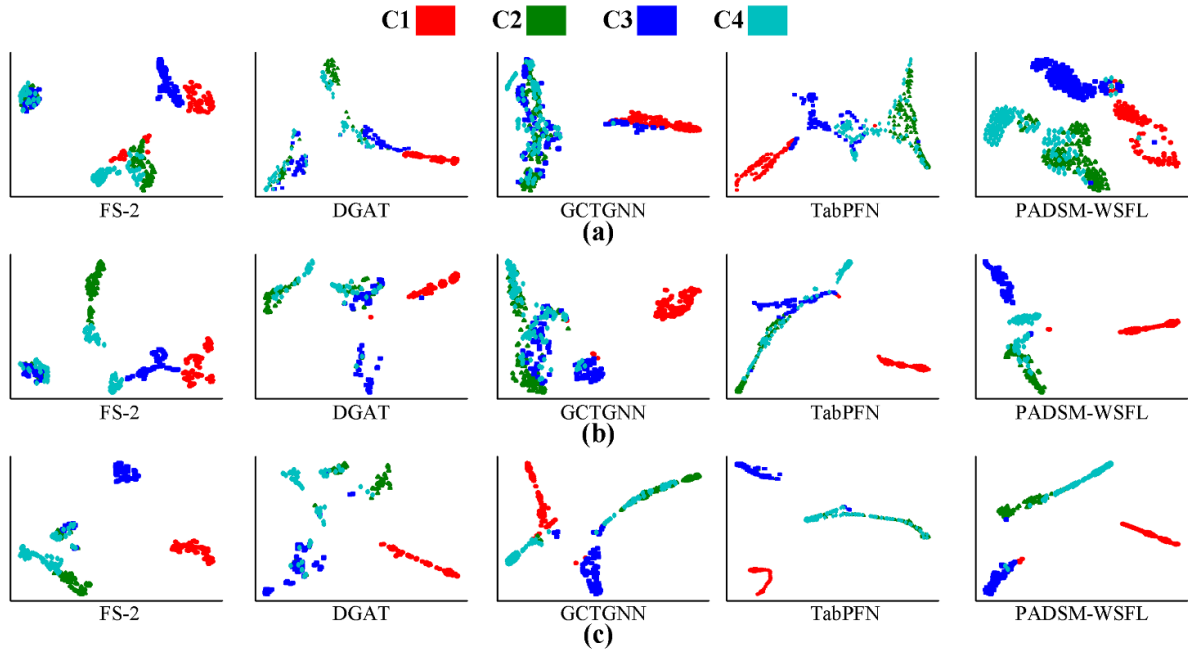
| Method     | Precision         | F1 score          | Recall            |
|------------|-------------------|-------------------|-------------------|
| FS-1       | 37.81±6.07        | 37.81±6.07        | 38.02±6.16        |
| FS-2       | 54.21±1.02        | 54.60±1.51        | 54.47±1.12        |
| SSGCDBN    | 64.29±9.25        | 56.78±14.11       | 55.42±6.19        |
| GCN        | 67.51±6.16        | 64.47±6.85        | 65.21±6.54        |
| GAT        | 73.69±10.44       | 66.95±11.28       | 67.99±11.95       |
| DGAT       | 78.85±3.23        | 77.56±2.04        | 77.71±2.98        |
| GCTGNN     | 80.10±3.82        | 79.64±3.68        | 79.93±3.44        |
| HyperFast  | 76.63±6.33        | 74.07±5.92        | 74.03±6.78        |
| TabPFN     | 79.75±3.32        | 73.91±8.67        | 74.44±8.04        |
| PADSM-WSFL | <b>91.57±1.01</b> | <b>90.84±1.36</b> | <b>90.91±1.41</b> |

Moreover, these experimental findings provide compelling evidence that the implementation of PADSM-WSFL networks can fundamentally enhance the robustness and reliability of fault diagnosis systems, especially in resource-constrained scenarios where only a limited number of training samples are available. Subsequently, upon expanding the training dataset to 15 samples, PADSM-WSFL demonstrated even more remarkable performance improvements, achieving accuracy rates that significantly surpassed both TabPFN (92.02%) and HyperFast (87.64%), thereby further validating its enhanced learning capabilities and superior diagnostic precision in few-shot learning contexts.

It is particularly noteworthy that, in contrast to the aforementioned semi-supervised learning models, the proposed PADSM-WSFL method achieves these results using only a small subset of labeled data, rather than relying on a

combination of limited labeled data and extensive unlabeled data, as is typical in semi-supervised learning approaches. The experimental outcomes thus provide strong evidence for the efficacy of weakly supervised learning and the enhanced distance metric function incorporated within the proposed framework.

In order to gain a more comprehensive understanding of the performance of various methodologies, T-SNE plots are generated and are depicted in figure 9. The results demonstrate that PADSM-WSFL exhibits superior capability in distinguishing between different types of fault signals. Specifically, when utilizing only one sample per class for model training, networks based on supervised and semi-supervised learning techniques demonstrate a marked inability to differentiate between various fault signals, only managing to distinguish fault type C1 (broken) from other types. Although DGAT



**Figure 9.** The T-SNE visualization of feature distributions for various methods in case study I. (a)–(c) Represent models trained with 1, 5 and 15 samples per class, respectively. Fault categories C1 through C4 correspond to broken, chipped, crack, and normal conditions.

shows marginally improved performance and can separate a greater number of fault types in the feature space, its capabilities remain limited. In contrast, PADSM-WSFL demonstrates a remarkable ability to differentiate almost all fault types, with only minor confusion occurring between fault types C2 (chipped) and C4 (normal). As the number of training samples increases to 15 per class, PADSM-WSFL achieves perfect separation of different fault signal types, as illustrated in figure 9(c). However, other models continue to struggle with the classification of certain fault types, particularly C2 (chipped), C3 (crack), and C4 (normal). These visualization results provide strong evidence that the PADSM-WSFL framework can effectively extract vibration signal features in few-shot scenarios and achieve superior separation of signal features for different fault types within the feature space.

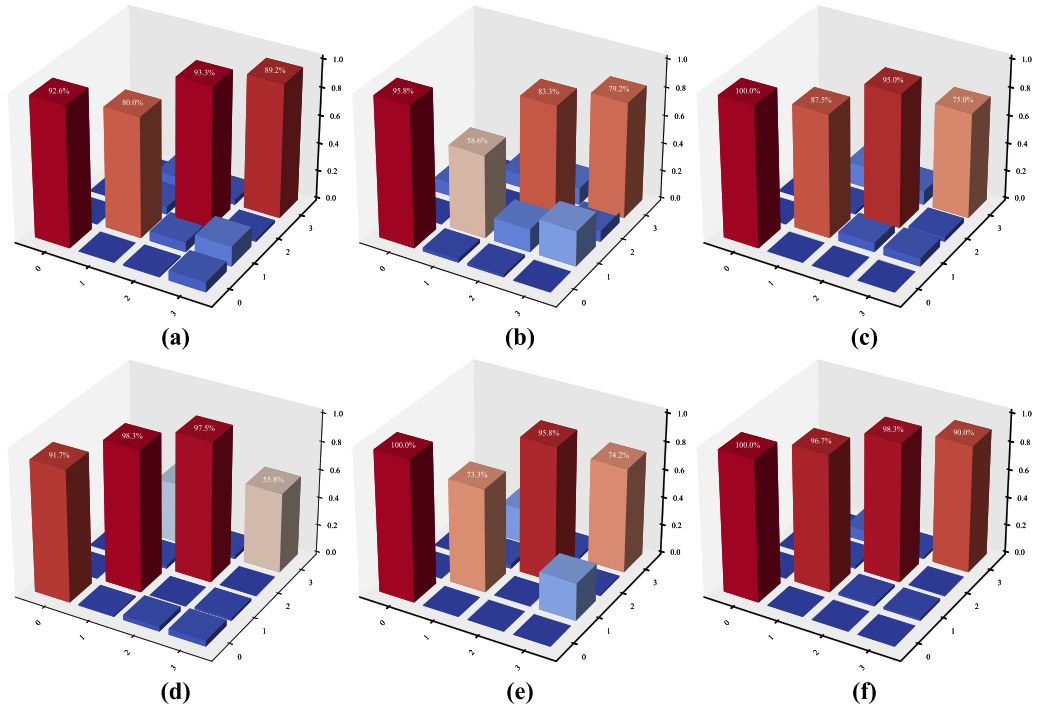
To conduct a more rigorous quantitative analysis of the diagnostic results, confusion matrices were employed in this case study, with the results depicted in figure 10. The data reveals that when using 15 samples per class for training, the proposed PADSM-WSFL method successfully recognizes all fault types with high accuracy. Furthermore, semi-supervised models including SSGCDBN, GAT, DGAT, and GCTGNN face significant challenges in differentiating between broken and chipped failures. The few-shot learning approaches, particularly methods such as TabPFN, encounter significant challenges when dealing with feature overlapping among crack, chipped, and normal classifications, especially when limited to only 15 training samples per class. This limitation in classification performance is clearly illustrated in figure 10, where the confusion matrix metrics reveal substantial misclassification patterns between these overlapping categories. Furthermore, the results underscore the inherent difficulties in distinguishing subtle variations between damage types when operating

under few-shot learning constraints. The proposed PADSM-WSFL method, however, demonstrates remarkable performance by accurately recognizing all four types of failures, achieving an impressive average accuracy of 95%. These findings underscore the efficacy of the PADSM-WSFL framework in fault diagnosis tasks, particularly in scenarios with limited training data.

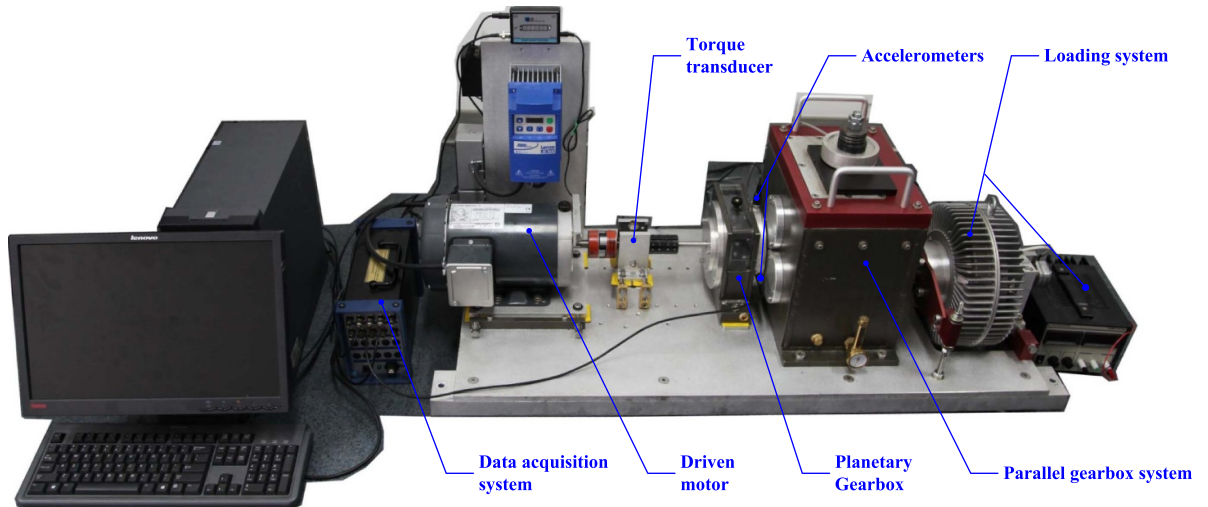
#### 4.2. Case study II

To further validate the effectiveness of the proposed method PADSM-WSFL, a test apparatus known as the drivetrain diagnostics simulation (DDS) is employed in this case study. The DDS, as illustrated in figure 11, comprises a complex array of interconnected components, each serving a specific function in the experimental process. These components include, but are not limited to, a high-precision data acquisition system, a driven motor for power input, a torque transducer for measuring rotational force, a planetary gearbox, multiple accelerometers for vibration measurement, a parallel gearbox system, and a loading system to simulate various operational conditions. For the purposes of this investigation, the operational parameters are carefully controlled and standardized. Specifically, the input speed was set at 20 Hz, while the output load was maintained at 0.35 A.

As discussed in case study I, this section further explores the effectiveness of the proposed PADSM-WSFL method in comparison to various established techniques, including FS-1, FS-2, SSGCDBN, GCN, GAT, DGAT, GCTGNN, HyperFast and TabPFN. The diagnostic results derived from the DDS datasets are detailed in table 5 and visually illustrated in figure 8(b). The values highlighted in bold denote the superior performance metrics attained through diverse methodological



**Figure 10.** Comparative analysis of confusion matrices for selected methods in case study I (15 samples per class): (a) SSGCDBN, (b) GAT, (c) DGAT, (d) GCTGNN, (e) TabPFN, (f) PADSM-WSFL. Fault categories 0~3 correspond to broken, chipped, crack, and normal conditions, respectively.



**Figure 11.** Diagram of the drivetrain diagnostics simulation (DDS).

approaches under experimental conditions. This study compares the proposed PADSM-WSFL method with well-known supervised learning models, such as FS-1 and FS-2. It is evident that PADSM-WSFL significantly enhances performance relative to these models. Specifically, when trained with a single sample for each faulty class, PADSM-WSFL attains an accuracy of 62.71% , substantially outperforming FS-1 and FS-2 by 29.58% and 41.07%, respectively. As the number of training samples increases to 15, PADSM-WSFL achieves a notable diagnostic accuracy of 95.83%, while the supervised methods struggle to exceed 68% accuracy. A comprehensive

comparison results in other metrics is also recorded in table 6.

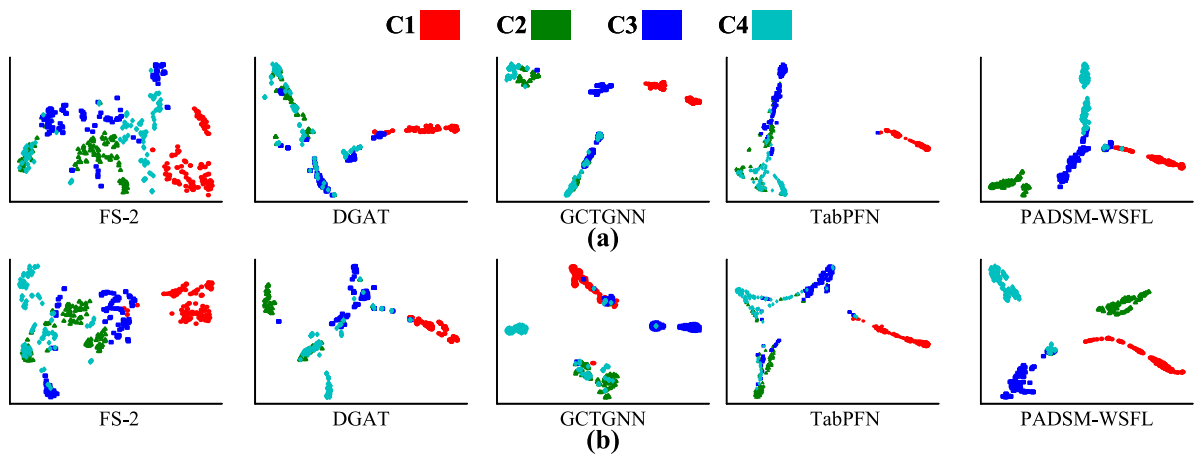
In addition, this study includes a comparative analysis with semi-supervised learning methods, such as SSGCDBN, GCN, GAT, DGAT and GCTGNN. The proposed PADSM-WSFL method consistently outperforms these models as well. Notably, when the number of training samples per class is increased to three, PADSM-WSFL continues to show superior performance, surpassing SSGCDBN, GCN, GAT, DGAT and GCTGNN by 46.80%, 52.57%, 48.05%, 53.05% and 55.76% in terms of accuracy, respectively. These experimental results

**Table 5.** The comparison results in accuracy (%) on case study II.

| Method     | 1                 | 3                 | 5                 | 10                | 15                |
|------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| FS-1       | 29.58±0.95        | 35.49±3.41        | 34.93±1.93        | 45.38±7.98        | 50.14±5.71        |
| FS-2       | 41.07±1.62        | 45.33±4.79        | 49.53±8.27        | 61.20±6.43        | 67.67±2.99        |
| SSGCDBN    | 45.83±4.35        | 46.80±1.84        | 60.27±3.67        | 70.07±2.90        | 81.11±3.79        |
| GCN        | 45.48±8.55        | 52.57±2.93        | 62.50±4.16        | 63.82±4.10        | 78.33±0.34        |
| GAT        | 45.97±8.44        | 48.05±5.24        | 55.62±9.92        | 60.62±2.96        | 77.49±4.90        |
| DGAT       | 49.86±9.88        | 53.05±4.64        | 65.35±6.26        | 66.39±2.66        | 78.75±2.64        |
| GCTGNN     | 53.12±4.48        | 55.76±6.34        | 62.22±4.66        | 67.84±2.05        | 76.70±5.01        |
| HyperFast  | 37.43±3.54        | 58.07±4.20        | 69.11±3.34        | 85.19±4.24        | 89.04±2.04        |
| TabPFN     | 58.90±4.00        | 69.26±3.06        | 75.18±1.94        | 74.46±4.06        | 82.91±2.13        |
| PADSM-WSFL | <b>62.71±2.98</b> | <b>81.11±2.56</b> | <b>86.94±1.47</b> | <b>92.71±1.46</b> | <b>95.83±0.68</b> |

**Table 6.** The comparison results in other metrics (%) on case study II.

| Method     | Precision         | F1 score          | Recall            |
|------------|-------------------|-------------------|-------------------|
| FS-1       | 29.43±4.17        | 29.29±4.12        | 29.38±4.27        |
| FS-2       | 60.56±8.75        | 60.59±8.62        | 61.09±8.44        |
| SSGCDBN    | 69.29±4.19        | 63.28±5.24        | 61.25±3.25        |
| GCN        | 72.78±3.90        | 71.05±4.77        | 71.88±4.16        |
| GAT        | 60.02±5.77        | 55.02±6.28        | 54.58±6.01        |
| DGAT       | 69.80±2.21        | 68.65±2.56        | 68.68±3.37        |
| GCTGNN     | 74.20±2.23        | 74.33±7.15        | 78.82±7.13        |
| HyperFast  | 74.82±11.04       | 70.37±12.75       | 70.21±12.82       |
| TabPFN     | 74.73±2.14        | 70.31±0.41        | 70.83±0.36        |
| PADSM-WSFL | <b>86.79±2.90</b> | <b>84.34±2.01</b> | <b>84.44±2.72</b> |

**Figure 12.** The T-SNE visualization of feature distributions for various methods in case study II. (a), (b) Represent models trained with 10 and 15 samples per class, respectively. Fault categories C1 through C4 correspond to broken, chipped, crack, and normal conditions.

provide compelling evidence that implementing PADSM-WSFL networks can significantly enhance the accuracy of fault diagnosis in scenarios where only limited training samples are available. With 10 training samples, PADSM-WSFL shows a remarkable improvement, outperforming DGAT and GCTGNN by 26.32% and 24.87%, respectively. This performance gap peaks at 15 training samples, where PADSM-WSFL achieves an impressive accuracy of 95.83%, while GCN and DGAT attain 78.33% and 78.75%, respectively. This significant advantage can also be illustrated in figure 12.

In the comprehensive comparative analysis against well-established few-shot learning methodologies, particularly TabPFN and HyperFast, the performance metrics and empirical outcomes of the respective approaches have been meticulously documented in table 5 and systematically visualized in figure 8(b). Through rigorous experimental validation, the empirical evidence conclusively demonstrates that the proposed PADSM-WSFL method not only consistently but also significantly outperforms these SOTA benchmark models across a diverse spectrum of testing scenarios. Of particular significance is the observation that when the training dataset is

incrementally expanded to incorporate three samples per class, PADSM-WSFL continues to exhibit exceptional diagnostic capabilities, thereby surpassing TabPFN and HyperFast by substantial margins of 69.26% and 58.07% in classification accuracy, respectively. Furthermore, these comprehensive experimental findings provide compelling and substantiated evidence that the strategic implementation of PADSM-WSFL networks can fundamentally enhance both the robustness and reliability of fault diagnosis systems, especially in challenging scenarios where only a limited number of training samples are accessible. Moreover, upon systematically expanding the training dataset to 15 samples, PADSM-WSFL exhibits even more pronounced and statistically significant performance enhancements, surpassing the baseline accuracies of TabPFN (82.91%) and HyperFast (89.04%) by considerable margins, thereby reinforcing its superior diagnostic capabilities in scenarios with moderately sized training sets. These results thus definitively underscore its superior learning efficiency and diagnostic precision, while simultaneously validating its effectiveness in real-world applications.

The results of the receiver operating characteristic (ROC) curves are presented in figure 13. These curves, which plot the true positive rate on the vertical axis against the false positive rate on the horizontal axis, illustrate the performance of classification models at various threshold settings. It is important to note that this evaluation metric provides valuable insights into model performance, with curves closer to 1 indicating superior fault diagnosis capabilities.

Figure 13(i) demonstrates that the ROC curves for the four distinct failure types—namely, broken, chipped, crack, and normal conditions—approach the optimal value of 1. This trend is particularly evident in the zoomed region, which offers a more detailed visualization of the results. In contrast, the supervised model FS-2, as depicted in figure 13(a), exhibits varying performance across different failure types. Specifically, the area under the curve (AUC) values for broken, chipped, crack, and normal states are 81%, 60%, 47%, and 30%, respectively. Interestingly, the semi-supervised models, such as GAT and DGAT, demonstrate enhanced performance compared to their supervised counterparts. As illustrated in figure 13(d), GAT achieves high accuracy in identifying failures, with AUC values of 100%, 83%, 93%, and 87% for the aforementioned failure types, respectively. Furthermore, DGAT, shown in figure 13(e), exhibits even more impressive results, with AUC values of 100%, 90%, 94%, and 92% for the same failure categories. Furthermore, the performance characteristics of two prominent few-shot learning methodologies, namely HyperFast and TabPFN, are comprehensively depicted in figures 13(g) and (h), respectively. Upon careful analysis of the experimental results, it becomes evident

that the HyperFast algorithm demonstrates remarkable discriminative capabilities, achieving AUC values of 100% for broken states, while maintaining robust performance with 95% for chipped states, 93% for crack states, and 95% for normal states. In parallel, the TabPFN methodology exhibits comparable yet slightly lower performance metrics, attaining an optimal AUC value of 100% for broken states, while yielding values of 91%, 92%, and 92% for chipped, crack, and normal states, respectively. These results underscore the efficacy of both approaches in accurately classifying various structural conditions, albeit with subtle variations in their discriminative capabilities. These comprehensive results provide strong evidence that the proposed PADSM-WSFL method can maintain a high fault detection rate while simultaneously reducing the false-positive rate across various fault states. This achievement underscores the potential of semi-supervised learning approaches in enhancing the accuracy and reliability of fault diagnosis systems.

#### 4.3. Discussions on the impact of noise

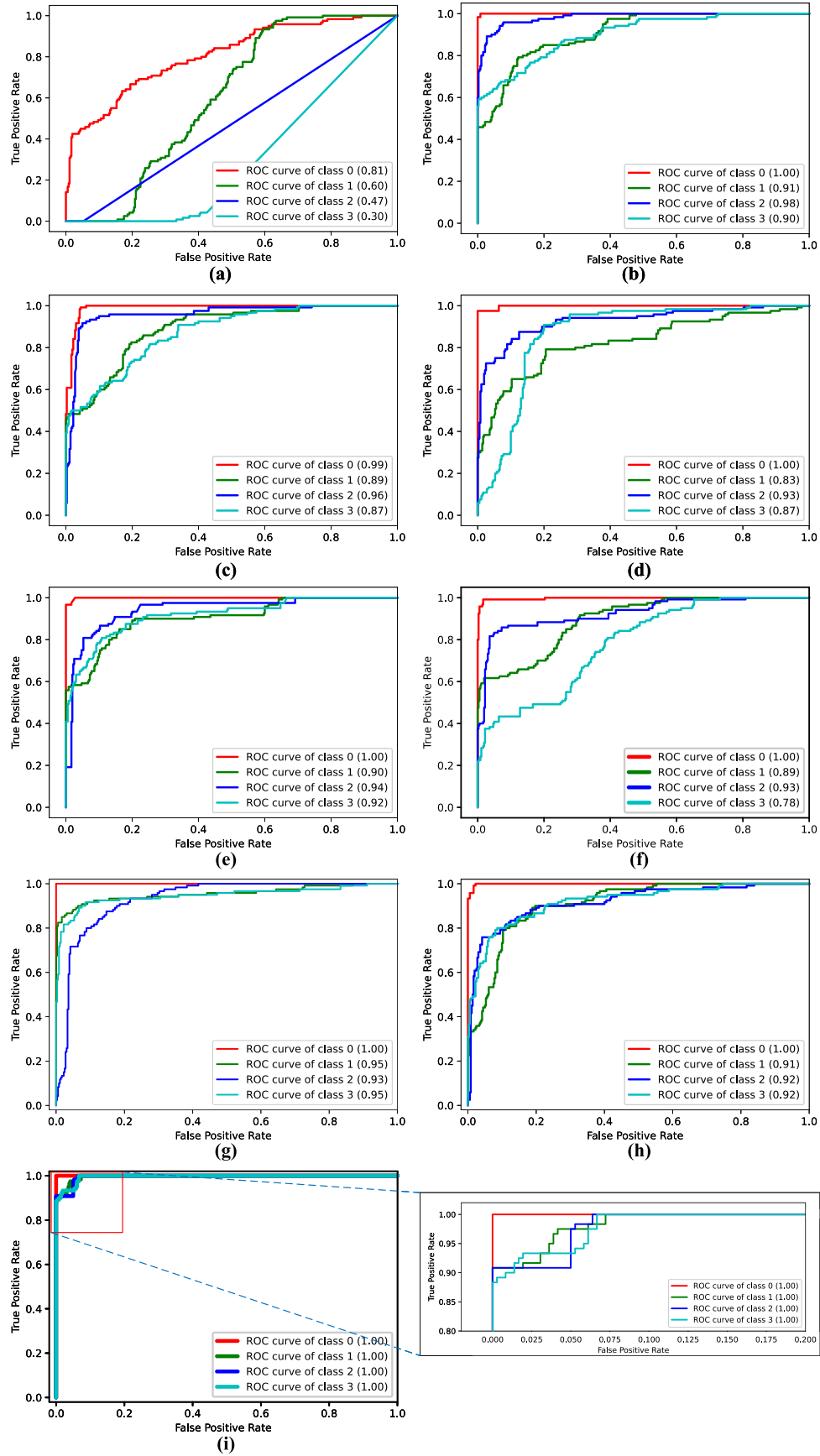
To verify the performance of processing noisy signals, different signal-to-noise ratios (SNRs) of Gaussian noise, including 0 dB, 5 dB and 10 dB, are added to the data in the case II dataset, according to the  $\text{SNR(dB)} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right)$ . The  $P_{\text{noise}}$  is the noise power,  $P_{\text{signal}}$  is the signal power.

The experimental results are given in table 7. It can be seen that the performance of the proposed method is still reliable when the SNR is greater than 5 dB, but the performance decreases when the SNR is 0 dB. This degradation can be attributed to the fact that the graph structure relies on the similarity between samples; noise can diminish this similarity, even among samples with same labels. Consequently, it disrupts the connections between similar samples, thereby deteriorating the overall quality of the graph.

#### 4.4. Discussions on $k$ values

The impact of neighborhood parameter selection on graph-based spectral analysis warrants systematic examination. Figure 14 shows diagnostic accuracies of the proposed method at different  $k$  values. As shown, different  $k$  values have certain influences on the diagnostic effects. For Case I,  $k = 2$  has the highest accuracy 92.71%; as for Case II, the highest accuracy 90.83% is obtained at  $k = 3$ .

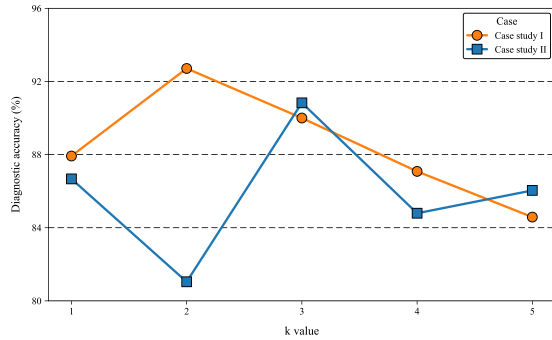
The diagnostic accuracy shows a fluctuating downward trend with the increase of  $k$  value, mainly because under the condition of few-shot samples, increasing  $k$  value will increase the probability of constructing edge connections between samples with different labels.



**Figure 13.** The ROC curves for comparative methods using 15 samples per class in case study II: (a) FS-2, (b) SSGCDBN, (c) GCN, (d) GAT, (e) DGAT, (f) GCTGNN, (g) HyperFast, (h) TabPFN, (i) PADSM-WSFL. Fault categories 0~3 correspond to broken, chipped, crack, and normal conditions, respectively.

**Table 7.** Experimental results under different SNR levels.

| Evaluation Metric | SNR (dB)         |                  |                  |                  |
|-------------------|------------------|------------------|------------------|------------------|
|                   | 10               | 5                | 0                | No noise         |
| Accuracy (%)      | 90.56 $\pm$ 1.22 | 90.21 $\pm$ 0.75 | 87.98 $\pm$ 0.79 | 92.71 $\pm$ 1.46 |
| F1 score (%)      | 90.31 $\pm$ 1.39 | 90.05 $\pm$ 0.88 | 87.42 $\pm$ 1.40 | 92.61 $\pm$ 0.14 |
| Precision (%)     | 91.02 $\pm$ 1.31 | 92.22 $\pm$ 1.05 | 90.32 $\pm$ 1.88 | 93.88 $\pm$ 0.55 |

**Figure 14.** Accuracies of the proposed method at different  $k$  values.

## 5. Conclusion

The proposed PADSM-WSFL framework effectively addresses the critical challenge of fault diagnosis in planetary gearboxes under conditions of limited labeled data availability. By integrating physics-aware modeling with weakly supervised few-shot learning, this innovative approach efficiently leverages both extremely limited labeled data and substantial unlabeled data resources. Consequently, the framework demonstrates superior robustness and generalization capabilities when compared to existing SOTA methods. This research offers a practical and adaptable solution for real-world industrial scenarios where the scarcity of labeled data has traditionally posed a substantial obstacle.

The framework shows promising potential for future developments and applications across multiple dimensions. First, the methodology can be extended to diverse mechanical systems beyond planetary gearboxes, including wind turbines, aircraft engines, and industrial robots. Second, incorporating transfer learning and online learning mechanisms will enable real-time adaptation and cross-domain knowledge transfer, enhancing the system's versatility. Third, the integration of interpretable AI components will provide clear explanations of diagnostic decisions, while the development of prognostic capabilities will enable remaining useful life prediction and early-stage fault detection. Finally, the framework can be applied in smart manufacturing, predictive maintenance systems, and Industry 4.0 initiatives where labeled data scarcity remains a significant challenge. However, several limitations of the current framework should be noted. The physics-aware

modeling component introduces additional computational overhead, which may affect real-time processing capabilities in some applications. The framework's performance heavily depends on the quality and accuracy of the underlying physics models.

Future work will focus on extending this approach to other mechanical systems, integrating transfer learning techniques to improve performance across various domains, and incorporating online learning capabilities to adapt in real time. Additionally, exploring interpretability methods will enhance the model's decision-making explainability, and conducting long-term studies will evaluate its effectiveness in predicting early-stage faults and estimating the remaining useful life of components.

## Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grants 52105111, and 52305085), the Guangdong Basic and Applied Basic Research Foundation (Grant 2025A1515012256), the Shantou University (STU) Scientific Research Initiation Grant (NTF21029), the Industry-Academia Cooperation Project from the Guangdong Institute of Special Equipment Inspection and Research Shunde Branch (XTJ-KY01-202503-030), the Natural Science Foundation of Sichuan Province (Grant 2023NSFSC0861), the China Postdoctoral Science Foundation (Grant 2023M740021), and the Natural Science Foundation of Anhui Province (Grant 2108085QE229).

## Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## ORCID iDs

Peng Chen  0000-0002-3265-3079  
 Jia Gao  0009-0001-9543-1051  
 Yaqiang Jin  0000-0001-9072-7948  
 Junyu Qi  0000-0003-1669-8987  
 Changbo He  0000-0003-4180-5334

## References

- [1] Goswami P and Rai R N 2025 Data-driven sensor selection for industrial gearbox fault diagnosis using principal component analysis *Meas. Sci. Technol.* **36** 036111
- [2] Shao H, Ming Y, Liu Y and Liu B 2024 Small sample gearbox fault diagnosis based on improved deep forest in noisy environments *Nondestruct. Test. Eval.* **40** 1–22
- [3] Chen P, Wu Y, Fan S, He C, Jin Y, Qi J and Zhou C 2025 Adaptive signal regime for identifying transient shifts: a novel approach toward fault diagnosis in wind turbine systems *Ocean Eng.* **325** 120798
- [4] Leaman F 2024 Phenomenological study of acoustic emissions generated by gear meshing in planetary gearboxes *Meas. Sci. Technol.* **36** 015136
- [5] Chen P, Wu Y, Xu C, Huang C-G, Zhang M and Yuan J 2025 Interference suppression of nonstationary signals for bearing diagnosis under transient noise measurements *IEEE Trans. Reliab.* **1–15**
- [6] Chen P, Li Y, Wang K and Zuo M J 2020 A novel knowledge transfer network with fluctuating operational condition adaptation for bearing fault pattern recognition *Measurement* **158** 107739
- [7] Chen P, Ma J, He C, Jin Y and Fan S 2025 Semi-supervised consistency models for automated defect detection in carbon fiber composite structures with limited data *Meas. Sci. Technol.* **36** 046109
- [8] Chen P, Li Y, Wang K, Zuo M J, Heyns P S and Baggeröhr S 2021 A threshold self-setting condition monitoring scheme for wind turbine generator bearings based on deep convolutional generative adversarial networks *Measurement* **167** 108234
- [9] Zhang C, Hu D and Yang T 2024 Research of artificial intelligence operations for wind turbines considering anomaly detection, root cause analysis and incremental training *Reliab. Eng. Syst. Saf.* **241** 109634
- [10] Chen P, Ma J, He C, Jin Y and Fan S 2025 Progressive contrastive representation learning for defect diagnosis in aluminum disk substrates with a bio-inspired vision sensor *Expert Syst. Appl.* **289** 128305
- [11] Tao W, Sun Z, Yang Z, Liang B, Wang G and Xiao S 2025 Transformer fault diagnosis technology based on AdaBoost enhanced transferred convolutional neural network *Expert Syst. Appl.* **264** 125972
- [12] Chen P, Zhang R, Fan S, Guo J and Yang X 2024 Step-wise contrastive representation learning for diagnosing unknown defective categories in planetary gearboxes *Knowl.-Based Syst.* **309** 112863
- [13] Wang C, Li H, Zhang K, Hu S and Sun B 2021 Intelligent fault diagnosis of planetary gearbox based on adaptive normalized CNN under complex variable working conditions and data imbalance *Measurement* **180** 109565
- [14] Shi J, Peng D, Peng Z, Zhang Z, Goebel K and Wu D 2022 Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks *Mech. Syst. Signal Process.* **162** 107996
- [15] Chen Y, Rao M, Feng K and Zuo M J 2022 Physics-informed LSTM hyperparameters selection for gearbox fault detection *Mech. Syst. Signal Process.* **171** 108907
- [16] Snyder Q, Jiang Q and Tripp E 2025 Integrating self-attention mechanisms in deep learning: a novel dual-head ensemble transformer with its application to bearing fault diagnosis *Signal Process.* **227** 109683
- [17] Chen P, Ma Z, Xu C, Jin Y and Zhou C 2024 Self-supervised transfer learning for remote wear evaluation in machine tool elements with imaging transmission attenuation *IEEE Internet Things J.* **11** 23045–54
- [18] Luo J, Huang J and Li H 2021 A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis *J. Intell. Manuf.* **32** 407–25
- [19] Ha J M and Fink O 2023 Domain knowledge-informed synthetic fault sample generation with health data map for cross-domain planetary gearbox fault diagnosis *Mech. Syst. Signal Process.* **202** 110680
- [20] Cheng W, Wang S, Liu Y, Chen X, Nie Z, Xing J, Zhang R and Huang Q 2023 A novel planetary gearbox fault diagnosis method for nuclear circulating water pump with class imbalance and data distribution shift *IEEE Trans. Instrum. Meas.* **72** 1–13
- [21] Jin Z, Xu Q, Jiang C, Wang X and Chen H 2023 Ordinal few-shot learning with applications to fault diagnosis of offshore wind turbines *Renew. Energy* **206** 1158–69
- [22] Yue K, Li J, Chen J, Huang R and Li W 2022 Multiscale wavelet prototypical network for cross-component few-shot intelligent fault diagnosis *IEEE Trans. Instrum. Meas.* **72** 1–11
- [23] Liu S, Chen J, He S, Shi Z and Zhou Z 2023 Few-shot learning under domain shift: attentional contrastive calibrated transformer of time series for fault diagnosis under sharp speed variation *Mech. Syst. Signal Process.* **189** 110071
- [24] Lao Z, He D, Jin Z, Liu C, Shang H and He Y 2023 Few-shot fault diagnosis of turnout switch machine based on semi-supervised weighted prototypical network *Knowl.-Based Syst.* **274** 110634
- [25] Su H, Yao Q, Xiang L and Hu A 2024 Semi-supervised temporal meta learning framework for wind turbine bearing fault diagnosis under limited annotation data *IEEE Trans. Instrum. Meas.* **73** 1–9
- [26] Ruan H, Wang Y, Li X, Qin Y and Tang B 2022 An enhanced non-local weakly supervised fault diagnosis method for rotating machinery *Measurement* **189** 110433
- [27] Guo L, Peng Y, Qin R and Liu B 2022 A new weakly supervised learning approach for real-time iron ore feed load estimation *Expert Syst. Appl.* **202** 117469
- [28] Yan S, Zhong X, Shao H, Ming Y, Liu C and Liu B 2023 Digital twin-assisted imbalanced fault diagnosis framework using subdomain adaptive mechanism and margin-aware regularization *Reliab. Eng. Syst. Saf.* **239** 109522
- [29] Qian M and Li Y-F 2022 A weakly supervised learning-based oversampling framework for class-imbalanced fault diagnosis *IEEE Trans. Reliab.* **71** 429–42
- [30] Liu S, Xu X, Zhang Y, Muhammad K and Fu W 2023 A reliable sample selection strategy for weakly supervised visual tracking *IEEE Trans. Reliab.* **72** 15–26
- [31] Jiao C, Zhang D, Fang X and Miao Q 2023 Ensemble of simplified graph wavelet neural networks for planetary gearbox fault diagnosis *IEEE Trans. Instrum. Meas.* **72** 1–10
- [32] Xu Y, Ji J, Ni Q, Feng K, Beer M and Chen H 2023 A graph-guided collaborative convolutional neural network for fault diagnosis of electromechanical systems *Mech. Syst. Signal Process.* **200** 110609
- [33] Chen H, Wang X-B and Yang Z-X 2023 Semi-supervised self-correcting graph neural network for intelligent fault diagnosis of rotating machinery *IEEE Trans. Instrum. Meas.* **72** 1–11
- [34] Yu X, Tang B and Deng L 2023 Fault diagnosis of rotating machinery based on graph weighted reinforcement

- networks under small samples and strong noise *Mech. Syst. Signal Process.* **186** 109848
- [35] Yin P, Nie J, Liang X, Yu S, Wang C, Nie W and Ding X 2023 A multi-scale graph convolutional neural network framework for fault diagnosis of rolling bearing *IEEE Trans. Instrum. Meas.* **72** 1–13
- [36] Yan S, Shao H, Xiao Y, Zhou J, Xu Y and Wan J 2022 Semi-supervised fault diagnosis of machinery using LPS-DGAT under speed fluctuation and extremely low labeled rates *Advanced Engineering Informatics* **53** 101648
- [37] Uddin S, Haque I, Lu H, Moni M A and Gide E 2022 Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction *Sci. Rep.* **12** 6256
- [38] Welch P 1967 The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms *IEEE Trans. Audio Electroacoust.* **15** 70–73
- [39] Li F, Wang L, Wang D, Wu J and Zhao H 2023 Transfer multiscale adaptive convolutional neural network for few-shot and cross-domain bearing fault diagnosis *Meas. Sci. Technol.* **34** 125002
- [40] Qiao L, Zhang Y and Wang Q 2023 Fault detection in wind turbine generators using a meta-learning-based convolutional neural network *Mech. Syst. Signal Process.* **200** 110528
- [41] Zhao X, Jia M and Liu Z 2020 Semisupervised graph convolution deep belief network for fault diagnosis of electromechanical system with limited labeled data *IEEE Trans. Ind. Inf.* **17** 5450–60
- [42] Yu X, Tang B and Zhang K 2021 Fault diagnosis of wind turbine gearbox using a novel method of fast deep graph convolutional networks *IEEE Trans. Instrum. Meas.* **70** 1–14
- [43] Cao S, Li H, Zhang K, Yang C, Xiang W and Sun F 2023 A novel spiking graph attention network for intelligent fault diagnosis of planetary gearboxes *IEEE Sens. J.* **23** 13140–54
- [44] Zhang Z and Wu L 2024 Graph neural network-based bearing fault diagnosis using granger causality test *Expert Syst. Appl.* **242** 122827
- [45] Hollmann N, Müller S, Eggensperger K, Hutter F 2022 TabPFN: a transformer that solves small tabular classification problems in a second (arXiv:2207.01848)
- [46] Bonet D, Montserrat D M, Giró-i-Nieto X and Ioannidis A 2024 HyperFast: instant classification for tabular data *Proc. AAAI Conf. on Artificial Intelligence* vol 38 pp 11114–23