

Physics-Guided GAN with Manhattan Attention: A Novel Approach for Imbalanced Bearing Vibration Fault Diagnosis

Lie Xu (徐冽), Daxiong Ji (冀大雄), Marcelo Ang, Yan Zhi Tan, Yuanchang Liu and Peng Wu

Abstract—This study introduces a novel Physics-Guided Generative Adversarial Network (PGAN) tailored explicitly for diagnosing rolling bearing faults under severely imbalanced datasets. PGAN integrates domain-specific physical constraints into the generative process, thereby enhancing both the physical realism and interpretability of the generated data. Key innovations include a generator conditioned on physics-based features and random noise, the introduction of dedicated physics-guided loss functions, and the incorporation of Manhattan Attention to improve the extraction of essential vibration features. Extensive experiments conducted using the CWRU and HUST datasets demonstrate PGAN’s superior performance compared to several state-of-the-art methods. Results indicate that PGAN significantly mitigates the impact of dataset imbalance, achieving robust diagnostic accuracy even in extreme conditions. Further visualization and comparative analyses underscore the method’s capability to deliver highly discriminative and interpretable features.

Index Terms—Fault diagnosis, imbalanced data, physics-guided, generative adversarial network, rolling bearing.

I. INTRODUCTION

FAULT diagnosis (FD) of rolling bearings is fundamental to ensuring the operational integrity, reliability, and efficiency of mechanical systems. Unexpected bearing failures can result in catastrophic consequences, including extended system downtime, substantial economic losses, and critical safety hazards [1], [2]. A significant challenge in this domain is the inherent imbalance in fault data collection—healthy conditions and certain fault types typically dominate datasets, while critical rare faults remain underrepresented. Although data-driven approaches, particularly deep learning (DL) models, have demonstrated remarkable advantages in bearing fault diagnosis through their ability to automatically extract hierarchical feature representations without extensive domain expertise [3], [4], these conventional DL techniques exhibit substantially diminished performance when confronted with limited and imbalanced fault datasets. This imbalance creates a pronounced bias toward majority classes while significantly reducing sensitivity to minority fault classes that often represent

the most critical failure modes [5], causing diagnostic systems trained on such imbalanced data to frequently misclassify rare but potentially catastrophic faults, thereby posing serious reliability concerns in industrial applications.

To address the data imbalance problem, Generative Adversarial Networks (GANs) have emerged as promising solutions by generating synthetic samples for minority fault classes [6]. Despite this advancement, conventional GAN architectures utilize random noise as inputs, producing synthetic samples that lack physical interpretability and realistic characteristics. This fundamental limitation significantly constrains their practical applicability and reliability in industrial scenarios [7]. Moreover, traditional GANs operate without leveraging domain-specific physical knowledge, resulting in generated samples that often deviate from realistic fault signatures and ultimately compromise the diagnostic system’s robustness, generalizability, and interpretability [8].

Several recent studies have explored innovative approaches to address the challenges in bearing fault diagnosis, each with specific strengths and limitations. A conditional GAN (BT-GAN) was developed in [13] for predicting stress fields in composite bolted joints, where superior performance was achieved compared to traditional finite element methods through multi-scale feature extraction and attention mechanisms. However, BT-GAN’s effectiveness is heavily contingent on training dataset completeness and quality, limiting its applicability in scenarios with sparse or imbalanced data. A dual-attention feature fusion network (DAFFN) was introduced in [14] that combines Wasserstein GAN with gradient penalty and K-means SMOTE to address imbalanced fault diagnosis in rotating machinery. While improved classification performance is demonstrated by DAFFN, its efficacy remains fundamentally dependent on the quality of hybrid data generation and the effectiveness of its dual-attention mechanism, presenting challenges in consistently handling severely imbalanced datasets. A hybrid fault data generation technique (MBAC-GAN) was proposed in [15] where physical bearing models are integrated with CycleGAN to bridge the simulation-to-reality distribution gap. Although the need for costly real-world fault experiments is reduced by MBAC-GAN, its performance is inherently limited by the simplified assumptions embedded in physical models and the inability to fully capture complex fault dynamics. A comprehensive review of hybrid physics-based and data-driven methodologies in smart manufacturing was conducted in [16], where the benefits of enhanced interpretability, reduced computational costs, and improved uncertainty management are highlighted. Additionally, To tackle incremental fault diagnosis and catastrophic forgetting in rotating machinery, an inverted transformer lifelong learning

This work was supported by the National Natural Science Foundation of China under Grant U24B20183, and the Pioneer Leading Goose + X Science and Technology Program of Zhejiang Province under Grant 2025C02018. The first author is supported by China Scholarship Council.

Lie Xu and Daxiong Ji are with the State Key Laboratory of Ocean Sensing & Ocean College, Zhejiang University, Zhoushan, Zhejiang, 316000, China (Corresponding author: Daxiong Ji. e-mail: xulie@zju.edu.cn; jidaxiong@zju.edu.cn)

Marcelo Ang and Yan Zhi Tan are with the Department of Mechanical Engineering, National University of Singapore, Singapore 119077, Singapore (e-mail: mpeangh@nus.edu.sg; mpetyz@nus.edu.sg)

Yuanchang Liu and Peng Wu are with the Department of Mechanical Engineering, University College London, WC1E 7JE London, U.K. (e-mail: yuanchang.liu@ucl.ac.uk; wu.peng@ucl.ac.uk).

method with a learnable pruning mechanism is proposed in [19]. The model embeds time-series signals from multiple sensors and dynamically adjusts its structure through learnable pruning, maintaining a balance between stability and plasticity as new faults are introduced. However, achieving an optimal balance between model accuracy and physical interpretability remains a persistent challenge across existing approaches.

These studies collectively reveal several critical limitations in current approaches:

- 1) Insufficient integration of domain-specific physical knowledge into generative models, leading to physically unrealistic synthetic samples;
- 2) Challenges in maintaining model interpretability while achieving high diagnostic accuracy;
- 3) Inadequate attention mechanisms for capturing fault-specific signal characteristics when dealing with severely imbalanced datasets;

Recently, Physics-informed Machine Learning (PIML) has garnered substantial attention for its capacity to seamlessly integrate prior physical knowledge into DL frameworks, thereby enhancing both performance and interpretability [9], [10]. The incorporation of physical constraints effectively guides neural networks toward capturing physically meaningful features and ensures that generated data remains consistent with known system behaviors and physical laws [11]. Additionally, specialized attention mechanisms, such as Manhattan Attention, have proven highly effective in selectively extracting critical diagnostic features from complex signals, substantially improving diagnostic accuracy and robustness, particularly in noisy industrial environments [12].

Motivated by these limitations and inspired by recent innovations, this paper proposes a novel Physics-Guided Generative Adversarial Network (PGAN), built upon the Wasserstein GAN with Gradient Penalty (WGAN-GP) [18] framework, specifically designed for rolling bearing fault diagnosis under imbalanced data conditions. Unlike traditional GAN architectures, PGAN leverages the stabilized training dynamics of WGAN-GP while integrating physics-informed principles directly into its architecture. The generator employs physical features as conditional inputs alongside random noise, explicitly incorporating domain knowledge to enhance the physical realism and interpretability of generated samples. To enforce adherence to physical constraints during training, physics-based loss functions are introduced, augmenting the adversarial loss with gradient penalty to ensure sample plausibility and improve model robustness. Furthermore, the Manhattan Attention mechanism is seamlessly integrated into the generator, enabling efficient capture of critical vibration features essential for accurate fault diagnosis across diverse operating conditions.

The primary contributions of this research are summarized as follows:

- 1) A novel Physics-Guided GAN is proposed that integrates domain-specific bearing knowledge into the generative process, producing physically realistic fault samples that significantly improve diagnostic accuracy for minority fault classes.

- 2) A comprehensive physics-based loss function framework is developed that enforces adherence to bearing dynamics principles, ensuring generated samples maintain physical consistency while addressing class imbalance challenges.
- 3) A specialized Manhattan Attention mechanism is designed within the generator architecture that effectively captures critical time-frequency features of bearing faults, enhancing feature extraction capabilities under noisy industrial conditions.

The remainder of this paper is structured as follows: Section II outlines the theoretical foundations of PGAN. Section III describes the detailed architecture and methods of the proposed PGAN framework. Section IV presents comprehensive experimental evaluations, verifying the effectiveness, interpretability, and robustness of the proposed method. Finally, Section V summarizes the findings and conclusions of this study.

II. PRELIMINARIES OF PGAN

A. Generative Adversarial Networks

Generative Adversarial Networks (GANs), proposed by Goodfellow et al. [17], comprise two neural networks: a *generator* G and a *discriminator* D , which are simultaneously trained through adversarial competition. The generator G transforms a random noise vector z , sampled from a prior distribution $p_z(z)$, into synthetic data samples $G(z)$ resembling real data, while the discriminator D evaluates both real data samples x , drawn from the true data distribution $p_{\text{data}}(x)$, and generated samples $G(z)$, outputting a probability $D(x)$ indicating the likelihood that the given input is real. The adversarial training process is governed by the following objective:

$$\min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

This formulation incentivizes the generator to produce samples indistinguishable from real data, while the discriminator aims to accurately differentiate between real and synthetic samples.

B. Wasserstein GAN with Gradient Penalty

Despite their success, standard GANs frequently encounter training instabilities and mode collapse issues. To mitigate these problems, Arjovsky et al. [7] introduced the Wasserstein GAN (WGAN), which employs the Wasserstein distance—also known as the Earth Mover's Distance—as a divergence metric between the real and generated distributions. The WGAN optimization problem is expressed as:

$$\min_G \max_{D \in \mathcal{D}} \mathbb{E}_{x \sim p_{\text{data}}(x)} [D(x)] - \mathbb{E}_{z \sim p_z(z)} [D(G(z))] \quad (2)$$

where $\mathcal{D}$ denotes the space of 1-Lipschitz continuous functions, and maintaining this Lipschitz continuity is crucial for the stable training of WGANs. To effectively enforce the Lipschitz constraint, Gulrajani et al. [18] introduced the

WGAN-GP, which modifies the discriminator's loss function by incorporating a gradient penalty term:

$$L = \mathbb{E}_{x \sim p_{\text{data}}} [D(x)] - \mathbb{E}_{z \sim p_z} [D(G(z))] + \lambda \mathbb{E}_{\hat{x} \sim p_{\hat{x}}} \left[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2 \right] \quad (3)$$

where $\hat{x}$ represents interpolations between real and generated data samples, and λ controls the strength of the gradient penalty. This additional term ensures that the discriminator's gradients remain close to unity, significantly enhancing training stability and convergence properties.

C. Physics-Informed Neural Networks

Physics-Informed Neural Networks (PINNs), proposed by Raissi et al. [8], integrate physical laws—typically described by partial differential equations (PDEs)—into the neural network's training process. By embedding these laws, PINNs ensure that predictions not only fit observed data but also conform to established physical principles. Considering a general nonlinear PDE of the form:

$$u_t + \mathcal{N}[u; \lambda] = 0, \quad x \in \Omega, \quad t \in [0, T] \quad (4)$$

where $u(t, x)$ is the PDE solution, $\mathcal{N}$ denotes a nonlinear operator parameterized by λ , and Ω is the spatial domain, PINNs define the PDE residual $f(t, x)$ as:

$$f(t, x) := u_t + \mathcal{N}[u; \lambda] \quad (5)$$

The training of PINNs involves minimizing a combined loss function:

$$L_{\text{total}} = L_{\text{data}} + L_{\text{phys}} \quad (6)$$

Here, L_{data} measures the difference between model predictions and empirical data, while L_{phys} penalizes deviations from the governing PDEs. Incorporating these constraints allows PINNs to produce physically consistent and robust predictions, particularly beneficial in scenarios with sparse observational data.

III. METHODOLOGY

A. Architecture Overview

The overview of the proposed PGAN architecture is presented in Fig. 1. It comprises two main components: the data generation module and the classifier module. In the data generation module, imbalanced real data samples collected from faulty bearings are provided as inputs to the discriminator. Unlike traditional GANs, the generator in PGAN incorporates physics-based conditions as inputs instead of the conventional noise vector. During training, the discriminator updates its network weights using a standard data-driven loss (data loss). In contrast, the generator optimization involves a composite loss function that includes both data loss and physics loss terms. This physics loss, computed using specific physical constraints elaborated in the subsequent subsection, guides the generator to produce physically consistent synthetic data.

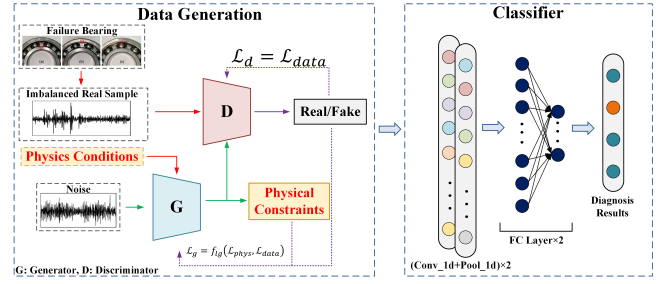


Fig. 1. Architecture of PGAN.

After training, the data generation module mitigates the imbalance by generating synthetic yet physically realistic data samples. Subsequently, the classifier module is trained on the balanced dataset produced by the data generation module, ultimately providing reliable diagnostic results.

B. Physics Guided Prior Knowledge

The detailed PGAN architecture comprises three main components: the generator, discriminator, and classifier. The generator and discriminator primarily employ fully-connected (FC) layers, as shown in Fig. 2. Whereas the classifier consists of convolutional (Conv) layers, max-pooling (Pool) layers, and FC layers. Additionally, the generator integrates Manhattan attention to extract intrinsic features effectively from each dimension.

Physics-Guided conditions for bearing fault data include fault frequencies specific to different fault types (inner race, outer race, and ball faults) and the corresponding mean amplitudes within fault frequency bands. The fault frequencies are computed as follows:

- Inner race fault frequency (f_{inner}):

$$f_{\text{inner}} = \frac{n}{2} f_r \left(1 + \frac{d}{D_{\text{pitch}}} \cos \alpha \right) \quad (7)$$

- Outer race fault frequency (f_{outer}):

$$f_{\text{outer}} = \frac{n}{2} f_r \left(1 - \frac{d}{D_{\text{pitch}}} \cos \alpha \right) \quad (8)$$

- Ball fault frequency (f_{ball}):

$$f_{\text{ball}} = \frac{D}{2d} f_r \left[1 - \left(\frac{d}{D_{\text{pitch}}} \cos \alpha \right)^2 \right] \quad (9)$$

Here, n is the number of rolling elements, f_r represents the rotational frequency, d denotes the rolling element diameter, D_{pitch} is the pitch diameter, and α is the contact angle of the bearing.

The mean amplitude within a specified fault frequency band is calculated by:

$$A_{\text{mean}} = \frac{1}{M} \sum_{f_i \in [f_{\text{fault}} - \frac{B}{2}, f_{\text{fault}} + \frac{B}{2}]} |X(f_i)| \quad (10)$$

where A_{mean} denotes the mean amplitude, $X(f_i)$ is the amplitude spectrum of the vibration signal at frequency f_i , f_{fault} is the targeted fault frequency (inner race, outer race,

or ball), B is the bandwidth around the fault frequency, and M represents the total number of frequency points within the selected band.

Unlike traditional GANs that rely solely on stochastic noise vectors as inputs, the input to our generator, denoted as x_{in} , consists of a fault-frequency-based modulation sine wave.

Let t denote time and let f_0 be the characteristic fault frequency (Hz) associated with the defect under a given operating condition. We encode the known periodic structure through a physics-conditioned carrier

$$c(t) = \sum_{k \in \mathcal{K}} w_k \sin(2\pi k f_0 t + \phi_k) \quad (11)$$

where $\mathcal{K}$ is the set of considered harmonic orders (e.g., $\{1, 2, 3\}$), $w_k \in \mathbb{R}_+$ are harmonic weights, and $\phi_k \in [-\pi, \pi]$ are phase offsets. In practice, $\{w_k, \phi_k\}$ are sampled from narrow priors to reflect modest operating variability (e.g., speed drift) while preserving the order structure.

Rather than synthesizing the signal directly from noise, the generator predicts physically meaningful terms relative to the carrier: (i) an amplitude envelope $a(t) \in \mathbb{R}$, (ii) a phase deviation $\Delta\phi(t) \in \mathbb{R}$, and (iii) a sparse event map $e(t) \in \mathbb{R}_{\geq 0}$ that excites structural resonances. The generated vibration is

$$\hat{x}(t) = (1 + a(t)) \left(t + \frac{\Delta\phi(t)}{2\pi f_0} \right) + \underbrace{(e * h_\theta)(t)}_{\text{impulse-driven resonance}} + n(t) \quad (12)$$

where $*$ denotes convolution in time, $(e * h_\theta)(t)$ is the resonant response obtained by filtering the event map $e(t)$ with a learned resonance kernel $h_\theta(t)$, and $n(t)$ is a small stochastic excitation accounting for unmodeled noise. The term $\Delta\phi(t)/(2\pi f_0)$ converts phase deviation (rad) into time (s), yielding an FM-style modulation of the carrier.

We parameterize the resonance kernel as a sum of exponentially decaying ring-downs:

$$h_\theta(t) = \sum_{m=1}^M b_m e^{-t/\tau_m} \sin(2\pi f_{r,m} t) u(t) \quad (13)$$

where M is the number of resonance modes; for each mode m , $b_m \in \mathbb{R}$ is the mode amplitude, $\tau_m > 0$ is the decay constant (s), $f_{r,m} > 0$ is the resonance frequency (Hz), and $u(t)$ is the unit step enforcing causality. To promote physically plausible, intermittent impacts, we regularize $e(t)$ with an ℓ_1 sparsity penalty (see next subsection).

This decomposition respects the classical impulse-excitation view of bearing faults (impacts that excite damped resonances), while the carrier $c(t)$ anchors periodicity at orders of f_0 . The generator's capacity is thereby focused on learning AM/FM modulation and impulse-driven resonant residuals, rather than rediscovering periodicity from scratch.

The carrier is not intended to reproduce the full fault mechanism; it provides a low-complexity basis that anchors the known periodicity (orders of f_0). The generator focuses on learning the physically salient variations—AM/FM, envelope evolution, and impulse-driven ring-downs—rather than discovering periodicity from scratch under extreme class imbalance. This inductive bias reduces the hypothesis space, stabilizes

adversarial training, and remains sufficiently expressive because envelope and phase deviations can represent narrow-band AM/FM processes in the characteristic bands, while $e * h_\theta$ models non-sinusoidal impulsive content.

By incorporating fault-frequency-based modulation sine waves as the generator input, the approach significantly enhances interpretability, as the generated signals are directly linked to specific physical fault mechanisms. This design also provides greater controllability, enabling precise manipulation of the generative process to simulate various fault scenarios through explicit adjustment of frequency components and related parameters. Furthermore, the integration of domain knowledge into the input facilitates domain-driven data augmentation, ensuring that the synthetic signals reflect the true nature of bearing faults rather than arbitrary patterns. Importantly, this strategy effectively addresses the challenge of imbalanced data by allowing the targeted generation of minority fault samples, thereby improving the balance and representativeness of training datasets. Collectively, these advancements lead to more transparent, robust, and practical diagnostic models for real-world bearing vibration analysis.

C. Physics-Consistent Losses

We regularize the generator with losses that (i) enforce envelope consistency in narrow harmonic bands around the known fault orders, (ii) match spectral content within the same bands, and (iii) promote sparsity of impulsive excitations. Throughout, $x(t)$ denotes the real (reference) vibration signal and $\hat{x}(t)$ the synthesized signal introduced previously; f_0 and the harmonic index k retain their earlier meanings (characteristic fault frequency and order, respectively).

a) *Band set around the fault orders.*: Let $\mathcal{B}$ denote the set of narrow frequency bands centered at the harmonic locations $\{k f_0 : k \in \mathcal{K}\}$. Concretely, each $B \in \mathcal{B}$ is an interval $[k f_0(1 - \delta), k f_0(1 + \delta)]$ with a small relative half-width $\delta > 0$ (e.g., 3%–6%), and $\mathcal{K}$ is the same order set used to construct the carrier.

b) *Envelope (analytic-signal) loss.*: Let $\mathcal{H}\{\cdot\}$ be the Hilbert transform and $|\cdot|_B$ the restriction (band-selection) operator that keeps only the frequency components inside B . We compare the analytic envelopes of x and $\hat{x}$ within the bands $\mathcal{B}$ using the ℓ_1 distance:

$$\mathcal{L}_{\text{env}} = \sum_{B \in \mathcal{B}} \left\| |\mathcal{H}\{x\}|_B - |\mathcal{H}\{\hat{x}\}|_B \right\|_1 \quad (14)$$

c) *Spectral (band-limited) loss.*: Let $P_x(f)$ and $P_{\hat{x}}(f)$ denote the (one-sided, consistently normalized) magnitude spectra of x and $\hat{x}$, respectively, obtained from the discrete Fourier transform with the same windowing and scaling. We penalize band-limited spectral mismatch as

$$\mathcal{L}_{\text{spec}} = \sum_{B \in \mathcal{B}} \left\| P_x(f)|_B - P_{\hat{x}}(f)|_B \right\|_1 \quad (15)$$

d) *Sparsity of impulsive events.*: As defined previously, $e(t)$ is the (nonnegative) event map that excites the resonance kernel. We encourage impulsiveness via an ℓ_1 penalty:

$$\mathcal{L}_{\text{sp}} = \|e\|_1 \quad (16)$$

e) *Total generator objective.*: Let $\mathcal{L}_{\text{GAN}}$ denote the adversarial term (e.g., WGAN-GP or hinge loss) and $\lambda_{\text{env}}, \lambda_{\text{spec}}, \lambda_{\text{sp}} > 0$ be scalar weights. The full generator loss is

$$\mathcal{L}_G = \mathcal{L}_{\text{GAN}} + \lambda_{\text{env}} \mathcal{L}_{\text{env}} + \lambda_{\text{spec}} \mathcal{L}_{\text{spec}} + \lambda_{\text{sp}} \mathcal{L}_{\text{sp}} \quad (17)$$

We use sinusoidal carriers as physics-conditioned basis functions: they inject the correct order structure a priori and let the generator devote capacity to envelope/phase modulation and impulsive residuals. Compared to random-noise inputs, carriers reduce search space and improve few-shot stability; compared to band-passed noise carriers, they provide phase-coherent references that facilitate learning of sideband spacing.

D. Detail Model Design

The processed inputs pass through two FC layers before further extraction via the Manhattan attention mechanism [12], defined as:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(-\frac{\|Q - K\|_1}{\sqrt{d_k}} \right) V \quad (18)$$

Here, Q , K , and V represent query, key, and value matrices, respectively, and d_k is the dimensionality of the key vectors. The Manhattan distance $\|Q - K\|_1$ emphasizes prominent features from the input data.

The discriminator subsequently distinguishes real from synthetic samples, enabling the generator to produce realistic, physically consistent data.

In the classifier component, one-dimensional convolutional (1-D Conv) layers are utilized to extract features from balanced datasets produced by the data generation module. Feature extraction using convolution is expressed as:

$$h_j^{(l)} = f \left(\sum_{i=1}^m h_i^{(l-1)} * W_{ij}^{(l)} + b_j^{(l)} \right) \quad (19)$$

where $h_j^{(l)}$ indicates the j -th feature map of the l -th convolutional layer, $h_i^{(l-1)}$ is the i -th input feature map from the previous layer, $W_{ij}^{(l)}$ denotes the convolution kernel connecting input and output feature maps, $b_j^{(l)}$ represents the bias term, m indicates the number of input feature maps, and $f(\cdot)$ is an activation function. In this study, ReLU is chosen as the activation function.

Each Conv layer is followed by a max-pooling layer to reduce dimensionality and computational complexity while retaining crucial features. The detailed model configurations are shown as Table I and II. In Table II, BN means 1d batch normalization, MP(2) means 1d Max Pooling with kernel size 2.

E. Training Procedure

The training of the proposed PGAN involves optimizing three distinct loss functions: the generator loss, discriminator loss, and classifier loss. The PGAN model adopts the WGAN-GP as its fundamental GAN framework.

TABLE I
GENERATOR AND CRITIC LAYER CONFIGURATIONS.

Net/Layer	Out Dim	Activation/Notes
Gen: FC1	128	ReLU; input: Physics guided wave
Gen: FC2	256	ReLU
Gen: ManhAttn	256	Softmax; param: 256-d
Gen: FC3	1024	–
Critic: FC1	256	LeakyReLU; input: 1024
Critic: FC2	128	LeakyReLU
Critic: FC3	1	–

TABLE II
CLASSIFIER CONFIGURATIONS.

Layer	Channels	Kernel	Activation	Notes
Conv1d	1→16	7, p=3	ReLU	BN, MP(2)
Conv1d	16→32	5, p=2	ReLU	BN, MP(2)
Conv1d	32→64	3, p=1	ReLU	BN, MP(2)
Flatten	–	–	–	–
FC1	8192→128	–	ReLU	Dropout(0.5)
FC2	128→4	–	–	–

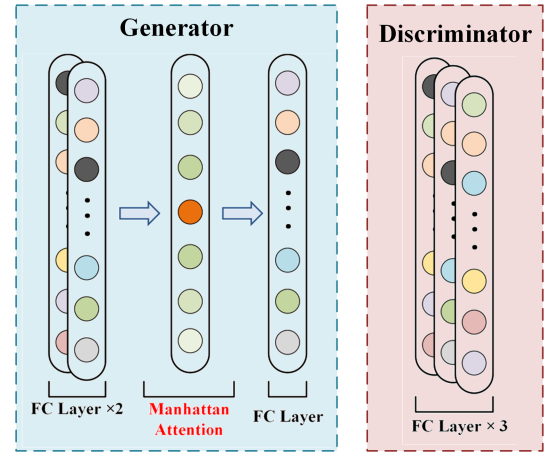


Fig. 2. Architecture of generator and discriminator in PGAN.

- 1) **Discriminator Loss:** The discriminator is trained to distinguish real from synthetic data accurately. Its loss function is purely data-driven, defined by the standard WGAN-GP objective:

$$\begin{aligned} \mathcal{L}_D = & \mathbb{E}_{z \sim p_z(z)} [D(G(z))] - \mathbb{E}_{x \sim p_{\text{data}}(x)} [D(x)] \\ & + \lambda_{\text{gp}} \mathbb{E}_{\hat{x} \sim p_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \end{aligned} \quad (20)$$

where $G(z)$ denotes the synthetic data generated from noise vector z , x represents real data samples, $\hat{x}$ is the interpolated sample between real and generated data, and λ_{gp} is the gradient penalty coefficient enforcing the Lipschitz constraint.

- 2) **Generator Loss:** The generator aims to produce data indistinguishable from real samples while adhering to physical constraints. The details about generator loss has been discussed in (17).
- 3) **Classifier Loss:** The classifier is optimized using balanced synthetic and real data from the generation mod-

ule. The classifier loss is purely data-driven, defined using the standard cross-entropy loss function:

$$L_C = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K y_{ik} \log(\hat{y}_{ik}) \quad (21)$$

where y_{ik} is the true label and $\hat{y}_{ik}$ denotes the predicted probability for class k , and K represents the number of classes.

Algorithm 1 provides detailed pseudo-code describing the training procedure of PGAN:

Algorithm 1 Training Procedure of PGAN

Require: Real imbalanced dataset X , physics conditions, batch size N , coefficients $w_{data}, w_{phys}, \lambda_{gp}$, epochs T

- 1: **for** epoch = 1 to T **do**
- 2: **for** each batch **do**
- 3: Sample real data $x \sim p_{data}(x)$
- 4: Sample physics guided wave $z \sim p_z(z)$
- 5: Generate synthetic data $\hat{x} = G(z)$
- 6: Update discriminator by minimizing L_D
- 7: Compute joint loss for generator L_G
- 8: Update generator by minimizing L_G
- 9: **end for**
- 10: **end for**
- 11: Generate balanced dataset using trained generator
- 12: **for** epoch = 1 to $T_{classifier}$ **do**
- 13: **for** each batch **do**
- 14: Sample balanced data
- 15: Update classifier by minimizing L_C
- 16: **end for**
- 17: **end for**
- 18: **return** Optimized Generator, Discriminator, Classifier

Through the outlined training procedure, PGAN effectively integrates data-driven and physics-guided learning, enhancing diagnostic performance in scenarios with imbalanced datasets.

IV. EXPERIMENTAL VALIDATION

In this section, two public real failure bearing datasets are used to validate the effectiveness of the proposed method.

A. Dataset Description

In this study, two benchmark datasets, the Case Western Reserve University (CWRU) dataset [20] and the Hanoi University of Science and Technology (HUST) dataset [21], are employed to evaluate the proposed PGAN method.

CWRU Dataset: The CWRU dataset is widely recognized for bearing fault diagnosis and was collected by the Bearing Data Center at Case Western Reserve University. It comprises vibration data captured at four rotational speeds (1797, 1772, 1750, and 1730 r/min) and covers faults of varying diameters (7, 14, and 21 mils) occurring at different bearing positions, including ball, inner race (IR), and outer race (OR). For this study, the dataset has been categorized into four distinct classes: normal, ball fault, inner race fault, and outer race fault.

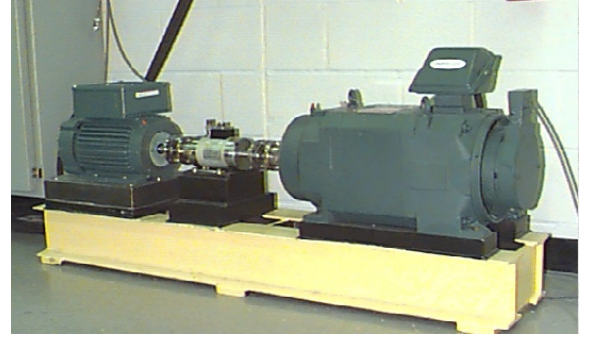


Fig. 3. CWRU dataset acquisition bench.

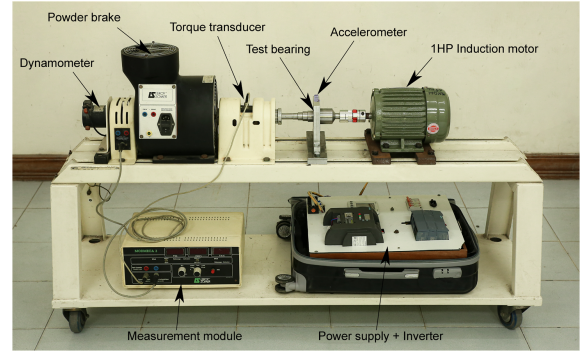


Fig. 4. HUST dataset acquisition bench.

Table III details the dataset used in this experiment, while the test bench used for data acquisition is illustrated in Fig. 3.

HUST Dataset: The HUST dataset is a recently developed dataset for bearing fault diagnosis, gathered by the School of Electrical and Electronic Engineering at Hanoi University of Science and Technology. It consists of vibration data from various ball bearing models (ID 6204, 6205, 6206, 6207, and 6208 from KG Bearing India) featuring typical single faults: inner race fault, outer race fault, and ball fault. Detailed specifications are shown in Table IV, and the corresponding experimental setup is presented in Fig. 4.

TABLE III
DETAILED INFORMATION OF DATASET CWRU

Fault type	Label	Fault diameter (mil)	Fault position	Load (HP)	Speed (r/min)
Normal	0	-	-	1	1797
Inner race	1	7/14/21	-	1	1797
Ball	2	7/14/21	-	1	1797
Outer race	3	7/14/21	6 o'clock	1	1797

TABLE IV
DETAILED INFORMATION OF DATASET HUST

Fault type	Label	Fault diameter (mm)	Fault position	Load (HP)	Speed (r/min)
Normal	0	-	-	1	1496
Inner race	1	0.2	-	1	1492
Ball	2	0.2	-	1	1496
Outer race	3	0.2	-	1	1492

B. Experimental Settings

To simulate realistic imbalanced conditions, data pre-processing is applied. Specifically, half of the available normal data samples are included in the training dataset, whereas fault data are intentionally underrepresented according to various predefined proportions (1:4, 1:10, 1:20, 1:25, 1:50, and 1:100 relative to normal data). Remaining samples from both categories are reserved exclusively for evaluating the classifier performance.

Data segmentation is performed using a window length of 1024 samples, resulting in 500 segments for normal data. Fault data are segmented proportionally based on the established imbalance proportions. The testing dataset includes 500 segments for each class, providing a comprehensive evaluation framework consisting of 2000 test samples (4 classes $\times$ 500 segments per class).

Model hyperparameters are optimized using grid search based on validation set performance, aiming to maximize diagnostic accuracy. Optimal parameters and their search spaces are presented in Table V. Across datasets, the response surfaces were smooth around the recommended region, indicating low hyperparameter brittleness. Accuracy rose when moving from negligible envelope consistency to moderate levels and then plateaued near $\lambda_{\text{env}} \approx 1.0$, while excessively large values conferred no consistent benefit and could slightly depress overall accuracy. Increasing the spectral band weight improved minority-class up to about $\lambda_{\text{spec}} \approx 0.3$, beyond which gains diminished and occasional trade-offs in macro-F1 emerged. For the sparsity prior, $\lambda_{\text{sp}} = 10^{-3}$ reliably suppressed spurious textures without attenuating salient transients; stronger regularization tended to over-constrain the generator. Regarding WGAN-GP stability, five critic steps with a gradient penalty of $\lambda_{\text{gp}} = 10$ achieved a favorable balance between stable adversarial dynamics and computational cost, whereas fewer penalties increased volatility and more critic steps yielded negligible accuracy improvements. Expanding the carrier from a single harmonic to a tri-harmonic set (1, 2, 3) enhanced separability of fault signatures—particularly for rolling-element and inner-race classes—without the noise amplification typically associated with higher-order expansions. Additionally, each experiment is conducted ten times to mitigate random initialization effects, and the average performance is reported.

C. Results

The diagnostic performance of the proposed PGAN under different data imbalance conditions is summarized in Table VI and Table VII. The analysis of the results demonstrates that model performance is significantly impacted by class imbalance. Both datasets initially exhibit high accuracy, precision, and recall at mild imbalance levels (1:4, 1:10), with metrics consistently above 99%. However, as the imbalance intensifies, performance notably deteriorates, with the HUST dataset experiencing a sharper decline than CWRU, reflecting lower robustness. Particularly at extreme imbalance (1:100), accuracies drop to 81.12% for CWRU and 74.35% for HUST, accompanied by higher variability.

TABLE V
SEARCHING SPACE AND OPTIMAL HYPERPARAMETERS OF THE PGAN

Hyperparameters	Searching space	Optimal
Input size	1×1024 (fixed)	
Batch size	32/64/128	64
Max epochs	100 (fixed)	
Optimizer	Adam (fixed)	
Learning rate	$1e-3/1e-4/1e-5$	$1e-4$
Number of hidden nodes in generator	$[32, 64]/[64, 128]/[128, 256]$	[128, 256]
Number of convolution kernels in classifier	$[32, 64]/[64, 128]/[128, 256]$	[128, 256]
Envelope loss weight	0.3/1.0/3.0	1.0
Spectral loss weight	0.1/0.3/1.0	0.3
Sparsity of impulsive events	$1e-4/1e-3/1e-2$	$1e-3$
Gradient penalty	1/5/10	10
Harmonic orders	$[1]/[1, 2]/[1, 2, 3]$	[1, 2, 3]

Confusion matrices in Fig. 5 and Fig. 6 further illustrate classification results across various imbalance scenarios. Normal state samples achieve consistently high classification accuracy (approximately 100%) due to adequate training data availability. Misclassification predominantly occurs between inner and outer race fault conditions, attributed primarily to the similarity in fault-induced vibration signatures and insufficient discriminative physical prior knowledge.

TABLE VI
PERFORMANCE METRICS (MEAN/STD) FOR CWRU UNDER DIFFERENT PROPORTIONS

Proportion	Accuracy		Precision		Recall	
	Mean	Std	Mean	Std	Mean	Std
1:4	99.99%	0.01	99.99%	0.01	99.99%	0.02
1:10	99.78%	0.59	99.66%	0.91	99.87%	0.34
1:20	99.21%	2.19	99.49%	1.26	98.73%	3.69
1:25	96.84%	4.89	96.81%	4.58	97.12%	4.54
1:50	92.71%	3.37	92.12%	3.94	91.87%	4.27
1:100	81.12%	5.61	76.69%	13.02	78.29%	10.82

TABLE VII
PERFORMANCE METRICS (MEAN/STD) FOR HUST UNDER DIFFERENT PROPORTIONS

Proportion	Accuracy		Precision		Recall	
	Mean	Std	Mean	Std	Mean	Std
1:4	99.97%	0.07	99.96%	0.07	99.96%	0.07
1:10	99.17%	1.15	99.22%	1.04	99.17%	1.15
1:20	88.30%	2.32	89.32%	1.94	88.30%	2.32
1:25	88.09%	5.85	88.35%	8.13	88.09%	7.77
1:50	79.14%	6.23	82.32%	9.04	79.14%	6.23
1:100	74.35%	6.36	81.43%	9.64	74.35%	6.36

These findings underscore PGAN's ability to effectively address imbalance challenges in fault diagnosis tasks, reinforcing the advantages of integrating physics-informed learning into data-driven frameworks.

D. Comparisons and Discussions

To comprehensively evaluate the effectiveness of the proposed PGAN, comparisons are performed with several state-of-the-art methods, including CMDGAN [22], CDCGAN [23],

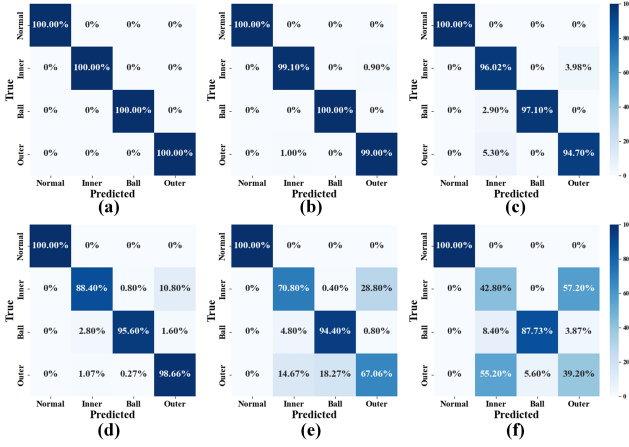


Fig. 5. Confusion martics on CWRU with different proportions. (a)1:4 (b)1:10 (c)1:20 (d)1:25 (e)1:50 (f)1:100

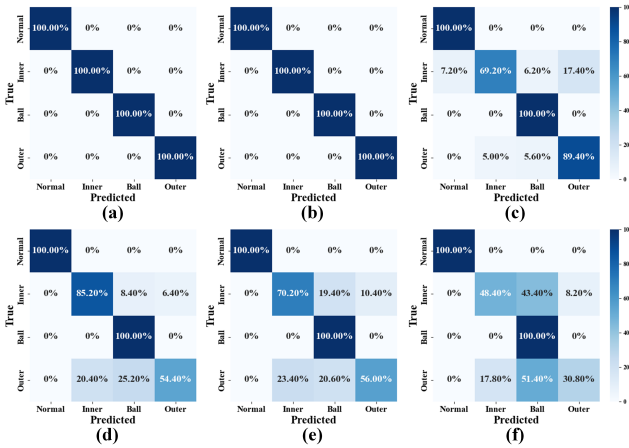


Fig. 6. Confusion martics on HUST with different proportions. (a)1:4 (b)1:10 (c)1:20 (d)1:25 (e)1:50 (f)1:100

DeepSMOTE [24], and DCWDN [25]. Among these methods, CMDGAN and CDCGAN are GAN-based techniques designed for handling data imbalance issues; DeepSMOTE leverages deep learning-enhanced SMOTE for data augmentation, and DCWDN utilizes an AutoEncoder-based architecture. Additionally, classical GAN-based methods such as standard GAN, Conditional GAN (CGAN), and WGAN-GP are included to validate the improvements brought by PGAN.

For fair and consistent comparisons, vibration signals for all methods undergo identical preprocessing steps. The mean diagnostic accuracy of 10 times test for varying data imbalance proportions on both CWRU and HUST datasets is reported in Tables VIII and IX, respectively.

From Table VIII, it is evident that the proposed PGAN consistently outperforms the comparison methods across all imbalance proportions on the CWRU dataset. Specifically, at moderate levels of imbalance (proportions of 1:4 to 1:25), PGAN achieves superior diagnostic accuracy, clearly benefiting from the integration of physics-informed constraints. Under extremely severe imbalance conditions (proportions of

1:50 and 1:100), PGAN maintains relatively high accuracy (92.71% and 81.12%, respectively), significantly exceeding other approaches. The DCWDN method, although competitive at lower imbalance proportions (1:4 and 1:10), exhibits substantial accuracy degradation at higher imbalance levels (1:20 and beyond). Other GAN-based methods and DeepSMOTE display consistently lower accuracy, highlighting limitations in effectively addressing extreme imbalance situations.

Similarly, analysis of results from the HUST dataset in Table IX reveals comparable trends. PGAN achieves excellent accuracy at mild imbalance levels (1:4 and 1:10), clearly outperforming all comparative methods. As imbalance severity increases, PGAN maintains strong performance, although it is slightly surpassed by DCWDN at the proportion of 1:50 (79.15% vs. 80.96%). However, PGAN outperforms all methods, including DCWDN, under the most challenging condition (1:100), achieving an accuracy of 74.35%. Other comparative methods experience significant performance drops at higher imbalance levels, further reinforcing the robustness and adaptability of the physics-informed constraints integrated within PGAN.

Fig. 7 illustrates the ablation study results for the proposed method and its variants across varying degrees of class imbalance (normal:fault ratios of 1:4, 1:10, 1:20, 1:25, 1:50, and 1:100). Each subplot in the figure presents the test accuracy of four models:

- **Proposed method (blue):** The full model incorporating both physics-guided and attention mechanisms.
- **Without physics guided (orange):** The model excluding the physics-guided module.
- **Without attention (green):** The model without the attention mechanism.
- **Vanilla WGAN-GP (red):** The baseline, which lacks both the physics-guided and attention modules.

As the class imbalance becomes more severe, from 1:4 to 1:100, all models demonstrate a decline in accuracy. However, the magnitude of this decline varies significantly across models.

Across all imbalance settings, the proposed method (blue bar) consistently achieves the highest accuracy. This demonstrates the effectiveness of jointly leveraging physics-guided constraints and attention mechanisms, particularly for challenging, highly imbalanced datasets.

The removal of the physics-guided module causes a noticeable reduction in accuracy at all imbalance ratios, with the performance gap widening as the data becomes more imbalanced. Omitting the attention mechanism also results in accuracy degradation, though the drop is generally less severe than when physics guidance is removed. The vanilla WGAN-GP (red bar), lacking both modules, consistently yields the lowest accuracy, especially in extreme imbalance cases (1:50, 1:100).

For attention mechanism ablations, as shown in Fig. 8, across all imbalance settings, Manhattan attention attains the highest accuracy, with SE attention second and vanilla self-attention slightly worse. The margin is small for mild imbalance (1:4–1:20) and widens modestly as imbalance increases

TABLE VIII
COMPARISONS ON CWRU WITH DIFFERENT PROPORTION (10-RUN AVERAGE ACCURACY)

Proportion	Proposed method	GAN	CGAN	WGAN-GP	CMDGAN	CDCGAN	DeepSMOTE	DCWDN
1:4	99.99%	99.91%	99.31%	98.69%	97.82%	98.96%	97.57%	99.60%
1:10	99.78%	95.05%	91.48%	89.62%	91.31%	93.95%	89.53%	99.06%
1:20	99.21%	89.13%	83.88%	83.13%	85.00%	86.20%	84.30%	85.30%
1:25	96.84%	86.54%	80.46%	79.23%	84.70%	82.68%	81.19%	76.34%
1:50	92.71%	81.07%	75.22%	72.57%	74.55%	78.17%	76.81%	72.32%
1:100	81.12%	77.16%	66.26%	70.19%	59.97%	64.60%	67.72%	70.47%

TABLE IX
COMPARISONS ON HUST WITH DIFFERENT PROPORTION (10-RUN AVERAGE ACCURACY)

Proportion	Proposed method	GAN	CGAN	WGAN-GP	CMDGAN	CDCGAN	DeepSMOTE	DCWDN
1:4	99.97%	93.31%	93.87%	93.01%	95.82%	95.08%	91.63%	94.73%
1:10	99.17%	80.80%	73.41%	72.16%	82.92%	82.62%	79.18%	91.07%
1:20	88.30%	65.22%	62.85%	62.42%	75.75%	71.64%	72.78%	88.33%
1:25	88.09%	63.47%	60.58%	58.06%	72.74%	68.20%	69.10%	86.39%
1:50	79.15%	57.72%	56.12%	55.70%	55.90%	64.87%	57.26%	80.96%
1:100	74.35%	57.27%	50.39%	46.16%	43.25%	54.96%	36.26%	68.35%

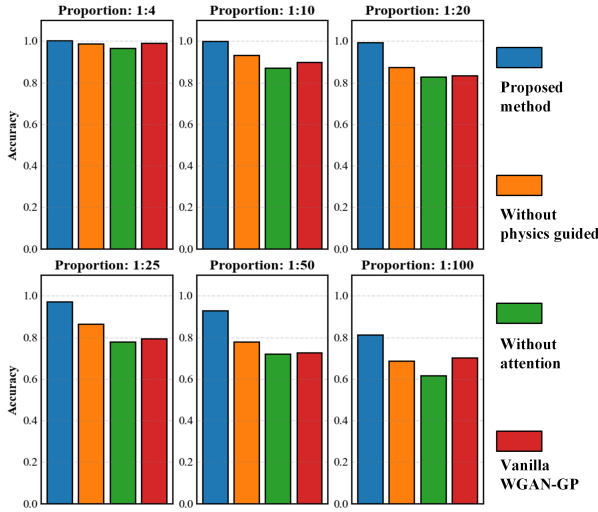


Fig. 7. Ablation study results for the proposed method and its variants under different data imbalance ratios.

(1:25–1:100), indicating that Manhattan attention confers robustness when minority data are scarce. We attribute this advantage to an inductive bias that aligns with bearing-fault physics: Manhattan attention imposes an explicit L1 distance decay (and, in our implementation, cycle-aware biasing), privileging dependencies within and across neighboring periods. This stabilizes training, preserves AM/FM structure and side-band spacing, and suppresses spurious long-range couplings that vanilla self-attention can learn under limited data; SE attention, being channel-wise recalibration, lacks temporal/phase selectivity. The advantage remains modest because the overall framework already embeds strong physics priors (carrier construction plus band-limited envelope/spectral losses), the data exhibit relatively signatures, and evaluation uses accuracy with ceiling effects, collectively reducing separability among attention variants.

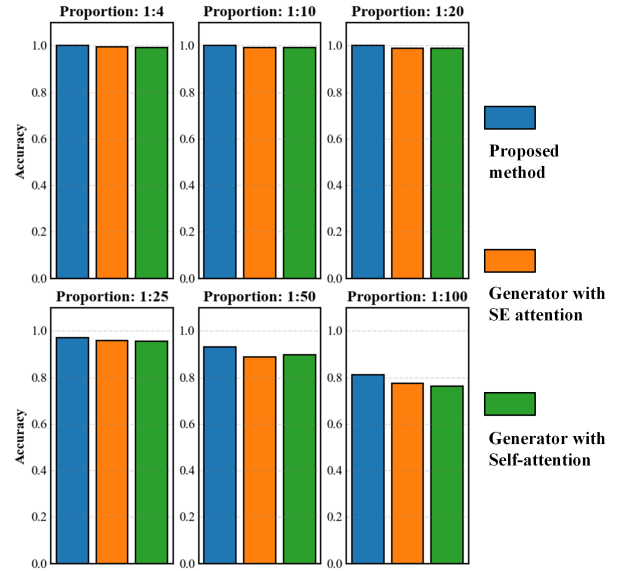


Fig. 8. Ablation study results for attention mechanisms under different data imbalance ratios.

E. Comprehensive Diagnostic-Sensitive Evaluation

To deeply evaluate the diagnostic quality of generated vibration signals, we propose a diagnostic-sensitive, harmonically weighted evaluation metric that quantitatively assesses the similarity between generated and real signals at all harmonic frequencies of the characteristic fault frequency within 0–4000 Hz.

Specifically, for a given characteristic fault frequency f_0 , we extract its N harmonics $\{f_n = nf_0 \mid f_n \leq 6000, n = 1, 2, \dots, N\}$. At each harmonic frequency f_n , we compute several normalized, dimensionless diagnostic-sensitive features for both the real and generated signals: (1) the energy ratio in the frequency band $[f_n - \Delta f, f_n + \Delta f]$ (where Δf is typically set to 30 Hz), (2) the ratio of the main peak to the mean in the

same band, (3) the kurtosis, and (4) the impulse factor after bandpass filtering and normalization. For each feature k , the absolute difference between the real and generated signals is calculated at each harmonic.

To reflect the actual diagnostic importance of each harmonic, we use the energy distribution of the real signal in each harmonic band as a weight w_n , defined as

$$w_n = \frac{E_n}{\sum_{i=1}^N E_i}, \quad (22)$$

where E_n is the envelope spectrum energy of the real signal in the band $[f_n - \Delta f, f_n + \Delta f]$. The final comprehensive diagnostic-sensitive evaluation score S is then calculated as

$$S = \sum_{n=1}^N w_n \cdot \sum_{k=1}^K |F_{n,k}^{(\text{real})} - F_{n,k}^{(\text{gen})}|, \quad (23)$$

where $F_{n,k}^{(\text{real})}$ and $F_{n,k}^{(\text{gen})}$ denote the k -th normalized feature (energy ratio, peak-to-mean, kurtosis, impulse factor) of the real and generated signals at the n -th harmonic, respectively.

Table X reports the comprehensive diagnostic-sensitive evaluation scores for different generative models.

TABLE X
COMPREHENSIVE DIAGNOSTIC-SENSITIVE EVALUATION SCORE FOR
DIFFERENT METHODS (LOWER IS BETTER)

Method	Score
Proposed	2.992
w/o Attention	3.698
w/o Physics-Guided	4.067
WGAN-GP (Baseline)	8.241

From Table X, it can be seen that our proposed physics-guided and attention-augmented GAN achieves the lowest comprehensive evaluation score, indicating that the generated signals most closely match the real signals in all diagnostically relevant harmonic bands. The ablation studies show that removing the attention mechanism or the physics-guided constraint both lead to increased scores, reflecting a reduction in diagnostic relevance and feature fidelity. The baseline WGAN-GP model yields the highest score, suggesting that without any diagnostic guidance, the generated signals fail to accurately reconstruct key fault-related characteristics, especially in the principal and lower-order harmonics that are most critical for fault identification. These results validate the effectiveness of our method in generating diagnostically sensitive and physically consistent vibration signals.

V. CONCLUSIONS

This paper introduced a novel PGAN specifically designed for rolling bearing fault diagnosis under severe data imbalance conditions. The proposed PGAN significantly advances beyond conventional GAN models by explicitly integrating domain-specific physical knowledge into the generative process. This approach not only enhances the physical realism and interpretability of synthesized data but also substantially mitigates the detrimental effects of dataset imbalance. Experimental validations performed on the CWRU and HUST

bearing datasets confirmed the superior performance of the proposed PGAN compared to state-of-the-art methods. Results indicated that PGAN consistently maintained high diagnostic accuracy, even under extreme imbalance conditions (up to a 1:100 imbalance proportion), demonstrating significant robustness and adaptability. Furthermore, feature visualization and comparative analyses emphasized the enhanced discriminative capabilities and interpretability provided by the physics-informed constraints and Manhattan Attention. Future research directions may explore further optimization of physics-based constraints, extension of the method to other industrial fault scenarios, and integration with real-time diagnostic platforms for broader practical applicability.

REFERENCES

- [1] Randall R B., *Vibration-based condition monitoring: industrial, automotive and aerospace applications*, John Wiley & Sons, 2021.
- [2] Smith W A, Randall R B, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mechanical systems and signal processing*, vol.64, pp. 100-131, 2015.
- [3] Haidong S, Hongkai J, Xingqiu L, et al. "Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine. *Knowledge-Based Systems*," vol.140, pp. 1-14, 2018.
- [4] H. Zhang, H. Guo, J. Shang and K. Peng, "Spatio—Temporal Generative Adversarial Network for Industrial Fault Diagnosis with Imbalanced Data," *IEEE Transactions on Instrumentation and Measurement*, 2025.
- [5] Z. Li, T. Zheng, Y. Wang, Z. Cao, Z. Guo and H. Fu, "A Novel Method for Imbalanced Fault Diagnosis of Rotating Machinery Based on Generative Adversarial Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-17, 2021.
- [6] G. Bai, W. Sun, C. Cao, D. Wang, Q. Sun and L. Sun, "GAN-Based Bearing Fault Diagnosis Method for Short and Imbalanced Vibration Signal," *IEEE Sensors Journal*, vol. 24, no. 2, pp. 1894-1904, 2024.
- [7] Arjovsky M, Chintala S, Bottou L., "Wasserstein generative adversarial networks," in *Proceedings of International conference on machine learning (PMLR)*, pp. 214-223, 2017.
- [8] Raissi M, Perdikaris P, Karniadakis G E., "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational physics*, vol. 378, 686-707, 2019.
- [9] Karniadakis G E, Kevrekidis I G, Lu L, et al. "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422-440, 2021.
- [10] Gao S, Liu J, Zhang Z, et al, "Physics-Guided Generative Adversarial Networks for fault detection of underwater thruster," *Ocean Engineering*, vol. 286: 115585, 2023.
- [11] S. Li, X. Lin, H. Shi, Y. Shi and K. Zhu, "Physics-Guided Deep Learning Method for Tool Condition Monitoring in Smart Machining System," *IEEE/ASME Transactions on Mechatronics*, vol. 29, no. 3, pp. 2327-2337, 2024.
- [12] Q. Fan, H. Huang, M. Chen, H. Liu and R. He, "RMT: Retentive Networks Meet Vision Transformers," 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5641-5651, 2024.
- [13] Zhao Y, Liu Y, Lin Q, et al. "A conditional generative model for end-to-end stress field prediction of composite bolted joints," *Engineering Applications of Artificial Intelligence*, vol. 134: 108692, 2024.
- [14] Wang C, Wang H, Liu M. "A dual-attention feature fusion network for imbalanced fault diagnosis with two-stream hybrid generated data," *Journal of Intelligent Manufacturing*, vol. 35, no. 4, pp. 1707-1719, 2024.
- [15] Y. Yu, L. Tang, Z. Liu and J. Xiang, "A Novel Bearing Fault Data Generation Strategy Combining Physical Modeling and CycleGAN Variant for Fault Diagnosis Without Real Samples," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1-17, 2025.
- [16] Wang J, Li Y, Gao R X, et al. "Hybrid physics-based and data-driven models for smart manufacturing: Modelling, simulation, and explainability," *Journal of Manufacturing Systems*, vol. 63, pp. 381-391, 2022.
- [17] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27*, Montreal, Quebec, Canada, pp. 2672-2680, 2014.
- [18] Gulrajani I, Ahmed F, Arjovsky M, et al. "Improved training of wasserstein gans," *Advances in neural information processing systems*, vol. 30, 2017.

- [19] Hai-hong T, Jia-wei L, Wu-wei F, et al. "Towards machinery incremental fault diagnosis based on inverted transformer lifelong learning with learnable pruning mechanism." *Engineering Applications of Artificial Intelligence*, vol. 152 (2025): 110763.
- [20] Smith W A, Randall R B. "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study". *Mechanical systems and signal processing*, vol. 64, pp. 100-131, 2015.
- [21] Thuan N D, Hong H S. "HUST bearing: a practical dataset for ball bearing fault diagnosis". *BMC research notes*, vol. 16, no. 1, pp. 138, 2023.
- [22] X. Liu, W. Sun, H. Li, Z. Wang and Q. Li, "Imbalanced Sample Fault Diagnosis of Rolling Bearing Using Deep Condition Multidomain Generative Adversarial Network," *IEEE Sensors Journal*, vol. 23, no. 2, pp. 1271-1285, 2023.
- [23] Luo J, Huang J, Li H. "A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis," *Journal of Intelligent Manufacturing*, vol. 32, no. 2, pp. 407-425, 2021.
- [24] D. Dablain, B. Krawczyk and N. V. Chawla, "DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 9, pp. 6390-6404, 2023.
- [25] R. Li et al., "Deep Complex Wavelet Denoising Network for Interpretable Fault Diagnosis of Industrial Robots With Noise Interference and Imbalanced Data," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1-11, 2025.
- [26] Van der Maaten L, Hinton G. "Visualizing data using t-SNE". *Journal of machine learning research*, vol. 9, no. 11, 2008.