RESEARCH ARTICLE

A New Method for Fault Diagnosis of Hydraulic System Based on Improved Empirical Wavelet Transform and Kernel Extreme Learning Machine

CHAO CHEN, MUHETAER KELIMU[ID], AND BO YANG
College of Mechanical Engineering, Xinjiang University, Ürümqi, Xinjiang 830017, China

Corresponding author: Muhetaer Kelimu (kmuhetar@xju.edu.cn)

ABSTRACT To address the problem that multivariable faults in hydraulic systems are difficult to diagnose accurately in a short time, a new signal processing method combining an improved empirical wavelet transform with a kernel extreme learning machine optimized by the Pelican optimization algorithm is proposed. First, the problem of unreasonable spectral partitioning of the empirical wavelet transform is solved using an improved k-means clustering method, which upgrades the empirical wavelet transform in the maximum scale space and then adaptively decomposes the acquired pressure signal to obtain a series of sub-signal components, thus reducing the dimensionality of multivariate faults and saving computation time. Second, 17 features of each sub-signal component are calculated and input to the kernel extreme learning machine for training, and the sequence-forward selection strategy is used to select the optimal features from the kernel extreme learning machine to ensure the basic accuracy and efficiency of prediction. Finally, a Pelican optimization algorithm improved kernel extreme learning machine algorithm is proposed for fast classification of faults. Experiments show that the method can diagnose a variety of faults in hydraulic systems quickly and accurately, with a diagnostic accuracy of 97.24 percent, which is better than other methods.

INDEX TERMS Improved empirical wavelet transform, hydraulic systems, multivariable fault diagnosis, pelican optimization algorithm, kernel extreme learning machine.

I. INTRODUCTION

Hydraulic transmission systems are widely used in various fields because of their ability to produce large forces or torques and their ease of control. Some scholars have conducted statistical analyzes on the location of failures of machinery and equipment containing hydraulic systems and found that most failures of machinery and equipment are caused by hydraulic systems [1], [2], [3]. The operation of the hydraulic system is so frequent that it is easy to break down, but if the fault location is not found in time and corresponding measures are not taken, it will cause serious economic losses at least and cause personal safety problems

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino[ID].

at worst [4]. Therefore, it is necessary to perform fault diagnosis for hydraulic systems so that the fault location can be quickly located and the type of fault can be determined to avoid unforeseen consequences.

In general, fault diagnosis is divided into analytical model-based methods and data-driven methods [5]. However, many components in hydraulic systems are complex rotating machines, and it is difficult to build fault models. Thanks to the improvement of industrial sensor accuracy and the speed of computer computation, data-driven hydraulic system fault diagnosis has become mainstream [6].

Popular data-driven fault diagnosis includes fault diagnosis methods based on artificial intelligence [7], [8], [9], [10], [11], [12] and signal processing [13], [14]. Both methods can also be combined for fault diagnosis [15], [16].

Artificial intelligence fault diagnosis methods are preferred by researchers because of their high computational accuracy and the fact that they rarely require expertise in fault objects. However, artificial intelligence diagnosis has a complex structure and is inefficient when diagnosing multiple faults. To solve this problem, different signal preprocessing methods combined with convolutional neural networks were utilized in the literature [17] and [18], respectively, for fault diagnosis of rotating machinery and to enhance the efficiency of fault diagnosis. Although this approach may not be directly applicable to non-rotating machinery, it demonstrates that signal preprocessing can improve the efficiency of fault diagnosis. Therefore, it is especially important to choose the appropriate signal processing algorithm.

In 1807, the Fourier transform was introduced, which greatly accelerated the development of the field of digital signal processing. The wavelet transform is a multi-resolution signal analysis method based on the Fourier transform, which can achieve the decomposition and reconstruction of a discrete-time signal through the fusion of approximate and detailed components [19]. Unfortunately, the choice of the wavelet transform basis function is determined by subjective experience, which is very unfriendly to inexperienced stakeholders. In addition, the wavelet transform produces great frequency leakage when dealing with complex signals, which makes it impossible to use the wavelet transform for hydraulic system fault diagnosis [20].

Empirical Modal Decomposition (EMD) is another classical signal processing method. EMD can adaptively decompose the vibration signal into many kinds of Intrinsic Mode Functions (IMFs), which solves the problem of empirically selecting appropriate wavelet bases for the wavelet transform. After the IMFs are merged, the noise interference existing in the vibration signal can be eliminated [21]. Yang et al. [22] achieved their purpose of reducing the noise within the signal through multipoint data fusion. EMD can decompose the signal according to the temporal characteristics of the processed object, so there is no need to set the basis function, and there is no problem similar to the wavelet transform. However, modal blending and breakpoint effects are very common in EMD, and it has no theoretical basis to support them.

The Empirical Wavelet Transform (EWT), which is similar in composition to the wavelet transform, was mentioned by Gilles in 2013 and is an emerging method for time-frequency analysis with the advantage that EMD can automatically adapt to the signal and has a rigorous theoretical basis [23]. EWT is widely used in the field of signal processing [24], [25], [26]. For the process of EWT analysis of vibration signals, the selection of meaningful boundaries is the key to the success of signal decomposition [27]. However, most studies on EWT use local maxima or local minima in the selection of boundaries, which has a large error in dealing with complex nonlinear nonsmooth signals because, with noisy signals, the amplitude of many noises has reached the same height as normal signals, so some of the boundaries determined by the minima are actually due to noise [28].

To overcome these problems of conventional EWT, Pan et al. [29] used the Gaussian kernel inner product to remove the noise from the target signal and thus decompose the normal signal in the target signal. Kong et al. [30] decomposed the fault signal under strong noise using the phenomenon of meshing resonance. Kim et al. [31] used the inverse-spectrum-assisted EWT to make the spectrum of the target signal more smooth, together with the Hilbert transform to average the envelope spectrum to obtain the results of gear faults. The order spectrum coherence is combined with historical data obtained from healthy machines to obtain the anomaly envelope spectrum, which is further processed for fault diagnosis. This anomaly envelope spectrum is further processed by smoothing operations to perform not only automatic fault detection but also to identify damaged components [32]. Kedadouche et al. [33] used operational modal analysis to determine the stable frequency of the signal, calculate the support boundary, apply the scaling function and wavelet function corresponding to each detection segment, and filter the signal with the constructed filter bank to obtain the IMF. The approach is then roughly the same as for EMD. Like Gilles to calculate the scale space [27], Zhao et al. [34] used several maximum lengths of the scale space as the boundaries for dividing the spectrum, obtained a series of IMF components after reconstructing the signal, and performed power spectrum analysis on these components to derive the fault frequency. Zhang et al. [35] used the Power Spectral Density (PSD) instead of the Fourier spectrum, and since the extreme value of the PSD points is smaller, there is no need to calculate the complex scale space; instead, the minimal value of PSD is used as the boundary to divide the spectrum to obtain the components of the original signal. Zhang et al. [36] proposed a variable spectrum splitting EWT to estimate the modes using multi-taper power spectral density, and finally a set of boundaries associated with the spectrum fluctuations to obtain the components of the original signal. Ding [37] obtained the upper and lower boundaries of the sidebands by searching left and right bidirectionally at the center of the spectrum and merged the upper and lower boundaries of the sidebands as the boundaries of the spectrum segmentation to achieve the fault detection of bearings. Zheng et al. [38] replaced the Fourier amplitude spectrum with the power spectrum to verify the effectiveness of the method by decomposing the fault signal of the hydraulic pump slide valve loosening. However, these studies seem to avoid the spectrum division of EWT, and do not fundamentally solve the problem of EWT, i.e., the unreasonable spectrum division. To solve this problem, Yu et al. [39] used the DBSCAN clustering method to cluster the largest scale space to obtain the boundary of the spectrum, and the decomposition of the vibration signals of the hydraulic pump in three directions proved the effectiveness of the method in the detection of weak faults in the pump. The DBSCAN method does have a more significant improvement for EWT of vibration signals. However, we found through experiments that this method seems to fail for components

in hydraulic systems where it is difficult to detect vibration signals.

In a hydraulic circuit, the pressure signal and the flow signal are easily detected for any component through which hydraulic fluid passes. The pressure and flow signals of hydraulic components contain a great deal of information, and the fault information is only the tip of the iceberg. The feature extractor can reflect the features of the fault information from many features [40], [41]. Li et al. [42] first conducted an empirical determination of the number of signal EWT decompositions. Subsequently, they selected the component with the highest energy intensity and utilized the inverse dispersion entropy features of this component as the features of the original signal for efficient feature extraction purposes. Lu et al. [43] took the first n-order components with high correlation coefficients between the decomposed components and the original signal to construct fused feature vectors, which were used to build a two-level diagnostic model based on the salp swarm algorithm optimized Kernel Extreme Learning Machine (KELM) to identify normal and abnormal states and the categories under the abnormal states. Ding et al. [44] used EWT to decompose the hydraulic pump vibration signal, used principal component analysis to reduce the dimensionality of the feature extraction of the decomposed signal, and finally input the feature vector containing the fault features into the Extreme Learning Machine (ELM) to obtain the fault classification results. Liu et al. [45] developed a novel personalized diagnosis method for gear fault detection using numerical simulation and ELM algorithm, which can diagnose the health condition of gears by extracting features and training ELM model based on large amount of vibration signals.

Many studies have been conducted at home and abroad for the fault diagnosis of vibration signals of hydraulic components, and vibration analysis is the most effective condition-monitoring technique for rotating systems [46]. It is very effective for fault detection of bearings, shafts, hydraulic pumps, hydraulic motors, etc. [47], [48]. However, many hydraulic components are non-rotating components and we cannot or have difficulty in measuring their vibration signals. In contrast, the pressure signals of the hydraulic system are easier to measure. According to our experiments, it is verified that the pressure signals can be applied to the fault diagnosis of the hydraulic system.

This paper proposes to combine the improved EWT and the Pelican Optimization Algorithm (POA) optimized KELM to preprocess the pressure signals to diagnose multivariable faults in hydraulic systems. The implementation of the proposed method in this study is shown in Fig. 1. The article makes significant contributions in the following ways:

1) By using the light-k-means clustering algorithm, the problem of unreasonable spectral division of pressure signals when performing EWT is solved.

2) Seventeen parameters are proposed to construct the feature pool, and a Sequence Forward Selection (SFS) strategy is used to select the features with the optimal ability to distinguish faults.

3) Optimizing the kernel parameters and regularization coefficients of KELM using POA has better fault classification results.

In the second part, the pertinent theory is introduced. The third part delineates the experimental setting and conducts an experimental analysis to validate and evaluate the proposed approach. The fourth part lays out the conclusions and outlines the future work.

II. METHODOLOGY

The following sections will introduce in detail the basic principles of EWT, the realization of the largest scale space, the steps to realize light-k-means, and the process of the KELM improved by the POA.

A. THE FUNDAMENTALS OF EWT

Inspired by the wavelet transform, Gills proposed the EWT in 2013, which has a similar definition to the transform.

From the perspective of Fourier, this method can adaptively select a band-pass filter bank according to the spectrum of the signal to be processed, divide the signal spectrum into different frequency bands, and extract a series of different Sub-Signal Components (SSCs).

We define ω_n as the middle point of the segment with width $2\tau_n$ and index n (where $1 \leq n \leq N$). According to the Shannon criterion, the values of ω_0 and ω_n are 0 and π, respectively. The partition of the EWT boundary is shown in Fig. 2. Divide a signal into N continuous frequency bands, marked as bands $\Lambda_n = [\omega_{n-1}, \omega_n]$, which limits the signal to $U_{n=1}^N \Lambda_n = [0, \pi]$ in the Fourier spectrum.

Inspired by the idea of Littlewood-Paley and Meyer wavelet construction, Gills used (1) and (2) formulas to construct empirical scaling functions $\varphi_n(\omega)$ and empirical wavelet functions $\psi_n(\omega)$ when the frequencies are at different positions [23].

$$\varphi_n(\omega) = \begin{cases} 1 \ if \ |\omega| \leq (1-\gamma)\omega_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega|-(1-\gamma)\omega_n)\right)\right] \\ \quad if \ (1-\gamma)\omega_n \leq |\omega| \leq (1+\gamma)\omega_n \\ 0 \ otherwise \end{cases} \quad (1)$$

$$\psi_n(\omega) = \begin{cases} 1 \ if \ (1+\gamma)\omega_n \leq |\omega| \leq (1-\gamma)\omega_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}(|\omega|-(1-\gamma)\omega_{n+1})\right)\right] \\ \quad if \ (1-\gamma)\omega_{n+1} \leq |\omega| \leq (1+\gamma)\omega_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega|-(1-\gamma)\omega_n)\right)\right] \\ \quad if \ (1-\gamma)\omega_{n+1} \leq |\omega| \leq (1+\gamma)\omega_n \\ 0 \ otherwise \end{cases}$$
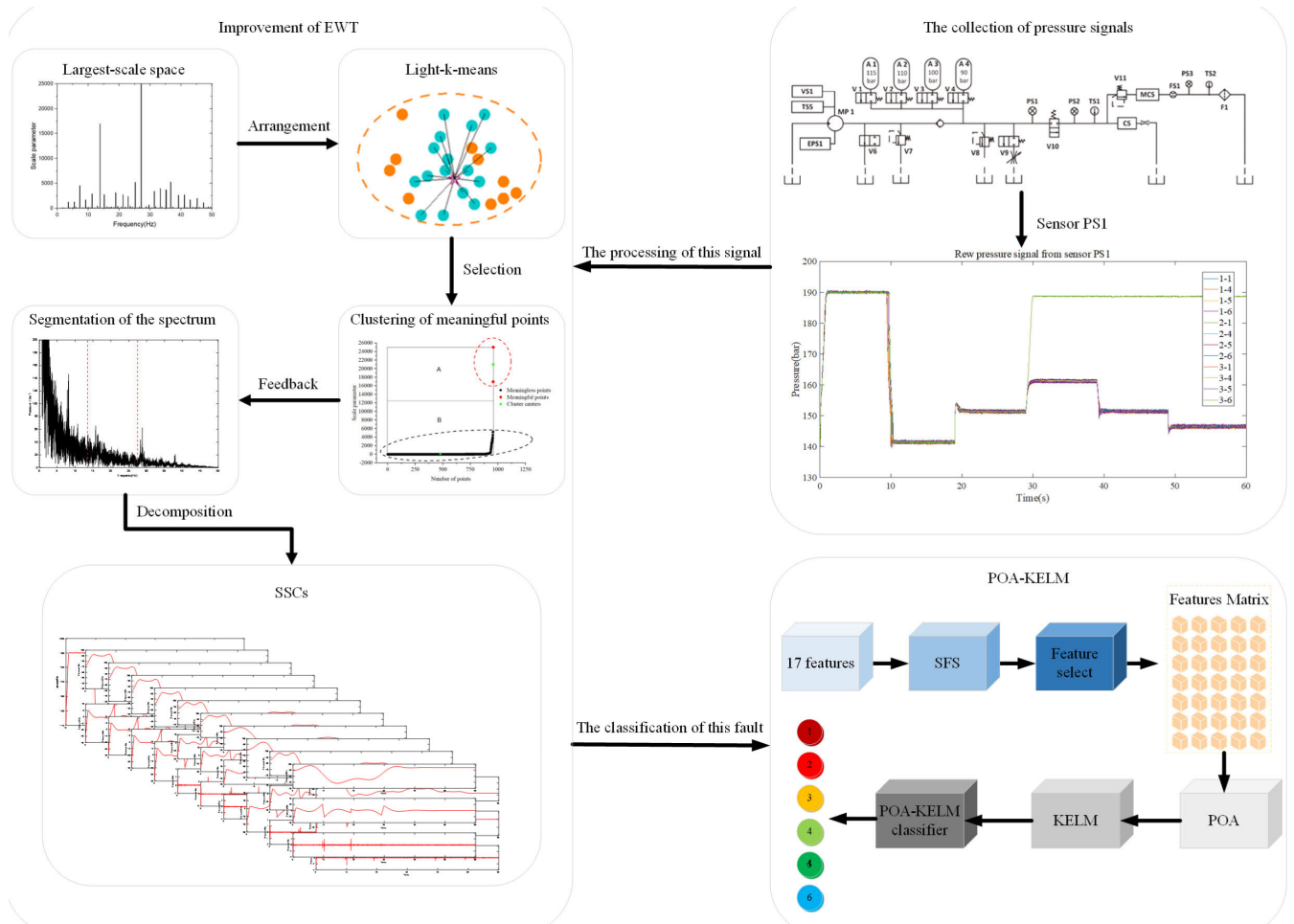
$$\quad (2)$$

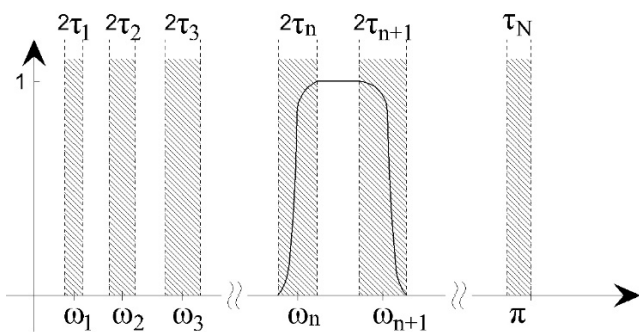**FIGURE 1.** The implementation flow of the proposed method.

FIGURE 2. Band segmentation of EWT.

where $\gamma = \tau_n/\omega_n$, and $\beta(x)$ is a function about $\forall C^k \in [0, 1]$ that is calculated by (3).

$$\beta(x) = \begin{cases} 0 \ if \ x \leq 0 \\ and \ \beta(x) + \beta(1-x) = 1 \ \forall \ x \in [0, 1] \\ 1 \ if \ x > 1 \end{cases} \quad (3)$$

Many functions satisfy this equation. According to document [49], (4) is the most used one.

$$\beta(x) = 35x^4 - 84x^5 + 70x^6 - 20x^7 \quad (4)$$

Let $f(\omega)$ and F^{-1} represent the Fourier transform and the inverse transform operator of the signal $f(t)$ to be decomposed, and then the detail coefficients $W_f^e(n, t)$ of the EWT are obtained by taking the inner product of the signal $f(t)$ and the empirical wavelets $\psi_n(t)$:

$$W_f^e(n, t) = \langle f(t), \psi_n(t) \rangle$$
$$= \int f(\tau)\overline{\psi_n(\tau - t)}d\tau = F^{-1}[f(\omega)\overline{\psi_n(\omega)}] \quad (5)$$

And the approximation coefficients $W_f^e(0, t)$ are calculated by the inner product of the signal $f(t)$ and the scaling function $\varphi_1(t)$:

$$W_f^e(0, t) = \langle f(t), \varphi_1(t) \rangle$$
$$= \int f(\tau)\overline{\varphi_1(\tau - t)}d\tau = F^{-1}[f(\omega)\overline{\varphi_1(\omega)}] \quad (6)$$

where $\overline{\psi_n(\tau)}$ and $\overline{\varphi_1(\tau)}$ represent the complex conjugate functions of $\psi_n(\tau)$ and $\varphi_1(\tau)$, respectively, while $\varphi_1(\omega)$ and $\psi_n(\omega)$ are given by (1) and (2).

The signal $f(t)$ is broken down into $N + 1$ SSCs $f_0(t), f_1(t), \ldots, f_N(t)$, and these SSCs are mathematically

represented by (7).

$$
\begin{cases}
f_0(t) = W_f^e(0, t) * \varphi_1(t) \\
f_k(t) = W_f^e(k, t) * \psi_k(t)
\end{cases}
\tag{7}
$$

where k is an index variable such that $1 \leq k \leq N$, and the reconstructed signal is calculated by (8).

$$
f(t) = W_f^e(0, t) * \varphi_1(t) + \sum_{k=1}^{N} W_f^e(k, t) * \psi_k(t)
$$
$$
= F^{-1}[W_f^e(0, \omega)\varphi_1(\omega) + \sum_{k=1}^{N} W_f^e(k, \omega)\psi_k(\omega)]
\tag{8}
$$

where $W_f^e(0, \omega)$, $\varphi_1(\omega)$, $W_f^e(k, \omega)$, and $\psi_k(\omega)$ are respectively the Fourier transforms of $W_f^e(0, t)$, $\varphi_1(t)$, $W_f^e(k, t)$, and $\psi_k(t)$.

B. IMPROVEMENT OF THE LARGEST SCALE SPACE

In the process of signal decomposition and reconstruction, segmenting the Fourier spectrum effectively and reasonably is key to the success of signal decomposition. Currently, the popular method is to use the maximum value of the spectrum interval as the boundary of the divided spectrum. However, when this method decomposes the signal containing noise interference, the local maximum value caused by the noise will be used as the boundary of the divided spectrum, which will lead to the spectrum. Over-segmentation means that the single signal is wrongly decomposed into two or more SSCs, which has a great negative impact on the final fault classification. Therefore, it is necessary to improve the EWT to achieve accurate and effective spectrum segmentation.

Gilles and Heal [27] proposed a parameter-free scale space method in 2014, projecting the spectrum onto the scale space and using the clustering method to divide the scale space into two categories (one is a meaningful spectrum segmentation boundary, and the other is a meaningless boundary). The classification criterion serves as a threshold for meaningful spectral segmentation.

Assuming that $f(x)$ is a continuous function defined on $[0, x_{max}]$, the scale space $L(x, t)$ of the function represented by (9).

$$
L(x, t) = g(x; t)^* f(x)
\tag{9}
$$

where the Gaussian kernel function $g(x; t) = \exp(-x^2/(2t))/\sqrt{2\pi t}$, $\sqrt{t}$ represents the scale parameter, $*$ represents the convolution product, and x represents the time series.

For a discrete function such as a signal, the scale space must also be discrete, so the scale space of a discrete signal is expressed as (10):

$$
L(x, t) = \sum_{n=-\infty}^{+\infty} f(x - n)g(n; t)
\tag{10}
$$

For real signals, a truncation threshold $M = C\sqrt{t} + 1$ is used to obtain a discrete scale space:

$$
L(x, t) = \sum_{n=-M}^{+M} f(x - n)g(n; t)
\tag{11}
$$

To ensure that the calculated scale space approximation error is less than 10^{-9}, $C = 6$ is usually taken. $\sqrt{t} = s\sqrt{t_0}$, $s \in [1, S_{max}]$ and s is an integer. The initial value $\sqrt{t_0} = 0.5$ is usually chosen because it corresponds to the midpoint value of two signal samples. $\sqrt{t_{max}} = s_{max}\sqrt{t_0}$, and $s_{max} = 2x_{max}$. Generally speaking, the scale curve is generated by the frequency corresponding to the minimum point in the frequency domain, and the length of the scale curve corresponding to other frequencies is 0.

Gilles and Heal [27] did not extend $\sqrt{t_{max}} = x_{max}$, which raises a problem: Since the scale parameter is too small, when the noise signal is large enough, its projection into the scale space fills the entire scale space with the same length as the normal signal. This will make the noise boundary and the meaningful boundary classified into one category, resulting in mistakenly using the noise boundary as the boundary of the divided spectrum. In order not to make such mistakes, the best way is to expand the scale space and display all the scale space of the signal. Therefore, under the premise of the full-time series, expand the scale parameter to the maximum value that can be calculated by the scaling function, and usually choose to add a constant ε to the maximum time series to test whether the scale space is fully displayed. The final modified scale parameter is $\sqrt{t_{max}} = x_{max} + \varepsilon$ [39]. The constant ε can be selected according to the following principles:

1) ε takes a constant much larger than the maximum time series, i.e., $\varepsilon \gg x_{max}$.
2) Calculate the largest scale parameter P_{max} that can be displayed, and then update ε, i.e., $\varepsilon = P_{max} - x_{max}$.

C. LIGHT-K-MEANS

Based on the selection of the above-mentioned largest scale space, we can know that the scale parameter is limited to the largest scale parameter that can be displayed, i.e., $\sqrt{t} \in [0, P_{max}]$. According to a large number of experiments, the extreme value of a signal is projected into the scale space, and when its resulting length value is greater than a given standard threshold, its corresponding frequency can be used as the boundary value of a meaningful segmented spectrum. The boundary of splitting the spectrum is nothing more than dividing into meaningful boundaries and meaningless boundaries, so boundary detection becomes a binary classification problem.

The frequencies in the largest scale space are divided into two categories according to the scale length, greater than or equal to the threshold T represents a meaningful spectrum division boundary, and others are represented as meaningless boundaries.

For clustering problems, the simplest and most effective method is the k-means algorithm [50]. For the k-means algorithm, the difficulty of clustering lies in the determination of the data set category (that is, the size of the k value) and the selection of the initial cluster center (once the initial value is not well selected, effective clustering results may not be obtained even give wrong results). For our binary

classification problem, it is obvious that k=2. Therefore, the biggest difficulty with the algorithm becomes the problem of initial cluster center selection.

Li et al. [51] proposed light-k-means under the premise that the number of categories k is known. The specific steps are as follows:

1) Randomly select p' data points from S, which are represented as H in the set, and the complement of H is H_c.
2) Apply the traditional k-means algorithm in the set H, and divide the set H into subsets.
3) Find the center point of the set.
4) Assign the set H_c to the subset closest to them.

However, there is a fatal problem when randomly selecting data points, if all the data points of the same category are selected, this method will consume more time than traditional k-means, and even get wrong results.

Inspired by this, we improved the light-k-means algorithm, the specific steps are as follows:

1) Constrain all data points to a rectangular window with sides a and b.
2) Compare the lengths of the side lengths a and b of the rectangular window, select the larger one, and record it as c.
3) Divide side c into k small windows on average.
4) Randomly select data points in each small window, and then perform the steps of the original light-k-means algorithm.
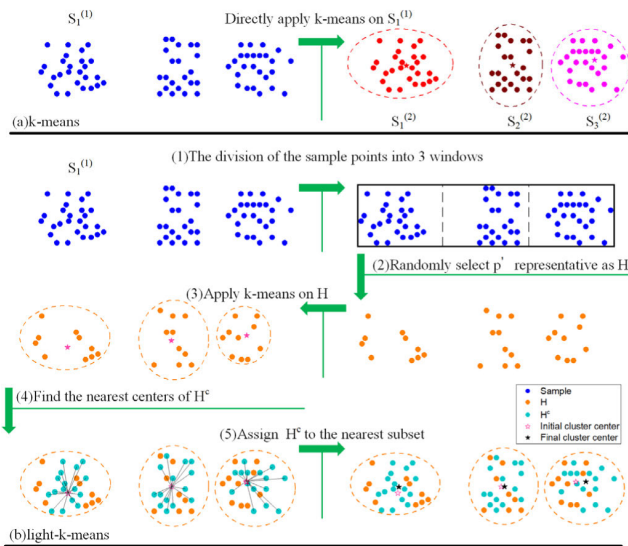

FIGURE 3. Comparison of the clustering process of k-means and light-k-means (a) The process of k-means clustering (b) The process of light-k-means clustering.

Fig. 3 shows the comparison of k-means and light-k-means clustering. It can be seen from the figure that light-k-means reduces the number of initial iterations by randomly selecting sample points, so that two calculations can greatly reduce the time cost, thereby greatly improving the calculation speed.

After all, we know that when the number of samples reaches a certain level, the number of iterations increases exponentially.

D. EXTRACTION OF FEATURES

The selection of the signal feature vector depends on the specific problem. Gamboa-Medina et al. [52] verified for the water network leakage problem that the feature vector based on the pressure signal depends on three features: energy (ENE), entropy (ENT), and the number of excess zeros (ZCC).

The dimensionless parameters are not affected by mechanical conditions and are widely used for the diagnosis of mechanical faults [53]. If the feature vector consisting of dimensionless parameters has poor ability to distinguish between different faults, the final achieved classification performance may be less than satisfactory, no matter how good the adopted learning algorithm is [54]. This means that we have to choose the appropriate dimensionless indicators in order to seek an accurate determination of the type of failure. Since the object of our study is a hydraulic system, its failure is not only the leakage problem, but also other failures such as slide valve failure. Inspired by the processing of raw data by Xiong et al. [55] using dimensionless indicators, we add 14 time-domain features of the mechanical principle to enhance the accuracy of feature extraction, including Mean Value, Standard Deviation, Variance, Peak-to-Peak Value, Square Root Amplitude, Average Amplitude, Mean Square Amplitude, Peak Value, Waveform Index, Peak Index, Impulsion Index, Clearance Factor, Degree of Skewness, and Kurtosis Value.

The above 17 features are combined, and a feature pool $S = \{F1, F2, \ldots, 17\}$ is created, whose expressions are shown in Table 1, where discrete data points $X = \{x_1, x_2, \ldots, x_n\}$. The subset that with the highest precision is found iteratively through the SFS [56]. The accuracy A can be expressed as:

$$A = \frac{\sum_{i=1}^{N} P(c_i = y_i)}{N} * 100\% \qquad (12)$$

where $P(s) = \begin{cases} 1 \ if \ s = true \\ 0 \ if \ s = false \end{cases}$, c_i is the classification prediction label, y_i is the actual fault label, and N is the total number of labels.

E. KELM

The output weight of the ELM [57] is calculated according to the input weight, and the weight is randomly generated, which makes the result unstable. Drawing on the experience of the successful application of kernel function in SVM, KELM [43] replaces the output matrix between the hidden layer and output layer with a kernel function. This not only avoids the uncertainty of the learning model but also retains the advantages of ELM. Therefore, the objective function F of KELM training is:

$$F = [K(x, x_1) \ldots K(x, x_N)](\frac{I}{C} + \Omega)^{-1}L \qquad (13)$$

TABLE 1. Expressions for 17 characteristic indicators.

Feature	Expression	Feature	Expression
Energy	$ENE(X) = \sum_{i=1}^{n} x_i^2$	Mean Square Amplitude	$MSA(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$
Entropy	$ENT(X) = -\sum_{i=1}^{n} \frac{x_i^2}{\sum_{i=1}^{n} x_i^2} \log_2 \frac{x_i^2}{\sum_{i=1}^{n} x_i^2}$	Peak Value	$PV(X) = \max(|MEAN(X)|)$
Zero Crossing Count	$ZCC(X) = \sum_{i=2}^{n} I(x_i x_{i-1} < 0)$	Waveform Index	$WI(X) = \frac{MSA(X)}{|MEAN(X)|}$
Mean	$MEAN(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$	Peak Index	$PI(X) = \frac{PV(X)}{MSA(X)}$
Standard Deviation	$SD(X) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - MEAN(X))^2}$	Impulsion Index	$II(X) = \frac{PV(X)}{MEAN(X)}$
Variance	$V(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - MEAN(X))^2$	Clearance Factor	$CF(X) = \frac{MSA(X)}{MEAN(X)}$
Peak-to-Peak Value	$PPV(X) = \max(x_i) - \min(x_i)$	Degree of Skewness	$DS(X) = \frac{\frac{1}{n}\sum_{i=1}^{n}(|x_i| - MEAN(X))^3}{SD(X)^3}$
Square Root Amplitude	$SRA(X) = \left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{|x_i|}\right)^2$	Kurtosis Value	$KV(X) = \frac{\frac{1}{n}\sum_{i=1}^{n}(|x_i| - MEAN(X))^4}{SD(X)^4}$
Average Amplitude	$AA(X) = \frac{1}{n}\sum_{i=1}^{n}|x_i|$		

where $K(x, x_i)$ is the kernel function, Ω is a matrix calculated by selecting different kernel functions, C is the regularization coefficient, I is the identity matrix, and L is the ideal output matrix.

F. POA
The POA is an intelligent optimization algorithm proposed by Trojovský and Dehghani [58] in 2022, which simulates the attack and hunting behavior of the pelican to build a model to solve the optimization problem. In POA, pelican hunting is divided into two processes, the approaching prey phase, and the surface flight phase. The position of the pelican changes with the position of the prey, and the steps to solve the optimization problem are as follows:

1) Determine the size of the pelican group and calculate the objective function value.
2) Approaching the prey phase: Calculate the position status of each pelican and update the group size.
3) Surface flight stage: Calculate whether the position state in 2 is to achieve a better objective function value, if not, keep the previous position state, and if so, update the position information and update the group size.
4) Keep the best candidate plan for each pelican, complete, and output the result.

G. POA-KELM
The regularization coefficients C and kernel parameters γ have a great influence on the prediction accuracy and test accuracy of KELM, and once their selection is not appropriate, it will greatly attenuate the performance of KELM.
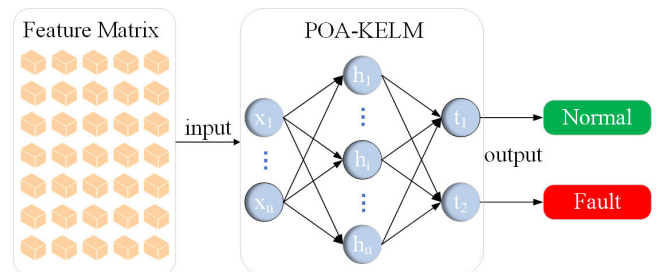

FIGURE 4. The workflow of POA-KELM.

Therefore, we use POA to optimize the regularization coefficients and kernel parameters of KELM, to make the accuracy of KELM higher. The workflow of POA-KELM is shown in Fig. 4. The specific optimization process is as follows:

1) Given an upper and lower bound, make a POA search in this interval.
2) Given the population range, calculate the objective function value and initialize KELM parameters.
3) Optimize C and γ of KLEM using POA.
4) Train KELM for each C and γ individually to derive the training accuracy.
5) Keep the C and γ that make the highest accuracy.

H. FRAMEWORK FOR FAULT CLASSIFICATION
The framework for fault classification using pressure signals is shown in Fig. 1, and some key steps are explained as follows:

1) Use the pressure signal that reflects the state of the hydraulic components as the initial signal.
2) Transferring the signal to the frequency domain by Fourier transform, and the spectrum is projected under the maximum scale space, and the maximum scale space is split into two parts using light-k-means, leaving the part larger than the threshold T as the boundary of the split Fourier spectrum. The improved EWT decomposes the initial signal to obtain a series of SSCs.
3) Calculate 17 feature indicators for each component, input each feature indicator into the original KELM and calculate its test accuracy, select the highest accuracy feature indicator in each round until the test accuracy no longer increases (i.e., SFS strategy), select the optimal feature for each SSC, and compose the feature vector as the input of POA-KELM.
4) Input the feature vector into POA-KELM, and use POA to update C and γ of KELM to derive the highest accuracy solution. Finally, classify the faults and get the results.
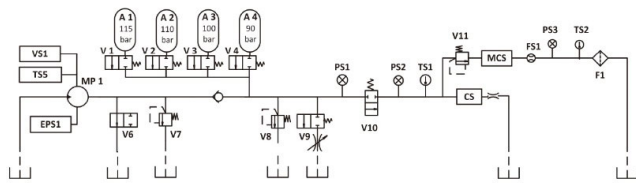

FIGURE 5. Working hydraulics for data acquisition with switchable orifice V9.

III. EXPERIMENT AND ANALYSIS

A. EXPERIMENTAL MODEL

To verify the validity of the improved EWT and POA-KELM, we use the pressure data sets collected by Helwig et al. [59] for multivariable faults of various components of the hydraulic system, which was experimented on a hydraulic test stand and obtained through sensors. The test stand consists of a primary working circuit and a secondary cooling-filtration circuit, connected through a tank, where the working hydraulic system principle is shown in Fig. 5. Because of the difference in the location of the sensors, they serve different objects. We know that sensor PS1 is very sensitive to the performance of the hydraulic pump and the hydraulic valve by the calculation of the characteristic values of Liu et al. [12] for the same component of the 3 pressure sensors. So we choose three kinds of failure samples of the hydraulic pump without leakage, weak leakage, and serious leakage, and four kinds of failure samples of the hydraulic valve without jamming, slight jamming, serious jamming, and near failure. The hydraulic pump does not leak and the hydraulic valve does not jam as a set of a normal state of the hydraulic system, then they have 12 combinations, the combination number 1-1, 1-4, ..., 3-6, each combination has 10 groups of the original signal, a total of 120 groups; fault type has a total of 6, the fault label is 1, 2, ..., 6. The pressure signals of groups 1-1

are not decomposed, groups 1-4, 1-5, 1-6, 2-1, and 3-1 are decomposed into 2 sub-signals, and the rest are decomposed into 3 sub-signals, so there are 290 groups of signals in total. The 17 eigenvalues of each of the 290 groups of signals are calculated separately to calculate their eigenvectors by the SFS strategy. The resulting eigenvectors are trained and validated in a 4:1 manner, and finally, the sample data are expanded using a five-fold cross-validation method to make the results more accurate. For the pressure data acquisition experiments, the system cycle was repeated with a constant load cycle (duration of 60 s) and a sampling frequency of 100 Hz. Table 2 describes the different fault types.

TABLE 2. Description of the fault.

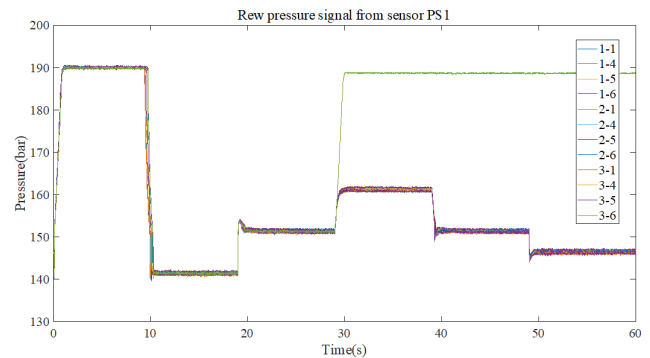| Fault part	Fault Type	Diameter/ valve lift	Classification label
Hydraulic pump MP1	No leakage	0mm	1
	Weak leakage	0.2mm	2
	Serious leakage	0.25mm	3
Hydraulic valve V10	No lag	100%	1
	Small lag	90%	4
	Severe lag	80%	5
	Near failure	73%	6

FIGURE 6. 12 types of raw pressure signals.

Fig. 6 indicates the original pressure signal for each fault type measured by sensor PS1. We can find that among the signals in groups 3-6, both the hydraulic pump and the hydraulic valve are the most serious failures, still working properly in the first 20s, but as time goes by, both the valve and the pump fail to the point of not following the laws of the hydraulic system. The remaining 11 groups of signals are not so different that the naked eye can not distinguish their categories, so it is necessary to improve the EWT decomposition of each pressure signal.

B. IMPROVED EWT DECOMPOSITION RESULTS

According to A, based on each sensor data, we can get 120 signals containing 12 types of faults. Next, we will introduce the decomposition process of the improved EWT and the decomposition results in detail, using groups 2-4 as examples. The sampling frequency is chosen to be 100 Hz and
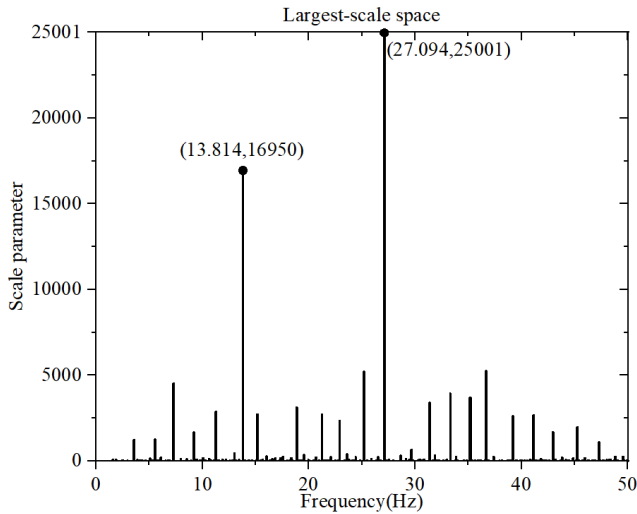
FIGURE 7. Largest-scale space of pressure signals.

FIGURE 8. Selection of meaningful points based on light-k-means.

the running time is 60 seconds. The working algorithm environment is MATLAB 2016b, the hardware central processing unit (CPU) is Intel ®Core™ i5-6500 CPU @ 3.20GHz, and the random access memory (RAM) of the computer is 4.00 GB. Assuming the initial constant $\varepsilon = 5x_{max} = 30000$, calculate the maximum scale parameter $P_{max} = 25001$, update the constant $\varepsilon = 19001$, and the signals are expanded under the largest scale space (as shown in Fig. 7). The scale parameters of each signal extremum are ordered in ascending order, from the lowest to the highest. The number of categories is known to be 2, so we constrain all data points to a 957×25000 rectangular space and divide it equally into 2 windows A and B. Randomly select data points in these two windows. Here we select 1 data point in window A and 10 data points in window B. The center coordinates (957, 25001) and (690.1, 209.3) of the data points in the two windows are obtained using the k-means algorithm, respectively. Using these two center coordinates as initial coordinates, the k-means algorithm is applied to all data points. The final updated clustering centers are (956.5, 20975.5) and (478, 82.8). Thus, under the original scale space, the two data points (27.094, 25001) and (13.814, 16950) are divided into one class, and the rest of the data points are divided into another class. The final clustering results are shown in Fig. 8. Therefore, we set the boundaries for dividing the Fourier spectrum as 27.094 and 13.814, as shown in Fig. 9. Since the amplitude of the Fourier spectrum at low frequencies is so large that the information at low amplitudes cannot be seen, we zoom in on the Fourier frequency domain to study the information at low amplitudes as shown in Fig. 10. Finally, 2 meaningful points are obtained to split the Fourier spectrum, and 3 SSCs are obtained (as shown in Fig. 11). It is obvious from Fig. 11 that the SSC (1) is very similar to the normal signal (groups 1-1), which shows that we successfully decompose the fault signal into a normal signal, a hydraulic pump leakage fault signal, and a hydraulic valve jamming
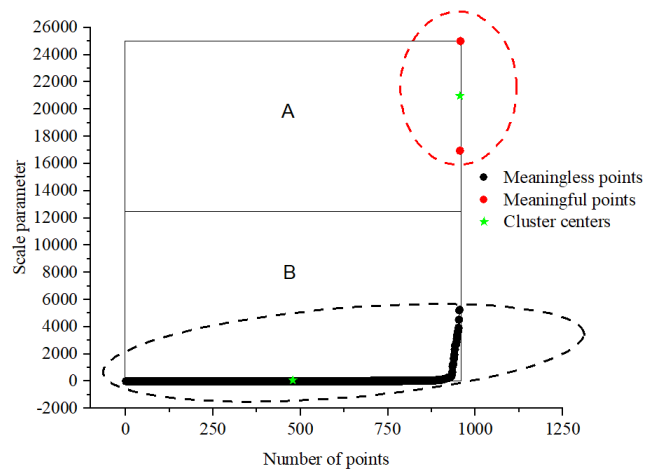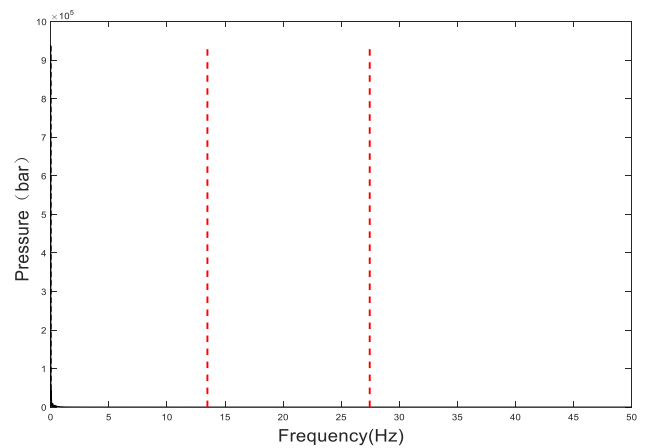

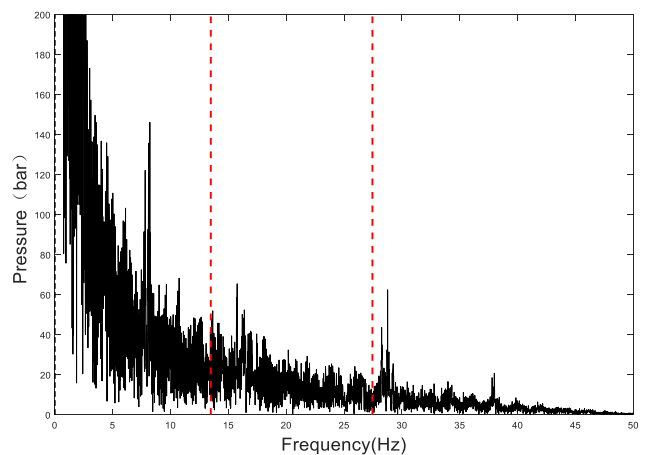FIGURE 9. Fourier spectrum division of pressure signals.

FIGURE 10. Local amplification of Fourier spectrum division of pressure signals.

fault signal. This also precisely illustrates the effectiveness of the improved EWT proposed in this study and also shows that
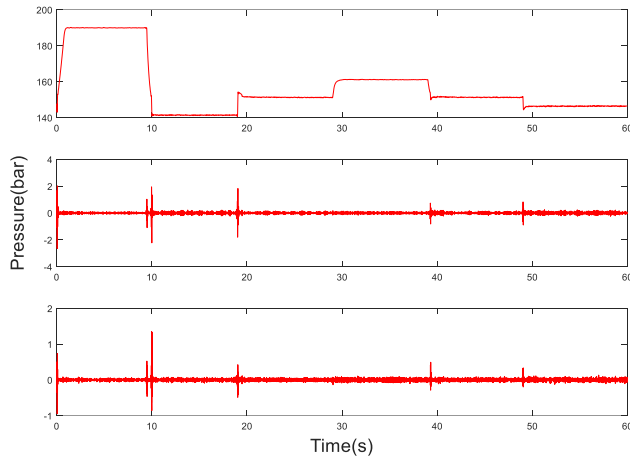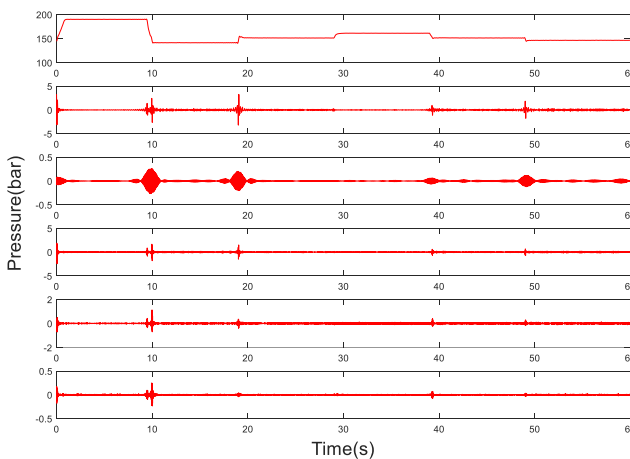
FIGURE 11. EWT based on light-k-means.

FIGURE 12. EWT based on empirical law.

the pressure signal characteristics can effectively detect faults in hydraulic systems. Through the processing of the signal by EWT, we transformed the multivariate fault signal into a univariate fault classification problem, which helped greatly in the subsequent work.

We compared light-k-means with empirical law and other clustering methods as a way to illustrate the superiority of the method proposed in this study. The comparison of the decomposed signals from Fig. 11 and Fig. 12 shows that the number of SCCs obtained by the improved EWT decomposition based on light-k-means increases from 3 to 6. This is because the EWT based on the empirical law clustering method divides the three scale parameters caused by hydraulic pump leakage failure or hydraulic valve jamming failure into a group of meaningful boundaries together. This leads to the over-decomposition of one of the fault signals into 4 signals when the signal is decomposed, so we are not clear about what these 4 signals represent, and therefore incorrectly classify them into 6 categories when the fault is classified. According to our attempts at other clustering methods, we found that the segmentation boundaries calculated

by all the methods except k-means and light-k-means are greater than 2. Compared with traditional k-means, light-k-means reduces the iteration time by orders of magnitude after randomly selecting samples, so it can do its job simply and efficiently in a very short time.

C. IMPROVING THE SUPERIORITY OF EWT

To further illustrate the superiority of the improved EWT, we compared the improved EWT with other clustering methods in terms of the number of segmentation boundaries, error statistics, precision, and time consumption. Among them, the error statistics include mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE), weighted mean absolute percentage error (WMAPE), and Fréchet distance (FD) [60].

FD defines a method to calculate the similarity of two curves considering the relationship between the positions of two points, which is commonly used for time series similarity measure and trajectory series. Eiter and Heikki [61] computed recursively the Fréchet distance of two discrete series. The expressions for the error statistics and FD can be found in Table 3, where y_i refers to the original signal sequence and y_i' refers to the signal reconstruction sequence. $c(i-1, j)$ denotes the distance from a point $(i-1, j)$ to point (i, j), and $d(ui, vj)$ denotes the Euclidean distance between points i and j.

The accuracy was calculated by the impulse factor proposed by Yu et al. [62]. The impulse factor metric is the ratio of the fault shock frequency amplitude E_f to the intrinsic pulse frequency amplitude E_0:

$$\alpha = E_f / E_0 \qquad (14)$$

The boundary detection methods used in the traditional EWT largest scale space include Empirical law, Otsu, Half-Normal law, and k-means, and to enhance the convincing power, we also added the DBSCAN method [39] for comparison. The reconstructed signal for each component is calculated according to (8), followed by the calculation of the comparison of each parameter mentioned above. The comparison results are shown in Table 4.

As can be seen from Table 4, the difference between light-k-means and k-means data is not large, but the difference in time spent between the two is quite large. This is because light-k-means randomly select samples to reduce the iteration time and the time reduces by orders of magnitude, thus greatly improving the computational speed. DBSCAN has obvious advantages in vibration signal-based fusion, but for pressure signals, it has obvious limitations. This is because the choice of the density radius r and the minimum number of Minpts is highly artificial, and the scale space under pressure signals is more dense than that under vibration signals. Therefore, if r or Minpts is too large or too small, it will lead to problems in scale space clustering. light-k-means is the most suitable spectral boundary selection method for EWT to decompose pressure signals, both in terms of accuracy and error statistics

TABLE 3. The expressions for the error statistics.

Error statistics	Expression
MAE	$\dfrac{1}{n}\sum_{i=1}^{n}\left\|y_i - y_i'\right\|$
MSE	$\dfrac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i'\right)^2$
MAPE	$\dfrac{1}{n}\sum_{i=1}^{n}\dfrac{\left\|y_i - y_i'\right\|}{y_i}$
WMAPE	$\dfrac{\sum_{i=1}^{n}\left\|y_i - y_i'\right\|}{\sum_{i=1}^{n}y_i}$
FD	$\max\{\min(c(i-1,j), c(i-1,j-1), c(i,j-1)), d(ui,vj)\}$

TABLE 4. Comparison of several detection methods.

principals	number of components	accuracy	computational cost	MAE	MSE	MAPE	WMAPE	FD
light-k-means	3	0.49517	1.4524245s	0.373	1.176	0.5296	1.866	0.09134
k-means	3	0.49517	17.542778s	0.373	1.176	0.5296	1.866	0.09134
means	17	0.36424	41.356736s	0.767	0.591	0.3594	5.729	1.82453
DBSCAN	10	0.27453	18.383977s	0.916	1.146	0.1846	4.792	0.76855
Empirical law	6	0.46452	8.3324524s	0.439	0.267	0.2443	3.368	0.03724
Otsu	35	0.20542	80.484945s	1.834	4.377	0.7095	7.343	1.42587
Half-Normal law	42	0.38643	99.35325s	3.394	9.521	1.0435	8.923	1.77535

and in terms of time consumption, thus demonstrating the effectiveness and rapidity of improved EWT.

D. CLASSIFICATION OF FAULTS

Because the radial basis function (RBF) can map samples to a higher-dimensional space to solve many nonlinear problems, the kernel function we choose is the Gaussian kernel function, where the Gaussian kernel function is expressed as:

$$K(x, x_i) = \exp(-\frac{(x - x_i)^2}{\gamma^2}) \qquad (15)$$

where γ is the parameter of the Gaussian kernel function.

The 290 groups of SSCs are divided into three groups training, testing, and verification according to 3:1:1 to obtain the feature vectors. The 17 feature values of the 290 groups of SSCs obtained by EWT decomposition are calculated separately, and each feature value is concatenated into the training group to train the original KELM (the initial Gaussian kernel parameter is set to 4, and the regularization coefficient is set to 2) to get the testing accuracy, and the SFS-based feature selection process is shown in Table 5. First, the feature pool is emptied, and in the first round, we can see that the highest test accuracy is obtained by training KELM with feature 11, so the pool is updated to {F11}, and then KELM is trained with the remaining 16 features except feature 11 and the above operation is repeated until the accuracy no longer increases. According to our experiments, starting from the fifth round, the accuracy of each feature input is lower than 93.79%, so we stop the loop after the fourth round and combine

TABLE 5. SFS-based feature selection.

	Round 1	Round 2	Round 3	Round 4
Feature 1	0.7069	0.6379	0.6897	0.6897
Feature 2	0.7931	0.7759	0.9138	-
Feature 3	0.4138	0	0	0
Feature 4	0.7759	0.9138	-	-
Feature 5	0.6034	0.8793	0.9138	0.9379
Feature 6	0.5862	0.7931	0.9138	0.9379
Feature 7	0.7931	0.9138	0.8966	0.9138
Feature 8	0.5517	0.569	0.8966	0.9379
Feature 9	0.5862	0.9138	0.8966	0.9379
Feature 10	0.6034	0.9138	0.9138	0.9138
Feature 11	0.8966	-	-	-
Feature 12	0.7413	0.8966	0.9138	0.9138
Feature 13	0.8103	0.7069	0.7586	0.7931
Feature 14	0.7414	0.569	0.6552	0.6724
Feature 15	0.7241	0.6379	0.6552	0.5517
Feature 16	0.7586	0.8793	0.8966	0.9138
Feature 17	0.569	0.569	0.431	0.4483
Selected	Feature 11	Feature 4	Feature 2	Feature 5
Feature Pool	{F11}	{F11, F4}	{F11, F4, F2}	{F11, F4, F2, F5}

{F11, F4, F2, F5} to form the feature vectors are used as input to POA-KELM.

To classify the results more accurately, we use a five-fold cross-validation to expand the samples. The above operation is cycled through five different compositions of test groups, and finally, 1450 data samples are obtained. It is found that no matter how we replace the compositions between the test and training groups, the final selected feature pool is still {F11, F4, F2, F5}, which illustrates the effectiveness of this feature extraction method from the side. Finally, the data are
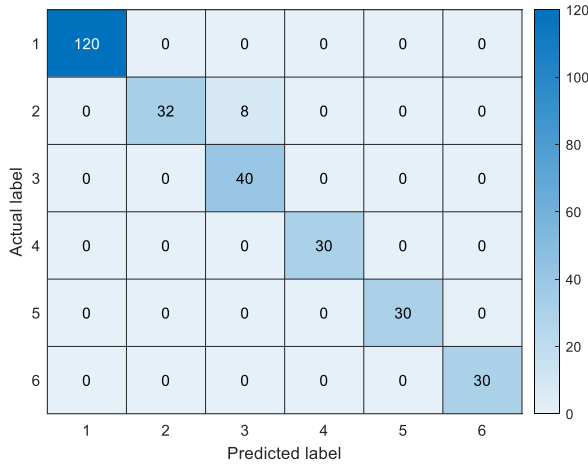
FIGURE 13. Confusion matrix of POA-KELM test set.

divided according to the ratio of training and testing of 4:1. The final confusion matrix is obtained as shown in Fig. 13. Thus the computational accuracy is 97.24%.
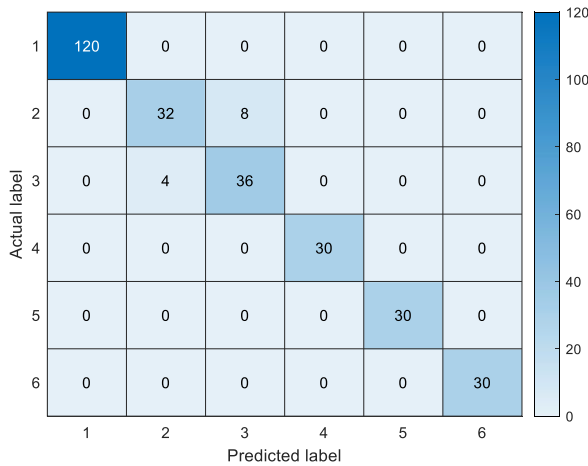

FIGURE 14. Confusion matrix of SOA-KELM test set.

E. COMPARISON OF CLASSIFIER PREDICTION ACCURACY
In order to demonstrate the supremacy of POA-KELM in multivariate fault diagnosis within hydraulic systems, identical signal sources were utilized and subjected to multi-variable fault classifiers, including the Seagull Optimization Algorithm enhanced KELM (SOA-KELM) and the original KELM. The outputs were then analysed for discrepancies in both efficiency and precision, thus indicating the dispar-ities between the aforementioned classifiers. Their confusion matrix is shown in the Fig. 13, Fig. 14 and Fig. 15, where the abscissa represents the predicted fault type, the ordinate represents the actual fault type, the value on the main diagonal represents the number of correct predictions, and the numbers in other positions represent the number of incorrect predic-tions the sum of numbers represents the number of samples in the test set. As can be seen in Fig. 13, Fig. 14 and Fig. 15, the second and third types of failures do not match the
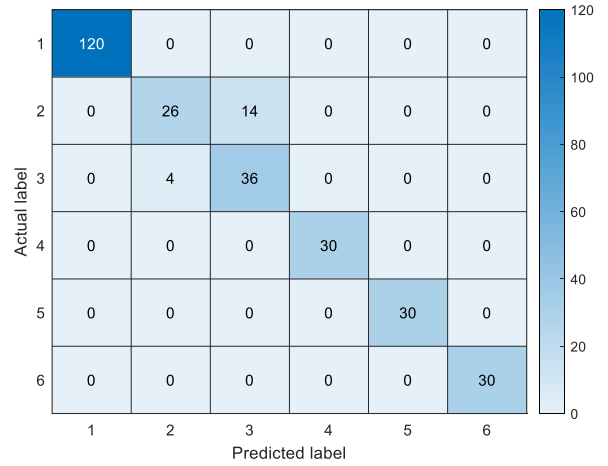

FIGURE 15. Confusion matrix of the KELM test set.

reality when predicted. We find that the signals of these two categories differ only in their peak values when performing the calculation of the eigenvectors. However, the change in the external temperature of the hydraulic system at certain moments makes the peak values of the second and third category of faults very close to each other. Therefore, the POA-KELM classifier incorrectly classifies the two classes of faults into the same category. With regards to the remaining categories of faults, there exists a considerable discrepancy among the four characteristics associated with these faults as well as an elevated resistance to interference, resulting in an exact correspondence between the predicted and actual fault types.
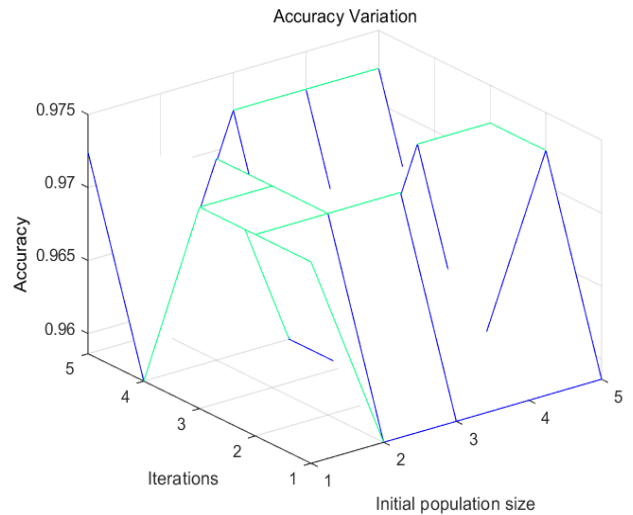

FIGURE 16. Variation of POA-KELM test accuracy with population size and iteration times.

For comparison, the kernel functions we chose were all RBF. The number of initial populations and the number of iterations of POA have a great influence on the accu-racy of the optimization, and they must be chosen appro-priately, otherwise, they will cause a decrease in accuracy. Fig. 16 and Fig. 17 show the variation in the accuracy of
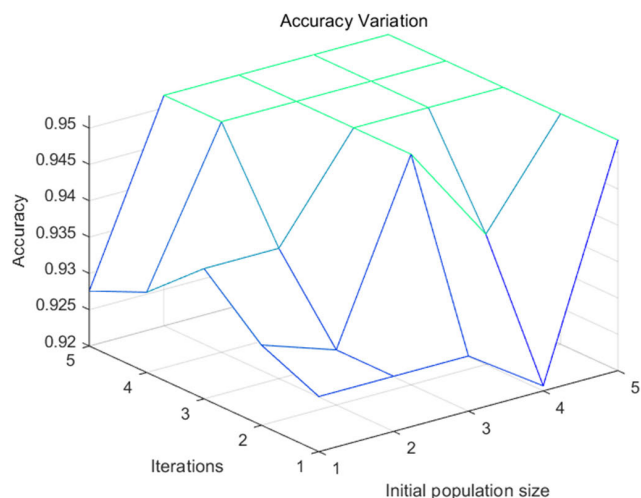
FIGURE 17. Variation of SOA-KELM test accuracy with population size and iteration times.

POA-KELM and SOA-KELM with the number of initial populations and the number of iterations. We can find that POA-KELM can achieve the highest accuracy of 97.24% when the initial population is 1, the initial population and the number of iterations are updated to 1 and 1, respectively, and the regularization coefficient and kernel parameters are 858.0409 and 23.7134, respectively, and the time required is 0.381905 seconds. In contrast, SOA-KELM achieves a maximum test accuracy of 95.17% when the initial number of populations and the number of iterations are 2 and 4. After updating the initial number of populations and the number of iterations, the obtained regularization coefficients and kernel parameters are 44.0626 and 0.1, respectively, and the time required is 0.565492 seconds. We give KELM initial regularization coefficients and kernel parameters of 2 and 4, respectively, and KELM The test accuracy obtained is 93.79% and the time required is 0.379686 seconds. The results show that POA-KELM is undoubtedly the most suitable method for multivariate fault diagnosis of hydraulic systems in terms of balance time and accuracy.

IV. CONCLUSION

In this paper, the improved EWT and POA optimized KELM are combined to solve the problem of low efficiency of multivariable fault diagnosis in hydraulic systems. The improved EWT is used to reduce the dimension of multivariable problems and greatly reduce the calculation time. Then POA-KELM is used to classify the decomposed signals, which improves the accuracy of fault classification. This article solves the problem of inaccurate spectral segmentation in EWT by expanding the scale space. Using the clustering method of light-k-means, the iteration time of data points is greatly shortened, and the computational efficiency is improved. Compared with other boundary detection methods, the improved EWT has high accuracy and high computational speed in signal reconstruction. The SFS method was used to

remove the useless features in the SSCs decomposed by EWT, and the feature vectors of the signals were selected, which not only improved the computation speed, but also retained the features that could best reflect the signal trend. Input the feature vector into the POA-KELM classifier to predict the fault type.

Due to the fast nature of the POA optimization method, the KELM classifier achieves the highest accuracy in very few iteration cycles, thus reducing the overall fault diagnosis time and benefiting even more from the accuracy of the decomposition of the improved EWT method, which makes the distinction between the characteristics of each type of fault more obvious. In conclusion, the combination of improved EWT and POA-KELM improves the diagnostic accuracy of multivariable faults and also reduces the diagnostic time, which is undoubtedly beneficial for scientific progress.

Despite the achievements obtained by this study, there are still some limitations. Firstly, the 17 features used in this study have a weak capability to identify the second and third types of faults. To address this issue, future studies should focus on improving the feature selection process to enhance the ability to resist interference. Secondly, this study did not consider the presence of sensor faults in hydraulic systems, which is a common problem. Thus, future studies should incorporate multi-sensor information fusion techniques to detect sensor faults and eliminate the effects of single-sensor errors.

Based on these limitations, future research can further explore how to improve the accuracy and efficiency of fault diagnosis. For instance, other machine learning algorithms can be applied to fault diagnosis, or more efficient feature extraction methods can be developed. In summary, the limitations of this study provide a vast space for future research to explore.

REFERENCES

[1] E. Artigao, S. Martín-Martínez, A. Honrubia-Escribano, and E. Gómez-Lázaro, "Wind turbine reliability: A comprehensive review towards effective condition monitoring development," *Appl. Energy*, vol. 228, pp. 1569–1583, Oct. 2018.

[2] Z. Chen, W. Wang, Z. Xu, and Y. Zhao, "Faults simulation and intelligent diagnosis analysis of hydraulic system of multi-functional laying trolley," in *Proc. IEEE 8th Int. Conf. Fluid Power Mechatronics (FPM)*, Wuhan, China, Apr. 2019, pp. 15–22.

[3] L. X. Hou, J. Kong, F. G. Li, C. X. Ju, and W. Xu, "Analysis and control of the fault in the hydraulic system of coal-machinery," *Adv. Mater. Res.*, vols. 1061–1062, pp. 921–926, Dec. 2014.

[4] X. Wang, S. Lin, S. Wang, Z. He, and C. Zhang, "Remaining useful life prediction based on the Wiener process for an aviation axial piston pump," *Chin. J. Aeronaut.*, vol. 29, no. 3, pp. 779–788, Jun. 2016.

[5] S. Yin, S. X. Ding, and D. Zhou, "Diagnosis and prognosis for complicated industrial systems—Part II," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3201–3204, May 2016.

[6] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6418–6428, Nov. 2014.

[7] S. Tang, Y. Zhu, and S. Yuan, "An improved convolutional neural network with an adaptable learning rate towards multi-signal fault diagnosis of hydraulic piston pump," *Adv. Eng. Inf.*, vol. 50, Oct. 2021, Art. no. 101406.

[8] S. Tang, Y. Zhu, and S. Yuan, "A novel adaptive convolutional neural network for fault diagnosis of hydraulic piston pump with acoustic images," *Adv. Eng. Informat.*, vol. 52, Apr. 2022, Art. no. 101554.

[9] Y. Zhu, G. Li, R. Wang, S. Tang, H. Su, and K. Cao, "Intelligent fault diagnosis of hydraulic piston pump combining improved LeNet-5 and PSO hyperparameter optimization," *Appl. Acoust.*, vol. 183, Dec. 2021, Art. no. 108336.

[10] X. Guo, K. Lu, Y. Cheng, W. Zhao, H. Wu, D. Li, J. Li, S. Yang, and Y. Zhang, "Research on fault diagnosis method for hydraulic system of CFETR blanket transfer device based on CNN-LSTM," *Fusion Eng. Des.*, vol. 185, Dec. 2022, Art. no. 113321.

[11] X. Ji, Y. Ren, H. Tang, C. Shi, and J. Xiang, "An intelligent fault diagnosis approach based on Dempster–Shafer theory for hydraulic valves," *Measurement*, vol. 165, Dec. 2020, Art. no. 108129.

[12] J. Liu, H. Xu, X. Peng, J. Wang, and C. He, "Reliable composite fault diagnosis of hydraulic systems based on linear discriminant analysis and multi-output hybrid kernel extreme learning machine," *Rel. Eng. Syst. Saf.*, vol. 234, Jun. 2023, Art. no. 109178.

[13] Z. Qiu, R. Min, D. Wang, and S. Fan, "Energy features fusion based hydraulic cylinder seal wear and internal leakage fault diagnosis method," *Measurement*, vol. 195, May 2022, Art. no. 111042.

[14] L. Meng, M. Zhao, Z. Cui, X. Zhang, and S. Zhong, "Empirical mode reconstruction: Preserving intrinsic components in data augmentation for intelligent fault diagnosis of civil aviation hydraulic pumps," *Comput. Ind.*, vol. 134, Jan. 2022, Art. no. 103557.

[15] S. Tang, Y. Zhu, and S. Yuan, "An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal," *Eng. Failure Anal.*, vol. 138, Aug. 2022, Art. no. 106300.

[16] J. Shi, J. Yi, Y. Ren, Y. Li, Q. Zhong, H. Tang, and L. Chen, "Fault diagnosis in a hydraulic directional valve using a two-stage multi-sensor information fusion," *Measurement*, vol. 179, Jul. 2021, Art. no. 109460.

[17] J. Xiong, M. Liu, C. Li, J. Cen, Q. Zhang, and Q. Liu, "A bearing fault diagnosis method based on improved mutual dimensionless and deep learning," *IEEE Sensors J.*, vol. 23, no. 16, pp. 18338–18348, Aug. 2023.

[18] J. Xiong, C. Li, C.-D. Wang, J. Cen, Q. Wang, and S. Wang, "Application of convolutional neural network and data preprocessing by mutual dimensionless and similar Gram matrix in fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1061–1071, Feb. 2022.

[19] T. Zhongda, L. Shujiang, W. Yanhong, and S. Yi, "A prediction method based on wavelet transform and multiple models fusion for chaotic time series," *Chaos, Solitons Fractals*, vol. 98, pp. 158–172, May 2017.

[20] G. Cheng, X. Chen, H. Li, P. Li, and H. Liu, "Study on planetary gear fault diagnosis based on entropy feature fusion of ensemble empirical mode decomposition," *Measurement*, vol. 91, pp. 140–154, Sep. 2016.

[21] C. Guo, Y. Wen, P. Li, and J. Wen, "Adaptive noise cancellation based on EMD in water-supply pipeline leak detection," *Measurement*, vol. 79, pp. 188–197, Feb. 2016.

[22] D. Yang, H. Li, Y. Hu, J. Zhao, H. Xiao, and Y. Lan, "Vibration condition monitoring system for wind turbine bearings based on noise suppression with multi-point data fusion," *Renew. Energy*, vol. 92, pp. 104–116, Jul. 2016.

[23] J. Gilles, "Empirical wavelet transform," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 3999–4010, Aug. 2013.

[24] J. Zheng, H. Pan, S. Yang, and J. Cheng, "Adaptive parameterless empirical wavelet transform based time-frequency analysis method and its application to rotor rubbing fault diagnosis," *Signal Process.*, vol. 130, pp. 305–314, Jan. 2017.

[25] J. Chen, J. Pan, Z. Li, Y. Zi, and X. Chen, "Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals," *Renew. Energy*, vol. 89, pp. 80–92, Apr. 2016.

[26] Y. Hu, X. Tu, F. Li, and G. Meng, "Joint high-order synchrosqueezing transform and multi-taper empirical wavelet transform for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions," *Sensors*, vol. 18, no. 1, p. 150, Jan. 2018.

[27] J. Gilles and K. Heal, "A parameterless scale-space approach to find meaningful modes in histograms—Application to image and spectrum segmentation," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 12, no. 6, Nov. 2014, Art. no. 1450044.

[28] Y. Hu, F. Li, H. Li, and C. Liu, "An enhanced empirical wavelet transform for noisy and non-stationary signal processing," *Digit. Signal Process.*, vol. 60, pp. 220–229, Jan. 2017.

[29] J. Pan, J. Chen, Y. Zi, Y. Li, and Z. He, "Mono-component feature extraction for mechanical fault diagnosis using modified empirical wavelet transform via data-driven adaptive Fourier spectrum segment," *Mech. Syst. Signal Process.*, vols. 72–73, pp. 160–183, May 2016.

[30] Y. Kong, T. Wang, and F. Chu, "Meshing frequency modulation assisted empirical wavelet transform for fault diagnosis of wind turbine planetary ring gear," *Renew. Energy*, vol. 132, pp. 1373–1388, Mar. 2019.

[31] Y. Kim, J. M. Ha, K. Na, J. Park, and B. D. Youn, "Cepstrum-assisted empirical wavelet transform (CEWT)-based improved demodulation analysis for fault diagnostics of planetary gearboxes," *Measurement*, vol. 183, Oct. 2021, Art. no. 109796.

[32] S. Schmidt and K. C. Gryllias, "The anomalous and smoothed anomalous envelope spectra for rotating machine fault diagnosis," *Mech. Syst. Signal Process.*, vol. 158, Sep. 2021, Art. no. 107770.

[33] M. Kedadouche, Z. Liu, and V.-H. Vu, "A new approach based on OMA-empirical wavelet transforms for bearing fault diagnosis," *Measurement*, vol. 90, pp. 292–308, Aug. 2016.

[34] H. Zhao, S. Zuo, M. Hou, W. Liu, L. Yu, X. Yang, and W. Deng, "A novel adaptive signal processing method based on enhanced empirical wavelet transform technology," *Sensors*, vol. 18, no. 10, p. 3323, Oct. 2018.

[35] K. Zhang, C. Ma, Y. Xu, P. Chen, and J. Du, "Feature extraction method based on adaptive and concise empirical wavelet transform and its applications in bearing fault diagnosis," *Measurement*, vol. 172, Feb. 2021, Art. no. 108976.

[36] K. Zhang, L. Shi, Y. Hu, P. Chen, and Y. Xu, "Variable spectral segmentation empirical wavelet transform for noisy signal processing," *Digit. Signal Process.*, vol. 117, Oct. 2021, Art. no. 103151.

[37] J. Ding, "A double impulsiveness measurement indices-bilaterally driven empirical wavelet transform and its application to wheelset-bearing-system compound fault detection," *Measurement*, vol. 175, Apr. 2021, Art. no. 109135.

[38] Z. Zheng, Z. Wang, Y. Zhu, S. Tang, and B. Wang, "Feature extraction method for hydraulic pump fault signal based on improved empirical wavelet transform," *Processes*, vol. 7, no. 11, p. 824, Nov. 2019.

[39] H. Yu, H. Li, and Y. Li, "Vibration signal fusion using improved empirical wavelet transform and variance contribution rate for weak fault detection of hydraulic pumps," *ISA Trans.*, vol. 107, pp. 385–401, Dec. 2020.

[40] X.-B. Fu, B. Liu, Y.-C. Zhang, and L.-N. Lian, "Fault diagnosis of hydraulic system in large forging hydraulic press," *Measurement*, vol. 49, pp. 390–396, Mar. 2014.

[41] C. Lu, S. Wang, and X. Wang, "A multi-source information fusion fault diagnosis for aviation hydraulic pump based on the new evidence similarity distance," *Aerosp. Sci. Technol.*, vol. 71, pp. 392–401, Dec. 2017.

[42] Y.-X. Li, S.-B. Jiao, and X. Gao, "A novel signal feature extraction technology based on empirical wavelet transform and reverse dispersion entropy," *Defence Technol.*, vol. 17, no. 5, pp. 1625–1635, Oct. 2021.

[43] S. Lu, W. Gao, C. Hong, and Y. Sun, "A newly-designed fault diagnostic method for transformers via improved empirical wavelet transform and kernel extreme learning machine," *Adv. Eng. Informat.*, vol. 49, Aug. 2021, Art. no. 101320.

[44] Y. Ding, L. Ma, C. Wang, and L. Tao, "An EWT-PCA and extreme learning machine based diagnosis approach for hydraulic pump," *IFAC-PapersOnLine*, vol. 53, no. 3, pp. 43–47, 2020.

[45] X. Liu, H. Huang, and J. Xiang, "A personalized diagnosis method to detect faults in gears using numerical simulation and extreme learning machine," *Knowl.-Based Syst.*, vol. 195, May 2020, Art. no. 105653.

[46] P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. Slaoui-Hasnaoui, T. Tameghe, and G. Ekemb, "Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges," *Energies*, vol. 7, no. 4, pp. 2595–2630, Apr. 2014.

[47] L. Zuo, F. Xu, C. Zhang, T. Xiahou, and Y. Liu, "A multi-layer spiking neural network-based approach to bearing fault diagnosis," *Rel. Eng. Syst. Saf.*, vol. 225, Sep. 2022, Art. no. 108561.

[48] A. Kumar, C. Parkash, G. Vashishtha, H. Tang, P. Kundu, and J. Xiang, "State-space modeling and novel entropy-based health indicator for dynamic degradation monitoring of rolling element bearing," *Rel. Eng. Syst. Saf.*, vol. 221, May 2022, Art. no. 108356.

[49] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992.

[50] J. Han, M. Kamber, and J. Pei, "Advanced cluster analysis," in *Data Mining: Concepts and Techniques*, vol. 3. The Netherlands: Elsevier, 2012, pp. 501–508.

[51] H. Li, X. Ye, A. Imakura, and T. Sakurai, "Divide-and-conquer based large-scale spectral clustering," *Neurocomputing*, vol. 501, pp. 664–678, Aug. 2022.

[52] M. M. Gamboa-Medina, L. F. R. Reis, and R. C. Guido, "Feature extraction in pressure signals for leak detection in water networks," *Proc. Eng.*, vol. 70, pp. 688–697, Jan. 2014.

[53] J. Xiong, Q. Liang, J. Wan, Q. Zhang, X. Chen, and R. Ma, "The order statistics correlation coefficient and PPMCC fuse non-dimension in fault diagnosis of rotating petrochemical unit," *IEEE Sensors J.*, vol. 18, no. 11, pp. 4704–4714, Jun. 2018.

[54] A. Qin, Q. Hu, Y. Lv, and Q. Zhang, "Concurrent fault diagnosis based on Bayesian discriminating analysis and time series analysis with dimensionless parameters," *IEEE Sensors J.*, vol. 19, no. 6, pp. 2254–2265, Mar. 2019.

[55] J. Xiong, Q. Zhang, J. Wan, L. Liang, P. Cheng, and Q. Liang, "Data fusion method based on mutual dimensionless," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 2, pp. 506–517, Apr. 2018.

[56] Z. Tan, J. Ning, K. Peng, Z. Xia, and D. Wu, "Logistic-ELM: A novel fault diagnosis method for rolling bearings," *J. Brazilian Soc. Mech. Sci. Eng.*, vol. 44, no. 11, p. 553, Nov. 2022.

[57] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Conf. Neural. Netw. Conf.*, Budapest, Hungary, Jul. 2004, pp. 985–990.

[58] P. Trojovský and M. Dehghani, "Pelican optimization algorithm: A novel nature-inspired algorithm for engineering applications," *Sensors*, vol. 22, no. 3, p. 855, Jan. 2022.

[59] N. Helwig, E. Pignanelli, and A. Schütze, "Condition monitoring of a complex hydraulic system using multivariate statistics," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, Pisa, Italy, May 2015, pp. 210–215.

[60] M. M. Fréchet, "Sur quelques points du calcul fonctionnel," *Rendiconti del Circolo Matematico di Palermo*, vol. 22, no. 1, pp. 1–72, Dec. 1906.

[61] T. Eiter and M. Heikki, "Computing discrete Fréchet distance," Tech. Rep., 1994. [Online]. Available: https://www.researchgate.net/publication/228723178_Computing_Discrete_Frechet_Distance

[62] H. Yu, H. Li, Y. Li, and Y. Li, "A novel improved full vector spectrum algorithm and its application in multi-sensor data fusion for hydraulic pumps," *Measurement*, vol. 133, pp. 145–161, Feb. 2019.

CHAO CHEN received the B.Eng. degree in vehicle engineering from Xihua University, China, in 2021. He is currently pursuing the M.Eng. degree in mechanical engineering with Xinjiang University. His research interests include signal processing, machine learning, and hydraulic system fault diagnosis.

MUHETAER KELIMU received the M.Eng. degree in biomechanics from the University of Paris XII, France, in 1999. He is currently a Professor and a Graduate Tutor with the School of Mechanical Engineering, Xinjiang University. He has published over 40 articles in his research-related fields. His major research interests include hydraulic transmission and control, hydraulic system design optimization, biomechanics, and biomedical engineering (medical image processing and blood rheology).

BO YANG received the B.Eng., M.Eng., and Ph.D. degrees from Chang'an University, China, in 2011, 2013, and 2018, respectively. He is currently a Graduate Tutor with the School of Mechanical Engineering, Xinjiang University. He has published over ten articles in his research-related fields. His major research interests include hydraulic transmission and control.

• • •