ELSEVIER

Full length article

SCG-GFFE: A Self-Constructed graph fault feature extractor based on graph Auto-encoder algorithm for unlabeled single-variable vibration signals of harmonic reducer

Shilong Sun [a,b,*], Hao Ding [a,b], Zida Zhao [a,b], Wenfu Xu [a,b,d], Dong Wang [c]

[a] *School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, 518055, China*
[b] *Guangdong Key Laboratory of Intelligent Morphing Mechanisms and Adaptive Robotics, Shenzhen, China*
[c] *The State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China*
[d] *Key University Laboratory of Mechanism & Machine Theory and Intelligent Unmanned Systems of Guangdong, China*

ARTICLE INFO

ABSTRACT

As a pivotal component in robotic systems, harmonic reducer fault diagnosis plays a crucial role in safe and stable operation; however, the lack of labelled fault samples hampers its effectiveness. This study introduces a self-constructed graph graph-autoencoder fault feature extractor (SCG-GFFE), a novel method that uses the Graph Auto-encoder (GAE) algorithm. SCG-GFFE leverages Graph Neural Networks (GNNs) for the unsupervised extraction of fault features, enhancing fault diagnosis in scenarios with limited labelled data. This approach overcomes the challenges of graph construction for fault diagnosis, particularly for single-sensor vibration signals. First, we developed the SCG-GFFE method for efficient fault feature extraction, which does not require complex domain expertise and is capable of generating graph structures. Second, we designed a suitably structured subgraph for dealing with single, unlabelled, non-multisource vibration signals. Third, a real-world harmonic reducer's vibrational signal is utilised to build an experimental study and demonstrate the effectiveness of the SCG-GFFE method. The results of SCG-GFFE demonstrate remarkable accuracy in fault diagnosis, consistently exceeding 98% across various classifiers, and indicate that the proposed method can solve the problem of limited labelled data in harmonic reducer diagnosis.

1. Introduction

In the context of Industry 4.0, traditional manufacturing is rapidly shifting towards intelligent and automated processes, with robots being a central focus. Owing to the collaborative nature of robots in industrial production, a single robot malfunction can disrupt the entire workflow, emphasising the need for preemptive maintenance. Harmonic reducers, known for their compactness, precision, high reduction ratio, and low noise, are integral components of robots [1]. Therefore, diagnosing faults in harmonic reducers is crucial owing to their significant role in robotic operations.

Currently, both academia and industry are yet to delve deeply into the field of harmonic reducer fault research. Existing theoretical achievements in this area are also scattered, resulting in a lack of robust theoretical underpinnings for the fault diagnosis of harmonic reducers. Recently, fault-diagnosis research based on deep learning has witnessed significant advancements. It encompasses various techniques such as convolutional neural networks (CNNs) [2], recurrent neural networks (RNNs) [3], autoencoders (AEs) [4], deep belief networks (DBNs) [5], and contrastive learning [6], all of which have demonstrated commendable outcomes. For example, Yao [7] proposed an adversarial domain adaptation network with pseudo-siamese feature extractors (PSFEN) to improve the performance of the cross-bearing fault transfer diagnosis. Wang [8] proposed a refined prototype and correlation weighting Manhattan distance (RPCMN) to solve the few samples and the inevitable noise from the vibration signals. Wang [9] proposed a Brownian correlation metric prototypical network (BCMPN) algorithm to solve the problem of few samples in the training source domain and zero samples in the test target domain (FST-ZST). All these three references investigated the common problem of vibration fault diagnosis, including various working conditions, few fault samples, cross-domains, and so on. However, intelligent fault diagnosis of harmonic reducers
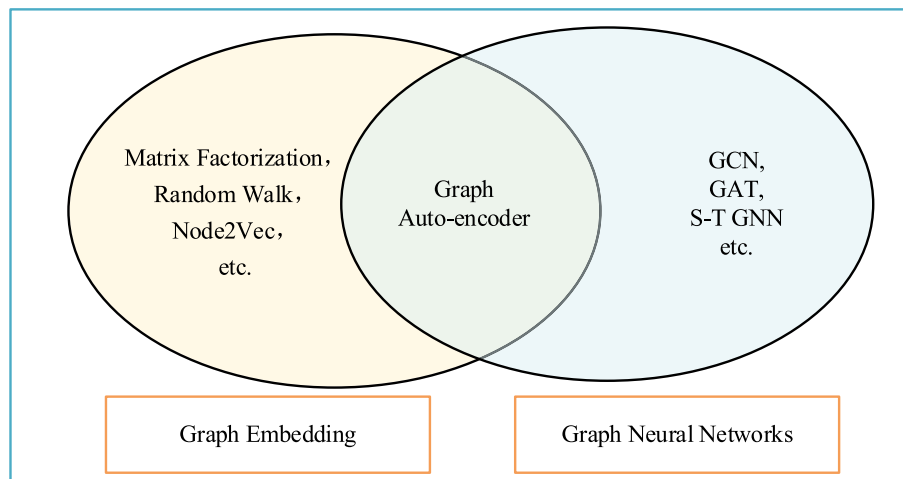

Fig. 1. Graph learning landscape: Interplay between graph embedding and GNN models.

remains limited. Zhi [10] introduced the WRCTD algorithm for denoising sensor data, followed by CNN-LSTM for accurate fault-type diagnosis. Liu [11] applied WRCTD to AE signals and then combined OMA and VMD for fault feature extraction. He [12] addressed multiscale harmonic reducers using MSMCNN. Zhou [13] designed an adaptive 1-D MCNN model for effective contextual feature extraction from continuous signal data. Yang [14] proposed SDP-ConvNeXt for converting vibration signals into SDP images and using ConvNeXt with Transformer for efficient fault classification. Yang [15] used GANs for imbalanced harmonic reducer fault data and MSCNN for high-accuracy multi-class fault classification. In sum, it can be concluded from refers [10] to [15] that the essence of traditional deep learning approaches in fault diagnosis, which often treat sample data as points in Euclidean space and process them as sequences or matrices, primarily through CNNs, focuses on leveraging the models' capability to extract intrinsic features from individual samples. However, these methods usually operate under the assumption that samples are independently and identically distributed, thereby neglecting the intrasample correlation information crucial for diagnosing complex or subtle fault patterns. To overcome these limitations, there's a shift towards advanced models like RNNs, LSTMs, GNNs, attention mechanisms, Transformers, and hybrid models. These approaches aim to better capture temporal or relational dependencies between samples, enhancing the ability to identify faults in complex systems by not only analyzing intrinsic features within individual samples but also considering the dynamic interrelationships among them, thus offering a more nuanced and comprehensive approach to fault diagnosis.

To better understand the interconnections between data samples, more focus is given to graph data and Graph Neural Network (GNN) methods. These methods are effective because they use graph data, where samples are linked through edges, making it easier to represent complex nonlinear relationships, particularly in non-Euclidean spaces. Noteworthy graph neural network architectures include Graph Convolutional Networks (GCNs) [16], GAEs [17], Graph Attention Networks (GATs), graph generative networks, and graph spatiotemporal networks (S-T GNNs) [18]. Currently, the primary domains for GNNs are still within recommender systems [16], bioinformatics [17], transportation and urban planning [18], and social network analysis [19]. However, their applications in fault diagnosis remain relatively limited. Nonetheless, there has been a gradual increase in research in this area in recent years. Zhang [19] proposed a multi-head graph neural network based on triplet metric combined with decoupled adversarial learning. This approach transforms multidimensional vibration signals into graph structures using L2 distance and triplet loss, along with multihead attention mechanisms and multidomain decoupled adversarial learning

to adapt to complex unknown conditions in the target domain, yielding significant performance improvements. Yu [20] introduced a novel rolling bearing fault diagnosis framework. This framework employs a CNN to extract fault features, coupled with dynamic graph embedding and a GNN for capturing fault feature sensitivities under fluctuating conditions. A node-voting mechanism is incorporated to optimise the recognition results. Tang [21] synergistically leveraged the strengths of semi-supervised Conditional Random Fields and Graph Attention Networks, proposing the CRF-GAT algorithm. By modelling label dependencies and learning object representations, this approach facilitates semi-supervised fault diagnosis and effectively addresses real-world problems with limited data. Yang [22] introduced a method based on Deep Capsule Graph Convolutional Networks (DCGCN). By integrating multisensor data spectral analysis and graph convolutional networks, this method effectively diagnoses complex composite faults in industrial robot harmonic transmissions under varying conditions. Chen [23] tackled difficult industrial process fault diagnosis by transforming sensor signals into heterogeneous graphs with multiple edge types. An Interactive-aware Graph Neural Network (IAGNN) was proposed, which adaptively learns edge weights through attention mechanisms. This approach exhibits a superior performance in fault diagnosis across various industrial processes. These studies have mainly focused on multidimensional signals or signals acquired from multiple sensors. They employed techniques based on prior knowledge and signal processing to assess the distances between faulty samples, construct graph data structures, and carefully consider and design the choice and utilisation of graph models for fault diagnosis. Furthermore, most current methods predominantly emphasise supervised learning, relying heavily on label quantity and quality, whereas research on semi-supervised or unsupervised methods to alleviate the demand for labelled samples remains limited.

When addressing the application of GNNs in fault diagnosis, two primary challenges emerge: the construction of graph data and the selection of appropriate graph models [24]. When selecting appropriate graph learning models, a targeted choice must be made based on specific task requirements and the particular characteristics of graph data. In real industrial settings, the volume of data required for intelligent fault diagnosis may be substantial, whereas the actual availability of meaningful data is restricted. In addition, many data points lack corresponding labels for the fault types [25]. Opting for data reacquisition or reannotation may incur substantial additional costs and pose inherent risks. Therefore, the direct application of semi-supervised or unsupervised fault diagnosis methods has profound significance because it can significantly reduce the demand for labelled data [26–28]. As illustrated in Fig. 1, in the realm of graph learning, graph embedding tasks

represent essential unsupervised learning endeavors [29]. In this context, the GAE algorithm is one of the most suitable graph neural network models [30].

Furthermore, because of the absence of inherent relational graphs among the sensor signal data used in fault diagnosis, unlike in social networks, the construction of relationships between data points is of paramount importance. This significantly influences the quality and upper limits of the subsequent graph-based learning outcomes. Only with a well-constructed graph data structure can the advantages of graph neural networks be fully realised to achieve effective feature extraction. Building a suitable graph data structure involves selecting data points as nodes and establishing the connections between them. Generally, for complex systems or mechanical equipment with multiple data sources [31–34], richer data availability makes it easier to explore the relationships between nodes and thereby establish connections. However, collecting multi-source data in practical industrial settings may require diverse types of specialised sensors, resulting in higher costs. Additionally, specific data collection processes may require equipment disassembly to instal sensors in specific locations, leading to increased time and complexity [35,36]. Consequently, in many cases, it is more convenient to collect only the vibration signals during equipment operation. This approach is not only time- and cost-effective but also more straightforward. Nevertheless, for single-vibration signal data, owing to the limited data types, the expressive power of graph data structures is relatively constrained. The space for constructing operable graph data is limited, making it more challenging to devise suitable graph-data structures. Currently, when utilising graph machine learning or GNN methods for fault diagnosis, typical approaches to graph construction can be broadly categorised into two types: those based on K-nearest neighbour (KNN) algorithms [37,38] and those based on prior or expert knowledge [39]. The former method is relatively straightforward but possesses a certain degree of randomness. It relies heavily on the classification effectiveness of KNN for data features, which may result in insufficient reliability. When the features between different data types are not sufficiently distinct and cannot be effectively separated using the KNN method, erroneous connection relationships may easily emerge, introducing significant risks for subsequent graph-based learning. The latter approach requires deep understanding of domain knowledge. It also necessitates specific analyses of different research subjects or scenarios involving the same research subject. However, this approach tends to be complex, challenging, and less universally applicable.

Therefore, this study focuses on exploring a simpler, more straightforward, and universally applicable approach for constructing graph data. Specifically, this study addresses the context of single-variable vibration signals (acquired from a single sensor, with each sample containing only one variable). The objective is to investigate a more direct and general method for constructing graph data, avoiding the limitations of both KNN and domain-specific knowledge approaches. This research strives to create an effective graph data structure without relying on complex signal processing or numerical computations, while still using the GAE network for unsupervised fault feature extraction. This study aims to utilise GNN algorithms for the processing and fault feature extraction of unlabelled, single-source vibration signal data from harmonic reducers. The main contributions of this study are as follows.

(1) We propose a GNN-based fault feature extraction method called SCG-GFFE, which enables both unsupervised and semi-supervised learning and facilitates swift and efficient extraction of fault features for diverse downstream tasks.

(2) A self-constructed graph approach for univariate signals is proposed, which can enable the creation of appropriate graph-structured data from singular non-multi-source vibration signals.

(3) An actual harmonic reducer testbed is designed and utilized, and can collect the experimental vibration signal under various fault conditions and operational scenarios.

The remainder of this paper is organised as follows. Section II presents an introduction to GCN and GAE networks. Section III outlines the specific operations of SCG-GFFE. Section IV presents details of the experimental research and its results. Section V summarises the work conducted in this study.

2. Preliminary related works-GCN and GAE algorithms

Consider an undirected graph denoted as $G(V, E)$, where V signifies the collection of nodes constituting the graph, and E symbolizes the collection of interconnecting edges among these nodes. Every node within this graphical framework is endowed with a unique feature vector. The entirety of these feature vectors, collectively forming a matrix, is designated as X. Specifically, $X \in \mathbb{R}^{N \times C}$, wherein N characterizes the count of nodes encompassed by the graph, and denotes the dimensionality of each node's corresponding feature vector. This matrix X is commonly recognized as the feature matrix, serving as a comprehensive embodiment of the distinct node attributes inherent to the graph.

Furthermore, for the purpose of capturing potential connections between nodes within the graph, an adjacency matrix denoted as $A \in \mathbb{R}^{N \times N}$ is introduced. In this matrix, if an edge connects the $i-th$ node u with the $j-th$ node v, the corresponding entry A_{ij} is set to 1; otherwise, it remains 0. Notably, in the context of an undirected graph, matrix A is naturally constrained to exhibit symmetry, implying that A_{ij} is equivalent to A_{ji}. In simple terms, matrix X captures the feature information of the graph, while matrix A encodes the structural information of the graph; it is only when these two matrices are combined that a complete representation of a graph data structure can be achieved.

Because graph structures exist in non-Euclidean spaces, their convolutional expressions differ from those in CNNs, making direct spatial-domain processing more challenging. Thus, to facilitate the convolution operation, both convolution signals must be transformed into the frequency domain. This transformation allows the convolution of the graph signals to be equated to a multiplication operation in the frequency domain. Following this, an inverse transformation is applied to restore the results back to the spatial domain, effectively realising the graph convolution operation, as illustrated in the following equation [16]:

$$g_{\theta} * x \approx \theta\left(I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}\right)x \tag{1}$$

where g_{θ} denotes the filter, $x \in \mathbb{R}^N$ represents the signal, I_N denotes the $N-th$ order identity matrix, and D symbolizes the degree matrix, which can be computed using the following equation:

$$D_{ij} = \begin{cases} \sum_k A_{ik}, i = j \\ 0, i \neq j \end{cases} \tag{2}$$

To mitigate the issue of gradient explosion, a renormalization operation is applied, transforming $\left(I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}\right)$ into $\widetilde{D}^{-\frac{1}{2}} \widetilde{A} \widetilde{D}^{-\frac{1}{2}}$. Moreover, extending the formula mentioned above from scalars to matrices, the expression for graph convolution can be derived as shown in the following equation:

$$X_{conv} = \widetilde{D}^{-\frac{1}{2}} \widetilde{A} \widetilde{D}^{-\frac{1}{2}} X \Theta \tag{3}$$

Here, $\Theta \in \mathbb{R}^{C \times F}$ represents the parameter matrix of the filter, $X_{conv} \in \mathbb{R}^{N \times F}$ stands for the convolved signal matrix, and the node vectors can be transformed from C dimensions to F dimensions through the convolution operation. Besides, $\widetilde{D}$ can be calculated using the following two equations:

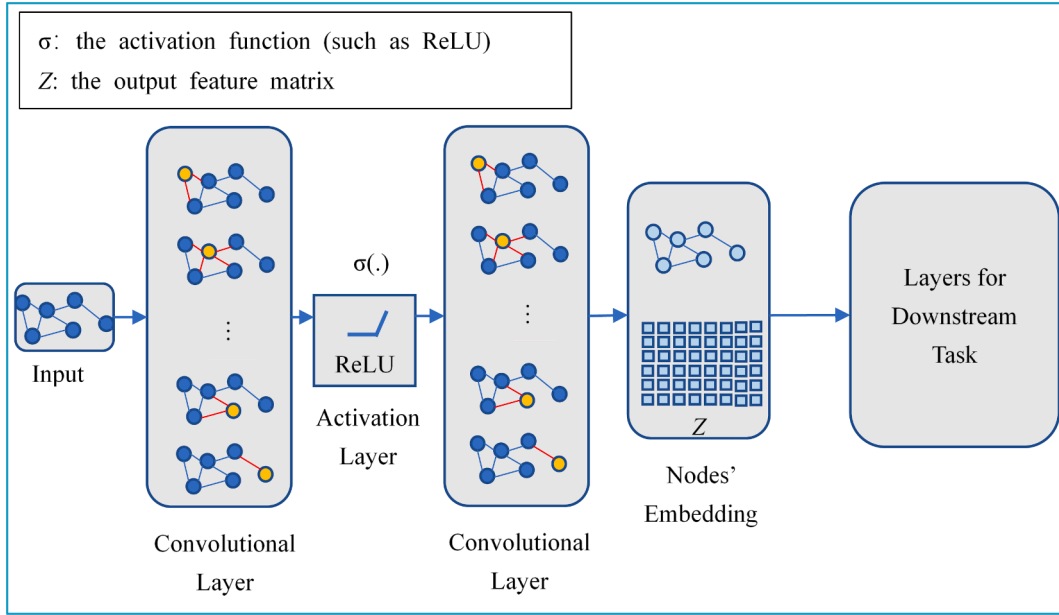$$\widetilde{D}_{ij} = \begin{cases} \sum_k \widetilde{A}_{ik}, i = j \\ 0, i \neq j \end{cases} \tag{4}$$

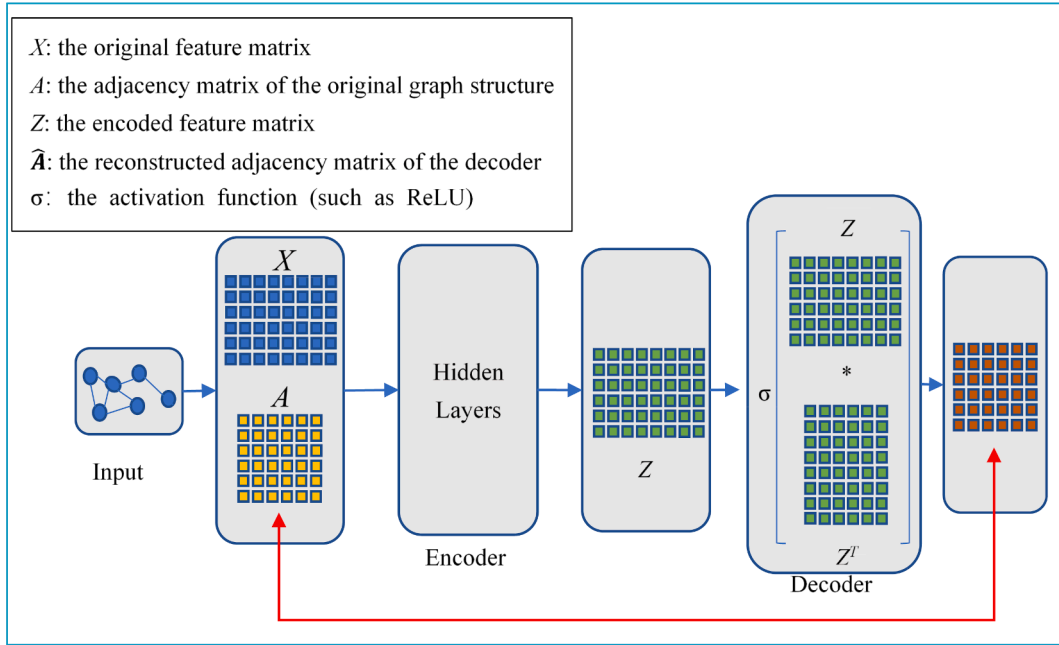**Fig. 2.** Intuitive illustration of graph convolution algorithm.

Fig. 3. Intuitive illustration of GAE algorithm.

$$\widetilde{A} = A + I_N \tag{5}$$

Hence, the final form of GCN can be represented by the following equation[16]:

$$X^{(l+1)} = f\left(X^{(l)}, A\right) = \sigma\left(\widehat{A}X^{(l)}W^{(l)}\right) = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}X^{(l)}W^{(l)}\right) \tag{6}$$

Here, $X^{(l)}$ and $X^{(l+1)}$ represent the feature matrices for the input and output of the $l-th$ convolutional layer, respectively. The function f represents the specific graph convolution operation for each layer. The $l-th$ layer's parameter matrix $W^{(l)}$ contains the trainable parameter set optimized during the training process. Within the architecture of each GCN layer, the computation of node features is achieved by combining A with $X^{(l)}$. Subsequently, these computed node features are fused with the

adaptive parameters $W^{(l)}$ to form a new feature matrix denoted as $X^{(l+1)}$. It is important to emphasise that, after the output of each graph convolutional layer, a nonlinear activation function is typically applied (often using *ReLU*, for example). Introducing this nonlinear activation function enhances the network's ability to recognise complex patterns and derive more detailed and information-rich representations from underlying graph data.

A more intuitive perspective emerges from Fig. 2, wherein graph convolution can be envisioned as a process that involves transforming the features of all neighbouring nodes (e.g. node u) and subsequently aggregating these transformed features to update the characteristics of node u[40].

Furthermore, as evident from Fig. 3, the architecture of the GAE is notably streamlined and aligns with the principles of auto-encoders. The
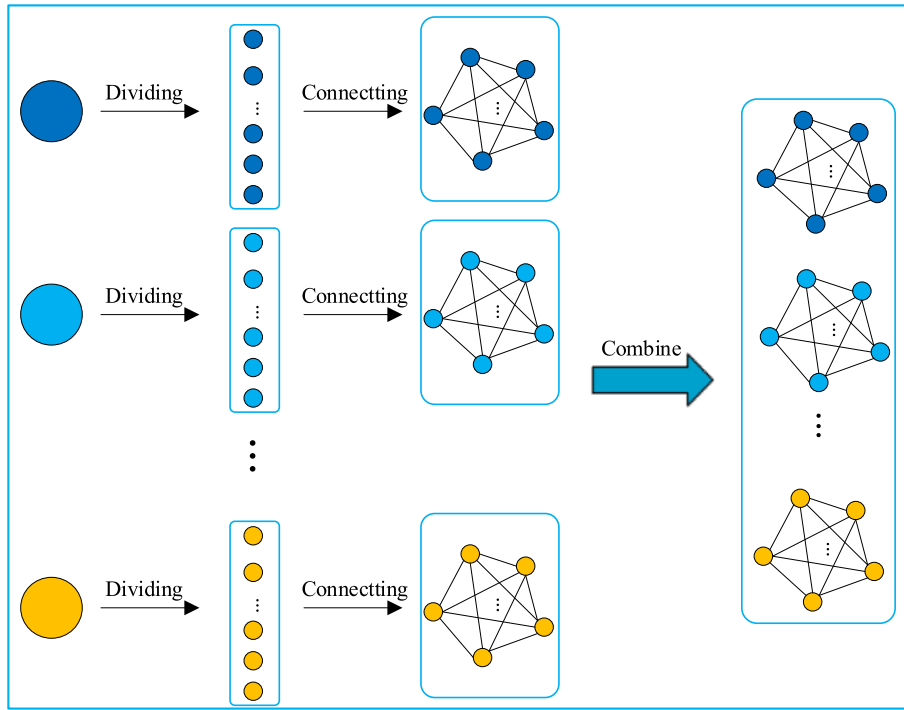
Fig. 4. Illustration of the self-constructed graph approach.

process entails obtaining node embeddings within the graph using the graph neural network algorithm (encoder phase), followed by reconstruction of the inherent features of the original graph based on the specific task (decoder phase). In general, the encoder phase adopts a two-layer GCN structure, whereas the decoder phase employs an inner-product structure to reconstruct the adjacency matrix of the graph.

3. The procedure of the proposed method

In this section, considering the context of practical industrial applications, we delve into the ideation and step-by-step implementation of the proposed fault feature extraction methodology for unlabelled harmonic reducer vibration signal data. Additionally, this section encompasses the strategy for evaluating the efficacy of feature extraction, providing a comprehensive examination of the experimental procedures undertaken in this study.

3.1. Self-construction of graph data

As depicted in Fig. 4, for vibration signal data samples, assuming preliminary data analysis and preprocessing steps have been performed (e.g., normalisation, Fourier transformation), each obtained sample is temporarily referred to as a "preprocessed sample". Subsequently, the construction of the graph data was necessary. Given our focus on unsupervised learning based on GNNs for extracting fault features, representing learning takes precedence. This entails a node-level task, commonly known as a graph or node embedding. Each sample (representing a segment of the vibration signal) was treated as a node. By establishing connections between nodes and utilising GNN methods, the vector representations for each node are derived. Considering the task requirements and challenges in constructing the graph model, we partition the vibration signal of each preprocessed sample into several segments in chronological order to avoid erroneous graph construction. Among these segments, a selective subset of K segments was chosen.

Consequently, each original sample was divided into K new samples of equal length (assuming length L). Next, these new K samples can be treated as K nodes, where each sample's vibration signal serves as an

attribute feature of the corresponding node. Subsequently, an undirected complete graph is constructed in which the weight of each edge is determined by the cosine distance between the two connected nodes. Suppose that nodes u and v are connected by an edge. The expression in the equation below defines the cosine distance between them:

$$\text{dist}(u, v) = 1 - \frac{Z_u \bullet Z_v}{\|Z_u\| \bullet \|Z_v\|} \tag{7}$$

where Z_u and Z_v are two vectors representing the attribute features of nodes u and v, denotes the vectors' dot product (inner product), $\|Z_u\|$ and $\|Z_v\|$ represent the norms (magnitudes) of vectors Z_u and Z_v.

Subsequently, utilising the Complete Graph derived from K nodes as a subgraph, each preprocessed sample yields an analogous subgraph. Ultimately, amalgamating all these subgraphs results in a comprehensive graph structure dataset, with no edges connecting the subgraphs.

3.2. GAE fault feature extractor

This section presents the GFFE designed for unsupervised learning of unlabelled data. The framework and overall flowchart of the proposed method are shown in Fig. 5. The structure of the GAE fault feature extractor described in this study is illustrated in the diagram. After obtaining the final graph data, it is used as the input for the forward propagation process. It undergoes the GCNConv1, BatchNorm, ReLU, and GCNConv2 layers sequentially (combined as the Encoder module) for feature processing, resulting in the extracted feature matrix Z. Subsequently, the matrix is fed into the Decoder module, where it is multiplied with its own transpose matrix Z^T and passed through a sigmoid activation function, yielding the reconstructed adjacency matrix $\widehat{A}$, as expressed by the following equation:

$$\widehat{A} = \text{sigmoid}(Z \bullet Z^T) \tag{8}$$

Finally, the loss function L can be computed as shown in the equations below:

$$L = L\_recon + \alpha * L\_kl + \beta * L\_reg \tag{9}$$
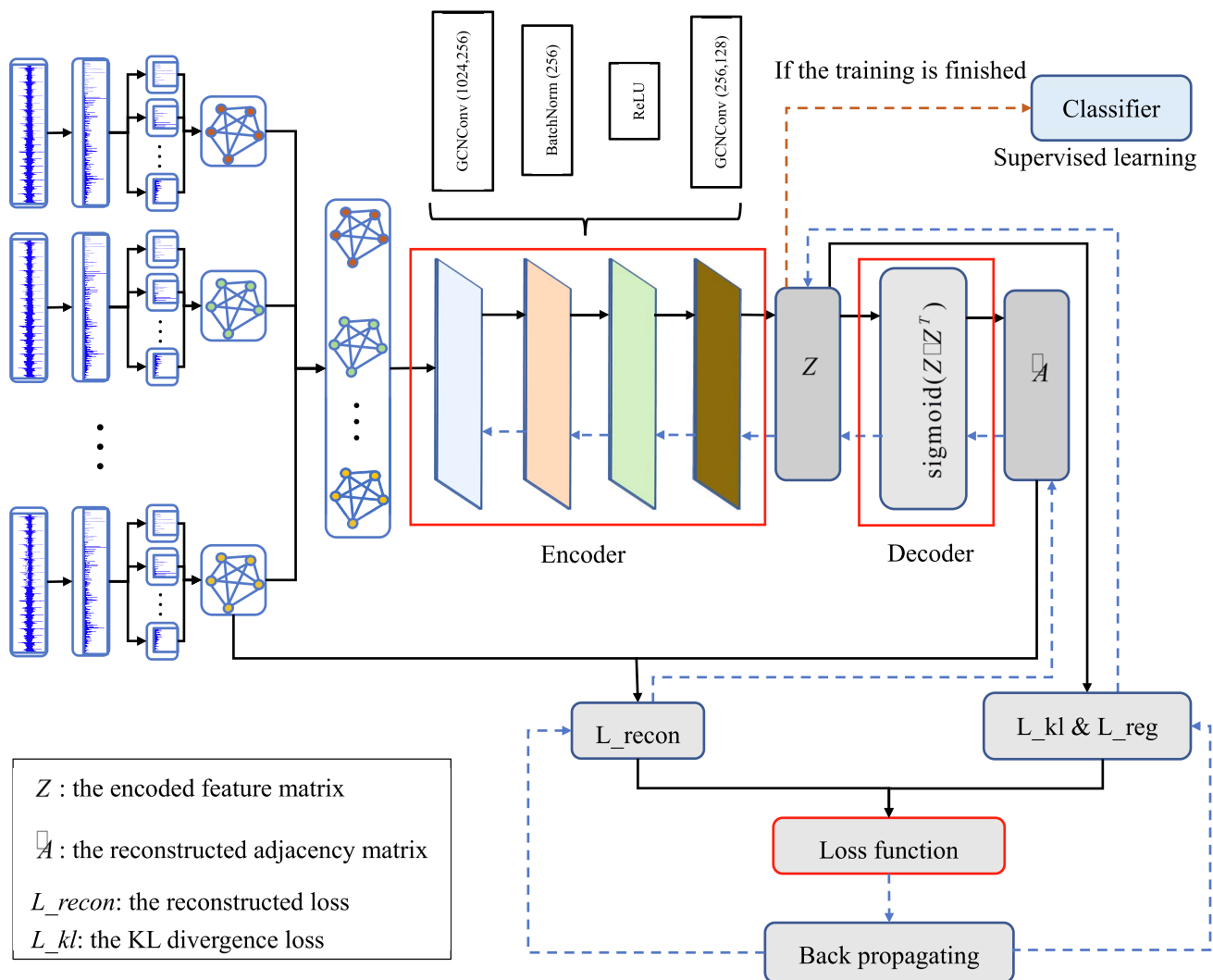
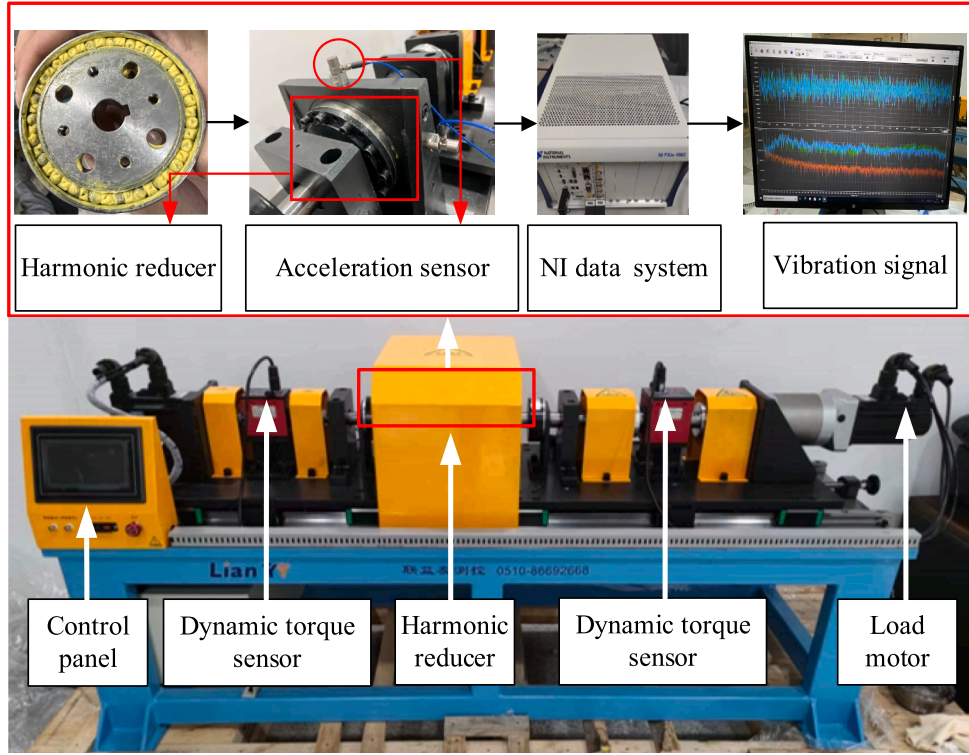**Fig. 5.** Framework and flowchart of the proposed SCG-GFFE method.

Fig. 6. Harmonic reducer test bench system setup.

$$L\_recon = -\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(A_{ij}\cdot\log\left(\widehat{A}_{ij}+10^{-10}\right)+\left(1-A_{ij}\right)\cdot\log\left(1-\widehat{A}_{ij}+10^{-10}\right)\right)}{N^2}$$

(10)

$$L\_kl = -0.5 \times \left(\frac{1}{N\times N}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(1+\log\left(Z_{ij}^2+10^{-10}\right)-Z_{ij}^2-h^2\right)\right)$$

(11)

$$L\_reg = \sum_{i=1}^{N}\sum_{j=1}^{N}Z_{ij}^2$$

(12)

where L_recon represents the reconstruction loss, N represents the number of nodes in the graph, and A_{ij} denotes the elements of the adjacency matrix of the graph structure data, and $\widehat{A}_{ij}$ denotes the elements of the reconstructed adjacency matrix. L_kl corresponds to the loss associated with the KL divergence, Z_{ij} represents the elements of the feature matrix obtained from the encoder, and h represents the dimension of the hidden layer in the encoder. Moreover, L_reg is the e regularization loss, where β is the regularization coefficient. Furthermore, both α and β are constant coefficients, which are used to balance the contributions of various loss terms.

Regarding the reconstruction loss, cross-entropy loss is employed here to quantify the disparity between $\widehat{A}$ and A. This served as a metric for evaluating the reconstruction capability of the model. In a GAE model, the objective of the reconstruction error term is to enable the decoder to accurately restore the input data, thereby preserving the essential information from the original data. By optimising the reconstruction loss, the model is compelled to extract meaningful features from graph data that are sensitive to the structure and information of the graph. This allows the model to capture the key patterns and correlations present in the data.

On the other hand, the Kullback-Leibler (KL) divergence is used to measure the discrepancy between the encoder's output Z and a predefined prior distribution. In unsupervised learning, KL divergence is commonly employed to learn the continuous distribution of the latent space. The objective of the L_kl term is to guide the encoder's output to adhere to the prior distribution in the latent space, promoting the learning of more discriminative and generalizable feature representations. Specifically, in this context, L_kl measures the disparity between Z and a standard Gaussian distribution. The learned features exhibit better continuity and a more interpretable structure when the encoder's output distribution is aligned with a standard Gaussian distribution. Moreover, the regularisation loss component aims to encourage elements within the encoded matrix to approach zero, thereby reducing the complexity of the model and preventing overfitting, thereby enhancing its generalisation capability. Throughout the training process, this loss introduces supplementary constraints on parameter gradients, nudging the model towards acquiring simpler, smoother representations that can better adapt to variations in unfamiliar data. Ultimately, by summing these three loss terms and applying weighting factors α and β, the reconstruction loss, regularisation loss, and KL divergence aspects are effectively balanced within the loss function[41]. The aim is to achieve enhanced model training and generation outcomes.

Therefore, after obtaining L, minimizing it through backpropagation is performed to train the parameter matrices within the GCN layers. This process yielded a well-suited GAE model. Subsequently, by inputting the constructed graph data into the trained GAE model, Z is extracted, which represents the fault features extracted by the feature extractor.

3.3. Classification head

After performing feature extraction, it is necessary to conduct actual classification validation using the extracted features to assess the effectiveness of the extracted fault features[42]. Considering the practical industrial context, purely unsupervised fault diagnosis may not be sufficiently reliable because of the absence of labelled references. In real industrial settings, discrepancies between fault diagnosis results and actual conditions can lead to potential accidents, resulting in production losses or even personal safety hazards. Additionally, although many data samples generally lack labels, not all samples are completely unlabelled.

Table 1
Overview of harmonic reducer dataset.

State	Sample frequency/kHz	Rotational speed/rpm	Load/%
Normal	25.6	600, 1200, 1500	0, 10, 20, 30
Flex spline pitting	25.6	600, 1200, 1500	0, 10, 20, 30
Wave generator stuttering	25.6	600, 1200, 1500	0, 10, 20, 30
Output shaft bearing misalignment	25.6	600, 1200, 1500	0, 10, 20, 30

Table 2
The parameters configuration of GAE.

First GCN layer	Second GCN layer	BatchNorm	α	β
GCNConv(1024, 256)	GCNConv(256,128)	256	0.03	0

Table 3
The experimental parameter settings.

Epoch	Learning rate	Optimizer
15	0.001	Adam

To address this, we assigned labels to all samples after feature extraction. Subsequently, a small proportion of each fault category sample was selected as the training set, while the remaining samples served as the testing set. Basic machine learning classification models were chosen for supervised training using the training set. The trained classifiers were then tested using the testing set to evaluate their diagnostic performance, thus assessing the quality of the extracted features. As the training set had limited data, while the testing set was more extensive, the testing results were considered reliable. This training method is effective because it matches real-world situations in which only a few labelled samples are available for supervised training. If a classifier performs well under these conditions, it suggests that the features it extracts are suitable for identifying faults. Additionally, it involves splitting the feature space into two parts for training and testing with similar classifiers. If a classifier trained on the newly extracted features performs better than a classifier trained on the original features, it indicates that the extracted features are more informative for fault detection. This outcome implies that the feature-extraction process was successful.

4. Experimental results and comparisons

4.1. Harmonic reducer test bench design and dataset

As illustrated in Fig. 6, in this study, we designed a practical and versatile harmonic reducer test system to collect authentic and reliable data on harmonic reducer faults. This system comprises a test bench with a faulty harmonic reducer, acceleration sensors, an NI data system, and fault-signal acquisition software, among other components. The test bench consists of a control system, drive motor, load motor, torque-speed meter, planetary gearbox, and dynamic torque sensor, among others. Customised faulty components can be manufactured during the experiments and replaced on the harmonic reducer of the test bench to simulate various fault scenarios. The control panel provides an intuitive interface for the parameter settings. The presence of the torque-speed meter and dynamic torque sensor enables real-time monitoring of input-rated torque, load torque, input speed, output speed, and other data. While the test bench is operational, the drive motor provides the rated input torque, and the load torque is supplied by the motor at the load end, simulating the operation of the harmonic reducer under realistic loads. For data collection, the digital control panel allowed us to intuitively control the input speed and load torque, facilitating data collection under various operating conditions. Load torque was measured as a percentage of the rated input torque.

However, it is essential to note that operating faulty components under high loads and high speeds may lead to part failure. The test bench was equipped with corresponding safety measures and an emergency stop button to ensure safety. Therefore, data collection practices for high-load and high-speed scenarios should not be overly prolonged.

Continuing from the previously described harmonic reducer test rig, several typical faulty components were manufactured for experimentation and data collection. Three fault types and one normal state data point were collected, as summarised in Table 1. The three fault types in this study were identified as "Wave generator stuttering", "Output shaft bearing misalignment", and "Flex spline pitting", in addition to the "Normal" state. These data types are briefly represented as "normal," "stuttering," "misalignment," and "pitting." The sampling frequency was fixed at 25600 Hz throughout the data collection process, with rotational speeds of 600 rpm, 1200 rpm, and 1500 rpm and varying loads of 0 %, 10 %, 20 %, and 30 %. By combining the different states, speeds, and loads randomly, 48 distinct subdatasets were generated. All data were acquired using the same accelerometer sensor to measure the vibration signals.

4.2. Implementation details of the experiment

Before constructing the graph data, we adopted the sliding window approach to segment the data samples, ensuring standardised sample lengths. Additionally, we applied Z-score normalisation to each data point, as shown in the following equation:

$$x^{'} = \frac{x - \mu}{\sigma} \tag{13}$$

where x represents the original data point, μ signifies the mean of the dataset, σ denotes the standard deviation of the dataset, and $x^{'}$ stands for the normalized data point. Standardising the data to follow a distribution with a mean of 0 and a standard deviation of 1 improves the model performance and enhances the stability of data processing. This normalisation process aids in eliminating scale differences among variables, resulting in a more balanced and reliable input for subsequent analyses.

Furthermore, to enhance the learning effectiveness of the GNN model while considering end-to-end processing and minimising the dependency on expert knowledge, the time-domain signal data underwent an FFT transformation. This additional step obviates the need for further signal processing operations. Subsequently, the data from the four different states can be selected by focusing solely on the samples corresponding to a rotational speed of 600 rpm and a load of 0 %. By employing the previously outlined approach, we constructed an undirected graph dataset, resulting in an input graph for the GAE. This graph encompasses 1800 nodes, including 180 Complete Subgraphs, as previously described. Subsequently, the input is fed into the trained GAE model for encoding, yielding the extracted features. This approach synergistically leverages the power of graph neural networks, end-to-end processing, and the advantageous properties of FFT to facilitate meaningful feature extraction while maintaining a minimised reliance on domain expertise.

During the specific experimental procedures, in conjunction with the PyTorch Geometric (PyG) library, we tailored the parameters of the GAE model to the particulars of our experimental setting, as presented in Table 2. Notably, α corresponds to the weight value within the aforementioned loss function.

The experimental parameter settings during the training of the extractor are outlined in Table 3.

Then, during supervised training using the newly extracted features with limited samples, specifically, 6 % of the new dataset was utilised as

Table 4
Configuration of each classifier.

SVC	decision_function_shape				
	'ovr'				
RFC	n_estimators	max_depth	max_features	random_state	/
	100	None	'sqrt'	42	/
XGB	n_estimators	learning_rate	max_depth	objective	random_state
	100	0.1	3	' multi:softprob '	42
KNN	n_neighbors				
	5				
MLP	hidden_layer_sizes	activation	solver	max_iter	random_state
	(50,)	'relu'	'adam'	500	123

Table 5
Comparison of testing results for a single operating condition before and after SCG-GFFE feature extraction.

Classifier	Accuracy with unextracted features	Accuracy with extracted features	Average enhancement
SVC	0.9369 ± 0.0136	0.9994 ± 0.0018	0.0625
RFC	0.9625 ± 0.0112	0.9994 ± 0.0018	0.0369
XGB	0.8536 ± 0.0382	0.9864 ± 0.0065	0.1328
KNN	0.8510 ± 0.0117	0.9989 ± 0.0022	0.1479
MLP	0.9919 ± 0.0032	1.0000 ± 0.0000	0.0081

Table 6
Multi-metric evaluation of few-labeled sample classification results after SCG-GFFE's feature extraction.

Classifier	macro-P	macro-R	accuracy	macro-F1	AUC
SVC	0.9991 ± 0.0019	0.9991 ± 0.0019	0.9991 ± 0.0019	0.9991 ± 0.0019	1.0000 ± 0.0000
RFC	0.9994 ± 0.0017	0.9994 ± 0.0018	0.9994 ± 0.0018	0.9994 ± 0.0018	1.0000 ± 0.0000
XGB	0.9817 ± 0.0086	0.9806 ± 0.0096	0.9806 ± 0.0096	0.9807 ± 0.0096	0.9997 ± 0.0003
KNN	0.9979 ± 0.0026	0.9978 ± 0.0027	0.9978 ± 0.0027	0.9978 ± 0.0027	1.0000 ± 0.0000
MLP	1.0000 ± 0.0000	1.0000 ± 0.0000	1.0000 ± 0.0000	1.0000 ± 0.0000	1.0000 ± 0.0000

Table 7
Multi-metric evaluation of few-labeled sample classification results without SCG-GFFE's feature extraction.

Classifier	macro-P	macro-R	accuracy	macro-F1	AUC
SVC	0.9267 ± 0.0153	0.9228 ± 0.0191	0.9228 ± 0.0191	0.9227 ± 0.0194	0.9926 ± 0.0022
RFC	0.9605 ± 0.0127	0.9595 ± 0.0143	0.9595 ± 0.0143	0.9595 ± 0.0143	0.9945 ± 0.0030
XGB	0.8692 ± 0.0185	0.8667 ± 0.0190	0.8667 ± 0.0190	0.8665 ± 0.0191	0.9759 ± 0.0074
KNN	0.8508 ± 0.0123	0.8359 ± 0.0139	0.8359 ± 0.0139	0.8358 ± 0.0142	0.9623 ± 0.0047
MLP	0.9889 ± 0.0044	0.9888 ± 0.0045	0.9888 ± 0.0045	0.9887 ± 0.0046	0.9917 ± 0.0025

Table 8
Average enhancements in multimetric evaluation of fault diagnosis results with and without SCG-GFFE feature extraction.

Classifier	macro-P	macro-R	accuracy	macro-F1	AUC
SVC	0.0724	0.0763	0.0763	0.0764	0.0074
RFC	0.0389	0.0399	0.0399	0.0399	0.0055
XGB	0.1125	0.1139	0.1139	0.1142	0.0238
KNN	0.1471	0.1619	0.1619	0.1620	0.0377
MLP	0.0111	0.0112	0.0112	0.0113	0.0083

prevalent classification models were chosen for this purpose: Support Vector Classifier (SVC), Random Forest Classifier (RFC), XGBoost Classifier (XGB), K-Nearest Neighbors Classifier (KNN), and Multi-Layer Perceptron (MLP).

The specific configurations of each classifier are listed in Table 4.

The feature extraction process described above operates under the condition of unsupervised learning, as neither the construction nor the learning of the graph data requires knowledge of the sample labels. This characteristic makes it suitable for scenarios in industrial contexts where labelled data are scarce. Moreover, the methodology in this study aligns well with real-world practices. In practical scenarios, data can be collected and diagnostic processes can be repeated multiple times for the same equipment or component to enhance the reliability and stability of fault diagnosis results while mitigating the risks associated with erroneous diagnoses. To some extent, when combined with GAE, the proposed graph construction method can be considered as a form of data filtering that effectively removes less relevant noise information. During utilisation, the diagnostic results obtained from multiple segments of the original sample, as partitioned by the method, can be integrated. The most frequent diagnostic outcome among these segments was selected as the final fault diagnosis of the original sample.

4.3. Classification and comparison results

During testing, to evaluate the outstanding performance of the proposed SCG-GFFE method, experiments were conducted following the method and procedure described earlier. Targeting a few-shot supervised learning scenario, the same dataset was utilized with controls placed on the utilization of the SCG-GFFE model, as depicted in Table 5 for comparison purposes. The experimental results demonstrate a significant improvement in the accuracy of fault diagnosis with the utilization of SCG-GFFE, thereby affirming its efficacy. Additionally, to ensure the reliability and consistency of the results, repeated experiments were conducted for each experimental setup, with the mean and variance calculated to portray result stability and mitigate experimental stochasticity. In terms of fault classifiers, experiments were carried out using five basic classifiers individually, thereby avoiding potential anomalies introduced by unique classifiers, emphasizing the effectiveness of the feature extractor, and validating the generality and reliability of SCG-GFFE through comparative experiments.

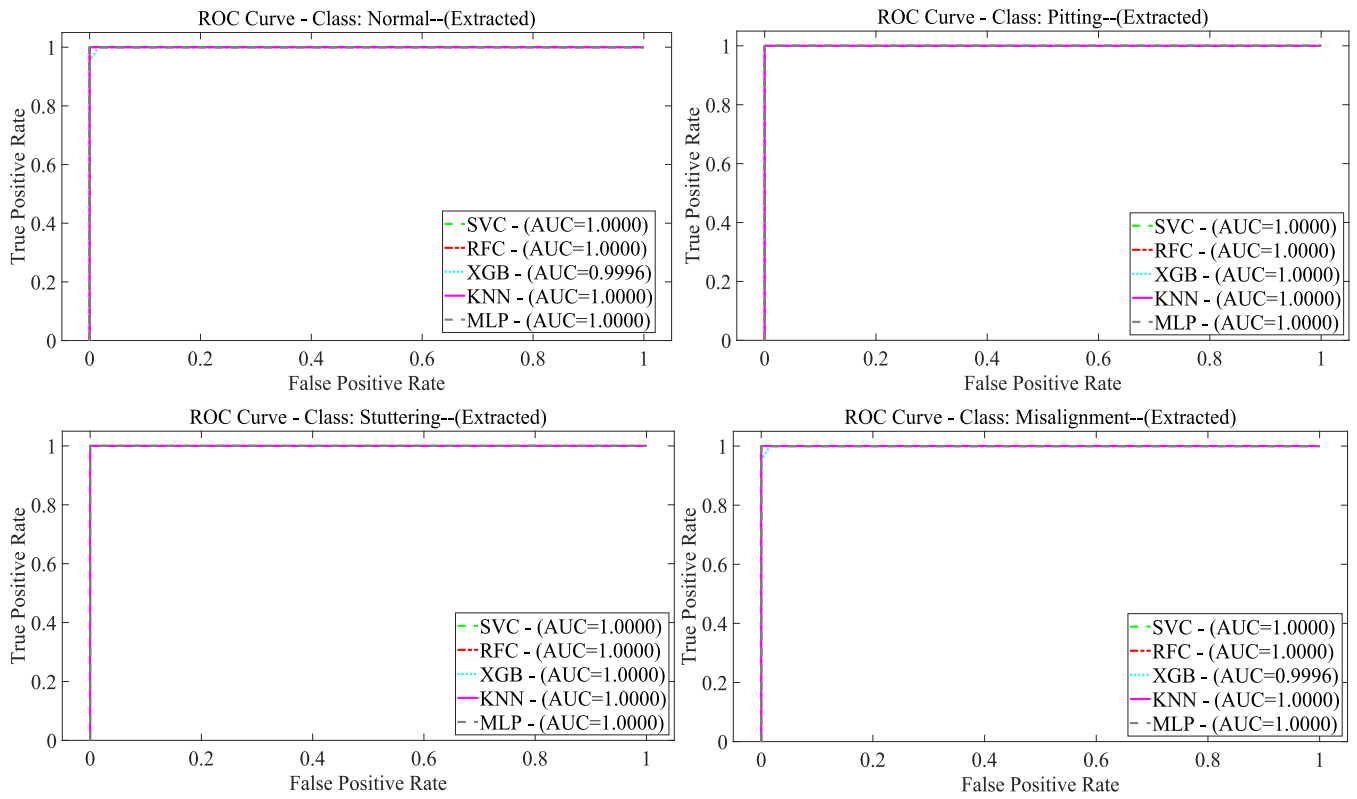In addition, to comprehensively evaluate the performance of the

the training set to train the classifiers in a supervised manner, simulating scenarios with limited labelled samples in industrial applications. The remaining 94 % of the new samples served as the test set. To thoroughly assess the quality of the new sample data, a selection of fundamental and commonly used machine learning classification models based on scikit-learn was employed. The training and testing processes were carried out separately for the new and old samples with the aim of comparing diagnostic outcomes. For the model training phase, parameter selection for the classifiers is performed arbitrarily, without deliberate tuning, to minimise the influence of the classifiers. To avoid randomness, five

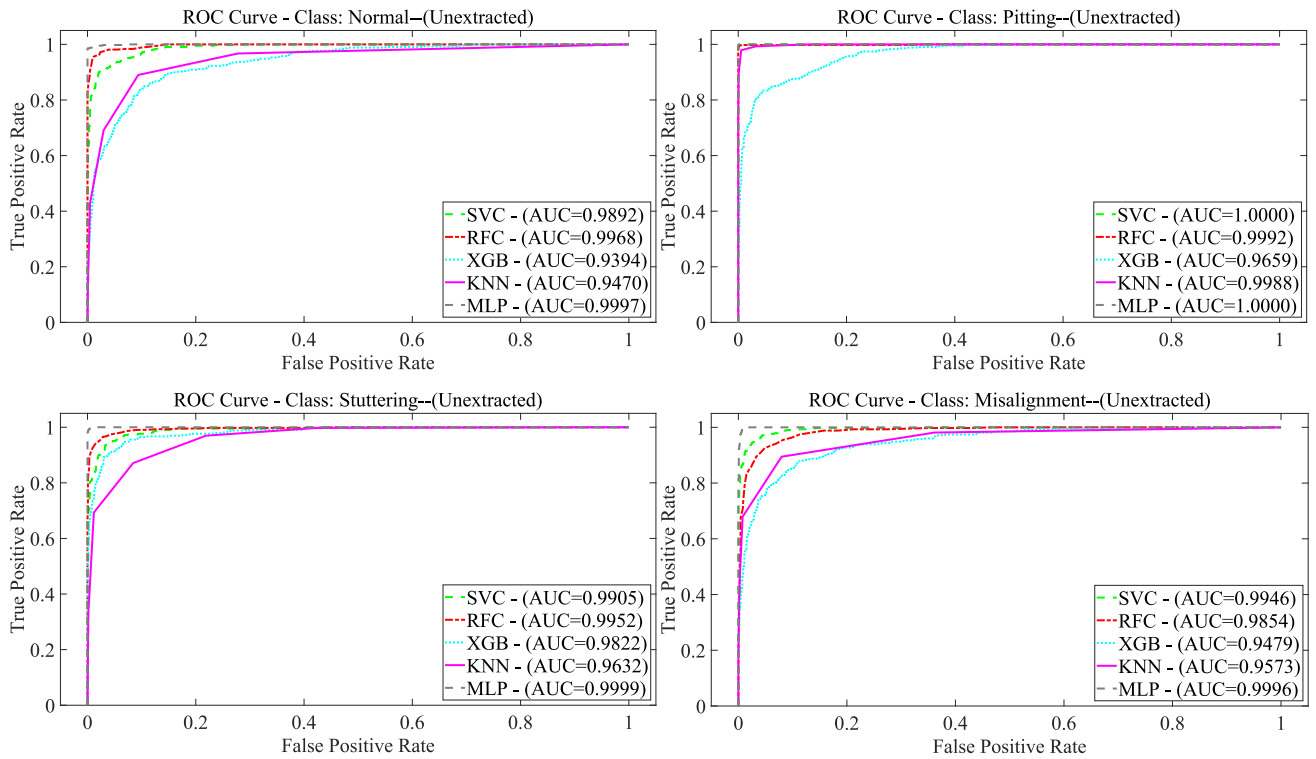**Fig. 7.** ROC curves for various classifiers in fault diagnosis with SCG-GFFE.

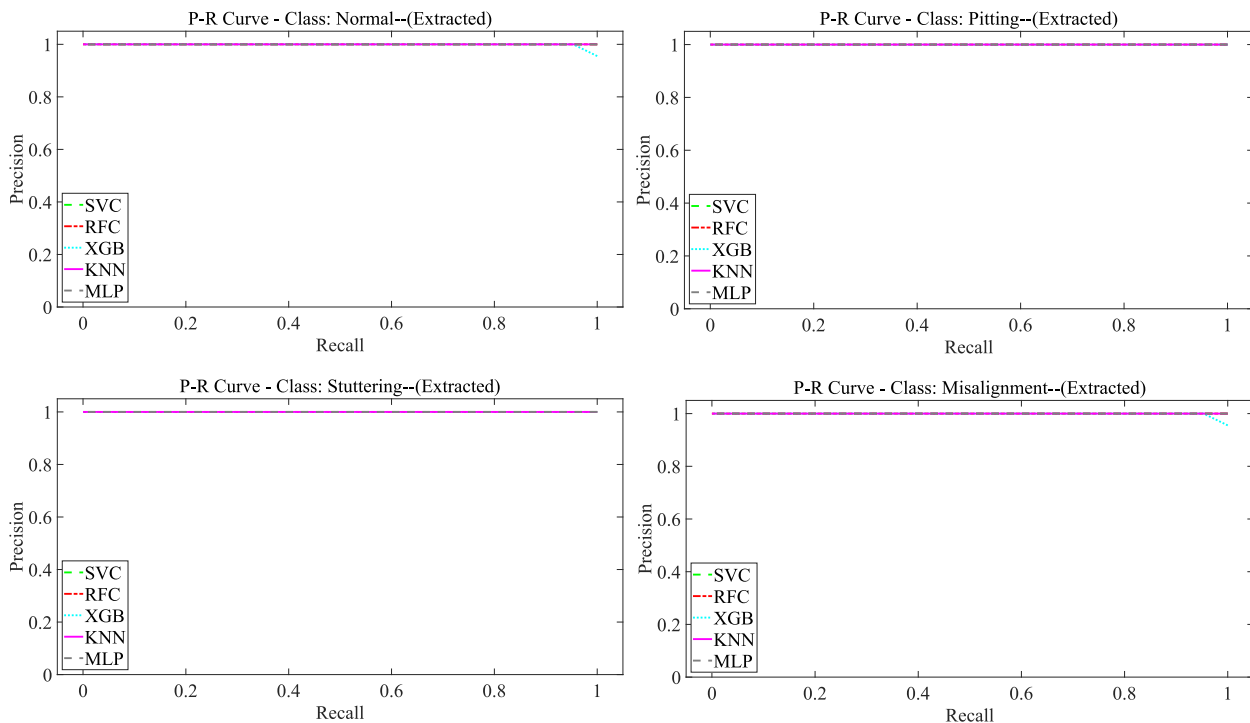**Fig. 8.** ROC curves for various classifiers in fault diagnosis without SCG-GFFE.

Fig. 9. P-R curves for various classifiers in fault diagnosis with SCG-GFFE.
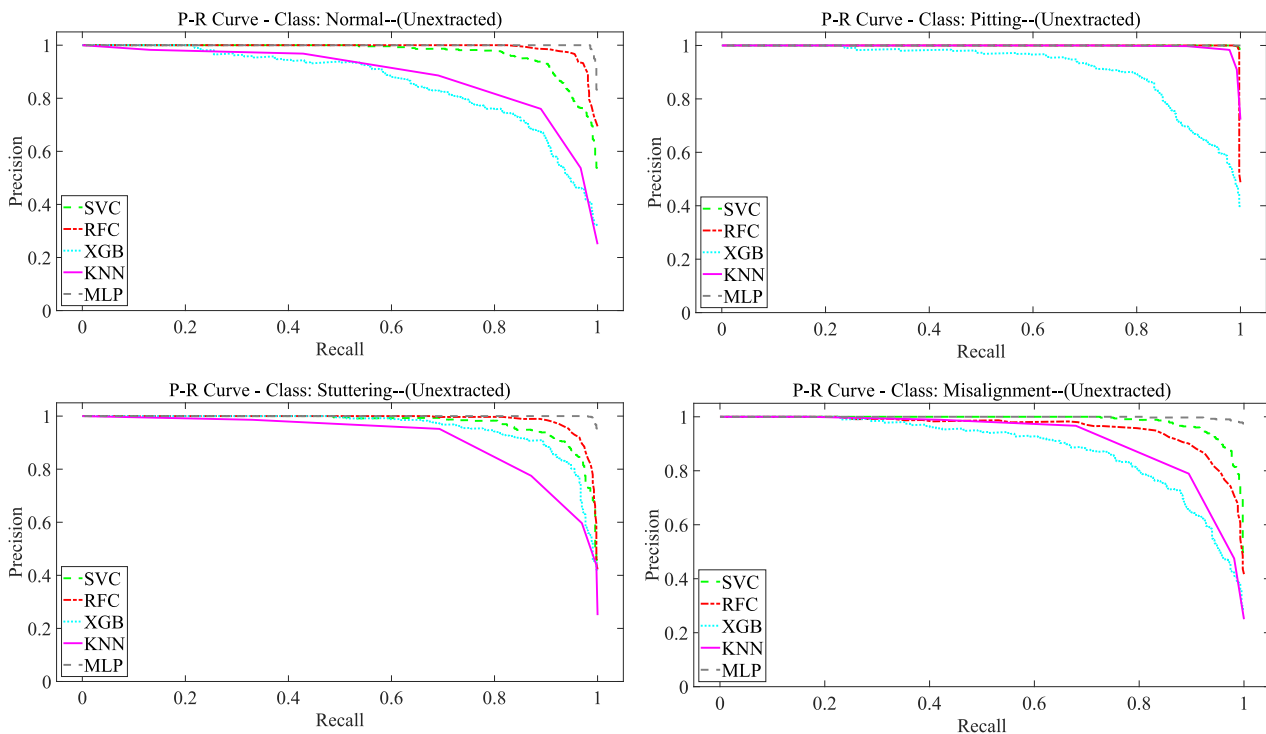

Fig. 10. P-R curves for various classifiers in fault diagnosis without SCG-GFFE.

Table 9
Comparison of testing results for 12 operating conditions before and after SCG-GFFE's feature extraction.

Classifier	Accuracy with unextracted features	Accuracy with extracted features	Average enhancement
SVC	0.9026 ± 0.0128	0.9708 ± 0.0140	0.0682
RFC	0.8807 ± 0.0331	0.9754 ± 0.0111	0.0947
XGB	0.7956 ± 0.0184	0.9638 ± 0.0104	0.1682
KNN	0.8293 ± 0.0116	0.9888 ± 0.0098	0.1595
MLP	0.9618 ± 0.0047	0.9973 ± 0.0043	0.0355

Table 10
Comparison of testing results based on SCG-GFFE and KNN-GFFE with different K.

Classifier	KNN-GFFE (K = 1)	KNN-GFFE (K = 5)	KNN-GFFE (K = 7)	KNN-GFFE (K = 9)	SCG-GFFE (Ours)
SVC	0.9157 ± 0.1740	0.9083 ± 0.0087	0.9028 ± 0.0216	0.8922 ± 0.0142	1.0000 ± 0.0000
RFC	0.8900 ± 0.0150	0.8823 ± 0.0113	0.8898 ± 0.0228	0.8827 ± 0.0103	0.9994 ± 0.0019
XGB	0.8504 ± 0.0209	0.8444 ± 0.0182	0.8395 ± 0.0208	0.8474 ± 0.0226	0.9690 ± 0.0236
KNN	0.8728 ± 0.0142	0.8666 ± 0.0122	0.8115 ± 0.0294	0.8310 ± 0.0219	0.9994 ± 0.0021
MLP	0.9289 ± 0.0094	0.9194 ± 0.0084	0.9244 ± 0.0156	0.9208 ± 0.0085	1.0000 ± 0.0000

model, precision, recall, and F1-Score were chosen as auxiliary metrics. Because these three metrics were originally designed for binary classification tasks, they were adapted here for the four-class classification task by treating them as four binary classifications. The metrics were averaged across each class to obtain macro-P, macro-R, and the corresponding macro-F1, as defined by the following formulas:

$$macro - P = \frac{1}{n}\sum_{i=1}^{n} P_i = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{TP}{TP + FP}\right)_i \tag{14}$$

$$macro - R = \frac{1}{n}\sum_{i=1}^{n} R_i = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{TP}{TP + FN}\right)_i \tag{15}$$

$$macro - F1 = \frac{2 \times macro - P \times macro - R}{macro - P + macro - R} \tag{16}$$

where n represents the number of classes in the classification task or the number of binary classification tasks when split, and TP, FP, and FN represent the numbers of true positives, false positives, and false negatives, respectively, in the binary classification tasks.

Since different indicators can capture different performance characteristics, considering the possible limitations of single "accuracy" as an evaluation indicator, the comprehensive evaluation of multiple indicators can make the evaluation of experimental results more comprehensive and reliable. The experimental results are shown in Tables 6-8. Table 6 is the result of using SCG-GFFE, Table 7 is the result of not using SCG-GFFE, and Table 8 statistics the improvement obtained by using SCG-GFFE. Obviously, from the results of various indicators, the use of SCG-GFFE makes the fault diagnosis effect better, which also verifies the effectiveness of SCG-GFFE from multiple perspectives.

To provide a more intuitive observation of the effect of using SCG-GFFE for feature extraction before and after the on-fault diagnosis of the few-labelled samples, we randomly selected one experimental result. For each fault category, we calculated the diagnostic performance separately. We plotted the precision-recall (P-R) curves and Receiver Operating Characteristic (ROC) curves for each classification model (Figs. 7-10 below). As can be clearly observed, when employing SCG-GFFE for fault diagnosis in harmonic reducers, followed by the use of basic classifiers, the results are excellent and significantly superior to those obtained without SCG-GFFE.

The previous experiments were conducted based on a specific operating condition for comparative validation, considering relatively simple scenarios. However, in practical applications, various operating conditions may be encountered. Therefore, comparative experiments were conducted under 12 different operating conditions, as shown in Table 9, with a focus on distinguishing different fault states in a four-class classification task. The experimental results demonstrate the continued effectiveness of the proposed method in complex operating conditions. Thus, considering single or multiple operating conditions, the effectiveness and stability of the feature extractor remain consistent, yielding expected results. This further validates the stability of SCG-GFFE and its applicability in a wide range of scenarios.

As mentioned in our earlier discussions, constructing graph data from vibration signals poses a significant challenge due to the absence of explicit structural information. For single-sensor measurements of vibration signals, one of the most direct and commonly used methods for graph construction is based on K-nearest neighbors (KNN). Therefore, after validating the performance of the feature extractor, experiments were designed to assess the performance of the proposed self-constructed graph method and compare it primarily with the KNN-based graph construction method. Building upon the GFFE proposed in this paper, the graph construction method was replaced, and the effects of the self-constructed graph method and the KNN-based graph method were compared, as presented in Table 10. The experimental results demonstrate that the self-constructed graph method proposed in this paper exhibits superior performance compared to the KNN-based
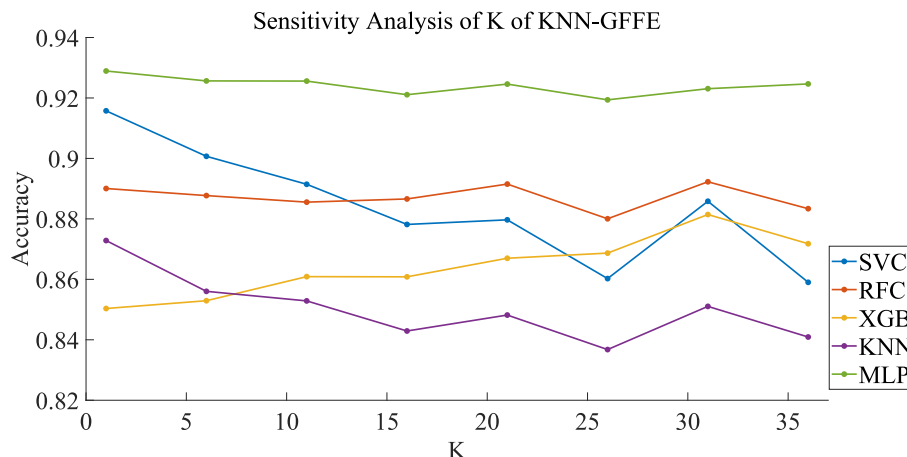

Fig. 11. Diagnostic results for different values of K in the KNN-GFFE-based fault diagnosis.
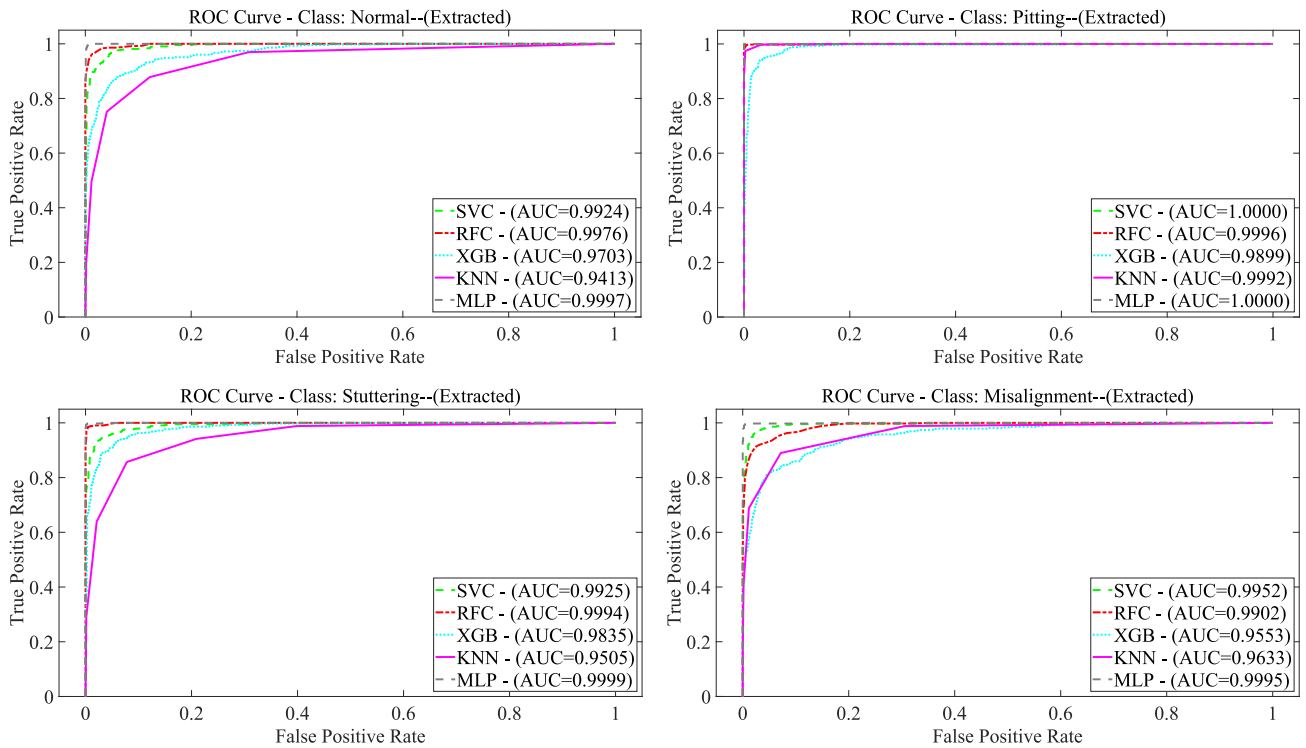
Fig. 12. ROC curves for various classifiers in fault diagnosis with KNN-GFFE (K = 1).
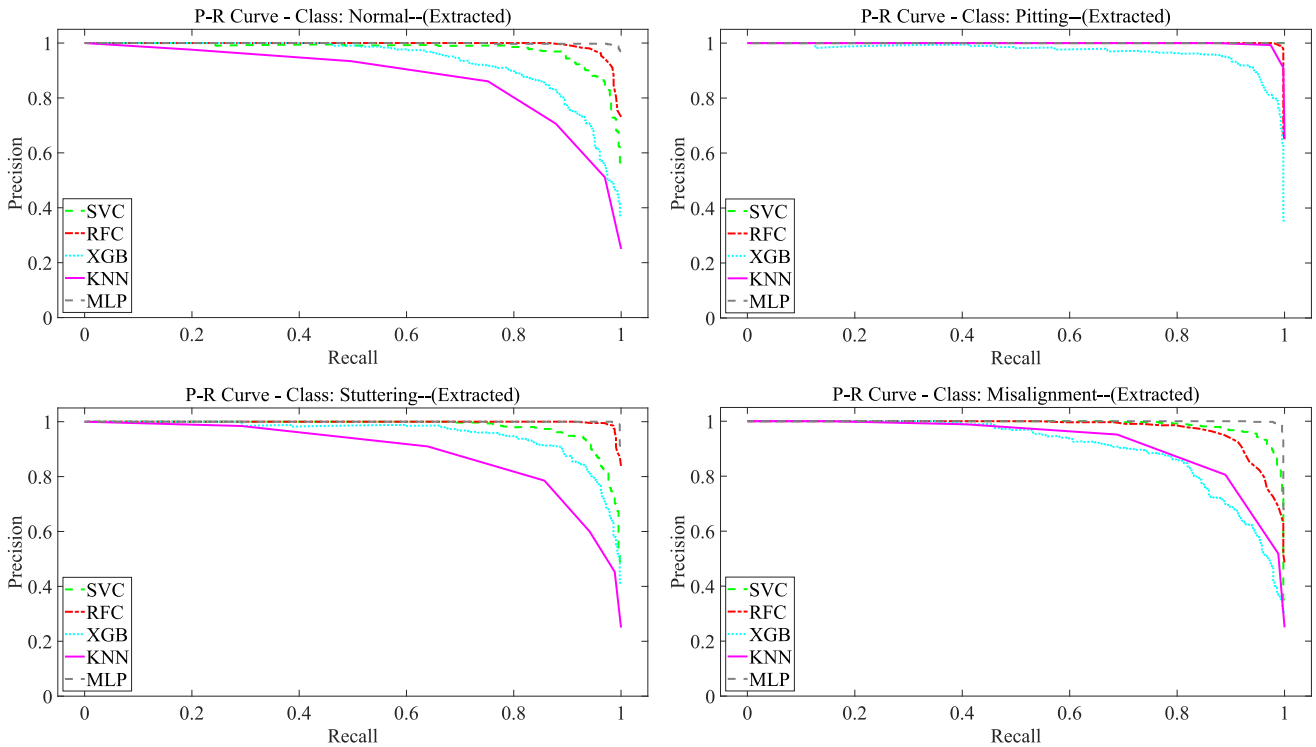

Fig. 13. P-R curves for various classifiers in fault diagnosis with KNN-GFFE (K = 1).

graph construction method.

Subsequently, for different values of K in KNN-GFFE, the corresponding results for fault diagnosis on a few-label sample dataset are shown in Fig. 11. It is evident that the results of the fault diagnosis using KNN-GFFE exhibit some fluctuations with varying K values. The results are notably inferior to the performance achieved by SCG-GFFE. It further

underscores the instability of the KNN-based method and its reliance on the choice of K value and the classification capability of the KNN algorithm itself, thereby highlighting the superior performance of the self-constructing graph method. Furthermore, as an example, when using K = 1 in KNN-GFFE for feature extraction, the performance of fault classification based on the extracted features is illustrated in Figs. 12 and

Table 11

Comparison of testing results using SCG-GFFE and WRCTD-CNN-LSTM.

Classifier	Accuracy
SCG-GFFE-SVC	0.9994 ± 0.0018
SCG-GFFE-RFC	0.9994 ± 0.0018
SCG-GFFE-XGB	0.9864 ± 0.0065
SCG-GFFE-KNN	0.9989 ± 0.0022
SCG-GFFE-MLP	1.0000 ± 0.0000
WRCTD-CNN-LSTM	0.8584 ± 0.0143

13, clearly demonstrating a significant performance gap compared with SCG-GFFE.

Finally, a comparison was also conducted with recent experimental results of a deep learning-based harmonic reducer model, WRCTD-CNN-LSTM [43], as shown in Table 11. The results indicate that the proposed approach outperforms the WRCTD-CNN-LSTM method significantly in the context of few-shot learning on this dataset. Considering the SCG-GFFE-MLP proposed in this paper, which achieves near-perfect accuracy of 100 % with only a simple multilayer perceptron and a mere 6 % of the data for training, its performance is remarkably impressive. Hence, it is inferred that even without further extensive comparisons with other models at the level of harmonic reducer fault diagnosis, the excellence of the proposed method is adequately demonstrated.

In summary, experimental validations were conducted on various aspects such as model feature extraction effectiveness, stability, and reliability. Multiple metrics and experimental conditions were provided to comprehensively verify the excellent performance of the proposed SCG-GFFE from various perspectives.

4.4. Visualization of extracted features and sensitivity analysis

Furthermore, to provide a more intuitive view of the feature extraction results, the t-SNE algorithm is employed to reduce the dimensionality of both the original and new features. Subsequently, these reduced-dimensional features were visualised on a two-dimensional plane to highlight the clustering effects. A comparison of the clustering effects of the embeddings before and after extraction is shown in Fig. 14.

Clearly, whether in a single operating condition or under the more complex scenario of twelve different conditions, it is evident from the visualizations that after applying the feature extraction method proposed in this study, the features of different fault types are effectively separated. This separation is conducive to enhancing the effectiveness of fault diagnosis and serves as a visual confirmation of the successful extraction of fault features.

Simultaneously, a sensitivity analysis was conducted for the two loss term coefficients α and β in the experimental settings. They were varied within the range of 0–0.1 with a step size of 0.01. Multiple experiments were performed, each with a fixed combination of α and β, and the average classification accuracy for the fault diagnosis was recorded. The sensitivity analysis results are shown in Fig. 15, where *kl_coef* and *reg_coef* represent the coefficients for the KL divergence loss term and the regularisation loss term, respectively, corresponding to the weighting factors α and β. It can be observed that the presence of both KL divergence loss and regularisation loss contributes to the improvement in the final performance, and variations in the loss coefficients have an impact on the outcome. In practical applications, it is advisable to set the loss coefficients flexibly based on the specific choice of the final classifier. However, overall, even with the suboptimal parameter choices discussed in this paper, the final fault diagnosis performance remained sufficiently impressive. This further substantiates the outstanding
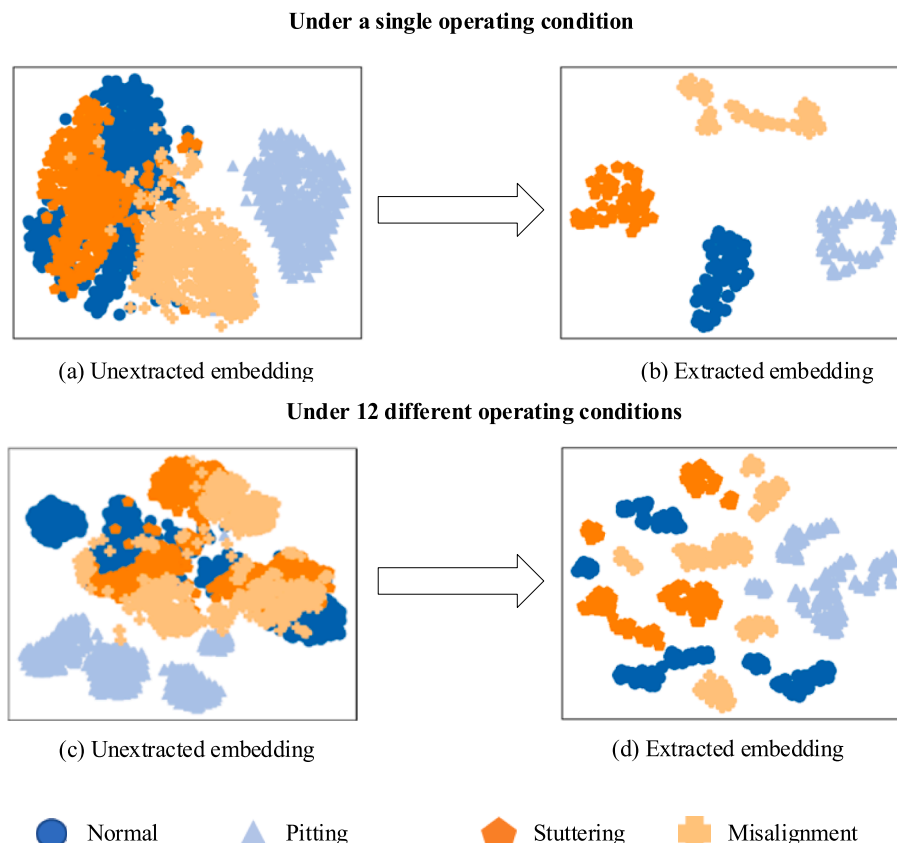
Under a single operating condition

(a) Unextracted embedding   (b) Extracted embedding

Under 12 different operating conditions

(c) Unextracted embedding   (d) Extracted embedding

● Normal   ▲ Pitting   ⬟ Stuttering   ⬟ Misalignment

Fig. 14. Visualisation of the clustering effects of embeddings before and after SCG-GFFE processing.
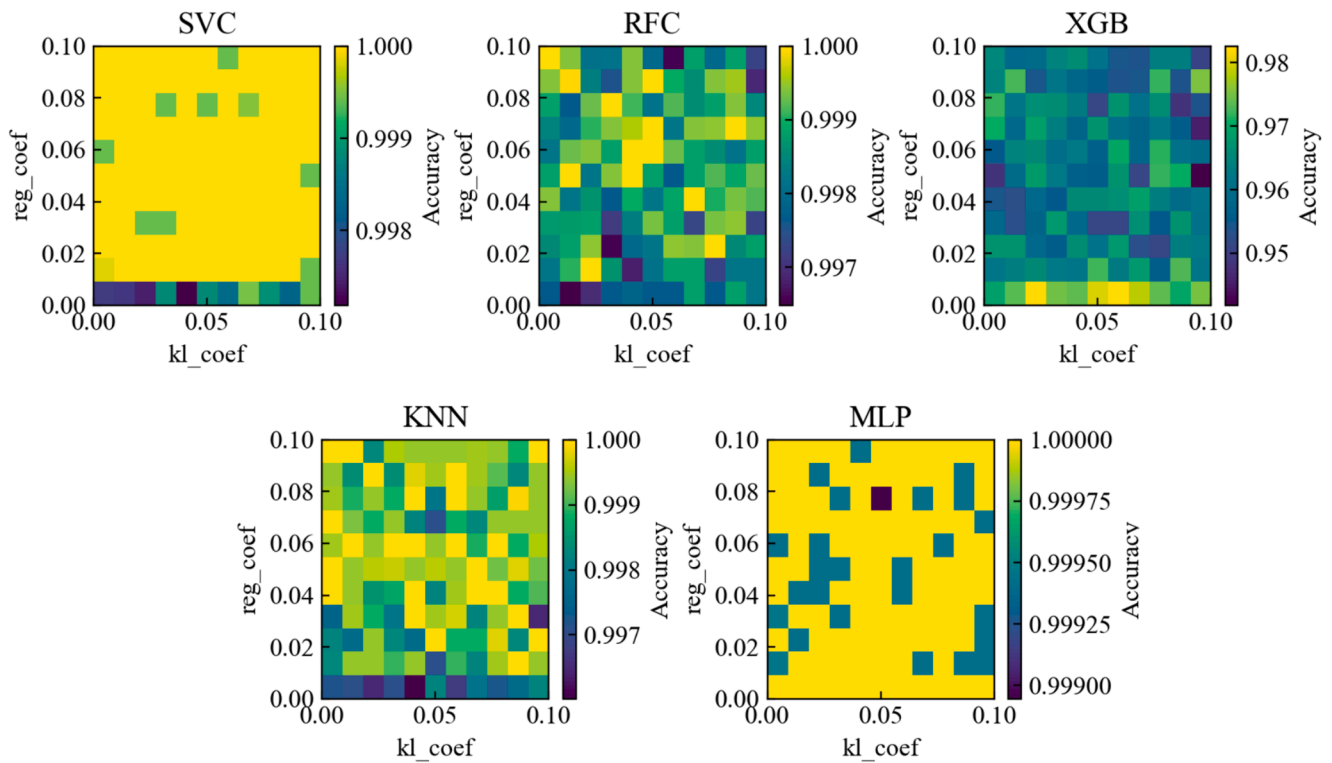
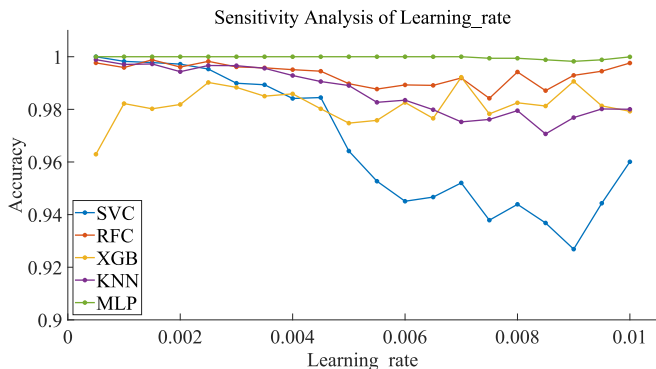**Fig. 15.** Sensitivity analysis of loss term coefficients α and β.

Fig. 16. Sensitivity analysis of learning rate.

feature extraction capability of SCG-GFFE. Add sensitivity analysis results regarding the learning rate are presented in Fig. 16.

The sensitivity analysis results, as shown in the figures, indicate that the value of β significantly affects the overall classification performance, with a preference for β = 0. Furthermore, the impact of α is also present but not as pronounced as that of β. The chosen parameter settings of α and β as 0.03 and 0, respectively, appear to be favourable. Moreover, when the learning rate became relatively large, the overall classification performance of the classifier tended to decrease. Thus, a learning rate of 0.001 was deemed to be appropriate in this context.

In specific experiments related to harmonic reducer fault diagnosis, we segmented a complete end-to-end fault diagnosis network into two parts: the fault feature extractor and the fault classifier. Different training and testing strategies were applied to each, and the testing performance of the fault classifier uniformly indicated the overall effectiveness of both fault diagnosis and fault feature extraction. Through feature visualization, along with various metrics such as fault classification accuracy, and conducting multiple sets of experiments using different fault classifiers, this method has been comprehensively validated from diverse

perspectives, affirming its reliability, effectiveness, and stability. Simultaneously, the comparison of fault diagnosis outcomes with or without the utilisation of SCG-GFFE provides robust evidence for the effectiveness and outstanding performance of SCG-GFFE. It adeptly extracts fault feature information from signals, significantly enhancing the salience of the fault features in the extracted new samples.

5. Conclusion

This study addresses a pivotal issue in the fault diagnosis of harmonic reducers by focusing on the extraction of fault features. A fault feature extractor, named SCG-GFFE, was devised based on graph neural networks. Tailored for unlabelled vibration signals collected from a single sensor, this fault feature extractor facilitates the unsupervised extraction of fault features, enabling an effective fault diagnosis with limited labelled samples. There are three main contributions. First, this method offers a practical and effective approach to constructing graph data for non-multisource signals, resulting in a self-constructed graph. Through organic integration and segmentation of the samples, the self-constructed graph demonstrated a reduced likelihood of erroneous edge connections, showcasing heightened reliability and stability. Second, compared to traditional and prevalent graph construction methods based on signal processing and expert knowledge, our approach does not rely on the integration of expert knowledge. It is more direct, effective, and user-friendly. Third, experiments on harmonic reducer fault diagnosis conclusively demonstrate that this fault feature extractor can be seamlessly integrated with other classification algorithms, enabling precise fault diagnosis of harmonic reducers even with limited labelled samples.

In summary, the method proposed in this study can construct appropriate graph-structured data from vibration signals collected by a single sensor, perform graph learning, and effectively extract fault features from unlabelled samples. This enhances the quality of the samples and contributes to downstream tasks, including improving the fault diagnosis accuracy with limited labelled samples. The simplicity and efficiency of this method make it a valuable standalone step for enhancing the sample quality. It can be seamlessly integrated with other

fault diagnosis methods, showing robust stability, generality, and scalability. Therefore, this method holds promising prospects for unsupervised fault feature extraction from harmonic reducer vibration signals and for conducting the fault diagnosis of harmonic reducers.

CRediT authorship contribution statement

Shilong Sun: Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Hao Ding:** Writing – original draft, Software, Methodology, Formal analysis. **Zida Zhao:** Software, Methodology, Formal analysis. **Wenfu Xu:** Supervision, Project administration. **Dong Wang:** Writing – review & editing, Validation, Supervision, Formal analysis.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

References

[1] K. Kaneko, et al., Humanoid robot HRP-5P: An electrically actuated humanoid robot with high-power and wide-range joints, IEEE Rob. Autom. Lett. 4 (2) (2019) 1431–1438.

[2] Z. Chen, A. Mauricio, W. Li, K. Gryllias, A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks, Mech. Syst. Sig. Process. 140 (2020) 106683.

[3] H. Lee, H. Jeong, G. Koo, J. Ban, S.W. Kim, Attention recurrent neural network-based severity estimation method for interturn short-circuit fault in permanent magnet synchronous machines, IEEE Trans. Ind. Electron. 68 (4) (2021) 3445–3453.

[4] X. Wu, Y. Zhang, C. Cheng, Z. Peng, A hybrid classification autoencoder for semi-supervised fault diagnosis in rotating machinery, Mech. Syst. Sig. Process. 149 (2021) 107327.

[5] S. Gao, L. Xu, Y. Zhang, Z. Pei, Rolling bearing fault diagnosis based on SSA optimized self-adaptive DBN, ISA Trans. 128 (2022) 485–502.

[6] S. Sun, T. Peng, Y. Zhou, X. Zhang, D. Wang, Contrastive learning and dynamics embedding neural network for label-free interpretable machine fault diagnosis, ISA Trans. (2023).

[7] Q. Yao, Q. Qian, Y. Qin, L. Guo, F. Wu, Adversarial domain adaptation network with pseudo-siamese feature extractors for cross-bearing fault transfer diagnosis, Eng. Appl. Artif. Intel. 113 (2022) 104932.

[8] C. Wang, J. Yang, B. Zhang, A fault diagnosis method using improved prototypical network and weighting similarity-Manhattan distance with insufficient noisy data, Measurement 226 (2024) 114171.

[9] J. Yang, C. Wang, C.a. Wei, A novel Brownian correlation metric prototypical network for rotating machinery fault diagnosis with few and zero shot learners, Adv. Eng. Inf. 54 (2022) 101815.

[10] Z. Zhi, L. Liu, D. Liu, C. Hu, Fault detection of the harmonic reducer based on CNN-LSTM with a novel denoising algorithm, IEEE Sens. J. 22 (3) (2022) 2572–2581.

[11] L. Liu, Z. Zhi, Y. Yang, S. Shirmohammadi, D. Liu, Harmonic reducer fault detection with acoustic emission, IEEE Trans. Instrum. Meas. 72 (2023) 1–12.

[12] Y. He, J. Chen, X. Zhou, S. Huang, In-situ fault diagnosis for the harmonic reducer of industrial robots via multi-scale mixed convolutional neural networks, J. Manuf. Syst. 66 (2023) 233–247.

[13] X. Zhou, H. Zhou, Y. He, S. Huang, Z. Zhu, J. Chen, Harmonic reducer in-situ fault diagnosis for industrial robots based on deep learning, Sci. China Technol. Sci. 65 (9) (2022) 2116–2126.

[14] S. Yang, Y. Xiang, Z. Long, X. Ma, Q. Ding, J. Jia, Fault diagnosis of harmonic drives based on an SDP-ConvNeXt joint methodology, IEEE Trans. Instrum. Meas. 72 (2023) 1–8.

[15] G. Yang, Y. Zhong, L. Yang, H. Tao, J. Li, R. Du, Fault diagnosis of harmonic drive with imbalanced data using generative adversarial network, IEEE Trans. Instrum. Meas. 70 (2021) 1–11.

[16] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv preprint arXiv:1609.02907, 2016.

[17] T. N. Kipf and M. Welling, "Variational graph auto-encoders," arXiv preprint arXiv: 1611.07308, 2016.

[18] G. Jin, et al., Spatio-temporal graph neural networks for predictive learning in urban computing: A survey, IEEE Trans. Knowl. Data Eng. (2023) 1–20.

[19] K. Zhang, J. Chen, S. He, F. Li, Y. Feng, Z. Zhou, 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 (2022) 1–16.

[20] Z. Yu, C. Zhang, C. Deng, An improved GNN using dynamic graph embedding mechanism: A novel end-to-end framework for rolling bearing fault diagnosis under variable working conditions, Mech. Syst. Sig. Process. 200 (2023) 110534.

[21] Y. Tang, et al., Rotating machine systems fault diagnosis using semisupervised conditional random field-based graph attention network, IEEE Trans. Instrum. Meas. 70 (2021) 1–10.

[22] G. Yang, H. Tao, R. Du, Y. Zhong, Compound fault diagnosis of harmonic drives using deep capsule graph convolutional network, IEEE Trans. Ind. Electron. 70 (4) (2023) 4186–4195.

[23] D. Chen, R. Liu, Q. Hu, S.X. Ding, Interaction-aware graph neural networks for fault diagnosis of complex industrial processes, IEEE Trans. Neural Networks Learn. Syst. 34 (9) (2023) 6015–6028.

[24] Z. Chen et al., "Graph neural network-based fault diagnosis: a review," arXiv preprint arXiv:2111.08185, 2021.

[25] Z. An, X. Jiang, J. Cao, R. Yang, X. Li, Self-learning transferable neural network for intelligent fault diagnosis of rotating machinery with unlabeled and imbalanced data, Knowl.-Based Syst. 230 (2021) 107374.

[26] J. Li, R. Huang, J. Chen, J. Xia, Z. Chen, W. Li, Deep self-supervised domain adaptation network for fault diagnosis of rotating machine with unlabeled data, IEEE Trans. Instrum. Meas. 71 (2022) 1–9.

[27] Y. Zhou, Y. Dong, G. Tang, Time-varying online transfer learning for intelligent bearing fault diagnosis with incomplete unlabeled target data, IEEE Trans. Ind. Inf. 19 (6) (2023) 7733–7741.

[28] B. Pang, Q. Liu, Z. Sun, Z. Xu, Z. Hao, Time-frequency supervised contrastive learning via pseudo-labeling: An unsupervised domain adaptation network for rolling bearing fault diagnosis under time-varying speeds, Adv. Eng. Inf. 59 (2024) 102304.

[29] M. Shi, et al., Deep hypergraph autoencoder embedding: An efficient intelligent approach for rotating machinery fault diagnosis, Knowl.-Based Syst. 260 (2023) 110172.

[30] R. Zhang, Y. Zhang, C. Lu, X. Li, Unsupervised graph embedding via adaptive graph learning, IEEE Trans. Pattern Anal. Mach. Intell. 45 (4) (2023) 5329–5336.

[31] Z. Wang, Z. Wu, X. Li, H. Shao, T. Han, M. Xie, "Attention-aware temporal–spatial graph neural network with multi-sensor information fusion for fault diagnosis,", Knowl.-Based Syst. 278 (2023) 110891.

[32] L. Wang, H. Cao, H. Xu, H. Liu, A gated graph convolutional network with multi-sensor signals for remaining useful life prediction, Knowl.-Based Syst. 252 (2022) 109340.

[33] W. Wan, J. Chen, J. Xie, MIM-Graph: A multi-sensor network approach for fault diagnosis of HSR Bogie bearings at the IoT edge via mutual information maximization, ISA Trans. 139 (2023) 574–585.

[34] Z. Wu, H. Jiang, X. Wang, H. Zhu, Knowledge correlation graph-guided multi-source interaction domain adaptation network for rotating machinery fault diagnosis, ISA Trans. (2023).

[35] S. Sun, H. Huang, T. Peng, C. Shen, D. Wang, A data privacy protection diagnosis framework for multiple machines vibration signals based on a swarm learning algorithm, IEEE Trans. Instrum. Meas. 72 (2023) 1–9.

[36] S. Sun, H. Huang, T. Peng, D. Wang, An improved data privacy diagnostic framework for multiple machinery components data based on swarm learning algorithm, IEEE Trans. Instrum. Meas. 72 (2023) 1–9.

[37] C. Yang, J. Liu, K. Zhou, M.-F. Ge, X. Jiang, Transferable graph features-driven cross-domain rotating machinery fault diagnosis, Knowl.-Based Syst. 250 (2022) 109069.

[38] T. Li, Z. Zhao, C. Sun, R. Yan, X. Chen, Multireceptive field graph convolutional networks for machine fault diagnosis, IEEE Trans. Ind. Electron. 68 (12) (2021) 12739–12749.

[39] Z. Chen, J. Xu, T. Peng, C. Yang, Graph convolutional network-based method for fault diagnosis using a hybrid of measurement and prior knowledge, IEEE Trans. Cybern. 52 (9) (2022) 9157–9169.

[40] M. Niepert, M. Ahmed, K. Kutzkov, Learning convolutional neural networks for graphs, International Conference on Machine Learning, 2016.

[41] J. Zhou, Y. Qin, D. Chen, F. Liu, Q. Qian, Remaining useful life prediction of bearings by a new reinforced memory GRU network, Adv. Eng. Inf. 53 (2022) 101682.

[42] T. Peng, C. Shen, S. Sun, D. Wang, Fault Feature Extractor Based on Bootstrap Your Own Latent and Data Augmentation Algorithm for Unlabeled Vibration Signals, IEEE Trans. Ind. Electron. 69 (9) (2022) 9547–9555.

[43] Z. Zhi, L. Liu, D. Liu, C. Hu, Fault detection of the harmonic reducer based on CNN-LSTM with a novel denoising algorithm, IEEE Sens. J. 22 (3) (2021) 2572–2581.