

# Physics-informed auto-encoder based on digital twin for rolling bearing fault diagnosis under imbalanced sample conditions

Zhiwu Shang<sup>a,b,\*</sup>, Ziyu Wang<sup>a,b</sup>, Cailu Pan<sup>a,b</sup>, Wanxiang Li<sup>c,d</sup>, Maosheng Gao<sup>e</sup>

<sup>a</sup> School of Mechanical Engineering, Tiangong University, Tianjin, 300387, China

<sup>b</sup> Tianjin Modern Electromechanical Equipment Technology Key Laboratory, Tianjin, 300387, China

<sup>c</sup> TBEA Beijing Tianjin Hebei Intelligent Technology Co., Ltd, Tianjin, 301701, China

<sup>d</sup> TBEA Tianjin Transformer Co., Ltd, Tianjin, 300300, China

<sup>e</sup> TBEA Electrical Equipment Group Co., Ltd, Tianjin, China

## ARTICLE INFO

### Keywords:

Data augmentation  
Digital twin  
Dynamic model  
Fault diagnosis  
Imbalanced sample

## ABSTRACT

In the fault diagnosis of rolling bearings, the scarcity of fault samples results in far fewer fault data than healthy ones, leading to sample imbalance that negatively affects diagnostic accuracy. Generative neural networks can produce additional fault data, but they cannot fully ensure that the generated data capture bearing dynamics, thereby undermining reliability. This paper proposes a Physics-Informed Auto-Encoder based on Digital Twin (PIAE-DT) for interpretable data augmentation under imbalanced conditions. A dimensionless physical information module incorporating bearing dynamic characteristics was embedded into the Auto-Encoder, enabling identification of dynamic model parameters and improving interpretability by ensuring that each learned feature corresponds to a physically meaningful parameter. The identified parameters were mapped to a probability distribution, from which a fault data augmentation method in the digital twin space was developed through resampling to increase the diversity of augmented data. Furthermore, reconstructing vibration signals from these parameters enhanced the reliability of the augmented data, and the results verified the effectiveness of the physics-informed module. Finally, by applying PIAE-DT to identify dynamic parameters of known samples, fault data were augmented in the digital twin space and used to alleviate imbalance in diagnosis. The method was evaluated on both public and private datasets, achieving maximum diagnostic accuracies of 97.60 % and 99.86 %, and it outperforms other data augmentation methods.

## 1. Introduction

Rolling bearings are indispensable transmission and support components in high-precision machinery such as aircraft engines, high-speed trains, and precision machine tools, where their operational condition exerts a decisive influence on system performance, stability, and safety (Lei et al., 2020; Sun et al., 2025b). Under harsh service conditions characterized by dynamic load fluctuations, bearings are susceptible to wear, fatigue cracks, and other defects, which can lead to performance degradation, safety incidents, and substantial economic losses (Zhao et al., 2025). Such consequences not only jeopardize the stable and reliable operation of the system but also impose significant challenges in terms of maintenance costs, fault prediction, and safety assurance. Consequently, the development of advanced bearing health monitoring and fault diagnosis has emerged as a prominent research focus in both

academia and industry (Xin et al., 2025).

In industrial practice, the deployment of fault diagnosis systems faces significant challenges, as fault samples are often scarce while normal samples are abundant, leading to imbalanced datasets that hinder model learning and generalization (Lu et al., 2024; Zhang et al., 2022a). Oversampling is a widely adopted strategy to address this imbalance in rolling bearing fault diagnosis by generating additional minority class samples without discarding existing data. However, traditional techniques such as the Synthetic Minority Oversampling Technique (SMOTE) may cause overfitting and are highly sensitive to noise in the original dataset, thereby reducing the generalization capability of diagnostic models (Yi et al., 2022; Zhu et al., 2017). To overcome these limitations, Han et al. proposed the Attention Feature Selection Oversampling (AFS-O) method, which incorporates feature selection and weighting prior to oversampling to generate high-quality synthetic data

\* Corresponding author. School of Mechanical Engineering, Tiangong University, Tianjin, 300387, China.

E-mail address: [shangzhiwu@tiangong.edu.cn](mailto:shangzhiwu@tiangong.edu.cn) (Z. Shang).

<https://doi.org/10.1016/j.engappai.2025.113017>

Received 3 June 2025; Received in revised form 28 August 2025; Accepted 28 October 2025

Available online 3 November 2025

0952-1976/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

and improve classification performance (Han et al., 2024). Similarly, Li et al. developed the Adaptive Clustering Weighted Oversampling (ACWOS) method, which adaptively determines clustering structures and assigns oversampling weights based on sample density and relative distances, thereby enhancing the diversity of minority samples under imbalanced and variable-speed conditions (Li et al., 2024). Both methods aim to mitigate the noise sensitivity and overfitting issues of conventional oversampling while enhancing model robustness.

While the improved oversampling techniques discussed above can effectively mitigate the adverse effects of data imbalance, their performance remains largely dependent on the quality of manually engineered features (Zhang et al., 2025). In recent years, deep learning-based intelligent diagnostic approaches have gained considerable attention due to their capability to automatically learn hierarchical feature representations directly from raw vibration signals, thereby reducing the reliance on handcrafted features (Ni et al., 2025; Zhang et al., 2020). Among them, architectures such as convolutional neural networks (CNNs) and deep neural networks (DNNs) have been extensively applied in rolling bearing fault diagnosis for their strong adaptive feature extraction capabilities (Sun et al., 2025a; Yan et al., 2025). Li et al. proposed a dual-stage attention-based recurrent neural network (DA-RNN) combined with a convolutional block attention module-embedded convolutional neural network (CBAM-CNN) to expand imbalanced datasets and improve diagnostic accuracy (Li et al., 2022). Yang et al. introduced an external attention-integrated CNN (EA-CNN) combined with continuous wavelet transform (CWT) to reduce model complexity and enhance performance under imbalanced data conditions (Yang et al., 2025). Chen et al. incorporated clustering- and density-based resampling with multi-domain feature extraction (time, frequency, and entropy) and attention-based temporal convolution, together with Gradient-weighted Class Activation Mapping (Grad-CAM) visualization, to further improve diagnostic accuracy (Chen et al., 2025). Collectively, these studies highlight the importance of addressing data imbalance in deep learning-based fault diagnosis to enhance both accuracy and robustness.

Although deep learning-based classification models have made remarkable progress in addressing data imbalance, their performance is still constrained by the limited availability of minority class samples. To overcome this limitation, generative models have emerged as a promising alternative, as they can synthesize realistic and diverse fault samples to directly balance the dataset (Liu et al., 2023). With the rapid development of deep learning, generative models such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) have become important tools for fault diagnosis under imbalanced data conditions due to their strong capability to generate diverse and representative samples (Zhou et al., 2020). Hou et al. proposed an Enhanced Generative Adversarial Network with a Data Selection Module (EGAN-DSM), which integrates a network enhancement module, Hilbert space distance-based data screening, and a wide first-layer kernel CNN to improve diagnostic accuracy under imbalanced conditions (Hou et al., 2023). Liu et al. developed an interpretable adversarial variational autoencoder with sequential attention (AVAE-SQA), combining reliable control limits, adversarial training, and attention mechanisms to enhance both diagnostic accuracy and interpretability (Liu et al., 2023). Wang et al. incorporated operating-condition information and cross-frequency attention into an enhanced GAN to generate longer vibration sequences under variable conditions, thereby improving robustness in imbalanced scenarios (Wang et al., 2025). Yin et al. proposed a variational auxiliary-classifier Wasserstein GAN (VACWGAN) that integrates a variational autoencoder with an ACGAN-style classifier, enriching scarce multi-fault samples, stabilizing training, and achieving higher diagnostic accuracy under severe imbalance (Yin et al., 2025). Concurrently, these generative-model-based approaches produce high-quality minority class samples, effectively mitigating the performance degradation of conventional methods under data scarcity and varying operating conditions, and providing new solutions for

imbalanced fault diagnosis.

In contrast to purely data-driven generative models, which can enrich minority class samples but often lack physical interpretability and consistency with real-world system dynamics, Digital Twin (DT) technology offers a physically consistent approach to data augmentation (Cui et al., 2024). DT technology, by contrast, enables physically consistent data augmentation through the integration of physics-based simulation and real sensor data, offering a promising pathway to address these limitations (Yi et al., 2023; Zhang et al., 2023). In rolling bearing fault diagnosis, DT has shown considerable potential (Qin et al., 2024a; Zhong et al., 2025). However, many existing approaches still focus on unidirectional mapping from virtual systems to physical systems, neglecting reverse calibration with real sensor data to capture dynamic variations. Moreover, empirical parameter settings and deterministic optimization limit simulation–real alignment and data diversity. To address these challenges, recent research has explored combining DT technology with machine learning, emphasizing the integration of physical knowledge of bearing dynamics and failure mechanisms into data-driven models, or the fusion of prior knowledge with learned representations. Such physics-informed and explainable fault diagnosis approaches have been shown to improve robustness, reliability, and generalization in complex industrial scenarios (Ali et al., 2024a, 2024b, 2025a, 2025b). For instance, Qin et al. proposed a multi-degree-of-freedom DT integrated with a frequency-domain Bidirectional Long Short-Term Memory (Bi-LSTM) Cycle-Consistent Generative Adversarial Network (CycleGAN) and spectrum-constraint loss to improve mapping performance, together with an inverse physics-informed neural network embedding system dynamics for parameter estimation and data generation under imbalanced conditions (Qin et al., 2024b). Shang et al. developed a finite-element-based DT coupled with a multi-scale attention CycleGAN and perceptual loss to achieve domain alignment and enhance the realism of generated data (Shang et al., 2025). Zhu et al. introduced a data–model fusion approach that integrates DT with a conditional deep convolutional GAN for physically consistent data synthesis (Zhu et al., 2025). In conjunction, these studies highlight that bidirectional virtual–real mapping, parameter calibration, and probabilistic modeling are crucial for improving DT-based fault diagnosis performance under imbalanced sample conditions.

Overall, current data augmentation strategies for bearing fault diagnosis often face the challenge of balancing diversity and interpretability. Methods focusing on diversity frequently produce data with limited or only partial interpretability, making it difficult to trace their origins or underlying mechanisms. Conversely, approaches that prioritize interpretability often result in insufficient diversity, thereby constraining the generalization capability of diagnostic models. Therefore, achieving a well-calibrated trade-off between diversity and interpretability under imbalanced sample conditions remains a critical challenge for developing reliable and robust bearing fault diagnosis systems.

To address this challenge, we propose a Physics-Informed Auto-Encoder based on Digital Twin (PIAE-DT) framework that integrates digital twin technology with deep learning, aiming to jointly improve data diversity and physical interpretability. A dimensionless physical information module embedding bearing dynamic characteristics is incorporated into an Auto-Encoder, enabling accurate estimation of dynamic model parameters. These parameters are mapped into a probabilistic distribution space, from which resampling techniques generate high-quality synthetic fault data within the digital twin domain. By reconstructing vibration signals from physically meaningful dynamic parameters, the proposed approach ensures physical consistency and reliability of the generated data while enriching their variability.

The main contributions of this study are as follows.

- (1) Physics-Informed Auto-Encoder based on Digital Twin Architecture – Proposes a new PIAE-DT framework that integrates digital

twin technology with physics-informed auto-encoders to address data imbalance in bearing fault diagnosis.

- (2) Data Generation with Diversity and Interpretability – Develops a mechanism within PIAE-DT to generate fault data with both high diversity and strong physical interpretability, expanding datasets and improving model training.
- (3) Physically Consistent Parameter Modeling – Establishes a method for deriving the dimensionless parameter distribution of target bearings, ensuring synthesized fault data remain consistent with underlying physical mechanisms.
- (4) Comprehensive Performance Evaluation – Conducts extensive experiments under various data imbalance scenarios, demonstrating that PIAE-DT-generated data significantly improve diagnostic accuracy and robustness.

The remainder of this paper is organized as follows. Section 2 reviews related work in three key areas: Auto-Encoders, Physics-Informed Neural Networks, and Digital Twin technology. Section 3 presents the proposed method in detail. Specifically, it describes the construction of the bearing digital twin model and its corresponding digital twin space, the design and principles of PIAE-DT, the loss function, the fault data generation process, the training strategy, and the overall diagnostic framework under imbalanced sample conditions. Section 4 reports the experimental validation, including dataset description, experimental setup, unbalanced dataset construction, analysis of the digital twin space, data quality assessment, and application to fault diagnosis. Finally, Section 5 concludes the paper and outlines potential directions for future research.

## 2. Related work

### 2.1. Auto-encoder

An autoencoder is a well-known unsupervised learning technique designed to reconstruct high-dimensional input data with minimal reconstruction error (Zhang et al., 2025). Fundamentally, an autoencoder comprises two core components: an encoder and a decoder. The encoder compresses the high-dimensional input data into a low-dimensional latent space, generating a compressed representation termed the latent feature vector (or hidden layer features). The encoding process can be mathematically formulated as (Li et al., 2021):

$$\mathbf{h} = \text{Encoder}(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad (1)$$

where  $\mathbf{x} \in \mathbb{R}^n$  denotes the input sample,  $\mathbf{W} \in \mathbb{R}^{m \times n}$  is the weight matrix,  $\mathbf{b} \in \mathbb{R}^m$  is the bias vector, and  $\mathbf{h} \in \mathbb{R}^m$  represents the latent feature vector.

The decoder aims to reconstruct the original high-dimensional data from the latent features. The reconstruction process is defined as:

$$\hat{\mathbf{x}} = \text{Decoder}(\mathbf{W}'\mathbf{h} + \mathbf{b}') \quad (2)$$

where  $\hat{\mathbf{x}} \in \mathbb{R}^n$  is the reconstructed output,  $\mathbf{W}' \in \mathbb{R}^{n \times m}$  and  $\mathbf{b}' \in \mathbb{R}^n$  are the decoder's weight matrix and bias vector, respectively.

The autoencoder is trained to minimize the discrepancy between the input  $\mathbf{x}$  and the reconstructed output  $\hat{\mathbf{x}}$  by optimizing the parameters  $\{\mathbf{W}, \mathbf{b}, \mathbf{W}', \mathbf{b}'\}$  through backpropagation. A commonly used reconstruction loss is the mean squared error (MSE), which quantifies the average deviation between the input and reconstructed data. The MSE loss is defined as:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{m} \sum_{i=1}^m (\hat{x}_i - x_i)^2 \quad (3)$$

where  $\mathcal{L}_{\text{MSE}}$  denotes the loss value, and  $m$  is the total number of training samples.

### 2.2. Physics-Informed Neural Networks

Physics-Informed Neural Networks (PINNs) can integrate empirical data with governing equations, such as partial differential equations (PDEs), by embedding physical laws into the neural network architecture (Wang et al., 2024). By incorporating physical principles as soft constraints within the optimization process, PINNs not only enhance their interpretability and generalization capabilities but also ensure that the solutions adhere closely to the underlying physics, making them more reliable in practical applications. The total loss function of a PINN is typically formulated as a weighted combination of multiple objectives (Cuomo et al., 2022):

$$\mathcal{L}_{\text{PINNs}} = \lambda_{\text{data}} \mathcal{L}_{\text{data}} + \lambda_{\text{BC}} \mathcal{L}_{\text{BC}} + \lambda_{\text{PDE}} \mathcal{L}_{\text{PDE}} \quad (4)$$

where  $\mathcal{L}_{\text{data}}$  quantifies the discrepancy between predicted and observed data,  $\mathcal{L}_{\text{BC}}$  enforces compliance with initial/boundary conditions, and  $\mathcal{L}_{\text{PDE}}$  penalizes deviations from the governing PDE. The weighting coefficients  $\lambda_{\text{data}}, \lambda_{\text{BC}}, \lambda_{\text{PDE}}$  balance these terms, which can be either manually tuned or automatically learned during training.

The parameters of the PINN are optimized through gradient-based methods such as Adam, where the gradients of the composite loss function are computed via automatic differentiation. While traditional PINN frameworks excel at solving isolated partial differential equations, they often fall short when dealing with complex systems involving multiple interacting components. This limitation highlights the necessity for extending these frameworks to encompass broader physical phenomena, thereby enhancing their applicability across various domains (Zhang et al., 2023; Zhong et al., 2025). In this work, we extend the physics-informed paradigm by integrating the physical characteristics of bearings into the neural network through physics-based constraints.

### 2.3. Digital twin

The digital twin framework primarily consists of physical space, virtual space, and the data flow or mapping relationship between them (Cui et al., 2024; Qin et al., 2024a, 2024b). Physical space encompasses the actual system or component under investigation, including all its physical attributes and operational conditions. Virtual space comprises sophisticated analytical models, such as lumped parameter dynamic models (LPDM), which simulate the behavior of the physical system under various conditions and provide insights into potential faults or performance issues (Hao et al., 2024; Yi et al., 2023). Data flow or mapping serves as the critical link between physical space and virtual space, ensuring that real-time data from the physical system can be accurately reflected in the virtual models, thereby enabling continuous monitoring and predictive maintenance capabilities.

To tackle the challenge of sample imbalance in bearing fault diagnosis, we leverage digital twin technology to create synthetic fault samples. This approach not only mitigates the scarcity of labeled fault data but also enhances the robustness and reliability of diagnostic algorithms by providing a more comprehensive dataset for training and validation.

## 3. Proposed method

### 3.1. Digital Twin Model of Bearing

Establishing a dynamic model of bearing failure is crucial for developing accurate bearing digital twins. This model allows for the simulation of various failure modes under realistic operating conditions, thereby enhancing predictive maintenance capabilities. By simplifying the bearing failure to a single degree of freedom nonlinear model, we minimize model complexity while maintaining adaptability to a wide range of bearing datasets.

As illustrated in Fig. 1, the bearing digital twin model developed in this study acquires defect displacement data across different categories by varying the size of faults, which is then embedded into the dynamic model.

As indicated on the left side of Fig. 1, the bearing model consists of an outer ring that does not rotate and an inner ring that rotates. The bearing has one degree of freedom, which is the vertical displacement of the outer ring.

The rotation speed of the cage is:

$$\omega_c = \frac{\omega_s}{2} \left( 1 - \frac{d \cos \alpha}{D} \right) \quad (5)$$

$$|\text{mod}(\theta_j, 2\pi) - \theta_D| < \beta_o, |\text{mod}(\theta_i, 2\pi) - \text{mod}(\theta_j, 2\pi)| < \beta_i, \begin{cases} |\text{mod}(\varphi_j, 2\pi) - \frac{\pi}{2}| < \beta_r \\ |\text{mod}(\varphi_j, 2\pi) - \frac{3\pi}{2}| < \beta_r \end{cases} \quad (9)$$

where  $\omega_s$  is the shaft speed,  $d$  is the rolling element diameter,  $\alpha$  is the contact angle,  $D$  is the pitch circle diameter.

The angular position of the  $j$ th rolling element is:

$$\theta_j = \omega_c t + \frac{2\pi(j-1)}{n_b} \quad (6)$$

where  $n_b$  is the number of rolling elements.

The rotational angular velocity of the rolling element is:

$$\varphi_j = \frac{\omega_s D}{2d} \left( 1 - \left( \frac{d}{D} \cos \alpha \right)^2 \right) t \quad (7)$$

As illustrated on the right side of Fig. 1, this dynamic model simulates three failure states: outer ring fault, inner ring fault, and roller fault. The three forms of local faults are all represented as rectangular peeling.

The circumferential span angles are defined respectively for the outer ring fault, inner ring fault, and rolling element fault:

$$\beta_o = \arcsin \frac{L_o}{D_o}, \beta_i = \arcsin \frac{L_i}{D_i}, \beta_r = \arcsin \frac{L_r}{d} \quad (8)$$

where  $L_o$  is the size of the outer ring defect,  $D_o$  is the diameter of the outer ring,  $L_i$  is the size of the inner ring defect,  $D_i$  is the inner ring diameter,  $L_r$  is the size of the rolling element defect,  $d$  is the diameter of the rolling element.

The position angles at which the  $j$ th rolling element enters the outer ring defect, inner ring defect, and rolling element defect are defined respectively:

where  $\theta_D$  is the position angle of the outer ring defect,  $\theta_i$  is the position of the inner ring defect angle,

The increased radial displacements caused by the outer ring defect, inner ring defect, and rolling element defect are defined respectively:

$$\delta_o = \frac{d}{2} - \sqrt{\left( \frac{d}{2} \right)^2 - \left( \frac{L_o}{2} \right)^2}, \delta_i = \frac{d}{2} - \sqrt{\left( \frac{d}{2} \right)^2 - \left( \frac{L_i}{2} \right)^2}, \delta_r = \frac{d}{2} - \sqrt{\left( \frac{d}{2} \right)^2 - \left( \frac{L_r}{2} \right)^2} \quad (10)$$

in summary, the contact deformation between the  $j$ th rolling element and the raceway can be described as:

$$\delta = \gamma + \lambda_o \delta_o + \lambda_i \delta_i + \lambda_r \delta_r \quad (11)$$

where  $\gamma$  is the radial clearance,  $\lambda_o$  indicates whether the outer ring is faulty,  $\lambda_i$  indicates whether the inner ring is faulty,  $\lambda_r$  indicates whether the rolling elements are faulty. The values of  $\lambda_o$ ,  $\lambda_i$ ,  $\lambda_r$  are 0 or 1, depending on the type of fault.

The stiffness in Hertz contact force depends on:

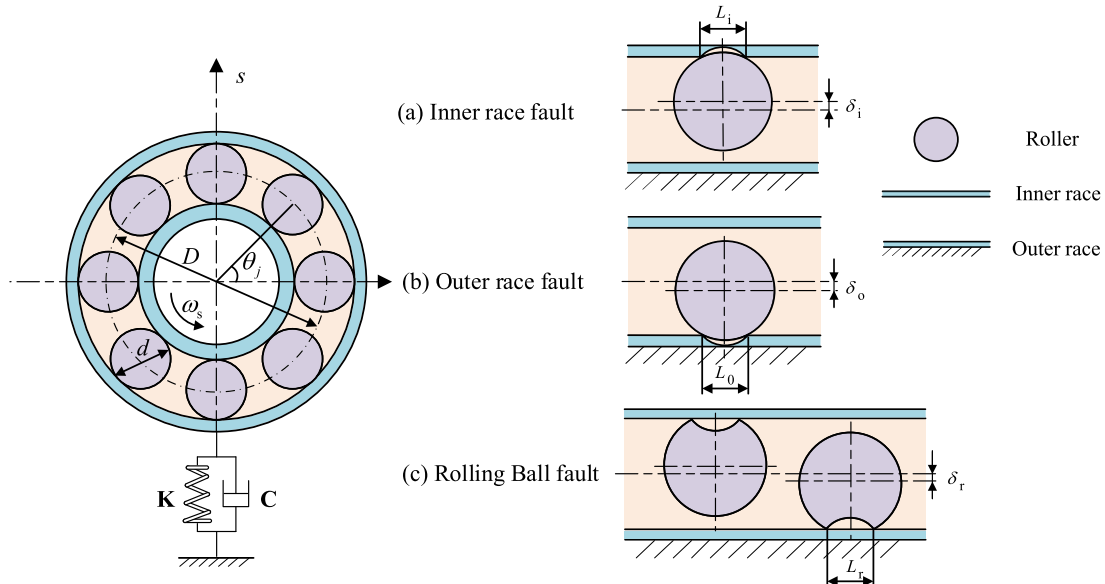


Fig. 1. Digital twin model of bearing.



$$K_b = \frac{2\sqrt{2}E}{3(1-\nu^2)\sqrt{\delta^{*3}}\sum\rho} \quad (12)$$

where  $E$  is the elastic modulus,  $\nu$  Poisson's ratio,  $\delta^*$  represents Hertz contact deformation coefficient (depending on curvature difference),  $\sum\rho$  is the curvature.

The differential equation of dynamics is expressed as (Ashtekar et al., 2010):

$$\mathbf{M}\ddot{s} + \mathbf{C}\dot{s} + \mathbf{K}s + K_b(\max(s - \delta, 0))^{1.5} = W \quad (13)$$

where  $\mathbf{M}$  is the mass matrix,  $\mathbf{C}$  is the damping matrix,  $\mathbf{K}$  is the stiffness matrix,  $K_b$  is the Hertz contact stiffness,  $\ddot{s}$  is the acceleration,  $\dot{s}$  is the velocity,  $s$  is the displacement,  $\delta$  is the defect,  $W$  is the external load.

Thus, this article has established a complete single degree of freedom digital twin bearing dynamic model.

### 3.2. The digital twin space of bearings

The stiffness and damping coefficient are two critical parameters in the dynamic modeling of bearing failures, as they directly govern the accuracy of simulated vibration responses. However, these parameters are often empirically determined, leading to variations in stiffness and damping coefficient across different vibration data samples, which may fluctuate within constrained ranges. While global optimization methods can identify ideal parameter values in a macroscopic sense, ensuring parameter-sample specificity remains challenging, thereby limiting the model's accuracy and generalizability.

To address these limitations, this study proposes a novel methodology that integrates probabilistic modeling with digital twin technology. First, we posit that the stiffness and damping coefficient are not fixed constants but instead follow distinct probability distributions across samples. Specifically, we assume:

$$(k_i, c_i) = \underset{k, c}{\operatorname{argmax}} p(k, c | x_i) \quad (14)$$

where  $p(k, c | x_i)$  denotes the posterior distribution of the dynamic parameters conditioned on the observed signal.

This paradigm shifts the optimization objective from deterministic parameter estimation to probabilistic distribution learning, thereby establishing a digital twin space capable of statistically representing the intrinsic physical property distributions.

Second, the dynamic parameters of known samples are inferred and mapped to a parameterized probability distribution.

$$\left\{ (k_j, c_j) \right\}_{j=1}^M \sim \hat{p}_{k, c}(k, c) \quad (15)$$

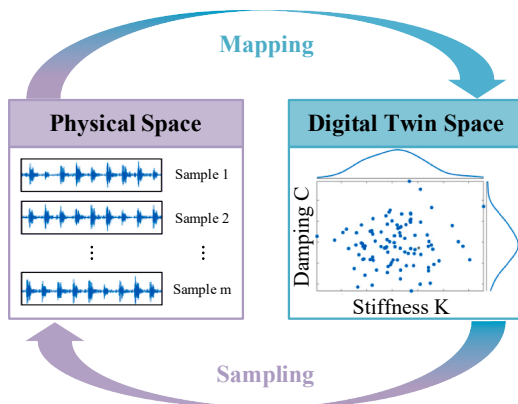


Fig. 2. Double mapping in digital twin space.

where  $\hat{p}_{k, c}(k, c)$  representing possible values of the dynamic model parameters for use in sampling-based inference.

Finally, new bearing fault data is generated by sampling from the derived distribution, enabling the creation of physically consistent fault scenarios.

Fig. 2 illustrates the bidirectional mapping mechanism of the bearing digital twin space, which consists of two key processes: (1) a physical-to-virtual mapping that extracts parameter distributions from real-world vibration signals through statistical learning, and (2) a virtual-to-physical mapping that generates realistic vibration signals from sampled parameter distributions.

### 3.3. Physics-informed auto-encoder based on digital twin

#### 3.3.1. Principle of physics-informed auto-encoder based on digital twin

This study redefines the optimization objective: rather than identifying fixed optimal dynamic parameters for all samples, we aim to learn the optimal dynamic parameter distributions through the bearing digital twin space. This ensures that the generated data maintains physical consistency with real-world bearing behaviors.

To achieve this, a neural network architecture is required to compress raw vibration data into low-dimensional representations within the digital twin space while enabling high-fidelity reconstruction of physical signals. However, conventional generative models are insufficient for these tasks due to their inability to incorporate domain-specific physical constraints during the generation process. To address this gap, we propose a Physics-Informed Auto-Encoder Based on Digital Twin (PIAE-DT), a novel architecture that synergizes data-driven learning with physics-based regularization.

As illustrated in Fig. 3, the PIAE-DT architecture integrates the data compression capabilities of an autoencoder with physics-informed constraints, aiming to enhance bearing fault diagnosis accuracy through two key mechanisms: (1) improving physical interpretability by embedding bearing dynamics into latent representations, and (2) enhancing sample diversity through probabilistic sampling in the digital twin space.

Specifically, the autoencoder is responsible for sample expansion and diversity preservation, while physics-informed modules replace the traditional hidden layer with a digital twin latent space that encodes bearing-specific dynamic parameters.

$$c, k = \text{Encoder}_\theta(x) \quad (16)$$

$$x' = \text{Decoder}_\theta(c, k) \quad (17)$$

$$s = \text{PIM}_\theta(\ddot{x}, t) \quad (18)$$

where the encoder serves as a network for extracting dynamic features, the decoder reconstructs the vibration signal from the extracted dynamic parameters, and PIM (Physics-Informed Module) is responsible for obtaining the displacement signal corresponding to the current acceleration signal.

In neural network training, maintaining consistent scale across output parameters is critical to facilitate stable gradient propagation and effective model convergence. However, in bearing fault diagnosis applications, physical parameters often exhibit extreme scale discrepancies—spanning several orders of magnitude—which severely hinder model training stability. For instance, the damping coefficient typically ranges from 10 to  $10^4$  N·s/m, while the stiffness varies between  $10^8$  and  $10^{10}$  N/m, resulting in a disparity of 6–8 orders of magnitude.

To address these scale-related challenges, we propose a dimensionless parameter transformation framework. Specifically, a physical scaling operation is applied to the neural network outputs, normalizing all parameters to a unified dimensionless scale. This is achieved by dividing each physical parameter by a corresponding baseline value derived from bearing specifications. For example, stiffness and damping

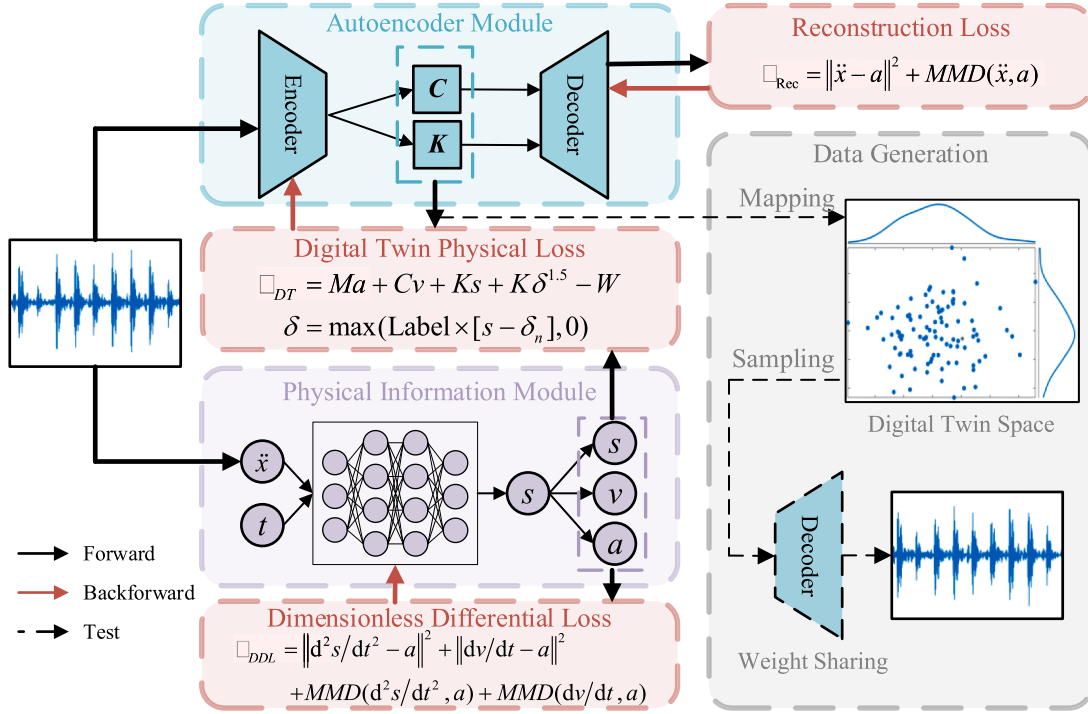


Fig. 3. Physics-informed auto-encoder based on digital twin.

are scaled as follows:

$$\tilde{k} = \frac{k}{10^8 \text{N/m}}, \tilde{c} = \frac{c}{10^2 \text{N}\cdot\text{s/m}} \quad (19)$$

where  $10^8 \text{N/m}$  and  $10^2 \text{N}\cdot\text{s/m}$  are baseline stiffness and damping values derived from bearing specifications.

This two-stage approach ensures that: (1) normalized parameters enable stable network training, and (2) rescaled outputs align with physical measurement scales, preserving engineering significance while maintaining numerical stability.

### 3.3.2. Design of PIAE-DT loss function

The composite loss function integrates three core components to balance data fidelity, physical plausibility, and digital twin consistency:

**Data Reconstruction Loss:** This component ensures pixel-level accuracy between the reconstructed and raw vibration signals.

$$\mathcal{L}_{\text{Rec}} = \|\ddot{x} - a\|^2 + \text{MMD}(\ddot{x}, a) \quad (20)$$

where  $\ddot{x}$  is real data,  $a$  is pseudo data,  $\text{MMD}$  is the maximum mean difference. This loss is to make the data distribution of real data and pseudo data consistent.

**Dimensionless Differential Loss:** This component encodes two dimensionless physical principles. First, it enforces dimensionless kinematic consistency by establishing differential relationships between normalized acceleration, velocity, and displacement. Second, it ensures distributional consistency by aligning the generated signals with real data through Maximum Mean Discrepancy (MMD).

$$\mathcal{L}_{\text{DDL}} = \left\| \frac{d^2s}{dt^2} - a \right\|^2 + \left\| \frac{dv}{dt} - a \right\|^2 + \text{MMD}\left(\frac{d^2s}{dt^2}, a\right) + \text{MMD}\left(\frac{dv}{dt}, a\right) \quad (21)$$

where  $s$  is dimensionless displacement,  $v$  is dimensionless velocity. This loss is to enable the physical information module to realize part of the integration function, so that the acceleration signal can be input and the displacement and velocity signals can be output. Therefore, not only the differential loss is required, but also the output displacement and ve-

locity signals need to be constrained in the data distribution, so that they are more like vibration signals.

**Digital Twin Physical Consistency Loss:** This component ensures that the normalized stiffness and damping parameters obey bearing dynamics.

$$\mathcal{L}_{\text{DT}} = Ma + Cv + Ks + K\delta^{1.5} - W, \delta = \max(\text{Label} \times [s - \delta_n], 0) \quad (22)$$

where  $\text{Label}$  is the label information (including all categories of the dataset), and  $\delta_n$  is the pre calculated column vector of defect size.

The total loss is formulated as (Aguilar et al., 2019):

$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_{\text{DDL}} + \beta \mathcal{L}_{\text{DT}} + \gamma \mathcal{L}_{\text{Rec}} - \frac{1}{2} (\log \alpha + \log \beta + \log \gamma) \quad (23)$$

where the final logarithmic regularization term mitigates degenerate solutions (e.g., excessively small or large weights) and promotes balanced contributions from all three loss components.

### 3.3.3. Training strategy and fault data generation of PIAE-DT

The proposed Physics-Informed Autoencoder with Digital Twin (PIAE-DT) framework employs a two-stage training paradigm designed to achieve both specialized module optimization and holistic parameter coordination. This strategy enables the network to simultaneously capture physical dynamics and data-driven representations, ultimately facilitating the generation of realistic and diverse fault data.

In Stage-1, the encoder, decoder, and physics-informed subnetwork are optimized independently, each guided by its own dedicated loss function to ensure targeted learning. Specifically, the reconstruction loss enforces consistency in the time-domain waveform between generated and measured vibration signals, thereby preserving fine-grained temporal features of the data. The digital twin loss integrates constraints from both the bearing dynamic equations and the latent feature distributions, ensuring that the encoded representations align with physically meaningful system behaviors. Meanwhile, the differential consistency loss evaluates the discrepancies between the second-order displacement derivatives and the first-order velocity derivatives of the generated and raw signals, promoting the preservation of dynamic response

characteristics in both amplitude and phase domains.

The parameters of each module are updated via the Adam optimizer with a learning rate according to:

$$\begin{cases} \theta_d \leftarrow \theta_d - \eta_1 \cdot \text{Adam}(\nabla_{\theta_d} L_{\text{Rec}}) \\ \theta_e \leftarrow \theta_e - \eta_1 \cdot \text{Adam}(\nabla_{\theta_e} L_{\text{DT}}) \\ \theta_{PI} \leftarrow \theta_{PI} - \eta_1 \cdot \text{Adam}(\nabla_{\theta_{PI}} L_{\text{DDL}}) \end{cases} \quad (24)$$

This modular optimization scheme allows each component to learn its respective functional role without interference from the others, which is crucial for achieving stable convergence in subsequent joint updates.

In Stage-2, the network transitions to a global optimization phase, wherein all parameters—along with the adaptive loss-weighting coefficients  $\alpha$ ,  $\beta$ , and  $\gamma$ —are updated jointly. The joint parameter update rules are given by:

$$\begin{cases} \theta \leftarrow \theta - \eta_2 \cdot \text{Adam}(\nabla_{\theta} L_{\text{total}}) \\ \alpha \leftarrow \alpha - \eta_2 \cdot \text{Adam}(\nabla_{\alpha} L_{\text{total}}) \\ \beta \leftarrow \beta - \eta_2 \cdot \text{Adam}(\nabla_{\beta} L_{\text{total}}) \\ \gamma \leftarrow \gamma - \eta_2 \cdot \text{Adam}(\nabla_{\gamma} L_{\text{total}}) \end{cases} \quad (25)$$

The training process alternates between Stage-1 and Stage-2 until the model converges or a pre-defined maximum number of iterations is reached.

Upon convergence, the learned dynamic parameters are subjected to statistical analysis to derive frequency-based probability distributions, thereby defining the digital twin parameter space  $\hat{p}(k, c)$ . New dynamic parameter pairs are generated through sampling-based inference and

subsequently decoded through the parameter-sharing decoder of PIAE:

$$\tilde{x}_j = \text{Decoder}_{\theta_{sh}}(k'_j, c'_j) \quad (26)$$

where  $\theta_{sh}$  denotes the shared decoder parameters within the PIAE architecture. This process produces vibration signals that are both physically plausible and statistically diverse.

This two-stage training paradigm offers several advantages. First, the independent module optimization in Stage-1 ensures that the encoder, decoder, and physics-informed components each learn task-specific features without mutual degradation. Second, the global optimization in Stage-2 harmonizes the objectives of all modules through adaptive weighting, enabling the network to dynamically adjust the importance of different physical and data-driven constraints during training. Finally, by leveraging the digital twin parameter space, the framework enables controlled and diverse fault data generation.

#### 3.4. Overall framework for fault diagnosis of imbalanced samples

To address the pervasive challenge of class imbalance in intelligent fault diagnosis of rotating machinery, this study proposes a digital-twin-based framework for generating synthetic minority-class samples and rebalancing the training dataset, as illustrated in Fig. 4. The overall goal is to alleviate the bias induced by dominant normal or major fault classes, thereby improving the recognition capability of diagnostic models on scarcely represented fault types. The framework consists of four main stages: imbalanced dataset construction, digital twin space

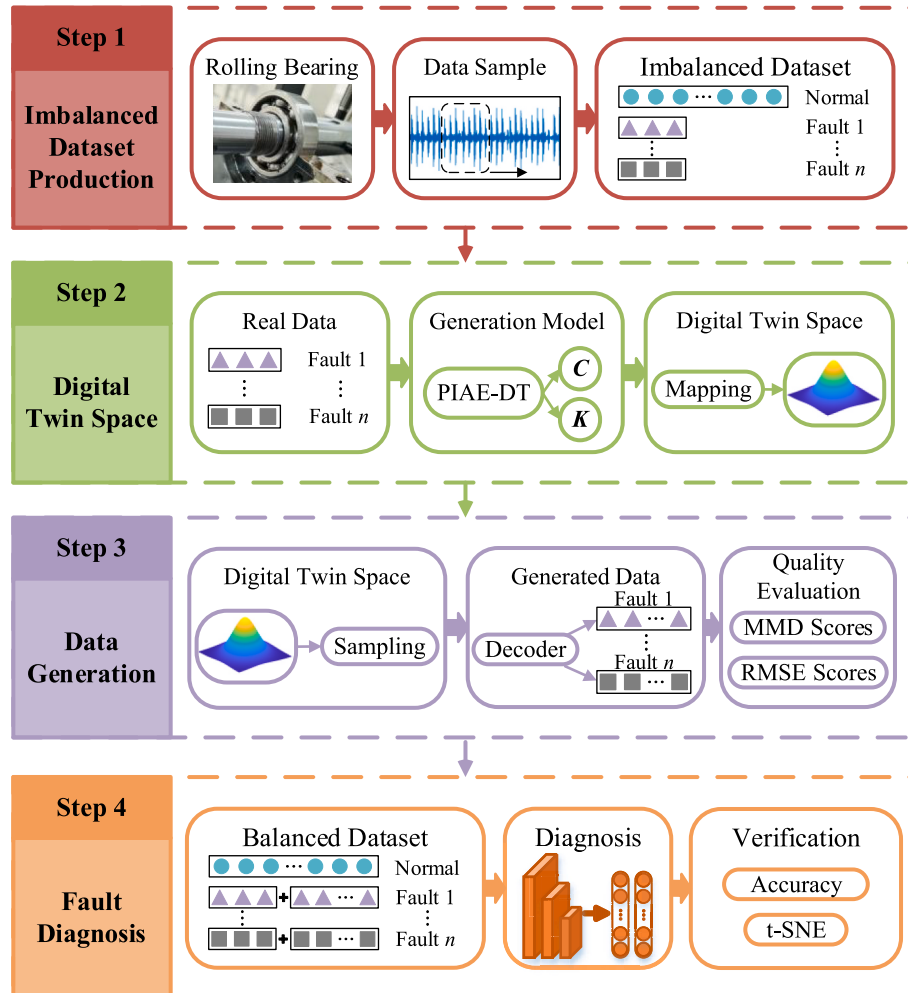


Fig. 4. Overall framework for unbalanced sample fault diagnosis.

modeling, sample generation with quality evaluation, and diagnostic validation.

In the first stage (Step 1), raw vibration signals collected from rolling bearings are processed using a sliding window approach to construct imbalanced datasets under prescribed class ratios. In Step 2, a physics-informed autoencoder digital twin (PIAE-DT) model is trained using real vibration data across multiple fault categories; the resulting model parameters construct a digital twin feature space that captures the latent dynamic distribution for each fault type. In Step 3, minority-class samples are generated by sampling within the digital twin space and decoding them back to the time domain. The authenticity and fidelity of these generated signals are assessed through time-series waveform similarity (RMSE) and feature space distribution similarity (MMD). Finally, in Step 4, the generated samples are merged with the original dataset at specific ratios to obtain balanced training datasets. A standard CNN diagnostic model is then trained and evaluated, where classification accuracy under varying imbalance ratios and t-SNE feature visualization serve as validation metrics to assess the efficacy of the synthetic data.

In summary, the proposed framework forms a closed-loop workflow from real-data-driven digital twin modeling to high-quality sample generation and diagnostic performance enhancement. By reconstructing balanced datasets through model-driven synthetic augmentation, this approach significantly mitigates the adverse impact of class imbalance and provides a practical, scalable paradigm for improving robustness and generalization in data-driven fault diagnosis under limited-sample conditions.

## 4. Experiment

### 4.1. Dataset description

The experiments are conducted on one publicly available dataset and one laboratory private dataset. The publicly available dataset is from Case Western Reserve University (CWRU) (Neupane and Seok, 2020), where the test bench mainly includes a driving motor, a torque sensor, a power meter, and an electronic controller (Fig. 5). The SKF6205 drive end bearing is tested with a downsampling frequency of 12 kHz at 1797 rpm, and the collected acceleration signals cover internal raceway faults, external raceway faults, rolling element faults, and normal categories with fault sizes of 7 mil, 14 mil, and 21 mil. The laboratory private dataset is obtained from the HD-FD-H-03P rotor gearbox fault simulation test bench (HFHP) (Fig. 6), which mainly consists of a magnetic powder brake, a bearing module, a flywheel, a torque speedometer, a one-way sensor, and a motor; the SKF6205 bearing is tested with a downsampling frequency of 12 kHz at 1500 rpm, and the collected acceleration signals include internal raceway faults, external raceway faults, and normal categories with fault sizes of 0.3 mm, 0.4 mm, and 1

mm.

The experimental bearing of the two data sets in this paper is SKF6205, and the specific parameters of the bearing are shown in Table 1.

### 4.2. Unbalanced dataset production

To better simulate the practical scenarios in industrial applications wherein fault samples are scarce, the original datasets were manually processed to construct imbalanced datasets. Initially, the raw vibration signals were segmented into samples using a sliding window with a window length of 1024 and a sliding step (stride) of 512. Following this, a total of  $10 \times 600$  and  $7 \times 600$  samples were obtained from the CWRU and HFHP datasets, respectively. Each dataset was randomly divided into a training set and a test set, with the latter denoted as Dataset A, consisting of  $10 \times 100$  and  $7 \times 100$  samples, correspondingly.

Within the training set, all normal (healthy) samples were designated as Dataset B ( $1 \times 500$ ), while 1 % of the fault samples were randomly selected and labeled as Dataset C ( $9 \times 5$  for CWRU and  $6 \times 5$  for HFHP). The remaining fault samples were labeled as Dataset D ( $9 \times 495$  for CWRU and  $6 \times 495$  for HFHP). Datasets B and C together formed an imbalanced dataset with an imbalance ratio of 1:100, which was utilized to train the fault diagnosis model. Meanwhile, Dataset C served as the known fault samples for training the generative model, and Dataset D was employed to evaluate the quality of generated data. The artificially generated fault samples were denoted as Dataset E. By adjusting the number of samples in Dataset E, Datasets B, C, and E were combined to create datasets with varying imbalance ratios for model training. The specific data partitioning is summarized in Table 2.

Through the above sample partitioning and processing strategy, both the CWRU and HFHP datasets were divided into multiple subsets (Datasets A–E), which were subsequently used for model training, testing, and evaluation of generated samples. This comprehensive configuration supports the experimental development and validation of fault diagnosis models, particularly with respect to their robustness and generalization capability under conditions of extreme class imbalance.

### 4.3. Experimental setup

The computational hardware configuration comprises 16 GB RAM and an NVIDIA RTX 4060 GPU. The implementation utilizes Python 3.10 with PyTorch 2.2 and NumPy as core dependencies. In order to ensure the robustness of the results, the performance was evaluated in 10 repeated independent experiments, so as to reduce the impact of any inherent randomness.

The model was trained for 200 epochs with a batch size of 5. The initial learning rate was set to 0.0005 and optimized using the Adam optimizer. To enhance convergence stability, a StepLR scheduler was applied, decaying the learning rate by a factor of 0.5 every 50 epochs. Gradient clipping with a maximum norm of 1.0 was employed to prevent gradient explosion. A penalty term was incorporated into the loss function as a form of regularization. For model storage, checkpoints were saved, and both the final model after 200 epochs and the model achieving the best training performance were retained.

To clearly present the network architecture, the structural design of the proposed PIAE-DT is summarized in Table 3, while the training hyperparameters adopted in this study are listed in Table 4.

### 4.4. Training results analysis

#### 4.4.1. Training loss results analysis

The training loss curves demonstrate that the proposed method exhibits good convergence. Specifically, all loss terms show a rapid decline at the early stage, indicating that the model quickly learns effective features from random initialization and significantly reduces error. After approximately 40–60 epochs, the curves gradually stabilize and remain

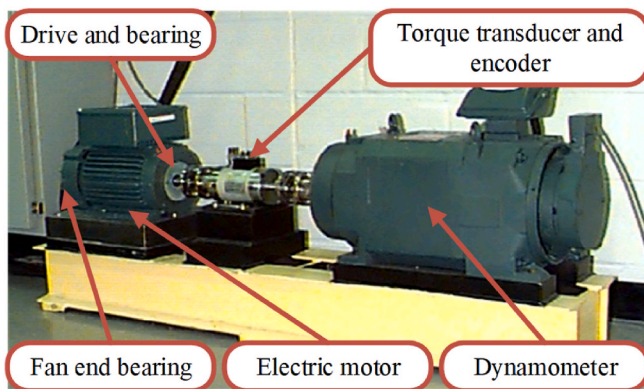


Fig. 5. Bearing test stand at case western reserve university (Neupane and Seok, 2020).



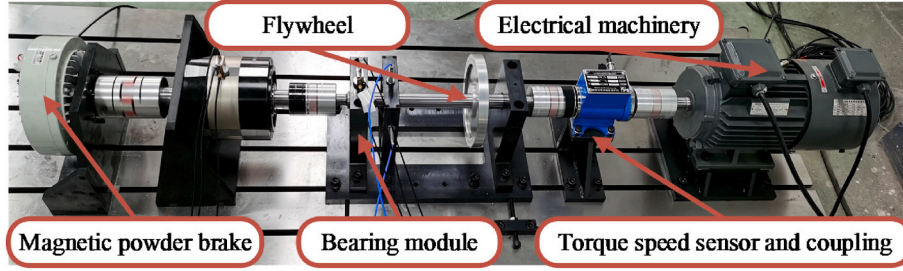


Fig. 6. HFHP Self built Bearing Fault Simulation Experiment Platform(Pan et al., 2023).

**Table 1**  
SKF6205 bearing parameters.

Name	Value	Name	Value
Number of rolling elements	9	Inner raceway diameter	31.06
Pitch Diameter	39	Outer raceway diameter	46.94
Rolling element diameter	7.94	contact angle	0

**Table 2**  
Usage of CWRU and HFHP datasets.

Datasets	Size	Purpose
CWRU Dataset A	$10 \times 100$	Test diagnostic model
HFHP Dataset A	$7 \times 100$	Training diagnostic models
CWRU Dataset B	$1 \times 500$	
HFHP Dataset B	$1 \times 500$	Train generative models, train diagnostic models
CWRU Dataset C	$9 \times 5$	
HFHP Dataset C	$6 \times 5$	Evaluate the quality of generated data
CWRU Dataset D	$9 \times 495$	
HFHP Dataset D	$6 \times 495$	Fill in imbalanced datasets and train diagnostic models
CWRU Dataset E	$9 \times (0-495)$	
HFHP Dataset E	$6 \times (0-495)$	

at a low level throughout the subsequent iterations, suggesting that the model reaches a favorable equilibrium and maintains training stability (see Fig. 7).

From the performance of different loss terms, the reconstruction loss  $L_{Rec}$  decreases rapidly and gradually stabilizes in the early stage, indicating that the model can effectively complete the reconstruction task; while  $L_{DDL}$  (dimensionless differential loss) and  $L_{DT}$  (digital twin loss) decrease relatively slowly, but show a similarly stable convergence

**Table 3**  
Parameters of PIAE-DT.

Network name	Layers	Input	Output
Encoder	Concat	$1024 + 10$	1034
	FC1 + ReLU	1034	512
	FC2 + ReLU	512	256
	FC3 + ReLU	256	128
	FC4	128	2
Decoder	Concat	$2 + 10$	12
	FC5 + ReLU	12	128
	FC6 + ReLU	128	256
	FC7 + ReLU	256	512
	FC8	512	1024
PI Module	FC9 + ReLU	1024	512
	FC10 + ReLU	512	256
	FC11 + ReLU	256	512
	FC12	512	1024

**Table 4**  
Training hyperparameters.

datasets	value	datasets	value
Batch size	5	Learning rate	0.0005
Epoch	200	Optimizer	Adam
LR scheduler	StepLR ( $\gamma = 0.5$ , step = 50)	Gradient clipping	Max-norm = 1.0
Checkpoint policy	Final/Best model	Regularization	Penalty term

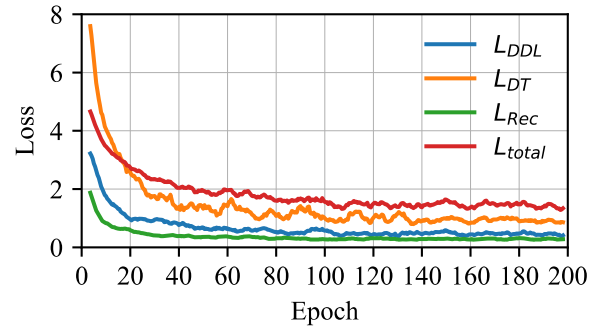


Fig. 7. Training loss curves of different loss terms during training.

trend, indicating that their regularization constraints are gradually being satisfied; the total loss  $L_{total}$  maintains a downward trend and eventually stabilizes at a relatively low level, further proving the coordination between each loss term. In summary, the proposed method can achieve effective error reduction during the training process and maintain a stable state over a long iteration period, fully verifying that the proposed method has good effectiveness and trainability.

As illustrated in Fig. 8, the overall loss function in this paper

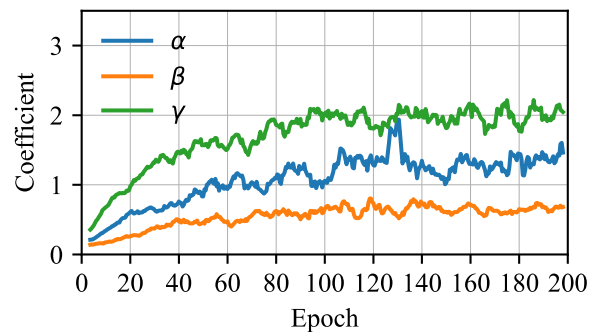


Fig. 8. Adaptive weighting coefficient curves during training.

introduces adaptive weighting coefficients  $\alpha$ ,  $\beta$  and  $\gamma$  for different sub-loss terms, which correspond to  $L_{DDL}$ ,  $L_{DT}$ , and  $L_{Rec}$ , respectively. As training progresses, these weighting coefficients can be dynamically adjusted to balance the contributions of each task to the overall optimization at different stages.

Specifically, the coefficient  $\gamma$  (corresponding to  $L_{Rec}$ ) remains at a relatively high level (close to 2) throughout the training process, indicating that at the optimization start, the model emphasizes the importance of the reconstruction loss, thereby ensuring that the generated results are consistent with the input quality; the coefficient  $\alpha$  (corresponding to  $L_{DDL}$ ) gradually increases in the early training stage and tends to stabilize at around 1 in the later stage, reflecting continuous attention to the discrimination constraint, which helps improve the model's discriminative capability; the coefficient  $\beta$  (corresponding to  $L_{DT}$ ) remains at a relatively low level overall, indicating that this constraint accounts for a relatively small proportion in the overall optimization and mainly plays a supporting role.

Such an adaptive weighting mechanism enables the loss function to automatically adjust the weights during training according to the optimization needs of different tasks, thereby avoiding the dominance of a single loss term, ensuring that the model takes into account both reconstruction quality and discriminative capability during convergence, and achieving more stable and efficient training results.

#### 4.4.2. Digital twin space result analysis

Dataset C, containing known fault samples, was used to train a generative model and obtain the probability distribution of key dynamic parameters, thereby constructing the CWRU digital twin space (Fig. 9). The results show that when defect size variation is small, the dimensionless stiffness and damping parameters remain essentially stable and exhibit little correlation with defect size or type, being more strongly influenced by the bearing model and acquisition environment. While using fixed parameter values does not compromise the realism of vibration responses, it limits sample diversity; this can be mitigated by converting parameter histograms into probability distributions and sampling from them. Following the same procedure, the dynamic parameter distributions for HFHP Dataset C were obtained (Fig. 10). Despite originating from the same bearing model, the two datasets yield different parameter distributions, indicating that dynamic parameters are not solely determined by defect characteristics but are also sensitive to experimental setup, environmental noise, and sensor properties.

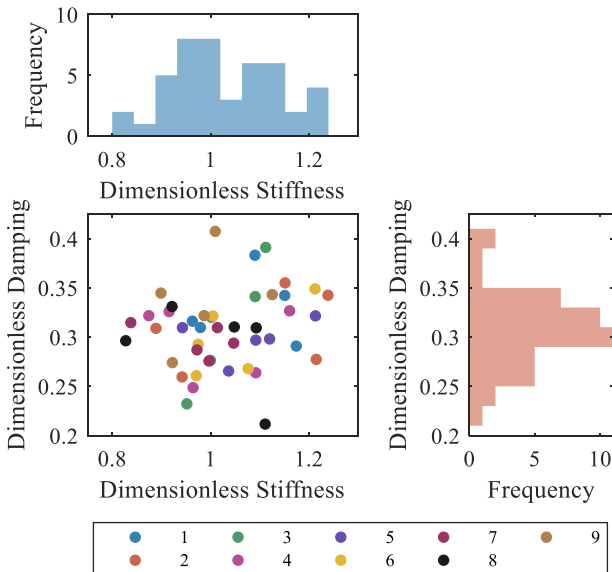


Fig. 9. CWRU digital twin space.

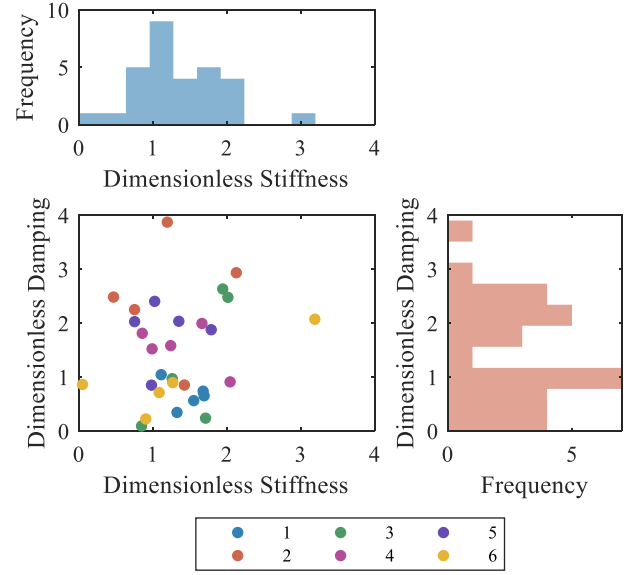


Fig. 10. HFHP digital twin space.

Overall, Figs. 9 and 10 suggest that these parameter distributions are weakly related to defect size and type but strongly reflective of the test environment.

In summary, constructing a data-driven digital twin space based on probabilistic dynamic parameters not only preserves the physical authenticity of vibration signals but also enhances diversity and generalization, laying the foundation for robust sample generation and downstream model training.

#### 4.4.3. Generated data reliability analysis

New fault data samples were generated by randomly sampling from the digital twin space of both datasets and combining the sampled parameters with their corresponding labels. These were then fed into a trained weight-sharing decoder to reconstruct vibration signals. As shown in Figs. 11 and 12, the generated signals closely match the original data in both datasets. For CWRU (Fig. 11), the generated data align well with the overall waveform shape and amplitude distribution of the real signals. In HFHP (Fig. 12), the generated signals exhibit comparable amplitude ranges and temporal patterns. Notably, more distinct vibration behaviors are observed in defect regions, while signals in non-defect areas remain flat and consistent. Interestingly, the generated signals in non-defect segments are even smoother than in the real

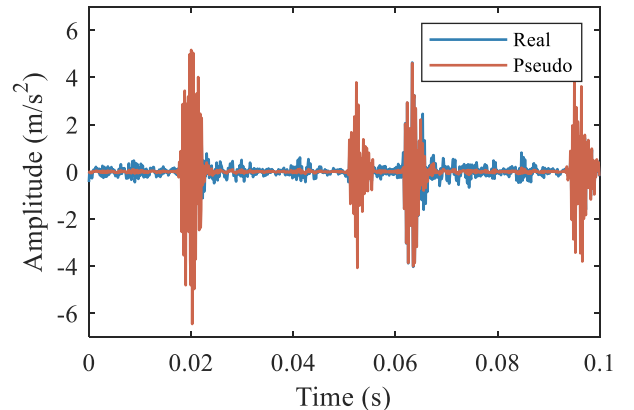


Fig. 11. Comparison between real data and generated data of CWRU.

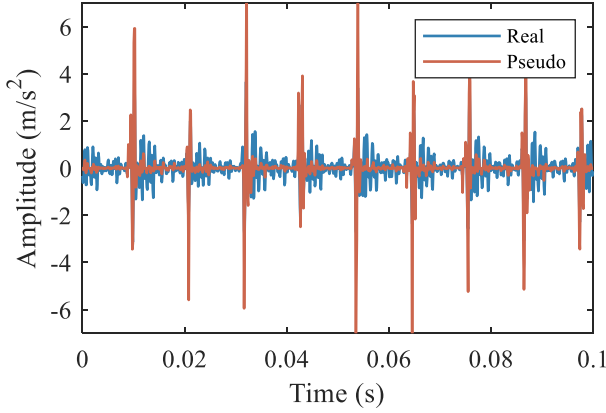


Fig. 12. Comparison between real data and generated data of HFHP.

data, indicating that the model not only captures key temporal characteristics but may also enhance them during synthesis.

In summary, the synthesized signals preserve essential fault-related features while improving smoothness and consistency, demonstrating the proposed method's capability in generating high-fidelity and reliable fault data for downstream tasks.

#### 4.5. Apply to fault diagnosis

##### 4.5.1. Quantitative evaluation metrics and imbalance ratio design

In order to quantify the difference between simulated and measured data, two metrics are used in this paper: Maximum Mean Discrepancy (MMD)(Zhang et al., 2022b) and Root Mean Square Error (RMSE)(Lei et al., 2012). MMD is employed to evaluate the difference in data distribution, whereas RMSE mainly focuses on the amplitude difference between simulated and measured data. The mathematical definitions are as follows:

$$MMD[F, p, q] = \sup_{f \in F} (E_{x \sim p}[f(x)] - E_{\hat{x} \sim q}[f(\hat{x})]) \quad (27)$$

$$RMSE = \sqrt{\sum_{n=1}^N (x(n) - \hat{x}(n))^2 / N} \quad (28)$$

in addition, the fault diagnosis performance is evaluated using classification accuracy, which is defined as:

$$Accuracy = \frac{\sum_{i=1}^C N_{ii}}{\sum_{i=1}^C \sum_{j=1}^C N_{ij}} \quad (29)$$

where  $N_{ii}$  denotes the number of samples whose true class is  $i$  and predicted class is  $i$ , and  $C$  represents the number of classes.

To construct different imbalance scenarios, the following sample ratios are designed, as summarized in Table 5.

Table 5  
Imbalance ratio.

Real normal sample	real fault samples	Generate fault samples	proportion
500	5	0	100:1
500	5	5	50:1
500	5	15	25:1
500	5	45	10:1
500	5	95	5:1
500	5	145	2:1
500	5	495	1:1

In the imbalanced classification experiments, 500 real normal samples from Dataset B and 5 real fault samples from each fault type in Dataset C are employed as the training set. The number of generated samples for each fault type is gradually increased from 0 to 495, resulting in a progressive reduction of the imbalance ratio from 100:1 to 1:1. All data in Dataset A are utilized as the test set for the classification network. To further verify the effectiveness of the proposed method in handling imbalanced samples, a uniform diagnostic network (WDCNN) is adopted for all experiments under identical parameter settings(Zhang et al., 2017).

##### 4.5.2. Baselines for extreme imbalance scenarios

To begin with, we evaluate the baseline diagnostic performance in the extremely imbalanced 100:1 scenario without any augmentation, so as to establish a reference point. As reported in Table 6 (CWRU) and Table 7 (HFHP), the diagnostic accuracies of all models hover around only 62 %, indicating a substantial degradation in fault recognition capability due to the severe class imbalance.

##### 4.5.3. Comparative experiments under different imbalanced ratios

In this work, the comparison models include SMOTE, VAE, WGAN-GP, and the dynamic model. The dynamic model employs the same simulation parameters as in the literature to solve the differential equations.

Next, we assess the quality of the generated samples under the most challenging imbalance level (100:1), as high-fidelity data generation is the foundation for effective augmentation. To this end, we adopt the MMD and RMSE metrics to quantitatively measure the distribution similarity and waveform fidelity between the synthesized and real samples. As depicted in Fig. 13(a), the proposed model consistently achieves the smallest MMD values across all nine fault categories on the CWRU dataset, indicating that the distribution of the generated samples is closest to that of the authentic data and thus demonstrating a superior capability in capturing the underlying data distribution. A similar observation can be made from Fig. 13(b), where the proposed model again records the lowest RMSE scores, further confirming that the generated samples most closely resemble the raw signals. Comparable trends are observed on the HFHP dataset: as illustrated in Fig. 14, PIAE-DT also obtains the lowest MMD and RMSE values, validating its excellent generalization ability across different datasets and highlighting the high fidelity of the synthesized samples.

Building upon the high-quality data generation capability demonstrated above, we further evaluate how the downstream diagnostic performance evolves after sample augmentation as the imbalance ratio is gradually relaxed from 50:1 to 1:1. As summarized in Tables 8 and 9, all augmentation strategies lead to noticeable improvements in accuracy, confirming the positive effect of synthetic sample supplementation on imbalance mitigation. Notably, the proposed PIAE-DT model consistently outperforms all competing methods across every imbalance setting. In particular, when the class distribution becomes perfectly balanced (1:1), PIAE-DT achieves peak accuracies of 97.6 % on the CWRU dataset and 99.86 % on the HFHP dataset, fully demonstrating its superior capability to alleviate data imbalance and enhance diagnostic performance through high-fidelity sample generation.

Building on these promising results, Fig. 15 illustrates the evolution of diagnostic accuracy improvement for different augmentation methods as the imbalance ratio is gradually relaxed from 50:1 to 1:1. As the class distribution becomes more balanced, all approaches exhibit progressively increasing performance gains, confirming the positive effect of synthetic sample supplementation. Among them, SMOTE and

Table 6  
Baseline diagnostic accuracy (100:1) on the CWRU dataset.

Method	100:1					Average (Baseline)
No processing	61.8	62.0	61.7	62.3	62.3	62.02

**Table 7**

Baseline diagnostic accuracy (100:1) on the HFHP dataset.

Method	100:1					Average (Baseline)
No processing	62.57	61.00	62.71	62.57	64.86	62.74

VAE show stable improvements under moderate imbalance, while WGANGP yields noticeable boosts in relatively balanced scenarios. Remarkably, the proposed PIAE-DT consistently achieves the largest accuracy gains across all settings and reaches peak improvements of 36.6 % and 37.1 % at the 1:1 ratio on the CWRU and HFHP datasets, respectively, thereby demonstrating its superior capability to mitigate data imbalance and enhance diagnostic performance through high-fidelity sample generation.

Under the highly imbalanced 100:1 setting, all models show a sharp drop in diagnostic accuracy to around 62 %, highlighting the detrimental impact of severe class imbalance on fault recognition. However, leveraging high-fidelity synthetic samples for data augmentation effectively mitigates this issue. Quantitative evaluations using MMD and RMSE demonstrate that PIAE-DT produces generated samples that are most consistent with real data in both distribution and waveform characteristics, indicating a powerful data modeling capability. Subsequent diagnostic experiments under varying imbalance ratios (from 50:1 to 1:1) reveal that all augmentation strategies improve accuracy, with performance gains increasing as the class distribution becomes more balanced. Notably, PIAE-DT consistently outperforms competing

**Table 8**

Diagnostic accuracy under different imbalance ratios on the CWRU dataset.

Method	50:1	25:1	10:1	5:1	2:1	1:1
Dynamics	68.2	71.2	77.4	81.3	91.4	95.4
SMOTE	74.7	79.7	81.0	85.9	92.8	96.3
VAE	78.0	78.6	80.8	83.2	91.2	95.8
WGANGP	78.5	80.4	81.2	89.1	93.5	96.8
PIAE-DT	80.0	82.0	85.4	90.6	93.7	97.6

methods, achieving peak accuracies of 97.6 % on the CWRU dataset and 99.86 % on the HFHP dataset, and yielding the largest accuracy gains of 36.6 % and 37.1 % at the 1:1 ratio. These results demonstrate the superior ability of PIAE-DT to alleviate data imbalance and enhance fault diagnosis through high-quality sample generation.

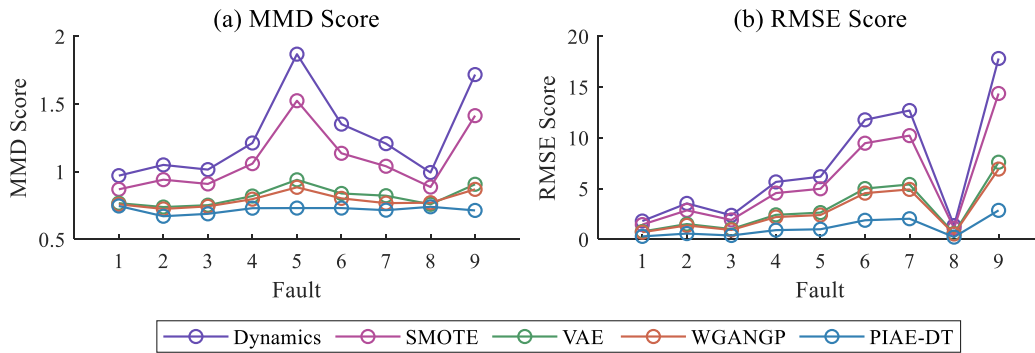
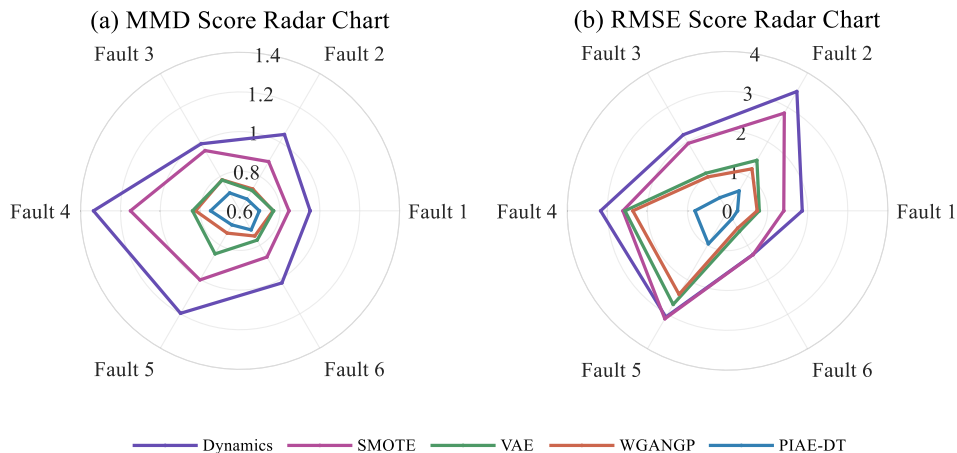
#### 4.5.4. Ablation experiments under different imbalanced ratios

To further elucidate the individual contribution of each structural component in the proposed framework toward mitigating sample

**Table 9**

Diagnostic accuracy under different imbalance ratios on the HFHP dataset.

Method	50:1	25:1	10:1	5:1	2:1	1:1
Dynamics	62.71	70.86	81.14	84.29	87.57	96.43
SMOTE	74.43	76.43	77.57	88.57	90.00	99.43
VAE	76.71	74.86	84.43	87.00	90.86	96.86
WGANGP	72.00	78.57	85.00	84.29	92.71	97.57
PIAE-DT	81.71	85.86	86.86	92.00	96.00	99.86

**Fig. 13.** MMD and RMSE scores of generated samples on the CWRU dataset.**Fig. 14.** MMD and RMSE scores of generated samples on the HFHP dataset.



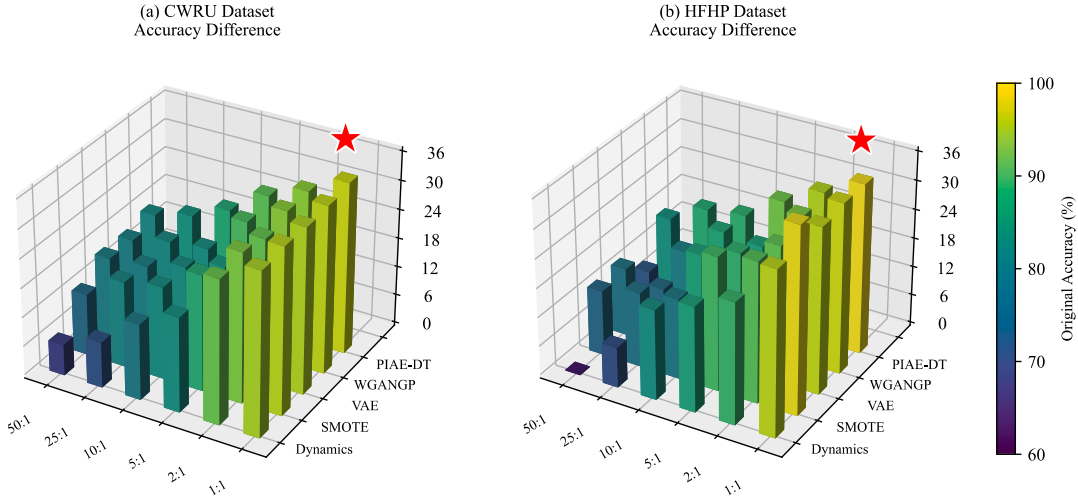


Fig. 15. Diagnostic accuracy improvements of different augmentation methods under various imbalance ratios.

imbalance and enhancing diagnostic performance, we conduct ablation studies by progressively removing or modifying key modules of the model. Such investigations are essential as they allow us to: (i) quantitatively assess the importance of each module to the overall performance; (ii) verify the rationality of the model design; and (iii) provide guidelines for future development and simplification of the model for practical deployment. By systematically comparing the diagnostic accuracies of different ablation configurations across various imbalance ratios, we aim to gain deeper insight into the effectiveness and necessity of each architectural element.

To this end, five comparative models are constructed as follows:

Model 1 (Full model) represents the original PIAE-DT framework, which integrates both the Physics-Informed (PI) module and the Autoencoder (AE) module. Model 2 (w/o PI) removes the PI module and retains only the AE, allowing evaluation of the contribution of Physics-Informed incorporation to sample generation fidelity and diagnostic performance. Model 3 (w/o AE) discards the AE module while keeping the PI, thereby examining the role of the AE in signal reconstruction and feature preservation. Model 4 (AE  $\rightarrow$  CNN) replaces the fully connected structure of the AE with a Convolutional Neural Network (CNN) to investigate the effect of local receptive field modeling on generation performance. Finally, Model 5 (AE  $\rightarrow$  RNN) substitutes the AE with a Recurrent Neural Network (RNN) to explore whether sequential modeling capability is better suited for vibration signal generation.

In the subsequent experiments, these five models are evaluated under different imbalance ratios ranging from 50:1 to 1:1, enabling a comprehensive analysis of the functionality and effectiveness of each component within the proposed fault diagnosis framework.

In the ablation study, we first assess the quality of the generated samples under the most severe imbalance scenario (100:1) using the quantitative MMD and RMSE metrics. As summarized in Figs. 16 and 17, Model 5 exhibits the highest generation quality, followed by Model 4, while Model 1 delivers a moderate performance, and Model 3 and Model 2 perform comparatively worse. These results indicate that replacing the fully connected AE structure with an RNN (Model 5) or CNN (Model 4) can enhance the fidelity of the generated samples relative to the original architecture (Model 1), due to their stronger capability in modeling temporal dynamics and local feature patterns.

However, it should be noted that the extremely low MMD/RMSE scores of Models 4 and 5 may result from overfitting to the limited minority fault samples, which does not necessarily translate to improved generalization in downstream diagnostic tasks. In contrast, Model 1 achieves a more balanced trade-off between generation fidelity and generalization ability, which is essential for practical data augmentation. The diagnostic performance results will be presented next to further support this observation (Fig. 18).

Based on the above observations of generation quality, we further examine whether these differences can translate into improvements in downstream diagnostic performance under varying imbalance ratios. The classification accuracies of the five models on the CWRU and HFHP datasets are summarized in Tables 10 and 11, respectively.

As seen, Model 1 consistently delivers superior diagnostic accuracy across most imbalance ratios on both datasets, particularly when the class distribution becomes relatively balanced (e.g., 2:1 and 1:1, where accuracies reach 93.7 %/97.6 % on CWRU and 96.00 %/99.86 % on HFHP). Although Model 4 and Model 5 occasionally achieve slightly

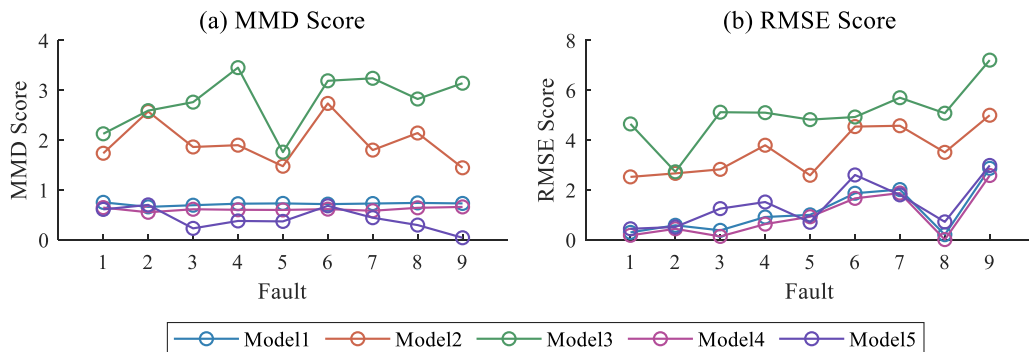


Fig. 16. MMD and RMSE scores of generated samples on the CWRU dataset.

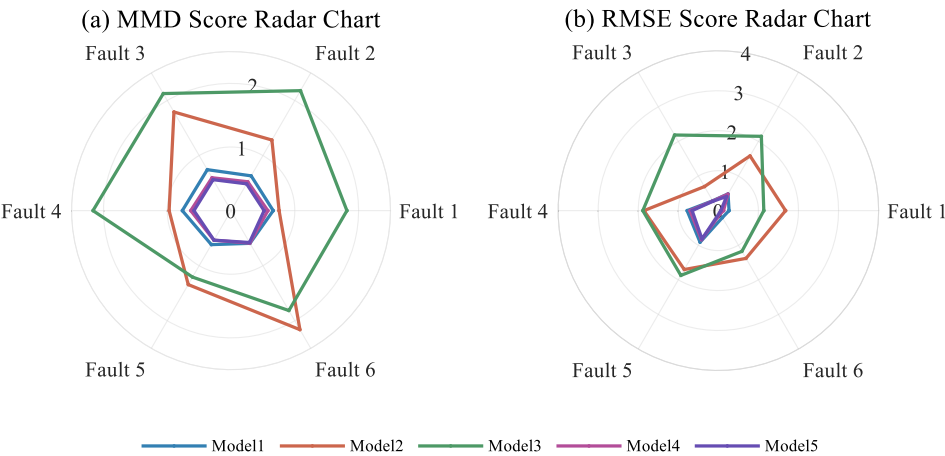


Fig. 17. MMD and RMSE scores of generated samples on the HFHP dataset.

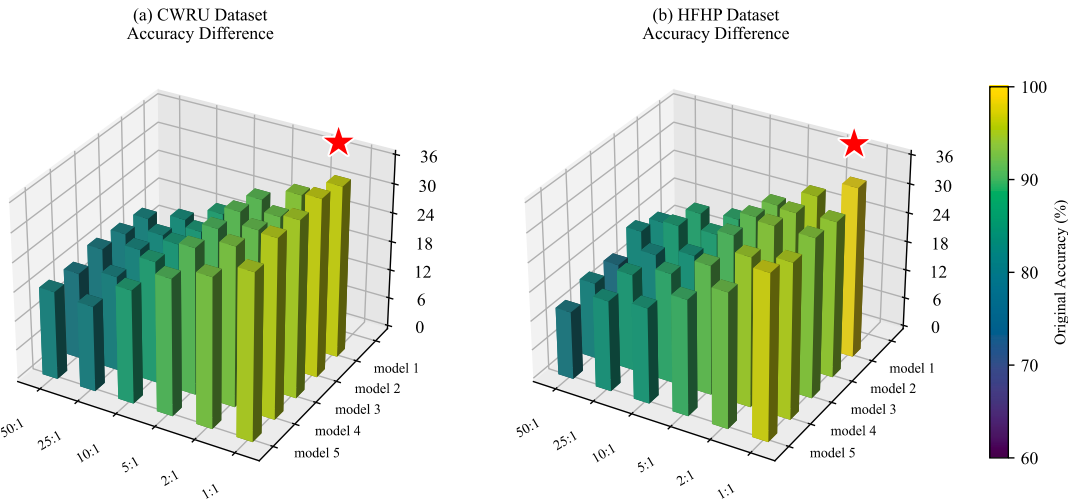


Fig. 18. Diagnostic accuracy improvements of different augmentation methods under various imbalance ratios.

**Table 10**  
Diagnostic accuracy under different imbalance ratios on the CWRU dataset.

Method	50:1	25:1	10:1	5:1	2:1	1:1
Model 1	80	82	85.4	90.6	93.7	97.6
Model 2	79.2	81.2	84.2	90.5	91.7	97.5
Model 3	78.4	80.6	84	89.4	91.4	95.6
Model 4	78	79.7	85.1	90.2	92.8	96.7
Model 5	80	79.4	85.1	89.8	92.5	95.8

better performance in the extremely imbalanced or moderately imbalanced scenarios (e.g., 50:1 on CWRU and 10:1 on HFHP), their diagnostic accuracies fluctuate more significantly as the ratio changes, which suggests a weaker stability and generalization capability. In contrast, Model 2 and Model 3, which remove either the PI or AE module, show a noticeable decline in diagnostic performance under all settings, highlighting the necessity of both components in the proposed architecture.

These results substantiate that a balanced generation–diagnosis trade-off, rather than the absolute fidelity of synthetic samples, is key to effective performance enhancement under class imbalance. The proposed Model 1 strikes this balance well, leading to overall better and more stable diagnostic outcomes.

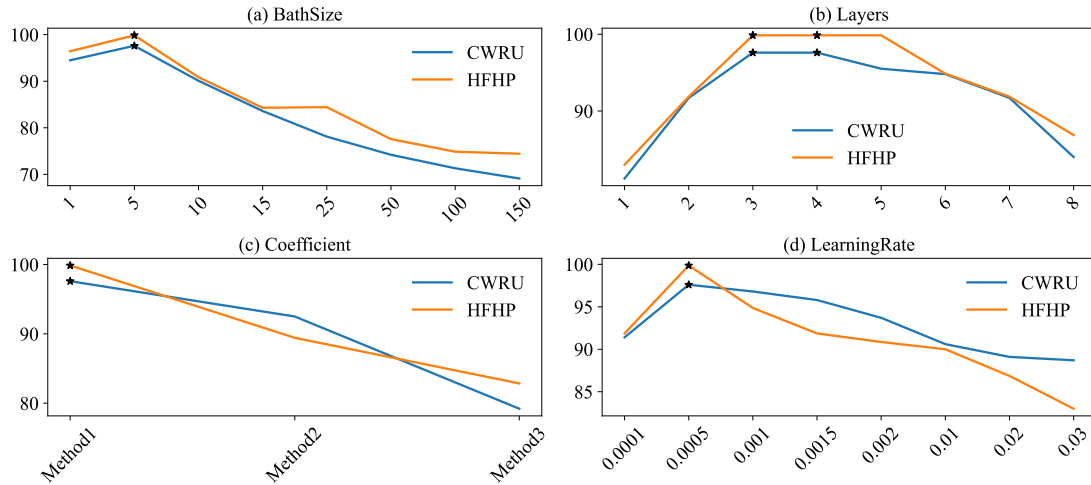
**4.5.5. Analysis of model structure and hyperparameters**

Under conditions of extreme class imbalance and limited sample size, model performance is highly sensitive to training hyperparameters and network architecture. Systematic experimental results demonstrate that batch size, network depth, loss-coefficient configuration, and learning rate all exert decisive influence on recognition accuracy, with each factor exhibiting an “optimal within moderation” pattern. These findings highlight the importance of carefully selecting architectural and training settings to ensure stability and generalization in such challenging scenarios.

Specifically (see Fig. 19), with respect to batch size, increasing it from 1 to 5 leads to gradual improvement in performance, reaching the optimum at a batch size of 5. Under this setting, each training batch

**Table 11**  
Diagnostic accuracy under different imbalance ratios on the HFHP dataset.

Method	50:1	25:1	10:1	5:1	2:1	1:1
Model 1	81.71	85.86	86.86	92.00	96.00	99.86
Model 2	81.14	84.57	85.00	90.86	93.86	94.29
Model 3	76.43	80.86	83.00	89.43	93.57	93.43
Model 4	80.14	84.29	86.86	90.86	94.86	96.14
Model 5	78.14	82.86	83.86	88.00	91.86	98.00

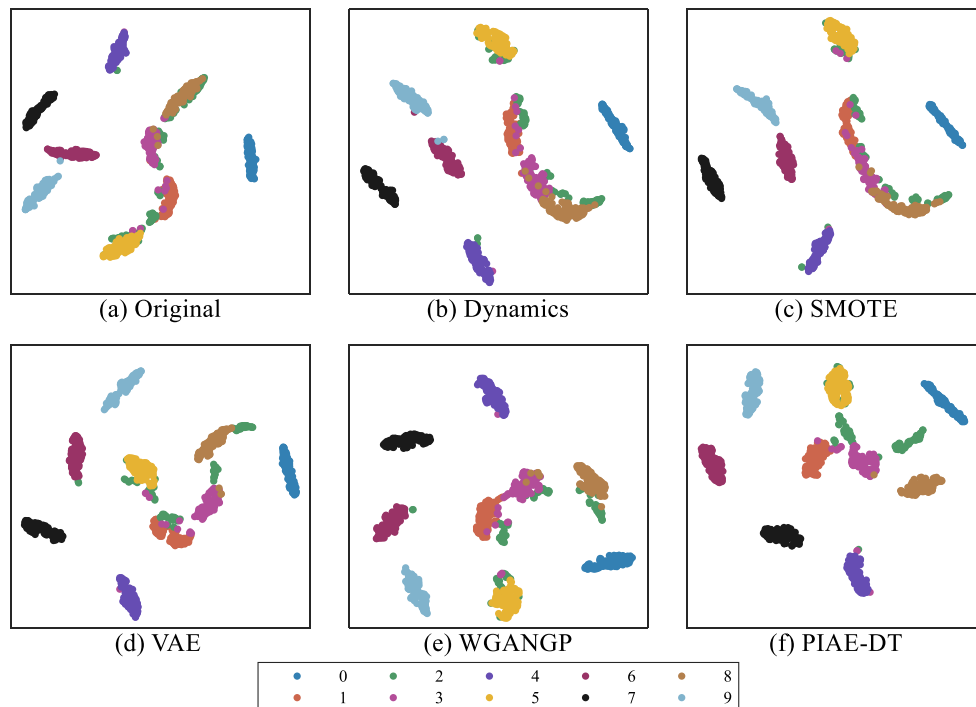


**Fig. 19.** Ablation study on hyper-parameter configurations. The effects of (a) batch size, (b) network depth, (c) loss coefficient strategies, and (d) learning rate on model performance are investigated.

encompasses all five available samples per class, thereby enabling more representative gradient updates. However, when the batch size is further increased ( $\geq 10$ ), performance markedly deteriorates, suggesting that small-batch training is more conducive to generalization. Regarding network depth, expanding from 1 to 3–4 layers substantially improves performance, indicating that moderate depth enhances feature extraction capacity; nonetheless, increasing the depth to 5 layers or more results in performance degradation, likely due to overfitting and gradient propagation difficulties under the small-sample regime. For loss-coefficient strategies, the adaptive scheme (Method1) achieves the best results, as dynamic adjustment of loss weights effectively balances the contribution of different terms during training. In contrast, removing the regularization term (Method2) degrades performance, underscoring its importance for stability and generalization, while fixed coefficients

(Method3) yield the worst outcomes, indicating that lack of flexibility hampers learning. Finally, learning rate experiments reveal a clear rising-then-falling trend: performance improves as the rate increases from 0.0001 to 0.0005, but sharply declines when raised to 0.001 or beyond. This suggests that overly large learning rates cause unstable convergence or divergence, whereas excessively small rates result in under-optimization; balancing convergence speed and stability, 0.0005 emerges as the optimal choice.

In summary, under small-sample and highly imbalanced conditions, model training should adhere to the principle of “optimal within moderation.” A small batch size (e.g., 5), moderate network depth (3–4 layers), a regularized adaptive loss-coefficient mechanism, and a medium learning rate (0.0005) together provide the most favorable performance and generalization capability. These results not only identify



**Fig. 20.** t-SNE diagrams of various generation methods at a 1:1 scale.

the optimal configuration ranges of critical factors but also offer valuable guidance for model design and optimization in similar problem settings.

#### 4.5.6. Visualization analysis

To further evaluate the effectiveness of the proposed model, t-distributed stochastic neighbor embedding (t-SNE) was employed to visualize the hidden and output layer features of the diagnostic classifier under the 100:1 imbalance condition. As illustrated in Fig. 20, the hidden-layer feature distributions of all networks reveal that fault types 2 (green), 5 (yellow), 3 (pink), and 1 (orange) tend to cluster closely, suggesting these categories are relatively difficult to distinguish. Nonetheless, the proposed model (Fig. 20(f)) exhibits markedly clearer cluster boundaries than the comparison networks (Fig. 20(a)–(e)), whose feature distributions overlap severely. Although the samples generated by each model provide some improvement over the non-augmented baseline (Fig. 20(a)), the synthetic data produced by PIAE-DT enables notably richer and more separable representations.

Similarly, Fig. 21 visualizes the output-layer features of the classifier. Without augmentation (Fig. 21(a)), most fault types are not clearly delineated, confirming that severe imbalance leads to poor diagnostic separability. Although Models 2–5 (Fig. 21(b)–(e)) improve the clustering to some extent, fault types 5 (yellow) and 4 (purple) remain difficult to distinguish, potentially due to overly similar generated samples. In contrast, the proposed model (Fig. 21(f)) forms well-separated, compact clusters across all fault categories, indicating that PIAE-DT not only improves diagnostic accuracy, but also enhances feature extraction quality by providing high-fidelity, discriminative synthetic data.

## 5. Conclusion

This study developed a Physics-Informed Auto-Encoder based on Digital Twin (PIAE-DT) to enhance bearing fault diagnosis under imbalanced sample conditions. By embedding a dimensionless physical information module into the auto-encoder framework, the method

identified dynamic model parameters with clear physical meaning, thereby improving interpretability. A probability distribution mapping of these parameters was used to generate diverse and dynamically consistent fault data in the digital twin space, and reconstructed vibration signals validated the physical plausibility of the augmented samples. Experiments on public and private datasets demonstrated that the proposed method achieved diagnostic accuracies of 97.60 % and 99.86 %, respectively, under imbalanced conditions, and outperformed existing augmentation approaches, indicating its strong potential to improve the reliability of bearing fault diagnosis. Across both public and private datasets, the proposed method improved diagnostic performance, with the most notable benefits observed in addressing class imbalance. The augmented signals closely matched the characteristics of real measurements, supporting the reliability and applicability of the approach under diverse conditions. These improvements were particularly evident for minority fault classes, suggesting that the physics-guided constraint helps preserve critical dynamic features and strengthens classifier robustness under severe imbalance.

Looking forward, bearing fault diagnosis is expected to move toward more advanced, physics-guided neural architectures that integrate the adaptability of data-driven methods with the interpretability of physical modeling. With the increasing adoption of digital twin technologies in industrial monitoring, future research will incorporate temporal convolutional networks, Transformers, or other efficient feature extractors into the PIAE-DT, and develop more refined bearing dynamic models to better capture complex operational behaviors and improve the accuracy of generated fault data. The current network structure, while effective, limits the ability to reconstruct signals with higher fidelity, and its performance has been validated mainly under conditions with clear and distinct fault features. In lower-fidelity digital twins or noisier environments, diagnostic performance may experience a slight reduction. Building on the envisioned integration of advanced architectures and refined dynamic models, the framework will be improved and may have the potential to be applied to both real-time monitoring and zero-sample fault diagnosis, while also making it possible to be deployed in environments where labeled fault data are scarce or unavailable. With

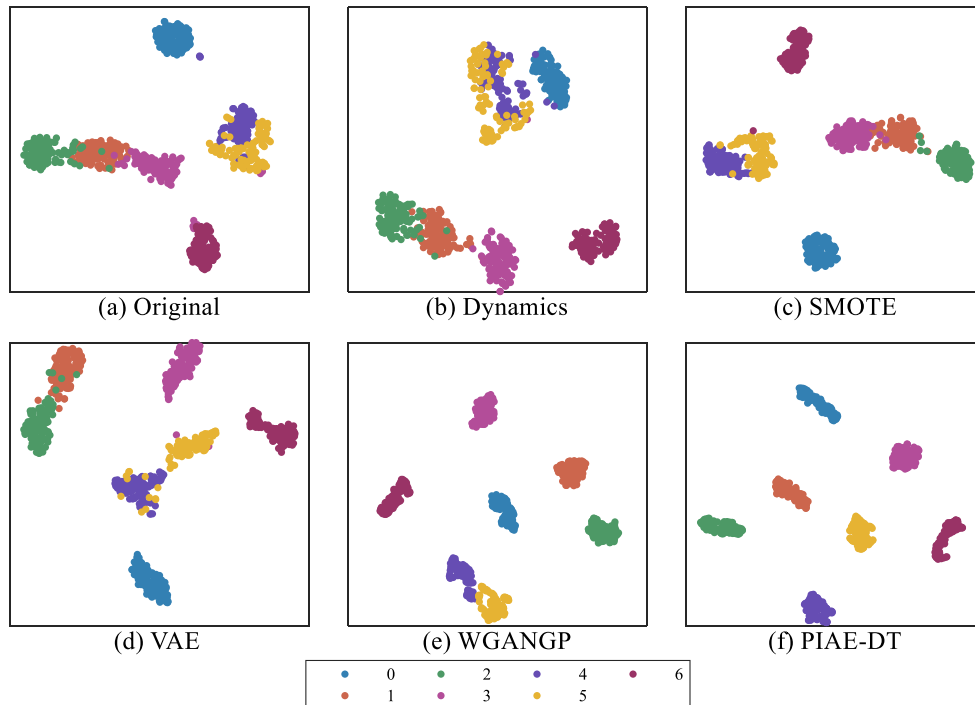


Fig. 21. t-SNE diagrams of various generation methods at a 1:1 scale.



additional improvements, the method may also be able to be tested under varying and previously unseen conditions to help ensure robustness and broaden its applicability in practical industrial scenarios.

### CRedit authorship contribution statement

**Zhiwu Shang:** Writing – review & editing, Project administration, Methodology, Conceptualization. **Ziyu Wang:** Writing – original draft, Visualization, Data curation. **Cailu Pan:** Supervision, Investigation. **Wanxiang Li:** Supervision, Investigation. **Maosheng Gao:** Project administration, Investigation, Conceptualization.

### Declaration of competing 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.

### Acknowledgements

This work was financially supported by The National Natural Science Foundation of China (62576244), The Key Program of Natural Science Foundation of Tianjin, China (21JCZDJJC00770) and the National Natural Science Foundation of China and the Civil Aviation Administration of China joint funded projects (U1733108).

### Data availability

The authors do not have permission to share data.

### References

- Aguilar, E., Bolaños, M., Radeva, P., 2019. Regularized uncertainty-based multi-task learning model for food analysis. *J. Vis. Commun. Image Represent.* 60, 360–370.
- Ali, H., Safdar, R., Ding, W.L., Zhou, Y.Q., Yao, Y., Yao, L., Gao, F.R., 2025a. Intelligent machine learning-based multi-model fusion monitoring: application to industrial physio-chemical systems. *Control Eng. Pract.* 162, 25.
- Ali, H., Safdar, R., Rasool, M.H., Anjum, H., Zhou, Y.Q., Yao, Y., Yao, L., Gao, F.R., 2024a. Advance industrial monitoring of physio-chemical processes using novel integrated machine learning approach. *J. Ind. Inf. Integr.* 42, 22.
- Ali, H., Safdar, R., Zhou, Y.Q., Yao, Y., Yao, L., Zhang, Z., Ding, W.L., Gao, F.R., 2025b. A novel dynamic machine learning-based explainable fusion monitoring: application to industrial and chemical processes. *Mach. Learn.-Sci. Technol.* 6, 38.
- Ali, H., Safdar, R., Zhou, Y.Q., Yao, Y., Yao, L., Zhang, Z., Rasool, M.H., Gao, F.R., 2024b. Robust statistical industrial fault monitoring: a machine learning-based distributed CCA and low frequency control charts. *Chem. Eng. Sci.* 299, 24.
- Ashtekar, A., Sadeghi, F., Stacked, L.E., 2010. Surface defects effects on bearing dynamics. *Proc. Inst. Mech. Eng. Part J-J. Eng. Tribol.* 224, 25–35.
- Chen, Y.X., Feng, G.X., Chen, H.T., Gou, L.F., Wang, S.J., 2025. A machine learning-based interpretable fault diagnosis method for rolling bearings under unbalanced samples. *Eng. Appl. Artif. Intell.* 158, 18.
- Cui, L.L., Xiao, Y.C., Liu, D.D., Han, H.G., 2024. Digital twin-driven graph domain adaptation neural network for remaining useful life prediction of rolling bearing. *Reliab. Eng. Syst. Saf.* 245, 17.
- Cuomo, S., Di Cola, V.S., Giampaolo, F., Rozza, G., Raissi, M., Piccialli, F., 2022. Scientific machine learning through physics-informed neural networks: where we are and what's next. *J. Sci. Comput.* 92, 62.
- Han, Z.Z., Wang, H.R., Shen, C., Song, X.W., Cao, L.C., Yu, L.Q., 2024. Attention features selection oversampling technique (AFS-O) for rolling bearing fault diagnosis with class imbalance. *Meas. Sci. Technol.* 35, 15.
- Hao, D., Dong-bo, Z., Ze-Xu, C., Xiao-Rong, H.Y., Xiao-long, Z., Xi-yao, L., 2024. Optimization design of load-sharing and light weight of the twin rotors concentric face gear power-split transmission system based on gear tooth matching condition. *Proc. Inst. Mech. Eng. Part C-J. Eng. Mech. Eng. Sci.* 238, 6018–6041.
- Hou, Y.D., Ma, J.L., Wang, J.J., Li, T.Z., Chen, Z.Q., 2023. Enhanced generative adversarial networks for bearing imbalanced fault diagnosis of rotating machinery. *Appl. Intell.* 53, 25201–25215.
- Lei, Y.G., Kong, D.T., Lin, J., Zuo, M.J., 2012. Fault detection of planetary gearboxes using new diagnostic parameters. *Meas. Sci. Technol.* 23, 10.
- Lei, Y.G., Yang, B., Jiang, X.W., Jia, F., Li, N.P., Nandi, A.K., 2020. Applications of machine learning to machine fault diagnosis: a review and roadmap. *Mech. Syst. Signal Process.* 138, 39.
- Li, J., Liu, Y.B., Li, Q.J., 2022. Intelligent fault diagnosis of rolling bearings under imbalanced data conditions using attention-based deep learning method. *Measurement* 189, 15.
- Li, S., Peng, Y.F., Shen, Y.P., Zhao, S.B., Shao, H.D., Bin, G.F., Guo, Y., Yang, X.K., Fan, C., 2024. Rolling bearing fault diagnosis under data imbalance and variable speed based on adaptive clustering weighted oversampling. *Reliab. Eng. Syst. Saf.* 244, 20.
- Li, W.X., Shang, Z.W., Gao, M.S., Qian, S.Q., Zhang, B.R., Zhang, J., 2021. A novel deep autoencoder and hyperparametric adaptive learning for imbalance intelligent fault diagnosis of rotating machinery. *Eng. Appl. Artif. Intell.* 102, 19.
- Liu, Y.P., Jiang, H.K., Yao, R.H., Zhu, H.X., 2023. Interpretable data-augmented adversarial variational autoencoder with sequential attention for imbalanced fault diagnosis. *J. Manuf. Syst.* 71, 342–359.
- Lu, J.T., Wu, W., Huang, X., Yin, Q.T., Yang, K.Z., Li, S.M., 2024. A modified active learning intelligent fault diagnosis method for rolling bearings with unbalanced samples. *Adv. Eng. Inform.* 60, 11.
- Neupane, D., Seok, J., 2020. Bearing fault detection and diagnosis using case Western reserve university dataset with deep learning approaches: a review. *IEEE Access* 8, 93155–93178.
- Ni, Y.F., Li, S., Guo, P., 2025. Discrete wavelet integrated convolutional residual network for bearing fault diagnosis under noise and variable operating conditions. *Sci. Rep.* 15, 26.
- Pan, C.L., Shang, Z.W., Liu, F., Li, W.X., Gao, M.S., 2023. Optimization of rolling bearing dynamic model based on improved golden jackal optimization algorithm and sensitive feature fusion. *Mech. Syst. Signal Process.* 204, 19.
- Qin, Y., Liu, H.Y., Mao, Y.F., 2024a. Faulty rolling bearing digital twin model and its application in fault diagnosis with imbalanced samples. *Adv. Eng. Inform.* 61, 13.
- Qin, Y., Liu, H.Y., Wang, Y., Mao, Y.F., 2024b. Inverse physics-informed neural networks for digital twin-based bearing fault diagnosis under imbalanced samples. *Knowl. Base Syst.* 292, 13.
- Shang, Z.W., Wang, X.B., Pan, C.L., Cheng, H.C., Wang, Z.Y., 2025. Research on digital twin-assisted bearing fault diagnosis method based on virtual-real mapping. *Meas. Sci. Technol.* 36, 23.
- Sun, B., Hu, W.T., Wang, H., Wang, L., Deng, C.Y., 2025a. Remaining useful life prediction of rolling bearings based on CBAM-CNN-LSTM. *Sensors* 25, 18.
- Sun, B.J., Sheng, Z.X., Song, P., Sun, H.L., Wang, F., Sun, X.G., Liu, J.Y., 2025b. State-of-the-Art detection and diagnosis methods for rolling bearing defects: a comprehensive review. *Appl. Sci.-Basel* 15, 24.
- Wang, F.J., Zhai, Z., Zhao, Z.B., Di, Y., Chen, X.F., 2024. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nat. Commun.* 15, 12.
- Wang, T., Ong, Z.C., Khoo, S.Y., Siow, P.Y., Zhang, J.L., Wang, T., 2025. An enhanced generative adversarial network for longer vibration time data generation under variable operating conditions for imbalanced bearing fault diagnosis. *Eng. Appl. Artif. Intell.* 151, 22.
- Xin, Y.C., Zhu, J.N., Cai, M.Y., Zhao, P.Y., Zuo, Q.Z., 2025. Machine learning based mechanical fault diagnosis and detection methods: a systematic review. *Meas. Sci. Technol.* 36, 26.
- Yan, H.Y., Zhou, H.D., Zheng, J., Zhou, Z.Z., 2025. Rolling bearing fault diagnosis based on 1D convolutional neural network and kolmogorov-arnold network for industrial internet. *CMC-Comput. Mat. Contin.* 83, 4659–4677.
- Yang, J., Luo, Y.B., Huang, G.J., 2025. Rolling bearing fault diagnosis model based on external attention integrated convolutional neural network under imbalanced data conditions. *Meas. Sci. Technol.* 36, 13.
- Yi, X.K., Xu, Y.Y., Hu, Q., Krishnamoorthy, S., Li, W., Tang, Z.Z., 2022. ASN-SMOTE: a synthetic minority oversampling method with adaptive qualified synthesizer selection. *Complex Intell. Syst.* 8, 2247–2272.
- Yi, Y.H., Xia, C.Y., Feng, C., Zhang, W.J., Fu, C.L., Qian, L.Q., Chen, S.Q., 2023. Digital twin-long short-term memory (LSTM) neural network based real-time temperature prediction and degradation model analysis for lithium-ion battery. *J. Energy Storage* 64, 12.
- Yin, Z.Y., Zhang, F.Q., Yin, C., Xu, G.Y., Liu, S.Y., 2025. A bearing fault diagnosis method for sample imbalance. *Eng. Appl. Artif. Intell.* 157, 12.
- Zhang, F.Y., Liu, F.Z., Liu, M.H., Zhang, Y.L., Jiang, M.S., Sui, Q.M., 2025. Novel sparse auto-encoder framework with pseudo-labeled reinforcement for cross domain fault diagnosis with imbalanced samples. *Meas. Sci. Technol.* 36, 18.
- Zhang, S., Zhang, S.B., Wang, B.N., Habetler, T.G., 2020. Deep learning algorithms for bearing fault DiagnosticsxA comprehensive review. *IEEE Access* 8, 29857–29881.
- Zhang, T.C., Chen, J.L., Li, F.D., Zhang, K.Y., Lv, H.X., He, S.L., Xu, E.Y., 2022a. Intelligent fault diagnosis of machines with small & imbalanced data: a state-of-the-art review and possible extensions. *ISA Trans.* 119, 152–171.
- Zhang, W., Peng, G.L., Li, C.H., Chen, Y.H., Zhang, Z.J., 2017. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors* 17, 21.
- Zhang, X.Y., He, L., Wang, X.K., Wang, J.Q., Cheng, P.F., 2022b. Transfer fault diagnosis based on local maximum mean difference and K-means. *Comput. Ind. Eng.* 172, 8.

- Zhang, Y.C., Ji, J.C., Ren, Z.H., Ni, Q., Gu, F.S., Feng, K., Yu, K., Ge, J., Lei, Z.H., Liu, Z., 2023. Digital twin-driven partial domain adaptation network for intelligent fault diagnosis of rolling bearing. *Reliab. Eng. Syst. Saf.* 234, 14.
- Zhao, J.D., Wang, W.M., Huang, J., Ma, X.L., 2025. A comprehensive review of deep learning-based fault diagnosis approaches for rolling bearings: advancements and challenges. *AIP Adv.* 15, 17.
- Zhong, J.S., Zheng, Y., Ruan, C.T., Chen, L., Bao, X.Y., Lyu, L., 2025. M-IPISincNet: an explainable multi-source physics-informed neural network based on improved SincNet for rolling bearings fault diagnosis. *Inf. Fusion* 115, 16.
- Zhou, F.N., Yang, S., Fujita, H., Chen, D.M., Wen, C.L., 2020. Deep learning fault diagnosis method based on global optimization GAN for unbalanced data. *Knowl. Base Syst.* 187, 19.
- Zhu, T.F., Lin, Y.P., Liu, Y.H., 2017. Synthetic minority oversampling technique for multiclass imbalance problems. *Pattern Recogn.* 72, 327–340.
- Zhu, Y.H., Cheng, J.F., Liu, Z.F., Zou, X.F., Cheng, Q., Xu, H., Wang, Y., Tao, F., 2025. Data generation approach based on data model fusion: an application for rolling bearings fault diagnosis with small samples. *IEEE Trans. Instrum. Meas.* 74, 16.