IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 15, 2018, accepted November 9, 2018, date of publication July 30, 2019, date of current version January 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2018.2888950

A Naive-Bayes-Based Fault Diagnosis Approach for Analog Circuit by Using Image-Oriented Feature Extraction and Selection Technique

WEI HE [1], YIGANG HE [1,2], BING LI [1], AND CHAOLONG ZHANG [1,3]

[1]School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China
[2]School of Electrical Engineering, Wuhan University, Wuhan 430072, China
[3]School of Physics and Electronic Engineering, Anqing Normal University, Anqing 246011, China

Corresponding author: Yigang He (yghe@hfut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51577046 and Grant 51777050, in part by the State Key Program of National Natural Science Foundation of China under Grant 51637004, in part by the National Key Research and Development Plan "Important Scientific Instruments and Equipment Development" under Grant 2016YFF0102200, in part by the Equipment Research Project in Advance under Grant 41402040301, in part by the Natural Science Foundation of Hunan Province under Grant 2017JJ2080, and in part by the Basic Research Service Fee Project of Central University under Grant JDK16TD01.

ABSTRACT Analog circuit is one of the most commonly used components in industrial equipment, and circuit failure may lead to significant causalities and even enormous financial losses. To address this problem, a novel scheme based on the wavelet spectrum features, feature selection, and Naive Bayes classifier is presented for the fault location of an analog system in this paper. The scheme mainly consists of three stages. First, the cross-wavelet transform (XWT) method is utilized to obtain the time–frequency representations of the raw signals of analog circuits. Second, the local optimal-oriented pattern is applied to all the XWT spectrum images to generate the original high-dimensional feature set. Then, an integration feature selection approach via joint Hilbert–Schmidt independence criterion and kernel Fisher linear discriminant analysis is proposed and utilized to obtain low-dimensional fault features, which are uncorrelated and distinctive. Finally, the training samples set is imported into the Naive Bayes classifier, and the fault diagnosis results can be drawn through inputting the testing samples set into the trained Naive Bayes classifier. The simulation results on two typical circuits have demonstrated that the proposed method is a promising means to detect and classify most analog circuit faults, achieving a better diagnosis accuracy than that of the other published works.

INDEX TERMS Analog circuit, fault diagnosis, cross-wavelet transform (XWT), local optimal oriented pattern (LOOP), Hilbert Schmidt independence criterion (HSIC), kernel Fisher linear discriminant analysis (KLDA), Naive Bayes.

I. INTRODUCTION

Analog circuit, a common component in most industrial system, plays a critical role in the reliable operation of the entire system. An unexpected failure of analog circuit may cause sudden breakdown of the whole equipment, bringing about enormous financial losses or even casualties [1], [2]. Therefore, fault diagnosis of analog circuit is of utmost significance for security and reliability in industrial manufacturing [3], [4].

Fault diagnosis method mainly involves two aspects: feature extraction and classification. It is well known that feature extraction is a vital stage that determines classification performance [1]. Considerable scholars have devoted themselves on feature extraction, among which time-domain analysis methods have been widely utilized. For instance, Yuan *et al.* [5] employed kurtosis and entropy of the sampled signal to locate circuit faults. Chang *et al.* [6] utilized the similarity-based metric technique to detect anomalies of light-emitting diode. Shahbazi *et al.* used the slopes of inductor currents as the fault features to detect the defect in dc-dc converters. Long *et al.* [7] proposed to employ six statistical characteristics as features for diagnosis of analog circuit. However, the features in time domain are manually chosen which heavily depend on the prior knowledge of field experts.

Frequency domain methods are utilized to detect the variation via spectral analysis, which are widely applied in many works. For example, via fast Fourier transform (FFT) [8],

The associate editor coordinating the review of this manuscript and approving it for publication was Jun Shi.

 This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

sweep frequency response analysis [9], short time Fourier transform (STFT) [10], [11] of output signal, the features of defect circuit were fetched, and they all achieved a better performance. Nevertheless, frequency-domain-based approaches discard the structure information embedded in raw data. Thus, it is difficult to find the most discriminative features while dealing with complex circuit.

Moreover, whether use a pure time-domain or frequency-domain method, it is incapable of providing sufficient information for fault detection. Time-frequency analysis (TFA) can not only reveal transient components distributed in all frequency bands, but also identify time-variant features [12]. Recently, various time-frequency joint analyses, including short-time Fourier transform, wavelet transform (WT) [13], S-transform (ST) [14], [15], Hilbert-Huang transform (HHT) [16] and Gabor analysis [17], have been proposed for fault diagnosis in electronic system. However, resultant matrices, time-frequency representations (TFR) significantly increase computation complexity due to its multi-dimensional characteristic. Moreover, TFR usually contains a large amount of redundant information. Therefore, it is necessary to conduct dimensionality reduction on TFR, and the information should be fetched as much as possible. Traditional matrix-based method always applied suitable projections to map the matrices in a feature subspace capturing high-discriminative fault information. Various projection approaches are commonly utilized in obtaining features of matrices such as two direction two dimensional linear discriminative analysis (TD2DLDA) [18], two-dimensional non-negative matrix factorization (2DNMF) [19] and two-dimensional principal component analysis (2DPCA) [20], and they all achieved desirable results.

In general, the classification method based on TRF can be considered as a kind of image recognition issues essentially. Compared with the methods based on projection, the image classification method based on texture features can ont only capture the global statistics of local features, but weaken the inference of the horizontal distribution variation [21]. The well-known local binary pattern (LBP) has been shown to be effective encoders of repeated local patterns for robust discrimination in several visual recognition tasks. However, a major limitation of LBP is the arbitrary sequence of binarization weights, that may result in poor performance which is susceptible to rotation variance. Consequently, a new descriptor, local optimal oriented pattern (LOOP) is proposed to overcome the drawback of LBP, via encapsulating rotation invariance into the main formulation of LBP [22], [23]. As a result, the processing time complexity and classification rate can be reduced.

After feature extraction, a classifier construction is another concerning issue. Artificial neural network (ANN) and Support vector machine are widely implemented in fault diagnosis. However, the ANN technique has disadvantages of falling into local extremum, slow convergence speed and overfitting [24]. Moreover, the SVM method has a limitation of high computation cost in hyper-parameters optimization procedure [25]. Owing to its simplicity structure and effective learning ability, Naive Bayes (NB) has been successfully applied in many fields, such as EEG recognition [26], text classification [27] and Software defect prediction [28]. Moreover, NB is capable of providing a probabilistic interpretation of its outputs which based on the assumptions that all properties are mutually independent and their weights are equally important.

Nevertheless, the original feature vectors is with high dimensionality and it not only contains relevant features but also redundant information. If all the features are imported into classifier directly without further processing, it may result in complicated classification model with poor performance. Therefore, it is of utmost significance to conduct an intelligent feature selection scheme to automatically determine representative features which obviously identify the circuit condition rather than applying whole features.

Generally, approaches for feature selection can be broadly divided into two categories: wrapper and filter [29]. The wrapper methods tend to achieve excellent performance because they chose the most distinctive features with the evaluation of classifier. Nevertheless, the wrapper methods is time consuming, which is not suitable for fault diagnosis [30]. Compared with wrapper methods, filter methods converge much faster and find the optimal features using an evaluation function, which is conducive for scholars to understanding the importance of the considered features [31]. However, the filter method has some drawbacks, which are described as below:1)The relationship between sets of features and the class labels is not jointly considered, as they usually use pairwise feature-label measurements, and 2) they do assume linear dependence between features [32]. To cope with these issues, a novel filter approach called Hilbert-Schmidt independence criterion (HSIC) is introduced to evaluate statistical dependence between features and class labels. Unlike most of the feature selection methods, HSIC criterion have a good ability of capturing high order relations between features and it is easy to conduct, providing interpretable results [33], [34]. Thus, the raw features with high dimensionality is pre-selected to a feature subset with use of HSIC. Furthermore, kernel fisher linear discriminant analysis (KLDA) [35] is implemented to further obtain the lower-dimension feature vectors.

The rest of the paper is organized as follows. Section II describes the procedure of spectrum image-oriented feature extraction based on cross-wavelet transform (XWT) and local optimal oriented pattern (LOOP). In Section III, a two-stage feature selection approach via joint Hilbert-Schmidt independence criterion (HSIC) and Kernel fisher linear discriminant analysis (KLDA) is given. Section IV describes the theory of Navie Bayes classifier. Section V presents the framework of the proposed intelligent fault diagnosis approach. Section VI gives the simulation results on two experimental circuits. Finally, Section VII presents the conclusion of the paper.

II. IMAGE-ORIENTED FEATURE EXTRACTION

A. CROSS-WAVELET TRANSFORM (XWT)

As mentioned above, Cross-wavelet transform can be considered as an extension of wavelet-based analysis, and it can be used to find a high common power between two time-domain signals in time-scale plane. Via the details about mathematical background described in the works [36], [37], a very brief overview of cross-wavelet transform is given as follows.

Assuming $a(t)$ denotes reference signal and $b(t)$ represents fault signal. The cross-wavelet transform is described below,

$$XWT^{ab}(s,\tau) = \frac{1}{p_\psi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W^a(u,v)W^{b*}(\frac{u}{s},\frac{v-\tau}{s})\frac{dudv}{u^2} \quad (1)$$

where $W^a(s,\tau)$ and $W^b(s,\tau)$ are the wavelet transform of $a(t)$ and $b(t)$ corresponding to a mother wavelet of $\psi(t)$. s and τ are usual 'dilation' and 'translation' parameters. p_ψ is a constant and given as $p_\psi = \int_{-\infty}^{+\infty}\frac{|\psi(\omega)|^2}{|\omega|}d\omega < \infty$

Hence, we can plot the cross-wavelet spectrum via using the magnitude XWT^{ab} and phase $\phi = tan^{-1}\frac{\Im\{W^{ab}\}}{\Re\{W^{ab}\}}$.

B. LOCAL BINARY PATTERN (LBP)

Assuming i_c denotes the intensity of an picture I at pixel (x_c, y_c), $i_n (n = 0, \cdots, 7)$ represents the picture element intensity in the 3×3 neighborhood of (x_c, y_c) excluding the center pixel i_c.

Then, we can obtain the LBP formulation for the pixel (x_c, y_c), which can be defined as below:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c)2^n \quad (2)$$

where

$$s(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & otherwise \end{cases} \quad (3)$$

However, LBP has a limitation that the sequence of binarization weights is arbitrary. According to the selected starting pixel of sequence of binary weights $(2^n, n = 0, \cdots, 7)$, the 8 elements of the output 3×3 area are determined subsequent weightage n order. There is not explicit logic behind the proper weight allocation. And, the obtained results is sensitive to rotation variance. As a result, a different binary sequence can be generated when the same mode rotated between pictures of the same pattern or even within different sections of the same picture.

C. LOCAL OPTIMAL ORIENTED PATTERN (LOOP)

To overcomes the limitations of LBP, LOOP is proposed to preserve the strengths of the sequence.

The 8 Kirsch masks are oriented in the direction of these 8 neighboring pixels $i_n \quad (n = 1, 2, \cdots, 7)$, thus giving a measurement of the strength of intensity variation in those directions, respectively.

This leads us to propose the incorporation of this information by assigning the binarization weight to each neighboring pixel corresponding to the strength of Kirsch output in the direction of that pixel. The underlying rationale behind this approach is that the Kirsch mask output in particular direction provides an indication of the probability of occurrence of an edge in that direction. Since the LBP indicates the intensity variation over the neighboring pixels in the same directions, the value of the Kirsch output is employed to assign the decimal-to-binary weightage.

As mentioned above, the 8 responses of the Kirsch masks are m_n corresponding to pixels with intensity i_n, $n = 0, \cdots, 7$. Each of these pixels are assigned an exponential w_n (a digit between 0 and 7) according to the rank of the magnitude of m_n among the 8 Kirsch mask outputs.

Then the LOOP value for the pixel (x_c, y_c) is explained as

$$LOOP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c)2^{w_n} \quad (4)$$

where

$$s(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & otherwise \end{cases} \quad (5)$$

Thus the LOOP descriptor encodes rotation invariance into the main formulation.

III. FEATURE SELECTION

A. HILBERT SCHMIDT INDEPENDENCE CRITERION (HSIC)

To estimate the independence between two sets of random variables, Hilbert Schmidt Independence Criterion (HSIC) is proposed. Supposing X and Z denote the two sequences of random variables, from which the sample datas (x, z) can be obtained from the probability density function X and Z. Let the function, which projects the $x \in X$ into the feature space F, be $\phi(x) \in F$. Hence, the inner product between the features can be obtained by using a kernel function $k(x, x') =< \phi(x), \phi(x') >$, and F denotes the associated reproducing kernel Hilbert space (RKHS). By a similar way, assuming G stands for the RKHS on Z with kernel $l(.,.)$ and the projection function $\psi(z)$. Then the cross-covariance operator $C_{xz} : G \rightarrow F$ between these two projection functions can be described as follows:

$$\begin{aligned} C_{xz} &= E_{xz}[(\phi(x) - \mu_x) \otimes (\psi(z) - \mu_z)] \\ &\Rightarrow E_{xz}[\phi(x) \otimes \psi(z)] - \mu_x \otimes \mu_z \end{aligned} \quad (6)$$

Here $\otimes$ denotes a tensor product. The HSIC is determined as the squared Hilbert Schmidt norm of (6). In view of kernel [38], the empirical estimate of the HSIC can be described as below:

$$HSIC(Z, F, G) = (m-1)^{-2}tr(K\ C\ L\ C) \quad (7)$$

where m represents the number of samples, $C, K, L \in \Re^{m \times M}$, $K_{ij} = k(x_i, x_j)$,$L_{ij} = l(z_i, z_j)$,$C_{ij} = \delta_{ij} - m^{-1}(\delta_{ij} = 1\ if\ i = j$,zero otherwise) denotes the centering matrix, and tr stands for the trace operator. By using Eq. 7, the independence between the two samples set in the Hilbert space can be given.

B. KERNEL FISHER LINEAR DISCRIMINANT ANALYSIS (KLDA)

Assuming Φ represents a nonlinear mapping to a kernel space F. Let S_W^Φ and S_B^Φ be the within-class and between-class scatter matrixes, respectively. $\{(x, y), x \in R^d,$ *y is the class label*$\}$ stands for a sample. n_i represents the samples number in the ith class and C denotes the number of classes; m_i^Φ is the mean vector for the ith class and m^Φ represents the mean vector for all the classes in F, where Φ separates the variables in F from those in R^d. Hence, S_B^Φ and S_W^Φ are defined as

$$S_B^\Phi = \sum_{i=1}^{c} n_i(m_i^\Phi - m^\Phi)(m_i^\Phi - m^\Phi)^T \tag{8}$$

$$S_W^\Phi = \sum_{i=1}^{c}\sum_{j=1}^{n_i}(\Phi(x_{ij}) - m_i^\Phi)(\Phi(x_{ij}) - m_i^\Phi)^T \tag{9}$$

Here, $m_i^\Phi = 1/n_i \sum_{j=1}^{n_i} \Phi(x_j^i)$. Then, in order to find the linear discriminant in F, we maximize

$$J(W) = \frac{W^T S_B^\Phi W}{W^T S_W^\Phi W} \tag{10}$$

Owing to the solution w must lie in the span of all training samples, w is defined as a linear combination of training samples Φx_i

$$w = \sum_{i=1}^{l} \alpha_i \Phi(x_i) = \Phi\alpha \tag{11}$$

According to Eq.11 into Eq. 10, the following equation can be given

$$J(\alpha) = \frac{\alpha^T \Phi^T S_B^\alpha \Phi\alpha}{\alpha^T \Phi^T S_W^\alpha \Phi\alpha} = \frac{\alpha^T M\alpha}{\alpha^T N\alpha} \tag{12}$$

In order to maximize the function $J(\alpha)$, the eigenvector problem of $N^{-1}M$ should be addressed. The optimal function can be explained as below

$$\alpha^* = \underset{\alpha}{arg\ max} \frac{\alpha^T M\alpha}{\alpha^T N\alpha} \tag{13}$$

where, α^* is a leading eigenvector with the largest eigenvalue. Hence, the projection of an input vector x onto w can be calculated by the following formulation.

$$x_p = (w \cdot \Phi(x)) = \sum_{i=1}^{l} \alpha_i^* k(x_i, x) \tag{14}$$

where x_p is the projection, $k(\cdot, \cdot)$ stands for the kernel function.

IV. NAIVE BAYES CLASSIFIER

On the basis of Bayesian theory, each feature of a particular class is irrelevant to any other features. Thus, Naive Bayes (NB) is proposed as a probabilistic classifier. Compared with other classification modes, NB consumes much less training time, and it can solve small samples learning problem effectively. Supposing D represents a training samples set for n-patterns with class labels, Y stands for event vector. The event Y, which is owned by the pattern with highest posterior probability, can be explained as below:

$$P(C_i|Y) > P(C_j|Y) \quad for\ 1 \le j \le n, j \ne i \tag{15}$$

where

$$P(C_i|Y) = \frac{P(Y|C_i)P(C_i)}{P(Y)} \tag{16}$$

where $P(C_i)$ stands for the class prior probability. $P(Y)$ is the prior probability of Y. $P(C_i|Y)$ represents the posterior probability. $P(Y|C_i)$ denotes the posterior probability of Y conditioned on C_i.

It should be pointed out that $P(Y)$ is constant for all classes, so it is only necessary to maximize the numerator of $P(C_i|Y)$. When class prior probabilities can be calculated, the $P(C_1) = P(C_2) = \cdots = P(C_n)$ and $P(Y|C_i)$ is maximized. Otherwise, the class prior probabilities can be obtained via $P(C_i) = |C_{i,D}/|D|$, where $|C_{i,D}|$ denotes the number of a training set of the class C_i in D.

In order to decrease the computing complexity of estimating $P(Y|C_i)$, the events that are independent conditionally of each other is utilized to construct classifier. Then

$$\begin{aligned} P(Y|C_i) &= \prod_{k=1}^{n} P(y_k|C_i) \\ &= P(y_1|C_i) \times P(y_2|C_i) \times \cdots P(y_n|C_i) \end{aligned} \tag{17}$$

The probabilities $P(y_1|C_i), P(y_2|C_i), \cdots, P(y_n|C_i)$ are calculated from the training set and y_k denotes the value of an event for the data set Y.

To obtain the class label of Y, $P(Y|C_i)P(C_i)$ is estimated for each class C_i. According to the condition in Eq. 18, the class label of event Y can be considered as C_i by using NB classifier.

$$P(Y|C_i)P(C_i) > P(Y|C_j)P(C_j) \quad for\ 1 \le j \le n, j \ne i \tag{18}$$

V. PROPOSED DIAGNOSTIC SCHEME

The implementation of the proposed diagnostic scheme is depicted in Fig.1. The whole fault diagnosis procedure is divided into four steps: data acquisition, feature extraction, feature selection and fault location, which are described as follows:

Step I: The fault signals and reference signal are sampled from the CUTs, and the sample set is divided into two sets: training set and testing set.

Step II: The cross-wavelet transform is employed to convert the original samples set into the time-frequency representations (TFR), then we utilized the Local Optimal oriented Pattern (LOOP) on TFRs to generate the original features set.

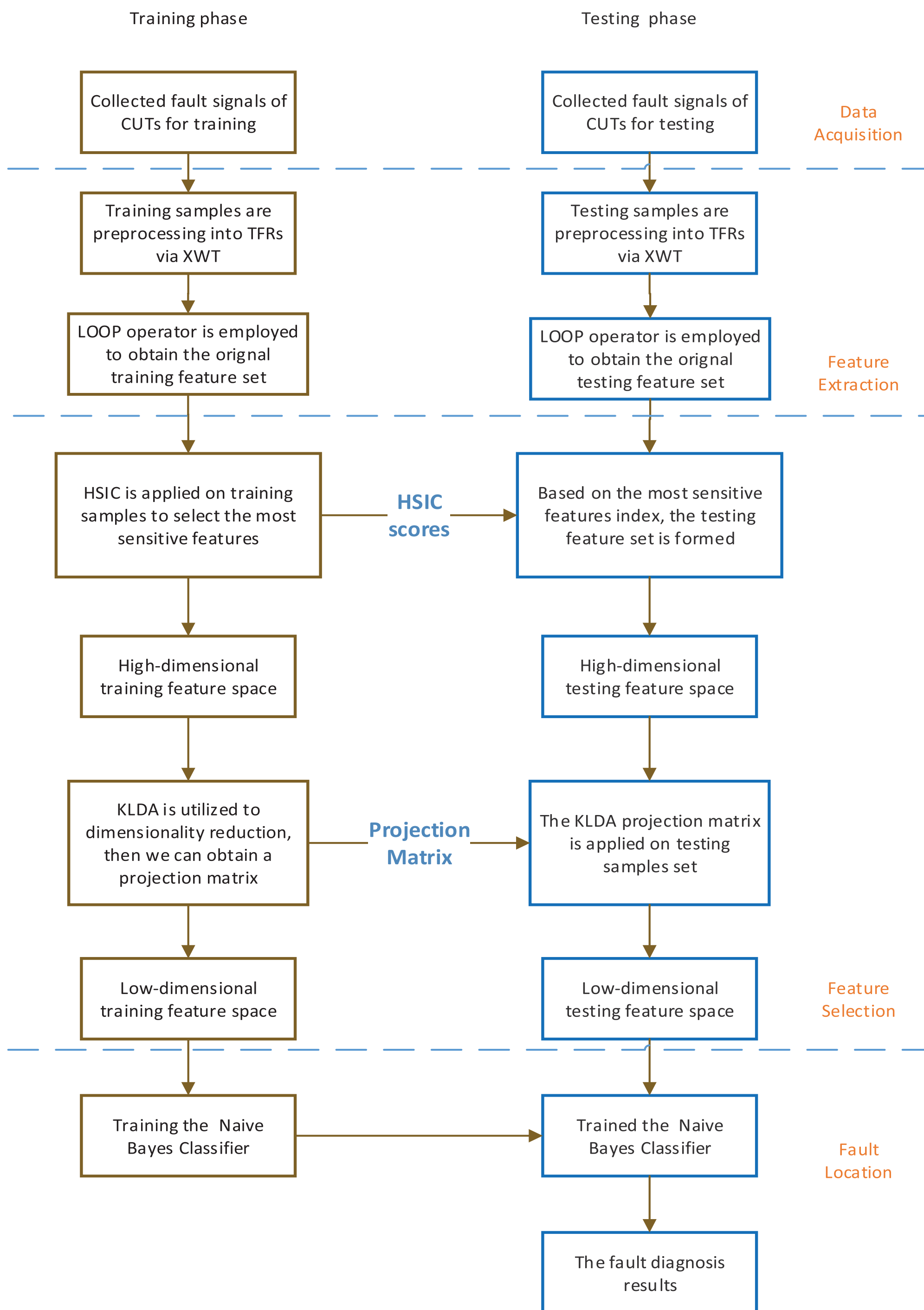


FIGURE 1. **Implementation of the proposed fault diagnostic technique.**

Step III: For feature selection, Hilbert Schmidt Independence Criterion (HSIC) is utilized to determine the most sensitive LOOP characteristics which have not high correlation. Then, with the use of KLDA, the low-dimensional representations are obtained from training and testing sample set.

Step IV: The low dimensional training feature set is utilized as the input vector of Naive Bayes classifier. Then, low-dimensional testing feature set is applied to conduct fault location via trained Naive Bayes classifier.

To validity the effectiveness of the proposed model, the K-fold cross validation (K-CV)is adopted to evaluate the

performance of presented scheme. The fundamental idea of K-CV is to divide the data-set into K groups, K-1 groups are the training samples and the other is the validation samples. Here, we set K = 5.

VI. RESULTS AND ANALYSIS

A. EXPERIMENTAL CIRCUITS

In this section, two circuits under test (CUTs) including a four opamp biquad highpass filter and a duffing chaotic circuit, are utilized to illustrate the diagnosis procedure of the proposed scheme. Multisim software is implemented to simulate circuits, and the proposed algorithm is conducted by using Matlab tool in a personal computer with a 2.5 GHz processor and 8-GB random access memory. Each fault class of testing circuits is conducted 60 Monte-Carlo analysis.

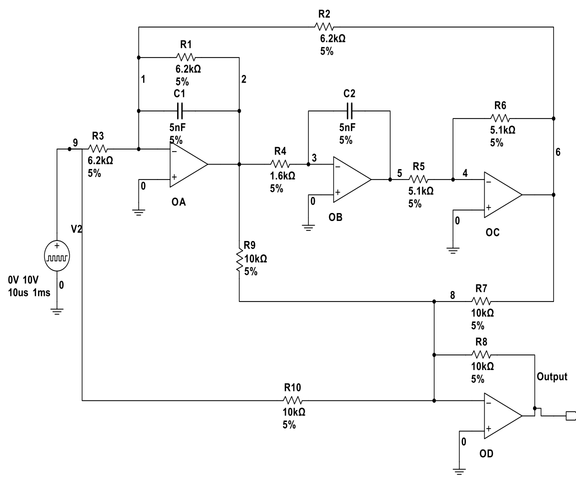


FIGURE 2. Opamp biquad filter circuit.

1) CUT 1:FOUR OPAMP BIQUAD HIGHPASS FILTER

The four opamp high-pass filter circuit depicted in Fig.2 is utilized as the first experimental circuit. The nominal value and the tolerance of each component are all labeled in the figure. As shown in Fig.2, the test pulse signal was designed as the excitation source. Meanwhile, the sampling time of output signal was assumed as $1ms$. Table 1 lists the whole single- and multiple- faults cases of opamp circuit. The normal working state of CUT was named as F0, which is not listed in Table 1. This test is labeled Case 1. Additionally, the incipient fault means that the parameter of components just beyond tolerance range and the performance of circuit begins to degrade from its normal working state [39]. Compared with abrupt fault (hard and soft fault), it is difficult to conduct incipient fault diagnosis because much more overlapping exists [7], [40]. Therefore, incipient fault diagnosis is also implemented in this section, and the incipient faults are also shown in Table. 2. This case is referred to as Case 2. Moreover, the size of raw signals is 1020. Among these samples, 50% are used for training and the last 50% are used for testing.

TABLE 1. Single and multiple faults of the opamp filter circuit.

Fault code	Fault class	Fault value
F1	$R1\downarrow$	3000Ω
F2	$R1\uparrow$	15000Ω
F3	$R2\downarrow$	2000Ω
F4	$R2\uparrow$	18000Ω
F5	$R3\downarrow$	2700Ω
F6	$R3\uparrow$	12000Ω
F7	$R4\downarrow$	500Ω
F8	$R4\uparrow$	2500Ω
F9	$C1\downarrow$	$2.5nF$
F10	$C1\uparrow$	$10nF$
F11	$C2\downarrow$	$1.5nF$
F12	$C2\uparrow$	$15nF$
F13	$C1\uparrow C2\downarrow$	above
F14	$C2\downarrow R1\uparrow$	above
F15	$R1\uparrow R2\downarrow$	above
F16	$R2\uparrow R4\downarrow$	above

TABLE 2. Incipient faults of the opamp filter circuit.

Fault code	Fault class	Fault value
F1	$R1\downarrow$	$4.96k\Omega$
F2	$R1\uparrow$	$7.44k\Omega$
F3	$R2\downarrow$	$4.96k\Omega$
F4	$R2\uparrow$	$7.44k\Omega$
F5	$R3\downarrow$	$4.96k\Omega$
F6	$R3\uparrow$	$7.44k\Omega$
F7	$R4\downarrow$	1.28Ω
F8	$R4\uparrow$	1.92Ω
F9	$C1\downarrow$	$4nF$
F10	$C1\uparrow$	$6nF$
F11	$C2\downarrow$	$4nF$
F12	$C2\uparrow$	$6nF$
F13	$OA(IN+,OUT)$	open
F14	$OA(IN+,IN-)$	short
F15	$OB(IN-,OUT)$	short
F16	$OD(IN-,IN+)$	$1M\Omega(leakage)$

2) CUT 2:DUFFING CHAOTIC CIRCUIT

To investigate the performance of proposed method in fault diagnosis of nonlinear circuit, a test of the duffing chaotic circuit shown in Fig.3 is conducted in this section. In this case, a excitation signal with the frequency of 0.155159Hz and the amplitude of 0.7414148V is chosen. The normal tolerance of resistor and capacitor is also assumed as 5%. We only collected the signals at output node, and a 30% deviation of nominal value was considered as a fault condition. The fault modes are listed in Table 3. In this work, the test is denoted as Case 3. After data acquisition, we obtain the original samples set with size of 1260. The size of training samples set and testing samples set are all equal to $630(30\times 21)$.

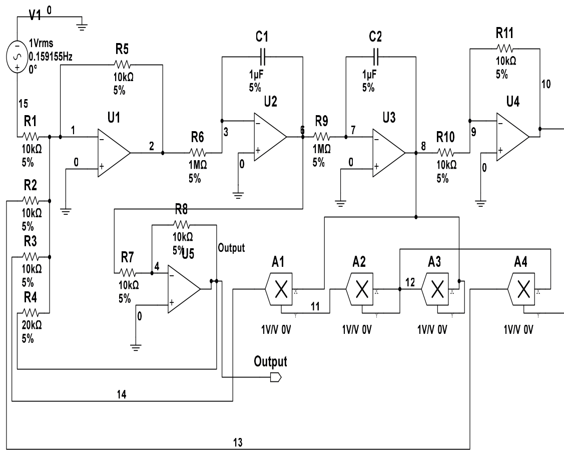


FIGURE 3. Duffing chaotic circuit.

TABLE 3. Fault modes of the duffing chaotic circuit.

Fault code	Fault class	Fault value
F0	–	–
F1	$R1\downarrow$	$7k\Omega$
F2	$R1\uparrow$	$13k\Omega$
F3	$R2\downarrow$	$7k\Omega$
F4	$R2\uparrow$	$13k\Omega$
F5	$R3\downarrow$	$7k\Omega$
F6	$R3\uparrow$	$13k\Omega$
F7	$R4\downarrow$	$14k\Omega$
F8	$R4\uparrow$	$26k\Omega$
F9	$R8\downarrow$	$7k\Omega$
F10	$R9\downarrow$	$0.7M\Omega$
F11	$R10\uparrow$	$13k\Omega$
F12	$C1\downarrow$	$0.7\mu F$
F13	$C2\uparrow$	$1.3\mu F$
F14	$R1\uparrow R2\uparrow$	$(13k\Omega)(13k\Omega)$
F15	$R1\downarrow R3\downarrow$	$(7k\Omega)(7k\Omega)$
F16	$R5\uparrow C1\uparrow$	$(7k\Omega)(1.3\mu F)$
F17	$R6\downarrow C2\downarrow$	$(0.7M\Omega)(0.7\mu F)$
F18	$U2(IN+, IN-)$	$100\Omega(Leakage)$
F19	$U4(IN-, OUT)$	open
F20	$U5(IN+, OUT)$	short

B. FEATURE EXTRACTION VIA XWT AND LOOP

Fig.4 depicts the XWT time-frequency spectrums of fault class F3, F7, F11 and F16 of opamp circuit in Case 1 experiment. And Fig.5 shows the TFRs of normal condition F0 and fault class F16 for duffing chaotic circuit in Case 3 experiment. As shown in Fig.4 and Fig.5, it is clear that the phases- and energy- distribution of TFRs in different fault cases are significant differences. Nevertheless, it also poses a challenging task to accurately recognize massive spectrum images of fault signals in various fault types. Hence, it is indispensable to develop an advance feature extraction tool to fetch sufficient information from TFRs.

With the utilization of LOOP implementation, spectrum images of fault signals are further processed to obtain a sequence of LOOP descriptors with size of 256. Fig.6 demonstrates the LOOP features of fault types F1 and F16 in Case 1. For case 2, the LOOP features of F3 and F13 are shown in Fig.7. And, Fig.8 illustrates the LOOP characteristics of normal condition F0 and fault class F12 in Case 3. Therefore, the size of original training feature set for Case 1, Case 2 and Case 3 are 510×256, 510×256 and 630×256, respectively. If these feature sets are imported into the classifiers directly without any further feature reduction, it will increase high computational complexity and storage capacity. Hence, it is necessary to develop a feature selection approach to find the most discriminative and concise feature candidates in the next step.

C. FEATURE SELECTION

As an novel filter method, HSIC can be used to estimate the independence score of each feature and preselect M high score feature to construct a candidate subset for next dimensionality reduction. Hence, the preliminary candidate feature set with low dimension can be drew from the original high dimensional feature set. Due to the different scales of each feature, which may bring out bias in fault location, z-scores are implemented on training samples set to eliminate the influence of bias. The Z-score mathematical formula is described below

$$x* = \frac{x - \mu}{\sigma} \tag{19}$$

where $x*$ denotes the normalized value. x is the original feature value. μ represents the mean value of feature. σ stands for standard deviation of feature. In addition, the test samples set is also normalized referring to the mean and standard deviation of the train sample sets.

Fig.9 shows the scores of the 256 features for Case 1. It implies that as many as 63 scores are larger than 0 which indicts that these features have a great influence on classification. Fig.10 demonstrates the scores of the 256 features for Case 2. As shown in Fig.10, we can observe that 54 feature candidates are greater than 0. Fig.11 depicts the scores of 256 features for Case 3. It is clear to find that 47 feature scores are larger than 0. Considering that retaining useful information as much as possible, the 63 features for Case 1, 54 features for Case 2 and 47 features for Case 3 are all chosen as the features reselection candidates. Then, 63, 54 and 47 greater score features for these three Cases are separately selected to construct three preliminary feature subsets. Consequently, the original 510×256, 510×256 and 630×256 high dimensionality can be reduced to 510×63, 510×54 and 630×47. However, the obtained feature subsets cannot guarantee desirable classification performance and the dimensions are still very high. In view of these issues, a traditional dimensionality reduction method, kernel linear discriminant analysis (KLDA) is conducted on the pre-selection feature subsets to obtain low dimensional representation.

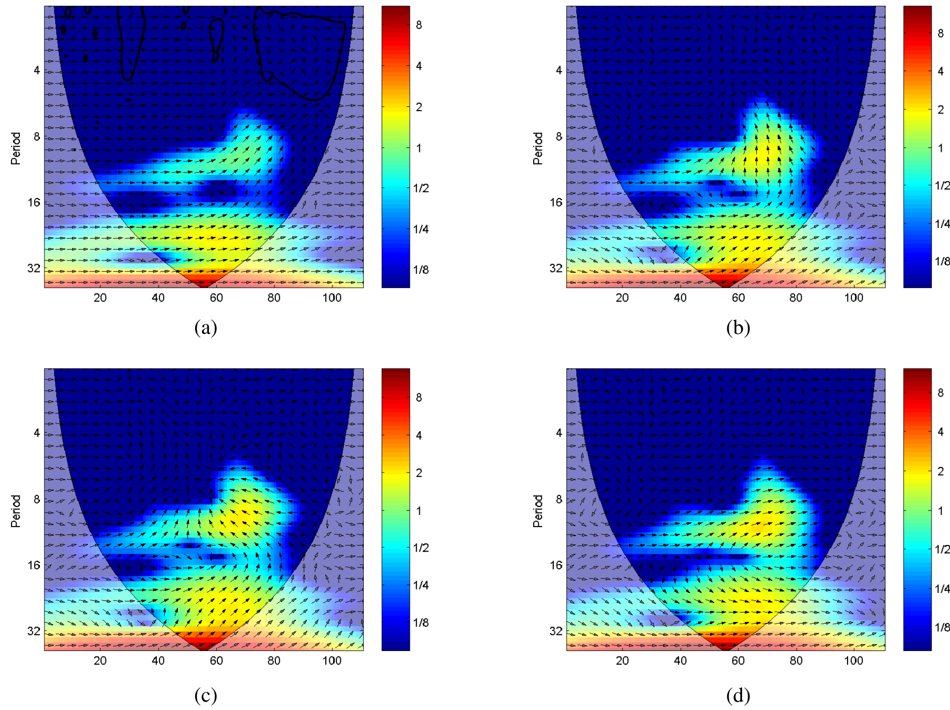


FIGURE 4. The XWT spectrum image of opamp circuit for Case 1. (a) F3. (b) F7. (c) F11. (d)F16.

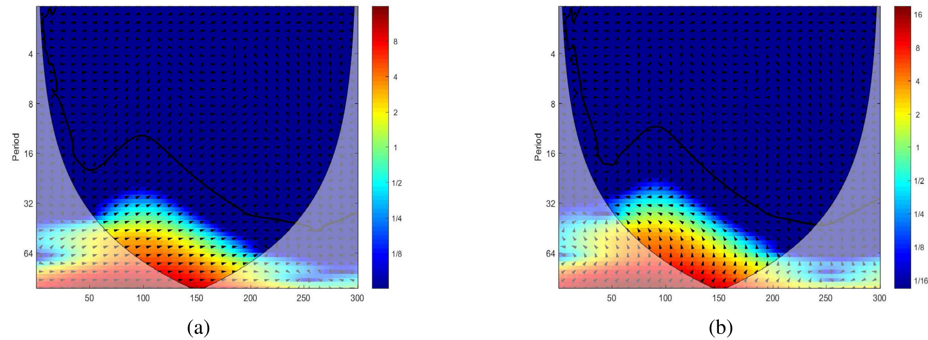


FIGURE 5. The XWT spectrum image of duffing circuit for Case 3. (a) F0. (b) F12.

The final feature vectors of training sample set obtained by KLDA are implemented as the input of Naive Bayes classifier. With regard to the classification performance of feature numbers, the simulations of total features generated via KLDA on recognition accuracy are conducted. Additionally, serial existing feature reduction methods including locality sensitive discriminant analysis (LSDA) [41], principal component analysis (PCA) [42] and isometric projection (IsoP) [43] are also implemented on data-sets.

Fig.12(a) and 12(b) depict the classification rate for Case 1 and Case 2 varying from the first two features to all features. It can be observed that the performance of HSIC-KLDA is superior than that of other methods for these two cases. The classification rate of the proposed method for Case 1 with first eight features achieves 99.80% which is the highest. It indicates that only one sample is classified into wrong class. Meanwhile, for Case 2, as the number of features increasing, a higher recognition accuracy is obtained and the highest rate of 98.8235% is achieved while the number of candidates is 16. Fig.12(c) shows the results with changing numbers of features when implementing diagnosis task of Case 3. With increasing number of principal features, the classification rate of these methodologies continue to increase, and the proposed scheme achieves the highest classification rate 97.9365% when the number of features equals to the 16. It demonstrates that the proposed framework based on HSIC and KLDA outperforms the other existing methods on the same data-set with considerably higher accuracy for Case 3.

For better visual presentation, the testing dataset 3-D distribution for Case 1 in the above feature selection and reduction scheme are plotted in Fig.13. As depicted in Fig. 13, it illustrates that the normal and fault conditions are well

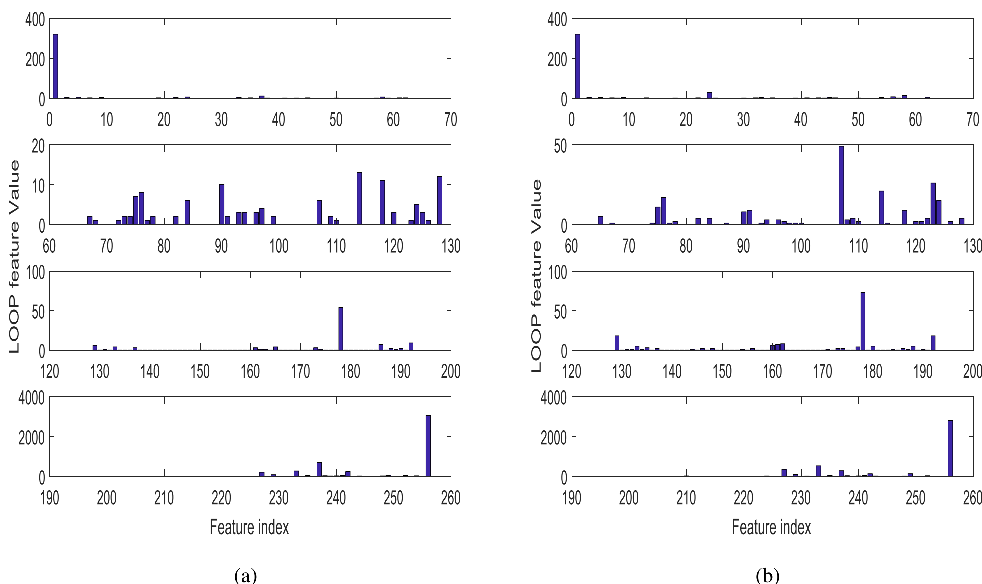


FIGURE 6. Two sample distributions of loop features in Case 1. (a) F1. (b) F16.

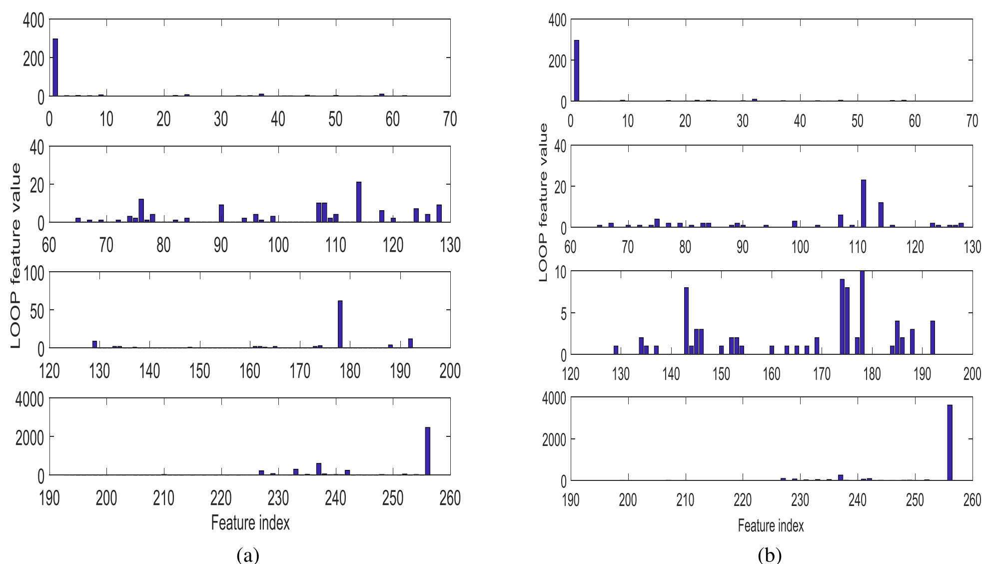


FIGURE 7. Two sample distributions of loop features in Case 2. (a) F3. (b) F13.

separated and is conducive to accurately classify. However, there also exist some weakly overlapping sections between some fault classes. Therefore, it is obvious that the proposed scheme can fetch the most discriminative features of the fault signals when dealing with single- and multiple- faults in opamp circuit.

D. CLASSIFICATION RESULTS

To investigate the advantage of Navie Bayes classifier, comparisons based on Navie Bayes and support vector machine (SVM) are discussed in this section. Fig.14, Fig.15 and Fig.16 show the confusion matrix of the proposed method for the Case 1, Case 2 and Case 3. The fact that most of diagonal elements in different fault classes in Case 1 are close to 1 suggests that the use of proposed method can improve the possibility of classifying the patterns into its actual class and lower the possibility of confusing between different classes. Moreover, the performance of SVM-based classification is inferior to that of Navie Bayes. Because of the fact that much more overlapping exists between different incipient fault classes in Case 2, the classification accuracies of 97.06% and 96.47% are achieved by using Naive Bayes and SVM classifiers. Although the rates are lower than that of Case 1, they are still desirable. Furthermore, when dealing with duffing chaotic circuit, there also exists performance degradation. It implies that proposed approach can not effectively fulfill

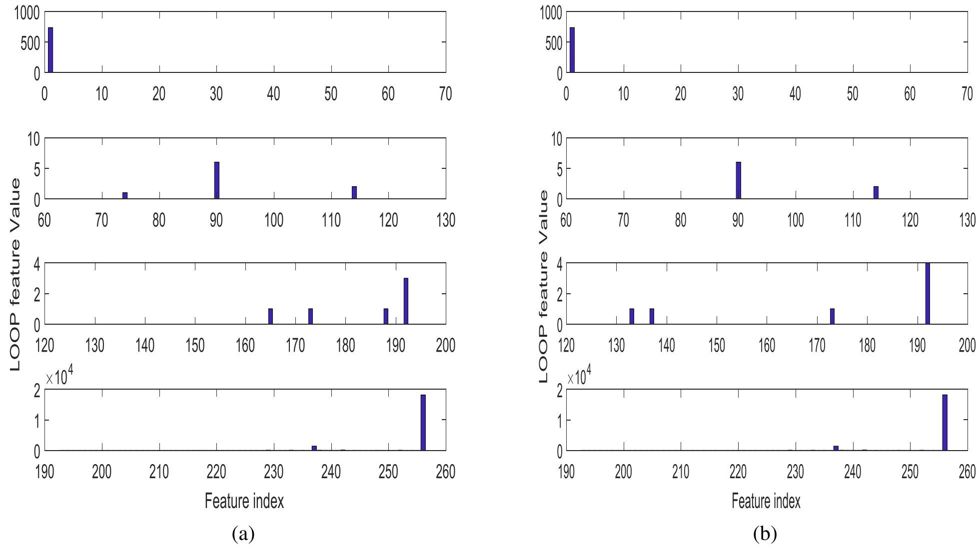


FIGURE 8. Two sample distributions of loop features in Case 3. (a) F0. (b) F12.

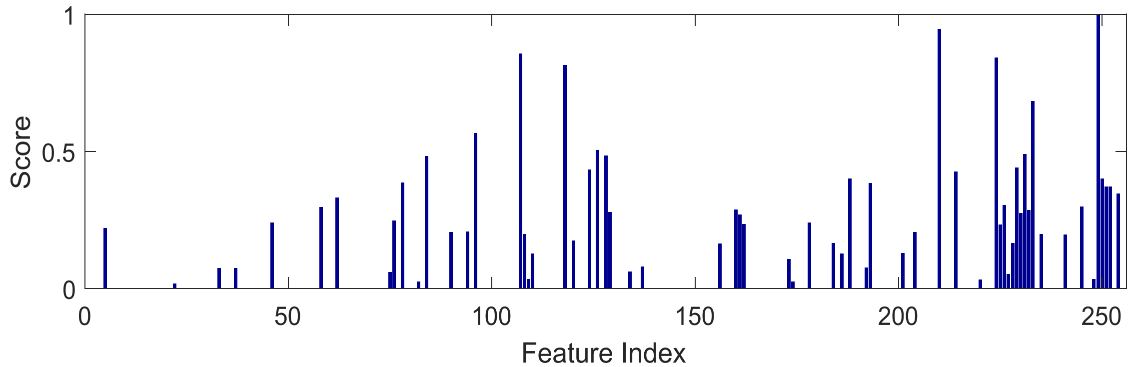


FIGURE 9. HSIC scores of all features for Case 1.

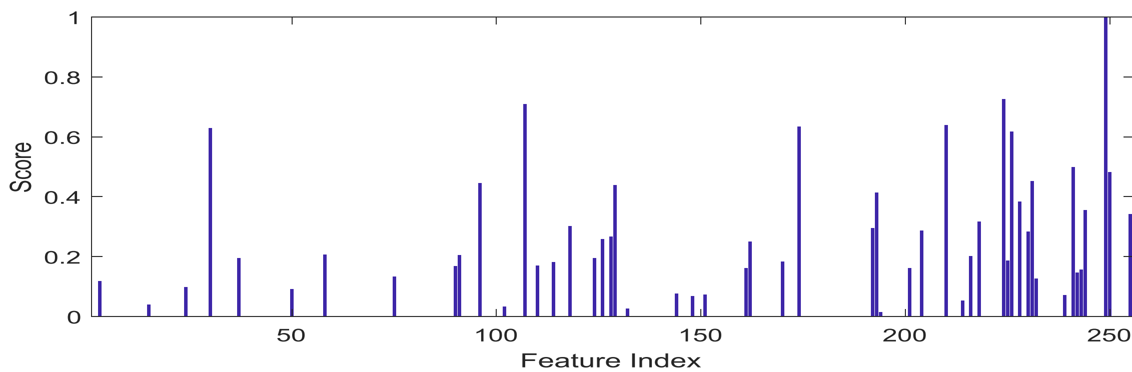


FIGURE 10. HSIC scores of all features for Case 2.

the task of fault diagnosis in complex nonlinear circuit, but still achieve satisfactory performance.

In order to compare the performance of proposed framework with other scheme, 5 extra feature extraction methods which have been reported in other works, are conducted to fault diagnosis, and described as follows: (1) XWT+uLBP+PCA: Uniform LBP is employed on XWT spectrum to obtain features. These features are utilized as inputs of classifier with PCA processing. (2) XWT+B2DLDA: The B2DLDA is performed on the resultant spectrums to generate four dimensional features for each sample. (3) ST+LOOP+KLDA: We implement

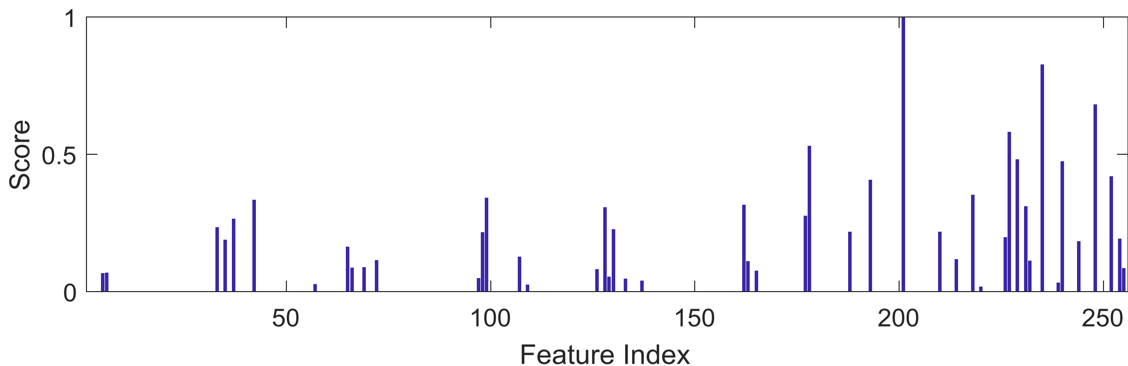


FIGURE 11. HSIC scores of all features for Case 3.

(a)

(b)

(c)

FIGURE 12. Classification accuracy(%)for different approaches with varying numbers of features of (a) Case 1. (b) Case 2. (c) Case 3.

Stockwell transform (ST) on raw signal of defect circuit to obtain TFRs, then LOOP is performed to construct original high-dimensional feature set. Afterward, KLDA is conducted to determine the lower-dimensional representations. (4) WT-energy features: the wavelet transform (WT) based decomposition node energy index are designed as

TABLE 4. Performance comparison with other fault diagnosis scheme.

Method	Feature Extraction	Feature Selction	Classifier	Case 1		Case 2		Case 3	
				Feature Number	accuracy	Feature Number	accuracy	Feature Number	accuracy
Proposed	XWT+LOOP	HSIC +KLDA	Navie Bayes	4	99.41%	10	97.25%	12	97.03%
Method 1	XWT+uLBP	PCA	SVM	4	93.53%	10	90.39%	12	90.63%
Method 2	ST+LOOP	KLDA	Navie Bayes	6	94.31%	10	94.31%	12	85.37%
Method 3	XWT	B2DLDA	VVRKFA	4	99.22%	8	94.71%	10	93.81%
Method 4	WT-energy features	PCA	SVM	8	91.18%	12	92.35%	12	87.78%
Method 5	FFT freuqency amplitude	PCA	PNN	6	92.35%	12	89.22%	12	85.24%

TABLE 5. Diagnosis results based on different size of training samples.

Case 1		Case 2		Case 3	
Size of training/ testing samples	testing accuracy	Size of training/ testing samples	testing accuracy	Size of training/ testing samples	testing accuracy
85/935	97.54%	85/935	92.30%	105/1155	81.99%
170/850	98.94%	170/850	93.65%	210/1050	90.19%
255/765	99.35%	255/765	94.90%	315/945	95.13%
340/680	99.41%	340/680	95.74%	420/840	95.60%
425/595	99.66%	425/595	96.13%	525/735	96.89%
510/510	99.41%	510/510	97.25%	630/630	97.30%

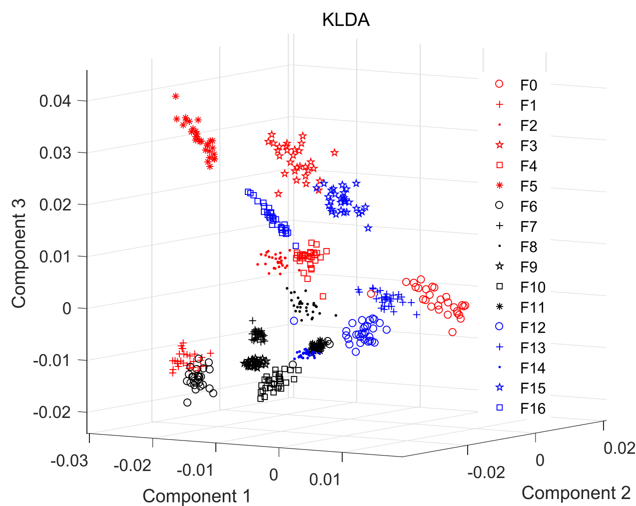


FIGURE 13. Visualization of the first three components for Case 1.

feature descriptors, and we extract the energy of nodes coefficient containing the most dominant fault information to construct the feature vectors. (5) FFT+PCA: the Fast Fourier Transform (FFT) is applied to extract the main information of the fault signal, then we propose Principle Component Analysis (PCA) to obtain a lower dimensional feature. Subsequently, they are imported into classifiers. The classification rate of these 5 methods for comparison and the proposed approach are displayed in Table 4. The influence of sample size on the performance of proposed methods is investigated in Table 5. Besides classification accuracy, the computation cost is another important measure used to evaluate the proposed scheme, especially when a large number of training samples are involved. The consuming times of the proposed algorithm and other methods in Table 4 for each case are reported in Fig.17.

It can be summed up that

1) In Table 4, the proposed approach is effective in fault diagnosis for these three cases. And it all achieve desirable performances with lower feature dimensionality. Thus, we can consider that the proposed approach is promising, which is not limited to linear circuit but could be employed in fault diagnosis of nonlinear circuit.
2) In Table 5, the accuracies of proposed method maintain increasing and eventually settle around 99.41% for Case 1, 97% for Case 2 and 97% for Case 3. Hence, the proposed method can effectively avoid overfitting. In general, larger size of the training samples will achieve better performance associated with testing accuracies.
3) As shown in Fig.17, we observe that the proposed method consumes longer calculation time than other methods. However, the majority of consumption is

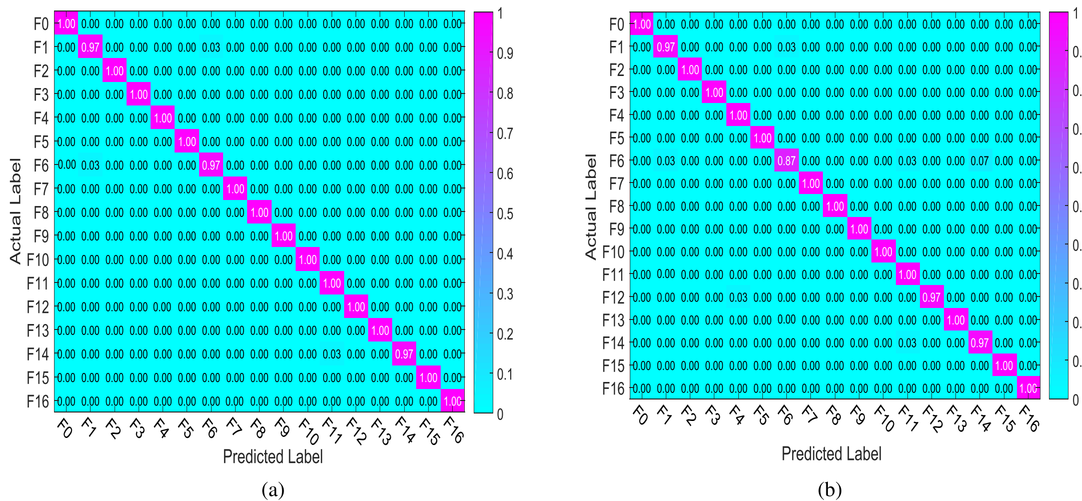


FIGURE 14. The Confusion matrixes for Case 1 via (a) Naive Bayes. (b) SVM.

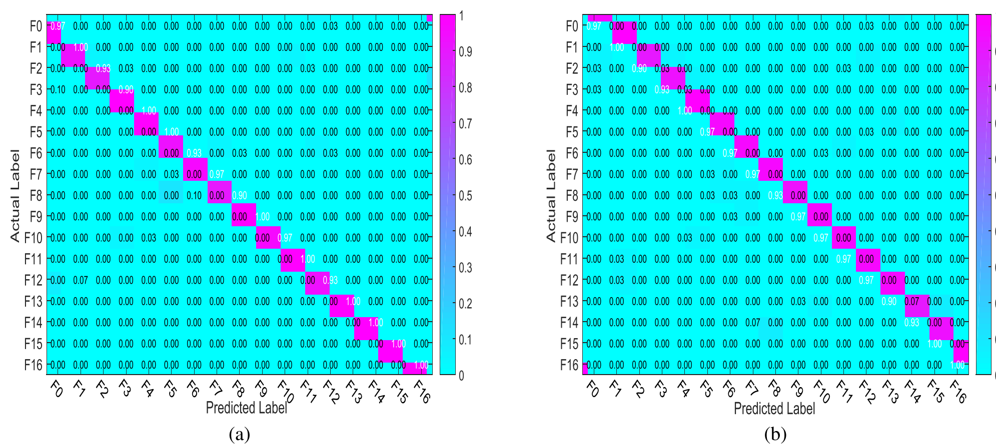


FIGURE 15. The Confusion matrixes for Case 2 via (a) Naive Bayes. (b) SVM.

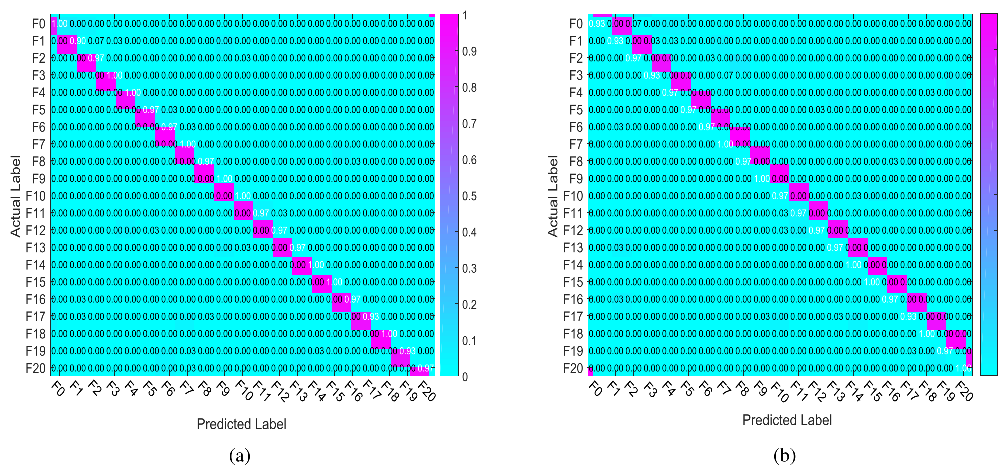


FIGURE 16. The Confusion matrixes for Case 3 via (a) Naive Bayes. (b) SVM.

devoted to the feature extraction. The reason is that the feature extraction algorithm based image requires much more computation resources. Besides, the computation cost is still in an acceptable range while the proposed algorithm achieves the highest classification rate.

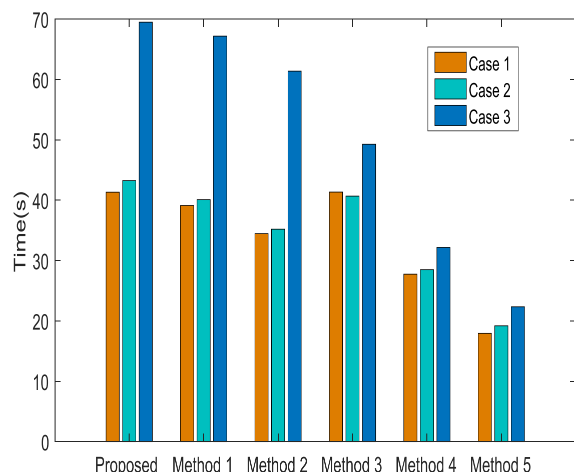


FIGURE 17. The consuming time of proposed algorithm for three cases.

In addition, big O notation is also adopted to analysis the time complexity of the proposed feature extraction algorithm. Naive Bayes classifier is not considered because it only consumes much fewer time. In this algorithm, XWT is utilized with the cost of $O(n^2\ log(n))$. Then The time complexity to apply LOOP operation is $O(n^2\ log(n))$. Finally, HSIC is employed to select the sensitive features with the cost of $O(n^2)$.

VII. CONCLUSIONS

This work has proposed an effective fault diagnosis scheme based on spectrum image processing and feature selection to achieve the optimal classification performance. The simulation results of two cases demonstrate the superiority of the presented methodology. The conclusions can be drawn as below:

(1) To obtain the sufficient information of defect circuits, we propose a time-frequency image feature extraction framework. In view of the above, XWT is applied on the raw fault signal to generate the TFRs, then LOOP is utilized to pick up characteristics. As a result, a high dimensional feature set, which contains both representative characteristics and useless information, is obtained for each fault signal. Therefore, it is necessary to implement further process on these high dimensional features to adaptively select the most discriminative ones.

(2) To address the above issues, HSIC is employed to preselect the features with high scores. These preselected features are utilized as candidates for further feature process phase. In this phase, KLDA is conducted on candidates set. By using the integration method of HSIC and KLDA, the final feature set with lower dimension and better, can be obtained.

(3) The simulation results of experimental opamp circuit and duffing circuit reveal that the proposed diagnosis framework can effectively and accurately classify different fault types. Comparative study demonstrate that the presented scheme can achieve a desirable performance with lower feature dimension. Therefore, it can be summed up that this methodology would be suitable and efficient for fault diagnosis of analog circuit.

However, the utilization of the proposed approach still has two fundamental limitations. First, the algorithm by using image-oriented feature extraction is time-consuming, that presents exceptional challenge to the real-time application of the proposed method. The second drawback is that the proposed method is unable to extract the most discriminative features from nonlinear circuit. In view of this, We should focus on these issues in the future work.

REFERENCES

[1] Z. Liu, Z. Jia, C.-M. Vong, S. Bu, and J. Han, "Capturing high-discriminative fault features for electronics-rich analog system via deep learning," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1213–1226, Jun. 2017.

[2] B. Long, S. Tian, and H. Wang, "Diagnostics of filtered analog circuits with tolerance based on LS-SVM using frequency features," *J. Electron. Test.*, vol. 28, no. 3, pp. 291–300, 2012.

[3] A. S. S. Vasan, B. Long, and M. Pecht, "Diagnostics and prognostics method for analog electronic circuits," *IEEE Trans. Ind. Electron.*, vol. 60, no. 11, pp. 5277–5291, Nov. 2013.

[4] Y. Deng, Y. Shi, and W. Zhang, "An approach to locate parametric faults in nonlinear analog circuits," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 2, pp. 358–367, Feb. 2012.

[5] L. Yuan, Y. He, J. Huang, and Y. Sun, "A new neural-network-based fault diagnosis approach for analog circuits by using kurtosis and entropy as a preprocessor," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 3, pp. 586–595, Mar. 2010.

[6] M.-H. Chang, C. Chen, D. Das, and M. Pecht, "Anomaly detection of light-emitting diodes using the similarity-based metric test," *IEEE Trans. Ind. Informat.*, vol. 10, no. 3, pp. 1852–1863, Aug. 2014.

[7] B. Long, W. Xian, M. Li, and H. Wang, "Improved diagnostics for the incipient faults in analog circuits using LSSVM based on PSO algorithm with Mahalanobis distance," *Neurocomputing*, vol. 133, pp. 237–248, Jun. 2014.

[8] T. Wang, J. Qi, H. Xu, Y. Wang, L. Liu, and D. Gao, "Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter," *ISA Trans.*, vol. 60, pp. 156–163, Jan. 2016.

[9] V. Behjat, A. Vahedi, A. Setayeshmehr, H. Borsi, and E. Gockenbach, "Sweep frequency response analysis for diagnosis of low level short circuit faults on the windings of power transformers: An experimental study," *Int. J. Elect. Power Energy Syst.*, vol. 42, no. 1, pp. 78–90, 2012.

[10] E. Cabal-Yepez, A. G. Garcia-Ramirez, R. J. Romero-Troncoso, A. Garcia-Perez, and R. A. Osornio-Rios, "Reconfigurable monitoring system for time-frequency analysis on industrial equipment through STFT and DWT," *IEEE Trans. Ind. Informat.*, vol. 9, no. 2, pp. 760–771, May 2013.

[11] X. Chen and Z. Feng, "Time-frequency analysis of torsional vibration signals in resonance region for planetary gearbox fault diagnosis under variable speed conditions," *IEEE Access*, vol. 5, pp. 21918–21926, 2017.

[12] P. Cao, S. Zhang, and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning," *IEEE Access*, vol. 6, pp. 26241–26253, 2018.

[13] W. Li, A. Monti, and F. Ponci, "Fault detection and classification in medium voltage DC shipboard power systems with wavelets and artificial neural networks," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 11, pp. 2651–2665, Nov. 2014.

[14] Y. Tan, Y. Sun, and X. Yin, "Analog fault diagnosis using S-transform preprocessor and a QNN classifier," *Measurement*, vol. 46, no. 7, pp. 2174–2183, Aug. 2013.

[15] P. K. Mishra, A. Yadav, and M. Pazoki, "A novel fault classification scheme for series capacitor compensated transmission line based on bagged tree ensemble classifier," *IEEE Access*, vol. 6, pp. 27373–27382, 2018.

[16] S. Tang, Z. Li, and L. Chen, "Fault detection in analog and mixed-signal circuits by using Hilbert–Huang transform and coherence analysis," *Microelectron. J.*, vol. 46, no. 10, pp. 893–899, 2015.

[17] M. Riera-Guasp, M. Pineda-Sanchez, J. Perez-Cruz, R. Puche-Panadero, J. Roger-Folch, and J. A. Antonino-Daviu, "Diagnosis of induction motor faults via Gabor analysis of the current in transient regime," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 6, pp. 1583–1596, Jun. 2012.

[18] B. Li, P.-L. Zhang, D.-S. Liu, S.-S. Mi, and P.-Y. Liu, "Classification of time–frequency representations based on two-direction 2DLDA for gear fault diagnosis," *Appl. Soft Comput.*, vol. 11, no. 8, pp. 5299–5305, 2011.

[19] K. Wang, R. Liao, L. Yang, J. Li, S. Grzybowski, and J. Hao, "Optimal features selected by NSGA-II for partial discharge pulses separation based on time-frequency representation and matrix decomposition," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, no. 3, pp. 825–838, Jun. 2013.

[20] W. Li, M. Qiu, Z. Zhu, B. Wu, and G. Zhou, "Bearing fault diagnosis based on spectrum images of vibration signals," *Meas. Sci. Technol.*, vol. 27, no. 3, p. 035005, 2016.

[21] H. Chen, J. Wang, J. Li, and B. Tang, "A texture-based rolling bearing fault diagnosis scheme using adaptive optimal kernel time frequency representation and uniform local binary patterns," *Meas. Sci. Technol.*, vol. 28, no. 3, p. 035903, 2017.

[22] J. Zhang, J. Liang, and H. Zhao, "Local energy pattern for texture classification using self-adaptive quantization thresholds," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 31–42, Jan. 2013.

[23] T. Chakraborti, B. McCane, S. Mills, and U. Pal. (2017). "LOOP descriptor: Local optimal oriented pattern." [Online]. Available: https://arxiv.org/abs/1710.09317

[24] M. Aminian and F. Aminian, "A modular fault-diagnostic system for analog electronic circuits using neural networks with wavelet transform as a preprocessor," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 5, pp. 1546–1554, Oct. 2007.

[25] C. Zhang, Y. He, L. Yuan, W. He, S. Xiang, and Z. Li, "A novel approach for diagnosis of analog circuit fault by using GMKL-SVM and PSO," *J. Electron. Test.*, vol. 32, no. 5, pp. 531–540, Oct. 2016.

[26] A. Sharmila and P. Geethanjali , "DWT based detection of epileptic seizure from eeg signals using Naive Bayes and k-NN classifiers," *IEEE Access*, vol. 4, pp. 7716–7727, 2016.

[27] B. Tang, S. Kay, and H. He, "Toward optimal feature selection in Naive Bayes for text categorization," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 9, pp. 2508–2521, Sep. 2016.

[28] Ö. F. Arar and K. Ayan, "A feature dependent Naive Bayes approach and its application to the software defect prediction problem," *Appl. Soft Comput.*, vol. 59, pp. 197–209, Oct. 2017.

[29] Y. Liu, J. Zhang, and L. Ma, "A fault diagnosis approach for diesel engines based on self-adaptive WVD, improved FCBF and PECOC-RVM," *Neurocomputing*, vol. 177, pp. 600–611, Dec. 2016.

[30] X. Zhang, Q. Zhang, M. Chen, Y. Sun, X. Qin, and H. Li, "A two-stage feature selection and intelligent fault diagnosis method for rotating machinery using hybrid filter and wrapper method," *Neurocomputing*, vol. 275, pp. 2426–2439, Nov. 2018.

[31] X. Liu, L. Ma, L. Song, Y. Zhao, X. Zhao, and C. Zhou, "Recognizing common CT imaging signs of lung diseases through a new feature selection method based on Fisher criterion and genetic optimization," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 2, pp. 635–647, Mar. 2015.

[32] G. Camps-Valls, J. Mooij, and B. Schölkopf, "Remote sensing feature selection by kernel dependence measures," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 587–591, Jul. 2010.

[33] B. B. Damodaran, N. Courty, and S. Lefèvre, "Sparse Hilbert Schmidt independence criterion and surrogate-kernel-based feature selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 4, pp. 2385–2398, Apr. 2017.

[34] L. Song, A. Smola, A. Gretton, J. Bedo, and K. Borgwardt, "Feature selection via dependence maximization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 1393–1434, Jan. 2012.

[35] Y. Xiao and L. Feng, "A novel neural-network approach of analog fault diagnosis based on kernel discriminant analysis and particle swarm optimization," *Appl. Soft Comput.*, vol. 12, no. 2, pp. 904–920, 2012.

[36] S. Banerjee and M. Mitra, "Application of cross wavelet transform for ECG pattern analysis and classification," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 2, pp. 326–333, Feb. 2014.

[37] M. I. Plett, "Transient detection with cross wavelet transforms and wavelet coherence," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1605–1611, May 2007.

[38] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf, "Measuring statistical dependence with Hilbert–Schmidt norms," in *Algorithmic Learning Theory*, S. Jain, H. U. Simon, and E. Tomita, Eds. Berlin, Germany: Springer, 2005, pp. 63–77.

[39] Y. Shi, Y. Deng, and W. Zhang, "Diagnosis of incipient faults in weak nonlinear analog circuits," *Circuits, Syst., Signal Process.*, vol. 32, no. 5, pp. 2151–2170, 2013.

[40] C. Zhang, Y. He, L. Yuan, and S. Xiang, "Analog circuit incipient fault diagnosis method using DBN based features extraction," *IEEE Access*, vol. 6, pp. 23053–23064, 2018.

[41] I. Attoui, N. Fergani, N. Boutasseta, B. Oudjani, and A. Deliou, "A new time–frequency method for identification and classification of ball bearing faults," *J. Sound Vib.*, vol. 397, pp. 241–265, 2017.

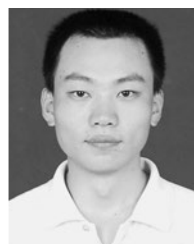
[42] Y. Xiao and L. Feng, "A novel linear ridgelet network approach for analog fault diagnosis using wavelet-based fractal analysis and kernel PCA as preprocessors," *Measurement*, vol. 45, no. 3, pp. 297–310, Apr. 2012.

[43] D. Cai, X. He, and J. Han, "Isometric projection," in *Proc. AAAI*, 2007, pp. 528–533.

WEI HE received the bachelor's degree in automation from the Hefei University of Technology, China, in 2009. He is currently pursuing the joint Ph.D. degree with the School of Electrical Engineering and Automation, Hefei University of Technology, and the University of Duisburg-Essen, Germany. His research interests include condition monitoring, fault diagnosis, machine learning, and signal processing.

YIGANG HE received the Ph.D. degree from Xian Jiaotong University, in 1996. He is currently a Professor and a Doctoral Supervisor with the School of Electrical Engineering, Wuhan University. His research interests are in the areas of circuit theory and its applications, testing and fault diagnosis of analog and mixed-signal circuits, smart grid, radio-frequency identification technology, and intelligent signal processing. He is one of the recipients of the National Distinguished Young Scientists Foundation.

BING LI received the B.E. degree in automobile engineering from the Chongqing University of Science and Technology, Chongqing, China, in 1995, and the M.E. and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 2006 and 2011, respectively. He was a Postdoctoral Researcher and a Visiting Scholar with the College of Electrical and Information Engineering, Hunan University, from 2011 to 2013. He has been an Associate Professor with the School of Electrical and Automation Engineering, Hefei University of Technology, Hefei, China, since 2013. His current research interests include radio-frequency identification technology, wireless sensor networks, and signal processing.



CHAOLONG ZHANG received the Ph.D. degree from the Hefei University of Technology, in 2018. He is currently an Associate Professor with the School of Physics and Electronic Engineering, Anqing Normal University. His current research interests include fault diagnostics and prognostic of analog and mixed-signal circuits, and battery capacity prognostic.

• • •