

Received February 20, 2020, accepted March 7, 2020, date of publication March 16, 2020, date of current version March 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981247

# A Fast and Intelligent Open-Circuit Fault Diagnosis Method for a Five-Level NNPP Converter Based on an Improved Feature Extraction and Selection Model

SHU YE<sup>1</sup>, (Student Member, IEEE), JIANGUO JIANG<sup>1</sup>,  
ZHONGZHENG ZHOU<sup>1</sup>, (Student Member, IEEE),  
CONG LIU<sup>1</sup>, (Student Member, IEEE), AND YUNLONG LIU<sup>2</sup>

<sup>1</sup>Key Laboratory of Control of Power Transmission and Conversion, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>2</sup>School of Mechanical Electronic and Information Engineering, China University of Mining and Technology, Beijing 221116, China

Corresponding author: Shu Ye (yeshu\_sjtu@163.com)

This work was supported in part by the National High-Tech Research and Development Program of China under Grant 2014AA052602, and in part by the National Basic Research Program of China under Grant 2014CB046306.

**ABSTRACT** The open-circuit faults of power semiconductor devices in multilevel converters are generally diagnosed by analyzing circuit signals. For converters with five or more levels, the difficulty of fault detection increases with increasing topological complexity, the number of switching devices and the number of candidate signal parameters. In this paper, a complete solution for open-circuit fault detection for a five-level nested neutral-point piloted (NNPP) converter is proposed based on improved unsupervised feature learning algorithms. Feature engineering and machine learning algorithms are applied for feature extraction and selection and the construction of classification models. Circuit signals are monitored, and their time-domain characteristics are extracted for fault recognition. An unsupervised feature learning selector, which combines a dependence-guided unsupervised feature selection (DGUFS) filter and a random forest feature selection (RFFS) wrapper to automatically select parameters and generate the optimal feature subset for fault detection, is proposed. The random forest (RF) algorithm is used to build a classifier. The experimental results show that the solution framework presented has the advantages of high efficiency, high flexibility, a superior fault recognition rate and good generalization ability.

**INDEX TERMS** Fault diagnosis, feature extraction and selection, filter-wrapper method, DGUFS-RFFS, open-circuit fault, five-level NNPP converter.

## I. INTRODUCTION

As the key equipment of power conversion systems, power electronic converters are widely used in motor drives and power systems. Power switching devices, with limited overload capacities, are inevitably exposed to high voltages and currents, which will lead to a high risk of damage and affect the safe and stable operation of the system. According to statistics, in the AC motor control systems for industrial applications, 38% failures derive from damage to power electronic devices [1]. The number of semiconductor devices increases with the voltage level of the converters, which will enhance

the risk of switch breakdown and reduce the reliability of the system [2]. The need for fast and accurate fault diagnosis technology continues to grow.

Open-circuit failure is one of the most common issues for switches in converters. By monitoring the signals of the operating system, malfunctions should be identified with as accurate classification and position information as possible [3]. Traditional methods of fault detection analysis depend more on cognitive-based comparisons of fault and normal signals. The open-circuit fault detection method for three-level T-type converters proposed in [4] uses the grid current amplitude and phase angle as the recognition parameters. An alternative approach presented in [5] uses irregular variations of neutral-point current, switching states and phase current for

The associate editor coordinating the review of this manuscript and approving it for publication was Okyay Kaynak<sup>1</sup>.

fault detection. In [6], the authors proposed an open-circuit fault diagnosis method for a four-wire T-type converter considering both unbalanced load and unbalanced input voltage conditions. The method was performed by decomposing the voltage into positive-, negative- and zero sequence components, and the faults were located by the law of positive- and negative-sequence voltage error offsets and zero sequence voltage error offsets. In [7], an open-circuit fault detection algorithm in single-phase 3L-NPC converters was presented. The grid voltage, the DC side voltages and the switching states were used to build a mixed logical dynamic model to estimate the grid current. By subtracting the estimated value from the measured current, the residual was calculated and used for fault diagnosis. Most analytical diagnosis procedures are intuitive, but the design can be overly complex [8]–[11]. However, in converters with five or more levels, the traditional methods appear incapable of fast and accurate fault detection since the complexity considerably increases with the number of switches. In [12], an open-switch fault-tolerant operation was presented for a multichannel voltage-source five-level power converter by monitoring the voltage across flying capacitors. The variations in the rotor currents and DC link voltage are also used for fault diagnosis in grid/machine-connected mode. The speed of fault detection depends on the frequency of the grid/machine-side rotor current, which tends to be very low. In recent years, machine learning algorithms have become effective tools for fault diagnosis in complex environments [13]. Since the performance of an algorithm is greatly influenced by the available data and features, feature engineering is widely used to extract characteristics from raw data and improve the corresponding fault recognition results. The authors of [14] proposed a concept that exploits the features extracted from circuit responses instead of component parameters for circuit health estimation using a kernel learning technique. In [15], a fault diagnosis method for a wind turbine planetary gearbox was developed. A stacked denoising autoencoder technique was adopted to learn robust and distinguishable features from measured signals, and a least squares support vector machine was employed for fault identification.

For multilevel converters, signals such as the phase or amplitude of currents and voltages can be used for fault detection. However, the direct use of these signals lacks identifiability and increases the data processing burden. Feature extraction is used to reduce the dimensions of large-scale data and summarize new characteristics with accurate classification. Time-domain features are effective indicators that reflect the state of operation and can be used individually or in combination for fault detection [16]. Designing classifiers with many unrelated features will result in considerable computational complexity and poor classification performance. Therefore, filtering redundant parameters through feature selection to further reduce dimensionality has very important practical significance [17]–[19].

Traditional feature selection methods are usually based on expert experience or enumeration. If the number of candidate

parameters is large, the approach may be too time consuming to find the optimal feature subset. In recent years, fault classification based on pattern recognition has been the main method for intelligent fault diagnosis. The validity of the pattern features directly affects the design and performance of the classifier. The amount of raw data obtained from the detected signals is considerable. To effectively perform classification and recognition tasks, the original data must be selected or transformed to obtain the essential characteristics that can best reflect the differences among modes. Optimization of the raw data is mainly achieved by dimensionality reduction, that is, conversion of a data space with a high dimension into one with a low dimension. The filter, wrapper and hybrid modes were introduced to reduce dimensionality and perform effective feature selection [20]. A filter uses a guideline of feature importance for the target attributes to weight and rank all features in the pool; this method is efficient but limited in accuracy [21]. A wrapper conducts feature selection based on specific evaluation criteria (mostly accuracy) and determines the optimal feature subset accordingly [22], [23]; this approach is high in accuracy but low in computational efficiency. Although the hybrid method inherits the advantages of both the filter and the wrapper, it inevitably falls to suboptimal solutions under certain conditions [24].

In this paper, a complete fault diagnosis solution for electronic power converters is proposed. A five-level nested neutral-point piloted (NNPP) converter is used to build the system for fault detection. A feature selection algorithm combining a dependence-guided unsupervised feature selection (DGUFS) filter and a random forest feature selection (RFFS) wrapper is also proposed. The proposed DGUFS-RFFS is a hybrid filter-wrapper method that not only avoids suboptimal solutions but also accelerates the feature selection process and generally achieves better performance than traditional methods. DGUFS, as presented in [25], is used to select parameters and evaluate feature combinations of a given dimension. The RFFS wrapper is used to determine the optimal feature subset from different dimensions. The currents of the flying capacitors  $I_{Cfa1}$  and  $I_{Cfa2}$ , the phase current  $I_a$  and the phase voltage  $V_{ao}$  of the five-level NNPP converter are selected as the original signals. The time-domain features are introduced as the extracted features, and a multidimensional feature pool is constructed. The optimal feature combination is selected by the DGUFS-RFFS model and input to the classifier for fault recognition.

## II. BRIEF INTRODUCTION TO THE OPEN-CIRCUIT FAULTS OF AN NNPP CONVERTER

Multilevel converters are considered the most attractive solutions in high-voltage and high-power applications due to their advantages, such as a low common-mode output voltage, low harmonics, and the use of low-voltage semiconductor devices. The five-level NNPP converter, which was proposed by GE in [26], was developed by nesting two or more medium-voltage 3-level neutral-point piloted (NPP) cells. This topology is easily scalable in a reasonable voltage range

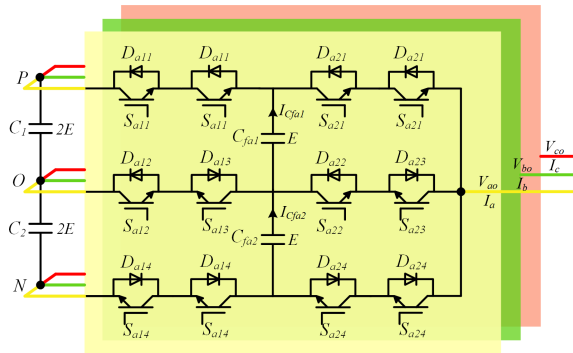


FIGURE 1. Topology of the five-level NNPP converter.

and at output voltage levels, and this approach provides a small filter, high power density and high efficiency, as shown in Fig. 1.

**A. MODULATION STRATEGY OF THE FIVE-LEVEL NNPP CONVERTER**

The five-level NNPP converter generates five-level phase voltages and nine-level phase-phase voltages. The output phase-phase voltages are supposed to be  $2E, E, 0, -E$  and  $-2E$ . The nested structure requires the modulation strategy to effectively balance the DC-side capacitor voltages and the flying capacitor voltages. The authors in [27] proposed an optimized space vector pulse width modulation (SVPWM) algorithm based on  $gh$  coordinates for this topology.

Taking phase A as an example, as shown in Fig. 1, each group of  $(S_{a11}, S_{a11}), (S_{a14}, S_{a14}), (S_{a21}, S_{a21})$  and  $(S_{a24}, S_{a24})$  should work under the same operating status. Each group of  $(S_{a11}, S_{a14}), (S_{a21}, S_{a24}), (S_{a11}, S_{a12}, S_{a13}), (S_{a21}, S_{a22}, S_{a23}), (S_{a12}, S_{a13}, S_{a14})$  and  $(S_{a22}, S_{a23}, S_{a24})$  cannot work in conduction mode simultaneously, thus preventing the formation of a short circuit. Each group of  $(S_{a11}, S_{a13}), (S_{a21}, S_{a23}), (S_{a12}, S_{a14})$  and  $(S_{a24}, S_{a24})$  should work in a complementary state to reduce the switching frequency and to ensure that the switching sequence is unaffected by dead zones or the direction of the load current. Therefore,  $S_{a11}, S_{a12}, S_{a21}$  and  $S_{a22}$  are four switches that need to be controlled independently.  $S_{a11}$  and  $S_{a13}, S_{a12}$  and  $S_{a14}, S_{a21}$  and  $S_{a23}, S_{a22}$  and  $S_{a24}$  should work in a complementary manner. The output voltage and the corresponding control strategy are shown in Table 1.

**B. OPEN-CIRCUIT FAULT ANALYSIS OF THE FIVE-LEVEL NNPP CONVERTER**

Fault diagnosis studies of the five-level NNPP converter are limited. In [28], a fault-tolerant solution for the five-level

TABLE 1. The control strategy of the 5-L NNPP converter.

Switch	$2E$	$E$	$E$	0	0	0	$-E$	$-E$	$-2E$
$S_{a11}$	1	0	1	0	0	1	0	0	0
$S_{a12}$	1	1	1	0	1	1	0	1	0
$S_{a21}$	1	1	0	1	0	0	0	0	0
$S_{a22}$	1	1	1	1	1	0	1	0	0

TABLE 2. The distorted output voltages when  $S_{a11}$  fails to open.

Switch	$2E$	$E$	$E$	0	0	0	$-E$	$-E$	$-2E$
$S_{a11}$	1	0	1	0	0	1	0	0	0
$S_{a12}$	1	1	1	0	1	1	0	1	0
$S_{a21}$	1	1	0	1	0	0	0	0	0
$S_{a22}$	1	1	1	1	1	0	1	0	0

TABLE 3. The distorted output voltages when  $S_{a11}$  and  $S_{a12}$  fail to open.

Switch	$2E$	$E$	$E$	0	0	0	$-E$	$-E$	$-2E$
$S_{a11}$	1	0	1	0	0	1	0	0	0
$S_{a12}$	1	1	1	0	1	1	0	1	0
$S_{a21}$	1	1	0	1	0	0	0	0	0
$S_{a22}$	1	1	1	1	1	0	1	0	0

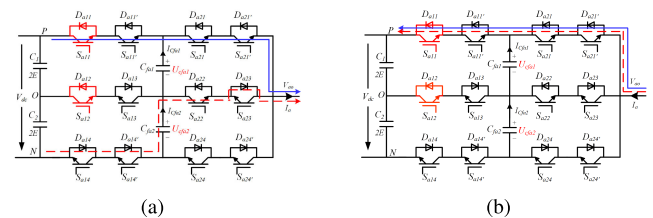


FIGURE 2. Change in the current path. The blue solid line represents the current path of the fault-free circuit, and the red dotted line represents the current path when  $S_{a11}$  and  $S_{a12}$  fail to open simultaneously: (a)  $I_a > 0$ . (b)  $I_a < 0$ .

TABLE 4. Classification of the open-circuit faults.

Status Label	Faulty Switch	Status Label	Faulty Switch	Status Label	Faulty Switch	Status Label	Faulty Switch
1	none						
2	$S_{a11}$	11	$S_{a11}, S_{a12}$	20	$S_{a21}, S_{a23}$	29	$S_{a13}, S_{a23}$
3	$S_{a12}$	12	$S_{a11}, S_{a13}$	21	$S_{a21}, S_{a14}$	30	$S_{a13}, S_{a14}$
4	$S_{a13}$	13	$S_{a11}, S_{a22}$	22	$S_{a21}, S_{a24}$	31	$S_{a13}, S_{a24}$
5	$S_{a14}$	14	$S_{a11}, S_{a23}$	23	$S_{a12}, S_{a13}$	32	$S_{a22}, S_{a23}$
6	$S_{a21}$	15	$S_{a11}, S_{a14}$	24	$S_{a12}, S_{a22}$	33	$S_{a22}, S_{a14}$
7	$S_{a22}$	16	$S_{a11}, S_{a24}$	25	$S_{a12}, S_{a23}$	34	$S_{a22}, S_{a24}$
8	$S_{a23}$	17	$S_{a21}, S_{a12}$	26	$S_{a12}, S_{a14}$	35	$S_{a23}, S_{a14}$
9	$S_{a24}$	18	$S_{a21}, S_{a13}$	27	$S_{a12}, S_{a24}$	36	$S_{a23}, S_{a24}$
10	$S_{a11}, S_{a12}$	19	$S_{a21}, S_{a22}$	28	$S_{a13}, S_{a22}$	37	$S_{a14}, S_{a24}$

NNPP topology was proposed, and it identified both the short- and open-circuit faults of a single IGBT. In this paper, the open-circuit failures of a single IGBT and dual IGBTs are analyzed. The probability of simultaneous failures of more than two IGBTs is relatively low; therefore, this topic is not discussed in this paper. As discussed in [28], when a single IGBT fails to open, taking  $S_{a11}$  as an example, the output voltages marked in red are distorted, as shown in Table 2.

For the simultaneous failure of two IGBTs, taking  $S_{a11}$  and  $S_{a12}$  as examples; when both of the switches fail to open, the output voltages marked in red are distorted, as shown in Table 3. Assuming that the breakdown of the IGBTs does not affect the reverse diode, Fig. 2 shows the change in the current path when the phase current  $I_a$  is positive or negative and when  $2E$  is the output phase voltage.

A total of 36 faulty states and one fault-free state are included for analysis and identification. The failures with classification labels are summarized in Table 4.

TABLE 5. Formulas of the time-domain feature extraction parameters.

$Ku$	$Ku = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4 / \left( \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^2$	A measure of the peakedness of a distribution that represents the combined weight of the tails relative to that of the rest of the distribution
$Sk$	$Sk = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3 / \left( \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^{3/2}$	A measure of the asymmetry of a distribution; a distribution with a negative skew has a longer left tail, and a positive skew has a longer right tail
$RMS$	$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2(t)}$	The square root of the arithmetic mean of the squares of the values
$Cf$	$Cf =  x_{peak}  / \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$	The instantaneous peak amplitude/RMS value; a statistical index identifying the impact of the signal
$Ff$	$Ff = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2(t) / \frac{1}{N} \sum_{i=1}^N x_i(t)}$	The ratio of the RMS value to the average, which cannot be smaller than 1 but theoretically has no upper limit when influenced by large deviations from the average

### III. FEATURE EXTRACTION AND SELECTION BASED ON THE HYBRID DGUFS-RFFS METHOD

In this paper, the time-domain signal parameters that are more representative of the fault characteristics are selected for feature extraction. Once the feature parameters are extracted and a feature space is constructed, feature selection is implemented to further achieve dimensionality reduction. The importance of the features is determined, and the redundant information is deleted. The proposed hybrid DGUFS-RFFS method is used to optimize the feature set, ensuring that the remaining features are reliable, relatively uncorrelated, retain as much information as possible, and reduce the amount of data as much as possible.

#### A. TIME-DOMAIN FEATURE EXTRACTION ALGORITHMS

The kurtosis ( $Ku$ ), skewness ( $Sk$ ), root mean square metric ( $RMS$ ), crest factor ( $Cf$ ) and form factor ( $Ff$ ) focus more on the extremes of a data set than on the average and are indicators of faulty signals. These individual feature parameters and combinations of parameters have different recognition effects for fault diagnosis. The formulas of the above feature parameters are listed in Table 5.

We added all these features into a feature pool as candidate parameters and then used certain rules to select the feature subsets that yielded the best performance (e.g., the highest fault recognition rate or the fastest detection speed). These feature subsets contain different features derived from different signals and of different dimensions; thus, different options will affect the accuracy and efficiency of the fault diagnosis process. Selecting the optimal feature subset has always been a concern of scholars.

#### B. DGUFS-RFFS FEATURE SELECTION METHOD

##### 1) HYBRID FILTER-WRAPPER METHODS

Feature selection is crucial for high-dimensional data classification problems. Dash and Liu proposed the basic framework of feature selection in 1997, which consisted of the following four parts: the generation of feature subsets, the evaluation of feature subsets, the stopping criteria and the verification of the results, as shown in Fig. 3 [29].

The commonly used feature subset generation methods include filters and wrappers. A filter is independent of the

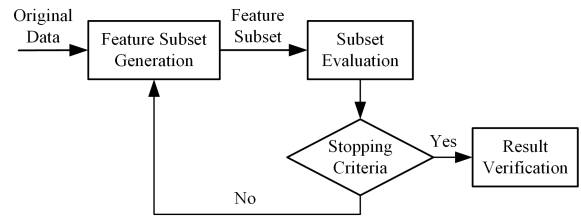


FIGURE 3. The basic framework of feature selection.

subsequent learning algorithms, uses evaluation criteria or evaluation functions to enhance the correlations among certain features and categories and reduces the correlations among other features. This approach has been widely used due to its fast processing speed and high computing efficiency. Unlike filters, wrappers use the accuracy of the subsequent classifiers as an evaluation index and are included in the learning algorithm. The feature subset selected by a wrapper is relatively small in size and high in prediction accuracy, although the algorithm has high complexity and low execution efficiency.

Combining the efficiency of a filter with the high accuracy of a wrapper can yield complementary advantages. Key feature recognition effectively reduces the dimensionality of the feature space while maintaining the accuracy of subsequent classification algorithms. The flowchart is shown in Fig 4.

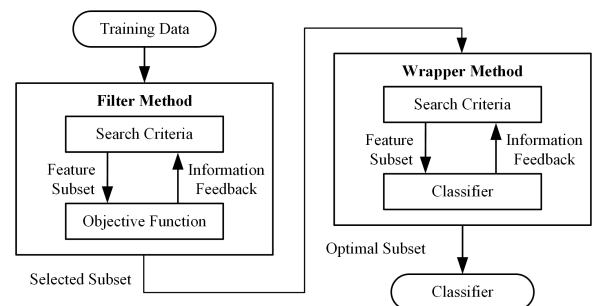


FIGURE 4. The hybrid filter-wrapper method.

##### 2) DGUFS FILTER

Suppose that the feature pool consists of  $d$  candidate features. The commonly used hybrid filter-wrapper methods mostly evaluate the importance of all  $d$  features based on certain criteria and then sort and select  $m$  prominent features from  $d$  to obtain the best classification results; this approach inevitably tends to fall to a suboptimal solution.

To solve this problem, a DGUFS method is used as a filter to select features. The aim of the algorithm is to directly select the most discriminative feature subset from  $d$  features for any given  $m$  instead of evaluating the importance of all  $d$  features. The selection based on this algorithm is targeted, which not only reduces the required computations but also enhances the recognition ability. The objective function is designed by combining two weighted dependence-guided

terms to achieve the maximum similarity among the original data  $X$ , the cluster label matrix  $V$  and the selected feature  $Y$ .

The overall DGUFS model can be expressed as:

$$\begin{aligned} \min_{Y, L} & -\beta \text{Tr}(S^T L) - (1 - \beta) \text{Tr}(Y^T Y H V^T V H) \\ \text{s. t. } & \|X - Y\|_{2,0} = d - m, \|Y\|_{2,0} = m, \\ & V \in \Omega, \quad L = V^T V, \text{rank}(L) = c, \\ & L \succeq 0, \quad L \in \{0, 1\}^{n \times n}, \text{diag}(L) = I. \end{aligned} \quad (1)$$

where  $X$  is a  $d \times n$  original data matrix with  $n$  samples,  $V$  is a  $c \times n$  cluster label matrix of  $X$ ,  $c$  is the cluster indicator,  $Y$  is the selected feature,  $\Omega$  is the candidate set of cluster label matrices that classify data into  $c$  groups,  $L$  is the linear kernel matrix of  $V$ ,  $S$  is the similarity matrix,  $H = I - \frac{1}{n}ee^T$  is the centering matrix, and  $e$  is an  $n$ -dimensional column vector with element values equal to 1.  $\text{Tr}$  denotes the trace, and  $\beta \in (0, 1)$  is a regularization parameter. Additionally,  $L, I, H \in \mathbb{R}^{n \times n}$ .

In the objective function, the first dependence-guided term  $-\text{Tr}(S^T L)$  is used to increase the dependence of the desired label matrix  $L$  on the original data  $X$  and is designed based on the geometrical structure of the data and the associated discriminative information; the second dependence-guided term  $-\text{Tr}(Y^T Y H V^T V H)$  is designed based on the matrix form of the Hilbert-Schmidt independence criterion (HSIC) to enhance the dependence of selected feature  $Y$  on the desired label matrix  $V$ . Both of the terms are designed using the trace norm (abbreviated as  $\text{Tr}$ ), and the method is based on the  $l_{2,0}$ -norm equality constraints, thereby effectively avoiding overfitting problems.

The DGUFS method increases the interdependence among selected features, raw data and cluster labels.  $m$  is considered in the process of obtaining the optimal subset. The features are no longer evaluated separately, and the subsets are treated as a whole considering the correlations and redundancy among features. The algorithm outperforms many of the leading methods of sparse learning-based unsupervised feature selection.

### 3) RFFS WRAPPER

The heuristic search strategy is one of the main wrapper strategies, and such methods include the individual optimal feature search strategy, sequence forward selection method, sequence backward selection method, etc., [30], [31]. Individual optimal feature search strategies have the advantages of low time complexity and high operational efficiency, and they are widely used for high-dimensional data sets. The sequence forward selection method is efficient, but the correlations among features to be added and the selected feature set are not considered. Once the features are added, they will not be deleted, which will result in redundancy in the feature subset. The sequence backward selection method involves an elimination algorithm based on a complete feature set, which requires many computations. However, this approach has

displayed good performance in practical applications because it considers the redundancy among features.

In this paper, an individual optimal feature selection wrapper based on the random forest (RF) approach was presented in combination with DGUFS to evaluate and compare the optimal feature subsets and acquire the final feature subset with the best recognition performance.

The RF, which was proposed by Leo Breiman in 2001, is an integrated classifier composed of a set of decision tree classifiers  $h(X, \theta_k)$ ,  $k = 1, 2, \dots, K$ , where  $\theta_k$  is a random vector that follows an independent distribution, and  $K$  represents the number of decision trees in the RF [32]. With a given independent variable  $X$ , each decision tree classifier determines the optimal classification result by voting.

Given a set of classifiers  $h_1(x), h_2(x), \dots, h_k(x)$ , the training set of each classifier is randomly sampled from the original randomly distributed data vector  $Y, X$ , and the margin function is defined as:

$$mg(X, Y) = av_k I(h_k(X) = Y) - \max_{j \neq Y} av_k I(h_k(X) = j) \quad (2)$$

where  $I(\cdot)$  is an indicator function.

The margin function is used to measure the degree to which the average number of correct classifications exceeds the average number of incorrect classifications. The larger the margin is, the more reliable the classification prediction.

The generalization error is defined as:

$$PE^* = P_{X, Y}(mg(X, Y) < 0) \quad (3)$$

where the subscripts  $X$  and  $Y$  represent the probability and  $P$  spans the  $X$  and  $Y$  spaces.

In an RF, when there are sufficient decision tree classifiers,  $h_k(X) = h(X, \theta_k)$  obeys the strong law of large numbers.

As the number of decision trees in the RF increases, all sequences  $\theta_1, \theta_2, \dots, \theta_k, PE^*$  converge almost everywhere:

$$P_{X, Y}(P_\theta(h(X, \theta) = Y) - \max_{j \neq Y} P_\theta(h(X, \theta) = j) < 0) \quad (4)$$

Formula (4) shows that RFs do not cause overfitting problems as the number of decision trees increases but may increase generalization errors within a certain limit.

The base classifier in the RF method proposed in this paper chooses the classification and regression tree (CART) algorithm. Assuming that the selected feature dimension is  $m$ , the complete feature set dimension is  $d$ , the number of training samples is  $n$ , and the time complexity is represented by  $O(\cdot)$ , the time complexity of the RF algorithm can be approximated as  $O(mn \log n^2)$ . In our experiment, the RFFS needs to be executed  $d$  times, and the total time complexity of the algorithm can be approximated as  $O(dmn \log n^2)$ . The time complexity of the RFFS algorithm is approximately linearly related to  $m$  and  $d$  and related to the approximate square of  $n$ .

### 4) THE DGUFS-RFFS METHOD

In this paper, the DGUFS method is used as a filter to determine the optimal solution of the subset in each feature

dimension, and the RFFS wrapper then is used to determine the dimension of the feature subset that yields the optimal solution. By combining the wrapper with the DGUFS algorithm, the RFFS wrapper only needs to implement individual optimal selection strategies for feature subsets of different dimensions  $m$  instead of performing  $C_d^m$  traverse optimization within the feature subset, thus considerably reducing the computational overhead. The complexity of the algorithm is effectively reduced in this approach. The feature subset obtained by the proposed DGUFS-RFFS method can achieve the optimal prediction effect within a given range of feature dimensions. The accuracy, detection efficiency and computational cost can be well balanced and flexibly adjusted according to the demand.

#### IV. FAULT DETECTION WITH THE FIVE-LEVEL NNPP CONVERTER

##### A. THE FRAMEWORK OF OPEN-CIRCUIT FAULT DIAGNOSIS USING DGUFS-RFFS

A general framework for fault identification is proposed, and it includes five modules, namely, original signal acquisition, signal preprocessing, feature extraction, feature selection and fault classification. Feature preprocessing includes sampling and abnormal sample cleaning.

The current of the flying capacitors  $I_{Cfa1}$  and  $I_{Cfa2}$ , the phase current  $I_a$  and the phase voltage  $V_{ao}$  are monitored and used for fault detection. The information contained in the signals above can instantly reflect changes in the modulation strategy and operating state; thus, these parameters are effective indicators for fault diagnosis. Five feature parameters, including  $Ku$ ,  $Sk$ ,  $RMS$ ,  $Cf$  and  $Ff$ , are extracted to form a feature pool with 20 features.

The DGUFS-RFFS method is used for feature selection, and another two feature selection methods are explored for comparison. MCFS (multiclustor feature selection) is a popular unsupervised learning algorithm with excellent classification ability based on manifold and  $l1$ -norm regularization models. The method uses the  $l1$ -norm to measure the ability of each feature to distinguish different categories, and the selection can maintain the various clustering structures of the data. The manual selection method is based on enumeration. Referring to the experience of experts and analyses of actual situations, the feature subset with the best possible classification effect is selected through hundreds of experiments. Although this method is time consuming and labor intensive, it is widely used in practice and can achieve good fault recognition results; thus, it is superior to many other prevailing algorithms.

The features selected by the above methods are input into the RF classifier for fault recognition. The DGUFS, MCFS and manual selection methods are compared in terms of the fault recognition rate, training time and testing time and are further analyzed based on evaluation indexes.

The flowchart of the complete solution is shown in Fig. 5.

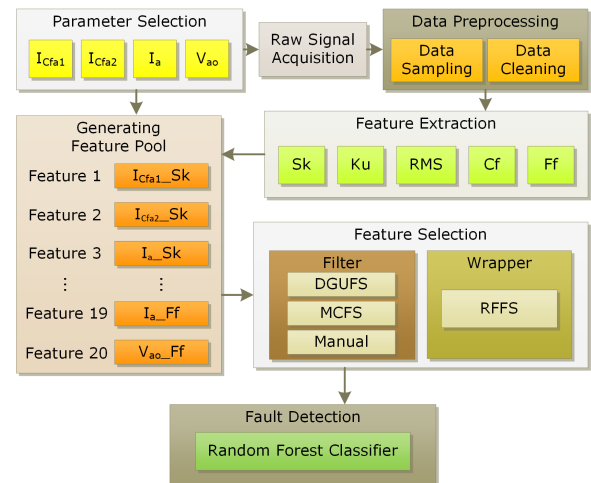


FIGURE 5. Flowchart of the complete solution for fault diagnosis.

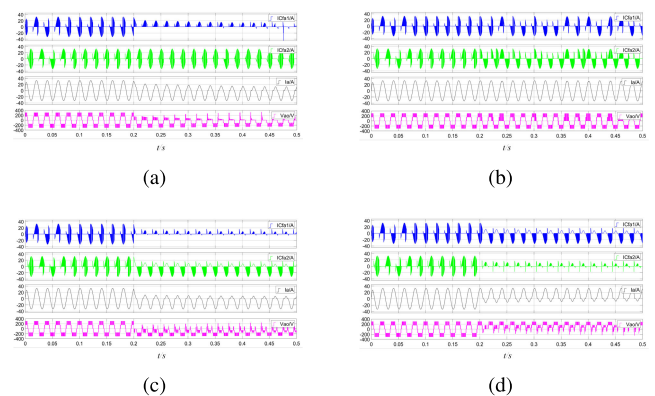


FIGURE 6. Waveforms of  $I_{Cfa1}$ ,  $I_{Cfa2}$ ,  $I_a$  and  $V_{ao}$  under open-circuit faults in the five-level NNPP converter: (a)  $S_{a11}$  fails to open. (b)  $S_{a12}$  fails to open. (c)  $S_{a11}$  and  $S_{a12}$  fail to open. (d)  $S_{a13}$  and  $S_{a14}$  fail to open.

##### B. IMPLEMENTATION AND SIMULATION

###### 1) SIMULATION OF AN OPEN-CIRCUIT FAULT FOR A FIVE-LEVEL NNPP CONVERTER

The simulation is conducted using MATLAB 2018 (a), and four failure modes are selected as representatives to simulate the current and voltage waveform changes in four failure states. The four modes include (a)  $S_{a11}$  fails to open, (b)  $S_{a12}$  fails to open, (c)  $S_{a11}$  and  $S_{a12}$  fail to open, and (d)  $S_{a13}$  and  $S_{a14}$  fail to open. Fig. 6 displays the simulation results for the waveforms of  $I_{Cfa1}$ ,  $I_{Cfa2}$ ,  $I_a$  and  $V_{ao}$  under the four faulty circumstances above. Failures occur at  $T = 0.2$  s.

###### 2) ANALYSIS OF FEATURE EXTRACTION AND SELECTION

After the faulty signals are obtained, time-domain feature extraction and dimensionality reduction are performed. Taking  $Sk$ ,  $Ku$  and  $Cf$  for instance, Fig. 7 shows the spatial distributions of the three feature parameters extracted from four original signals  $I_{Cfa1}$ ,  $I_{Cfa2}$ ,  $I_a$  and  $V_{ao}$ , including nine operating states of one fault-free state (represented as OK)

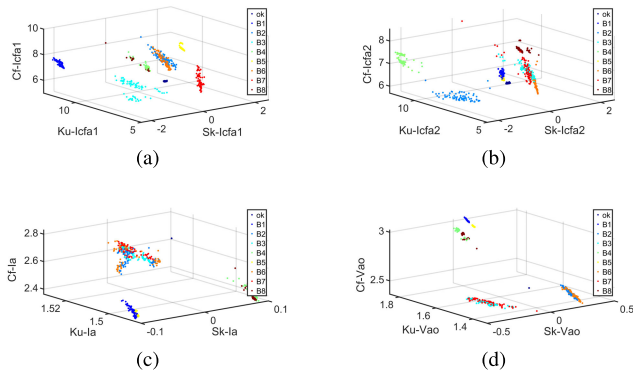


FIGURE 7. Spatial distributions of  $Sk$ ,  $Ku$  and  $Cf$  extracted from the original signals: (a) Features of  $I_{cfa1}$ . (b) Features of  $I_{cfa2}$ . (c) Features of  $I_a$ . (d) Features of  $V_{a0}$ .

TABLE 6. Device parameters of the converter.

DC voltage	100 V
DC-link capacitance	13600 $\mu F$
flying capacitance	6800 $\mu F$
load inductance	10 $\mu H$
load resistance	12 $\Omega$

and eight faulty states when the open-circuit failure of a single IGBT (from  $S_{a11}$  to  $S_{a24}$ , represented as  $B1$  to  $B8$ ) occurs.

The above spatial distributions are only a few representatives of many combinations of feature parameters. The spatial distribution and concentration of different features of the same signal exhibit differences. Different feature combinations will have different effects on fault identification.

## V. EXPERIMENTAL RESULTS AND DISCUSSION

### A. EXPERIMENTAL SETUP

An experimental prototype of the five-level NNPP converter was built, as shown in Fig. 8. The semiconductor power switches adopted were Infineon FF100R12RT4 switches. The system was controlled by a TI TMS320F28335 digital signal processor (DSP) and ACTEL A3P250 field-programmable gate array (FPGA). The device parameters of the converter were set as follows in Table 6.

Fig. 9 shows the output waveforms of phase A when the modulation ratio  $M = 0.9$ ; the line voltage  $V_{ab}$ , the phase voltage  $V_{a0}$ , the DC link voltages  $U_{C1}$  and  $U_{C2}$ , the flying capacitor voltages  $U_{Cfa1}$  and  $U_{Cfa2}$  and the phase current  $I_a$  are shown. The DC link voltages and the flying capacitor voltages are balanced by the control strategies proposed in [25], [26].

### B. EXPERIMENTAL DESIGN AND IMPLEMENTATION

The signals of the open-circuit faults of the NNPP converter are acquired, including those for 1 fault-free state, 8 single IGBT failures and 28 double IGBT failures. The frequency of the output voltage  $f = 50$  Hz. The raw sampled signals obtained include 480,000 values used for training and 80,000 values used for testing. Each feature extraction was

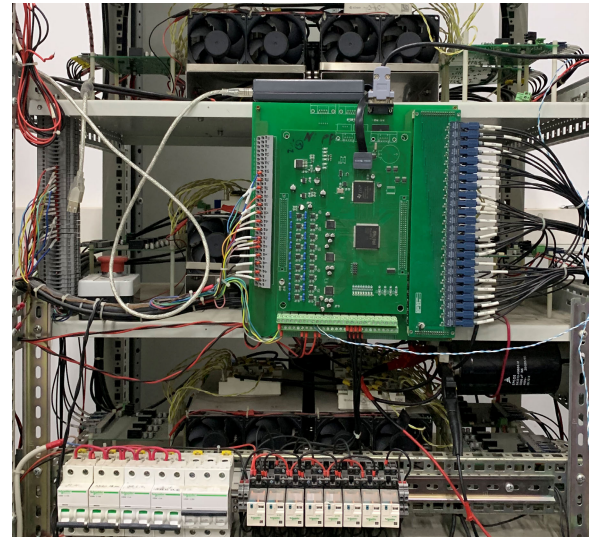


FIGURE 8. Experimental prototype of the five-level NNPP converter.

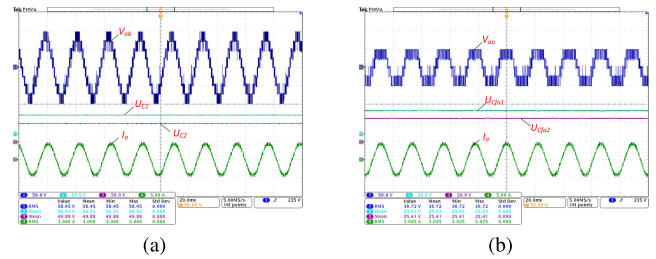


FIGURE 9. The output waveforms of the five-level NNPP converter under normal condition when  $M=0.9$ : (a)  $V_{ab}$ ,  $U_{C1}$ ,  $U_{C2}$  and  $I_a$ . (b)  $V_{a0}$ ,  $U_{Cfa1}$ ,  $U_{Cfa2}$  and  $I_a$ .

based on 4000 sampled values, which constituted a training feature set with a maximum dimension of  $120 \times m$  and a test feature set with a dimension of  $20 \times m$ , where  $m$  is the number of selected features ( $m = 1, 2, \dots, 20$ ). Table 7 shows the experimental results when the training feature set dimension is  $40 \times m$ , including the fault recognition accuracy, training time, and testing time. When the DGUFS-RFFS method is adopted, the accuracy increases with the number of features and reaches a peak when the feature number is 7. Then, the fault recognition rate remains relatively stable as the number of features continues to increase and finally decreases when the number exceeds 10. The main reason for the decrease in accuracy is the introduction of redundant information and overfitting. The arithmetic means and the root mean square error (RMSE) of the results were calculated based on 20 repeated experiments. The MCFS and the manual feature selection method are used for comparison with the DGUFS. The experimental results show that the hybrid DGUFS-RFFS method is generally superior to the other two methods, with higher fault recognition rates and, in most cases, shorter training and testing times. Using the three methods, the number of features required to reach the

**TABLE 7. Comparison of experimental results for three different feature selection methods.**

Feature Number	DGUFS-RFFS (AVG±RMSE)			MCFS-RFFS (AVG±RMSE)			MANUAL-RFFS (AVG±RMSE)		
	ACC (%)	Training Time(s)	Testing Time(s)	ACC (%)	Training Time(s)	Testing Time(s)	ACC (%)	Training Time(s)	Testing Time(s)
1	39.46±0	1.81±0.66	1.14±0.25	32.97±5.6E-15	1.54±0.062	1.08±0.18	32.97±5.55E-15	1.71±0.0894	1.10±0.20
2	55.97±0.35	3.41±0.27	1.08±0.22	67.05±0.57	3.51±0.22	1.04±0.15	66.96±0.43	3.38±0.20	0.97±0.14
3	78.24±1.2	3.28±0.23	1.05±0.17	71.22±1.1	3.65±0.25	1.09±0.13	77.44±0.36	2.77±0.16	0.97±0.13
4	92.87±0.28	2.67±0.14	0.98±0.17	71.44±1.1	3.40±0.17	1.13±0.18	88.39±0.47	2.58±0.11	1.03±0.16
5	93.24±0.44	3.01±0.17	0.96±0.18	76.72±0.77	3.50±0.17	1.08±0.13	86.08±0.37	2.94±0.11	0.95±0.15
6	94.76±0.74	2.96±0.18	0.98±0.19	84.18±2.0	3.27±0.15	1.05±0.12	87.86±0.62	3.02±0.13	0.96±0.14
7	96.02±0.98	2.97±0.16	0.87±0.14	82.47±1.5	3.40±0.16	0.99±0.13	86.37±0.84	2.94±0.17	0.99±0.14
8	94.20±0.39	3.01±0.18	0.95±0.18	82.74±1.5	3.27±0.22	0.99±0.14	87.26±1.3	3.02±0.15	0.96±0.13
9	93.61±1.0	3.00±0.14	0.94±0.11	84.45±1.1	3.12±0.27	1.02±0.13	85.95±0.67	2.97±0.12	0.98±0.16
10	92.55±1.8	3.35±0.18	0.95±0.09	84.52±1.5	3.47±0.23	0.93±0.16	84.12±0.61	3.23±0.13	0.94±0.14
11	92.16±1.8	3.41±0.15	0.88±0.15	86.02±1.6	3.32±0.19	0.99±0.19	84.24±0.72	3.22±0.12	1±0.11
12	87.98±1.5	3.42±0.21	0.95±0.12	84.38±1.7	3.49±0.16	0.98±0.16	84.03±0.62	3.12±0.091	0.92±0.15

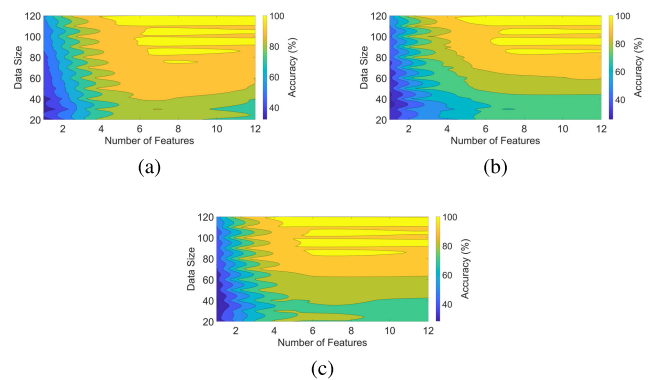
**TABLE 8. Optimum feature subsets based on different feature selection methods.**

Number of features	DGUFS-RFFS	MCFS-RFFS	MANUAL-RFFS
7	$I_{Cfa1-Ff}$	$I_{Cfa1-Sk}$	$I_{Cfa1-Sk}$
	$V_{ao-RMS}$	$I_{Cfa2-Sk}$	$I_{Cfa2-Sk}$
	$I_{Cfa2-Ku}$	$I_a-Sk$	$I_{Cfa1-Ku}$
	$I_{Cfa1-Ku}$	$V_{ao-Sk}$	$I_{Cfa2-Ku}$
	$I_a-RMS$	$I_{Cfa1-Ku}$	$I_{Cfa1-Cf}$
	$I_{Cfa2-RMS}$	$I_{Cfa2-Ku}$	$I_{Cfa2-Cf}$
	$I_{Cfa1-RMS}$	$I_a-Ku$	$I_{Cfa1-RMS}$
8	$I_{Cfa1-Ff}$	$I_{Cfa1-Sk}$	$I_{Cfa1-Sk}$
	$V_{ao-RMS}$	$I_{Cfa2-Sk}$	$I_{Cfa2-Sk}$
	$I_{Cfa2-Ku}$	$I_a-Sk$	$I_{Cfa1-Ku}$
	$I_{Cfa1-Ku}$	$V_{ao-Sk}$	$I_{Cfa2-Ku}$
	$I_a-RMS$	$I_{Cfa1-Ku}$	$I_{Cfa1-Cf}$
	$I_{Cfa2-RMS}$	$I_{Cfa2-Ku}$	$I_{Cfa2-Cf}$
	$I_{Cfa1-RMS}$	$I_a-Ku$	$I_{Cfa1-RMS}$
	$I_{Cfa2-Cf}$	$V_{ao-Ku}$	$I_{Cfa2-RMS}$

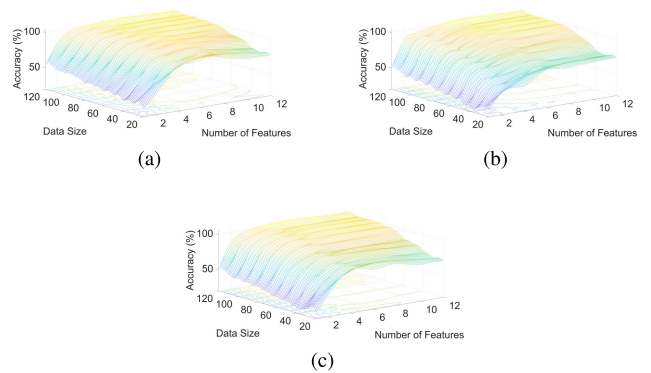
highest accuracy is 7, 11, and 4, with accuracies of  $96.02 \pm 0.98$ ,  $86.02 \pm 1.6$ , and  $88.39 \pm 0.47$ , respectively.

Fig. 10 and Fig. 11 show the relationships among the number of selected features, the training data set size and the fault detection accuracy using three feature selection methods. Notably, when m is the same in each method, the fault detection accuracy using the DGUFS-RFFS method is significantly higher than that of MCFS and the manual method. Figs. 12 and 13 illustrate the relationship among the number of selected features, the training data set size and the training time. The training speed is not sensitive to the number of features but is greatly influenced by the training data set size. For the same m, the training time of the DGUFS-RFFS method is slightly shorter than that of the other two methods.

Table 8 lists the optimal feature subsets selected by the three methods when the number of features is 7 and 8. The raw signals and the feature parameters can be selected according to different situations. In this paper, a feature pool with 20 candidate features is constructed using 4 circuit signals and 5 feature parameters. The DGUFS-RFFS method greatly shortens the computational process and provides the possibility for numerous attempts based on multiple features. In the case of large numbers of candidate signals and features, the superiority of the solution can be clearly illustrated.



**FIGURE 10. Contour map of the relationship among the fault detection accuracy, the number of selected features and the size of the training data set: (a) DGUFS-RFFS method. (b) MCFS-RFFS method. (c) Manual-RFFS method.**



**FIGURE 11. 3D mesh grid of the relationship among the fault detection accuracy, the number of selected features and the size of the training data set: (a) DGUFS-RFFS method. (b) MCFS-RFFS method. (c) Manual-RFFS method.**

Not only can this approach greatly reduce the computational cost, but it can also provide an improved and stable fault recognition rate compared to the rates obtained by other state-of-the-art algorithms. The proposed solution provides good extension and generalization ability. Although the DGUFS-RFFS method achieves superior performance, this result does not mean that there is no better choice.



TABLE 9. Comparison of evaluation indicators for three different feature selection methods.

Feature Number	DGUFS-RFFS		MCFS-RFFS		MANUAL-RFFS	
	NMI	ARI	NMI	ARI	NMI	ARI
1	0.77±0.0004	0.49±0.0007	0.63±5.63E-05	0.32±5.34E-05	0.63±1.11E-16	0.32±1.11E-16
2	0.87±0.0038	0.63±0.0083	0.88±0.0031	0.66±0.0052	0.88±0.002	0.66±0.0036
3	0.93±0.0043	0.77±0.0173	0.9±0.0036	0.69±0.010	0.93±0.0033	0.77±0.0064
4	0.98±0.0018	0.90±0.0048	0.91±0.0036	0.71±0.0098	0.97±0.0037	0.88±0.0081
5	0.98±0.0025	0.91±0.0073	0.93±0.0014	0.74±0.005	0.97±0.0013	0.86±0.0031
6	0.98±0.0015	0.93±0.0062	0.96±0.0071	0.84±0.023	0.97±0.0030	0.87±0.0087
7	0.99±0.0027	0.95±0.0112	0.96±0.0051	0.83±0.014	0.97±0.0030	0.87±0.0088
8	0.98±0.0005	0.93±0.003	0.96±0.0032	0.83±0.010	0.97±0.0031	0.87±0.011
9	0.98±0.0029	0.92±0.010	0.96±0.0042	0.84±0.012	0.97±0.0025	0.86±0.0067
10	0.98±0.0055	0.91±0.020	0.96±0.0042	0.84±0.013	0.96±0.0019	0.85±0.0036
11	0.98±0.0051	0.92±0.019	0.97±0.0042	0.86±0.014	0.96±0.0023	0.85±0.0063
12	0.97±0.0028	0.88±0.012	0.96±0.0043	0.85±0.014	0.96±0.0019	0.85±0.0052

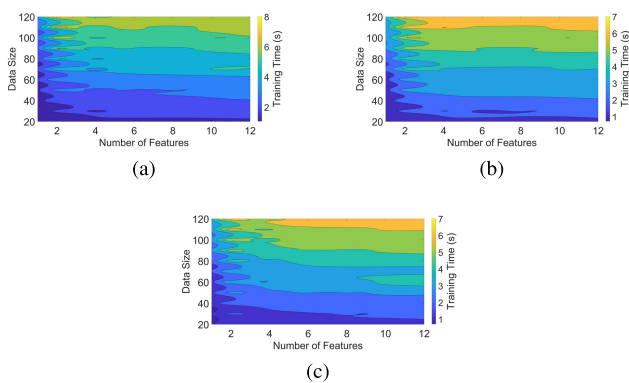


FIGURE 12. Contour map of the relationship among the training time, the number of selected features and the size of the training data set: (a) DGUFS-RFFS method. (b) MCFS-RFFS method. (c) Manual-RFFS method.

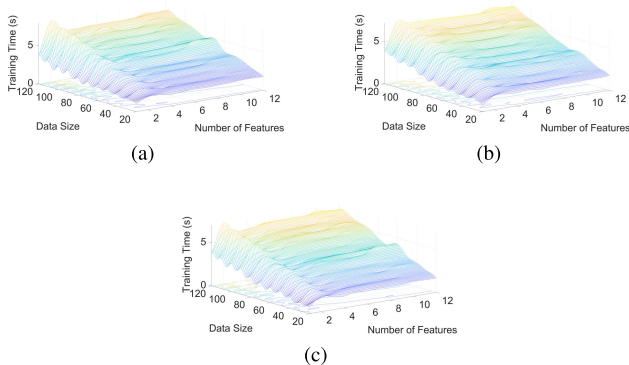


FIGURE 13. 3D mesh grid of the relationship among the training time, the number of selected features and the size of the training data set: (a) DGUFS-RFFS method. (b) MCFS-RFFS method. (c) Manual-RFFS method.

By selecting additional circuit signals or extracting other features, candidate feature parameters can be added to the feature pool to provide more alternatives for fault recognition. This issue can be discussed in future work.

Table 9 lists two of the important evaluation indexes for clustering: the mutual information index (NMI) and adjusted

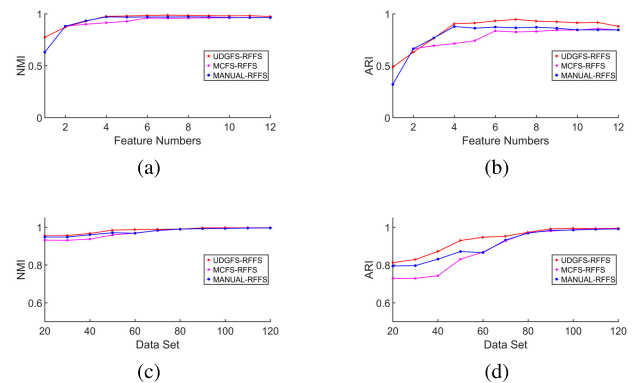


FIGURE 14. Comparison of the NMI and ARI for three feature selection methods: (a) NMI relative to the selected feature number. (b) ARI relative to the selected feature number. (c) NMI relative to the training data set size. (d) ARI relative to the training data set size.

rand index (ARI). Figure 14 shows a comparison of the NMI and ARI for the three feature selectors. When the number of features is small, compared with the DGUFS-RFFS-based NMI and ARI, the indexes based on the MCFS-RFFS and the MANUAL-RFFS methods are generally smaller, and as the number of features increases, the convergence rate of the indicators towards 1 is also slower. The NMI and ARI indicators based on the UDUFSS-RFFS method are generally stable and close to 1, indicating a satisfactory clustering effect.

## VI. CONCLUSION

In this paper, a complete fault detection solution for a five-level NNPP converter is proposed. Through signal acquisition, feature extraction, feature selection and classification, the effective diagnosis of 36 open-circuit faults is achieved. This method does not rely on past experience and does not require massive amounts of data for training. Failures can be discovered with limited data sets through signal processing and feature selection, which greatly shortens the time for fault diagnosis. The solution is flexible and can be easily used by other types of converters; moreover, this approach is not limited by topologies, the levels of converters or the device parameters in a given system. Different circuit signals

and feature parameters can be selected according to specific circumstances. The remarkable advantage of the solution is that an increase in the number of candidate features will not result in considerable computational complexity but will provide more choices for optimal feature subset selection. The optimal feature subset of a certain dimension can be directly determined by the DGUFS filter, thereby reducing the computational overhead of the subsequent RFFS wrappers. Compared with other leading feature selection algorithms, the proposed DGUFS-RFFS feature selector exhibits better recognition performance and efficiency. The effectiveness and practicality of the solution are verified by simulations and experiments.

## REFERENCES

- [1] F. W. Fuchs, "Some diagnosis methods for voltage source inverters in variable speed drives with induction machines a survey," in *Proc. 29th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, vol. 2, 2003, pp. 1378–1385.
- [2] Y. Song and B. Wang, "Survey on reliability of power electronic systems," *IEEE Trans. Power Electron.*, vol. 28, no. 1, pp. 591–604, Jan. 2013.
- [3] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.
- [4] S.-K. Baek, H.-U. Shin, S.-Y. Kang, C.-S. Park, and K.-B. Lee, "Open fault detection and tolerant control for a five phase inverter driving system," *Energies*, vol. 9, no. 5, p. 355, 2016.
- [5] H.-K. Ku, W.-S. Im, J.-M. Kim, and Y.-S. Suh, "Fault detection and tolerant control of 3-phase NPC active rectifier," in *Proc. IEEE Energy Convers. Congr. Exposit. (ECCE)*, Sep. 2012, pp. 4519–4524.
- [6] K. Wang, Y. Tang, and C.-J. Zhang, "Open-circuit fault diagnosis and tolerance strategy applied to four-wire T-type converter systems," *IEEE Trans. Power Electron.*, vol. 34, no. 6, pp. 5764–5778, Jun. 2019.
- [7] X. Ge, J. Pu, B. Gou, and Y.-C. Liu, "An open-circuit fault diagnosis approach for single-phase three-level neutral-point-clamped converters," *IEEE Trans. Power Electron.*, vol. 33, no. 3, pp. 2559–2570, Mar. 2018.
- [8] B. Li, S. Shi, B. Wang, G. Wang, W. Wang, and D. Xu, "Fault diagnosis and tolerant control of single IGBT open-circuit failure in modular multilevel converters," *IEEE Trans. Power Electron.*, vol. 31, no. 4, pp. 3165–3176, Apr. 2016.
- [9] S. M. A. Cruz, M. Ferreira, A. M. S. Mendes, and A. J. M. Cardoso, "Analysis and diagnosis of open-circuit faults in matrix converters," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1648–1661, May 2011.
- [10] J.-S. Lee and K.-B. Lee, "An open-switch fault detection method and tolerance controls based on SVM in a grid-connected T-type rectifier with unity power factor," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 7092–7104, Dec. 2014.
- [11] J. He, N. A. O. Demerdash, N. Weise, and R. Katebi, "A fast on-line diagnostic method for open-circuit switch faults in SiC-MOSFET-based T-type multilevel inverters," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 2948–2958, May 2017.
- [12] A. Joseph, T. R. Chelliah, R. Selvaraj, and K.-B. Lee, "Fault diagnosis and fault-tolerant control of megawatt power electronic converter-fed large-rated asynchronous hydrogenerator," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 7, no. 4, pp. 2403–2416, Dec. 2019.
- [13] F. Lv, C. Wen, Z. Bao, and M. Liu, "Fault diagnosis based on deep learning," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2016, pp. 6851–6856.
- [14] A. S. S. Vasan and M. G. Pecht, "Electronic circuit health estimation through kernel learning," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1585–1594, Feb. 2018.
- [15] L. Lu, Y. He, T. Wang, T. Shi, and B. Li, "Self-powered wireless sensor for fault diagnosis of wind turbine planetary gearbox," *IEEE Access*, vol. 7, pp. 87382–87395, 2019.
- [16] J. Tian, C. Morillo, M. H. Azarian, and M. Pecht, "Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1793–1803, Mar. 2016.
- [17] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [18] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, Feb. 1997.
- [19] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
- [20] D. Rodrigues, L. A. M. Pereira, R. Y. M. Nakamura, K. A. P. Costa, X.-S. Yang, A. N. Souza, and J. P. Papa, "A wrapper approach for feature selection based on bat algorithm and optimum-path forest," *Expert Syst. Appl.*, vol. 41, no. 5, pp. 2250–2258, Apr. 2014.
- [21] G. Roffo, S. Melzi, and M. Cristani, "Infinite feature selection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 4202–4210.
- [22] L. Wang, N. Zhou, and F. Chu, "A general wrapper approach to selection of class-dependent features," *IEEE Trans. Neural Netw.*, vol. 19, no. 7, pp. 1267–1278, Jul. 2008.
- [23] L. Ma, M. Li, Y. Gao, T. Chen, X. Ma, and L. Qu, "A novel wrapper approach for feature selection in object-based image classification using polygon-based cross-validation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 409–413, Mar. 2017.
- [24] 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, Jan. 2018.
- [25] J. Guo and W. Zhu, "Dependence guided unsupervised feature selection," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 1–8.
- [26] L. Meysenc, P. Stefanutti, P. Noisette, N. Hugo, and A. Akdag, "Multi-level AC/DC converter for traction applications," U.S. Patent 7558 087, Jul. 7, 2009.
- [27] J. Li, J. Jiang, and S. Qiao, "A space vector pulse width modulation for five-level nested neutral point piloted converter," *IEEE Trans. Power Electron.*, vol. 32, no. 8, pp. 5991–6004, Aug. 2017.
- [28] S. Ye, J. Jiang, J. Li, Y. Liu, Z. Zhou, and C. Liu, "Fault diagnosis and tolerance control of five-level nested NPP converter using wavelet packet and LSTM," *IEEE Trans. Power Electron.*, vol. 35, no. 2, pp. 1907–1921, Feb. 2020.
- [29] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, nos. 1–4, pp. 131–156, 1997.
- [30] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, nos. 1–2, pp. 273–324, Dec. 1997.
- [31] A. Arauzo-Azofra, J. L. Aznarte M, and J. M. Benítez, "Empirical study of individual feature evaluators and cutting criteria for feature selection in classification," in *Proc. 9th Int. Conf. Intell. Syst. Design Appl.*, Nov./Dec. 2009, pp. 541–546.
- [32] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.



**SHU YE** (Student Member, IEEE) received the B.S. degree in information and telecommunication engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2009. She is currently pursuing the Ph.D. degree in power electronics and power drives with the Key Laboratory of Control of Power Transmission and Conversion, Shanghai Jiao Tong University, Shanghai, China.

Her current research interests include multilevel converters, control algorithms, and reliability of power electronic systems.



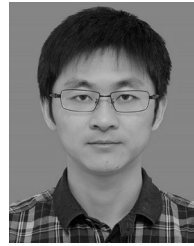
**JIANGUO JIANG** was born in Jiangsu, China, in 1956. He received the Ph.D. degree in electrical engineering from the China University of Mining and Technology, Xuzhou, China, in 1988. He is currently a Professor of electrical engineering with Shanghai Jiao Tong University, Shanghai, China.

His research interest includes high-power motor drive and control.



**ZHONGZHENG ZHOU** (Student Member, IEEE) was born in Shandong, China, in 1989. He received the B.S. degree in electrical engineering and automation from the China University of Mining and Technology, Xuzhou, China, in 2014. He is currently pursuing the Ph.D. degree in power electronics and power drives with the Key Laboratory of Control of Power Transmission and Conversion, Shanghai Jiao Tong University, Shanghai, China.

His current research interests include dc circuit breakers, multilevel converters, and high-power motor drives.



**YUNLONG LIU** received the M.S. degree in power electronics and power drives from the School of Electrical and Information Engineering, Jiangsu University, in 2015. He is currently pursuing the Ph.D. degree with the School of Mechanical Electronic and Information Engineering, China University of Mining and Technology, Beijing.

His current research interest includes multilevel converters.

...



**CONG LIU** (Student Member, IEEE) was born in Jiangsu, China, in 1988. He received the B.S. degree in electrical engineering and automation from the China University of Mining and Technology, Xuzhou, China, in 2010. He is currently pursuing the Ph.D. degree in power electronics and power drives with the Key Laboratory of Control of Power Transmission and Conversion, Shanghai Jiao Tong University, Shanghai, China.

His current research interests include multilevel converters, control algorithms, and intelligent industrial equipment.