Journal of Reliability
Science and
Engineering

PAPER • OPEN ACCESS

Metric learning-based two-stage imbalanced fault diagnosis model for water injection pump

To cite this article: Zhiwen Yao *et al* 2025 *J. Reliab. Sci. Eng.* **1** 045402

View the article online for updates and enhancements.

You may also like

- On uniformly staggered testing strategy for redundant safety instrumented systems
Sun-Keun Seo and Won Young Yun
- Intelligent reliability assurance methodologies for engineering systems: advances and challenges
Zhiliang Liu, Qin Zhang, Yu Liu et al.
- A new Kalman filter algorithm based on the interval process model
Jinyang Li, Jinwu Li, Chengkai Shen et al.

This content was downloaded from IP address 118.123.172.214 on 15/12/2025 at 02:48

OPEN ACCESS
IOP Publishing | ISECAEP
Journal of Reliability Science and Engineering
J. Reliab. Sci. Eng. 1 (2025) 045402 (15pp)
https://doi.org/10.1088/3050-2454/ae227a

Metric learning-based two-stage imbalanced fault diagnosis model for water injection pump

Zhiwen Yao[1], Xiwen Cui[1], Nagi Gebraeel[2], Tangbin Xia[1,3,*] and Lifeng Xi[1,3]

[1] State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, SJTU-Fraunhofer Center, Shanghai 200240, People's Republic of China
[2] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332, United States of America
[3] Special Environment Digital Manufacturing Equipment Technology Invitation Center, Mianyang, Sichuan 621900, People's Republic of China

E-mail: xtbxtb@sjtu.edu.cn

Received 25 June 2025, revised 26 October 2025
Accepted for publication 20 November 2025
Published 3 December 2025

Abstract

To address the severe data imbalance and the difficulty of feature extraction in fault diagnosis of plunger-type injection pumps, we propose a fault diagnosis algorithm that integrates metric learning with a sampling-adaptation mechanism. First, we design a feature extractor based on Laplacian wavelet convolution and a multi-scale attention mechanism to capture critical fault features and adaptively assign feature weights. Second, we introduce a metric-learning module driven by dual-center loss, imposing inter-class and intra-class distance constraints in a high-dimensional embedding space to promote clustering of samples from the same fault class and thus enhance fault recognition capability. Finally, we adopt a class-balanced batch sampling mechanism to adaptively strengthen the model, overcoming the performance limitations imposed by data imbalance and further improving overall accuracy. Validated on a real, imbalanced injection-pump dataset, the proposed method achieves a diagnostic accuracy of 97.11%.

Keywords: imbalanced fault diagnosis, Laplacian wavelet transform, metric learning, dual-center loss

1. Introduction

As the heart of oilfield water injection systems, water injection pumps critically determine production efficiency and operational safety across oil extraction sites. Among various pump types, plunger-type water injection pumps stand out due to their high operational efficiency, making them particularly suitable for high-pressure, high-power scenarios requiring precise flow regulation, especially in medium-small oilfields during mid-to-late exploitation phases that demand sustained injection pressures. The intricate mechanical design of plunger-type water injection pumps, combined with prolonged operation under high-pressure, high-vibration, and fluctuating load conditions, as well as frequent exposure to corrosive substances and solid impurities in injected fluids, contributes to frequent pump failures.

To overcome the limitations of existing maintenance strategy, the petroleum exploration and development industry urgently requires the implementation of predictive maintenance for water injection pumps [1]. This method employs real-time equipment monitoring to proactively identify potential faults, perform fault localization and root-cause analysis, and

* Author to whom any correspondence should be addressed.

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

© 2025 The Author(s). Published by IOP Publishing Ltd on behalf of the Institute of Systems Engineering, China Academy of Engineering Physics

establish a scientific basis for formulating efficient and precise maintenance strategies. Consequently, it reduces downtime, lowers maintenance costs, and extends equipment service life. Prognostics and health management (PHM) technology [2–5] is widely recognized as an effective solution to achieve these goals and has been extensively adopted across production and manufacturing industries. By appropriately applying PHM technology, operators can not only ensure system reliability and safety but also significantly enhance operational efficiency and reduce operational costs, making it a critical competitive advantage in the industry.

Similar to fault diagnosis methods for other mechanical equipment, fault diagnosis methods for water injection pumps can be broadly categorized into two types: mechanism-based models and data-driven approaches [6]. Mechanism-based fault diagnosis methods [7] involve modeling the physical characteristics and operational patterns of the water injection pump system. While these methods offer strong interpretability, their implementation is constrained by the numerous factors influencing pump operation in oilfield environments, limiting the applicability of these knowledge-dependent approaches. Data-driven methods, on the other hand, leverage the vast amounts of data generated during the operation of water injection pumps to develop and optimize diagnostic models through machine learning [8] or deep learning algorithms [9, 10], establishing a direct mapping between historical operational data and fault types for fault diagnosis. For instance, Li *et al* [11] proposed converting one-dimensional vibration signals collected from the hydraulic end of plunger pumps into three-channel three-dimensional images, which were then classified using the deep neural network AlexNet, achieving an average diagnostic accuracy of 98.00% for multiple fault types. Tang *et al* [12] transformed vibration signals into time-frequency images and utilized a Gaussian process-based Bayesian optimization algorithm to automatically select the hyperparameters of convolutional neural networks (CNN), enabling efficient fault diagnosis of axial plunger pumps.

While these methods demonstrate satisfactory performance under experimental conditions, their deployment in real-world water injection pump operations remains challenging.

Although research on data imbalance fault diagnosis techniques for plunger-type water injection pumps is relatively limited, the issue of data imbalance has consistently attracted widespread attention from both academia and industry, with numerous studies on other mechanical equipment providing valuable insights. Currently, solutions to this issue are primarily approached from three perspectives: data optimization, feature optimization, and training process optimization [13, 14].

Data optimization methods aim to balance datasets by directly adjusting the sample sizes, which include various sampling techniques and generative adversarial networks (GANs) [15]. For instance, Wang *et al* [16] proposed an improved technique which weights the samples of the majority class while oversampling the samples of the minority class and validated the effectiveness and innovativeness of the method on three different datasets, turbine rotors, bearings and gears. Wang *et al* [17] integrated Fourier transforms with GANs for data augmentation, achieving simultaneous fitting of data distributions in both the time and frequency domains, with experimental results demonstrating that, compared to alternative methods, the bearing fault signals generated by this approach were closer to real data in both domains. However, data optimization methods have certain limitations in practical applications. The quality of the generated samples is difficult to guarantee, and they may introduce noise or outliers.

Feature optimization methods address the issue of dominant feature space generation caused by the overrepresentation of majority class samples by imposing constraints on the feature space constructed for fault diagnosis models, ensuring that the features of minority class samples are adequately expressed. For instance, Zhou *et al* [18] proposed a local feature selection algorithm based on multi-objective binary particle swarm optimization, which directly utilizes the local distribution of data to obtain near-optimal feature subsets, achieving optimal results in experiments across various simulation and real-world datasets with different imbalance ratios. Jian and Ao [19] introduced a semi-supervised ensemble learning method that combines multiple base models to extract information from unlabeled data, and employs weighted voting and dynamic subset adjustment strategies to ensure model diversity and stability, effectively preventing the drowning out of minority class features. However, feature optimization methods generally lack interpretability.

Training process optimization methods primarily focus on model training strategies, with cost-sensitive learning being a representative technique. Specifically, cost-sensitive learning mitigates the impact of data imbalance by assigning different weights to different classes or samples and adjusting the loss function, ensuring that the model gives adequate attention to minority class samples during training. For instance, Zhao *et al* [20] designed a misclassification cost-sensitive matrix that considers both data quantity and practical industrial costs, and utilized a multi-objective multi-universe optimization algorithm to automatically search for the optimal hyperparameter combination for an improved weighted extreme learning machine, demonstrating its effectiveness on reciprocating pump and bearing datasets. Mao *et al* [21] proposed a transferable dynamic-enhancement cost-sensitive network, which integrates maximum classifier discrepancy with Wasserstein distance to match the feature distributions between the source and target domains, and introduces dynamic-enhanced focal loss, significantly improving the model's focus on hard-to-classify and minority class samples, thereby enhancing the model's generalization ability and diagnostic accuracy. However, while training process optimization methods are intuitive and straightforward, they are prone to introducing the risk of overfitting.

When applying metric learning methods to industrial sites for end-to-end imbalanced fault diagnosis of water injection pumps, the disparity in fault severity and the complexity of the operating environment hinder the full utilization of metric learning capabilities, leading to diminished algorithm performance. In contrast, metric-learning-based methods [22], which combine the advantages of feature optimization and training-process optimization, design loss constraints in the

embedding space so that the distance between samples of the same class is reduced while the distance between samples of different classes is enlarged. This has become one of the effective approaches to addressing the difficulty of extracting fault features for mechanical equipment and its components.

Accordingly, this study proposes a method that integrates metric learning with a sampling-adaptation mechanism. First, Laplacian wavelet convolution combined with a multi-scale spatiotemporal attention mechanism is employed to capture key fault features across multiple temporal and spatial scales, thereby optimizing the model at the feature level. Next, a metric-learning module based on dual-center loss is introduced to constrain the sample distribution in the embedding space and thus enlarge decision boundaries between classes, optimizing the model at the training level. Finally, a class-balanced sampling-adaptation mechanism is adopted to fine-tune the model, further enhancing the focus on minority-class samples and optimizing the model at the data level. Together, these components effectively alleviate the adverse impact of data imbalance on the fault-diagnosis performance of water injection pumps.

2. Methodology

2.1. Framework

The proposed two-stage model comprises two phases: pre-training and sampling adaptation, facilitating end-to-end fault diagnosis of water injection pumps. The input signals are the raw vibration data collected from the operational environment of water injection pumps, while the output is the fault category determined by the model.

The pre-training stage consists of three modules: a feature extractor based on Laplace wavelet convolution (LWC) and multi-scale convolutional attention mechanism (MCAM) (referred to as the LWC and MCAM-based feature extractor, LMFE), a metric learning module based on dual center loss (DCL), and a classifier. Initially, the vibration signals pass through the LMFE, where fault information across different scales is thoroughly extracted, and weights are adaptively allocated, resulting in an embedded vector mapped onto a high-dimensional embedding space. Simultaneously, the DCL-based metric learning module is applied to compute the DCL of the feature vectors in the embedding space, guiding the samples to cluster more effectively and optimizing the mapping relationship between the samples and the embedding space. Subsequently, a fully connected classifier is employed to convert the embedded vectors into fault classification labels, with parameters updated through multiple iterations of training to obtain the optimal pre-trained model.

In the sampling adaptation stage, the Class-Balanced Batch Sampling Mechanism (CBBSM) is utilized to fine-tune the parameters of the feature extractor and classifier obtained from the pre-training phase, leading to the construction of the final fault diagnosis model for water injection pumps.

In this methodology, the optimization strategy is systematically structured across three distinct levels: feature extraction, training, and data optimization. Firstly, the integration of LWC with the multi-scale spatiotemporal attention mechanism effectively captures critical fault features across diverse spatiotemporal scales. This synergistic approach enhances model performance by optimizing the extraction of pertinent features, thereby enriching the representation of input signals. Secondly, the introduction of a metric learning module utilizing DCL establishes a robust framework for constraining the sample distribution within the embedding space. This mechanism facilitates tighter clustering of similar fault categories while broadening the decision boundaries between different classes, thereby optimizing the training process and improving classification accuracy. Lastly, the implementation of a class-balanced sampling adaptation mechanism strategically concentrates attention on minority class samples during training. This methodology not only strengthens the representation of underrepresented classes but also mitigates the adverse effects of data imbalance, fostering a more equitable learning process. Consequently, it enhances the overall performance of the fault diagnosis model for water injection pumps across multiple dimensions. This comprehensive optimization approach significantly augments diagnostic capability, equipping the system to effectively address the complexities inherent in imbalanced fault diagnosis scenarios. The flowchart of the model is shown in figure 1.

In real-world scenarios where only streaming operational signals are available and fault labels are absent, the proposed model supports online adaptation by freezing the pre-trained LMFE feature extractor to preserve established fault feature representations. During deployment, only the classifier and attention module parameters are updated using new unlabeled data, thus avoiding disruption of the original diagnostic knowledge. This can be achieved through strategies such as pseudo-labeling, where the model iteratively uses its own high-confidence predictions as temporary labels, or self-supervised clustering, which groups incoming feature embeddings to reflect evolving operational conditions. By selectively fine-tuning the decision layers while keeping the feature extractor fixed, the model can incrementally adapt to changes in the data distribution, maintaining reliable diagnostic performance without manual annotation or retraining from scratch.

2.2. Pre-training stage

2.2.1. LWC and MCAM-based feature extractor. The proposed LMFE primarily consists of two components: LWC and MCAM, which are designed to comprehensively capture fault information within the input vibration signals.

The first convolutional layer of a CNN significantly influences the network's performance, as this layer directly processes high-dimensional data and is particularly susceptible to the vanishing gradient problem in deeper network architectures [23]. Furthermore, the convolutional kernels learned during the training process of a CNN often exhibit noisy and incoherent multi-band shapes, which can compromise their interpretability. Therefore, it is essential to carefully design the convolutional kernel in the initial layer of the model, enhancing feature extraction capabilities while

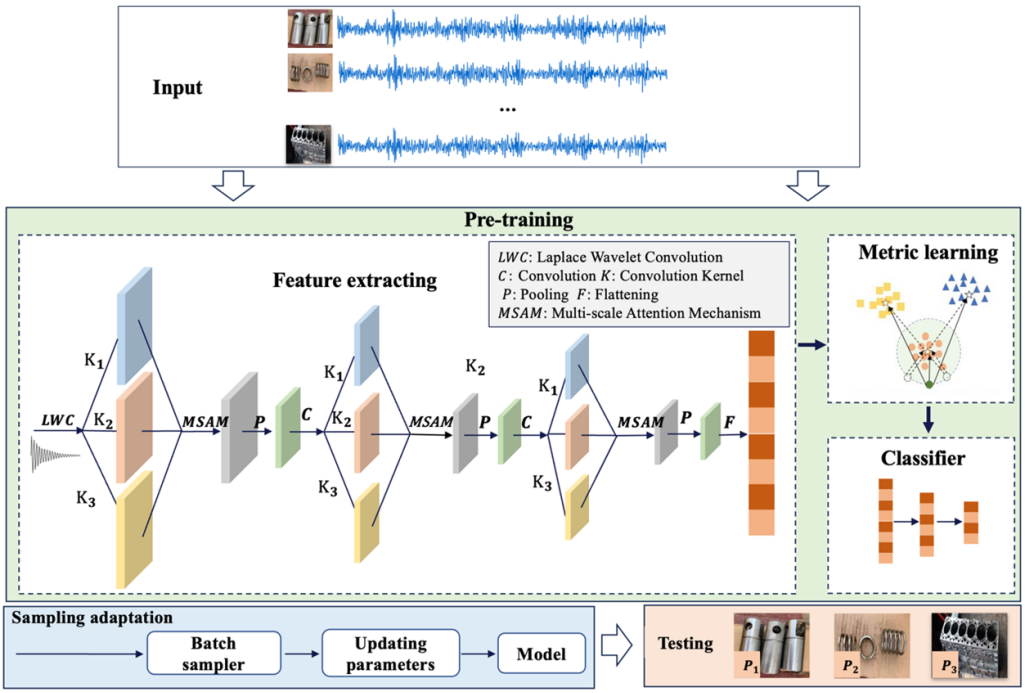


Figure 1. Imbalanced fault diagnosis model fusing metric learning and sampling adaptation mechanism.

ensuring that the convolutional waveforms possess physical significance.

Continuous wavelet transform (CWT) is a widely employed technique in the field of signal processing. By selecting appropriate wavelet basis functions, it decomposes time-domain signals into sub-bands of varying scales and frequencies, effectively representing non-stationary signals and revealing detailed information at different scales. Wavelet Convolution incorporates the concept of CWT within the CNN framework by utilizing predefined wavelet basis functions as convolutional kernels. As a result, wavelet convolution not only inherits the effective representation capabilities of wavelet transforms for non-stationary signals but also facilitates automatic feature extraction through the convolution operation. Compared to traditional convolutional layers, wavelet convolution enables the model to achieve higher diagnostic accuracy with fewer parameters, while generating more interpretable feature maps [24].

Common fault types in water injection pumps, such as wear and loosening, frequently manifest as transient impacts within the vibration signals, characterized by distinct unidirectional decaying oscillations. To address this phenomenon, the Laplace wavelet basis function is identified as an ideal candidate for the wavelet convolution kernel due to its distinctive waveform characteristics. As illustrated in figure 2, the waveform of the function aligns well with the form of impact signals caused by faults in the water-injection pump. The function can be expressed as follows:

$$\psi_{\mathrm{u,s}}(t) = Ae^{\frac{-\zeta}{\sqrt{1-\zeta^2}}2\pi f\left(\frac{t-u}{s}-\tau\right)}\sin\left(2\pi f\left(\frac{t-u}{s}-\tau\right)\right) \quad (1)$$

where $\psi_{\mathrm{u,s}}(\cdot)$ represents Laplacian wavelet basis function, t represents time, A represents wavelet normalization function, ξ represents damping ratio, τ represents time parameter, u represents translation parameter, s represents scale parameter, f represents signal frequency.

The waveform of this function exhibits a strong correlation with the impact signals induced by faults in water injection pumps, as shown in the figure 3, thereby facilitating the accurate capture of both time-domain and frequency-domain features within the vibration signals. Consequently, LWC is designated as the initial convolutional layer of the LMFE, and the output generated by the LWC layer can be mathematically represented as follows:

$$L = B\left(\delta\left(\psi_{\mathrm{u,s}}(t) * x_t\right)\right) \quad (2)$$

where L represents LWC output, x_{t} represents input vibration signal, B represents batch normalization operation, δ represents *RELU* activation function, $*$ represents convolution.

Traditional CNN typically employ fixed-size convolutional kernels for feature extraction; however, this configuration presents significant limitations when applied to fault diagnosis in plunger-type water injection pumps [25]. Firstly, the diverse fault signals in plunger-type water injection pumps exhibit high heterogeneity, necessitating different kernel scales for various fault modes. For instance, high-frequency vibration signals often signify micro-defects or localized damage, making smaller-scale convolutional kernels more suitable for precise local feature extraction. In contrast, low-frequency vibration signals reflect global patterns, revealing the overall operational state and structural characteristics of the water injection pump, which are better suited to larger-scale convolutional kernels.

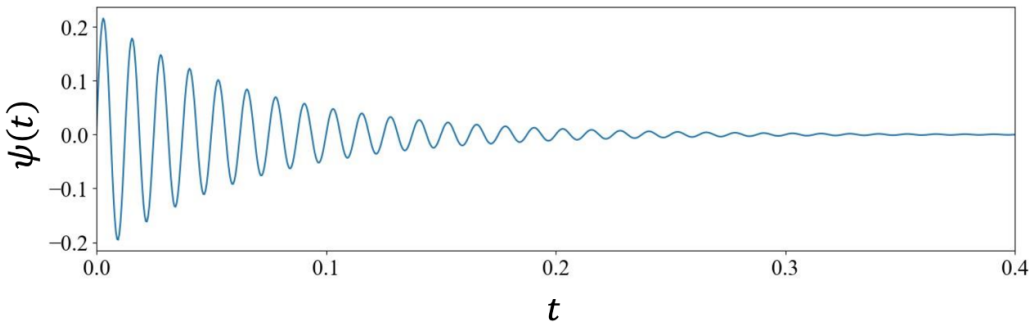


Figure 2. Illustration of the Laplace wavelet basis function.

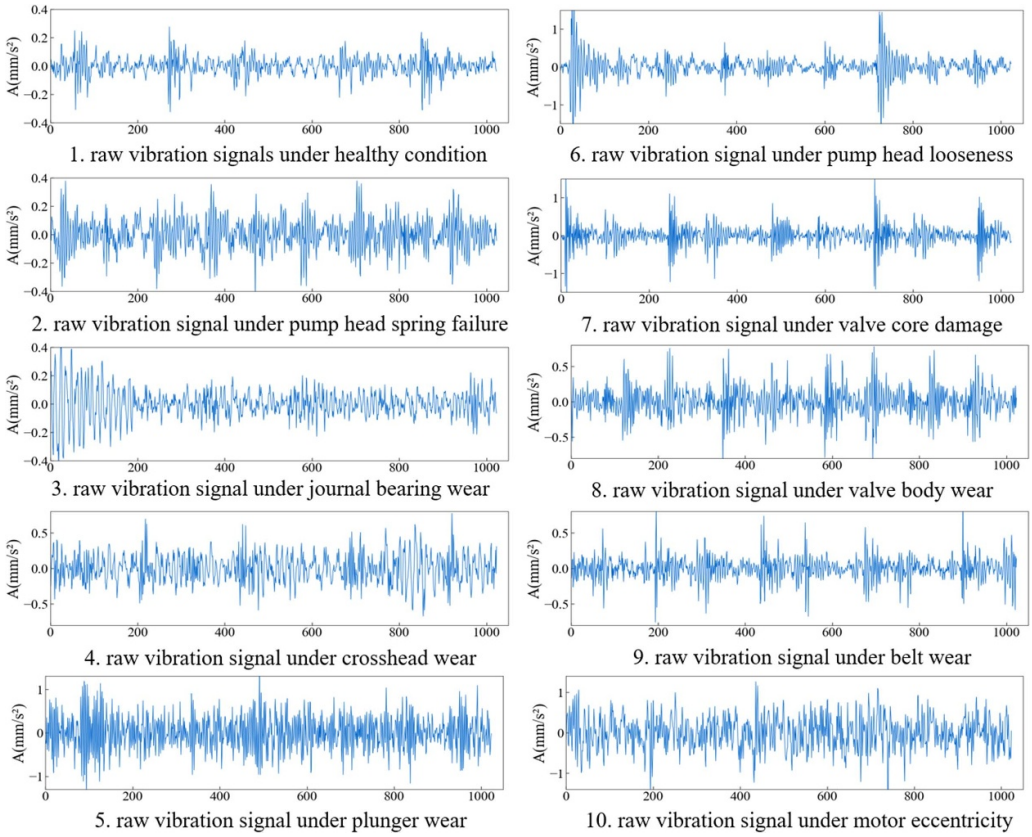


Figure 3. Time-domain vibration signals under different fault conditions.

As a result, a single-scale convolutional kernel struggles to capture features across multiple scales adequately, thus impairing the accuracy and generalization capabilities of model. Additionally, as operational parameters of the water injection pump fluctuate, the scale of fault characteristics also varies, rendering a single-scale convolutional kernel ineffective in adapting to these dynamic changes. This dynamic situation heightens noise sensitivity and compromises the reliability of fault diagnosis. Therefore, after performing initial feature extraction using LWC, the proposed method incorporates MCAM. The whole process is demonstrated in figure 4.

Specifically, MCAM implements convolutional operations utilizing kernels of varying sizes [26, 27], thereby facilitating the hierarchical extraction of features from fault signals and effectively capturing fault information present in the input signals across different scales, as demonstrated below:

$$M_i = B\left(\delta\left(f_i\left(\boldsymbol{P}\right)\right)\right) \tag{3}$$

where $\boldsymbol{P}$ represents input feature map, f_i represents the ith convolution operation, M_i represents the feature vector obtained by the ith convolution.

Upon the successful extraction of multi-scale feature vectors, it is imperative to address the limitations of conventional methods that rely on fixed-weight feature fusion techniques, such as concatenation or element-wise addition. To enhance the feature fusion process, the convolutional block attention module (CBAM) [28] is implemented. Initially, CBAM utilizes a channel attention mechanism to assess the significance

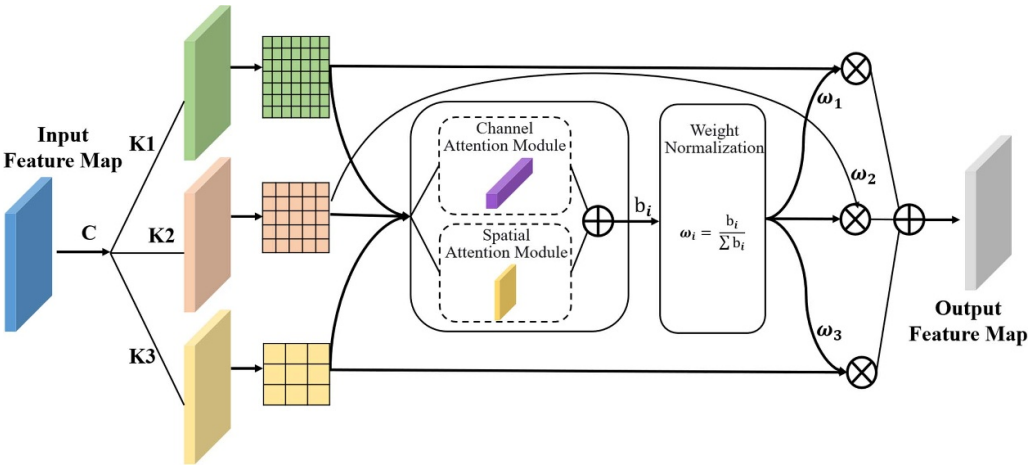


Figure 4. Multi-scale convolutional attention mechanism.

of each channel. Subsequently, a spatial attention mechanism is employed to further refine the contribution of each spatial location. This dual attention mechanism dynamically adjusts the weight relationships among the various features, enabling the network to adaptively prioritize the most salient features while suppressing less relevant information, thereby augmenting the effectiveness of feature fusion and enhancing the overall performance of the network. The weights of M_i after processing by CBAM can be expressed as follows:

$$b_i = M_c(M_i) \oplus M_s(M_i) \tag{4}$$

where M_c represents channel attention mechanism, M_s represents spatial attention mechanism, $\oplus$ represents broadcast addition operation.

In addition to improving the discriminative capability of feature fusion, the incorporation of channel and spatial attention mechanisms offers valuable physical interpretability for fault diagnosis. Specifically, the weights assigned by the attention modules can be regarded as quantitative indicators reflecting the diagnostic importance of features extracted at different scales or locations. For vibration signals in water injection pumps, the pulse-like transients captured in the preprocessing stage (by the LWC) are closely associated with underlying mechanical faults. The subsequent attention mechanism, by adaptively assigning higher weights to those features corresponding to impulse events, provides an interpretable measure of the significance of these transient phenomena in the diagnosis process.

As a result, the model not only benefits from enhanced diagnostic accuracy but also enables the quantitative assessment of the relevance of pulse-type features, facilitating a more transparent understanding of the fault characteristics embedded in the raw signals. This interpretability bridges the gap between data-driven model outputs and physical signal properties, offering practitioners actionable insights into which signal components most strongly influence the final diagnostic decision.

Furthermore, to mitigate the risk of certain feature vectors exhibiting disproportionately large weights that may induce excessive dependence and contribute to overfitting, the LMFE integrates a weight normalization operation to ensure that the contributions of all feature vectors to the model remain within a reasonable range, as illustrated below:

$$Q = \sum \frac{b_i}{\sum b_i} \otimes M_i \tag{5}$$

where $\otimes$ represents pointwise multiplication operation, b_i represents the initial weights of the feature vector obtained by the ith convolution, $\boldsymbol{Q}$ represents final output feature map.

The mentioned plug-and-play MCAM module, as detailed above, replaces traditional fixed-size convolution and pooling operations within the network and, in conjunction with LWC, constitutes the proposed LMFE.

2.2.2. DCL-based metric learning module. In the pretraining stage, the DCL-based metric learning module aims to further optimize the distribution of samples within the embedding space, enhancing the discriminability among the embedded vectors of different classes. Traditional metric loss functions, such as contrastive loss [29] and triplet loss [30], typically focus on inter-sample distances without directly optimizing intra-class compactness and inter-class separability. Consequently, their limitations become particularly apparent in scenarios characterized by high noise levels, substantial intra-class variance, or complex data distributions. For instance, contrastive loss depends on pair construction and is easily biased in imbalanced data [31], negatives are mostly trivial majority samples, while hard-negative mining can overfit to outliers/noise. The pairwise objective shapes only local separations, offering little global control over class margins or intra-class compactness, and is prone to false negatives when classes partially overlap, while triplet loss training hinges on costly triplet mining [32], most triplets are uninformative without mining, but aggressive mining overemphasizes noisy cases. Its constraints are local, often yielding fragmented intra-class structure and under-specified decision boundaries.

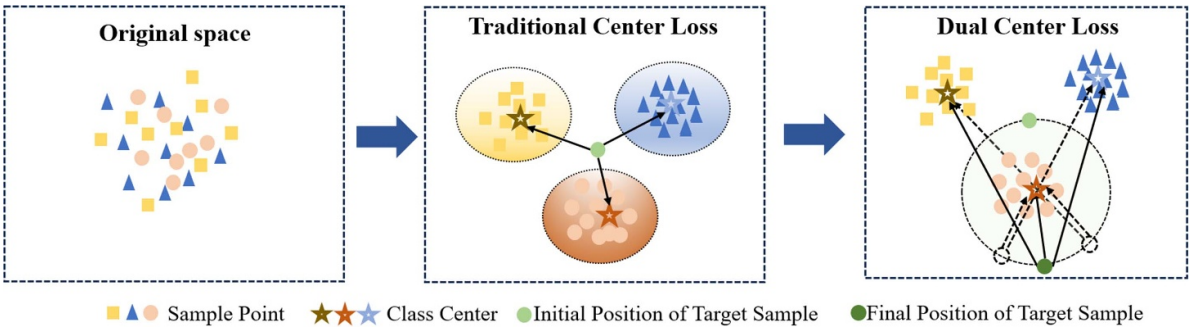


Figure 5. Illustration for dual center loss.

Performance is highly sensitive to batch margins, and long-tailed classes lack stable anchor–positive pairs, causing high variance and overfitting.

In contrast, center loss [33] adopts a class-centric perspective by introducing a center point for each class to optimize the clustering compactness of samples belonging to that class, thereby capturing global information about the categories more effectively. However, traditional center loss has certain shortcomings in constraining inter-class distances and may lead to ambiguous decision boundaries between specific classes. To address this limitation, the proposed metric learning module based on DCL can enhance both intra-class compactness and inter-class separability, thereby improving the classification performance and robustness of the model in complex scenarios.

Fundamentally, traditional center loss functions as an intra-class loss L_{intra} by minimizing the Euclidean distance between samples and their corresponding class centers, ensuring that samples of the same class cluster tightly within the embedding space and reducing intra-class variance, as mathematically expressed as follows:

$$L_{\text{intra}} = \frac{1}{N}\sum_{i=1}^{N} \|C_{y_i} - E_i\|_2 \tag{6}$$

where N represents batch size, Y_i represents the label of sample X_i, C_{Y_i} represents the embedding vector of the class center that the sample belongs to in the embedding space, E_i represents the embedding vector of the sample in the embedding space.

The proposed DCL builds upon traditional L_{intra} by incorporating an additional inter-class loss constraint L_{inter}, utilizing cosine distance as the metric. This approach aims to constrain the angular relationship between sample points and their corresponding class centers, ensuring that the sample points are positioned close to the center of their respective classes (i.e. positive class center) while maximizing their distance from the centers of classes that are not associated with the labels (i.e. non-labeled classes). Consequently, this mechanism reduces the overlap between the distributions of different classes. A geometric illustration of L_{inter} in a two-dimensional plane is presented in figure 5.

Under the condition that the L_{intra} remains constant, the distance from the sample points to the positive class center is fixed. At this point, the angle between the vector connecting the sample points to the positive class center and the vector to the negative class center reflects the relative spatial positioning of the sample points within the embedding space. The definition of cosine similarity is as follows:

$$\cos(\theta) = \frac{(E_i - C_{y_i})^{\top}(E_i - C_m)}{\| E_i - C_{y_i}\|_2 \, \| E_i - C_m\|_2}. \tag{7}$$

Cosine distance exhibits a negative correlation with cosine similarity. Therefore, a smaller angle indicates a reduced cosine distance between the vectors, signifying that the sample points are further from the negative class center. Conversely, a larger angle results in an increased cosine distance, indicating proximity to the negative class center. Notably, when the angle approaches 0°, the sample point aligns along the extended line connecting the positive and negative class centers, resulting in the maximum distance to the negative class center.

Consequently, during the optimization iterations, as the cosine distance-based L_{inter} decreases, the sample points move progressively further away from the negative class center, thereby clarifying the decision boundaries and effectively enhancing the discriminability between classes. Ultimately, the mathematical expression of the weighted L_{inter} can be written as:

$$L_{\text{inter}} = \sum_{i=1}^{N}\sum_{m\neq y_i} k_m \left(1 - \cos(E_i - C_{y_i}, E_i - C_m)\right) \tag{8}$$

Furthermore, considering that fault diagnosis of water injection pumps typically involves multi-class classification, this study innovatively introduces a normalized weight term to further adjust the cosine distance between sample points and various negative class centers. Specifically, the weight term allows negative class centers that are closer to exert a greater influence on the overall L_{inter}, while those that are farther away have a relatively diminished impact. This weighting design enables the model to prioritize the differentiation of categories that are nearer to the current sample points, thereby enhancing optimization efficiency while maintaining methodological rigor. The mathematical representation of the weights is presented as follows:

$$k_m = \frac{e^{-\|C_m - E_i\|_2}}{\sum_{j\neq y_i} e^{-\|C_j - E_i\|_2}} \tag{9}$$

where $\boldsymbol{E}_i$ represents feature vector of the sample point in the embedding space, $\boldsymbol{C}_{y_i}$ represents feature vector of the positive-class center in the embedding space, $\boldsymbol{C}_m$ represents feature vector of the negative-class center in the embedding space, k_m represents weight corresponding to the loss term of the sample point.

L_{intra} and L_{inter} complement each other, and through their synergistic interaction, the intra-class compactness of each fault category is enhanced, the decision boundaries between classes are broadened, and the overall distribution of samples within the embedding space is further optimized.

2.3. Sampling adaptation stage

In contrast to directly applying techniques such as sample generation, data augmentation, or down-sampling on the original imbalanced dataset, fine-tuning the already pre-trained model using the sampling adaptation mechanism allows for the effective utilization of data while mitigating the risk of overfitting without significantly increasing the training burden [34]. This approach proves to be an effective solution for water injection pump fault diagnosis scenarios, as it enhances both the stability and performance of the model, thereby demonstrating considerable engineering practicality.

In the sampling adaptation phase of the fault diagnosis algorithm proposed in this research, the CBBSM is employed to partition the original dataset, thereby enhancing the model's ability to recognize minority class fault samples. The model is then fine-tuned using the newly partitioned dataset. Specifically, the original water injection pump fault dataset D is divided into M subsets based on the class labels y_i, with each subset D_m containing all samples belonging to class m, which can be expressed as follows:

$$D_m = \{(x_i, y_i) \in D | y_i = m\}, \forall m \in \{1, 2, \ldots, M\}. \tag{10}$$

Subsequently, to achieve a balanced sampling of samples from each subset D_m within each batch, the inverse of the number of samples in each class is employed as the sample weight. Accordingly, the weight W_m corresponding to the samples in class m is defined as follows:

$$W_m = \frac{1}{|D_m|}. \tag{11}$$

The weight vector W corresponding to the dataset D is illustrated as follows, and it contains the sampling weights for each sample $\{(x_i, y_i)\}$.

$$W = \{W_{y_i}\}_{i=1}^{|D|}. \tag{12}$$

Regarding batch l, the CBBSM randomly samples N samples from the dataset according to the weight vector W, resulting in a sample set for updating the model parameters, as shown below:

$$B_l = \{(x_i, y_i) \mid (x_i, y_i) \sim R(\mathcal{D}, W, N)\} \tag{13}$$

where $R(\cdot)$ represents random sampling function.

2.4. Model training process

In the pre-training stage, the overall loss function of the model is a weighted sum of the classification loss and the DCL. To mitigate the influence of outliers on the overall loss function, a hyperparameter α is introduced to constrain the scale of the loss function. Therefore, L_{total} can be defined as follows:

$$L_{\text{total}} = L_{\text{cls}} + \alpha L_{\text{dc}} \tag{14}$$

L_{cls} evaluates the effectiveness of classification by computing the cross-entropy loss between the classifier output and the true fault labels, which can be calculated as follows:

$$L_{\text{cls}} = \sum_{i=1}^{N} \sum_{m=1}^{M} y_i \log p_{im} \tag{15}$$

where, M represents number of classes, p_{im} represents the probability that the model predicts sample x_i belongs to class m.

The L_{dc} incorporates both the intra-class loss and inter-class loss introduced in section 2.2.2, with a hyperparameter β implemented to achieve balance among the magnitudes. Notably, employing the mean of the embedded features from the entire training set as class centers can impose significant computational burdens, which necessitates defining these class centers as learnable parameters that are updated in each batch.

$$L_{\text{dc}} = L_{\text{intra}} + \beta L_{\text{inter}}. \tag{16}$$

In the sampling adaptation stage, no calculations or updates are conducted on the class centers within the metric learning module. Additionally, the parameters of LWC in the feature extractor remain fixed, allowing only the parameters of the MCAM and the classifier to be updated. The comprehensive training process of the model is detailed in algorithm 1.

3. Experiments

3.1. Dataset and parameter settings

The dataset used in this experiment was collected from a three-plunger reciprocating water-injection pump at an oilfield in East China, as shown in figure 6.

During data acquisition, the injection pump was operated at a working pressure of 60 MPa and a working frequency of 50 Hz. A vibration sensor was mounted on the pump head, with a sampling frequency of 8 192 Hz. The dataset comprises operating data of the injection pump under 10 health conditions; detailed fault-label information and the number of samples used in the training and testing phases are given in table 1.

In the experiments, the hyperparameters α and β during the pre-training phase were set to 0.05 and 10, respectively. The Adam optimizer was utilized to optimize the feature extractor, metric learning module, and classifier. The pre-training phase consisted of 300 epochs with an initial learning rate of 0.01. To mitigate overfitting, a learning rate scheduler was applied, halving the learning rate if no decrease in loss was observed for

Algorithm 1. Training process of an imbalanced-dataset fault diagnosis model integrating metric learning and sampling-adaptation mechanisms.

Input:
Training dataset $\mathcal{D} = \{X, Y\}$
Hyperparameter α
Number of pre-training iterations $epoch_{\text{pre}}$
Pre-training learning rate for feature extractor & classifier η_1
Pre-training learning rate for class centers η_2,
Number of sampling-adaptation iterations $epoch_{\text{sa}}$,
Sampling-adaptation learning rate for feature extractor & classifier η_3
Output:
Optimized feature extractor $\widetilde{\mathscr{F}}$, Optimized classifier parameters $\widetilde{\mathcal{G}}$
Training process:
Initialize randomly Feature extractor $\mathscr{F}$ parameters $\theta_{\mathscr{F}}$, Classifier $\mathcal{G}$ parameters $\theta_{\mathcal{G}}$, Class centers $\mathcal{C} = \{\mathcal{C}_m\}_{m=1}^{M}$;
Split the training set $\mathcal{D}$, into K mini−batches $\mathscr{B}^p = \{\mathscr{B}_k\}_{k=1}^{K}$ $ofsize N$
while $t < epoch_{\text{pre}}$ do #Pre-training phase
  for $\mathscr{B}_k$ in $\mathscr{B}^p$ do
    For each sample X_i, compute embedding $Z_i = \mathscr{F}(X_i, \theta_{\mathscr{F}})$
    for embedding Z_i, predict class probabilities by classifier $\hat{Y}_i = \mathcal{G}(Z_i, \theta_{\mathcal{G}})$
    compute L_{class} by equations (3)–(9), compute L_{center} by equations (3)–(10), compute L_{total} by equations (3)–(11)
    Update parameters of Feature extractor $\theta_{\mathscr{F}} \leftarrow \theta_{\mathscr{F}} - \eta_1 \cdot \nabla_{\theta_{\mathscr{F}}} L_{\text{total}}$
    Update parameters of classifier $\theta_{\mathcal{G}} \leftarrow \theta_{\mathcal{G}} - \eta_1 \cdot \nabla_{\theta_{\mathcal{G}}} L_{\text{class}}$
    Update class center parameters $\mathcal{C} \leftarrow \mathcal{C} - \eta_2 \cdot \nabla_{\mathcal{C}} (\alpha L_{\text{center}})$;
  end for
end while
Obtain the feature extractor parameters $\tilde{\mathscr{F}}$, classifier parameters $\tilde{\mathcal{G}}$, class center parameters $\tilde{\mathcal{C}}$ optimized during the pre-training phase
while $t < epoch_{\text{sa}}$ do # sampling-adaptation phase
  Using equations (4)–(6), sample $\boldsymbol{L}$ batches from the training set $\mathcal{D}\,\mathscr{B}^s = \{\mathscr{B}_l\}_{l=1}^{L}$;
  for $\mathscr{B}_l$ in $\mathscr{B}^s$ do
    for each sample in batch X_i, generate embedding $Z_i = \tilde{\mathscr{F}}(X_i, \theta_{\tilde{\mathscr{F}}})$ by feature extractor

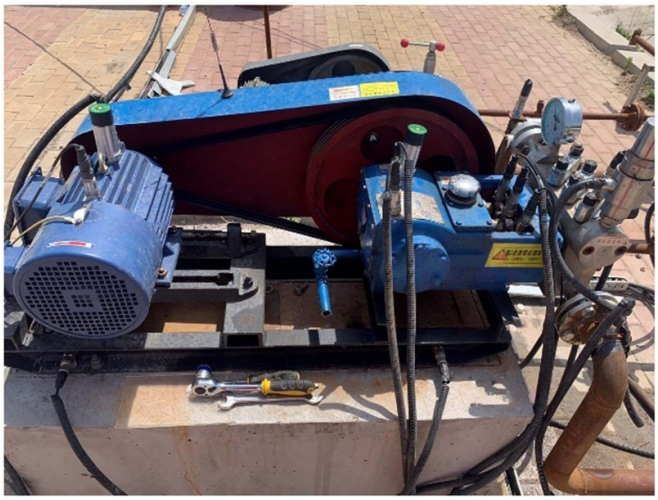


Figure 6. Three-piston reciprocating water injection pump.

15 consecutive epochs. During the sampling adaptation phase, the entire dataset was again utilized as the training set, using the Adam optimizer with a learning rate of 0.001, and training for 50 epochs. To simulate a fault diagnosis scenario in a real industrial environment, Gaussian white noise with a signal-to-noise ratio (SNR) of 1 is added to the original vibration signals.

Table 1. Details of the water injection pump dataset.

Label	Fault type	Training samples	Testing samples
C0	No Fault	2200	550
C1	Pump head spring failure	2200	550
C2	Journal bearing wear	2200	550
C3	Crosshead wear	2200	550
C4	Plunger wear	2200	550
C5	Pump head looseness	2200	550
C6	Valve core damage	1500	375
C7	Valve body wear	800	200
C8	Belt wear	800	200
C9	Motor eccentricity	800	200

In the experiments, the batch size was set to 64, and the model's detailed network parameters are listed in table 2.

To investigate the influence of the hyperparameter α, we performed a series of sensitivity experiments to balance the cross-entropy classification loss and the dual-center metric loss. The results are shown in figure 7. When α is relatively small, the impact of the metric loss is weak and the model focuses mainly on classification accuracy, while the structural constraints on the embedding space are insufficient. As a result, the feature distribution remains blurred, inter-class overlap is obvious, and misclassification is more likely,

Table 2. Parameter settings and network layer of the model.

Network layer	Parameter settings	Output size
Input signal	/	(64, 1024)
Multi–scale convolutional attention layer	Kernel size: 3/5/7 Output channels: 16	(64, 16, 1024)
Pooling layer	Pool size: 4	(64, 16, 256)
Multi–scale convolutional attention layer	Kernel size: 3/5/7 Output channels: 32	(64, 32, 256)
Pooling layer	Pool size: 4	(64, 32, 64)
Multi–scale convolutional attention layer	Kernel size: 3/5/7 Output channels: 64	(64, 64, 64)
Convolution layer	Kernel size: 3 output channels: 32	(64, 32, 64)
Convolution layer	Kernel size: 3 output channels: 16	(64, 16, 64)
Pooling layer	Pool size: 4	(64, 16, 16)
Flatten layer	/	(64, 256)
Fully Connected layer	Output neurons: 100	(64, 100)
Output layer	Output neurons: 6	(64, 6)

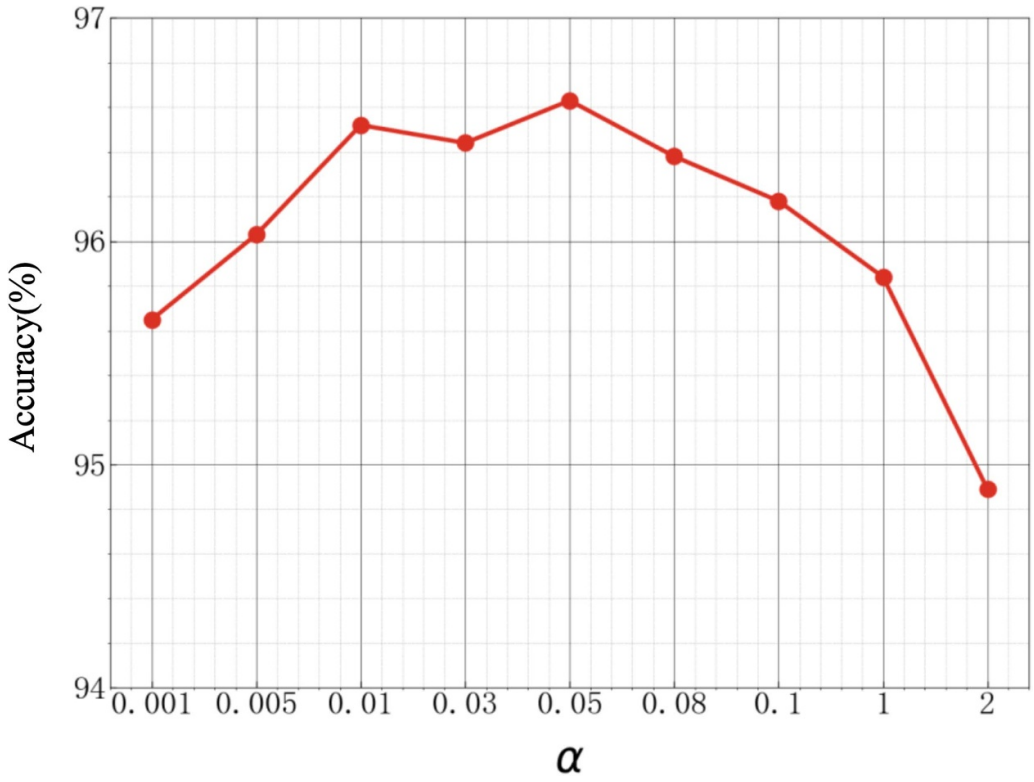


Figure 7. Effect of hyperparameter α on model accuracy.

especially for complex samples. Conversely, when α is too large, the model relies excessively on the metric loss, over-emphasizing intra-class compactness and inter-class separation. This not only compresses feature diversity within classes but may also spread class boundaries too far apart, leading to sparse embeddings and a reduced ability to describe boundary samples. In extreme cases, such over-optimization damages the adaptability of the model to the classification task and introduces redundant representations, which reduces efficiency and practical performance. Taking these findings into account, we determined that setting $\alpha = 0.05$ achieves the best trade-off, with the model reaching an accuracy of 97.11% while effectively balancing classification objectives and feature-space structuring.

3.2. Comparative experiments

In this section, accuracy, precision, recall, and the $F1$ score are selected as evaluation metrics. The comparative experiments include four state-of-the-art fault-diagnosis methods for

Table 3. Results of comparison experiments.

Model	Accuracy	Precision	Recall	$F1$ Score
CNN-BiLSTM	0.9381	0.9371	0.9566	0.9399
MSCNN	0.9421	0.9452	0.9463	0.9461
SMOTE-CNN	0.9372	0.9359	0.9428	0.9322
MWKCNN	0.9453	0.9427	0.9446	0.9446
WavCapsNet	0.9527	0.9321	0.9536	0.9421
The proposed model	**0.9711**	0.9732	0.9695	0.9704

mechanical equipment: the multi-scale CNN (MSCNN), the CNN–bidirectional long short-term memory network (CNN-BiLSTM), the SMOTE-sampled CNN (SMOTE-CNN), the multi-wavelet-kernel CNN (MWKCNN) [35] and wavelet-capsule network (WavCapsNet) [36].

The experimental results in table 3 show that the method proposed in this paper outperforms all comparison methods across every metric. Specifically, it achieves an accuracy of 97.11%, representing improvements of 2.90% and 2.71% over the traditional CNN-BiLSTM and MSCNN algorithms, respectively; a 3.38% gain over the SMOTE-CNN that uses sampling to address data imbalance; and a 2.28% increase compared to the state-of-the-art MWKCNN integrating multiple wavelet convolution kernels. On the real imbalanced injection-pump dataset, this method delivers the best overall classification performance.

The detailed classification confusion matrix of the proposed method is shown in figure 8. From the results, it can be seen that, except for C8 (belt wear), whose diagnostic accuracy is relatively lower, the model achieves over 95% accuracy on all other fault categories. The reduced performance on C8 can be attributed to the inherently weak and gradual vibration characteristics of belt wear, which are more easily obscured by background noise compared with the strong impact signatures of bearing or gear faults. Moreover, the spectral features of belt wear partially overlap with those of other rotating faults, leading to confusion in the embedding space, while the relatively limited number of C8 samples constrains the model's ability to capture sufficient intra-class variability. Together, these factors explain the increased misclassification rate for belt wear. Overall, despite this challenge, the method proposed in this paper delivers excellent fault-diagnosis performance on the real injection-pump dataset.

In addition, to more intuitively evaluate the model's performance in feature extraction and classification at different stages, this section employs the T-SNE algorithm for data visualization, as shown in figure 9.

In the T-SNE plot of the raw data, samples from different classes are scattered, highly mixed, and show large intra-class dispersion, indicating severe intra-class variance and inter-class overlap in the high-dimensional space, which makes direct classification difficult.

After feature extraction and optimization by the metric learning module, the T-SNE visualization in the embedding space shows a clear change: samples from the same class gradually cluster together, while the boundaries between different classes become clearer. This demonstrates that the

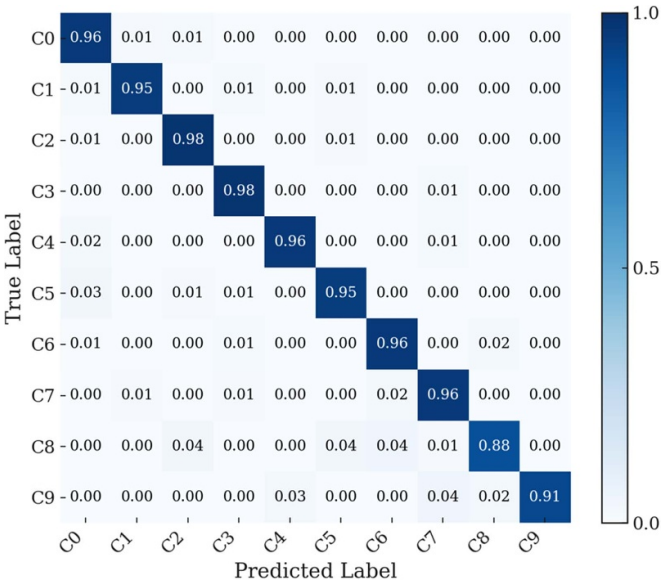


Figure 8. Confusion matrix of proposed method.

model effectively captures intrinsic class-specific features while enhancing inter-class separability, confirming the role of the dual-center loss in improving the embedding distribution.

Finally, in the T-SNE visualization of the fully connected layer, data points form compact clusters with further enlarged inter-class margins. The separability between health states reaches its best, indicating that the model ultimately achieves strong clustering and class separation, enabling accurate fault recognition and classification.

To further investigate the contribution of each component in the proposed method, we conducted ablation studies. LMFE denotes the baseline model that retains only the feature extractor while removing both the metric learning module and the sampling-adaptation mechanism. LMFE-CBBSM denotes the model without the metric learning module, and LMFE-DCL denotes the model without the sampling-adaptation mechanism.

As shown in table 4, when only the proposed feature extractor is employed, the model achieves an accuracy of 96.06%, already surpassing the comparison models, demonstrating that this module greatly enhances fault feature extraction capability. On this basis, incorporating the metric learning module and the sampling-adaptation mechanism further improves accuracy by 0.75% and 0.49%, respectively. When both modules are integrated simultaneously, the accuracy improves by 1.05%, which verifies the effectiveness of each component in the proposed framework.

3.3. Impact of the degree of dataset imbalance on model performance

To investigate the effect of dataset imbalance on model performance, this subsection randomly samples the original injection-pump dataset at different imbalance ratios for comparative experiments. Here, the imbalance ratio is defined as

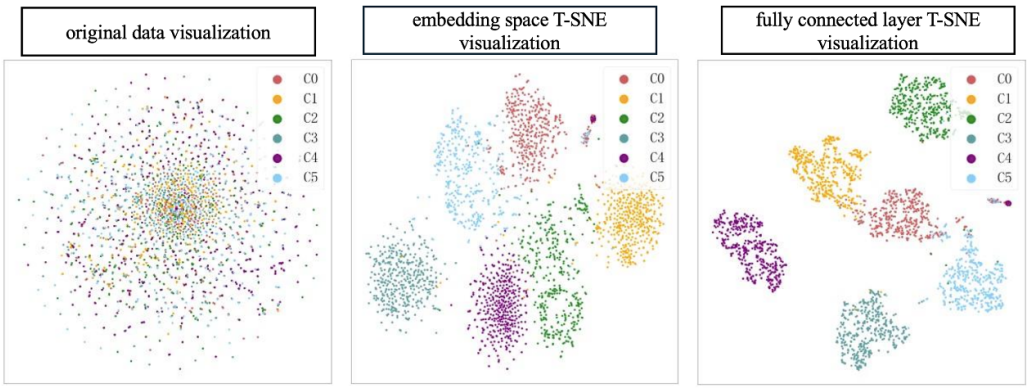


Figure 9. T-SNE visualization of feature extraction and classification performance at different stages.

Table 4. Results of ablation experiments.

Model	Accuracy	Precision	Recall	$F1$ Score
LMFE	0.9606	0.9647	0.9601	0.9613
LMFE-CBBSM	0.9655	0.9676	0.9657	0.9663
LMFE-DCL	0.9681	0.9744	0.9653	0.9697
The proposed model	0.9711	0.9732	0.9695	0.9704

the ratio of the number of samples in the majority classes (C0–C5) to that in the minority classes (C6–C9). In the experiments, the test-set sample counts for all classes are kept the same as before. The training-set sample count for the majority classes is held constant at 2200, while the training-set sample count for the minority classes varies with the imbalance ratio.

Figure 10 shows the diagnostic accuracy and $F1$ score for each model, the model TripletNet [37] using traditional metric learning method is also introduced to compare. As the imbalance ratio increases, both accuracy and $F1$ score of all models exhibit a downward trend. However, under all imbalance conditions, the proposed method outperforms the comparison models-especially at high imbalance ratios, where its diagnostic performance advantage is most pronounced-demonstrating the robustness of our model and its capability to address imbalanced fault diagnosis in real-world injection-pump scenarios.

Among all the comparison methods, the classical fault diagnosis model CNN-BiLSTM showed the most pronounced performance degradation. When the SNR dropped to 3 dB or below, its accuracy and $F1$-score rapidly fell below 70%, indicating that traditional deep learning models have limited robustness under strong noise interference. In contrast, TripletNet and DSML, both metric-learning–based methods, performed relatively better, demonstrating the effectiveness of metric learning in fault diagnosis tasks, especially under noisy conditions where noise resistance was partially improved. Notably, the proposed dual-center-loss–based metric learning method exhibited strong noise robustness under all SNR levels, with smaller declines in accuracy and $F1$-score. Even when the SNR was reduced to 0 dB, our method still maintained a clear advantage, outperforming the current best baseline by 2.88% in accuracy and 3.22% in $F1$-score. These results verify the superiority of the proposed method and highlight its potential for practical engineering applications in complex noisy environments.

3.4. Model performance evaluation on public datasets

To comprehensively evaluate the model's generalization capability, this subsection conducts detailed performance tests using the publicly available centrifugal multistage impeller blower dataset [38, 39]. The physical test rig, shown in figure 11, is a well-established experimental platform that simulates typical rotating machinery faults under controlled operating conditions, making it suitable for benchmarking intelligent fault-diagnosis methods. This dataset contains five health states: healthy condition (C0), bearing outer-race fault (C1), bearing inner-race fault (C2), bearing rolling-element fault (C3), and gear fault (C4). Among them, C0–C2 are defined as majority-class samples, while C3 and C4 are relatively rare and serve as minority-class samples, which closely reflects the class-imbalance characteristics encountered in real-world industrial scenarios. The vibration signals are collected using acceleration sensors installed on the blower housing, covering both normal and fault operating states under stable speed and load conditions.

Detailed fault-label information and the number of training and testing samples for each category are summarized in table 5. Specifically, each majority class (C0–C2) contains 720 training samples and 180 testing samples, while each minority class (C3–C4) includes 240 training samples and 60 testing samples. This balanced experimental design allows us to rigorously assess the model's performance not only on dominant classes but also on underrepresented fault types, thereby highlighting its robustness in imbalanced learning scenarios.

In the generalization experiments, the hyperparameters α and β during the pre-training phase were set to 0.05 and 10, respectively. The Adam optimizer was utilized to optimize the feature extractor, metric learning module, and classifier. The pre-training phase consisted of 300 epochs with an initial learning rate of 0.01. To mitigate overfitting, a learning rate scheduler was applied, halving the learning rate if no decrease in loss was observed for 15 consecutive epochs.

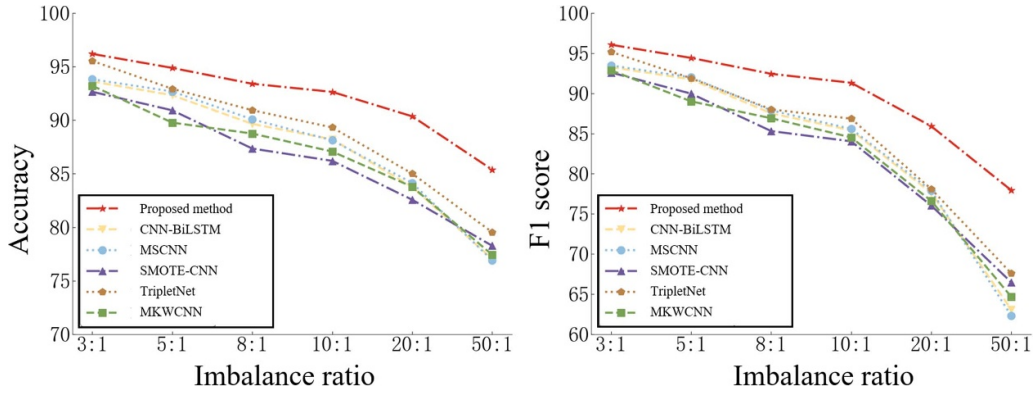


Figure 10. Performance of the models at different imbalance rates.



Figure 11. Test rig of centrifugal multi-stage impeller blower.

Table 5. Details of the public dataset.

Label	Fault type	Training samples	Testing samples
C0	No fault	720	180
C1	Bearing outer race fault	720	180
C2	Bearing inner race fault	720	180
C3	Bearing Rolling Element Fault	240	60
C4	Gear fault	240	60

During the sampling adaptation phase, the entire dataset was again utilized as the training set, using the Adam optimizer with a learning rate of 0.001, and training for 50 epochs. Furthermore, to simulate fault-diagnosis scenarios in real industrial environments, Gaussian white noise with a SNR of 1 was added to the raw vibration signals. The experimental results are presented in table 6. The proposed method achieved an accuracy of 99.29%, representing improvements of 3.52%, 2.71%, 4.14%, and 0.81% over CNN-BiLSTM, MSCNN, SMOTE-CNN, and MWKCNN, respectively, and also attained the highest performance on all other evaluation metrics, fully demonstrating its effectiveness on the public dataset.

Table 6. Results of comparison experiments on public dataset.

Model	Accuracy	Precision	Recall	*F*1 Score
CNN-BiLSTM	0.9577	0.9570	0.9577	0.9578
MSCNN	0.9658	0.9661	0.9658	0.9658
SMOTE-CNN	0.9515	0.9520	0.9515	0.9514
MWKCNN	0.9632	0.9652	0.9744	0.9542
WavCapsNet	0.9757	0.9498	0.9676	0.9618
The proposed method	0.9929	0.9931	0.9929	0.9929

4. Conclusion

To address the challenges of extracting key fault features in real-world injection-pump scenarios and the severe imbalance in historical datasets, this paper proposes an imbalanced-fault-diagnosis method that integrates metric learning with a sampling-adaptation mechanism. The effectiveness and superiority of this approach are validated on both a real injection-pump dataset and a public dataset. The main contributions are:

1. Development of a feature extractor based on Laplacian wavelet convolution and a MCAM, which effectively captures fault information at different scales of the input signal and adaptively assigns weights, greatly enhancing feature fusion and overall network performance. Experimental results show that the proposed method achieves an accuracy of 99.29% on the public dataset, demonstrating its ability to handle real-world industrial scenarios with added Gaussian noise, and outperforming traditional models in accuracy.
2. Introduction of a metric learning module under the metric-learning framework, driven by a dual-center loss. By adding inter-class loss terms for samples, it effectively expands decision boundaries in the embedding space, improving the model's fault-distinction capability in complex data environments. Under a severe imbalance ratio of 50:1, the model achieves a fault diagnosis accuracy of 86.94%, showcasing its robustness and superior performance even under challenging data imbalance conditions.
3. Proposal of a CBBSM that, without adding excessive computational burden, significantly reduces the negative impact

of dataset imbalance on model performance. This mechanism offers an important solution for injection-pump fault diagnosis in engineering applications. On a real injection-pump dataset, the method achieves 97.11% accuracy, demonstrating clear advantages over mainstream models, further confirming the proposed method's efficacy in real-world applications.

Future work will extend to injection-pump fault diagnosis in multi-sensor scenarios. Building on this method, the model will be further optimized to enrich its ability to process multi-source information, further enhancing its value for diagnosing faults in industrial-site injection pumps.

Funding

This research is supported by National Natural Science Foundation of China (72571173), Natural Science Foundation of Shanghai (25ZR1401196)

ORCID iD

Tangbin Xia 0000-0001-9121-1716

References

[1] Zhou X, Van Gelder P H A J M, Liang Y and Zhang H 2020 An integrated methodology for the supply reliability analysis of multi-product pipeline systems under pumps failure *Reliab. Eng. Syst. Saf.* **204** 107185
[2] Huang H *et al* 2025 Theory and application of possibility and evidence in reliability analysis and design optimization *J. Reliab. Sci. Eng.* **1** 15007
[3] Zio E 2022 Prognostics and health management (PHM): where are we and where do we (need to) go in theory and practice *Reliab. Eng. Syst. Saf.* **218** 108119
[4] Jin T and Abedin A 2025 Co-allocation of reliability-redundancy, spare parts and spacecraft trajectory for lunar habitation *J. Reliab. Sci. Eng.* **1** 15005
[5] Xia T, Dong Y, Xiao L, Du S, Pan E and Xi L 2018 Recent advances in prognostics and health management for advanced manufacturing paradigms *Reliab. Eng. Syst. Saf.* **178** 255–68
[6] Xia T, Xing X, Yan T, Wang D, Pan E and Xi L 2024 Interpretable temporal degradation state chain based fusion graph for intelligent bearing fault detection *Adv. Eng. Inform.* **59** 102342
[7] Wang J, Zhang Z, Liu Z, Han B, Bao H and Ji S 2023 Digital twin aided adversarial transfer learning method for domain adaptation fault diagnosis *Reliab. Eng. Syst. Saf.* **234** 109152
[8] Tan Q, Mu X, Fu M, Yuan H, Sun J, Liang G and Sun L 2022 A new sensor fault diagnosis method for gas leakage monitoring based on the naive Bayes and probabilistic neural network classifier *Measurement* **194** 111037
[9] Tang S, Zhu Y and Yuan S 2022 A novel adaptive convolutional neural network for fault diagnosis of hydraulic piston pump with acoustic images *Adv. Eng. Inform.* **52** 101554
[10] Wang S and Xiang J 2020 A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps *Soft Comput.* **24** 2983–97
[11] Li G, Hu J, Shan D, Ao J, Huang B and Huang Z 2023 A CNN model based on innovative expansion operation improving the fault diagnosis accuracy of drilling pump fluid end *Mech. Syst. Signal Process.* **187** 109974
[12] Tang S, Zhu Y and Yuan S 2022 Intelligent fault diagnosis of hydraulic piston pump based on deep learning and Bayesian optimization *ISA Trans.* **129** 555–63
[13] Mei J, Zhu M, Liu W, Fu M and Tang Q 2023 Conditional variational encoder classifier for open set fault classification of rotating machinery vibration signals *IEEE Trans. Ind. Inform.* **20** 3038–49
[14] Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H and Bing G 2017 Learning from class-imbalanced data: review of methods and applications *Expert Syst. Appl.* **73** 220–39
[15] Mao W, Liu Y, Ding L and Li Y 2019 Imbalanced fault diagnosis of rolling bearing based on generative adversarial network: a comparative study *IEEE Access* **7** 9515–30
[16] Wang J, Wei J, Huang H, Wen L, Yuan Y, Chen H, Wu R and Wu J 2024 IMWMOTE: a novel oversampling technique for fault diagnosis in heterogeneous imbalanced data *Expert Syst. Appl.* **251** 123987
[17] Wang H, Li P, Lang X, Tao D, Ma J and Li X 2023 FTGAN: a novel GAN-based data augmentation method coupled time–frequency domain for imbalanced bearing fault diagnosis *IEEE Trans. Instrum. Meas.* **72** 1–14
[18] Zhou Y, Gao L, Wang D, Wu W, Zhou Z and Ye T 2023 Imbalanced multi-fault diagnosis via improved localized feature selection *IEEE Trans. Instrum. Meas.* **72** 1–11
[19] Jian C and Ao Y 2023 Imbalanced fault diagnosis based on semi-supervised ensemble learning *J. Intell. Manuf.* **34** 3143–58
[20] Zhao Y, Yang X, Huang J, Gao J and Cui J 2024 Improved weighted extreme learning machine with adaptive cost-sensitive strategy for imbalanced fault diagnosis of rotating machinery *Mech. Syst. Signal Process.* **217** 111526
[21] Mao G, Li Y, Cai Z, Qiao B and Jia S 2023 Transferable dynamic enhanced cost-sensitive network for cross-domain intelligent diagnosis of rotating machinery under imbalanced datasets *Eng. Appl. Artif. Intell.* **125** 106670
[22] Gao W, Xu Z and Akoudad Y 2025 Metric-based meta-learning relation network for cross-domain few-shot bearing fault diagnosis *IEEE Sens. J.* **25** 13632–47
[23] Li T, Zhao Z, Sun C, Cheng L, Chen X, Yan R and Gao R X 2021 WaveletKernelNet: an interpretable deep neural network for industrial intelligent diagnosis *IEEE Trans. Syst. Man Cybern. Syst.* **52** 2302–12
[24] Jiang G, Wang J, Wang L, Xie P, Li Y and Li X 2023 An interpretable convolutional neural network with multi-wavelet kernel fusion for intelligent fault diagnosis *J. Manuf. Syst.* **70** 18–30
[25] Chen J, Huang R, Zhao K, Wang W, Liu L and Li W 2021 Multiscale convolutional neural network with feature alignment for bearing fault diagnosis *IEEE Trans. Instrum. Meas.* **70** 1–10
[26] Xu Y, Chen Y, Zhang H, Feng K, Wang Y, Yang C and Ni Q 2023 Global contextual feature aggregation networks with multiscale attention mechanism for mechanical fault diagnosis under non-stationary conditions *Mech. Syst. Signal Process.* **203** 110724
[27] Li F, Wang L, Wang D, Wu J and Zhao H 2023 An adaptive multiscale fully convolutional network for bearing fault diagnosis under noisy environments *Measurement* **216** 112993
[28] Woo S, Park J, Lee J Y and Kweon I S 2018 CBAM: convolutional block attention module *Proc. European Conf. on Computer Vision (ECCV) (Munich)* (Springer) pp 3–19
[29] Zhang Y, Liu Z and Huang Q 2023 A contrastive learning-based fault diagnosis method for rotating

machinery with limited and imbalanced labels *IEEE Sens. J.* **23** 16402–12
[30] Zhang K, Chen J, He S, Li F, Feng Y and Zhou Z 2022 Triplet metric driven multi-head GNN augmented with decoupling adversarial learning for intelligent fault diagnosis of machines under varying working condition *J. Manuf. Syst.* **62** 1–16
[31] Elharrouss O, Mahmood Y, Bechqito Y, Adel Serhani M, Badidi E, Riffi J and Tairi H 2025 Loss functions in deep learning: a comprehensive review (arXiv:2504.04242)
[32] Qian Z *et al* 2025 Enhanced diagnosis of thyroid-associated eye diseases based on deep learning: a novel triplet loss design strategy *Biomed. Signal Process. Control* **100** 107161
[33] Liu Z, Fu Y, Pan Q and Zhang Z W 2022 Orientational distribution learning with hierarchical spatial attention for open set recognition *IEEE Trans. Pattern Anal. Mach. Intell.* **45** 8757–72
[34] Zhong H, Yu S, Trinh H, Lv Y, Yuan R and Wang Y 2023 Fine-tuning transfer learning based on DCGAN integrated with self-attention and spectral normalization for bearing fault diagnosis *Measurement* **210** 112421
[35] Wang Y, Lei Y, Li N, Li X, Yang B, Gao X and Liu X 2025 Multimodal parallel bilinear co-attention fusion with full-dimensional dynamic wavelet-kernel structure for machinery RUL prediction *Measurement* **247** 116685
[36] Li W, Lan H, Chen J, Feng K and Huang R 2023 WavCapsNet: an interpretable intelligent compound fault diagnosis method by backward tracking *IEEE Trans. Instrum. Meas.* **72** 1–11
[37] Witmer A, Theagarajan R and Bhanu B 2023 Triplet-net classification of contiguous stem cell microscopy images *IEEE/ACM Trans. Comput. Biol. Bioinform.* **20** 2314–27
[38] Liu Z, Zhang J and He X 2023 A discrimination-guided active learning method based on marginal representations for industrial compound fault diagnosis *IEEE Trans. Autom. Sci. Eng.* **21** 6411–22
[39] Liu Z, Zhang J, He X, Zhang Q, Sun G and Zhou D 2022 Fault diagnosis of rotating machinery with limited expert interaction: a multicriteria active learning approach based on broad learning system *IEEE Trans. Control Syst. Technol.* **31** 953–60



Mr Zhiwen Yao
Zhiwen Yao received the BS degree in Mechanical Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2024, where he is currently pursuing a Master's degree in Industrial Engineering and Management with the School of Mechanical Engineering.

His research interests include fault diagnosis and prognostics and health management of industrial equipment using deep learning.



Ms Xiwen Cui
Xiwen Cui received the BS degree in Mechanical Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2022, where she is currently pursuing a Master's degree in Industrial Engineering and Management with the School of Mechanical Engineering.

Her research interests focus on applying deep learning methods for fault detection, prognosis, and health management of industrial machinery.



Prof Nagi Gebraeel
Nagi Gebraeel received MS and PhD degrees in Industrial Engineering from Purdue University, West Lafayette, Indiana, in 1998 and 2003, respectively. Currently, he is a Professor at the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. His research focuses on improving the accuracy of predicting unexpected failures of engineering systems by leveraging sensor-based data streams. His major research interests are in the areas of degradation modeling and sensor-based prognostics, reliability engineering, and maintenance operations and logistics. He is a member of IISE and INFORMS.



Prof Tangbin Xia
Tangbin Xia received a PhD degree in Mechanical Engineering (Industrial Engineering) from Shanghai Jiao Tong University in 2014 and was a co-cultured PhD with the University of Michigan. Currently, he is a Professor and serves as Vice Dean of the School of Mechanical Engineering at Shanghai Jiao Tong University, State Key Laboratory of Mechanical System and Vibration. His major research interests are intelligent maintenance systems, prognostics & health management, and advanced manufacturing. He is a member of IEEE, ASME, IISE and INFORMS.



Prof Lifeng Xi
Lifeng Xi received a PhD degree in Mechanical Engineering from Shanghai Jiao Tong University in 1995. He is currently a Professor and serves as the Executive Vice President of Shanghai Jiao Tong University. He also serves as the Fellow of State Key Laboratory of Mechanical System and Vibration, and the Executive Director of the Chinese Quality Association. He has published more than 200 papers in international journals. His interest areas are in the fields of production system design and planning, quality management, and reliability engineering.