Sound & Vibration 2025, 59(5), 3658.
doi: https://doi.org/10.59400/sv3658

Article

Intelligent fault detection of zero-sample rolling bearings driven by combined time-frequency analysis and multimodal knowledge

Haifeng Wei[1], Huijian Que[1], Rongxiang Zheng[1], Jianbin Liao[2], Guoqiang Li[2], Jin Yan[2,*]

[1] Longyan Tobacco Industrial Co., Ltd., Longyan 364021, China
[2] School of Marine Engineering, Jimei University, Xiamen 361021, China
*** Corresponding author:** Jin Yan, jasonyan@jmu.edu.cn

Abstract: To meet the demand for intelligent monitoring of rolling bearings and overcome the limitation of scarce fault samples in the real world, this paper proposes a zero-fault sample-based condition detection method that integrates the time-frequency analysis of rolling bearings with modal knowledge, which mainly includes: 1) The HHT envelope analysis and high-frequency filtering is introduced to reduce the interference of the base-frequency information of rolling bearings and to enhance the frequency of the fault information. 2) A novel zero-fault sample-based loss function is designed by combining the strong temporal sequence of rolling bearing monitoring data with the a priori knowledge of information mutualism to realize the effective training of the data-driven model. 3) An intelligent fault detection algorithm for rolling bearings is established based on a trained data-driven model. The proposed method is validated using a constructed rolling bearing experimental platform. The validation results show that the proposed method is able to build an effective intelligent fault detection model with zero fault samples of rolling bearings, which shows better fault detection performance than other supervised and unsupervised learning-based methods. The proposed anomaly detection method based on zero-fault samples can quickly establish its state detection model for important rolling bearings applied in engineering practice, providing a new perspective for data-driven fault diagnosis methods for rolling bearings.

Keywords: gearboxes; zero fault sample; prior knowledge; compound fault; contrastive loss function

CITATION

Wei H, Que H, Zheng R, et al. Intelligent fault detection of zero-sample rolling bearings driven by combined time-frequency analysis and multimodal knowledge. Sound & Vibration. 2025; 59(5): 3658. https://doi.org/10.59400/sv3658

ARTICLE INFO

Received: 22 August 2025
Revised: 16 September 2025
Accepted: 26 September 2025
Available online: 30 September 2025

1. Introduction

As a core functional component of modern mechanical equipment, the dynamic characteristics of rolling bearings directly determine the reliability, efficiency and service life of the system [1–3]. From the contact mechanics analysis, the Hertzian contact stress distribution of the bearing and the microscopic crack extension mechanism together constitute its failure threshold, while studies have shown that about 40% of rotating machinery failures originate from bearing failures (e.g., fatigue spalling, wear and tear, etc.), which lead to significant economic losses [4,5].

In recent years, with the deep development of intelligent manufacturing and industrial Internet of Things technology, intelligent fault diagnosis methods have become one of the cores means to improve the health management level and production efficiency of equipment [6]. Data-driven diagnosis methods based on deep learning and other artificial intelligence technologies have become a research hotspot in the field of rolling bearing fault detection due to their excellent feature extraction and fault identification capabilities. A large number of studies have shown that deep convolutional neural networks (CNN) [7], recurrent neural networks (RNN) [8], graphical neural networks (GNN) [9], and hybrid structural models can efficiently and automatically learn complex spatial and temporal features from a large number of labelled monitoring signals, and achieve the accurate identification of a wide range of typical faults. However, fault data of rolling bearings in practical engineering applications are extremely scarce, especially in new equipment, critical parts or early stages of service, and often face the dilemma of zero or very few samples [10, 11]. Traditional deep learning methods relying on large-scale labelled data are thus greatly restricted, and the models are difficult to generalize to real industrial environments, which seriously affects the practicality and reliability of intelligent diagnosis systems. Therefore, how to carry out intelligent fault diagnosis of rolling bearings with high efficiency and strong generalization ability under the condition of fewer fault-free samples or even zero samples has become a key scientific challenge that needs to be broken through in the field of intelligent maintenance and health management [12,13].

Aiming at the practical difficulties of rolling bearings with very few or no-fault samples, scholars have proposed a variety of innovative solutions based on migration learning and generated data. In the research of intelligent fault diagnosis of rolling bearings based on migration learning, the main idea is to migrate the acquired knowledge from the source domain (a large amount of labeled data in a laboratory environment, under simulated working conditions, or in a publicly available dataset) to the target domain (a task with a lack of fault samples in a real-world industrial scenario, under different working conditions, or with a new type of equipment) [14,15]. Current mainstream methods include feature space alignment, domain adaptation, parameter fine-tuning, domain adversarial network, etc., aiming to alleviate the problem of "domain drift" caused by different acquisition conditions, working condition changes and sample distribution differences. Through migration learning, the model can make full use of the existing health and fault knowledge to effectively identify the rolling bearing condition in new equipment or under new working conditions. At the same time, the method of generating data has also become an important means to improve rolling bearing fault diagnosis capability. Deep generative models such as Generative Adversarial Network (GAN) [16] and Variable Auto-Encoder (VAE) [17] are widely used for the synthetic expansion of fault samples. By constructing high-quality virtual fault signals, researchers have effectively alleviated the problems of overfitting and insufficient generalization ability caused by the scarcity of real fault data. Some studies have also combined strategies such as data augmentation, sample less learning, pseudo-labelling and self-supervised learning to further improve the intelligent fault recognition of rolling bearings under very few samples. However, both migration learning and generative data methods face many challenges, such as large distributional differences between source and target domains, difficulty in guaranteeing the authenticity of generated samples, and insufficient model robustness. Especially in the real industrial environments of cross-equipment, cross-condition, and even cross-platform, how to ensure that the migrated and generated diagnostic models have a high degree of generalization ability and physical consistency is still the difficulty and hotspot of current research [18, 19].

With the integration of artificial intelligence and physical modelling, physical information-driven intelligent fault diagnosis methods for rolling bearings have gradually become the focus of attention in academia and industry. Although traditional data-driven methods have made significant progress under large sample conditions, their "black box" characteristics and neglect of physical mechanisms have led to insufficient robustness and interpretability of the models under scarce samples and abnormal working conditions. For this reason, more and more researchers try to integrate the a priori knowledge of rolling bearing dynamics mechanism, material damage evolution, and physical characteristics of vibration signals into intelligent diagnostic models. Currently, physical information-driven diagnostic methods mainly include neural networks based on physical constraint optimization, deep coupling of physical models and generative models, and knowledge-driven decision reasoning mechanisms [20–22]. Some scholars have effectively improved the interpretability and generalization ability of diagnostic models by introducing physical constraint loss functions and jointly training physical models and data models. Another study combines physical mechanism modelling such as finite element analysis, system identification, friction contact theory, etc., with deep generative networks to generate high-quality fault samples with physical consistency, which provides a solid data foundation for intelligent diagnosis under extreme operating conditions and zero sample conditions. In addition, knowledge graph [23], expert system [24], symbolic reasoning [25] and other soft-hard technologies also provide multi-source information fusion and reasoning capability for rolling bearing health management. Physical information-driven intelligent fault diagnosis of rolling bearings can not only make up for the shortcomings of pure data-driven methods under data scarcity, but also significantly improve the adaptive ability of the model to new types of faults, unknown working conditions and complex environments. Therefore, facing the future, in-depth excavation of the coupling relationship between physical mechanism and data characteristics of rolling bearings, and constructing a new paradigm of zero-sample intelligent fault diagnosis with physical consistency, high robustness, and strong generalization ability have become the inevitable trend and core goal to promote the theoretical innovation and practical application of high-end equipment intelligent health management.

Considering that in real industrial environments, rolling bearings have multi-source monitoring signals (e.g., vibration, acoustic emission, temperature, current, etc.), and these signals can reflect the operating characteristics and potential fault mechanisms of bearings from different physical levels [26–28]. Therefore, the complementary information of multi-source monitoring signals is fully exploited, combined with the knowledge-driven multimodal feature fusion and inference mechanism, to construct an intelligent model that can realize rolling bearing fault detection under the condition of zero fault samples. Based on this, this paper introduces the rolling bearing monitoring signal analysis knowledge such as Hilbert envelope analysis and high frequency filtering, and at the same time, combines the information entropy reciprocity between the multi-source monitoring signals and the a priori knowledge of the strong temporal sequencing of monitoring signals, to design the autonomous training method of the intelligent model under the zero-fault samples, and combine the rolling bearing fault detection algorithms, and ultimately to obtain the intelligent abnormality detection model of the rolling bearing under the zero-sample conditions. The main contributions are as follows:

(1) The Hilbert envelope analysis and high-frequency filtering are introduced to achieve the enhancement of the inherent fault spectrum information in monitored signals of rolling bearings.
(2) A novel loss function is designed based on the multimodal monitoring signals and the strong temporal sequence of monitoring signals, which can realize the effective optimization of the deep network model with zero-fault samples
(3) An intelligent fault detection algorithm for rolling bearings is constructed based on the defined prior knowledge of the mutual difference in information entropy and the trained network-based feature encoder.

This paper is arranged as follows: the second part is the problem definition, the third part is the proposed method and its main content, the fourth part is the experimental validation of the proposed method and the discussion and analysis of the results, and the fifth part is the conclusion.

2. Problem definition

For the actual mechanical equipment of engineering, most of the time it is in reliable operation, so the monitoring of its state, the monitoring data obtained is mostly normal state data, with a large amount of data, but with low-density characteristics. And with the development of artificial intelligence technology, a large amount of data is needed for training and learning, which is compatible with the characteristics of mechanical equipment, with a large amount of condition monitoring data. However, AI methods have high requirements on data quality, which makes it challenging to apply AI technology to the learning of massive monitoring data of mechanical equipment to establish its intelligent condition monitoring, sensing, and recognition models. Therefore, how to carry out autonomous deep learning on massive monitoring data of mechanical equipment has become a key issue to be solved in this paper. Taking rolling bearings as an example, there are various ways to monitor their state, and the state can be characterized from vibration, sound, and other modal data. Currently, most of the proposed fault diagnosis methods for rolling bearings are implemented by constructing supervised learning tasks to achieve fault diagnosis modelling of rolling bearings, however, this type of method requires the support of different state data samples, and it is obvious that for rolling bearings in engineering practice, the states represented by their monitoring data are mostly normal, which leads to difficulties in fault diagnosis model modelling through supervised learning. Therefore, how to establish an artificial intelligence model to obtain an intelligent fault diagnosis model for rolling bearings using multimodal monitoring data of rolling bearings under zero fault samples is the key problem to be solved in this paper.

3. Proposed method

Due to the scarcity of available fault data in mechanical equipment, the requirements of data-driven methods for comprehensive datasets are often unmet. To address this limitation, this paper leverages the abundance of sensor monitoring data in mechanical systems and proposes an end-to-end intelligent fault detection method that integrates time-frequency analysis knowledge of rolling bearing vibration signals with multi-modal prior information. As illustrated in **Figure 1**, the proposed method consists of five main components: (a) multi-modal data acquisition, (b) embedding of time-frequency analysis and multi-modal prior knowledge, (c) design of a contrastive loss function for the deep network model under zero-fault-sample conditions, (d) construction of a feature encoder based on EfficientFormer, and (e) development of an end-to-end intelligent fault detection algorithm for rolling bearings utilizing the obtained feature encoder.

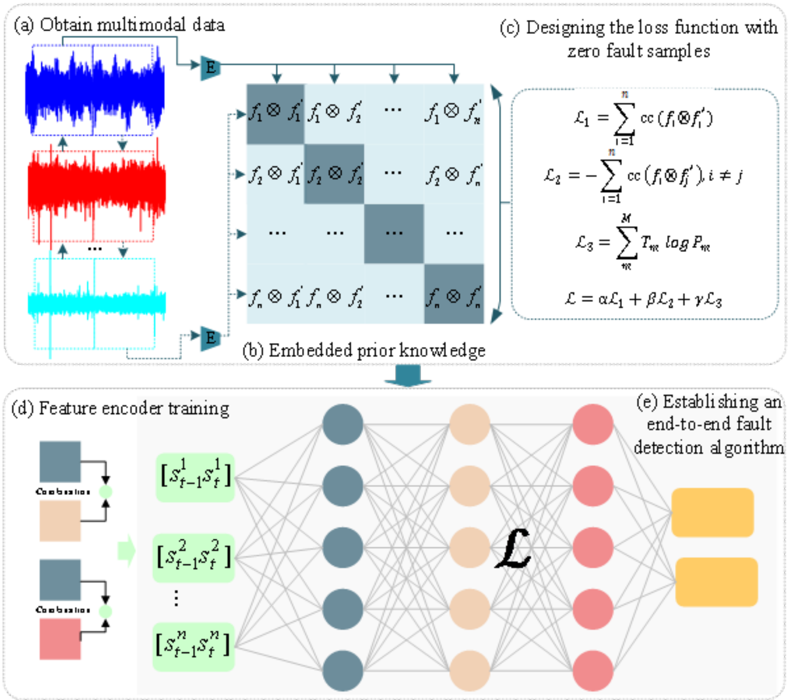


Figure 1. The framework for diagnostic methods under zero compound fault samples.

3.1. Hilbert envelope analysis

In real industrial scenarios, it tends to have a very low. A key consideration in this study is that rolling bearings, as essential components in rotating machinery, inherently exhibit periodic motion characteristics, and their fundamental frequency significantly influences the monitoring signals. Particularly for data-driven approaches, the fundamental frequency acts as a dominant feature and can affect the training of intelligent models. To address this issue, the Hilbert envelope is employed to analyze combined with high-frequency filtering to preprocess the monitoring signals of rolling bearings. Specifically, after performing envelope analysis on the bearing signals, a high-frequency filter is applied to further process the envelope, thereby enhancing the fault frequency information while suppressing the influence of the fundamental frequency. The detailed procedure is as follows: let the vibration signal be denoted as $x(t)$, and the modulation signal as $a(t)$. The relationship between them can be expressed as follows:

$$x(t) = a(t) \cos\left(2\pi f_s t + \phi(t)\right) \tag{1}$$

where f_s denotes the carrier frequency and $\phi(t)$ represents the phase modulation. The modulation process described above mainly involves the following steps:

Step 1: Perform the Hilbert transform on the signal and construct the analytic signal by introducing a 90-degree phase shift.

$$\tilde{x}(t) = H\left\{x(t)\right\} \tag{2}$$

Step 2: Perform envelope analysis. Specifically, compute the modulus (magnitude) of the analytic signal.

$$a(t) = |z(t)| = \sqrt{x^2(t) + \tilde{x}^2(t)} \tag{3}$$

Through the above operations, the low-frequency modulation signal $a(t)$ can be separated from the high-frequency carrier.

Step 3: Conduct envelope spectrum analysis. Specifically, apply the Fourier transform to the envelope signal to obtain its spectral information.

$$A(t) = fft\left(a(t)\right) \tag{4}$$

Step 4: As the fundamental frequency modulates the fault frequency in the low-frequency band, whilst the fault frequency continually manifests as harmonic frequencies, and given the finite energy of noise frequencies in the high-frequency band, a high-pass filter is applied to the envelope spectrum to further eliminate the influence of low-frequency disturbances, thereby accentuating the fault frequency in the high-frequency band.

Step 5: The envelope spectrum after high-pass filtering undergoes an inverse Fourier transform to yield the processed signal samples.

3.2. EfficientFormer

To tackle the challenge of developing intelligent models trained solely on monitoring signals acquired from rotating machinery under normal operating conditions, we propose a lightweight network architecture designed to mitigate the inherent risk of overfitting. While the integration of time-frequency analysis and temporal prior knowledge effectively addresses the limitations posed by imbalanced training data, the exclusive reliance on normal-state monitoring signals heightens the propensity for overfitting. To circumvent this issue, the adoption of a lightweight network structure is imperative. The EfficientFormer, a re-engineered Vision Transformer (ViT) architecture optimized for latency, leverages dimension-consistent Transformer blocks as its core components, coupled with a lightweight attention mechanism. This design substantially reduces the risk of overfitting while maintaining robust representational capacity.

A pivotal aspect of the proposed methodology lies in the judicious selection of a feature encoder. With the advent of generative artificial intelligence, Transformer-based models have demonstrated superior performance compared to conventional convolutional neural networks (CNNs) and recurrent neural networks (RNNs), underscoring their substantial potential for advanced applications. Nevertheless, the computational complexity of Transformer models, which necessitates significant resource allocation, presents a considerable barrier to their deployment in practical engineering contexts. To address this limitation, we build upon the foundational Transformer architecture [29] by introducing the EfficientFormer, a streamlined model tailored for computational efficiency. The EfficientFormer employs a latency-driven optimization strategy, substituting large convolutional kernels with multiple smaller kernels and refining the attention mechanism through techniques such as low-rank approximations or depth-wise separable convolutions. These modifications significantly reduce computational overhead while preserving model efficacy, thereby facilitating the practical implementation of the proposed approach in resource-constrained engineering environments.

$$Attention_{light} == Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{5}$$

where Q, K, and V reduce the number of parameters by parameter sharing.

In addition, EfficientFormer employs a hybrid architectural design, using a CNN architecture (e.g., *MBConv*) at the shallow layer and a lightweight Transformer block at the deep layer to balance local and global feature extraction.

$$Block_i = \begin{cases} MBConv, & i \leq \frac{N}{2} \\ Transformer_{light}, & i \geq \frac{N}{2} \end{cases} \tag{6}$$

3.3. Embedded multimodal and physical prior knowledge zero-fault sample loss functions

For data-driven intelligent fault diagnosis modelling, the key lies in the effective training of the intelligent model. However, the fault data of rolling bearings in engineering practice is extremely scarce, with even zero fault samples, which leads to the difficulty of achieving intelligent fault diagnosis modelling of rolling bearings under an unbalanced training set. Through literature research, it is known that the training of intelligent models lies in the construction of the loss function. Therefore, how to construct a suitable loss function under zero-fault samples of rolling bearings is the key to achieving intelligent diagnostic modelling under zero-fault samples. Considering that most of the actual rolling bearings are in normal service, the normal state monitoring data are sufficient, and with the continuous development of sensing technology, rolling bearings have a variety of state monitoring means, and a large number of normal state data in different modes can be obtained. At the same time, taking into account the strong temporal sequence of rolling bearing monitoring data, this a priori knowledge is not applied in the previous intelligent model modelling process. Based on this, this paper introduces the a priori knowledge of multi-modal data and strong temporal sequence to design a new type of loss function with zero fault samples.

Let the monitoring data of rolling bearings in different modes under normal state be $\{a_1^0, a_2^0, \cdots, a_n^0\}$, $\{b_1^0, b_2^0, \cdots, b_n^0\}$, $\cdots$, $\{x_1^0, x_2^0, \cdots, x_n^0\}$, in which '0' denotes the normal state, *n* denotes the number of samples in the normal state. Let the feature encoder based on EfficientFormer be $Effenet$, then the encoding process of the normal state samples of rolling bearings with different modes is as follows:

$$\{f_1^0, f_2^0, \cdots, f_n^0\} = Effenet(\{x_1^0, x_2^0, \cdots, x_n^0\}) \tag{7}$$

Since rolling bearings have multiple modes of monitoring signals, multiple high-dimensional features can be obtained from the above equation. Also, from the strong temporal order of the rolling bearing monitoring signals, it is known that:

$$x_t^0 \otimes x_{t-1}^0 \approx 1 \tag{8}$$

where, $t-1$ and t are denoted as the sample labels collected by the rolling bearing normal condition monitoring signals under continuous time series. Through the above equation, it can also be obtained that the high-dimensional features obtained from the two samples through $Effenet$ should satisfy:

$$f_t^0 \otimes \tilde{f}_{t-1}^0 \approx 1 \tag{9}$$

At the same time, it is evident from the monitoring principles and methodologies of sensors that signals acquired by different sensors exhibit modal variations, with discrepancies also existing in the information content being monitored. It should be noted that for triaxial vibration sensors, monitoring in different directions demonstrates varying sensitivities to distinct vibration orientations, particularly manifesting in amplitude differences, which further underscores their inherent distinctiveness. Consequently, to reflect this differentiation, a limit expression is introduced.

$$x_t^a \otimes x_t^b \approx 0,\ \tilde{f}_t \otimes f_t \approx 0 \tag{10}$$

where a and b denote the different sensors.

The expression of different modal a priori knowledge with a strong temporal sequence of monitoring signals can be achieved by Equations (8), (9), and (10). The innovation of this paper is to design the effective optimization of the feature encoder under zero-fault samples based on the above equation to get rid of the dependence on rolling bearing fault samples. The specific process is as follows:

Considering above expressions are all similarity and dissimilarity relationships, it is necessary to mathematically make use of these relationships, the comparison matrix can be constructed:

$$M = \begin{pmatrix} f_a \otimes \tilde{f}_a & \cdots & f_a \otimes \tilde{f}_x \\ \vdots & \ddots & \vdots \\ f_x \otimes \tilde{f}_a & \cdots & f_x \otimes \tilde{f}_x \end{pmatrix} \tag{11}$$

where the subscripts $a, b, \cdots x$ denote the different modes, and f and $\tilde{f}$ denote the high-dimensional representations of the signal samples collected under the continuous time series of the monitoring signal, respectively. From the above equation, the correlation calculation of $f \otimes \tilde{f}_x$ yields the correlation coefficient matrix M^{cc}:

$$M^{cc} = \begin{pmatrix} cc\left(f_a \otimes \tilde{f}_a\right) & \cdots & cc\left(f_a \otimes \tilde{f}_x\right) \\ \vdots & \ddots & \vdots \\ cc\left(f_x \otimes \tilde{f}_a\right) & \cdots & cc\left(f_x \otimes \tilde{f}_x\right) \end{pmatrix} \tag{12}$$

For M^{cc} its embedded in a variety of physical prior knowledge. There is a maximum similarity between high-dimensional features in the same modality, i.e., for M^{cc} the diagonal elements are maximal, based on which it can be obtained that:

$$\mathcal{L}_1 = \sum_{i=1}^{n} \mathrm{cc}\left(f_i \otimes \tilde{f}_i\right), i \in (a, b, \cdots, x) \tag{13}$$

Minimum similarity between the high-dimensional representations of different samples of modal monitoring signals, based on which it can be obtained:

$$\mathcal{L}_2 = \sum_{i,j=1}^{n} \mathrm{cc}\left(f_i \otimes \tilde{f}_j\right), i \neq j \; and \; i, j \in (a, b, \cdots, x) \tag{14}$$

Furthermore, from the comparison matrix, $\mathrm{cc}(f_i \otimes \tilde{f}_j)$ in row *i* of matrix M^{cc} should be the largest in the row when $i = j$, which means that the best optimisation result of the comparison matrix is:

$$T = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \tag{15}$$

Therefore, an argmax operation on each row of M^{cc} should yield the matrix of the above equation, which can be further expressed as:

$$P = [p_1, p_2, \cdots, p_r\,] \tag{16}$$

where *r* denotes the number of different modal monitoring data available for rolling bearings and p_r is the row vector.

$$p_r = \left[\mathrm{cc}\left(f_x \otimes \tilde{f}_a\right), \cdots, \mathrm{cc}\left(f_x \otimes \tilde{f}_x\right)\right]_{\text{argmax}} \tag{17}$$

The cross-entropy loss function is obtained by performing cross-entropy calculations on P and T:

$$\mathcal{L}_3 = \sum_{m}^{M} T_m log P_m \tag{18}$$

Through Equations (16), (17) and (18), a novel loss function with embedded multimodal and temporal monitoring signals first verifies that knowledge can be obtained:

$$\mathcal{L} = \alpha\mathcal{L}_1 + \beta\mathcal{L}_2 + \gamma\mathcal{L}_3,\ \alpha < 0,\ \beta, \gamma > 0 \tag{19}$$

3.4. Intelligent fault detection algorithms for rolling bearings

Based on the introduced Hilbert envelope analysis and the designed novel loss function embedded with the a priori knowledge of multimodal and timing monitoring signals, the deep network model training can be realized under the zero fault samples of rolling bearing, so the key of this section is how to get the deep network encoder based on the training to realize the intelligent detection of rolling bearings. Specific steps include:

Step 1: Select rolling bearing benchmark state samples $(a_1, b_1, \cdots, x_1)$, i.e., select samples from different modal monitoring data as benchmark state data respectively.

Step 2: Multimodal monitoring data acquisition of rolling bearings, intercepting the state monitoring signal samples, using Hilbert envelope analysis and high-frequency filtering methods to process the baseline signals and the acquired signals, and using morl mother wavelet to carry out the wavelet transform, to obtain the time-frequency mapping.

Step 3: The EfficientFormer-based feature encoder obtained by training, performs high-dimensional feature characterization of the benchmark wavelet spectrum and the monitoring signal wavelet spectrum, and constructs a high-dimensional characterization similarity comparison matrix:

$$M_x^{cc} = \begin{pmatrix} \mathrm{cc}\left(f_{a_1}^b \otimes \tilde{f}_{a_1}^1\right) & \cdots & \mathrm{cc}\left(f_{a_1}^b \otimes \tilde{f}_{x_1}^1\right) \\ \vdots & \ddots & \vdots \\ \mathrm{cc}\left(f_{x_1}^b \otimes \tilde{f}_{a_1}^1\right) & \cdots & \mathrm{cc}\left(f_{x_1}^b \otimes \tilde{f}_{x_1}^1\right) \end{pmatrix} \tag{20}$$

Step 4, normalize M_x^{cc}, map each row of elements in the matrix to the interval [−0.5, 0.5] respectively, and use the argmax operation to obtain the Hot code P_x for each row of elements in M_x^{cc} as follows:

$$r_m^x = \left[\mathrm{cc}\left(f_m^b \otimes f_1^x\right), \cdots, \mathrm{cc}\left(f_m^b \otimes f_n^x\right)\right] \tag{21}$$

$$v_i = \frac{r_m^x[i] - min(r_m^x)}{max(r_m^x) - min(r_m^x)} - 0.5 \tag{22}$$

$$P_m^x = [v_1^m, v_2^m, \cdots, v_n^m]_{\text{argmax}} \tag{23}$$

$$P_x = [P_1^x, \cdots, P_m^x, \cdots, P_n^x] \tag{24}$$

After normalization, the values are scaled uniformly to weaken the differences brought by absolute values and highlight the relative comparative advantages of feature similarity. Based on the normalized comparison matrix, one threshold is set. Generally, when P_m^x meets abs(sum(P_m^x)) $\geq$ threshold, the state is judged to be abnormal. When abs(sum(P_m^x)) $<$ threshold and $P_x = [0, 1, \cdots, n]$, it is judged to be normal, otherwise it is faulty. It is worth noting that threshold selection is determined based on the detection results of the normal state, and the thresholds are typically set higher, while for faults that are subtle yet potentially severe, thresholds are set lower to enhance sensitivity to such faults. Furthermore, the similarity comparison results between time-frequency spectrum are also used to achieve the fault detection, when the similarity with the normal state of time-frequency maps is greater than 0.5, it will also be set as a fault.

However, it should be noted that the proposed fault detection algorithm is based on the premise that the load and rotational speed of rolling bearings do not undergo significant variation during normal operation. This is because such variations would test the generalizability of the proposed method. Furthermore, during modelling, signal samples under different load and speed conditions were not incorporated for training, inevitably leading to a degree of overfitting in the model.

4. Case studies and experimental results

4.1. Rolling bearing experimental platform and dataset description

The main feature of the method proposed in this paper is that it can achieve intelligent fault detection under zero fault samples by fusing multimodal data of rolling bearings. In order to verify the effectiveness of the proposed method, it is necessary to build an experimental platform that can collect multimodal data of rolling bearings. To this end, this paper builds a rolling bearing experimental platform as shown in **Figure 2**, which consists of a motor, a coupling, two test bearings and a pulley. It should be pointed out that this platform needs to cooperate with the multimodal data acquisition device to ensure the alignment of the collected data. The acquisition instrument is Siemens Sim-center test-lab, equipped with sensors including single-axis vibration sensors, three-axis (XYZ) vibration sensors, sound level meter and so on.

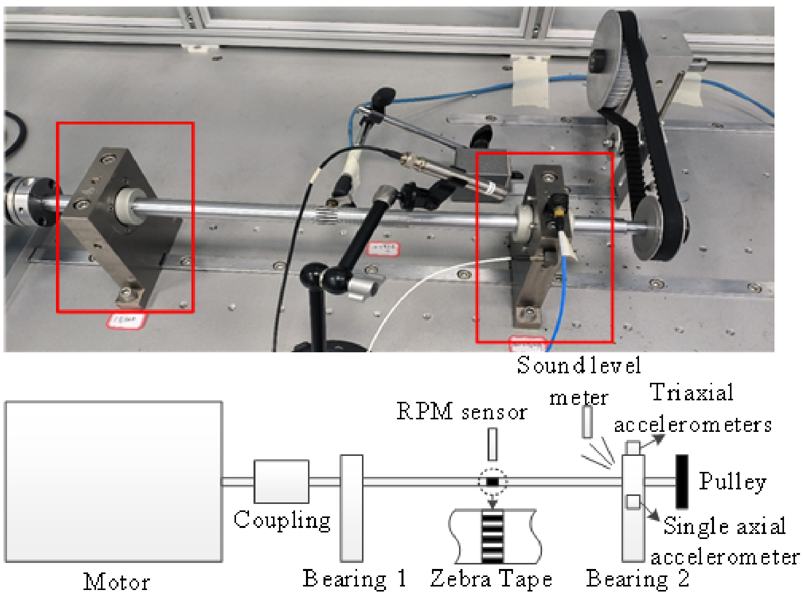


Figure 2. The rolling bearing fault experimental platform.

This test bench is set up a total of four types of failure states, respectively, the outer ring damage, inner ring damage bearing, rolling body damage, mixed damage. Experiments are mainly conducted through the adjustment of belt tension, setting the load size, and each time the replacement of bearings was used to determine the tensiometer, tensiometer size is set to −11.60 mm, through the tachometer to measure the actual speed of different operating conditions. A set speed of 500–1200 rmp/min was carried out in eight operating conditions, the sampling frequency of the collector was set at 102,400 Hz, and the test time under each condition was 900 s. The details of the data collected under a selected operating condition as the validation data set are shown in **Table 1**. XYZ-axis and uniaxial acceleration signals and sound signals were acquired for each experimental condition. In the case validation, multiple signals were acquired and divided into two groups to verify the effectiveness of the proposed method. Firstly, two groups of XYZ-axis acceleration signals were randomly selected for validation of the proposed method, including X-axis acceleration signals and Y-axis acceleration signals, X-axis acceleration signals and Z-axis acceleration signals, and Z-axis acceleration signals and Y-axis acceleration signals. Secondly, the X-axis acceleration signal, the single-axis acceleration signal and the sound signal are also combined two by two respectively. Therefore, a total of six sets of proposed method validation can be performed.

Table 1. Details of the rolling bearing fault detection method validation dataset.

Labels	**Fault name**	**Conditions**	**Description**	**Monitored signals**
Normal	Normal	500 rmp/min 55.30 N	—	XYZ-axis, single-axis acceleration signal and sound signal.
Anomaly1	Inner fault	500 rmp/min 55.30 N	Two scratches on the inner ring	XYZ-axis, single-axis acceleration signal and sound signal.
Anomaly2	Outer fault	500 rmp/min 55.30 N	Two small holes on the outer ring	XYZ-axis, single-axis acceleration signal and sound signal.
Anomaly3	Rolling body fault	500 rmp/min 55.30 N	Rolling body scratches	XYZ-axis, single-axis acceleration signal and sound signal.
Anomaly4	Compound fault	500 rmp/min 55.30 N	Mixed three faults.	XYZ-axis, single-axis acceleration signal and sound signal.

4.2. Discussion and analysis of results

The multimodal monitoring data of rolling bearings collected from actual experiments are used to carry out experimental validation of the method proposed in this paper. The normal state bearing vibration data under the working condition of 500 rmp/min is used as the training set, and the length of the training samples is set to 0.15 s, which results in a total of 1000 training samples. The vibration data collected under the four fault states is used as the test set, and the test samples for each type of state are all 1000.

Firstly, the effectiveness of the introduced Hilbert envelopment analysis and high-frequency filtering for the rolling bearing monitoring signals is verified. The cut-off frequency of the high-pass filter is 8000 Hz. The spectral analysis of the vibration signals after embedding the HHT envelope analysis in the proposed method is shown in **Figure 3**. Without HHT envelope analysis, the spectral energy is high and the fundamental frequency energy is more prominent. After HHT envelope analysis and high frequency filtering, the fault frequency is more prominent, which effectively guarantees the quality of the subsequent time-frequency analysis.

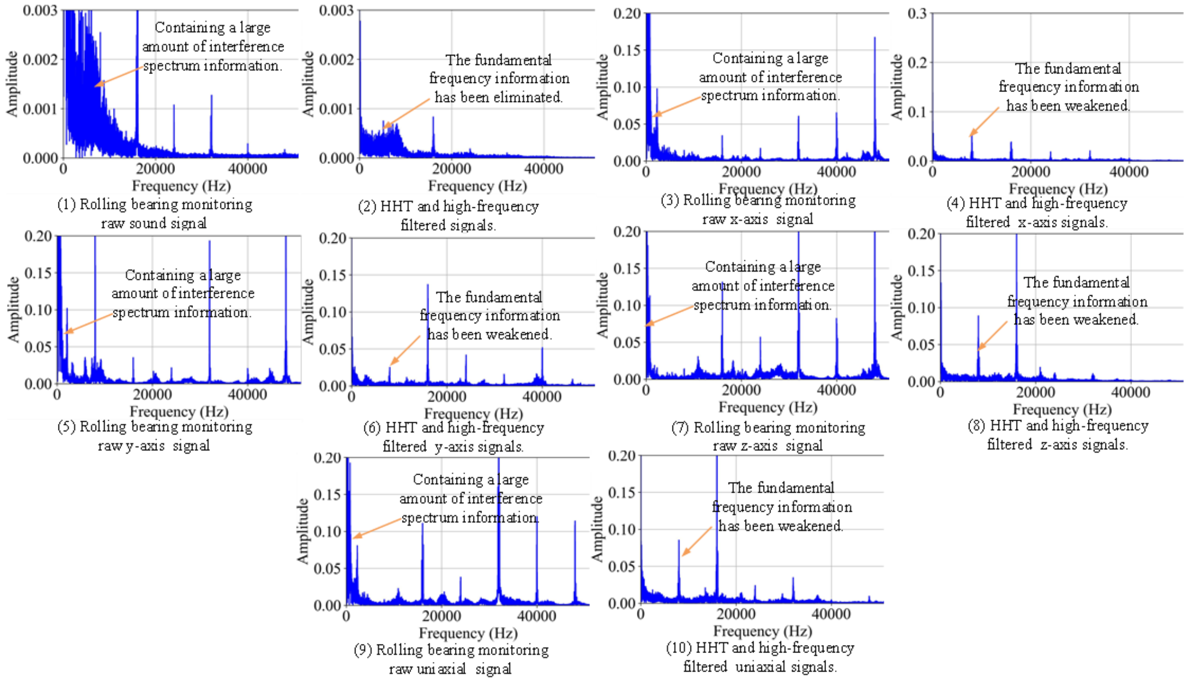


Figure 3. The Hilbert envelope is introduced to compare the spectrum with the high-frequency filtered spectrum.

One of the major advantages of the proposed method is that it can achieve effective training of deep network encoders by embedding multimodal and timing signal a priori knowledge under the condition of zero-fault samples of rolling bearings. In order to verify the effectiveness of the proposed method, two types of multimodal data collected under the normal condition of rolling bearings are selected to train the EfficientFormer-based feature encoder. A total of six sets of method validations were performed. **Figure 4** shows the optimization process of the loss function designed during the training of the proposed zero-fault sample-dependent feature encoder on six different sets of modal monitoring data of rolling bearings in normal state. From the results, it can be seen that the loss functions involved are able to achieve the descent under zero-fault samples and have the convergence property, which indicates that the designed loss functions are able to optimize the network parameters of EfficientFormer under the combination of six sets of multimodal data, and all of them are able to achieve the optimization of the network parameters. The above six sets of experiments show that the proposed loss function with embedded multimodal data and strong a priori knowledge of timing signals is effective and efficient, which provides a practical path to get rid of the dependence on rolling bearing fault data, and shows the potential of intelligent fault diagnosis modelling under zero fault samples of rolling bearings.

To further validate the designed loss function, three weights were adjusted to examine their impact on the proposed method. Three weights were set to 0.95 and 0.5 respectively, and the threshold is set to 0.2. Six sets of experiments were conducted using uniaxial vibration signals and acoustic signals. The test results are shown in **Table 2**. It indicates that when the weights of the three loss functions are set to 0.95, their impact on the proposed method is negligible, and the test accuracies are 99.25%, 98.75% and 100, respectively. However, the test accuracy decreases to varying degrees when the weights α is set to -0.5, or β, γ to 0.5. Finding the optimal parameters near 1 clearly requires substantial computational effort. Therefore, it suggests that α is set to -1, and β, γ to 1.

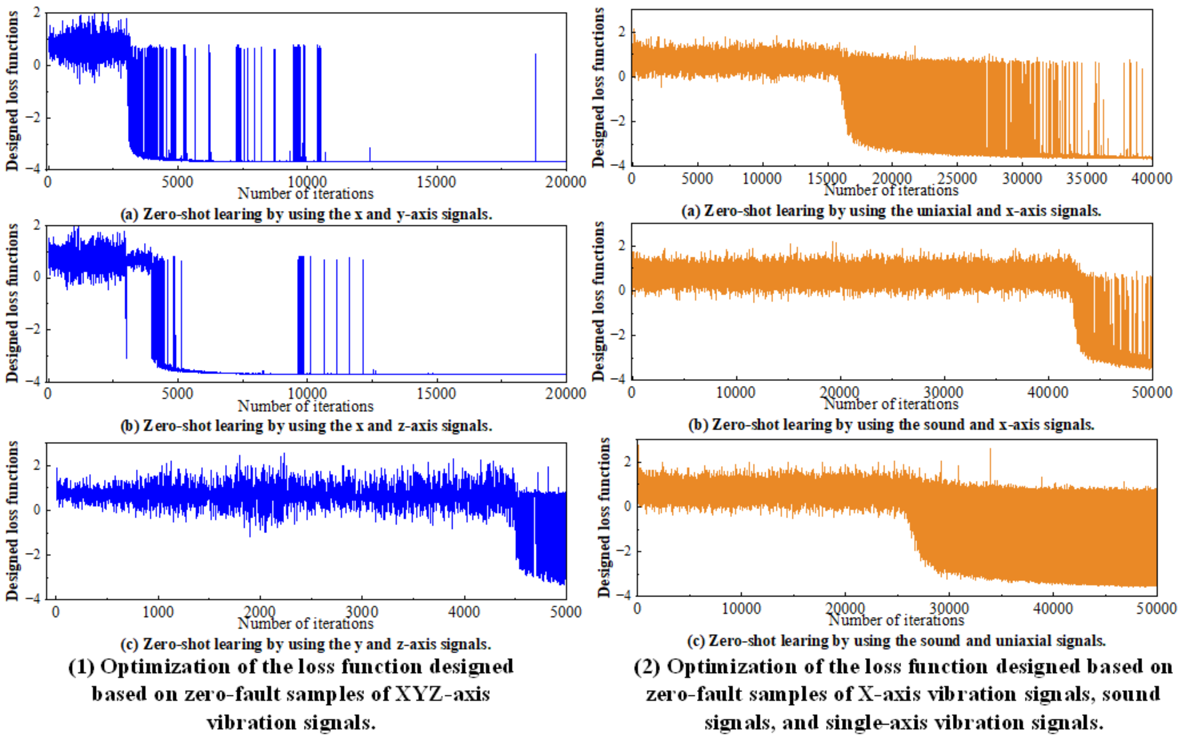


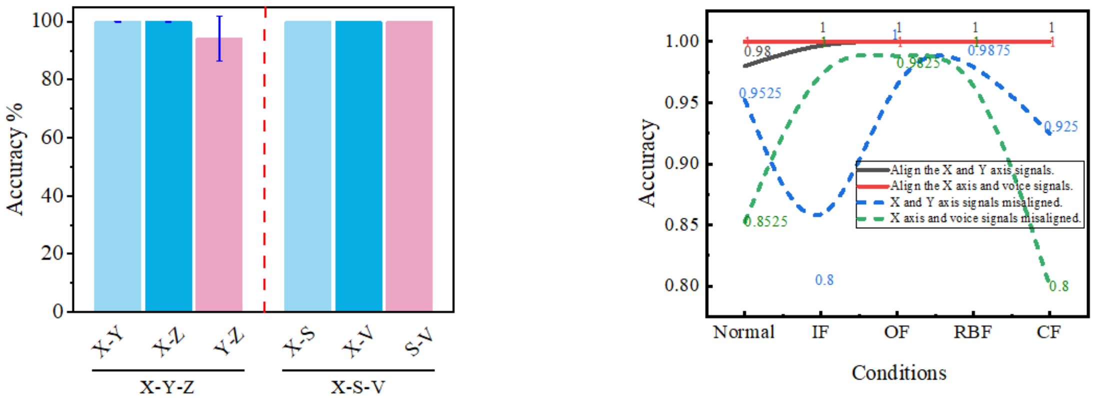
Figure 4. Embedded knowledge of multiple time-frequency analyses for designed loss function in the training process.

Table 2. Details of the rolling bearing fault detection method validation dataset.

Scenario	α	β	γ	Test accuracy (%)
1	−1	1	1	100
2	−0.95	1	1	99.25
3	−1	0.95	1	98.75
4	−1	1	0.95	100
5	−0.5	1	1	95.68
6	−1	0.5	1	80.33
7	−1	1	0.5	90.56

Figure 5a shows the results of ten validations under different combinations of modal information for the rolling bearing with zero fault samples. From the results, it can be seen that by using the triaxial vibration sensor in the normal state of the rolling bearing, the combination of the two can achieve the establishment of an effective intelligent fault detection model. In addition, the use of triaxial vibration signals, single-axis vibration signals, sound signals, two combinations can also achieve good results, and better than the three-axis vibration signals of the two combinations. In this paper, it should be pointed out that the method proposed in this paper is also applicable to the zero-fault sample modelling under the three modes of different signals, but when

the introduction of the three modal signals, the design of the zero-fault sample model training task will become more complex, which will lead to the collapse of the model training when the difference in the information between the different modal signals is not very obvious. Furthermore, considering that misalignment between multimodal signals impacts the established Equation (10), misalignment between modal signals leads to greater divergence among signal samples, thereby simplifying the defined learning task and hindering model optimization. **Figure 5b** presents test results for both aligned and misaligned signals. The results demonstrate that when modality signals are misaligned, test performance deteriorates. This confirms that misalignment simplifies the learning task for deep representation models, hindering their optimization. The complexity of the defined learning task must be considered to ensure effective training of deep representation models in practical applications of proposed method.

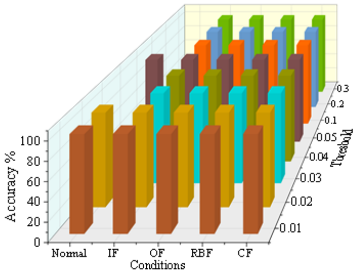


(a) Ten validation results of the proposed method under different model signal combinations.

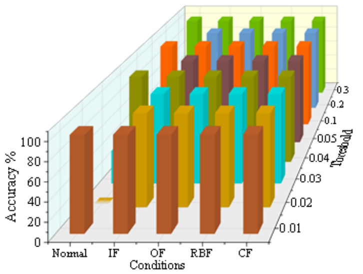
(b) Validation results for multimodal signals with 2 ms misalignment.

Figure 5. Validation and discussion of experiment result.

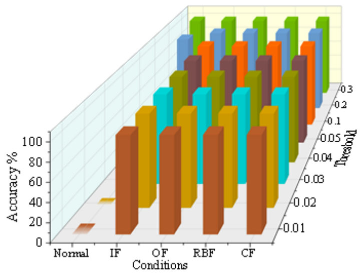
To further validate the performance of the proposed rolling bearing zero-sample fault detection method, the intelligent fault detection method is modelled under triaxial vibration signals, and the proposed method is validated under different fault detection thresholds for the detection of different states of rolling bearings. **Figure 6** shows the validation results of the proposed method under different thresholds for the detection of different states. The result shows that the validation performance of the proposed method with different detection thresholds is different under different combinations of triaxial vibration signals, which is mainly reflected in the detection of the normal state of rolling bearings. However, with the increase of threshold, as shown in **Figure 6a,b**, the verification of various states of rolling bearings tends to be optimal. Under the three-axis vibration signals X-axis and Y-axis, the performance of the constructed fault detection model is fluctuating, and the best results are obtained at the detection thresholds of 0.01, 0.02 and 0.1, respectively. This indicates that the similarity between the X-axis and Y-axis vibration signals makes it difficult to ensure that the constructed loss function is optimized to be optimal for zero fault samples. Additionally, when setting the threshold in the designed fault detection algorithm, it is recommended to initially set it higher. This ensures accurate detection of the rolling bearing's normal state. This is because the proposed method, which only uses normal state signals for training, is prone to overfitting. Experimental validation results also demonstrate that when the threshold is set to 0.3, the detection accuracy for all states in **Figure 6b,c** approaches 100%. However, setting the threshold lower tends to misclassify normal states as faults, as seen in **Figure 6c** with 0.01.



(a) Average results of ten trials of the proposed method for fault detection at different thresholds under X-Y axis vibration on data.



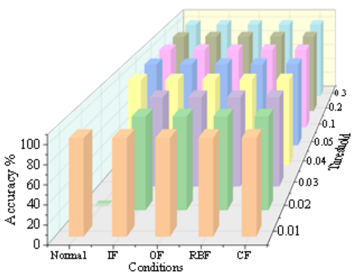
(b) Average results of ten trials of the proposed method for fault detection at different thresholds under X-Z axis vibration on data.



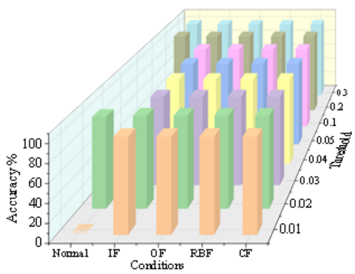
(c) Average results of ten trials of the proposed method for fault detection at different thresholds under Y-Z axis vibration on data.

Figure 6. Results of ten trials under different combinations of X-Y-Z vibration with different detection thresholds.

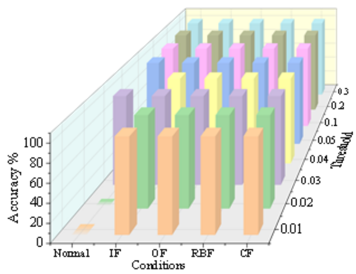
The rolling bearing zero-sample fault detection method constructed under triaxial vibration signals is also validated, uniaxial vibration signals and acoustic signals for ten tests respectively. From **Figure 7**, it can be seen that under the combination of three multimodal signals, the proposed method is able to establish an intelligent fault detection model with zero fault samples of faulty bearings, and with the fault detection threshold, it is able to effectively detect various fault states of rolling bearings. Unlike the model constructed under the triaxial vibration signal, the intelligent fault detection model obtained by combining the triaxial vibration signal, the uniaxial vibration signal and the acoustic signal can achieve the best detection performance when the detection threshold is set above 0.2, which is obviously more in line with the assumption of the difference between different modal signals introduced. When the threshold is set to 0.01, the results in **Figure 7b,c** demonstrate that the fault detection threshold should not be set too low. Conversely, when set higher such as at thresholds of 0.2 and 0.3 the test results prove more favorable.



(a) Average results of ten trials of the proposed method for fault detection at different thresholds with X and single axis vibration.



(b) Average results of ten trials of the proposed method for fault detection at different thresholds with X and voice data.



(c) Average results of ten trials of the proposed method for fault detection at different thresholds with single axis vibration and voice data.

Figure 7. Results of ten trials with different combinations of vibrations and voice with different detection thresholds.

To show the advancement of the method proposed, the method advancement is demonstrated at two levels. Firstly, advanced deep network models such as CNN, Resnet, Desennet and Vision Transformer are selected as rolling bearing intelligent fault detection models, respectively. Moreover, all the deep network models are trained by supervised learning, i.e., 1000 labeled samples of rolling bearings in various states are used as the training set, and the obtained models are verified against the rolling bearing test samples. **Figure 8** gives the comparison results between the four advanced deep network models under supervised learning and the proposed method, and the test errors between them are described by histograms. Description. Considering that the monitoring signals of five different modes of rolling bearings are captured in this experiment, the supervised learning of deep network models is carried out under various signals separately. From the comparison results, it can be seen that the performance of the proposed rolling bearing fault detection method with zero fault samples is not much different from that of the fault detection model obtained under supervised learning, and it also shows more than that of the deep network model trained by supervised learning under the monitoring signals of different modes, such as **Figure 8a,c–e**. Under triaxial Y-axis vibration signals, the proposed method does not show an advantage and its performance is 5.82% lower in terms of testing accuracy compared to Resnet.

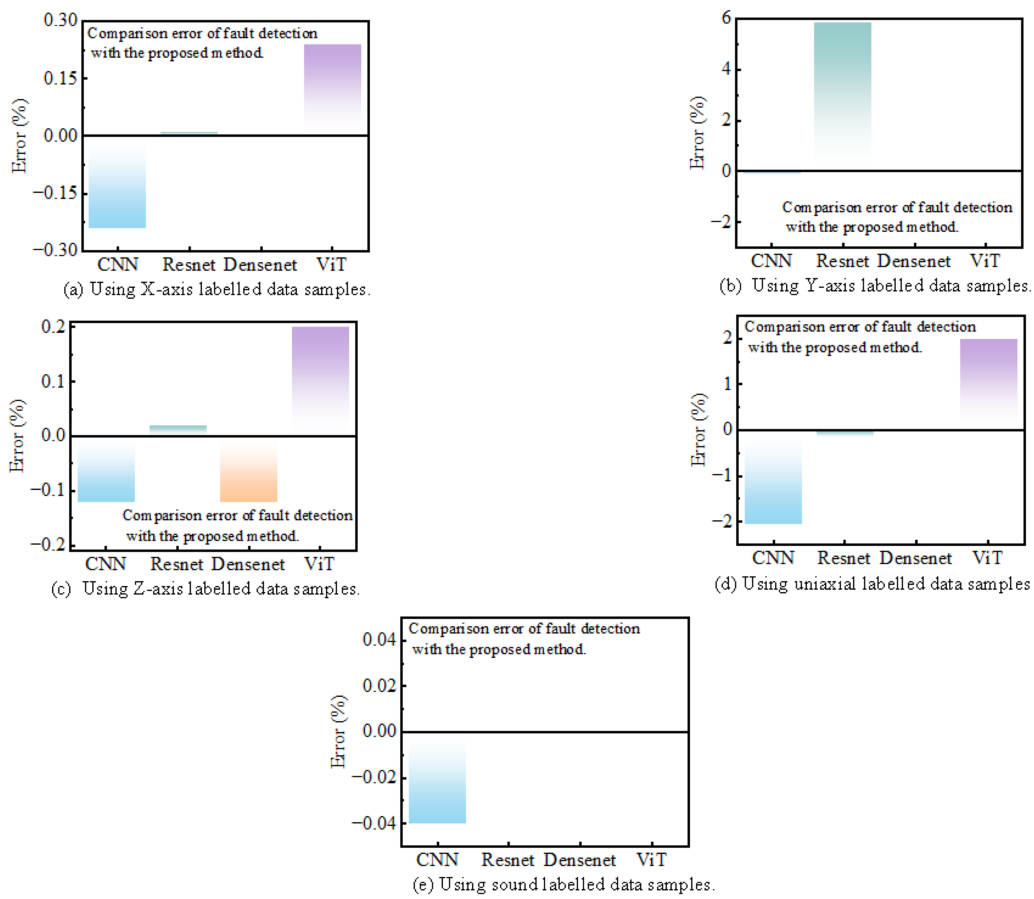


Figure 8. Comparison of the proposed method with different supervised learning advanced network models for fault detection.

To further validate the advancement of the proposed method, it is compared with advanced machine learning models such as XGBoost, linear discriminator (LAD), plain Bayesian network (BNB), decision tree (DT), and Gaussian Bayesian network (GNB). All of the above advanced machine learning models are trained using supervised learning, i.e., the rolling bearing fault samples are sufficient. Additionally, other unsupervised anomaly detection algorithms such as autoencoders (AE), one-class support vector machines (OC-SVM), and isolated forest (IF) were employed for comparison with the proposed method to further validate its superiority. **Table 3** gives the comparison results between different methods in terms of fault detection accuracy, modelling training time and model real-time performance. From the comparison results, it can be seen that compared with the traditional machine learning model the proposed method achieves the advantage in fault detection test, which exceeds the test accuracy of OC-SVM 94.26% and XGBoost 91.72%. Overall, the proposed methods represent advanced approaches equivalent to fully supervised and unsupervised learning techniques. Furthermore, to address the issue of excessive signal sample data points input into the aforementioned machine learning models, the principal component analysis (PCA) is introduced to downscale the signal and subsequently train the machine learning model. Specifically, it makes the length of each test sample larger due to the high sampling rate used, which makes the time required for the training of the traditional machine learning model longer. The proposed method converts the signal samples into time-frequency maps by combining the wavelet time-frequency analysis method, making it more suitable for EfficientFormer processing, and which makes the training time shorter. Moreover, the training and testing time of the comparison models is shortened when the PCA is introduced, but this also sacrifices performance to a certain extent.

Table 3. Comparative results between the proposed method and other state-of-the-art methods.

Methods	Fault detection (%)	Training time	Testing time
Ours	98.97%	214.65 s	0.040 ms
AE	85.64%	1256.24 s	200.5 ms
OC-SVM	94.26%	564.23 s	25.24 ms
IF	86.75%	2056.23 s	560.2 ms
SVM [30]	80.00%	464.20 s	11.13 ms
GMM [31]	50.92%	182.99 s	0.084 ms
LAD [32]	64.16%	1.10 s	0.391 ms
BNB [33]	77.64%	1.13 s	65.61 ms
GNB [34]	87.88%	2449.6 s	547.5 ms
DT [35]	78.16%	201.24 s	0.072 ms
XGBoost [36]	91.72%	85.93 s	0.110 ms
PCA + SVM	26.04%	1.87 s	0.0004 ms
PCA + GMM	80.00%	3.20 s	0.0002 ms
PCA + LAD	80.00%	2.80 s	0.0012 ms
PCA + BNB	74.94%	2.69 s	0.0008 ms
PCA + GNB	80.00%	3.40 s	0.353 ms
PCA + DT	83.30%	2.91 s	0.0004 ms
PCA + XGBoost	90.40%	2.96 s	0.0016 ms

5. Conclusion

The fault samples of rolling bearings are extremely difficult to obtain in the real world, which makes the existing intelligent methods difficult to model. To address the above problem, an intelligent fault detection method with zero fault samples was proposed by embedding the prior knowledge of multimodal monitoring signal reciprocity and strong temporal sequencing of monitoring signals. The Hilbert envelope was used to decrease the inherent fundamental frequency information interference in monitored signals. Three novel loss functions were constructed through the mutual heterogeneity of multimodal monitoring signals and the strong temporal order of sampled signals to obtain a trained network-based model, and then a fault detection algorithm was established. Compared with the advanced deep network model and machine learning model under supervised learning, the fault detection performance of the proposed method is still satisfactory, and it has better application value in engineering practice because of the zero-fault sample dependence. However, the effectiveness of the proposed method under varying operating conditions remains to be verified, particularly with regard to enhancing its performance across different rotational speeds of rolling bearings. The research focus lies in establishing a model capable of adapting to anomaly detection tasks across varying rotational speeds. Moreover, the future challenge lies in achieving intelligent fault diagnosis for rolling bearings under zero-fault-sample conditions based on the proposed method, enabling the identification of different fault types and varying degrees of the same fault.

Author contributions: HW: Methodology, Conceptualization, Investigation, Writing original draft, Writing review & editing. HQ and RZ: Conceptualization, Software. GL: Visualization. JL: Visualization. JY: Methodology, assisting in conducting bearing fault experiments. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially supported by the Natural Science Foundation of Fujian Province (2024J08064), the Natural Science Foundation of Xiamen under Grant (3502Z202471042), and the Key Laboratory of Marine Power Engineering Technology of the Ministry of Transport (Wuhan University of Technology) (KLMPET2023-02).

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Data availability statement: Data would be made available on reasonable request.

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

References

1. Zhang M, Jiang Z, Feng K. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. Mechanical Systems and Signal Processing. 2017; 93: 460–493. doi: 10.1016/j.ymssp.2017.02.013
2. Li C, De Oliveira JV, Cerrada M, et al. A Systematic Review of Fuzzy Formalisms for Bearing Fault Diagnosis. IEEE Transactions on Fuzzy Systems. 2019; 27(7): 1362–1382. doi: 10.1109/TFUZZ.2018.2878200
3. Chen X, Yang R, Xue Y, et al. Deep Transfer Learning for Bearing Fault Diagnosis: A Systematic Review Since 2016. IEEE Transactions on Instrumentation and Measurement. 2023; 72: 1–21. doi: 10.1109/TIM.2023.3244237
4. Wang D, Miao Q, Fan X, et al. Rolling element bearing fault detection using an improved combination of Hilbert

and wavelet transforms. Journal of Mechanical Science and Technology. 2009; 23(12): 3292–3301. doi: 10.1007/s12206-009-0807-4

5. Wang D, Zhao X, Kou L-L, et al. A simple and fast guideline for generating enhanced/squared envelope spectra from spectral coherence for bearing fault diagnosis. Mechanical Systems and Signal Processing. 2019; 122: 754–768. doi: 10.1016/j.ymssp.2018.12.055
6. Meng D, Nie P, Yang S, et al. Reliability analysis of wind turbine gearboxes: past, progress and future prospects. International Journal of Structural Integrity. 2025; 16(1): 4–38. doi: 10.1108/IJSI-08-2024-0129
7. Dalian Y, Junjun Z, Hui L. Capsule networks for intelligent fault diagnosis: a roadmap of recent advancements and challenges. Expert Systems with Applications. 2026; 296: 128814. doi: 10.1016/j.eswa.2025.128814
8. Rehman AU, Jiao W, Jiang Y, et al. Deep learning in industrial machinery: A critical review of bearing fault classification methods. Applied Soft Computing. 2025; 171: 112785. doi: 10.1016/j.asoc.2025.112785
9. Tang J, Hu Y, Zhou X, et al. Self-powered sensor for online monitoring of eccentricity faults in rotating machinery and its application in spindle eccentricity monitoring of machine tools. Nano Energy. 2024; 130: 110084. doi: 10.1016/j.nanoen.2024.110084
10. Liu K, Li Y, Cui Z, et al. Adaptive frequency attention-based interpretable Transformer network for few-shot fault diagnosis of rolling bearings. Reliability Engineering & System Safety. 2025; 263: 111271. doi: 10.1016/j.ress.2025.111271
11. Siavash-Abkenari N, Rahmani-Sane G, Torkaman H, et al. A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations. Ain Shams Engineering Journal. 2023; 14(4): 101945. doi: 10.1016/j.asej.2022.101945
12. Xu Y, Fan Y, Bao Y, et al. Few-shot learning for structural health diagnosis of civil infrastructure. Advanced Engineering Informatics. 2024; 62: 102650. doi: 10.1016/j.aei.2024.102650
13. Pan H, Xu H, Cheng J, et al. A Multiclass Graph Embedding Matrix Classification Method for Roller Bearing State Identification Under Limited Sample. IEEE Transactions on Reliability. 2025; 74(3): 3824–3832. doi: 10.1109/TR.2025.3530441
14. Siavash-Abkenari N, Rahmani-Sane G, Torkaman H, et al. Exploring a Cutting-Edge Framework for Bearing Fault Detection: A Synergistic Approach Integrating Statistical Analysis and Deep Learning Methods. IEEE Journal of Emerging and Selected Topics in Industrial Electronics. 2024; 5(3): 1226–1233. doi: 10.1109/JESTIE.2024.3373313
15. Choudhary A, Mian T, Fatima S, et al. Passive Thermography Based Bearing Fault Diagnosis Using Transfer Learning With Varying Working Conditions. IEEE Sensors Journal. 2023; 23(5): 4628–4637. doi: 10.1109/JSEN.2022.3164430
16. Liu X, Liu S, Xiang J, et al. A transfer learning strategy based on numerical simulation driving 1D Cycle-GAN for bearing fault diagnosis. Information Sciences. 2023; 642: 119175. doi: 10.1016/j.ins.2023.119175
17. Wang Y, Li D, Li L, et al. A novel deep learning framework for rolling bearing fault diagnosis enhancement using VAE-augmented CNN model. Heliyon. 2024; 10(15): e35407. doi: 10.1016/j.heliyon.2024.e35407
18. Ren Z, Lin T, Feng K, et al. A Systematic Review on Imbalanced Learning Methods in Intelligent Fault Diagnosis. IEEE Transactions on Instrumentation and Measurement. 2023; 72: 1–35. doi: 10.1109/TIM.2023.3246470
19. Sun Y, Song H, Guo L, et al. A transfer learning method: Universal domain adaptation with noisy samples for bearing fault diagnosis. Advanced Engineering Informatics. 2025; 65: 103243. doi: 10.1016/j.aei.2025.103243
20. Zhong J, Zheng Y, Ruan C, et al. M-IPISincNet: An explainable multi-source physics-informed neural network based on improved SincNet for rolling bearings fault diagnosis. Information Fusion. 2025; 115: 102761. doi: 10.1016/j.inffus.2024.102761
21. Qin Y, Liu H, Wang Y, et al. Inverse physics–informed neural networks for digital twin–based bearing fault diagnosis under imbalanced samples. Knowledge-Based Systems. 2024; 292: 111641. doi: 10.1016/j.knosys.2024.111641
22. Herwig N, Borghesani P, Smith W, et al. Signal processing- and physics-informed neural network for explainable bearing condition monitoring. Mechanical Systems and Signal Processing. 2025; 235: 112925. doi: 10.1016/j.ymssp.2025.112925
23. Peng C, Sheng Y, Gui W, et al. A Rolling Bearing Fault Diagnosis Method Based on Multimodal Knowledge Graph. IEEE Transactions on Industrial Informatics. 2024; 20(11): 13047–13057. doi: 10.1109/TII.2024.3431074
24. Berredjem T, Benidir M. Bearing faults diagnosis using fuzzy expert system relying on an Improved Range Overlaps and Similarity method. Expert Systems with Applications. 2018; 108: 134–142. doi: 10.1016/j.eswa.2018.04.025
25. Wang Z, Yu X, Guo Y, et al. A noise-enhanced feature extraction method combined with tunable Q-factor wavelet transform and its application to planet-bearing fault diagnosis. Applied Acoustics. 2025; 239: 110845. doi: 10.1016/j.apacoust.2025.110845

26. Ying W, Li Y, Noman K, et al. Stockwell transform spectral amplitude modulation method for rotating machinery fault diagnosis. Mechanical Systems and Signal Processing. 2025; 223: 111884. doi: 10.1016/j.ymssp.2024.111884
27. Zhao D, Wang H, Cui L. Frequency-chirprate synchrosqueezing-based scaling chirplet transform for wind turbine nonstationary fault feature time–frequency representation. Mechanical Systems and Signal Processing. 2024; 209: 111112. doi: 10.1016/j.ymssp.2024.111112
28. Lv Y, Hao J, Tang M, et al. Multimodal data fusion-based intelligent fault diagnosis for ship rotating machinery: Status quo and perspectives. Engineering Applications of Artificial Intelligence. 2025; 160: 111767. doi: 10.1016/j.engappai.2025.111767
29. Han K, Wang Y, Chen H, et al. A Survey on Vision Transformer. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2023; 45(1): 87–110. doi: 10.1109/TPAMI.2022.3152247
30. Hosseini NK, Toshani H, Abdi Jalebi S, et al. Enhanced Bearing Fault Detection in Induction Motors Using Projection-Based SVM. IEEE Transactions on Industry Applications. 2025; 61(3): 3623–3636. doi: 10.1109/TIA.2025.3536425
31. Wang Z, Liu T, Wu X, et al. Application of an Oversampling Method Based on GMM and Boundary Optimization in Imbalance-Bearing Fault Diagnosis. IEEE Transactions on Industrial Informatics. 2024; 20(2): 1931–1940. doi: 10.1109/TII.2023.3282236
32. Kim HH, Choi JY. Hierarchical multi-class LAD based on OvA-binary tree using genetic algorithm. Expert Systems with Applications. 2015; 42(21): 8134–8145. doi: 10.1016/j.eswa.2015.06.037
33. Muralidharan V, Sugumaran V. A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. Applied Soft Computing. 2012; 12(8): 2023–2029. doi: 10.1016/j.asoc.2012.03.021
34. Li G, Wu J, Deng C, et al. Convolutional Neural Network-Based Bayesian Gaussian Mixture for Intelligent Fault Diagnosis of Rotating Machinery. IEEE Transactions on Instrumentation and Measurement. 2021; 70: 1–10. doi: 10.1109/TIM.2021.3080402
35. Cui L, Dong Z, Xu H, et al. Triplet attention-enhanced residual tree-inspired decision network: A hierarchical fault diagnosis model for unbalanced bearing datasets. Advanced Engineering Informatics. 2024; 59: 102322. doi: 10.1016/j.aei.2023.102322
36. Trizoglou P, Liu X, Lin Z. Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines. Renewable Energy. 2021; 179: 945–962. doi: 10.1016/j.renene.2021.07.085