





Data Privacy Protection Diagnostic Algorithm for Industrial Robot Joint Harmonic Reducers Based on Swarm Learning

Haodong Huang , Shilong Sun , *Member, IEEE*, Dong Wang , *Member, IEEE*, and Wenfu Xu , *Senior Member, IEEE*

Abstract—Harmonic reducers play a crucial role in industrial robots. Their high load capacity and low friction performance make them highly favored. However, obtaining a large amount of high-quality data on all factory faults is not easy in actual industrial applications. At the same time, data sharing between factories is limited due to privacy concerns. To address this challenge, this article proposes an innovative solution by integrating convolutional neural networks (CNNs) into a swarm learning (SL) framework. In this framework, multiple factories act as edge computing nodes, sharing data features through the fusion of network parameters without directly sharing the data itself. First, we use CNNs to train each node and select a decision-maker before training to merge the model parameters. Secondly, the decision-maker chosen by SL collects the models from other nodes. Finally, the decision-maker disseminates the integrated model to the other nodes. We validated the proposed method using a harmonic reducer dataset and confirmed its reliability. The experimental results show that the proposed framework can improve computational efficiency without relying on a central server, and the shared model can also improve the fault diagnosis accuracy of each edge node.

Index Terms—Data privacy protection, fault diagnosis, harmonic reducer, swarm learning (SL) algorithm.

I. INTRODUCTION

As Intelligent manufacturing technology deepens its application in the automotive and aerospace industries, the

Received 25 October 2024; accepted 5 January 2025. Recommended by Technical Editor W. Wang and Senior Editor D. Chen. This work was supported in part by the National Natural Science Foundation of China under Grant 52475112, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515012041, in part by the Basic Research Program of Shenzhen under Grant JCYJ20220818102415034, and in part by the Shenzhen Higher Education Stability Support Plan under Grant GXWD20231130195340002. (Corresponding author: Shilong Sun.)

Haodong Huang, Shilong Sun, and Wenfu Xu are with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China, and also with Shenzhen and Guangdong Provincial Key Laboratory of Intelligent Morphing Mechanisms and Adaptive Robotics, Shenzhen 518055, China (e-mail: hhd1340201839@163.com; sunshilong@hit.edu.cn; wfxu@hit.edu.cn).

Dong Wang is with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: dongwang4-c@sjtu.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMECH.2025.3528212>.

Digital Object Identifier 10.1109/TMECH.2025.3528212

1083-4435 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

reliability of industrial robot systems becomes exceptionally crucial. Minor mechanical failures or collisions can cause robot systems to halt temporarily, leading to significant economic losses for the entire production system. Against this backdrop, the timely diagnosis and prediction of robot malfunctions emerge as urgent issues that modern industry needs to address, especially against the backdrop of future manufacturing and Industry 4.0. As an indispensable component of industrial robot systems, harmonic reducers are widely used in numerous fields, including aerospace, due to their compact size, outstanding load-bearing capacity, and high precision [1], [2], [3], [4]. However, the complex internal structure of harmonic reducers, combined with long-term operation under high load and torque conditions, leads to a higher rate of failure. Over long-term operation, harmonic reducers may experience various failures, such as wave generator stuttering, flexible wheel pitting, and output end bearing misalignment. Moreover, the rotation of industrial robot joints within specific angular ranges, frequent acceleration and deceleration, and incomplete rotation operations are common factors that increase the risk of gear failures. These factors collectively impact the reliability and repeat precision of robots. Given this, conducting research on fault diagnosis methods for harmonic reducers is of vital importance for enhancing the overall reliability of industrial robots.

In recent years, with the growing demand for intelligent manufacturing in the industry, research on fault diagnosis technology for a key component of industrial robots—the harmonic reducer—has also been significantly intensified. These studies aim to ensure the efficient and reliable operation of robot systems through precise monitoring and diagnosis of faults. Among the many methods, acceleration sensors are widely used to collect vibration signals, thereby implementing condition monitoring and fault diagnosis [5]. Currently, fault diagnosis technologies for harmonic reducers are mainly divided into two categories: data-driven and model-driven. For instance, Zhou et al. [6] proposed a fault diagnosis method for industrial robot harmonic reducers based on deep learning of time-domain continuous vibration signals. This method combines a one-dimensional (1-D) convolutional neural network with a matrix kernel adaptive model, resulting in higher prediction accuracy compared to traditional 2-D CNNs, support vector machine (SVM), and others. Mo et al. [7] proposed a dynamic bandit tree algorithm to improve adaptive filters and frequency band searching,

optimizing the boundaries of Meyer wavelet filters to alleviate parameter tuning burden and enhance the effectiveness of machine fault diagnosis. Zhi et al. [8] introduced a joint fault detection method combining wavelet region correlated threshold denoizing algorithm and convolutional neural network-long short term memory, which also outperforms traditional methods in detection results. Yang et al. [9] proposed an intelligent fault detection method for complete harmonic drives under various working conditions, introducing a new multiscale convolutional neural network architecture. It allows for automatic feature extraction and classification without relying on manual signal processing and empirical knowledge. Li et al. [10] proposed a novel fault diagnosis model based on capsule neural networks and deep reinforcement learning, aiming to address the issue of domain mismatch in industrial applications under different operating conditions, thereby achieving better diagnostic performance and adaptability. Jia et al. [11] proposed a harmonic reducer fault diagnosis scheme that utilizes Hidden Markov models to describe dynamic laws, extracting feature state transitions and observation probabilities to establish the mapping relationship between excitation and monitoring signals. Lu et al. [12] proposes a novel dynamic modeling approach, termed as graph-modeled wavelet packet coefficients, to address early warning detection and fault identification in rolling element bearings, validated through experiments. He et al. [13] introduced a field fault diagnosis method based on a multiscale mixed convolutional neural network model (MSMNN). This MSMNN model is specifically designed for multijoint industrial robot harmonic reducers with multiscale characteristics capable of extracting comprehensive and complementary fault features from complex onsite multichannel signals. Lee et al. [14] analyzed the failure modes and stress effects on the main components of harmonic reducers, extrapolating accelerated test data under three stress levels and using statistical methods to estimate their lifetime under rated conditions. Qiao et al. [15] proposed a nonlinear spectral feature fusion method for fault diagnosis of industrial robot rotational vector (RV) reducers. Nonlinear output frequency response functions are used to extract nonlinear spectra, accurately describing the nonlinear mechanisms of fault phenomena.

In practical application scenarios, harmonic reducers may encounter numerous faults that have the potential to halt the operation of robotic arms temporarily. Given cost considerations, the duration for maintaining these faulty states is often very limited, which severely restricts the volume of high-quality data collected. This situation poses a significant challenge to the widely adopted fault diagnosis techniques based on deep learning today. These techniques tend to rely on large datasets and may exhibit weaker recognition capabilities when faced with fault types characterized by smaller data volumes.

To overcome the challenge of data scarcity and improve the effectiveness of deep learning-based techniques, data augmentation has become a necessary approach. Currently, the main strategies can be categorized into two broad types, one of which involves data generation and expansion through generative adversarial networks (GANs) [16], [17]. Yang et al. [18] proposed a fusion diagnostic model CGAN-2-D-CNN that combines

conditional generative adversarial networks (CGANs) with 2-D convolutional neural networks (2-D-CNNs) for small-sample bearing fault diagnosis. Lyu et al. [19] developed a general anomaly detection method for catenary support components based on GANs, combining deep convolutional neural networks (DCNNs) with GANs to estimate fault occurrences and issue alerts to stop accidents. Pu et al. [20] introduced a semisupervised learning framework, OCGAD, capable of handling multiple fault diagnosis tasks, including fault detection and novelty detection using only normal data, as well as fault classification for unlabeled data. Zhou et al. [21] proposed a new GAN model aimed at generating more discriminative fault samples through a global optimization scheme, addressing the issue of inaccurate fault feature extraction due to imbalanced data in the field of fault diagnosis.

The second approach to address data insufficiency involves aggregating data from multiple locations. However, given that data has become a significant virtual asset for enterprises, few are willing to share it with external parties. In this context, the concept of federated learning (FL) emerged as a solution to this dilemma. Lu et al. [22] proposed a class-imbalanced privacy-preserving federated learning, in response to industrial big data privacy protection and wind turbine data. Additionally, the data volume owned by different servers varies, and a noise gradient mechanism is introduced to prevent client gradient tracking. Liu et al. [23] proposed an asynchronous FL algorithm aimed at addressing the challenge of data islands in photovoltaic power stations. Compared to individual photovoltaic stations, data aggregated from multiple stations provides a richer set of fault samples. Meanwhile, the asynchronous nature of the algorithm effectively resolves the issue of computing power disparities among different photovoltaic station servers. Despite these advantages, FL faces the issue of high communication costs. Fig. 1(a) illustrates the FL network structure.

From the research mentioned above, we can identify the drawbacks of FL. Warnat-Herresthal et al. [24] introduced an innovative learning approach, swarm learning (SL), which significantly reduces the cost of communication between nodes and operates without a central server, ensuring equal rights among nodes, with each having the opportunity to become a temporary central server. Building on this, Sun et al. [25] proposed an SL framework that combines adversarial domain networks with CNNs to address the issue of data insufficiency. Furthermore, in the same study, Sun et al. [26] proposed modifying the models of each edge computing node so that different nodes have different models, further protecting data privacy while also endowing the network with a certain level of interpretability. Fig. 1(b) illustrates the SL network structure.

To achieve centralized aggregation of data while ensuring that data privacy is adequately protected, this article adopts an SL framework for the effective integration of information. Moreover, this research focuses on the harmonic reducers of industrial robot joints as the core subject of article. For each node in the network, we have adopted a unique approach to sample and study both identical and different models. The core innovations and main contributions of this article are summarized as follows.

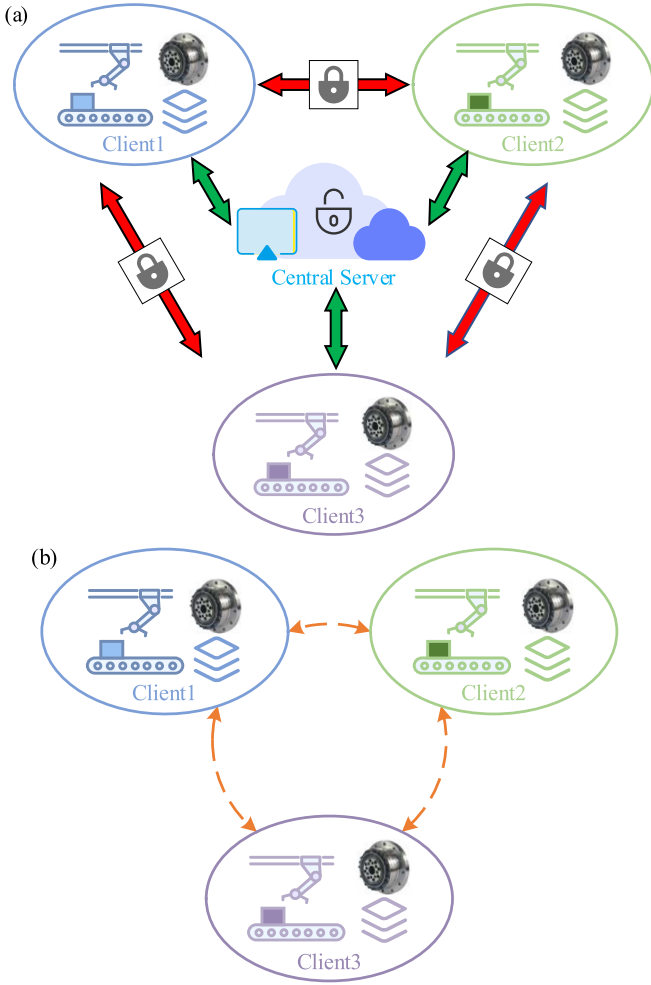


Fig. 1. Structural concept diagrams of two methods. (a) FL: with a central server. (b) SL: without a central server.

- 1) By adopting SL as the framework for information integration, this article effectively addresses the issues of data privacy protection and the “data island” phenomenon, promoting the rational utilization of data resources.
- 2) This article investigates models under two scenarios: one where each node in the network uses the same model, and another where each node applies different models. The aim is to explore the impact of model diversity on the effectiveness of model fusion.
- 3) Through experimental validation on harmonic reducers, the results demonstrate that the proposed method can effectively process signals from different devices, further proving the universality of the algorithm and showcasing its potential applicability to other devices or components.

The rest of this article is organized as follows. Section II provides a brief overview of SL and its related research. Section III elaborates on the method proposed in this article, including the local training model architecture for each node and the process of aggregating model parameters through the SL framework. Section IV validates the effectiveness of the proposed method with vibration signal data from harmonic reducers collected in the laboratory. Finally, Section V concludes this article.

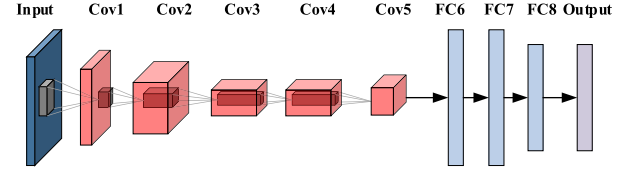


Fig. 2. Diagram of the AlexNet architecture, which consists of five convolutional layers and three fully connected layers.

II. PRELIMINARIES

A. AlexNet

AlexNet [27] is a DCNN that won the championship in the 2012 ImageNet large scale visual recognition challenge, significantly improving the accuracy of image classification tasks. The structure of AlexNet includes five convolutional layers, three fully connected layers, and the rectified linear unit (ReLU) function for nonlinear activation. It also incorporates several innovative techniques to enhance performance and reduce overfitting, such as max pooling, local response normalization, dropout, and a wide range of data augmentation techniques. The network structure of AlexNet is shown in Fig. 2.

In this article, the data processed is from industrial robot harmonic reducers, which are 1-D data. Therefore, a 1-D AlexNet is used to extract features from the vibration data. The computation of the 1-D convolutional layer can be represented as follows:

$$a_i^l = f \left(\sum_m w_m^l \cdot x_{i+m}^{l-1} + b^l \right) \quad (1)$$

where a_i^l is the activation value at position i in layer l , f is the activation function, and w_m^l represents the weight at position m in layer l . x_{i+m}^{l-1} is the input value at position $i+m$ in layer $l-1$, and b^l is the bias for layer l . After the convolution operation, a large number of features are obtained, necessitating the introduction of a 1-D pooling layer to reduce the feature dimensionality and simultaneously enhance the model’s generalization capability. Furthermore, by reducing the number of parameters and the model’s sensitivity to specific data, the pooling layer helps improve the model’s generalization ability, thereby reducing the risk of overfitting. The formula for max pooling can be represented as follows:

$$p_i^l = \max_{m \in M} x_{i+m}^l \quad (2)$$

where p_i^l is the pooling output at position i in layer l . M is the size of the pooling window and x_{i+m}^l is the input value within the pooling window at layer l .

The fully connected layer connects all neurons between the input and output data, enabling feature extraction, mapping to the output space, introducing nonlinearity, and learning parameters. This allows the network to learn complex data patterns to perform classification tasks. The computation formula for the fully connected layer is as follows:

$$a^l = f(\mathbf{W}^l \cdot \mathbf{a}^{l-1} + \mathbf{b}^l) \quad (3)$$

where a^l is the activation value vector for layer l , f is the activation function, \mathbf{W}^l is the weight matrix for layer l , and b^l is the bias vector for layer l .

B. Swarm Learning (SL)

SL is an advanced distributed artificial intelligence technology designed to share learning outcomes between nodes rather than raw data, to protect data privacy and achieve more effective decision-making and prediction. This technology mimics the phenomenon of swarm intelligence observed in nature, such as flocks of birds and colonies of ants, by enabling individual nodes to learn locally and exchange trained model parameters without centralized data storage. This not only reduces dependence on data transmission, enhancing processing efficiency, but also significantly improves data security and privacy protection. SL overcomes the privacy and security limitations of traditional centralized learning models by utilizing distributed learning and aggregated updates, thereby increasing the system's scalability, flexibility, and robustness. Thus, SL is considered a powerful tool capable of ensuring data privacy while enabling effective learning and decision-making across domains. Fig. 1(b) shows a schematic diagram of the SL network structure.

In SL, before each epoch of training, a decision-maker is randomly selected. After each epoch of training, this decision-maker needs to collect the models from other nodes participating in the training, and integrate their own model parameters with those of the others. After the model integration, the integrated model is distributed to each edge node, followed by the next round of training. This process is repeated until the training concludes. The formula for model updating is as follows:

$$\mathbf{m}_{j+1} = \frac{1}{n} \sum_{i=1}^n \mathbf{m}_{j,i} \quad (4)$$

where \mathbf{m}_{j+1} represents the integrated model parameters, which \mathbf{m} consists of two parts (\mathbf{W} and \mathbf{b}). i denotes the number of nodes, and $\mathbf{m}_{j,i}$ represents the model parameters obtained by the node after the j th round of training. The model integration process is shown in Fig. 3.

C. Laplace Wavelet

Wavelet analysis is a powerful tool for localized processing in the time-frequency domain, contrasting with the Fourier transform. The Fourier transform focuses on the frequency components of a signal, whereas the wavelet transform captures changes in both time and frequency. By scaling (affecting frequency) and shifting (affecting time) the original signal, the wavelet transform achieves multi-scale analysis of the signal. Through its unique properties of decay and oscillation, the wavelet transform allows for detailed analysis and reconstruction of signals at various scales. Wavelets can be represented as follows:

$$WT(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt. \quad (5)$$

From the formula, it is evident that wavelets differ from the Fourier transform. The Fourier transform has only one variable,

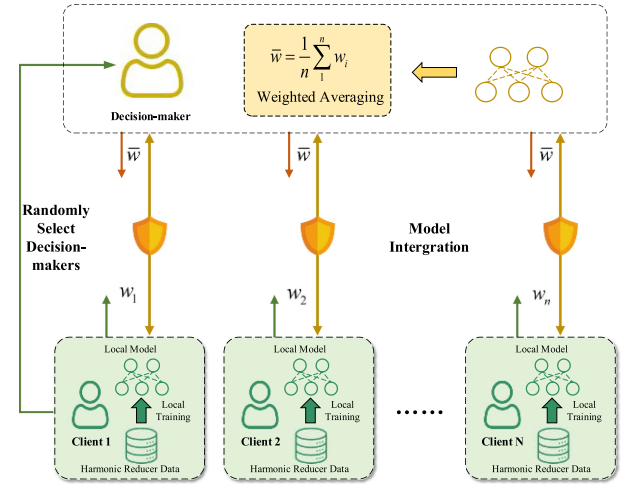


Fig. 3. Model integration process.

frequency, while wavelets have two variables: the scale factor a and the translation factor b . $\psi^*(\cdot)$ is the complex conjugate of the mother function $\psi(\cdot)$. Li et al. [28] conducted comparative experiments by using wavelets as the first layer in CNNs and found that the Laplace wavelet, with its spiral decay characteristic, can better extract the impact part. It can be represented as follows:

$$\begin{aligned} \psi(\omega, \zeta, \tau, t) &= \psi_\gamma(t) \\ &= \begin{cases} A e^{-\frac{\zeta}{\sqrt{1-\zeta^2}} \omega(t-\tau)} e^{-j\omega(t-\tau)}, & t \in [\tau, \tau + W] \\ 0, & \text{others} \end{cases} \end{aligned} \quad (6)$$

where $\gamma = \langle \omega, \zeta, \tau \rangle$ represents the parameters of the Laplace wavelet, ω represents the frequency, determining the oscillation frequency of the Laplace wavelet, ζ is the damping ratio which makes the Laplace wavelet decay rapidly, τ is the time parameter, A is used to normalize the wavelet function, and W indicates the width of the wavelet's support interval.

D. Chebyshev I Filter

The amplitude response formula of the Chebyshev I filter is used to describe the filter's gain or attenuation at different frequencies. Chebyshev filters have a steeper roll-OFF characteristic compared to Butterworth filters, but at the cost of ripple in the passband. For the Chebyshev I filter, the expression for its amplitude response is as follows:

$$|H_n(j\omega)| = \frac{1}{\sqrt{1 + \varepsilon^2 T_n^2 \left(\frac{\omega}{\omega_c} \right)}}. \quad (7)$$

$|H_n(j\omega)|$ is the amplitude response of the filter, ω is the angular frequency of the input signal, ω_c is the cutoff frequency of the filter, and ε is the ripple factor, which determines the maximum ripple amplitude within the passband. T_n represents the n th order Chebyshev polynomial, where n is the order of the filter that determines the steepness of the filter's roll-OFF. In a Chebyshev

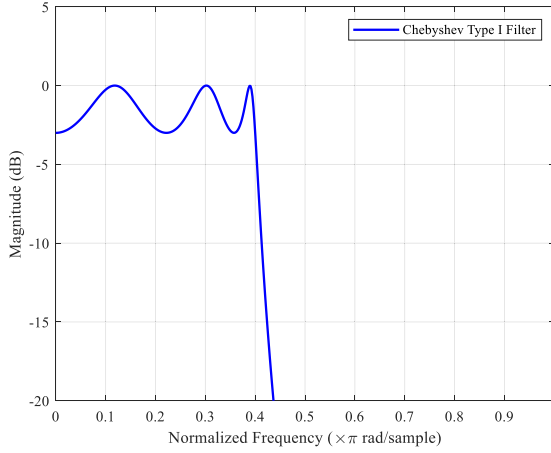


Fig. 4. Chebyshev I filter amplitude-frequency response curve.

I filter, the ripple within the passband is determined by the ε value, with the ripple amplitude increasing as ε increases. The higher the order n , the faster the filter attenuates near the cutoff frequency, but the more pronounced the ripple within the passband. This type of filter achieves rapid attenuation over a short frequency range. The filter's amplitude-frequency curve is shown in Fig. 4.

III. PROCEDURE OF THE PROPOSED METHOD

This section describes the proposed method for diagnosing faults in industrial robot joint harmonic reducers in scenarios with limited samples. The proposed method can automatically identify the state of the reducer from the signal. The main focus of this article is the three-node problem, and the flowchart of the proposed method is shown in Fig. 5. The training algorithm is presented in Algorithm 1.

Algorithm 1: Fault Diagnosis of Industrial Robot Harmonic Reducers Based on SL.

- I. Data preprocessing
 - II. Model preparation w, b
 - III. Building an SL framework
 - IV. For any node, $i = 1, 2, 3$
 - Choose $i = k_{th} \in (1, 2, 3)$ as the decision-maker
 - V. Model training
 - 1) For decision-maker $m_{j,k_{th}} = (w_{th}|b_{th})$, for the other two nodes

$$m_{j,2} = (w_{j,1}, w_{j,2}|b_{j,1}, b_{j,2})$$
 - 2) Model updates

$$m_j = \frac{1}{3}(m_{j,1} + m_{j,2} + m_{j,k_{th}})$$
 - 3) Model distribution

$$m_{j+1} = m_j$$
 - VI. Model validation
-

IV. CASE STUDY

To investigate the collaborative fault diagnosis method for industrial robot harmonic reducers further based on SL, this

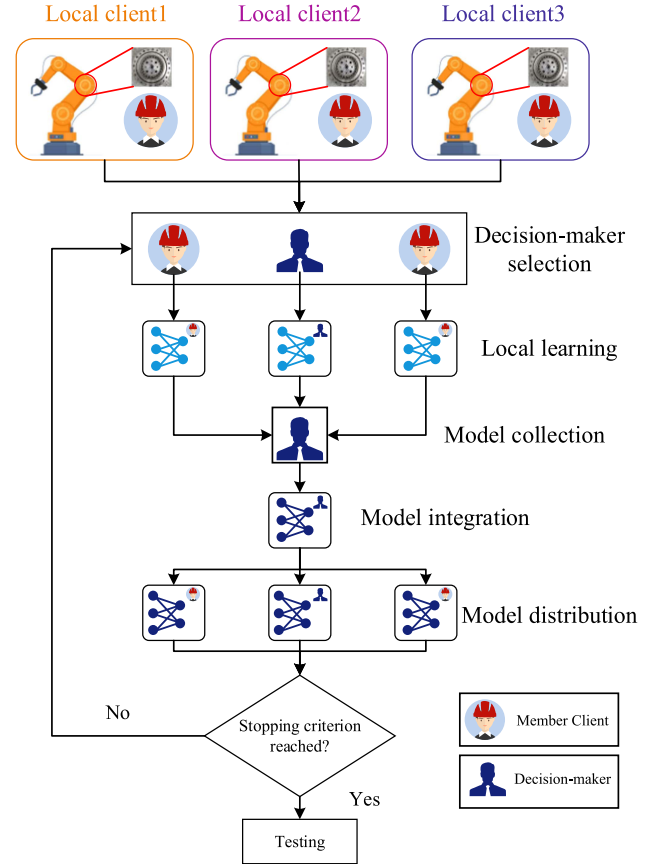


Fig. 5. Workflow of fault diagnosis for harmonic reducers is based on SL network. It is divided into two parts: local training and model integration. First, the model is trained locally, then the model is passed to the decision maker, who integrates the model and distributes it to clients.

article collected data by conducting experiments on a dedicated harmonic reducer fault testing platform established in the laboratory.

A. Dataset Preparation

The experiment involves three nodes, all of which are taken from the harmonic reducer fault test bench in the laboratory. Details of the harmonic reducer fault test bench are shown in Fig. 6. There are a total of four health states for the harmonic reducer: healthy, wave generator stuttering, flexible wheel pitting, and output end bearing misalignment. Fig. 7 illustrates some of the prefabricated faults. Together with the fault-free harmonic reducers, there are a total of four types of labels.

The time-domain and frequency-domain plots of the harmonic reducer data are shown in Fig. 8. The health condition of the harmonic reducer is divided into four categories. Considering the limited sample problem faced in this article, the dataset for each node contains 404 sets of data, each with a length of 2400. In each node, 200 sets of data are used for model training, while the remaining 204 sets are used for model testing. For each type of fault, we use 50 sets of data to train the diagnostic model and 51 sets to test the diagnostic efficacy of the model. The conditions of each dataset are given in Table I.

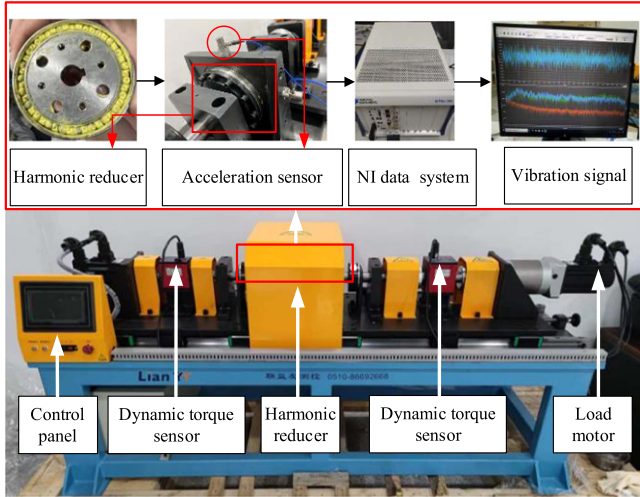


Fig. 6. Harmonic reducer fault test system diagram. The fault test bench includes NI acquisition equipment, harmonic reducer, motor, and other components.

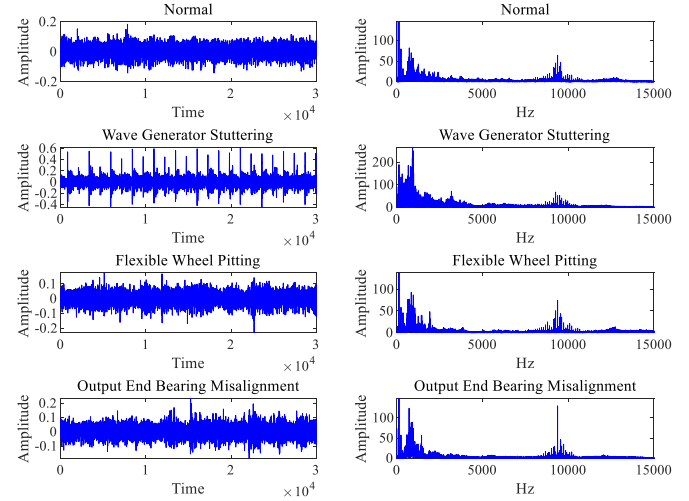


Fig. 8. Waveform diagram of the harmonic reducer is shown below: the left side is the time-domain graph, and the right side is the frequency-domain graph.

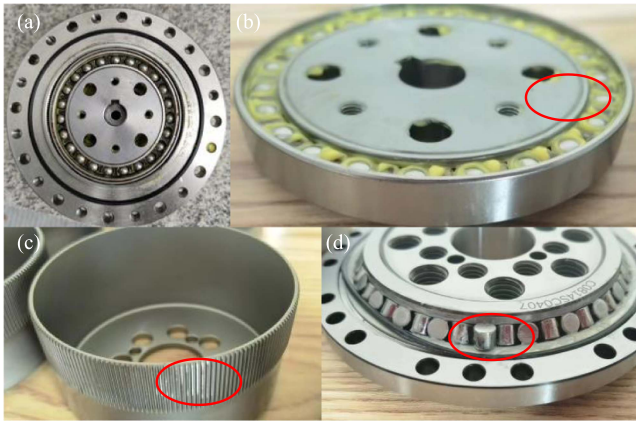


Fig. 7. Harmonic reducer fault images. (a) Normal. (b) Wave generator stuttering. (c) Flexible wheel pitting. (d) Output end bearing misalignment.

TABLE I
OPERATING CONDITIONS OF A HARMONIC REDUCER DATASET
WITH THREE NODES

Node Number	Load	Speed (rpm)	Sampling Frequency (Hz)
1	0	600	25.6k
2	0	1200	25.6k
3	0	1500	25.6k

B. Data Preprocessing

Due to the high gear ratio of the harmonic reducer, the information contained in the 2400 data points is limited. Additionally, the data from the harmonic reducer is influenced by factors, such as flexible thin-wall bearings and other noise, resulting in insufficiently distinct differences in the collected data. Therefore, it is necessary to preprocess the data. Envelope spectrum analysis is selected as the preprocessing method for the data. The formula

TABLE II
STRUCTURAL PARAMETERS OF ALEXNET

Layer	Tied parameter	Activation function
Input	/	/
Convolutional layer1	Kernel size: 32;Filters: 64;Stride: 1	ReLU
Pooling layer1	Pool_size: 3;Stride: 2	/
Convolutional layer2	Kernel size: 5;Filters: 192;Stride: 2	ReLU
Pooling layer2	Pool_size: 3;Stride: 2	/
Convolutional layer3	Kernel size: 3;Filters: 384;Stride: 1	ReLU
Convolutional layer4	Kernel size: 3;Filters: 256;Stride: 1	ReLU
Convolutional layer5	Kernel size: 3;Filters: 256;Stride: 1	ReLU
Pooling layer3	Pool_size: 3;Stride: 2	/
flatten	/	/
Dense1	Units: 1024	ReLU
Dense2	Units: 1024	ReLU
Dense3	Units: 4	SoftMax

is as follows:

$$\hat{x}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(t-\tau)}{\tau} d\tau = x(t) \cdot \frac{1}{\pi} \quad (8)$$

$$\tilde{x}(t) = x(t) + j\hat{x}(t) = A(t) e^{j\varphi(t)}. \quad (9)$$

$A(t)$ is the Hilbert envelope of $x(t)$. Performing Fourier transform on this $A(t)$ yields the envelope spectrum. The collected signal is divided into sub-signals of size 404×2400 . After obtaining the envelope, the data becomes 404×1200 , where each label's data consists of 101×1200 . Then, the data is normalized. Table II shows the details of the structural parameters of AlexNet network.

TABLE III
HYPERPARAMETER VALUES

Hyperparameter	Value
epoch	300
learning rate	0.0002
N_{train}	50
batch size	16
N_{test}	51

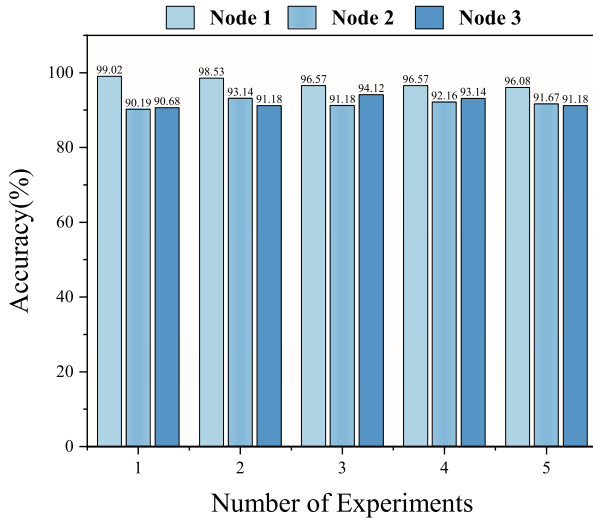


Fig. 9. Diagnostic experimental results of the harmonic reducer from 5 repeated trials within the SL framework.

C. AlexNet Architecture Parameters

D. Hyperparameter Settings

The implementation of harmonic reducer fault diagnosis based on the SL framework uses the Python-based PyTorch framework. The experimental hyperparameter settings are given in Table III.

E. Experimental Results and Comparisons

By placing the above dataset into the SL framework and conducting repeated experiments, the results obtained are shown in the Fig. 9.

Compared to the bearing data, the data from the harmonic reducer exhibit higher complexity. Nevertheless, the experimental results demonstrate high accuracy, confirming the feasibility of the experimental approach. Additionally, the short runtime of the experiments allows for swift identification and localization of any faults in the robotic arm joints once they occur.

To demonstrate the advantages of the proposed method in this article, we will conduct a series of comparative experiments for further analysis. These experiments will involve examining the dataset sizes and comparing the different model SL (DSL) with the local learning strategy.

1) *Experiment I: The Impact of Dataset Size on Experimental Results:* This experiment investigated the impact of training set size on the performance of fault diagnosis for harmonic

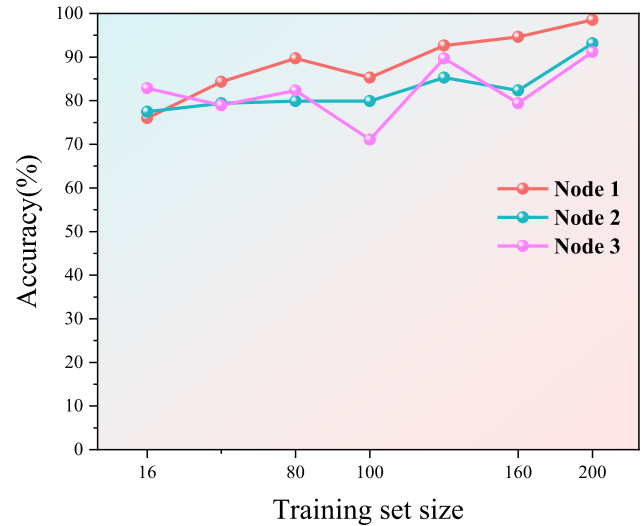


Fig. 10. Impact of the training dataset size on the diagnostic results of the harmonic reducer. As the size of the training set increases, it incorporates more fault information, leading to more accurate diagnostic outcomes.

reducer within the SL framework. Envelope spectrum was still adopted as the input to the network for this article. Considering a batch size set to 16, dataset sizes were varied to 16, 40, 80, 100, 120, 160, and 200 to observe the specific influence of different data scales on diagnostic accuracy. The experimental results are illustrated in Fig. 10.

The experimental results clearly show that as the size of the training set increases, the overall diagnostic accuracy significantly improves. Specifically, when there are only four sets of data for each fault type at each node, the accuracy at all nodes exceeds 75%. As the data volume further increases, the accuracy improvement is most pronounced for the first node, while the second and third nodes show relatively slower improvements. This phenomenon suggests that the proposed SL framework exhibits strong diagnostic capabilities even with small training set sizes. However, under conditions of extremely limited data, although the framework maintains high accuracy, its performance improvement becomes relatively constrained. This further validates the effectiveness and adaptability of the method in data-scarce environments.

2) *Experiment II: Comparison With DSL [26]:* In this experiment, we compare the proposed SL method with DSL [26]. The DSL method replaces the first layer of the AlexNet model with Laplace wavelet and Chebyshev Type I filter, as shown in Fig. 11. During the experiment, except for the replacement of the first layer, all other experimental parameters and network structures remain unchanged. The experimental results are given in Table IV.

From Table IV, it can be observed that the SL-based method proposed in this article achieves higher diagnostic accuracy compared to the DSL-based method. This result indicates that the data complexity of the harmonic reducer is high, and the application of the Laplace transform and filters has certain limitations in effectively capturing key fault features of the harmonic

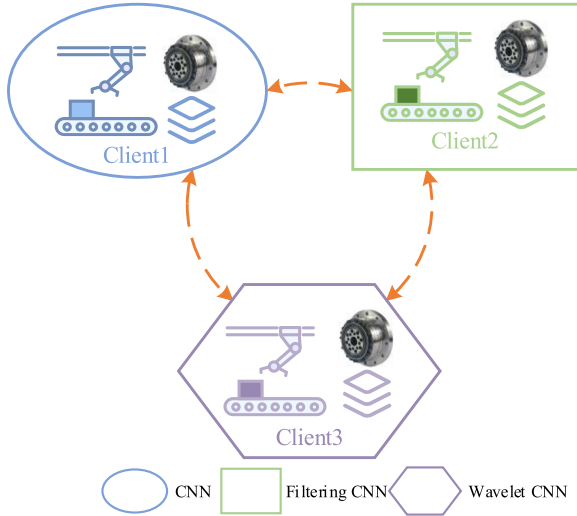


Fig. 11. DSL network architecture diagram. The local model is different for each node.

TABLE IV
REPEATING EXPERIMENTS FOR SL AND DSL

Node\ACC(%)	1	2	3	4	5	Average/Standard Deviation
SL node1	99.01	98.53	96.57	96.57	96.08	97.352 ± 1.32
DSL node1	85.78	91.18	86.76	85.78	84.80	86.86 ± 2.51
SL node2	90.19	93.14	91.18	92.16	91.67	91.668 ± 1.10
DSL node2	82.84	85.78	83.82	84.80	93.63	86.174 ± 4.31
SL node3	90.68	91.18	94.12	93.14	91.18	92.06 ± 1.49
DSL node3	91.67	91.18	89.22	93.63	90.69	91.278 ± 1.60

TABLE V
COMPARISON BETWEEN SL AND LOCAL LEARNING

Accuracy (%)	SL	Local learning	Model 1	Model 2	Model 3
Node 1	98.53	92.15	92.15	90.68	44.12
Node 2	93.14	81.37	78.92	82.35	49.02
Node 3	91.18	76.96	79.01	51.47	79.90

The bold values represents the results of our proposed methods SL.

reducer. This inadequacy in feature extraction affects effective communication between different models, thereby constraining the diagnostic accuracy of the DSL-based approach. In contrast, the SL method better addresses the data complexity, demonstrating stronger feature extraction capability and higher diagnostic accuracy, further validating the advantages of this method in the fault diagnosis of harmonic reducers.

3) Experiment III: Comparison With Local Learning: The experiment compares SL with local learning, which is divided into two parts. The first part adopts the proposed SL method, while the second part uses local training, where there is no model interaction among the three nodes. All training and testing processes are completed locally on each node. The experimental results are given in Table V.

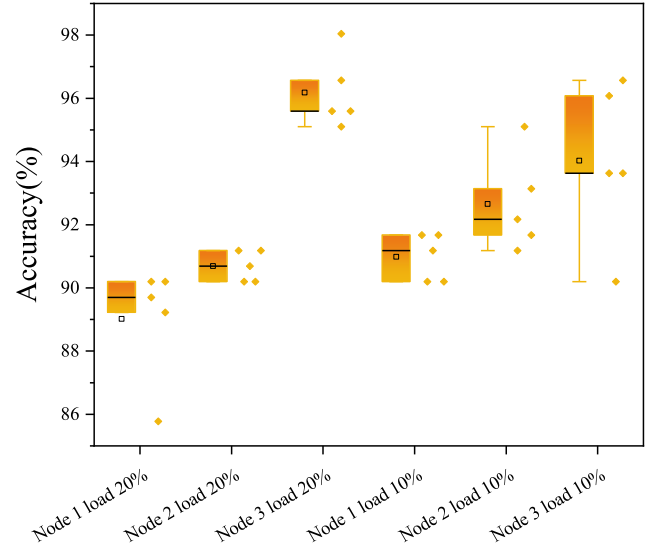


Fig. 12. Diagnostic results of three nodes in SL under load.

Table V gives experimental results comparing SL with local learning methods. Model 1, model 2, and model 3 represent the models obtained through local training on nodes 1, 2, and 3, respectively. Since both the training and testing data in local learning originate locally, it typically results in higher accuracy models. However, the experimental results demonstrate that the outcomes obtained using the SL methods is significantly superior to those based on local learning across all three nodes. This indicates that SL can significantly improve model testing accuracy by enhancing the quantity of data features in situations with limited and nonshared data.

In contrast, the local learning approach not only fails to achieve higher accuracy locally but also exhibits poorer model transferability to other nodes. It is noteworthy that the data used by these three nodes all originate from the same harmonic reducer fault testbed. If the data were sourced from different devices, the transfer accuracy of local models might further decrease. These findings further underscore the effectiveness of SL in improving fault diagnosis accuracy, especially when facing challenges associated with data sharing limitations.

4) Experiment IV: Load Data Experiment: To simulate the working conditions of a real robotic arm, load testing and data collection were conducted on the harmonic reducer fault test bench. During the loading process, it was observed that when the load reached 30%, the test bench experienced sudden stalling, indicating that it could not handle loads exceeding 30%. Therefore, four different health states of the harmonic reducer were selected, starting from no load and gradually increasing to 10% and 20% loads to simulate the actual working conditions of industrial robotic arms. To ensure data reliability, each experiment was repeated five times. The detailed experimental results are shown in Fig. 12.

From the box plot, it can be observed that after adding the load, the diagnostic accuracy of each node remains above 90%. However, when the load increases from 10% to 20%, there is a slight decline in accuracy for nodes 1 and 2, while node 3 shows

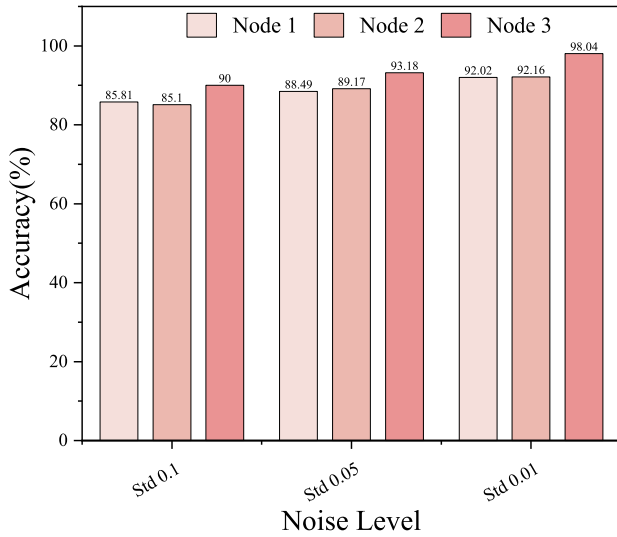


Fig. 13. Impact of noise on experimental results.

an improvement. This indicates that nodes 1 and 2 are more significantly affected by the load, whereas node 3, due to the averaging nature of the model, exhibits an increase in accuracy. Overall, the method proposed in this manuscript demonstrates good diagnostic performance under both loaded and unloaded conditions.

5) Experiment V: Add Noise Experiment: In real-world industrial robotic arm operations, various environmental noises are inevitable. To simulate these working conditions, Gaussian noise was artificially added to the collected data. To reflect environmental variations, we designed three sets of experiments by introducing Gaussian noise with standard deviations of 0.01, 0.05, and 0.1 into the data. The experimental results after adding noise are shown in Fig. 13.

Fig. 13 shows that as the noise level increases, diagnostic accuracy experiences a slight decline. When the standard deviation is 0.01, the accuracy of each node remains above 90%. However, as the noise level rises, the accuracy of the first two nodes decreases to above 80%, while the accuracy of the third node remains above 90%. This indicates that the fault diagnosis of the harmonic reducer based on the SL framework has a certain degree of noise resistance. However, if the noise is too high, the robustness of the diagnosis will be compromised. Therefore, in cases of excessive noise, it is necessary to perform some noise reduction data processing.

6) Experiment VI: Comparison With FL: We compared the proposed SL-based method with FL. To ensure fairness, we kept the training hyperparameters consistent and conducted five repeated experiments for FL as well. The experimental results are given in Table VI.

From the table, FL has lower accuracy at nodes 1 and 2 compared to SL, indicating that the features were not well preserved during the gradient integration process. Additionally, we will discuss the computational costs based on the model parameters and training time.

From Table VII, we can observe that the number of local model parameters in SL is the same as that in FL. However,

TABLE VI
REPEATING EXPERIMENTS FOR FL

Acc(%)	Node 1	Node 2	Node 3
Number			
1	82.84	83.82	93.62
2	81.86	81.86	94.11
3	81.37	83.82	92.64
4	80.88	86.27	93.63
5	80.88	88.73	94.12
Avg	81.56±0.01138	84.90±0.02097	93.62±0.00379

TABLE VII
COMPARISON OF COMPUTATIONAL COST

	Number of parameters	Data transfer times
FL	10679796	6/epoch
SL	10679796	4/epoch

our case involves three nodes, so an additional central server needs to be set up when using FL. In each training epoch, SL randomly selects a decision node, resulting in four model data transmissions. In the case of FL, the central server must collect models from the three nodes, integrate them, and then distribute the combined model, which requires six model data transmissions. Clearly, SL reduces the number of data transmissions. In terms of training time, the differences among the methods are not significant due to the relatively low complexity of the problems addressed in this article.

V. CONCLUSION

This article proposes a collaborative fault diagnosis algorithm for industrial robot joint harmonic reducers based on the SL framework. The uniqueness of this diagnostic framework lies in its protection of data privacy, short training time, and effective handling of limited data challenges. The main contributions of this article can be summarized as follows.

- 1) Achieved data privacy protection by sharing the characteristics of data without sharing the actual data.
- 2) Investigated the properties of SL-based fault diagnosis for harmonic reducers.
- 3) Achieved high-precision diagnosis on complex data sets involving the relatively novel harmonic reducers.

Through a series of comparative experiments, this article validates the effectiveness and feasibility of the proposed method. Therefore, this research not only provides a forward-looking solution for fault diagnosis of industrial robot joint harmonic reducers but also its generality allows it to be applied to the fault diagnosis of other devices or components, demonstrating its broad application potential.

REFERENCES

- [1] H. Li, S. Yang, L. Kong, and T. Wen, "High-precision angular speed tracking control of gimbal system with harmonic reducer," *IEEE Trans. Ind. Electron.*, vol. 69, no. 8, pp. 8168–8177, Aug. 2022.
- [2] H. Dong, B. Dong, C. Zhang, and D. Wang, "An equivalent mechanism model for kinematic accuracy analysis of harmonic drive," *Mechanism Mach. Theory*, vol. 173, Jul. 2022, Art. no. 104825.

- [3] M. A. Costa, B. Wullt, M. Norrlöf, and S. Gunnarsson, "Failure detection in robotic arms using statistical modeling, machine learning and hybrid gradient boosting," *Measurement*, vol. 146, pp. 425–436, Nov. 2019.
- [4] B. j. Jung, B. Kim, J. C. Koo, H. R. Choi, and H. Moon, "Joint torque sensor embedded in harmonic drive using order tracking method for robotic application," *IEEE/ASME Trans. Mechatron.*, vol. 22, no. 4, pp. 1594–1599, Aug. 2017.
- [5] Y. Lv, W. Zhao, Z. Zhao, W. Li, and K. K. H. Ng, "Vibration signal-based early fault prognosis: Status quo and applications," *Adv. Eng. Inform.*, vol. 52, Apr. 2022, Art. no. 101609.
- [6] X. Zhou, H. Zhou, Y. He, S. Huang, Z. Zhu, and J. Chen, "Harmonic reducer in-situ fault diagnosis for industrial robots based on deep learning," *Sci. China Technol. Sci.*, vol. 65, no. 9, pp. 2116–2126, Sep. 2022.
- [7] Z. Mo, Z. Zhang, Q. Miao, and K.-L. Tsui, "Intelligent informative frequency band searching assisted by a dynamic bandit tree method for machine fault diagnosis," *IEEE/ASME Trans. Mechatron.*, vol. 28, no. 2, pp. 770–780, Apr. 2023.
- [8] Z. Zhi, L. Liu, D. Liu, and C. Hu, "Fault detection of the harmonic reducer based on CNN-LSTM with a novel denoising algorithm," *IEEE Sensors J.*, vol. 22, no. 3, pp. 2572–2581, Feb. 2022.
- [9] G. Yang, Y. Zhong, L. Yang, and R. Du, "Fault detection of harmonic drive using multiscale convolutional neural network," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 3502411.
- [10] G. Li, J. Wu, C. Deng, X. Xu, and X. Shao, "Deep reinforcement learning-based online domain adaptation method for fault diagnosis of rotating machinery," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 5, pp. 2796–2805, Oct. 2022.
- [11] Y. Jia, Y. Li, M. Xu, Y. Cheng, and R. Wang, "A fault diagnosis scheme for harmonic reducer under practical operating conditions," *Measurement*, vol. 227, Mar. 2024, Art. no. 114234.
- [12] G. Lu, X. Wen, G. He, X. Yi, and P. Yan, "Early fault warning and identification in condition monitoring of bearing via wavelet packet decomposition coupled with graph," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 5, pp. 3155–3164, Oct. 2022.
- [13] Y. He, J. Chen, X. Zhou, and S. Huang, "In-situ fault diagnosis for the harmonic reducer of industrial robots via multi-scale mixed convolutional neural networks," *J. Manuf. Syst.*, vol. 66, pp. 233–247, Feb. 2023.
- [14] J.-H. Lee, Y.-H. Cho, D.-S. Kim, and J.-W. Park, "Lifetime estimation of harmonic reducer for manufacturing robot using accelerated life test," *J. Mech. Sci. Technol.*, vol. 36, no. 6, pp. 2879–2887, Jun. 2022.
- [15] Y. Qiao, H. Wang, J. Cao, Y. Lei, and H. Liu, "Nonlinear spectrum feature fusion diagnosis method for RV reducer of industrial robots," *Mech. Syst. Signal Process.*, vol. 204, Dec. 2023, Art. no. 110750.
- [16] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," *Comput. Ind.*, vol. 106, pp. 85–93, 2019.
- [17] W. Zhang, X. Li, X.-D. Jia, H. Ma, Z. Luo, and X. Li, "Machinery fault diagnosis with imbalanced data using deep generative adversarial networks," *Measurement*, vol. 152, 2020, Art. no. 107377.
- [18] J. Yang, J. Liu, J. Xie, C. Wang, and T. Ding, "Conditional GAN and 2-D CNN for bearing fault diagnosis with small samples," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 3525712.
- [19] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, "A generic anomaly detection of catenary support components based on generative adversarial networks," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 5, pp. 2439–2448, May 2020.
- [20] Z. Pu, D. Cabrera, Y. Bai, and C. Li, "A one-class generative adversarial detection framework for multifunctional fault diagnoses," *IEEE Trans. Ind. Electron.*, vol. 69, no. 8, pp. 8411–8419, Aug. 2022.
- [21] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization GAN for unbalanced data," *Knowl.-Based Syst.*, vol. 187, 2020, Art. no. 104837.
- [22] S. Lu, Z. Gao, Q. Xu, C. Jiang, A. Zhang, and X. Wang, "Class-imbalance privacy-preserving federated learning for decentralized fault diagnosis with biometric authentication," *IEEE Trans. Ind. Inform.*, vol. 18, no. 12, pp. 9101–9111, Dec. 2022.
- [23] Q. Liu et al., "Asynchronous decentralized federated learning for collaborative fault diagnosis of PV stations," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 3, pp. 1680–1696, May/Jun. 2022.
- [24] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, Jun. 2021.
- [25] S. Sun, H. Huang, T. Peng, C. Shen, and D. Wang, "A data privacy protection diagnosis framework for multiple machines vibration signals based on a swarm learning algorithm," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 3501309.
- [26] S. Sun, H. Huang, T. Peng, and D. Wang, "An improved data privacy diagnostic framework for multiple machinery components data based on swarm learning algorithm," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 3529009.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, pp. 84–90, 2012.
- [28] T. Li et al., "WaveletKernelNet: An interpretable deep neural network for industrial intelligent diagnosis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 4, pp. 2302–2312, Apr. 2022.



Haodong Huang was born in Jingmen, China. He received the bachelor's degree in mechanical engineering from the School of Mechanical Engineering and Automation with Northeastern University, Shenyang, China, in 2021 and the M.S. degree in mechanical engineering in 2024 from Harbin Institute of Technology, Shenzhen, China, where he is currently working toward the Ph.D. degree in mechanical engineering with the School of Mechanical Engineering and Automation.

His research interests include distributed training and humanoid robots.



Shilong Sun (Member, IEEE) received the Ph.D. degree in engineering from the City University of Hong Kong, Hong Kong, in 2018.

He is currently an Assistant Professor with Harbin Institute of Technology, Shenzhen, China. He nurtures keen interests in vibration energy harvesting design, fault diagnosis and prognosis, decision-making with Artificial Intelligence, and deep learning for industrial data. Now, he focuses on the remaining equipment life estimation research with deep learning, smart energy harvesting techniques and humanoid robots.



Dong Wang (Member, IEEE) received the Ph.D. degree in engineering from the City University of Hong Kong, Hong Kong, in 2015. He was a Senior Research Assistant, a Postdoctoral Fellow, and a Research Fellow with the City University of Hong Kong. He is currently an Associate Professor with the Department of Industrial Engineering and Management, Shanghai Jiao Tong University, Shanghai, China, where he is also with the State Key Laboratory of Mechanical System and Vibration. His research interests

include sparse and complex measures, signal processing, prognostics and health management, condition monitoring and fault diagnosis, statistical learning and machine learning, statistical process control, and nondestructive testing.

Dr. Wang is an Editorial Board Member for Mechanical Systems and Signal Processing. He is an Associate Editor for IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, *Measurement*, IEEE SENSORS JOURNAL and *Journal of Dynamics Monitoring and Diagnostics*.



Wenfu Xu (Senior Member, IEEE) received the B.E. and M.E. degrees in control engineering from the Hefei University of Technology, Hefei, China, in 2001 and 2003, respectively, and the Ph.D. degree in the control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2007.

He was a Research Associate with the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong. He is currently a Professor with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, China. His research interests include reconfigurable robots, space robots, and bionic robots.