


## Research Article

# Series Arc Fault Detection in a Low-Voltage Power System Based on CEEMDAN Decomposition and Sensitive IMF Selection

Guixia Fu,<sup>1</sup> Guizhen Chen,<sup>1</sup> Wei Wang,<sup>1</sup> Qinbing Wang,<sup>2</sup> and Guofeng Zou<sup>1</sup> 

<sup>1</sup>School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo, Shandong 255049, China

<sup>2</sup>Gaoqing County Power Supply Company, State Grid Shandong Electric Power Corporation, Gaoqing, Shandong 255000, China

Correspondence should be addressed to Guofeng Zou; [zgf841122@163.com](mailto:zgf841122@163.com)

Received 7 December 2021; Revised 8 January 2022; Accepted 17 January 2022; Published 31 January 2022

Academic Editor: Shao Haidong

Copyright © 2022 Guixia Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the series arc fault detection of a low-voltage distribution network, the features of the fault current signal are easily submerged and arc fault features are difficult to be represented, which greatly increases the difficulty of fault arc detection based on current signals. To solve these problems, a series arc fault detection method combining CEEMDAN decomposition and sensitive IMF selection is proposed. In this paper, the CEEMDAN algorithm is first applied to complete decomposition of the arc current in series faults. Then, 12 feature indicators of the arc current are defined and the frequency band division of the IMF component is realized according to the kurtosis index and energy feature which are more sensitive. The time window-based feature calculation method is proposed to obtain the local features of the time scale of each high-frequency IMF component. Accurate selection of sensitive IMF components is realized by comparing feature indexes such as the variance and root mean square value. Finally, for the current feature set, the second dimension reduction is realized by the subspace transformation algorithm and the series arc fault detection is realized based on the SVM. The actual experiments show that the optimal detection accuracy of the proposed method is 91.67% and the average accuracy of 10 crossvalidation experiments is 88.33%. It shows that the proposed sensitive IMF selection method can effectively capture the fault component signals in the current and the proposed fault feature description method has good representation and discrimination ability.

## 1. Introduction

In the low-voltage distribution network, arc fault is easily caused by line insulation damage and loose terminal. Local high temperature associated with arc fault can easily lead to electrical fire accident. Arc faults are divided into series arc, parallel arc, and ground arc. When a series arc fault occurs, it is equivalent to a series of time-varying resistor in the circuit, which will easily lead to the fault current similar to the load current. Sometimes, the waveform characteristics of the fault current are difficult to distinguish from the characteristics of the nonlinear load current [1–3]. It is the above factors that bring great difficulties to series arc detection and make it become a hot and difficult research area of arc detection [4, 5].

Arc fault detection methods can be divided into two categories: (1) arc detection based on the physical characteristics of arc light, arc sound, and temperature and (2) arc

detection based on time-frequency domain analysis of the arc voltage or current signal [6, 7]. Due to the randomness of the location of the arc fault, the first detection method is mostly used in electrical switchgear and its application in line arc fault detection is limited [8, 9]. The time-frequency analysis method based on current and voltage signals of monitoring points has become a research hotspot in arc fault detection. The current detection method can protect the downstream branch arc fault by installing the monitoring point in the upstream of the line. Therefore, compared with the voltage detection method [10], its applicability and flexibility is stronger and more favored by researchers [11, 12].

Fourier transform is widely used in early arc fault detection [13]. The essence of this method is to decompose the electrical signal into the superposition sum of multifrequency sine waves, which transform the time domain problem into the frequency domain for analysis. Fourier transform realizes the correlation between the time domain

and frequency domain of electrical signals, but signal analysis can only be implemented independently in the time domain or frequency domain, not simultaneously. Due to the sound adaptability, wavelet transform can realize multi-scale time-frequency analysis of the signal, which has been applied in the analysis of mechanical fault signals [14] and arc fault signals [15, 16]. Wang et al. [17] proposed a hybrid time and frequency analysis and fully connected neural network- (HTFNN-) based method to identify the series AC arc fault. Firstly, the samples are roughly divided into the resistive category, capacitive-inductive category, and switching category. Then, in each category, a separate fully connected neural network is used to identify the fault and the method achieves high identification accuracy. Chu et al. [18] proposed a novel high-frequency coupling sensor for extracting the features of low-voltage series arc faults. In the method, high-frequency feature signals under different loads are collected and transformed into two-dimensional feature gray images, which are used to train the convolutional neural network to realize series arc fault detection. Experiments show that the method is stable and universal.

Hilbert-Huang transform (HHT) [19] is a typical non-linear and nonstationary signal processing method, and its key step is empirical mode decomposition (EMD). EMD can decompose complex signals adaptively into several intrinsic mode functions (IMF), but this method has a serious mode mixing problem, which affects the performance of HHT [20]. Therefore, Wu and Huang [21] proposed the ensemble empirical mode decomposition (EEMD). By introducing Gaussian white noise with uniform frequency distribution into the signal to be decomposed, EEMD overcomes the problem of intermittent signal and avoids mode mixing. However, due to the interference of white noise, the reconstructed signal is easy to be distorted. Cheng et al. [22] proposed an enhanced periodic mode decomposition (EPMD) algorithm for accurate extraction of periodic pulses from rolling bearing composite fault signals, which effectively improved the accuracy of bearing fault diagnosis.

In 2011, Torres et al. [23] proposed the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), which further improves the accuracy and completeness of decomposition signals and effectively overcomes the problem of mode mixing. However, the CEEMDAN algorithm has not been applied in arc current signal analysis. CEEMDAN can achieve complete signal decomposition, but in arc fault detection, usually, only a few IMF components are sensitive to the arc fault and can reflect the characteristics of the fault arc. Most of the other IMF components are invalid for arc detection and even contain more interference information. Therefore, it is extremely difficult to extract fault identification features from all IMF components obtained by CEEMDAN decomposition and the interference features easily affect the accuracy of arc fault detection.

Based on the above analysis, in this paper, a series arc fault detection method based on CEEMDAN decomposition and sensitive IMF selection is proposed. The CEEMDAN algorithm is used to decompose the current signal and obtain the complete IMF components. Then, this paper proposes a strategy to automatically select the sensitive IMF

from the all IMF components. In this strategy, the kurtosis index and energy feature are taken as the basis to determine the fundamental frequency boundary and achieve frequency division. For high-frequency IMF, we design a local feature extraction method based on the time window. Using the number of fundamental frequency periods as the interval, IMF in the high-frequency band is divided into some non-overlapping time windows and the feature indexes of the signal are calculated in each time window. The sensitive IMF component with the strongest discriminability is selected adaptively based on the feature indexes such as variance and root-mean-square amplitude. After sensitive IMF selection, the local features of the best IMF component are used as the feature description of the current signal, which can be used to construct the current feature database. Finally, the subspace transformation algorithm is used to implement secondary dimension reduction for current features and the support vector machine (SVM) is used to the series arc fault detection. Experimental results show that the combination of CEEMDAN decomposition and sensitive IMF selection strategy, as well as the local feature construction method based on the time window, can effectively capture the discriminant features of the series arc, which realize the reliable detection of the arc fault.

The main highlights of the proposed method are generalized as follows:

- (1) To obtain complete decomposition results of fault current signals, the CEEMDAN decomposition algorithm is first applied to current signal decomposition
- (2) To extract the strongest discriminative IMF component, a method of frequency division and an accurate selection method of sensitive IMF are proposed
- (3) To overcome the difficulty in fault feature representation, a local feature calculation method based on the time window is proposed and 12 feature indexes are defined to express fault features
- (4) To better improve the effectiveness of series arc fault features and fault detection accuracy, this paper adopts subspace transformation for feature compression and SVM for fault detection

The remainder of this paper is structured as follows: in Section 2, the CEEMDAN algorithm is described. The feature calculation methods of the current signal are illustrated in Section 3. In Section 4, the selection method of the sensitive IMF component and the series arc fault detection method are proposed. Detailed experiments and analyses are performed in Section 5. In Section 6, the conclusions are drawn.

## 2. CEEMDAN Algorithm

Mode mixing refers to the phenomenon that a single IMF component contains multiple components with different frequencies or the same frequency component is decomposed into different IMF components. Mode mixing is usually

caused by the intermittency of the signal. Therefore, the EEMD algorithm introduces Gaussian white noise into the signal to be decomposed, which makes the signal become continuous at different scales and alleviates the mode mixing problem. However, the EEMD algorithm cannot completely eliminate the introduced noise interference, which makes the reconstructed signal distortion. In each stage of EMD decomposition of signals, the CEEMDAN algorithm adaptively adjusts the noise coefficient to generate Gaussian noises with different SNR introduced into the signals to be decomposed, which can avoid mode mixing and eliminate the interference of false information. The algorithm steps are as follows:

- (1) Gaussian white noise  $n_i(t)$  is added to the original signal  $x(t)$ . The signal with added noise is  $x(t) + \gamma_0 n_i(t)$ , where  $\gamma_0$  is the noise coefficient. EMD is used to perform  $I$  decomposition for the signal with noise. The first IMF component  $\text{IMF}_1(t)$  and the corresponding residual component  $r_1(t)$  of CEEMDAN are obtained through integration averaging:

$$\begin{aligned} \text{IMF}_1(t) &= \frac{1}{I} \sum_{i=1}^I \text{IMF}_{i1}(t), \\ r_1(t) &= x(t) - \text{IMF}_1(t), \end{aligned} \quad (1)$$

where  $I$  is the number of decomposition

- (2) Assume that  $\text{EMD}_j(\cdot)$  is the  $j$ th mode function of EMD decomposition. Decompose the signal  $r_1(t) + \gamma_1 \cdot \text{EMD}_1[n_i(t)]$  to obtain the second IMF component of CEEMDAN:

$$\text{IMF}_2(t) = \frac{1}{I} \sum_{i=1}^I \text{EMD}_1(r_1(t) + \gamma_1 \cdot \text{EMD}_1[n_i(t)]) \quad (2)$$

- (3) Calculate the  $k$ -order residual component:

$$r_k(t) = r_{k-1}(t) - \text{IMF}_k(t), \quad (3)$$

where  $\text{IMF}_k(t)$  is the  $k$ th IMF component. EMD decomposition is performed for the  $k$ th signal  $r_k(t) + \gamma_k \text{EMD}_k(n_i(t))$  until the first IMF component is obtained. On this basis, the  $(k+1)$ th IMF component of CEEMDAN is calculated:

$$\text{IMF}_{k+1}(t) = \frac{1}{I} \sum_{i=1}^I \text{EMD}_1(r_k(t) + \gamma_k \cdot \text{EMD}_k[n_i(t)]) \quad (4)$$

- (4) The above calculation steps are repeated until the residual components can no longer be decomposed

and all  $K$  IMF components of CEEMDAN are obtained; the remaining residual  $R(t)$  is

$$R(t) = x(t) - \sum_{k=1}^K \text{IMF}_k(t). \quad (5)$$

Therefore, after decomposition, the initial signal can be expressed as

$$x(t) = \sum_{k=1}^K \text{IMF}_k(t) + R(t). \quad (6)$$

The CEEMDAN method can realize complete reconstruction of the original signal based on noise-assisted analysis. Gaussian noises with different SNR are adjusted adaptively by noise coefficient and introduced into the signal to be decomposed, which improves the decomposition effect of EMD effectively

### 3. Feature Calculation of the Current Signal

Learn from the feature calculation method commonly used in mechanical fault diagnosis [24, 25], for each IMF component of the current signal; this paper defines 9 statistical feature indicators, as well as the energy feature, entropy feature, and energy entropy feature to form a 12-dimensional current feature vector. Assume that  $K$  IMF components are obtained after CEEMDAN decomposition and each component sequence contains  $N$  sampling points, i.e.,  $x_1, x_2, \dots, x_i, \dots, x_N$ , where  $i$  represents the  $i$ th sampling point. The 12 feature indicators are defined as follows:

The mean value is

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (7)$$

The variance is

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2 \quad (8)$$

The root-mean-square value is

$$X_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i)^2} \quad (9)$$

The root amplitude is

$$X_r = \left( \frac{1}{N} \sum_{i=1}^N |x_i| \right)^2 \quad (10)$$

The average amplitude is

$$\bar{X}_a = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (11)$$

The peak is

$$X_p = \max(|x_i|) \quad (12)$$

The kurtosis is

$$K = \frac{1}{N-1} \sum_{i=1}^N \frac{(x_i - \bar{X})^4}{\sigma^4} \quad (13)$$

The kurtosis index is

$$K_f = \frac{K}{X_{\text{RMS}}^4} \quad (14)$$

The margin index is

$$L_f = \frac{X_p}{X_r} \quad (15)$$

The energy feature is

$$E = \sum_{i=1}^N (x_i)^2 \quad (16)$$

The entropy feature is

$$H = - \sum_{i=1}^N P(x_i) \lg P(x_i), \quad (17)$$

where  $P(\cdot)$  is the probability.

The energy entropy features are

$$H_{EN} = - \sum_{i=1}^N P_i \lg P_i, \quad (18)$$

where  $P_i = E_i / \sum_{i=1}^K E_i$  and  $E_i$  represents the energy of the  $i$ th IMF component.

After the above calculation, each IMF component containing  $N$  sampling points is transformed into a 12-dimensional eigenvector, which reflects the overall state of the signal sequence.

#### 4. Selection of the Sensitive IMF Component and Series Arc Fault Detection

**4.1. Fundamental Frequency Determination and Frequency Division.** The 12 feature indexes defined in this paper reflect the characteristics of IMF components at different frequencies, among which the kurtosis index and energy feature have strong sensitivity to current signal fluctuations. Therefore, based on these two indicators, this paper divides the

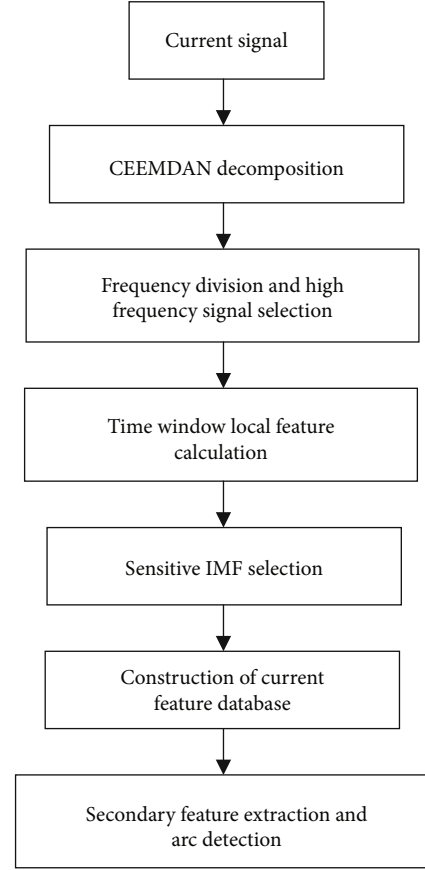
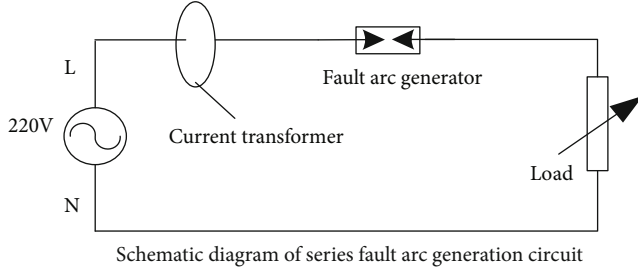


FIGURE 1: The overall flow chart of the proposed algorithm.

frequency bands for all IMF components. In our study, it is found that when the kurtosis index is the minimum value and the energy feature is the maximum value, the corresponding IMF component has the smallest fluctuation range and contains the most information. This component is the fundamental frequency component of the arc current signal, which is selected as the boundary of division. Therefore, the IMF component above this frequency is the high-frequency signal and the IMF component below this frequency is the low-frequency signal.

Since the current that causes the arc fault is usually a high-frequency signal, so after frequency division, feature index calculation and feature extraction only need to be carried out for the initially selected high-frequency signals. It can not only reduce the complexity of feature index calculation, feature extraction, and fault detection but also avoid the interference of low-frequency signals and improve the accuracy of arc detection.

**4.2. Local Feature Calculation and Selection of Sensitive IMF Components.** The global feature obtained from all  $N$  sampling points of IMF component data participating in index calculation can reflect the overall change of signals, which reflects the characteristics of signals of different frequency components in a macroscopic view. In the high-frequency band, the global features of some IMF components with similar frequencies are almost the same and the crucial features



(a) Schematic diagram of the series fault arc generation circuit

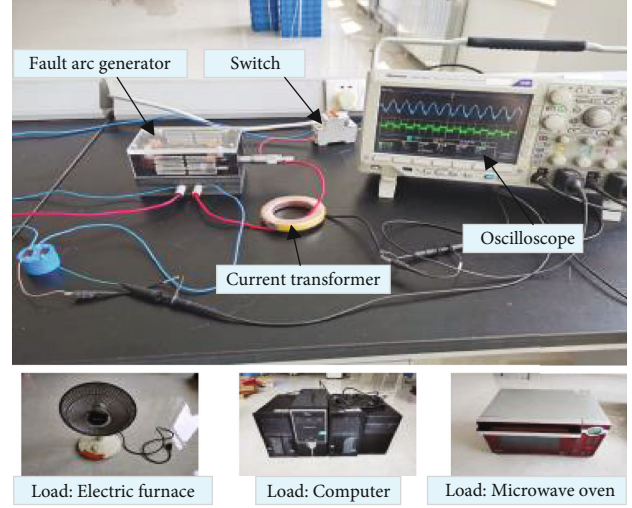


FIGURE 2: Series fault arc generation platform.

that distinguish fault current signals are easily covered up. Therefore, from the microviewpoint that the feature indexes need to fully reflect the local changes of the arc current signal, a calculation method of local feature indexes based on the time window is proposed. Then, the most sensitive IMF component is accurately selected from the high-frequency IMF components based on the time window local features. The algorithm steps are as follows:

- (1) Determine the period number  $M$  of the fundamental frequency IMF
- (2) According to period number  $M$ , the high-frequency IMF components are divided into  $M$  nonoverlapping time windows, namely,  $IMF_1^i, IMF_2^i, \dots, IMF_M^i$ , where  $i$  represents the  $i$ th IMF high-frequency component
- (3) For  $M$  time windows of the  $i$ th IMF component, 12 feature values are calculated according to the index definition formula, e.g., the feature of the  $j$ th window  $F_j^i = [\bar{X}, \sigma^2, X_{RMS}, X_r, \bar{X}_a, X_p, K, K_f, L_f, E, H, H_{EN}]$ . Therefore, the feature formed by the  $i$ th IMF component is  $M \times 12$  dimensions, i.e.,  $F^i = [F_1^i, F_2^i, \dots, F_M^i]$
- (4) The variance  $\sigma^2$ , root mean square value  $X_{RMS}$ , root amplitude  $X_r$ , average amplitude  $\bar{X}_a$ , energy feature  $E$ , and energy entropy feature  $H_{EN}$  of each time window in the IMF component are selected as the judgment indexes

If the fluctuation of the above judgment indexes in the front and back time windows (such as the  $j$ th and  $(j+1)$ th windows) is less than the threshold  $\theta$ , it indicates that the signals in each window of the IMF component are stable. That is, the IMF does not contain the current information causing the fault arc.

If the judgment indexes in the front and back time windows (such as the  $j$ th and  $(j+1)$ th windows) have obvious

TABLE 1: Current dataset under different loads.

Load type	Normal data (group)	Fault data (group)
Electric furnace (400 W)	10	30
Electric furnace (800 W)	10	30
Computer	10	30
Microwave oven	10	30

jump changes and the jump amplitude is greater than or equal to the threshold  $\theta$ , it indicates that the signal in the IMF component is not stable. That is, the IMF contains the current information causing the fault arc.

- (5) Through above calculation and comparison, the IMF component with the largest jump amplitude is selected from the IMF components whose jump value of the feature indexes in the front and back time windows exceeded the threshold  $\theta$ , which is the sensitive IMF component. So, this component has the strongest discrimination in the series arc fault detection

Local features with periodicity are captured by local feature calculation based on the time window, which can express the signal change of the time scale in more detail. It has obvious advantages and distinct degree for analyzing the high-frequency signal fluctuation caused by the arc fault. Based on this work, the selected sensitive IMF component can better describe the key features of signals and improve the accuracy of arc fault detection.

**4.3. Series Arc Fault Detection.** After the CEEMDAN decomposition, sensitive IMF selection, and local feature calculation, the feature information of the best IMF component in

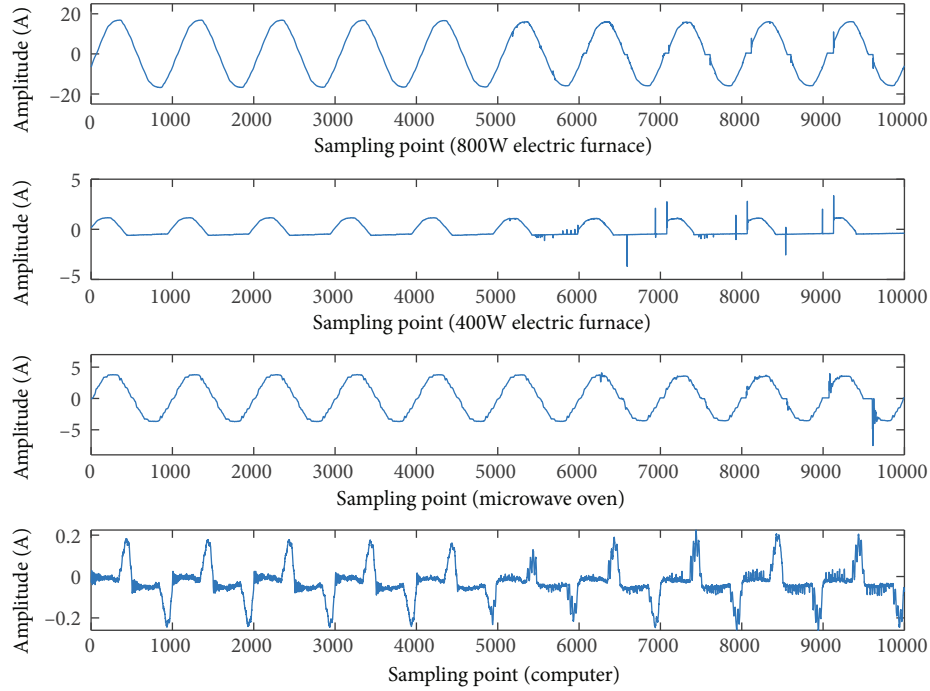


FIGURE 3: Fault arc current waveforms under different loads.

the arc current is obtained. However, these features still contain a lot of redundant information, even interference information. Therefore, to improve the discriminant ability of the features, it is very important to perform secondary feature extraction and dimension reduction for the preliminary arc current features.

Subspace mapping is to map feature vectors from the original space to the new space by mathematical transformation, and the feature vectors in the new space have lower dimension and more significant discriminant ability. Now, the classic subspace feature extraction methods in machine learning include PCA, LDA, ICA, KPCA, and KLDA [26]. In this work, the linear subspace mapping PCA and LDA algorithms and the nonlinear subspace mapping KPCA and KLDA algorithms are used for the secondary extraction and dimension reduction of arc current features.

The support vector machine (SVM) [27] is a typical binary classifier, which has unique advantages in solving small sample, nonlinear, and high dimension pattern recognition problems. It can mine the hidden decision information of sample features to the maximum extent, and it is widely used in the field of fault diagnosis. Therefore, in this paper, the SVM is selected as the classifier of series arc fault detection.

In summary, the overall flow of the proposed algorithm is shown in Figure 1.

## 5. Experiment and Analysis

**5.1. Current Signal Acquisition Platform.** The circuit principle of the series fault arc generation platform is shown in Figure 2. The input voltage is 220 V and LPCT with a bandwidth of 200 kHz is used as the current sensor. The arc fault

generator uses a copper electrode with a diameter of 10.0 mm as the moving contact and a graphite electrode with a diameter of 8.0 mm as the reference static contact.

**5.2. Current Dataset Construction.** Based on the fault arc generation circuit shown in Figure 2, the current signal is collected by a recording device. The sampling frequency is set to 50 kHz, and the sampling length is set to 20 ms. The current dataset constructed under different loads is shown in Table 1.

Considering the load in the actual circuit can be divided into three types: pure resistor load, resistor-inductance load, and nonlinear load. The 800 W electric furnace is selected as the pure resistor load. When the furnace works at 400 W, it works in the state of half-wave rectification. Therefore, 800 W and 400 W electric furnaces basically cover the current characteristics of resistor loads. A computer is a nonlinear load, and its current waveform can represent most of the switching power supply loads. A microwave oven belongs to the nonlinear load with more inductance, which can represent the current characteristics of most resistor-inductance loads. The above 4 kinds of loads are commonly used household appliances with higher frequency in life and can represent most loads, with typicality. In the experiment, 10 groups of normal current data and 30 groups of fault current data are collected under each load. Figure 3 shows the comparison of current waveforms before and after the arc fault, in which the first 5 periods are fault-free signals and the last 5 periods are fault signals.

**5.2.1. Training Dataset.** In the experiment, we randomly select 5 groups of normal current data and 20 groups of fault

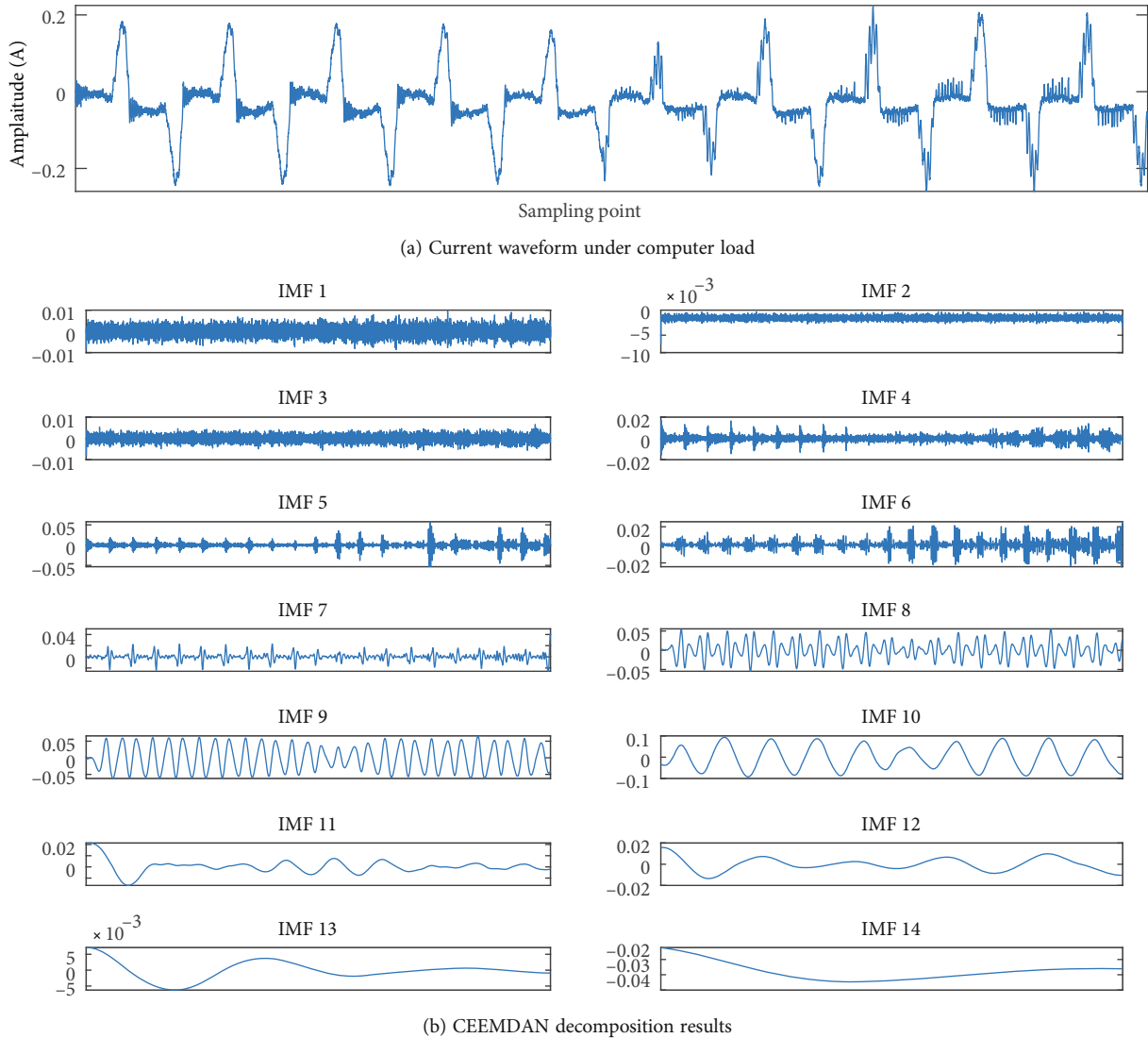


FIGURE 4: Fault arc current and decomposition results under computer load.

arc current data under 4 kinds of loads as training samples. Therefore, the training sample set size is 100.

**5.2.2. Test Dataset.** The remaining 5 groups of normal current and 10 groups of fault arc current under 4 kinds of loads are taken as the test sample set, and the total number of test samples is 60. In order to ensure reliable detection results, the average accuracy of the proposed algorithm is evaluated by 10 cross tests.

**5.3. Decomposition Experiment of the Current Signal.** In this work, the CEEMDAN algorithm is used to decompose the arc current signal and the results are shown in Figures 4 and 5. Figures 4(a) and 5(a) show the arc fault current waveform when the computer is used as the load and the microwave oven is used as the load. The first 5 periods are normal current waveform and the last 5 periods are fault current waveform.

Experiments show that the CEEMDAN algorithm adaptively decomposes the current signals under computer load

and microwave load into 14 components and realizes the detailed and complete decomposition of arc current signals in different frequency ranges, which can effectively overcome the mode mixing problem of the EMD decomposition algorithm. Observing the decomposition results of Figures 4(b) and 5(b), the component IMF10 contains 10 complete sinusoidal periodic signals whose frequency is consistent with the original current signal. Therefore, IMF10 corresponds to the fundamental frequency component. With IMF10 as the boundary, IMF1 to IMF9 are classified as high-frequency signals, and IMF11 to IMF14 are classified as low-frequency signals. The simple and clear division of each frequency band indicates that the signal decomposition is complete and no mode mixing occurs.

In addition, under computer load, high-frequency decomposition signals show that IMF4, IMF5, and IMF6 components have significant differences in the time scale, especially IMF5 and IMF6 which have good discrimination. Under microwave load, the waveforms from IMF4 to IMF7 are significantly different before and after arc occurrence,

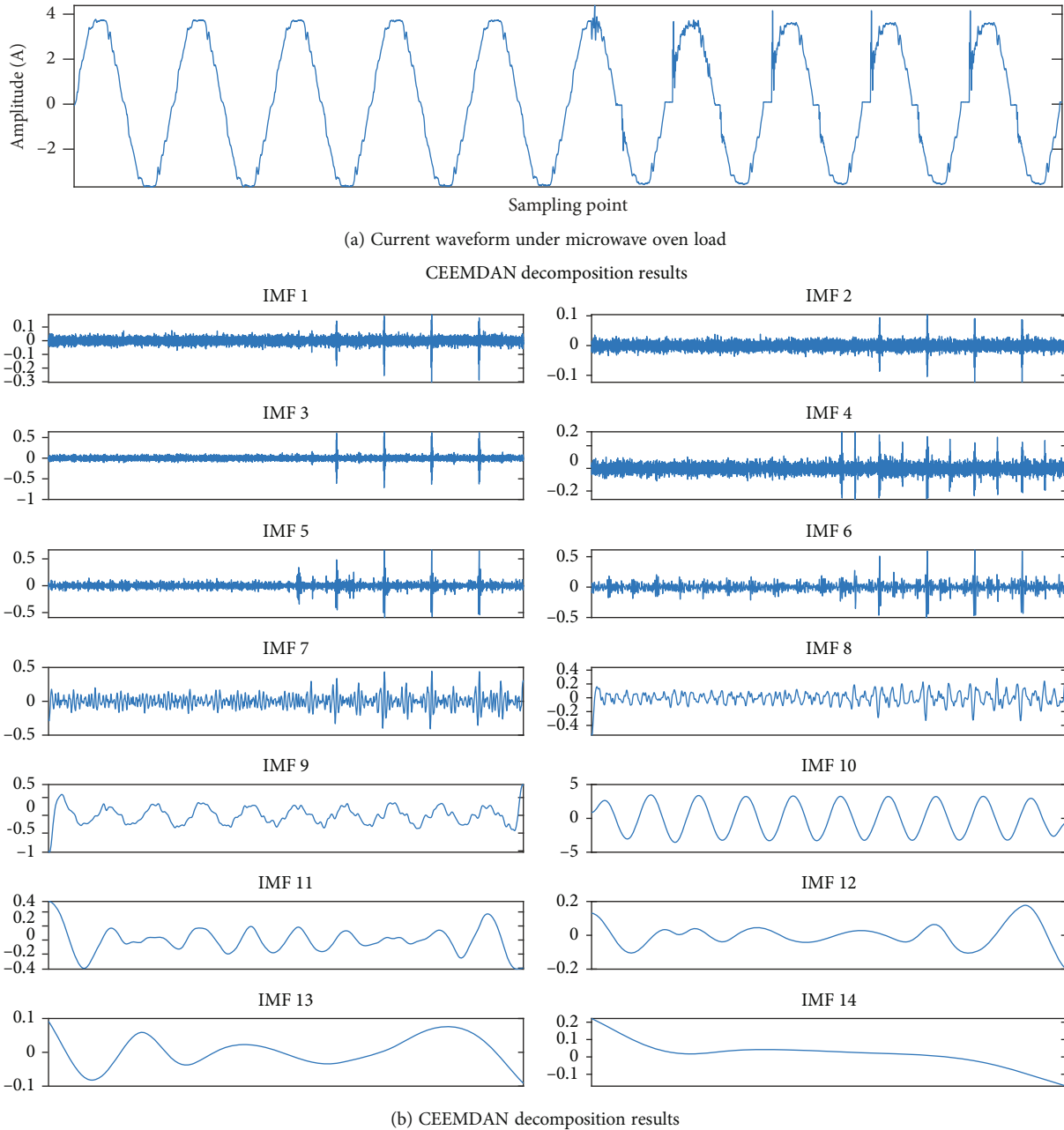


FIGURE 5: Fault arc current and decomposition results under microwave oven load.

especially IMF4 and IMF5 which have good differentiation which contributes greatly to arc fault detection.

In conclusion, the CEEMDAN decomposition strategy has a significant advantage in overcoming mode mixing and the discriminant of decomposed component signals is strong. Thus, it is feasible to select the CEEMDAN algorithm for arc current signal decomposition in this paper.

**5.4. Feature Calculation of the Arc Fault Current and Sensitive IMF Selection.** In order to make a better mathematical description of each IMF component and reduce the complexity of the fault detection operation, we adopt statistical feature indexes such as the mean and variance, as well as the energy feature, entropy feature, and energy entropy

feature, a total of 12 feature values as the feature description of each IMF component. Taking the decomposition results shown in Figure 5(b) as an example, 12 features of 14 IMF components are calculated and the feature matrix of  $14 \times 12$  dimension is constructed as shown in Table 2.

The data in Table 2 show that the 12 feature indexes reflect different characteristics of multiscale IMF components, among which the kurtosis index and energy feature have strong discrimination. Therefore, by comparing the kurtosis index and energy feature of each IMF component, it can be clearly judged that IMF10 is the fundamental frequency component. This component can be used as the boundary of frequency division. IMF1–IMF9 correspond to the high-frequency signals and IMF11–IMF14 correspond



TABLE 2: Calculation results of global eigenvalues of different IMF components.

IMF name	Mean	Variance	Root mean square	Square root amplitude	Average amplitude	Peak
IMF1	$-1.03e-4$	$4.75e-4$	0.0218	$2.48e-4$	0.0157	0.3038
IMF2	$-1.61e-4$	$1.41e-4$	0.0119	$7.45e-5$	0.0088	0.1235
IMF3	$2.37e-5$	0.0027	0.0517	0.0010	0.0316	0.7145
IMF4	$1.02e-4$	0.0016	0.0405	$8.39e-4$	0.0290	0.3235
IMF5	$1.37e-4$	0.0047	0.0682	0.0017	0.0408	0.6673
IMF6	$-3.57e-4$	0.0065	0.0803	0.0028	0.0528	0.6124
IMF7	$2.58e-4$	0.0090	0.0950	0.0050	0.0704	0.4450
IMF8	$-2.21e-4$	0.0084	0.0918	0.0046	0.0682	0.5403
IMF9	0.0012	0.0572	0.2391	0.0369	0.1920	1.0379
IMF10	<b>0.0129</b>	<b>5.1018</b>	<b>2.2588</b>	<b>4.1063</b>	<b>2.2064</b>	<b>3.5617</b>
IMF11	-0.0034	0.0300	0.1731	0.0165	0.1286	0.5487
IMF12	$9.22e-4$	0.0043	0.0653	0.0024	0.0488	0.2004
IMF13	0.0034	0.0017	0.0415	0.0012	0.0342	0.0917
IMF14	0.0160	0.0048	0.0711	0.0027	0.0520	0.2212
IMF name	Kurtosis	Kurtosis index	Margin index	Energy feature	Entropy feature	Energy entropy feature
IMF1	19.441	$8.6e+7$	$1.23e+3$	4.754	1.099	0.000841
IMF2	11.857	$5.99e+8$	$1.59e+3$	1.407	1.211	0.00028
IMF3	45.485	$6.37e+6$	713.93	26.72	0.731	0.00376
IMF4	10.377	$3.85e+6$	385.19	16.408	1.516	0.00263
IMF5	27.132	$1.25e+6$	401.82	46.523	1.046	0.00626
IMF6	13.898	$3.34e+5$	219.92	64.537	1.438	0.00828
IMF7	5.347	$6.57e+4$	89.717	90.227	2.199	0.01071
IMF8	6.987	$9.85e+4$	116.22	84.25	1.954	0.00986
IMF9	4.039	$1.243+3$	28.16	571.788	2.336	0.04906
IMF10	<b>1.539</b>	<b>0.0591</b>	<b>0.8674</b>	<b><math>5.102e+4</math></b>	<b>3.264</b>	<b>0.02517</b>
IMF11	3.99	$4.44e+3$	33.18	299.706	2.796	0.03512
IMF12	3.847	$2.11e+5$	84.24	42.66	2.691	0.00541
IMF13	2.323	$7.82e+5$	78.51	17.23	3.104	0.00095
IMF14	4.5012	$1.76e+5$	81.89	50.51	2.548	0.00336

to the low-frequency signals. This conclusion is identical with that obtained from the direct observation of Figure 5(b) in Section 5.3. It further proves that the proposed frequency division strategy is effective.

The current signals collected in our experiment all contain 10 periods. Therefore, according to the local feature calculation method of the time window, each IMF component in the high-frequency band can be divided into 10 windows. Then, 12 feature indexes such as the mean and variance are calculated for each window. Table 3 shows the feature calculation results of IMF5 divided into 10 time windows in Figure 5(b).

According to the data in Table 3, by comparing the indexes such as the variance, root mean square, root amplitude, average amplitude, energy feature, and energy entropy feature, we can see that the index values of the first 5 periods are significantly different from those of the last 5 periods. It shows that this IMF component contains the information

that causes the arc fault, so this IMF component can be used as one of the candidates of sensitive IMF components.

To accurately select the most sensitive IMF component, in our experiment, the decomposition results under microwave oven load are taken as an example. We calculate the local features of the time window of 9 high-frequency components including IMF1 to IMF9. After comprehensive comparison of all feature values, it is found that the features of IMF4 in the first 5 periods and the last 5 periods had the largest variation amplitude and the strongest distinguishing significance. The feature values of other IMF components in the first and last 5 periods also changed to some extent, but the amplitude of change is lower than IMF4, and the differentiation degree is weak. Therefore, the IMF4 component is selected as the most sensitive component according to the principle of significant differentiation and maximum variation amplitude. Similarly, in the decomposition results under the computer load in Figure 4(b), IMF6 is the most sensitive component.

TABLE 3: Calculation results of local eigenvalues of IMF5 components in the time window.

Window	Mean	Variance	Root mean square	Square root amplitude	Average amplitude	Peak
Window 1	$-5.96e-4$	0.0014	0.0372	0.00084	0.029	0.1433
Window 2	$-1.49e-4$	0.0017	0.0416	0.0011	0.033	0.1155
Window 3	$1.98e-4$	0.0015	0.0390	0.00097	0.0312	0.1363
Window 4	$3.05e-4$	0.0017	0.0416	0.0011	0.0335	0.1367
Window 5	<b><math>3.53e-4</math></b>	<b>0.0014</b>	<b>0.0381</b>	<b>0.00092</b>	<b>0.0303</b>	<b>0.1157</b>
Window 6	<b><math>9.08e-4</math></b>	<b>0.0046</b>	<b>0.0681</b>	<b>0.0021</b>	<b>0.0475</b>	<b>0.3939</b>
Window 7	$1.98e-4$	0.0079	0.0866	0.0031	0.0556	0.4836
Window 8	$-2.20e-6$	0.0090	0.0947	0.0025	0.0503	0.6613
Window 9	$5.45e-6$	0.0085	0.0924	0.0024	0.0493	0.6673
Window 10	$1.47e-4$	0.0087	0.0933	0.0025	0.0499	0.6508
Window	Kurtosis	Kurtosis index	Margin index	Energy feature	Entropy feature	Energy entropy feature
Window 1	3.127	$1.795e+6$	170.022	1.382	2.616	0.1045
Window 2	2.839	$0.951e+6$	136.19	1.728	2.951	0.1223
Window 3	2.989	$1.292e+6$	140.39	1.521	2.721	0.1118
Window 4	2.967	$0.991e+6$	121.95	1.7306	2.767	0.1224
Window 5	<b>2.900</b>	<b><math>1.38e+6</math></b>	<b>125.75</b>	<b>1.4501</b>	<b>2.792</b>	<b>0.1081</b>
Window 6	<b>9.207</b>	<b><math>3.821e+5</math></b>	<b>166.12</b>	<b>4.634</b>	<b>1.868</b>	<b>0.2597</b>
Window 7	11.549	$1.873e+5$	156.64	7.851	1.815	0.3003
Window 8	21.454	$2.66e+5$	261.49	8.974	1.277	0.3174
Window 9	22.105	$3.03e+5$	274.82	8.541	1.208	0.3120
Window 10	21.142	$1.783e+5$	261.49	8.709	1.195	0.3126

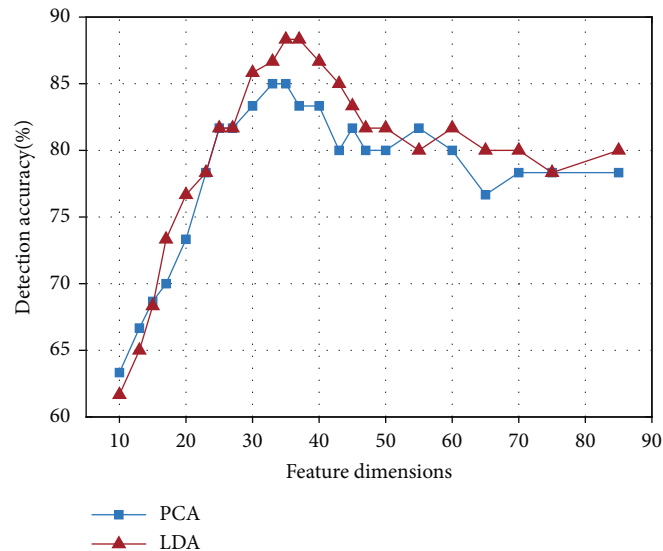
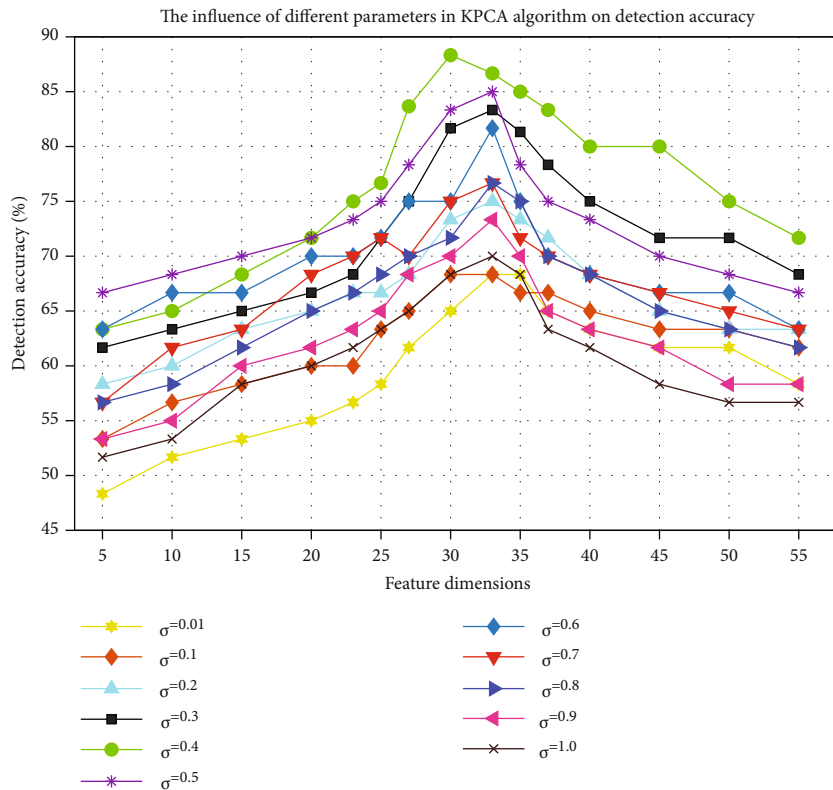


FIGURE 6: The influence of PCA and LDA feature dimensions on detection accuracy.

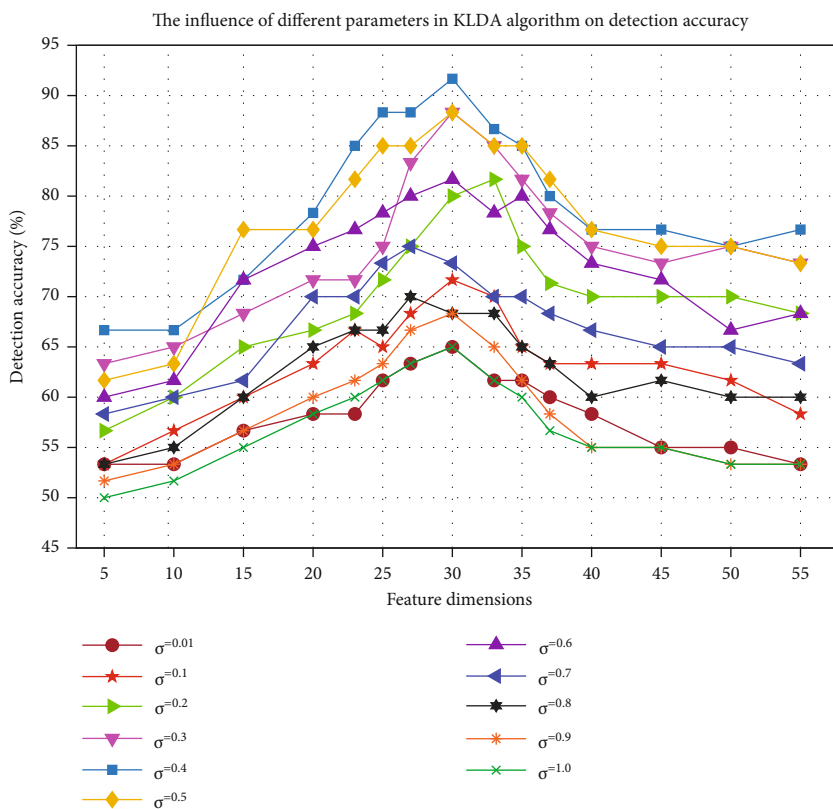
In summary, in this paper, we accurately select the most discriminating sensitive IMF component, and then, the fault arc current signal is characterized by 10 time windows and 12 feature indexes. Thus, the original current signal with each set of 10,000 sampling points is converted into a vector with 120 feature values. After local feature extraction of the

time window, 160 groups of normal current and fault arc current data under 4 load types finally form a current feature database with a scale of  $160 \times 120$ .

*5.5. Fault Arc Detection Experiment.* The  $160 \times 120$  dimension current feature database still contains some redundant



(a) The influence of different parameters in the KPCA algorithm on detection accuracy



(b) The influence of different parameters in the KLDA algorithm on detection accuracy

FIGURE 7: Detection accuracy of KPCA and KLDA under different parameters.

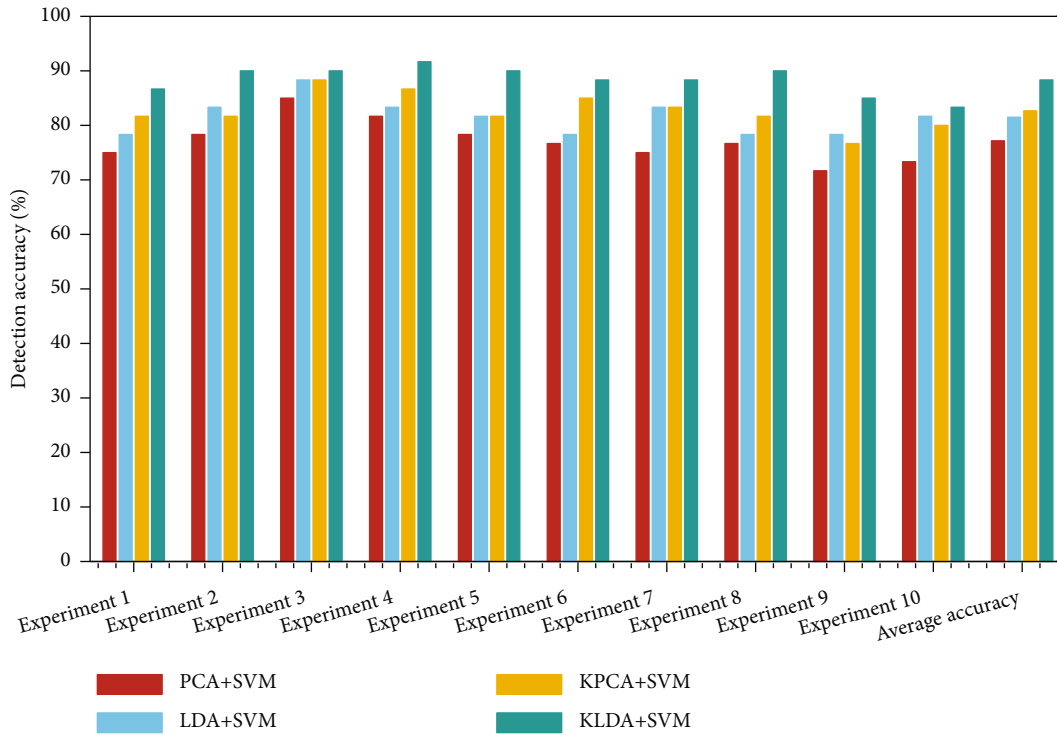


FIGURE 8: The detection accuracy of crossvalidation experiments.

information and interference information, which not only increases the complexity of arc fault detection but also affects the accuracy of detection. Therefore, to mine the features with higher discrimination, the subspace transformation method is used to perform secondary feature extraction for the current feature. The subspace transformation methods adopted in this paper include linear subspace methods PCA and LDA and nonlinear subspace methods KPCA and KLDA.

Series fault arc detection is a binary classification problem and the SVM is used as the fault detector in this paper. In the experiment, Gaussian kernel is used in the kernel function of the SVM and the optimal penalty factor of the SVM is determined to be 10 and the adjustable parameter of kernel function is determined to be 0.55 through the cross verification grid search method.

The main influencing factor of PCA and LDA is the retained feature dimension  $d$  during feature extraction. The influence factors of KPCA and KLDA include reserved feature dimension  $d$  and adjustable parameter  $\sigma$  in kernel function. Therefore, comparison experiments are performed to determine the parameter settings for optimal performance of each algorithm. Figure 6 shows the influence of PCA and LDA feature dimensions on detection accuracy in a cross experiment. Figure 7 shows the influence of KPCA and KLDA feature dimensions and the kernel function adjustable parameter on detection accuracy in a cross experiment.

As shown in Figure 6, the curves of arc fault detection accuracy have similar trends. When the dimension is lower than 25, the detection accuracy is generally low. At this stage, as the dimension gradually increases, the effective information contained in the feature gradually increases, so

the detection accuracy increases rapidly. When the dimension is 30–45, the detection accuracy reaches a high value but there is a small fluctuation affected by the validity of the feature. When the dimension exceeds 45, interference information will be introduced into retained features, so the detection accuracy will decrease slightly. At this stage, as the dimension continues to increase, the detection accuracy is generally stable, remaining at around 80%.

In addition, because LDA is a supervised algorithm, while PCA is an unsupervised feature extraction algorithm, LDA is easier for capturing strong discriminant features than PCA. Therefore, LDA shows more excellent detection performance. In this experiment, when the feature dimension is 34, the optimal detection accuracy of PCA-SVM is 85%. When the feature dimension is 35, the optimal detection accuracy of LDA-SVM is 88.33%. After 10 crossvalidation experiments, the average detection accuracy of PCA-SVM is 77.2%. The average detection accuracy of the LDA-SVM algorithm is 81.5%.

Figure 7 shows the influence of feature dimensions on detection accuracy when the kernel function adjustable parameter has different values. It is found that the effect of the feature-retained dimension on detection accuracy is similar to that of the linear subspace algorithm. The adjustable parameter of kernel function plays a key role in improving the detection accuracy. As shown in Figure 7, the optimal value of the adjustable parameter of the KPCA algorithm is 0.4 and the optimal feature-retained dimension is 30 and the highest detection accuracy is 88.33%. The optimal value range of the adjustable parameter of the KLDA algorithm is 0.3–0.4, and the optimal feature-retained dimension is 30,

and the highest detection accuracy reaches 91.67%. After comprehensive analysis, the kernel function adjustable parameter of the KLDA algorithm is set to 0.35. The optimal feature dimension of both algorithms is 30.

In order to prove the reliability and effectiveness of the proposed algorithm, 10 crossvalidation experiments are performed and the average detection accuracy of all experiments is calculated as the final performance evaluation. The results and average accuracy of 10 crossvalidation experiments are shown in Figure 8.

In general, the performance of the nonlinear subspace transformation algorithm is better than that of the linear algorithm. Although the detection performance of the LDA-SVM algorithm is better than the KPCA-SVM algorithm in the 1st, 9th, and 10th experiments, the overall average detection accuracy of KPCA-SVM is 82.7%, which is better than that of LDA-SVM 81.5%. It shows that the feature extraction ability of nonlinear subspace transformation is better than that of linear subspace transformation and the feature set contains more nonlinear information. In addition, the KLDA-SVM algorithm has the highest average detection accuracy of 88.33%. The advantage of the KLDA algorithm lies in the guidance of supervision information. Under the same condition of retaining 30-dimensional features, the KLDA algorithm can capture features with more significant discrimination and stronger classification ability. Therefore, compared with the KPCA-SVM algorithm, the detection accuracy of the KLDA-SVM improved by about 5%.

## 6. Conclusions

In order to realize arc detection of series faults accurately and efficiently, a detection algorithm based on CEEMDAN decomposition and sensitive IMF selection is proposed. In this paper, a series arc generation platform is built and the current data of four kinds of loads are collected. Based on the CEEMDAN algorithm, arc current decomposition is implemented and a frequency division strategy is proposed to realize high-frequency signal rough selection. Then, an accurate selection strategy for the sensitive IMF component is proposed, which eliminates the interference of invalid IMF components and reduces the complexity of fault detection. A local feature construction method based on the time window is proposed to realize local feature extraction of the sensitive IMF component and enhance the contrast and discrimination of arc current features. Subspace transformation is used to extract secondary feature and the reduce dimension, and the support vector machine is used to detect the series fault arc. The optimal average detection accuracy of the proposed algorithm is 88.33%, which proves the effectiveness of the proposed algorithm and provides an important reference for fault arc detection technology and device design.

## Data Availability

The dataset used in the experiment can be obtained by contacting the corresponding author.

## Conflicts of Interest

The authors declare no conflict of interest.

## Authors' Contributions

Fu Guixia proposed the original idea and wrote this paper; Chen Guizhen analyzed the data and corrected the translation; Wang Wei and Wang Qinbing collected the experimental data. Zou Guofeng participated in the reasonableness demonstration and manuscript modification.

## Acknowledgments

This research was funded by the Integration Funds of Shandong University of Technology and Zhangdian District (no. 118228), the National Natural Science Foundation of China (no. 61801272), and the Natural Science Foundation of Shandong Province of China (no. ZR2016FL14).

## References

- [1] G. G. Liu, S. J. Du SH, and X. Han, "Research on LV arc fault protection and its development trends," *Power System Technology*, vol. 1, pp. 321–329, 2017.
- [2] B. Basnet, H. Chun, and J. Bang, "An intelligent fault detection model for fault detection in photovoltaic systems," vol. 2020, pp. 1–11, 2020.
- [3] N. Qu, J. Wang, J. Liu, and Z. Wang, "An arc fault detection method based on multidictionary learning," *Mathematical Problems in Engineering*, vol. 2018, 8 pages, 2018.
- [4] Y. Kai, Z. Rencheng, Y. Jianhong, D. Jianhua, C. Shouhong, and T. Ran, "Series arc fault diagnostic method based on fractal dimension and support vector machine," *Transactions of China Electrotechnical Society*, vol. 31, no. 2, pp. 70–77, 2016.
- [5] T. J. Zhao, J. Meng, Y. Q. Song, X. Y. Xie, and M. C. Zhang, "Series arc detection and protection on the DC side of string-type PVs," *Power System Protection and Control*, vol. 48, no. 20, pp. 74–82, 2020.
- [6] Q. W. Lu, T. Wang, Z. R. Li, and C. Wang, "Detection method of series arcing fault based on wavelet transform and singular value decomposition," *Transactions of China Electrotechnical Society*, vol. 17, pp. 212–221, 2017.
- [7] N. Qu, J. Chen, J. Zuo, and J. Liu, "PSO-SOM neural network algorithm for series arc fault detection," *Adv. Math. Phys.*, vol. 2020, pp. 1–8, 2020.
- [8] J. G. Wang, W. Lin, Z. Wang, J. Li, W. He, and P. Wang, "Online detecting device for switchgear arc based on ultraviolet detection," *Power System Protection and Control*, vol. 39, no. 5, pp. 128–133, 2011.
- [9] B. R. Pan, C. Wei, X. Z. Gui, and L. P. Zhang, "Development and application of a portable arc protection tester," *Power System Protection and Control*, vol. 48, no. 13, pp. 149–155, 2020.
- [10] J. C. Kim, D. O. Neacșu, B. Lehman, and R. Ball, "Series AC arc fault detection using only voltage waveforms," in *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*, pp. 2385–2389, Anaheim, CA, USA, 2019.
- [11] Y. Wang, L. Hou, K. C. Paul, Y. Ban, C. Chen, and T. Zhao, "ArcNet: series AC arc fault detection based on raw current and convolutional neural network," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 77–86, 2022.

- [12] J. Jiang, W. Li, Z. Wen, Y. Bie, H. Schwarz, and C. Zhang, "Series arc fault detection based on random forest and deep neural network," *IEEE Sensors Journal*, vol. 21, no. 15, pp. 17171–17179, 2021.
- [13] N. Hadziefendic, M. Kostic, and Z. Radakovic, "Detection of series arcing in low-voltage electrical installations," *European Transactions on Electrical Power*, vol. 19, no. 3, pp. 423–432, 2009.
- [14] A. G. Kavaz and B. Barutcu, "Fault detection of wind turbine sensors using artificial neural networks," *Journal of Sensors*, vol. 2018, 11 pages, 2018.
- [15] G. Artale, A. Cataliotti, V. C. Nuccio, D. Di Cara, G. Tinè, and G. Privitera, "A set of indicators for arc faults detection based on low frequency harmonic analysis," in *2016 IEEE international instrumentation and measurement technology conference*, pp. 1–6, Taipei, 2016.
- [16] Y. Kun, M. Shang, W. Wei, D. Peng, and X. Zebao, "Simulation and research of low voltage series arc discharge detection and fault protection based on Cassie model and wavelet analysis," *Electrical & Energy Management Technology*, vol. 18, pp. 48–52, 2019.
- [17] Y. Wang, F. Zhang, X. Zhang, and S. Zhang, "Series AC arc fault detection method based on hybrid time and frequency analysis and fully connected neural network," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6210–6219, 2019.
- [18] R. Chu, P. Schweitzer, and R. Zhang, "Series AC arc fault detection method based on high-frequency coupling sensor and convolution neural network," *Sensors*, vol. 20, no. 17, pp. 4910–4919, 2020.
- [19] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society*, vol. 454, no. 1971, pp. 903–995, 1998.
- [20] G. Rilling, N. P. Flandrin, and P. Goncalves, "On empirical mode decomposition and its algorithms," in *IEEE-EURASIP workshop on nonlinear signal and image processing, NSIP-03, Grado (I)*, pp. 9–11, Italy, 2003.
- [21] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.
- [22] J. Cheng, H. S. Yang, H. Shao, H. Pan, J. Zheng, and J. Cheng, "Enhanced periodic mode decomposition and its application to composite fault diagnosis of rolling bearings," *ISA Transactions*, vol. 113, 2021.
- [23] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4144–4147, Prague, Czech Republic, 2011.
- [24] G. Zhiqiang and L. Yuedie, "Fault diagnosis of rolling bearing based on improved LMD and comprehensive characteristic index," *Journal of Hefei University of Technology: Natural Science*, vol. 44, no. 2, pp. 145–150, 2021.
- [25] J. Lingli, C. Liman, T. Hongchuang, and L. Xuejun, "Fault diagnosis of spiral bevel gear based on CEEMDAN permutation entropy and SVM," *Journal of Vibration, Measurement & Diagnosis*, vol. 41, no. 1, pp. 33–40, 2021.
- [26] K. C. Kempfert, Y. Wang, C. Chen, and S. W. K. Wong, "A comparison study on nonlinear dimension reduction methods with kernel variations: visualization, optimization and classification," *Intelligent Data Analysis*, vol. 24, no. 2, pp. 267–290, 2020.
- [27] W. Miao, Q. Xu, K. H. Lam, P. W. T. Pong, and H. V. Poor, "DC arc-fault detection based on empirical mode decomposition of arc signatures and support vector machine," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 7024–7033, 2021.