A Curvature-Guided Shape DTW for Anomaly Detection of Industrial Robot Joints Under Variable Operating Conditions

Jiabing Gu, Xing Wu, Xiaoqin Liu, *Member, IEEE*, Kai Xu, Dongxiao Wang, and Ran Ren

Abstract—As a core component of intelligent manufacturing equipment, it is particularly important to detect the abnormality of industrial robot. However, the variable motion conditions of robot joints lead to strong time-varying and discontinuous monitoring signals, which pose a challenge for anomaly detection. To address this, this study proposes an anomaly detection method based on current time–frequency ridge alignment. First, the current time–frequency ridge is preprocessed based on peak-frequency normalization, which effectively mitigates the speed variation and ridge extraction error interference. Subsequently, an improved ridge alignment algorithm named curvature-guided shape dynamic time warping (CSDTW) is proposed to further eliminate the impact of variable operating conditions. This algorithm innovatively integrates local curvature into the shape Dynamic time warping (DTW), this advancement solves the overfitting and feature destruction problems of conventional DTW. Finally, the aligned ridges are transported to the convolutional neural network (CNN) for discriminative feature learning and fault classification. Experimental validation on an industrial robot joint dataset demonstrates that CSDTW has the highest ridge alignment accuracy under variable operating conditions compared to the traditional DTW method, improving the anomaly detection accuracy of the CNN for multiple fault types. The method provides a reliable solution for detecting anomalies in industrial robots under variable operating conditions.

Index Terms—Anomaly detection, dynamic time warping (DTW), industrial robot, time–frequency ridge, variable operating conditions.
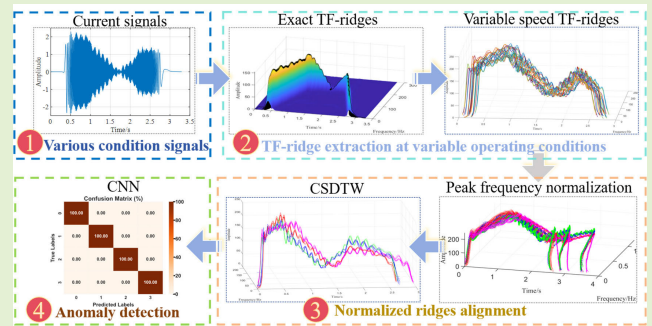
I. INTRODUCTION

WITH the continuous development of robotics and automation technology, industrial robots have become an indispensable part of industrial production due to their high efficiency and reliability [1], [2]. However, the execution of tasks involving multiple motion trajectories, frequent acceleration and deceleration, and impact loads makes the robot joints extremely susceptible to abnormalities [3], [4]. Failure to detect an abnormality in time will lead to the decreased stability and reliability of industrial robots, resulting in huge economic losses and even casualties [5]. Therefore, anomaly detection technology under variable operating conditions is of great value in ensuring the safe operation of industrial robots [6], [7].

Currently, researchers have conducted a lot of studies on anomaly detection methods, and a variety of research neighborhoods have emerged, such as machine learning-based methods [8], [9], signal processing-based methods [10]. In the machine learning neighborhood, deep learning is developing rapidly. For example, non-contact detection and neuro-fuzzy system-guided cross-modal zero-sample diagnosis [11] have become emerging trends in research. This further indicates that cross-fertilization of method systems between different neighborhoods is gradually becoming a new research trend. Lei et al. [12] used a meta-learning framework incorporating prior knowledge to successfully realize the diagnosis of scarce samples under variable operating conditions. Feng et al. [13] proposed an intelligent assessment method based on digital twin and migration learning, established a gear degradation

model, and combined with a cross-domain feature alignment algorithm to realize the migration assessment of gear wear state. Yu et al. [14] designed a deep fusion feature extraction network for extracting and fusing local and global features of multi-scale time–frequency information. The local and global time–frequency features are fused with migration learning to successfully realize the diagnosis task under variable operating conditions. Zhang et al. [15] proposed a new multimodal cross-domain fusion network, which introduces the maximum mean difference loss and realizes the cross-domain feature alignment within modes, and ultimately realizes the fault diagnosis of gearboxes under variable operating conditions. In the signal processing neighborhood, anomaly detection for different speed conditions mainly includes signal processing methods based on time–frequency analysis (TFA) [16], [17] and order tracking (OT) [18], [19], etc. Chen and Feng [20] successfully realized gear fault detection by capturing the frequency sideband information of gear faults under variable rotational speed conditions using a high-resolution polynomial chirp transform. Chen et al. [21] obtain a high-precision time–frequency representation through an iterative envelope tracking filter, which extracts the fault eigenfrequencies for anomaly detection. However, TFA methods often exhibit a strong dependence on the time–frequency resolution, which limits their adaptability across varying signal conditions. Furthermore, these methods generally lack robustness to noise, posing significant challenges in practical applications. In order to solve these problems, Xu et al. [22] proposed to enhance the fault characterization of rolling bearings under variable rotational speed conditions using the stochastic resonance technique and bistable system. However, the system parameters of this method are difficult to optimize, and the speed of the target shaft needs to be obtained from a tachometer. Li et al. [23] then proposed to utilize the tacholess OT method for fault detection, and the main idea is to estimate the instantaneous angular speed by extracting the time–frequency ridges. However, the OT-based algorithm does not completely eliminate the amplitude modulation effect due to excessive speed variations [24]. Moreover, the aforementioned studies have primarily focused on fault detection under single operating conditions, while anomaly detection under variable operating conditions needs further research.

The inherent limitations of TFA and OT methods in handling complex signal variations and noise have driven scholarly attention toward robust temporal alignment techniques. Dynamic time warping (DTW) is a time series alignment algorithm, and is widely used in many fields such as speech recognition [25], gesture recognition [26], and time series classification [27], among numerous other fields [28], [29], [30]. DTW eliminates the deviation of two sequences from variable operating conditions on the time scale by dynamically searching for matching paths to align the two sequences together [31], [32]. Han et al. [33] proposed to utilize multi-scale DTW to align the time-domain waveforms to eliminate the length variations due to variable rotational speeds, which ultimately successfully identified the faults. Sun et al. [34] proposed current-assisted DTW to align the

variable-speed vibration signal with the uniform signal template and finally successfully detect planetary gearbox faults by envelope order spectrum. However, this method requires the simultaneous acquisition of vibration and current signals, and it is difficult to ensure reliable and accurate anomaly detection results due to the large error in the results caused by signal transmission and noise interference. Sharma and Parey [35] align the intrinsic modal function by DTW to obtain the frequency domain average to extract the weak fault characteristics at variable speed. Hong et al. [36] align the shaft rotational harmonic signals with the template signals using DTW to estimate the shaft rotational speed and finally identify the gearbox faults under variable speed conditions. However, for signals with complex variations, DTW produces an overfitting phenomenon when aligning, i.e., one point matches multiple points, which makes the aligned portion behave as a straight line. This undoubtedly destroys the characteristics of the time series and makes it difficult to detect anomalies. Zhao and Itti [37] and others proposed an improved alignment algorithm, Shape DTW, to increase the ability of local alignment through shape descriptors. However, this algorithm ignores the overall continuity when facing complex and variable local shapes, and is prone to shape feature destruction at the connection of local shapes. Therefore, many researches have been carried out in the direction of constraining the search region and improving the search pattern, but there is a lack of methods that retain the feature continuity and at the same time provide a good constraint on the path offset.

In summary, the following research gaps exist in this study: insufficient dynamic adaptation and algorithmic robustness under multiple operating conditions; high dependence on time–frequency resolution and rotational speed stability; and overfitting and localized feature destruction generated by DTW.

To address these limitations, this study proposes a curvature-guided shape DTW (CSDTW) algorithm for robust time–frequency ridge alignment under variable operating conditions. The CSDTW dynamically sets the step pattern by the curvature of the ridges and searches them under the distance matrix computed by the Shape DTW. The aligned ridges are used in a convolutional neural network (CNN) for training and ultimately for anomaly detection. Simulation analyses validate the effectiveness of the CSDTW, while experiments on the RV reducer swing arm test bench of an industrial robot have shown that multiple fault types under variable operating conditions are accurately detected based on CSDTW. The contributions of this work are as follows.

1) In this study, a framework for anomaly detection under variable operating conditions is proposed. The framework eliminates the effect of variable conditions by peak frequency normalization and CDSTW alignment. Multiple fault types under three different operating conditions were successfully recognized.

2) The proposed CSDTW takes into account the variation of the local curvature of the signal, which improves the overfitting and feature destruction problems of DTW and Shape DTW while eliminating the influence of the working conditions. The optimization of the input signal
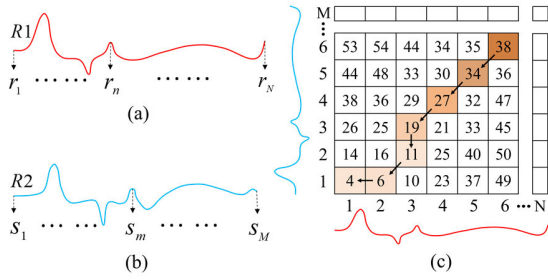
Fig. 1. DTW algorithm. (a) Time series $R1$. (b) Time series $R2$. (c) DTW match path.

Fig. 2. DTW produces overfitting when dealing with two sequences with large local shape differences.

quality of the CNN model is achieved, and the anomaly detection accuracy is improved.

3) 3-D time–frequency ridges of motor currents are extracted in this study, containing both frequency modulation and amplitude modulation, which enhances the reliability of fault detection.

This article is structured as follows. Section II introduces the principle of the DTW algorithm and mathematically derives the CSDTW. Section III presents the time–frequency ridge alignment anomaly detection framework based on the CSDTW. Section IV presents simulation experiments and analysis. Section V describes the robotic arm test bench experiments and results. Section VI concludes.

II. Theoretical Background

A. Dynamic Time Warping

DTW aligns two time series by calculating the distance matrix between them and searching for matching paths from the matrix by step pattern [38]. As shown in Fig. 1(a) and (b), there are two time series $R1 = (r_1, \ldots, r_n, \ldots, r_N)$ and $R2 = (s_1, \ldots, s_m, \ldots, s_M)$ with the different lengths of N and M ($n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$). Fig. 1(c) is the DTW cumulative distance matrix, where each element is calculated by

$$D_{n,m} = E_{n,m} + \min\{D_{n-1,m}, D_{n-1,m-1}, D_{n,m-1}\} \quad (1)$$

where $D_{n,m}$ denotes an element in the cumulative distance matrix D ($R1$, $R2$), the subscript denotes its position in the matrix. $E_{n,m}$ represents the squared Euclidean distance between the nth point on $R1$ and the mth point on $R2$, calculated as

$$E_{n,m} = \sum_{i=1}^{v} (r_n - s_m)^2 \quad (2)$$

where v denotes the dimension of the time series.

The first value of the matrix is represented as $D_{1,1} = E_{1,1}$. As shown by the arrows in Fig. 1(c), the algorithm starting from $D(N, M)$ at the end of the matrix, the path with the smallest cumulative distance is reversed and gradually returned to the starting point $D(1,1)$. The search process is step-by-step and is performed by the following rules:

$$D_{n,m}^* = \min\{D_{n,m}, D_{n+1,m}, D_{n,m+1}\} \quad (3)$$

where $D_{n,m}^*$ is a point on the match path and the subscript indicates the index on the original sequence. Therefore, the
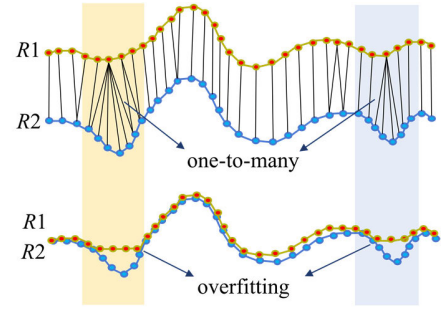
matching relationship can be represented by indexing the value of the original sequence

$$P_t = \{(r_1, s_1), \ldots, (r_n, s_m), \ldots, (r_N, s_M)\} \quad (4)$$

where P is the matching relationship, (r_n, s_m) represents the match between the nth point on $R1$ and the mth point on $R2$, ultimately realizing that all points of the two sequences are aligned together.

As shown in Fig. 2, DTW often results in a phenomenon where a single point matches multiple points when aligning curves with complex local shapes, leading to aligned local curves that become straight line segments. The reason for this is because DTW focuses on the global optimum and completely ignores the local alignment case. This overfitting leads to loss of information of sequences, and it should be avoided.

B. Curvature-Guided Shape DTW

DTW due to its local overfitting, researchers proposed Shape DTW to improve the local matching performance. Shape DTW converts the traditional point-to-point matching to high-dimensional feature-to-feature matching through shape descriptors. A shape descriptor is a mapping function about a time series that produces values representing local shape features of the time series. Given a point T of a time series, the subsequence R_T at that point has specific local shape structure information, and the shape descriptor encodes R_T to represent the local shape structure of the point T, the result of this encoding is called the shape descriptor d of the point T, and the encoding process is a mapping function $F(\cdot)$, i.e., $d = F(R_T)$.

Given two time series $R1 = (r_1, \ldots, r_n, \ldots, r_N)$ and $R2 = (s_1, \ldots, s_m, \ldots, s_M)$, computing the descriptors for each point produces descriptor sequence $d^{R1} = (d^{r1}, \ldots, d^{rn}, \ldots, d^{rN})$ and $d^{R2} = (d^{s1}, \ldots, {}^{sm}, \ldots, d^{sM})$. Shape DTW achieves alignment by resolving the optimization problem

$$\arg\min_{l, P^{R1} \in \{0,1\}^{l \times N}, l, P^{R2} \in \{0,1\}^{l \times M}} \sum_{i=1}^{k} \left\| P^{R1} \cdot d^{R1} - P^{R2} \cdot d^{R2} \right\|_2 \quad (5)$$

where l is the length of matching path, P^{R1} and P^{R2} are DTW matching paths for d^{R1} and d^{R2}, respectively. $\| \cdot \|_2$ is Euclidean norm and k is the number of rows of the matrix
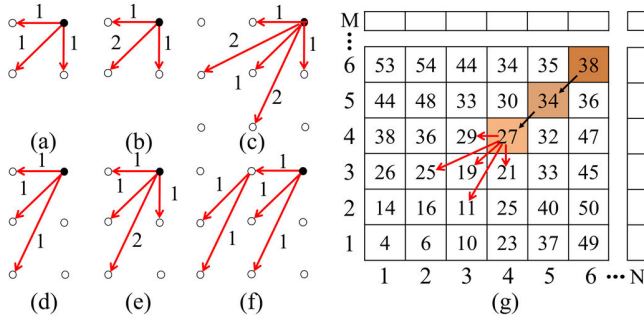
Fig. 3. Different step pattern. (a)–(c) Symmetric class. (d)–(f) Asymmetric class. (g) Search process for matching path.

Fig. 4. CSDTW selects different step patterns based on sequence shape characteristics.

of $[P^{R1} \cdot d^{R1} - P^{R2} \cdot d^{R2}]$. Finally, the matched paths are transferred to the original sequences to achieve alignment.

However, Shape DTW tends to produce feature and structure destruction in the connecting part of two complex local shapes when confronted with consecutive complex local shapes, the reason for this phenomenon is that the matching of Shape DTW to the local shapes is superior to that of the whole, resulting in the sacrifice of the smoothness and integrity of the curve itself. Meanwhile, Shape DTW is more sensitive to parameters, and when the parameters are not set properly, the destruction of features is more obvious. Therefore, to satisfy both the overall smoothness of the alignment and to produce good alignment locally, this study proposes that CSDTW dynamically adjust the search of matching paths by calculating the curvature.

CSDTW introduces step patterns Sp based on Shape DTW, as shown in Fig. 3, where different Sp and weights ω affect the searched paths. To simplify the calculation, Fig. 3(a) and (b) are used as the main search patterns in this study. The Sp in Fig. 3(a) includes three search directions: horizontal, vertical, and diagonal, with different weight values. The step pattern is defined as: Sp = $[\omega_1, \omega_2, \omega_3]$, it is used to search for matching path in $D(R1, R2)$

$$D^*_{n,m} = \min \left\{ Sp \left[D_{n+1,m}, D_{n,m+1}, D_{n,m} \right] \right\}. \tag{6}$$

Fig. 4 illustrates the optimization idea of CSDTW. When the local shapes of $R1$ and $R2$ change violently, increasing the values of ω_1 and ω_2 and decreasing ω_3 will ensure excellent local alignment. When the local shapes of $R1$ and $R2$ change gently, increasing the values of ω_3 and decreasing ω_1 and ω_2 will accelerate the matching speed, thus avoiding the destruction of features. The size of the weight increase or decrease needs to be decided according to the specific time series. Therefore, three step patterns are set for selection: Sp1 = $[\omega_{11}, \omega_{21}, \omega_{31}]$, Sp2 = $[\omega_{21}, \omega_{22}, \omega_{23}]$, and Sp3 = $[\omega_{31}, \omega_{32}, \omega_{33}]$, to adapt to the local variations of the signal.

CSDTW uses curvature as a measure of the degree of change in the shape of a sequence, and thus to select the appropriate step pattern for each moment. The greater the curvature, the more violent the change in shape; as the curvature approaches zero, the sequence tends to be linear or flat. For a univariate time series $R1$, the curvature at point r_N
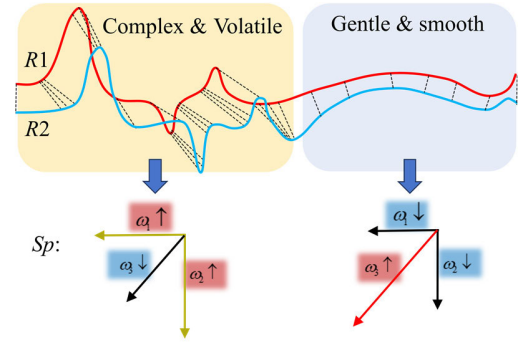
on $R1$ is calculated as follows:

$$C_{r_N} = \frac{\left| R1''(r_N) \right|}{\left(1 + (R1'(r_N))^2 \right)^{3/2}} \tag{7}$$

where $R1'(r_N)$ and $R1''(r_N)$ denote the first derivative and the second derivative at point r_N, respectively.

CSDTW selects the Sp by the normalized C_{r_N} at point r_N

$$Sp_{r_N} = \begin{cases} Sp1, \mu_1 \leq C_{r_N} \leq \mu_2 \\ Sp2, 0 \leq C_{r_N} \leq \mu_1 \\ Sp3, \mu_2 \leq C_{r_N} \leq 1 \end{cases}, \mu_1, \mu_2 \in [0, 1] \tag{8}$$

where μ_1 and μ_2 are thresholds, which are determined by the range of variation in the curvature of the specific time series.

Since μ_1 and $\mu_2 \in [0, 1]$, curvature below 0.3 is considered as slow change and curvature above 0.6 is considered as drastic change. In this study, the thresholds are set to $\mu_1 = 0.3$ and $\mu_2 = 0.6$, three step patterns are set to Sp1 = [1, 1, 1], Sp2 = [0.8, 0.8, 1.5], Sp3 = [1.3, 1.3, 0.8].

C. Algorithm Overview

In this section, the pseudo-code of CSDTW is provided in Algorithm 1. Lines 1–3 go through the process of computing the distance matrix by Shape DTW. Lines 4–13 select the step pattern dynamically by calculating the curvature. Line 14 searches for matching path in the Shape DTW distance matrix by step patterns.

III. ANOMALY DETECTION FRAMEWORK UNDER VARIABLE OPERATING CONDITIONS

This study proposes an anomaly detection framework based on CSDTW. The current signal was used as the monitoring object because it has the advantages of being non-intrusive, low cost, and can be integrated with the robot control system. Peak frequency normalization of current time–frequency ridges is proposed. This method can effectively reduce the effect of variable speed, and the effect of ridge error is avoided to some extent. The effect of different operating conditions is then further eliminated by using CSDTW to align the ridges. Finally, anomaly detection is performed on the aligned ridges using CNN.

Algorithm 1 CSDTW

Inputs: Two sequence $R1$, $R2$ of length N and M; Descriptor function F; Step pattern $Sp1$, $Sp2$, $Sp3$; Thresholds μ_1 and μ_2

Output: matching path P_t

1.   Sample subsequences $r1$, $r2$
2.   Compute descriptors of $r1$ $r2 \rightarrow d^r = F(r1)$, $d^{r2} = F(r2)$
3.   Compute distance matrix of d^{r1}, $d^{r2} \rightarrow D(R1, R2)$
4.   Compute normalized curvature of $R1 \rightarrow C_{R1}$
5.   **for** $i = 1$ to N **do**
6.   **if** $\mu_1 \leq C_{R1}(i) \leq \mu_2$ **then**
7. $Sp = Sp1$;
8.   **if** $0 \leq C_{R1}(i) \leq \mu_1$ **then**
9. $Sp = Sp2$;
10.   **else if** $\mu_2 \leq C_{R1}(i) \leq 1$ **then**
11. $Sp = Sp3$;
12.   **end if**
13. **end for**
14. Compute matching path use Sp in $D(R1, R2) \rightarrow P_t$

A. Peak Frequency Normalization

The instantaneous frequency of the current signal is proportional to the motor speed, and normalization of the current time–frequency ridges removes the effect of motor speed variations. The relationship between them is expressed as follows:

$$n_i = \frac{f_i}{p} \tag{9}$$

where n_i is the motor speed, f_i is the current fundamental frequency, and p is the number of motor pole pairs.

Industrial robot joints often perform reciprocating start-stop motions, resulting in strong time-varying amplitude and frequency of the current signals. Due to the high acceleration of the robot joints, the motor speeds during the execution of the task usually consist of acceleration, uniform speed, and deceleration phases. Fig. 5 shows the motor current signals for two typical operating conditions of the robot joints. We use a dynamic programming algorithm to extract the ridges as a way to ensure the effectiveness and accuracy of the extracted ridges. The reciprocating start and stop of the joint results in low current amplitude and time–frequency energy during the commutation phase, making the extraction of the time–frequency ridges susceptible to errors due to noise, as shown in Fig. 6. Variable operating conditions lead to inconsistencies in the extracted time–frequency ridge lengths and frequency amplitudes. As shown in Fig. 7(a), there is a frequency gap between the time–frequency ridges extracted at two different operating conditions. In order to reduce the effect of different speed conditions, traditional DTW methods often use normalization to preprocess the data. However, the common normalization does not completely eliminate the gap between the current time–frequency ridges, as shown in Fig. 7(b). The reason is that the complex working conditions lead to errors in ridge extraction, which prevents the normalization algorithm from accurately identifying the maximum ridge frequency.

Therefore, this study proposes to use the peak frequency of the spectrum to replace the maximum frequency. It can be
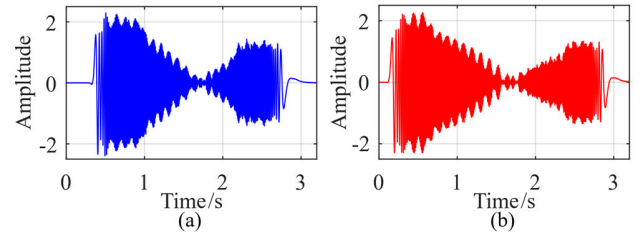

Fig. 5. Robot joint current signals under two typical operating conditions: (a) condition 1 and (b) condition 2.
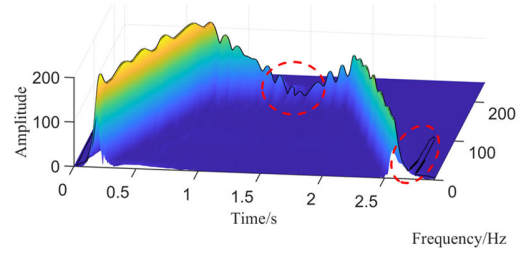

Fig. 6. Schematic representation of the time–frequency ridge extraction error: the part in the dotted circle is prone to ridge extraction error.
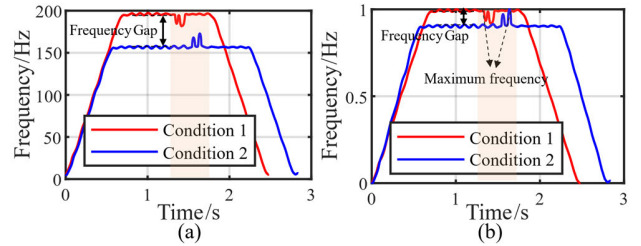

Fig. 7. Time–frequency ridges of the current signal for two operating conditions. (a) Frequency gap. (b) Normalization result.

present as

$$f^*(t) = \frac{f(t) - \min[f(t)]}{f_{\text{Peak}} - \min[f(t)]} \tag{10}$$

where f_{Peak} is peak frequency, $f(t)$ is instantaneous frequency. $f^*(t)$ is the normalized instantaneous frequency.

The peak frequency is the frequency corresponding to the maximum value of the spectrum obtained after the current signal has been Fourier transformed, as shown in Fig. 8. The peak frequency represents the fundamental frequency component of the motor current and reflects the stable operating frequency of the motor current during the uniform phase. Notably, local low amplitude or localized signal anomalies do not perturb the dominant peak frequency of the current, avoiding the problem of maximum frequency identification error. As shown in Fig. 9, The normalization of the peak frequency greatly reduces the effect of ridge extraction errors and eliminates frequency gaps, providing a promising foundation for subsequent ridge alignment. After peak frequency normalization, the difference in frequency amplitude between ridges for different operating conditions is reduced, while the inconsistency in ridge lengths due to variable-speed operating conditions will be compensated by the ridge alignment algorithm, further eliminating the effect of variable operating conditions.
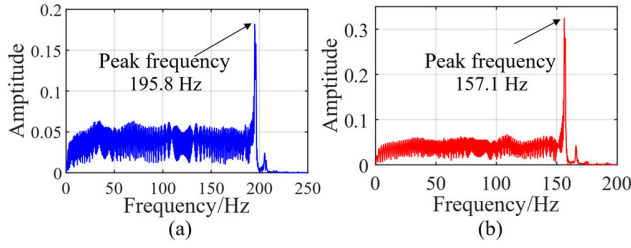
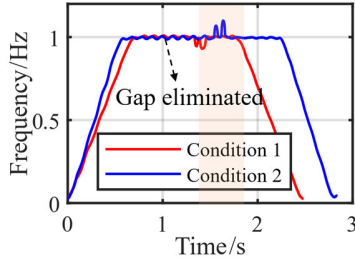Fig. 8. Current signal spectrum: (a) condition 1 and (b) condition 2.

Fig. 9. Frequency gap elimination for different operating conditions by peak frequency normalization.
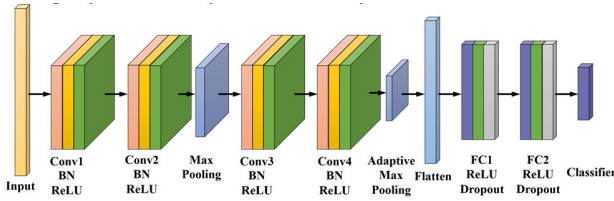

Fig. 10. CNN structure.

B. Convolutional Neural Network

In the field of anomaly detection, CNN has good feature extraction and pattern recognition capabilities. As shown in Fig. 10, the CNN structure mainly consists of convolutional layer, pooling layer, and fully connected layer [39], [40].

In the convolutional layer, each kernel processes the input features through a convolutional operation. This operation effectively extracts local features of the image and recognizes different feature patterns with different kernel functions. Before applying the nonlinear activation function, the value of the feature mapping can be expressed as

$$h_i^l = w_i^l * x + b_i^l \tag{11}$$

where w_i^l denotes the weight of the convolution kernel, and b_i^l denotes the bias term of the convolution kernel. x denotes the input signal and $*$ is the convolution operation.

After the convolution operation, the feature mapping needs to be nonlinearly transformed by an activation function whose expression can be written as

$$z_i^l = f\left(h_i^l\right) \tag{12}$$

where z_i^l represents the feature mapping after activation.

The convolutional layer outputs a multitude of feature maps, which are subsequently subjected to dimensionality reduction using a pooling layer designed to reduce the number of model

parameters while retaining key features. In the mathematical formulation, the pooling operation is defined as follows:

$$y_i^{l+1} = f\left(\beta^{l+1} dw\left(z_i^l\right) + b^{l+1}\right) \tag{13}$$

where dw denotes the down sampling operation; β^{l+1} denotes the weight of the down sampling operation.

However, CNN performance is highly dependent on large-scale labeled data, and the ability to generalize variable working conditions is insufficient. Therefore, the ability of CSDTW to eliminate the influence of operating conditions and retain small feature variations can enhance the detection accuracy of CNN with good interpretability.

The 3-D information of ridges is used as input to the CNN, which is preprocessed as low-pass filtering, intercepting segments, and constructing a time–frequency ridge anomaly dataset.

C. Anomaly Detection Based on CSDTW

DTW and Shape DTW cause overfitting and feature destruction when facing complex-shaped ridges. CSDTW improves Shape DTW to realize better alignment of ridges, which effectively eliminates the effect of variable operating conditions and realizes more efficient and reliable anomaly detection for industrial robots. As illustrated in Fig. 11, this study proposed an anomaly detection framework of industrial robot joints under variable operating conditions based on CSDTW. The main steps are as follows.

Step 1: 3-D ridge extraction and normalization. Time–frequency ridges are extracted from the current signal of variable operating conditions, and the energy of the ridges is also extracted, which contains rich amplitude modulation information. The extracted ridges are normalized using the peak frequency to accurately unify the amplitude of the frequency components of the ridges to the same level in preparation for ridge alignment.

Step 2: Align the ridges of variable operating conditions with the template ridges by CSDTW.

1) The normal ridges serve as templates, and the other ridges are aligned with the templates.
2) The curvature of the ridges is computed, and the step pattern is selected by the magnitude of the curvature value. When the curvature value is large, the search weight is increased for both horizontal and vertical directions. The purpose is to increase the search effort when the shape changes are complex and to find more suitable matching points. Conversely, when the curvature is small, the search weight for diagonal direction is increased to get a faster match.
3) The distance matrix between the ridges and the template ridges is computed, and different step patterns are used to search for matching paths in the matrix. The match path aligns the two ridges by index.

Step 3: The aligned ridges eliminate the effect of variable rotational speed and runtime cycle length and are used for CNN training and testing to obtain accurate and reliable detection results.
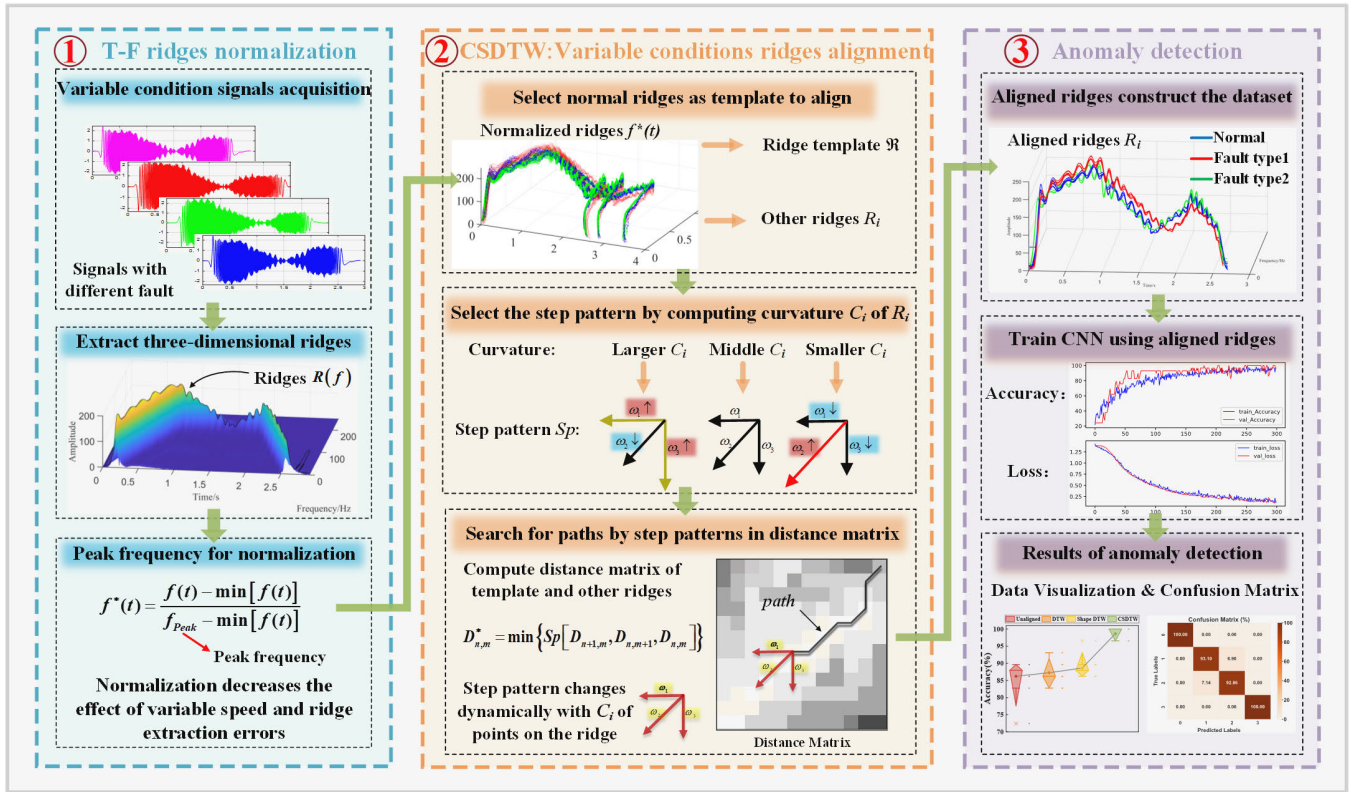
Fig. 11. Anomaly detection process under variable operating conditions based on CSDTW.

IV. SIMULATION ANALYSIS

Simulation experiments are designed to verify the effectiveness of the CSDTW algorithm. The signals acquired from the industrial robot under variable operating conditions are non-stationary current signals, and the amplitude and frequency of the signals are modulated [41]. Therefore, the amplitude and frequency modulation numerical simulation signal are constructed as

$$X(t) = (1 + 0.05t)\cos(2\pi \times 10t(\sin(2\pi t) + 0.1t)) + n(t) \tag{14}$$

where $X(t)$ denotes the final signal and $n(t)$ denotes Gaussian noise.

The simulated signal is a single-component signal with a fundamental frequency of 10 Hz, and the modulation frequency is determined by $\sin(2\pi t) + 0.1t$, the amplitude modulation is $1 + 0.5t$. Nonlinear frequency and amplitude modulation leads to large variations in the curvature of the time–frequency ridges of the signal in the local shape. To simulate different operating conditions, white noise of different intensities is added to the signal. As shown in Fig. 12, 20 and 10 dB of noise are added to the signal.

Different noise levels cause the ridges to have different curvatures locally. As shown in Fig. 13(a), ridges are extracted from the two signals and normalized. The A, B, and C represent the points on the ridges, and the subscripts represent the different ridges. The standard for ridge alignment requires that the crests or troughs are aligned at the same point in time, on the condition that the ridge features remain intact.
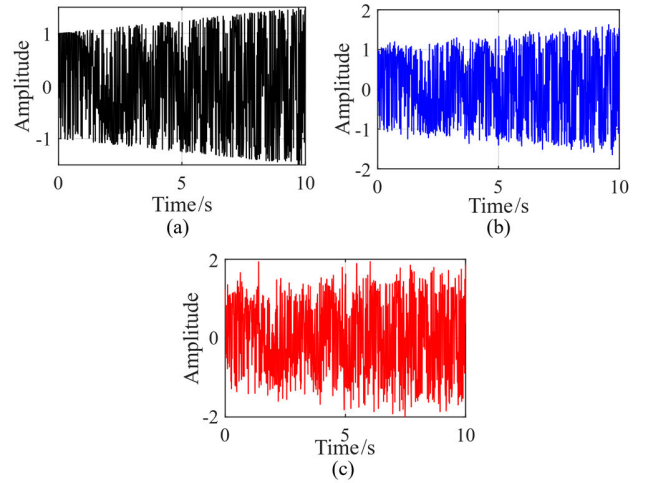

Fig. 12. Simulated signals. (a) No noise. (b) SNR is 20 dB. (c) SNR is 10 dB.

A1 and A2 should be aligned, as well as B1 and B2 or C1 and C2. As shown in Fig. 13(b), the DTW algorithm produces overfitting in the areas where A, B, and C are located. Overfitting refers to the occurrence of a one-to-many matching relationship, resulting in the same value at multiple points in time, which behaves as a straight line. As shown in Fig. 13(c), Shape DTW aligns points A, B, and C without inducing overfitting. However, ridge alignment in the yellow circle is suboptimal, and ridge features are destructed in different degrees. The reason for this is that Shape DTW
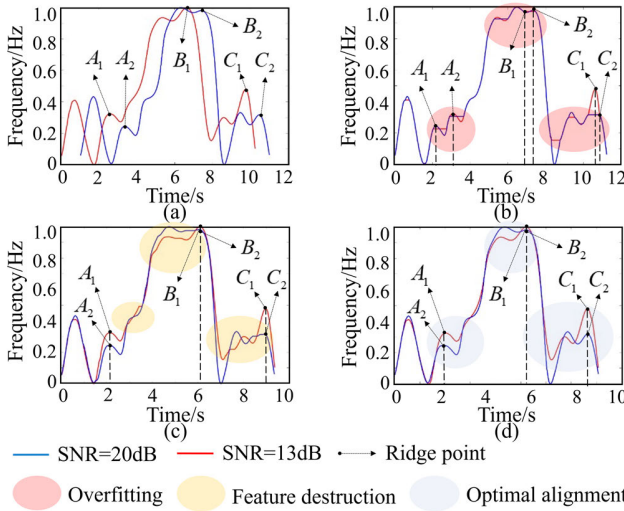
Fig. 13. Alignment result. (a) Before alignment. (b) DTW. (c) Shape DTW. (d) CSDTW.
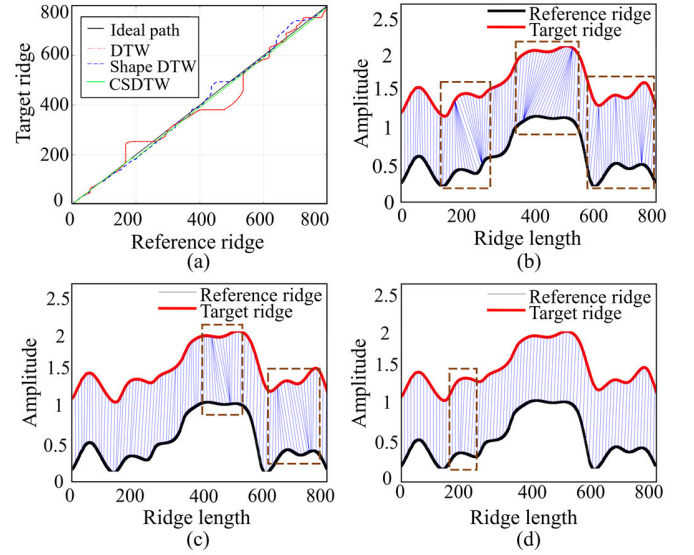

Fig. 14. Visualization of alignment paths and matching relations. (a) Alignment paths for the target and reference ridge. (b) DTW. (c) Shape DTW. (d) CSDTW.
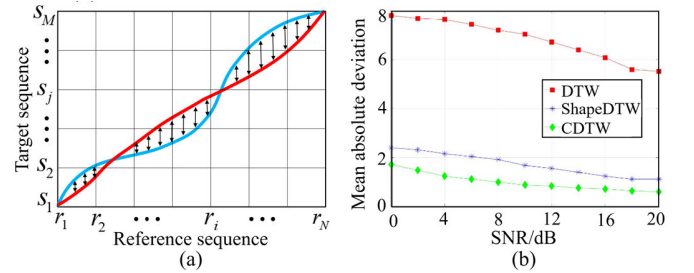

Fig. 15. Alignment quality evaluation: (a) mean absolute deviation and (b) alignment quality at different SNR.

prioritizes matching local shapes at the expense of matching accuracy in other parts. The CSDTW proposed in this study is shown in Fig. 13(d). The CSDTW fully takes into account the curvature changes information of the ridges and adopts different search step patterns when the curvature is changing, achieving excellent alignment in regions A, B, and C while maintaining the characteristics of the ridges to a large extent in the smaller curvature areas.

The alignment paths for each method are depicted in Fig. 14(a), the ideal path indicates the alignment path for identical ridges, it represents an ideal alignment path. That illustrates CSDTW achieves alignment closest to the ideal, signifying its strong alignment capability. The matching relationships obtained by each method are displayed in Fig. 14(b)–(d), DTW shows one-to-many matching when the curvature difference is large, which leads to alignment over fitting. Shape DTW addresses the over fitting problem of DTW and improves the alignment capability. However, when dealing with complex and variable shapes of ridges, Shape DTW still prioritizes the matching of local features, which often leads to the destruction of other local features. As shown in Fig. 14(d), CSDTW further improves the alignment capability of complex shapes based on Shape DTW. When the curvature changes drastically, excellent alignment results are obtained and curve features are preserved to the maximum extent.

In order to quantitatively assess the alignment performance of several methods under different noise levels, the average decision deviation in [37] is used as an evaluation metric

$$\delta_m = \frac{1}{m} \sum_{i=1}^{m} |y_i - \bar{y}| \tag{15}$$

where δ_m is the mean absolute deviation, m is the number of data points. y_i denotes the point on the path and $\bar{y}$ denotes the mean value. The index can sensitively capture the cumulative effect of local path deviation by quantifying the area difference between matched paths, overcoming the shortcoming of the traditional Euclidean distance that is sensitive to global scaling.

The mean absolute deviation refers to the area between two matching paths, and the error is proportional to the area, as shown in Fig. 15(a). The lower the average absolute deviation, the closer the matching path is to the ideal matching path, indicating a better alignment performance. The alignment error is larger when the SNR is low and smaller when the SNR is high. It can be clearly seen from Fig. 15(b), that the mean absolute deviation of CSDTW at different SNRs is smaller than that of both DTW and Shape DTW. When the noise level is high, such as when the SNR is 0 dB, the CSDTW still has better alignment quality.

V. EXPERIMENTAL VERIFICATION

A. Experimental Settings

In order to simulate the joint motion of industrial robots, a robotic arm test bench was established. As illustrated in Fig. 16, the test bench is mainly composed of PMSM, RV reducer, robotics swing arm, and pedestal, where the RV reducer is composed of a planetary gear reducer and a cycloid reducer, and the structure is shown in Fig. 17(a). The experiment is set up with sun gear faults and planetary gear faults. The type of fault is shown in Fig. 17(b) and (c). The single-tooth crack fault had a root cut depth of 0.3 mm, and the multi-tooth wear involved three tooth surfaces with
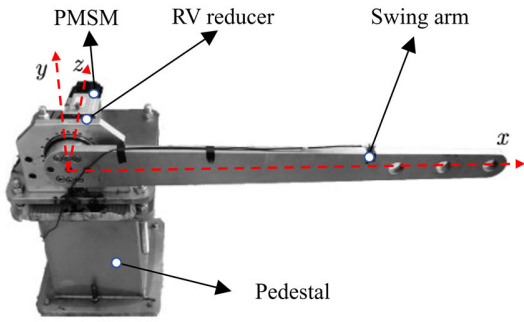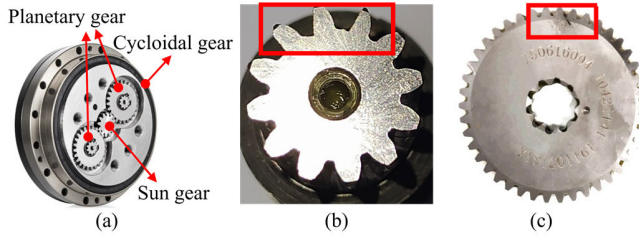
Fig. 16. Robotic arm test bench.

Fig. 17. RV reducer and gear failure. (a) RV reducer. (b) Sun gear l multi-tooth wear. (c) Planetary gear single tooth crack.

TABLE I
EXPERIMENTAL PARAMETERS OF ROBOTIC ARM TEST BENCH

Objects	Parameters	Objects	Parameters
Motor rated speed	2350 r/min	Reduction ratio	141:1
Motor rated power	1000 w	Sun gear number	12
Motor moment of inertia	2.1×10^{-4} kg·m^2	planetary gear number	42
Motor pole number	5	Cycloid disk teeth number	39
Motor stator resistance	0.8 Ω	Pin gear number	40
Inductors	3.27×10^{-7} H	Load quality	25 kg

depths of 0.5, 0.3, and 0.1 mm. The relevant parameters of the experimental bench are listed in Table I. The signals are collected using current sensors. The parameters of the acquisition system and the RV gearbox are shown in Table II.

The motion of the robotic arm from 0° to 180° is set as a cycle, and a cycle includes three stages of acceleration, uniform speed, and deceleration. The experiment was set up with three different operating conditions, denoted as operating condition 1, operating condition 2, and operating condition 3. The acceleration for all three conditions is 100°/s. The maximum speed of the operating condition 1 is 2350 r/min, acceleration time is 0.7 s, and the cycle time is 2.5 s. The maximum speed of the operating condition 2 is 1880 r/min, acceleration time is 0.5 s, and the cycle time is 2.8 s. The maximum speed of the operating condition 3 is 1410 r/min, acceleration time is 0.3 s, and the cycle time is 3.5 s. For

TABLE II
SYSTEM RELATED PARAMETERS

Parameters	Value	Parameters	Value
Number of current sensors	3	Sensor type	ZHTK25
Current sensor sensitivity	20A/3.5V	Motor type	SV-X2MH100C-B2LN
Sampling frequency	25.6 kHz	Driver type	SV-X3EB100A-A2
Capture card	NI9234	Driver software	Twin Cat3

each operating condition, 50 samples were collected, each containing a complete cycle of acceleration, uniform speed, and deceleration.

B. Time–Frequency Ridge Alignment Based on CSDTW

The signals acquired at the three conditions are shown in Fig. 18, It can be seen that the amplitude and frequency of the signal change sharply, with non-stationarity and strong time-variation. Different operating conditions result in increased signal lengths and more complex waveform variations. The template ridges are extracted from the normal signal in condition 1.

Fig. 19 illustrates the 3-D ridge extraction and normalization results. The extracted 3-D ridges are all plotted in the same coordinate system, as shown in Fig. 19(a) and (b). Fig. 19(c) and (d) shows the results of the proposed peak frequency normalization method. The normalization is used as a pre-processing method for ridge alignment, which effectively eliminates the effect of different frequency amplitudes due to variable operating conditions. Fig. 19(d) shows that using the peak of the signal spectrum to replace the ridge frequency maximum for normalization can achieve the accurate processing of ridges at variable operating conditions to the same frequency scale. The peak frequency normalization provides a solid foundation for the subsequent ridge alignment process.

Fig. 20(a)–(f) shows the alignment results of different methods. To be able to show the effect of alignment in a more concise and clear way, we have selected only 12 ridges for each operating condition to plot the alignment results. From the results, it can be seen that CSDTW achieves ridge alignment with inconsistent data lengths on the basis of normalized ridges, considers curvature information of complex local shapes of ridges for alignment, and maximally maintains the ridge structure from destruction. Fig. 20(a) and (b) show that DTW reduces the length difference brought by different operating conditions, but in the top view it can be observed that the length difference between different operating conditions is not completely eliminated. As Fig. 20(c) and (d) show Shape DTW improves the alignment performance of DTW and further expands the distance between the sun wheel fault and other faults. However, it is still difficult to distinguish between planetary wheel faults and normal faults. In Fig. 20(e) and (f),
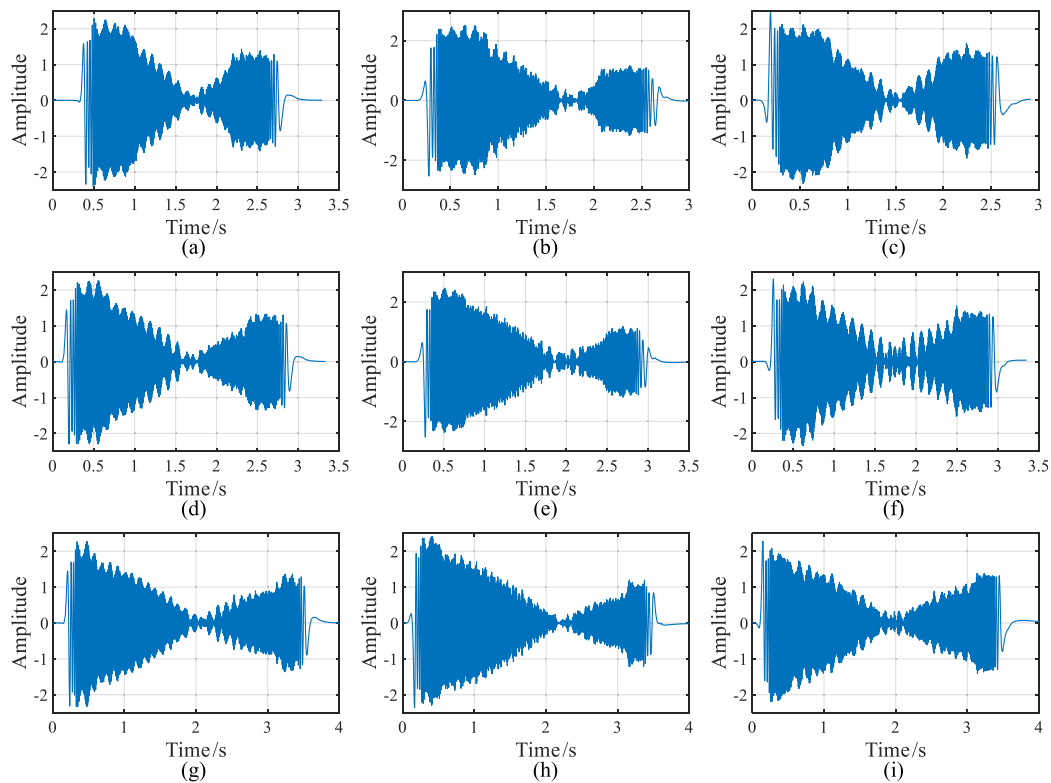
Fig. 18. Current signals under different operating conditions. (a)–(c) Operating condition 1. (d)–(f) Operating condition 2. (g)–(i) Operating condition 3. Each column indicates three gear conditions in turn: normal, sun gear failure, and planetary gear failure.
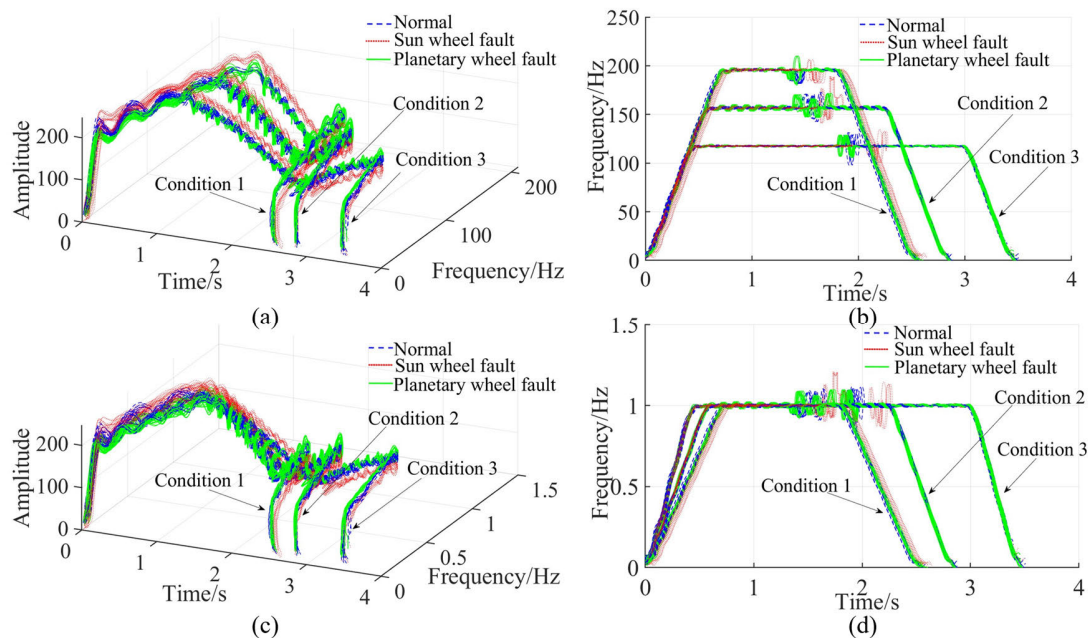

Fig. 19. Ridge normalization. (a) Original ridges. (b) Time–frequency plane of original ridges. (c) Peak frequency normalization. (d) Time–frequency plane of frequency normalization.

compared with DTW and Shape DTW, CSDTW has a better ridge aggregation effect: the distance between different faults increases, and the distance between the same faults decreases. CSDTW almost eliminates the length difference caused by

variable operating conditions and has the ability to distinguish between different faults.

To further quantify the effectiveness of different methods for ridge alignment, the mean absolute deviation was used as
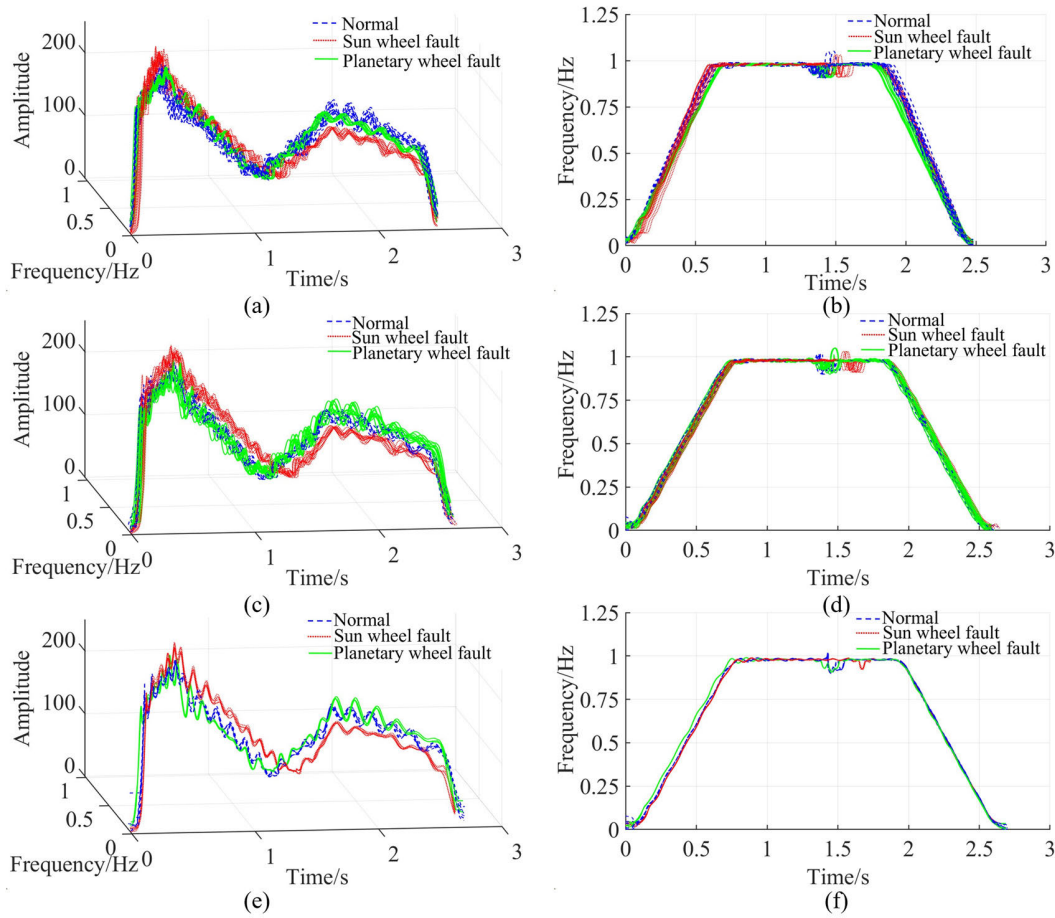
Fig. 20. Different alignment methods are used for the normalized ridges for the three operating conditions. (a) DTW. (c) Shape DTW. (e) CSDTW. (b), (d), and (f) Time–frequency plane of (a), (c), and (e).
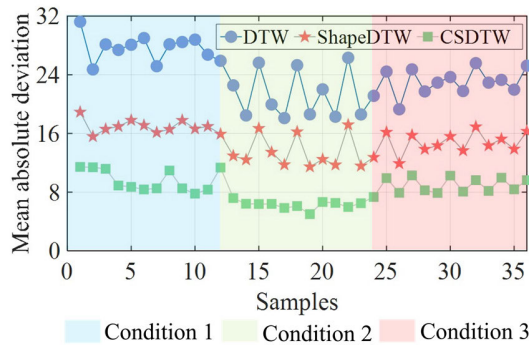

Fig. 21. Mean absolute deviation.

TABLE III
MAIN PARAMETER SETTINGS OF EXPERIMENTS

Parameters	Initial learning rate	Batch size	Epochs	Ridge point number
Values	0.0001	64	300	6000
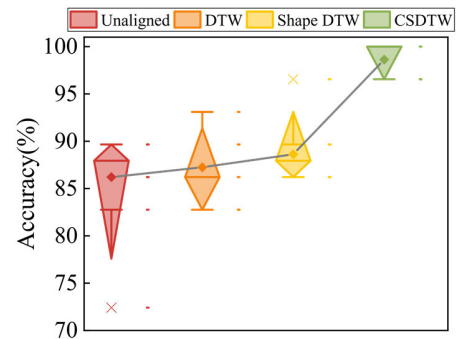
Fig. 22. Boxplot of ridge test data under different methods.

an evaluation metric. As shown in Fig. 21, the data points obtained from the mean absolute deviation of CSDTW are generally located below DTW and Shape DTW under the three working conditions, indicating that CSDTW has a smaller alignment error as well as a better alignment effect.

C. Result of Anomaly Detection

Since CNN is excellent in processing high dimensional data, this study inputs all points on the ridges as data points into the CNN for fault classification to obtain more accurate and reliable detection results. There are four states in the experiment, including normal, sun wheel fault, planet wheel fault, and compound fault. A total of 600 samples
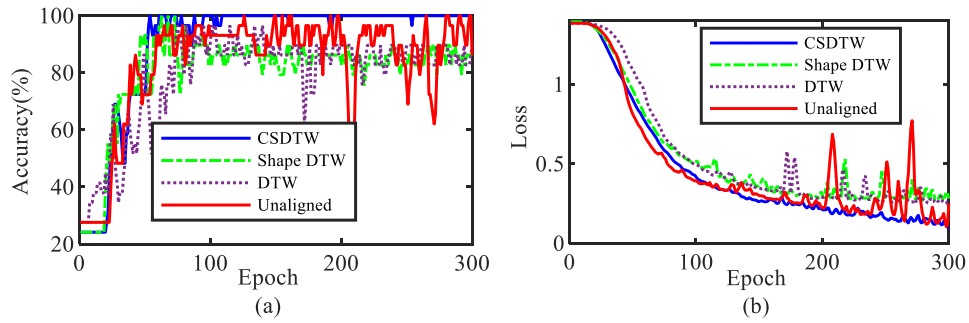
Fig. 23. (a) Accuracy curve with the number of iterations. (b) Loss variation curve.
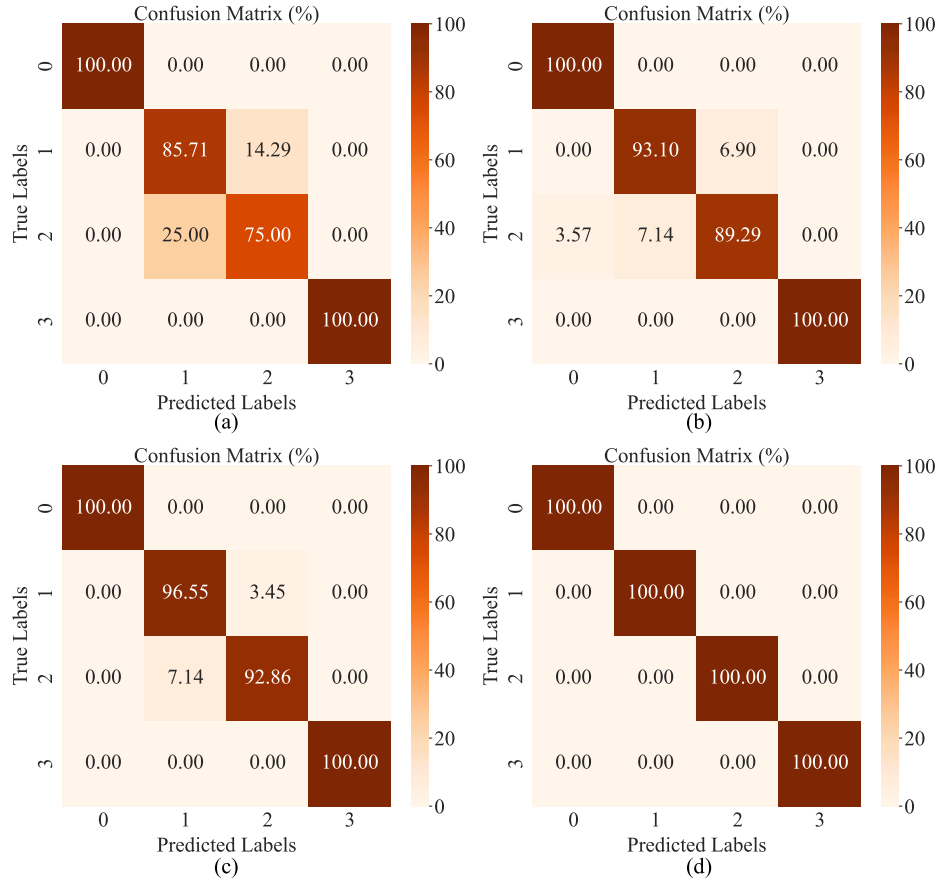

Fig. 24. Confusion matrices for different comparison methods. (a) Unaligned. (b) DTW. (c) Shape DTW. (d) CSDTW.

are included in the three conditions. Each sample contains a ridge, 80% of the data is randomly selected as the training set, and the remaining 20% is used to construct the test set. The relevant parameter configurations are detailed in Table III.

Gearbox anomaly detection is performed by aligned ridges of CSDTW and compared with the results of unaligned ridges, DTW-based aligned ridges, and shape DTW-based aligned ridges detection. The data from each method are analyzed by performing ten repetitions of the experiment and calculating the mean value to guarantee the accuracy and confidence of the results. The statistical results are shown in Table IV. The results show that the average detection accuracy of the unaligned ridges is the lowest due to the influence of the

variable working conditions, but the test accuracy under DTW is even lower than that of the unaligned data in tests 2 and 8, which is due to the fact that the overfitting of the DTW has seriously damaged the ridge shape characteristics, resulting in the fluctuation of its accuracy becoming larger. In contrast, the detection accuracy of Shape DTW is improved. However, Shape DTW only focuses on local alignment, and the feature destruction it causes also affects the classification accuracy to some extent. Therefore, the improvement in detection accuracy of both alignment methods is not obvious. Overall, the proposed CSDTW has the highest detection accuracy. Compared with the above methods, the detection accuracy is improved by more than 15%, which demonstrates the superiority and effectiveness of the CSDTW method.

TABLE IV
STATISTICAL RESULTS OF DIFFERENT METHODS (%)

Methods	Test1	Test2	Test3	Test4	Test5
Unaligned	89.66	89.66	89.66	89.66	82.76
DTW	93.10	86.21	82.76	86.21	89.66
Shape DTW	89.66	86.21	89.66	86.21	86.21
CSDTW	**96.55**	**96.55**	**96.55**	**100.00**	**100.00**

Methods	Test6	Test7	Test8	Test9	Test10
Unaligned	72.41	86.21	89.66	86.21	86.21
DTW	86.21	89.66	82.76	86.21	89.66
Shape DTW	86.21	**96.55**	86.21	89.66	89.66
CSDTW	**100.00**	**96.55**	**100.00**	**100.00**	**100.00**

In order to demonstrate the characteristics of the data distribution of the different methods, including its center location, degree of dispersion, and outliers, we plotted a box plot as shown in Fig. 22. The diamond shape represents the box, whose top and bottom boundaries correspond to the 90% and 10% quantiles of the data, respectively. The lines extending from the top and bottom of the box point to the maximum and minimum values of the data, respectively, and data points that exceed these boundaries are labeled as outliers. The horizontal line in the graph pinpoints the location of the median of the data, while the solid squares indicate the mean of the data. The scatterplot to the right of the box-and-line plot demonstrates the actual distribution of the data. It is clear from the figure that among the four comparative methods, the method of this study is significantly ahead in terms of accuracy and has a higher concentration of data.

In order to further demonstrate the superiority of the proposed CSDTW over the rest of the methods, Fig. 23 shows the diagnostic accuracy and loss variation curves of all the compared methods. It can be seen that the training process of the unaligned method is very unstable, with many large oscillations during the training process, making it difficult to converge to the optimal solution. In contrast, DTW and Shape DTW, which adopt the alignment strategy, are more effective but still suffer from training instability. Overall, CSDTW has the most stable training process and the best detection performance. As the number of training epochs increases, the detection accuracy reaches 100.00%. It shows that the proposed method gets a better alignment effect by introducing ridge curvature information. The effect of different operating conditions is eliminated, which improves the detection accuracy.

In order to understand the detection performance of each fault category, confusion matrices are given as shown in Fig. 24 for four different methods, with 0 representing normal, 1 representing composite faults, 2 representing sun wheel faults, and 3 representing planetary wheel faults. The results show that the unaligned method leads to poor classification of composite faults and solar wheel faults due to the impact of different operating conditions. The classification accuracy of DTW is improved compared to the unaligned method, but there are still cases of confusing composite faults with solar wheel faults. Compared with the previous two methods, Shape DTW has better detection results, but there are cases

of misclassifying a small number of composite faults as solar faults. Overall, the detection performance of CSDTW in each fault category is typical, with 100.00% detection accuracy and superior detection performance.

VI. CONCLUSION

To realize more accurate and reliable anomaly detection for industrial robots under variable working conditions, this study proposes a method based on 3-D time–frequency ridge alignment. The main conclusions are as follows.

1) A new ridge alignment method is proposed, and the alignment effect is better than DTW and Shape DTW on both simulation and real experiment. The advantages of the method are demonstrated through simulations and experiments.
2) Peak frequency normalization accurately unifies the amplitude of ridges to the same scale and avoids the effects of ridge extraction errors to some extent. CSDTW eliminates the effect of inconsistent ridge lengths due to variable operating conditions, and fully considers the ridge shape characteristics for alignment.
3) The alignment algorithm based on CSDTW well eliminates the influence of the variation of operating conditions and retains the accurate fault features, which improves the accuracy of the CNN fault detection results.

Further, future research could consider improving the algorithm to reduce the time complexity and dependence on parameters, as well as incorporating methods such as high-precision TFA to further optimize the results of ridge extraction. Finally, more innovative and application scenarios of time–frequency ridges in industrial robot anomaly detection are explored.

REFERENCES

[1] B. Xia, K. Wang, A. Xu, P. Zeng, N. Yang, and B. Li, "Intelligent fault diagnosis for bearings of industrial robot joints under varying working conditions based on deep adversarial domain adaptation," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022.

[2] G. Zhang, Y. Tao, J. Wang, K. Feng, and X. Han, "A motor current signal-based fault diagnosis method for harmonic drive of industrial robot under time-varying speed conditions," *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 1–10, 2025.

[3] H. Wan et al., "Time–frequency RWGAN for machine anomaly detection under varying working conditions," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–11, 2024.

[4] J. Tang, G. Zheng, C. Wei, W. Huang, and X. Ding, "Signal-transformer: A robust and interpretable method for rotating machinery intelligent fault diagnosis under variable operating conditions," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[5] K. Xu, X. Wu, D. Wang, and X. Liu, "Adaptive high-order multi-synchrosqueezing transform for monitoring of industrial robot under reciprocating start-stop condition," *Measurement*, vol. 247, Apr. 2025, Art. no. 116866.

[6] G. Song, S. H. Hong, T. Kyzer, and Y. Wang, "Energy consumption auditing based on a generative adversarial network for anomaly detection of robotic manipulators," *Future Gener. Comput. Syst.*, vol. 149, pp. 376–389, Dec. 2023.

[7] M. Castellano-Quero, M. Castillo-López, J.-A. Fernández-Madrigal, V. Arévalo-Espejo, H. Voos, and A. García-Cerezo, "A multidimensional Bayesian architecture for real-time anomaly detection and recovery in mobile robot sensory systems," *Eng. Appl. Artif. Intell.*, vol. 125, Oct. 2023, Art. no. 106673.

[8] Y. Xu et al., "Cross-modal fusion convolutional neural networks with online soft-label training strategy for mechanical fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 20, no. 1, pp. 73–84, Jan. 2024, doi: 10.1109/TII.2023.3256400.

[9] K. Choi, J. Yi, C. Park, and S. Yoon, "Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines," *IEEE Access*, vol. 9, pp. 120043–120065, 2021, doi: 10.1109/ACCESS.2021.3107975.

[10] M. Rao, X. Yang, Z. Cao, M. J. Zuo, and F. Chu, "Normalizing vibration signals with a novel piecewise power fitting method for intelligent fault detection of rotating machinery," *Mech. Syst. Signal Process.*, vol. 221, Dec. 2024, Art. no. 111751, doi: 10.1016/J.YMSSP.2024.111751.

[11] S. Li, J. Ji, K. Feng, K. Zhang, Q. Ni, and Y. Xu, "Composite neuro-fuzzy system-guided cross-modal zero-sample diagnostic framework using multisource heterogeneous noncontact sensing data," *IEEE Trans. Fuzzy Syst.*, vol. 33, no. 1, pp. 302–313, Jan. 2025, doi: 10.1109/TFUZZ.2024.3470960.

[12] Z. Lei et al., "Prior knowledge-embedded meta-transfer learning for few-shot fault diagnosis under variable operating conditions," *Mech. Syst. Signal Process.*, vol. 200, Oct. 2023, Art. no. 110491, doi: 10.1016/j.ymssp.2023.110491.

[13] K. Feng, J. C. Ji, Y. Zhang, Q. Ni, Z. Liu, and M. Beer, "Digital twin-driven intelligent assessment of gear surface degradation," *Mech. Syst. Signal Process.*, vol. 186, Mar. 2023, Art. no. 109896, doi: 10.1016/J.YMSSP.2022.109896.

[14] X. Yu et al., "Intelligent fault diagnosis of rotating machinery under variable working conditions based on deep transfer learning with fusion of local and global time–frequency features," *Struct. Health Monitor.*, vol. 23, no. 4, pp. 2238–2254, Jul. 2024, doi: 10.1177/14759217231199427.

[15] Y. Zhang, J. Ding, Y. Li, Z. Ren, and K. Feng, "Multimodal data cross-domain fusion network for gearbox fault diagnosis under variable operating conditions," *Eng. Appl. Artif. Intell.*, vol. 133, Jul. 2024, Art. no. 108236, doi: 10.1016/J.ENGAPPAI.2024.108236.

[16] J. P. Salameh, S. Cauet, E. Etien, A. Sakout, and L. Rambault, "Gearbox condition monitoring in wind turbines: A review," *Mech. Syst. Signal Process.*, vol. 111, pp. 251–264, Oct. 2018.

[17] A. Prudhom, J. Antonino-Daviu, H. Razik, and V. Climente-Alarcon, "Time–frequency vibration analysis for the detection of motor damages caused by bearing currents," *Mech. Syst. Signal Process.*, vol. 84, pp. 747–762, Feb. 2017.

[18] T. Wang, Q. Han, F. Chu, and Z. Feng, "Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review," *Mech. Syst. Signal Process.*, vol. 126, pp. 662–685, Jul. 2019.

[19] M. D. Choudhury, L. Hong, and J. S. Dhupia, "A novel tacholess order analysis method for bearings operating under time-varying speed conditions," *Measurement*, vol. 186, Dec. 2021, Art. no. 110127.

[20] X. Chen and Z. Feng, "Time–frequency space vector modulus analysis of motor current for planetary gearbox fault diagnosis under variable speed conditions," *Mech. Syst. Signal Process.*, vol. 121, pp. 636–654, Apr. 2019.

[21] S. Chen, M. Du, Z. Peng, M. Liang, Q. He, and W. Zhang, "High-accuracy fault feature extraction for rolling bearings under time-varying speed conditions using an iterative envelope-tracking filter," *J. Sound Vib.*, vol. 448, pp. 211–229, May 2019.

[22] H. Xu, S. Zhou, and T. Yang, "Stochastic resonance of a high-order-degradation bistable system and its application in fault diagnosis with variable speed condition," *Mech. Syst. Signal Process.*, vol. 186, Mar. 2023, Art. no. 109852.

[23] Y. Li, X. Zhang, Z. Chen, Y. Yang, C. Geng, and M. J. Zuo, "Time–frequency ridge estimation: An effective tool for gear and bearing fault diagnosis at time-varying speeds," *Mech. Syst. Signal Process.*, vol. 189, Apr. 2023, Art. no. 110108.

[24] G. He, K. Ding, W. Li, and X. Jiao, "A novel order tracking method for wind turbine planetary gearbox vibration analysis based on discrete spectrum correction technique," *Renew. Energy*, vol. 87, pp. 364–375, Mar. 2016.

[25] S. Jiang and Z. Chen, "Application of dynamic time warping optimization algorithm in speech recognition of machine translation," *Heliyon*, vol. 9, no. 11, Nov. 2023, Art. no. e21625.

[26] J. Tang, H. Cheng, Y. Zhao, and H. Guo, "Structured dynamic time warping for continuous hand trajectory gesture recognition," *Pattern Recognit.*, vol. 80, pp. 21–31, Aug. 2018.

[27] Y. Liu, Y.-A. Zhang, M. Zeng, and J. Zhao, "A novel distance measure based on dynamic time warping to improve time series classification," *Inf. Sci.*, vol. 656, Jan. 2024, Art. no. 119921.

[28] Z. Wei, Y. Gao, X. Zhang, X. Li, and Z. Han, "Adaptive marine traffic behaviour pattern recognition based on multidimensional dynamic time warping and DBSCAN algorithm," *Expert Syst. Appl.*, vol. 238, Mar. 2024, Art. no. 122229.

[29] M. Potrykus, "Dot-com and AI bubbles: Can data from the past be helpful to match the price bubble euphoria phase using dynamic time warping?" *Finance Res. Lett.*, vol. 67, Sep. 2024, Art. no. 105799.

[30] J.-L. Li, W.-L. Huang, and R.-X. Zhang, "Seismic data extrapolation based on multi-scale dynamic time warping," *Petroleum Sci.*, vol. 21, no. 6, pp. 3981–4000, Dec. 2024.

[31] L. Qiu, C. Qiu, and C. Song, "ESDTW: extrema-based shape dynamic time warping," *Expert Syst. Appl.*, vol. 239, Apr. 2024, Art. no. 122432.

[32] J. Guo, Z. Si, Y. Liu, J. Li, Y. Li, and J. Xiang, "Dynamic time warping using graph similarity guided symplectic geometry mode decomposition to detect bearing faults," *Reliab. Eng. Syst. Saf.*, vol. 224, Aug. 2022, Art. no. 108533.

[33] T. Han, X. Liu, and A. C. C. Tan, "Fault diagnosis of rolling element bearings based on multiscale dynamic time warping," *Measurement*, vol. 95, pp. 355–366, Jan. 2017.

[34] B. Sun, H. Li, C. Wang, K. Zhang, and S. Chen, "Current-aided dynamic time warping for planetary gearbox fault detection at time-varying speeds," *IEEE Sensors J.*, vol. 24, no. 1, pp. 390–402, Jan. 2024.

[35] V. Sharma and A. Parey, "Frequency domain averaging based experimental evaluation of gear fault without tachometer for fluctuating speed conditions," *Mech. Syst. Signal Process.*, vol. 85, pp. 278–295, Feb. 2017.

[36] L. Hong, Y. Qu, J. S. Dhupia, S. Sheng, Y. Tan, and Z. Zhou, "A novel vibration-based fault diagnostic algorithm for gearboxes under speed fluctuations without rotational speed measurement," *Mech. Syst. Signal Process.*, vol. 94, pp. 14–32, Sep. 2017.

[37] J. Zhao and L. Itti, "ShapeDTW: Shape dynamic time warping," *Pattern Recognit.*, vol. 74, pp. 171–184, Feb. 2018.

[38] M. Kloska, G. Grmanova, and V. Rozinajova, "Expert enhanced dynamic time warping based anomaly detection," *Expert Syst. Appl.*, vol. 225, Sep. 2023, Art. no. 120030.

[39] S. Zhao, Y. Duan, N. Roy, and B. Zhang, "A deep learning methodology based on adaptive multiscale CNN and enhanced highway LSTM for industrial process fault diagnosis," *Rel. Eng. Syst. Saf.*, vol. 249, Sep. 2024, Art. no. 110208.

[40] S. Tang, Y. Zhu, and S. Yuan, "Intelligent fault identification of hydraulic pump using deep adaptive normalized CNN and synchrosqueezed wavelet transform," *Rel. Eng. Syst. Saf.*, vol. 224, Aug. 2022, Art. no. 108560.

[41] G. Yu, Z. Wang, P. Zhao, and Z. Li, "Local maximum synchrosqueezing transform: An energy-concentrated time–frequency analysis tool," *Mech. Syst. Signal Process.*, vol. 117, pp. 537–552, Feb. 2019.

Jiabing Gu was born in Yunnan, China, in 2000. He is currently pursuing the M.D. degree in mechanical engineering with the School of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, China.

His research interests include anomaly detection and fault diagnosis of industrial robots.

Xing Wu received the Ph.D. degree in engineering majoring in mechanical design and theory from the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, China, in 2005.

He is currently the President of West Yunnan University of Applied Sciences, Dali, China. His current research interests include acoustic fault diagnosis theory and method, intelligent diagnosis method, and system application.

Dongxiao Wang received the Ph.D. degree in engineering from Kunming University of Science and Technology, Kunming, China, in 2023.

He is currently a Lecturer with Kunming University of Science and Technology. His research field is fault diagnosis for robots.

Xiaoqin Liu (Member, IEEE) received the Ph.D. degree in engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in December 2009.

His main research directions include new methods of condition monitoring and intelligent diagnosis, sound field measurement and source identification, dynamic testing, and dynamic analysis.

Kai Xu received the Ph.D. degree in mechanical design and theory from Kunming University of Science and Technology, Kunming, China, in 2024.

He is currently a Lecturer with Kunming University of Science and Technology. His research interests include electromechanical condition monitoring and fault diagnosis, and signal processing.

Ran Ren was born in Yunnan, China, in 2001. She is currently pursuing the M.D. degree in mechanical engineering with the School of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, China.

Her research field is fault diagnosis of rotating machinery based on transfer learning.