IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 13, 2019, accepted May 27, 2019, date of publication June 14, 2019, date of current version July 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2923017

Incipient Fault Diagnosis Method for IGBT Drive Circuit Based on Improved SAE

YIGANG HE, (Member, IEEE), CHENCHEN LI, TAO WANG, TIANCHENG SHI, LIN TAO, AND WEIBO YUAN

School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China

Corresponding authors: Yigang He (18655136887@163.com) and Chenchen Li (18226615994@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51577046, in part by the State Key Program of National Natural Science Foundation of China under Grant 51637004, and in part by the National Key Research and Development Plan ''Important Scientific Instruments and Equipment Development'' under Grant 2016YFF0102200.

ABSTRACT An incipient fault diagnosis method devised for insulated gate bipolar transistor (IGBT) drive circuit based on improved stack auto-encoder (SAE) is recommended. First, the Monte Carlo method is applied to extracting the time domain response signal of the circuit under test as sample data. Then, with SAE used to extract essential features of data, the SAE is employed to extract features of sample data. Meanwhile, multi-classification relevant vector machine (RVM) is involved for fault diagnosis of the acquired features. As the structure of the hidden layer in SAE and the learning rate could exert a significant effect on the feature extraction performance, in this paper, the quantum particle swarm optimization (QPSO) algorithm is used to optimize the above parameters. As revealed by the experimental results, the improved SAE method is effective in the extraction of the essential characteristics of the incipient faults for the IGBT drive circuit. Further with this, the incipient fault multi-classification RVM of the IGBT drive circuit is capable of achieving 100% diagnostic accuracy.

INDEX TERMS IGBT drive circuit, incipient diagnosis, deep learning, stacked auto-encoder, multi-classification relevant vector machine.

I. INTRODUCTION

IGBT is extensively utilized in various applications such as industrial drives, power transmission, aerospace, induction heating equipment and so on. Generally, high-power IGBT requires the use of a separate drive circuit, and the performance of the drive circuit exerts a substantial effect on the operating state of the IGBT. As revealed by statistics, the faults of medium and high voltage inverters resulting from high-power IGBT faults make up over 90% of all faults. As for IGBT faults, more than 30% result from driving circuits. In this sense, fault diagnosis of the IGBT drive circuit and keeping track of its operating status online are extremely crucial to enhancing the reliability of equipment while reducing maintenance costs and boosting the overall efficiency of the system.

Currently, there is little research focused on IGBT drive circuit faults, despite that IGBT drive circuits are analog circuits in essence. Therefore, it is possible to make reference to the related research content of analog circuit fault diagnosis to facilitate the diagnosis of IGBT drive circuit. The research into analog circuit fault diagnosis has been conducted for many years with many remarkable results obtained. Tan *et al.* [1] have suggested quantum neural network as processor, and S-transform to process the raw data. Zhang *et al.* [2] recommended an early fault diagnosis method for analog circuits on the basis of deep belief network (DBN) feature extraction. In the diagnosis scheme, measurement is performed of the time responses of analog circuits, prior to features being extracted with the use of the DBN method. According to the characteristics exhibited by the circuit, Xiao and He devised a neural-network fault diagnosis approach of analog circuits with maximal class separability based kernel principal components analysis (MCSKPCA) taken as preprocessor. The proposed approach is capable of detecting and diagnosing faulty components efficiently in the analog circuits through the analysis of their time response [3]. By analyzing the testability of analog circuits, the improved Mahalanobis distance is involved to assist analog circuit fault diagnosis [4]. By using a combinatorial optimization

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

diagnosis method, this literature presented an innovative method of analog circuit fault diagnosis that is based on dependence matrices and intelligent classifiers (DM-IC) [5]. Aimed at the existing analog circuit fault diagnosis methods, two fault diagnosis algorithms based on SVM and global SVM are suggested [6]. However, these methods show some defects. Although literature [1] achieves multiple types of fault identification, the accuracy of fault vibration is not as high as expected. Literature [2] and literature [6] have made significant improvement to the accuracy of fault diagnosis. Despite this, it provides no assistance for multi- classification fault diagnosis. In addition, the existing fault diagnosis technology is purposed to diagnose the fault that has already occurred. At this time, the fault has resulted in a certain loss. In order to minimize the economic loss incurred by the fault occurring to the drive circuit, the drive circuit needs to be identified accurately prior to the occurrence of the fault. That is to say, it has to enable incipient fault diagnosis of the drive circuit. The accuracy of fault diagnosis is closely related to the effect of feature extraction [7]. Effective feature extraction is considered requisite for making accurate fault diagnosis. The common feature extraction methods include principal component analysis [8], kernel principal component analysis [9], blind source separation [10], and wavelet analysis [11]. However, these methods are restricted to getting the shallow features in the original signal, for which incipient faults are often unable to be identified based on shallow features. Therefore, there is a necessity to explore the deep features in the original data so as to ensure an effective identification of incipient faults.

For the most recent years, deep learning has been extensively applied to speech recognition [12], facial recognition [13] and other fields owning to its excellent feature extraction performance. Therefore, in this paper, the SAE is introduced to extract the incipient fault signal of the drive circuit, and then the RVM is involved to make diagnosis of the fault. Given SAE's hidden layer structure, learning rate and kernel function parameters in RVM could exert significant effect on the final fault diagnosis results, for which the QPSO algorithm is applied in this paper to optimize the above three parameters.

II. FAULT DIAGNOSIS ALGORITHM FOR IGBT DRIVING CIRCUIT BASED ON SAE

A. DEEP LEARNING

Deep learning is a fault diagnosis method, It refers to the specific training of large sample data through unsupervised training and supervised learning, and multiple hidden layers have strong representational ability. Compared with other algorithms, it has a strong advantage in feature extraction [14]–[16]. Deep learning was first proposed by Hinton *et al* in 2006. It is a new field in machine learning. Its purpose is to analyze and interpret data, such as sound, image and text by simulating the structure of human brain, and to describe the data features through the model. Among them, the SAE is an important model of

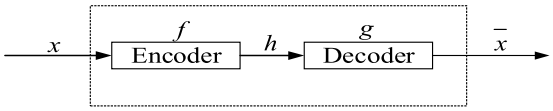


FIGURE 1. The principle of auto-encoder.

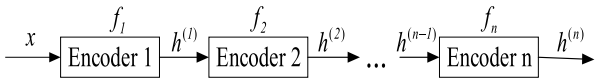


FIGURE 2. The principle of stack auto-encoder.

deep learning. The "depth" of deep learning refers to the multi-layer learning model, which is the structure of deep learning. It discovers the distributed characteristics of data by combining the underlying features of the data to form more abstract high-level features. In many models of deep learning, the number of levels of nonlinear operation is relatively large. Deep learning realizes complex and high-dimensional function representation by learning deep nonlinear network. Deep learning can transform the original fault signal layer by layer. In order to make the classification and prediction of samples easier, the feature representation of the sample in the original space is transformed into another new feature space. Compared with shallow learning, deep learning utilizes big data learning features and has achieved classification and recognition performance beyond existing algorithms. As machine learning, deep learning has been successfully applied for speech recognition, memory network, natural language processing, computer vision and other fields.

B. STACK AUTO-ENCODER

An auto-encoder is a type of neural network that uses a back-propagation algorithm to make the input value equal to the output value. After training, the input is compressed into a potential spatial representation to reconstruct the output. The principle is shown in fig.1. It consists of an encoder and a decoder. The final result is to implement $x - h - \bar{x}$. The auto-encoder is essentially a transformation of the input signal. As can be seen from fig.1, the encoder converts the input signal x into the encoded signal h, and the decoder converts the encoded signal h into the output signal $\bar{x}$.

As can be seen from fig.2, the stack auto-encoder is formed by stacking the coding parts of the encoder. That is, the process coding h of the upper auto-encoder is used as the input of the next layer of the auto-encoder. After several iterations of learning and pre-training, the training of the entire network is carried out layer by layer, and finally stacked into SAE. This is the process of layer by layer unsupervised pre-training.

The structure diagram of SAE is shown in fig.3. The variable x represent the input signal and inputs the signal to the first encoder. N is the number of stacked layers of SAE; where $Z^{(N)}$ represents the coded information on the Nth layer; AEN is the automatic encoder after the network training.

The following formula represents the coding process of SAE:

$$z = f(w_z x + b_z) \tag{1}$$

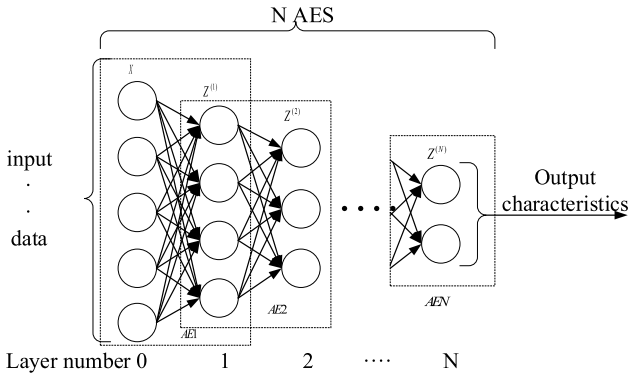


FIGURE 3. The structure schematic of Stack-AE.

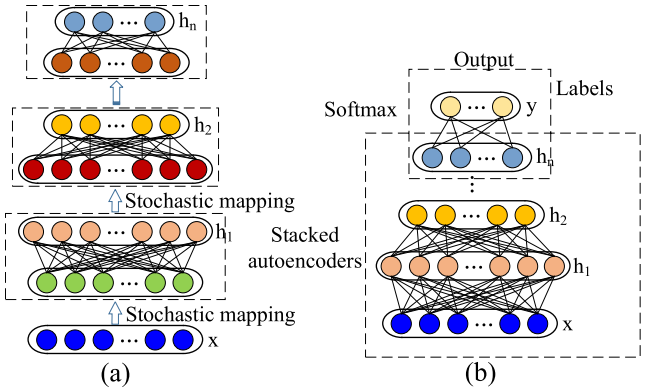


FIGURE 4. (a)Unsupervised training of SAE, (b)Supervised fine-tuning of SAE.

The following formula represents the decoding process of SAE:

$$\hat{x} = f(w_{\hat{x}} z + b_{\hat{x}}) \tag{2}$$

where: w_z is the weight value that reaches the hidden layer through the input layer;w_x is the weight value of the hidden layer to the reconstruction layer; b_z is the offset value of the hidden layer; b_x is the offset size of the output unit; $f(\cdot)$ represents the activation function, which is usually represented by sigmoid function or sine-cosine function. Generally there is $w_z = w_x^T = w$, where T is transposition.

Let $e(x, x)$ be the reconstructed error function of the reconstructed signal x and the input data x, Then the optimal solution objective function of w, b_z, b_x can be setting as follows:

$$T = \arg\min[e(x, \hat{x})] \tag{3}$$

When the input signal x is known, the reconstructed signal x is depends on the parameters w_r, b_z, b_x.According to the gradient descent method and the chain rule, the updating rule of the parameter set can be obtained:

$$w_r = w_r - \eta \frac{\partial e(x, \hat{x})}{\partial w_r} \tag{4}$$

$$b_z = b_z - \eta \frac{\partial e(x, \hat{x})}{\partial b_z} \tag{5}$$

$$b_{\hat{x}} = b_{\hat{x}} - \eta \frac{\partial e(x, \hat{x})}{\partial b_{\hat{x}}} \tag{6}$$

where η is the learning rate. The weight update formula is used to update the parameter set in turn. If the reconstruction error is the smallest or the convergence range is met, then the corresponding parameter set w^*, b_z^*, b_x^* is the parameter extracted by the encoder.

The training process of SAE is divided into two stages. The first stage is unsupervised learning, and the second stage is supervised global fine-tuning. The network structure of SAE is shown in fig.4, where figure 4(a) is the unsupervised learning stage and figure 4(b) is the global fine-tuning stage with supervision. First, the original data was ''destroyed'' by random mapping, and the AE at the first layer was trained by minimizing the average reconstruction error to obtain the corresponding parameter set; Then use the output of layer 1 as the input data of layer 2 to train layer 2 AE, obtain corresponding AE parameters, and repeat this process until all AE training is completed.

As can be seen from fig.4, the coding part of the auto-encoder is stacked in a certain manner, so that the encoder and the decoder have a multi-layer network, and then the features extracted by one encoder are passed as input to the next encoder to form a stack. SAE can be considered as a network of layer-by-layer pre-training to minimize reconstruction errors, or each encoder/decoder can be pre-trained using an unsupervised method. Unsupervised learning does not require sample labels. After pre-training, only a small number of sample labels are needed to fine-tune the entire network. Better results can be achieved through pre-training. Because the single-layer encoder is shallow deep learning, it avoids network local optimization [17].

III. RVM ALGORITHM AND PARAMETER OPTIMIZATION

A. RVM ALGORITHM

Relevant Vector Machine (RVM) [18] is a machine learning model derived from Bayesian algorithm. It is a new supervised learning method and is a sparse probability model similar to SVM. Moreover, it can provide a posterior probability distribution of the same accuracy as SVM to achieve the classification effect. Let the given training data set be $\{x_u, t_u\}_{u=1}^N$, $t_u \in \{0, 1\}$, where x_u is the N-dimensional input vector, and the model of RVM can be obtained from the following formula:

$$y(x) = \sum_{u=1}^{N} w_u K(x) \tag{7}$$

where: w_u is the weight vector, $K_u(x)$ is the kernel function.

For binary classification problem, the RVM classifier is usually used to project $y(x)$ through the Sigmoid function into the (0,1) interval, and then make the classification decision on

the following functions:

$$p\{t=1|x\} = \frac{1}{1+e^{-y(x)}} \tag{8}$$

When $p\{t=1|x\} < 0.5$, the classification result can be regarded as the first category. When $p\{t=1|x\} \geq 0.5$ the classification result can be regarded as the second category. $p\{t|x\}$ is the Bernoulli distribution, and the posterior probability likelihood of the expected result is estimated as:

$$p\{t|x\} = \prod_{u=1}^{N} p(t_i|x_i)^{t_i}(1-p(t_i|x_i)^{1-t_i}) \tag{9}$$

To avoid overfitting, RVM defines the prior probability of each weight:

$$\begin{aligned} p\{\omega|\alpha\} &= \prod_{u=1}^{N} N(\omega_u|0,\alpha_u^{-1}) \\ &= \prod_{u=1}^{N} \sqrt{\frac{\alpha_u}{2\pi}} \exp(-\frac{\alpha_u\omega_u^2}{2}) \end{aligned} \tag{10}$$

where: $\alpha = (\alpha_u)_{u=1}^{N}$ is a hyperparametric vector, $N(\cdot)$ represents a normal distribution function.

Compared with Support vector machine(SVM), the relevant vector machine (RVM) has high sparsity, effectively avoids complicated parameter setting problems, and can obtain high prediction accuracy. It also reduces the time for predictive calculations and greatly reduces the number of kernel functions involved in predictive calculations. At the same time, the kernel function of RVM does not need to meet Mercer's conditions, and its number is much smaller than the number of SVM. Since SVM can only diagnose the samples qualitatively, RVM can describe the diagnosis quantitatively through formula (8). In contrast, RVM is more suitable for engineering practice than SVM.

In essence, both SVM and RVM are a binary classifier that cannot be used directly for fault diagnosis. Therefore, this paper expands the RVM into multiple categories, enabling it to achieve the following two functions to meet the troubleshooting requirements we need: 1) Effective evaluation of the reliability of diagnostic results; 2) Formation of a decision-directed acyclic graph (DAG) to optimize diagnostic performance.

B. MULTI- CLASSIFICATION RVM

2001Tipping began to explore the K classification of RVM, rewriting equation (9) into the following formula:

$$P(t|w) = \prod_{i=1}^{N}\prod_{k=1}^{K} P(t_i|x_i)^{t_iK} \tag{11}$$

In the process of diagnosis, as the value of K increases, the amount of calculation will increase drastically. Therefore, directly using this method is not applicable to engineering practice. In order to improve their applicability, we generally use the method of combined learning to extend

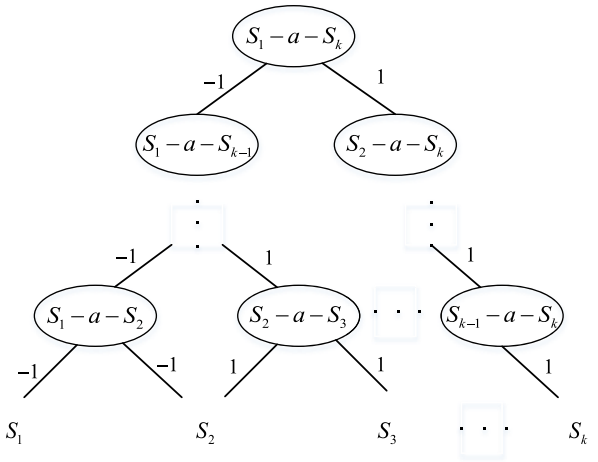


FIGURE 5. The decision-directed acyclic graph.

the multi-classifier of the binary classifier [19]–[21]. Several methods have been proposed to extend RVM for multi-classification. It includes one-to-rest (1-*v*-*r*) [22] and one-to-one (1-*v*-1) [23].

When using the 1-*v*-*r* method in training phases and the decision phase, each RVM classifier requires all training samples to perform operations, which still has a large amount of calculation. However, the 1-*v*-1 method only uses the second type of data for calculation, and then compares all the diagnostic results to complete the classification of the test samples. From $K \cdot O(x) < O(K \cdot x)$, It can be seen that the calculation complexity of the 1-*v*-1 method is smaller. Compared with the 1-*v*-*r* method, the number of RVMs in the decision phase of the 1-*v*-1 method is increased from k-1 to $k(k\text{-}1)/2$. In 1998, Platt *et al.* [24].introduce the idea of decision-oriented acyclic graphs into the SVM multi-classification extension field. For the problem of K-classification, our proposed DAG-SVM method is to first establish $k(k\text{-}1)/2$ binary classifiers, which will be distributed as different nodes in the k-1 layer decision structure. As shown in the figure below, each node in the figure corresponds to an SVM binary classifier. $n(i,j)$ is the j th node in the ith decision structure.

Assuming that the order for output decision-making values at the k-1 layer is $[S_1, S_2, \cdots, S_K]$ then there is:

$$n(i,j) = R_{S_j,S_{k+j-i}} \tag{12}$$

where: $R_{S_j,S_{k+j-i}}$ represents the classifier used to distinguish the sample categories S_j and S_{k+j-i}, denoted by S_j-a-S_{k+j-i} in the figure. Given the decision value $V(i,j) \in \{-1,1\}$ of $n(i,j)$, in the $i+1$ th layer, the DAG-SVM will send the sample to the node $n(i+1, j+(1+V(i,j)/2))$ for decision. For any test sample, the DAG-SVM can classify the sample of multiple categories through k-1 decisions. In 2009, W. Huilan et al. replace the original SVM binary classifier with the RVM binary classifier [25], and proposed the DAG-RVM multi-classifier, which successfully realized image recognition.

The original multi- classification extension method uses a fixed decision path to diagnose test samples. At the same

time, because the decision path of multi-classification expansion will have a huge impact on the diagnosis result, and the reliability of the diagnosis cannot be guaranteed. The decision-oriented acyclic graph of the above figure is proposed to select the optimal decision path based on the diagnosis result. This can greatly improve the reliability of diagnosis [26], [27]. This also increases the reliability of the entire IGBT drive circuit system.

C. QPSO ALGORITHM AND ITS PARAMETER SELECTION PROCESS

The Particle Swarm Optimization (PSO) algorithm is a bionic optimization algorithm [28]. The algorithm uses particles to mimic the foraging habits of birds, and the position of particles is dynamically adjusted to evolution.

According to the characteristic that PSO algorithm is prone to fall into local optimal, Dr. Sun proposed the QPSO algorithm and introduced the convergence mechanism of the algorithm with quantum characteristics [29]. By relying on the uncertainty and polymorphism of quantum, the particle can obtain any position with different probabilities, and guide the particle to explore the global optimal position more efficiently. Good results have been achieved in parameter optimization (30-31).

When the next generation of particle positions is generated, it is determined by the following formula whether each particle needs to evolve its individual optimal position:

$$P_i = \begin{cases} P_i, & f(P_i) \geq f(X_i(t+1)) \\ X_i(t+1), & f(P_i) \leq f(X_i(t+1)) \end{cases} \tag{13}$$

where: Pi is the individual optimal position of the i particle, $i = 1, \ldots, N, N$ is the size of the population; $f(x)$ is the fitness function that needs to be maximized; t is the current evolutionary generation; $Xi(t+1)$ is the position of the i particle at the t+1th evolution.

Among all the particles, choose the optimal particle position Pg, that is, the global optimal position:

$$P_g(t) = \arg\max\{f(P_i)\} \tag{14}$$

The next generation positions generation method of each particle of the QPSO algorithm is as follows:

$$nbest = \frac{1}{N}\sum_{i=1}^{N} P_i \tag{15}$$

$$P = sP_i + (1-s)P_g \tag{16}$$

$$X_i(t+1) = P \pm a\,|nbest - X_i(t)|\, In(\tfrac{1}{u}) \tag{17}$$

where: N represents the number of particles in the population, u and s are random numbers uniformly distributed in [0, 1], $nbest$ is the center calculated based on the individual optimal position of each particle, a is the compression expansion factor, which is generally set to gradually decrease from 1 to 0.3 in evolution. $X(t)$ represents the position of particle i at the tth iteration. Pi and Pg are respectively the global optimal position and individual optimal position of particle i. The kernel functions width, learning rate, and number of SAE layers need to be predefined. These three parameters are extremely important to the performance of the entire monitoring system. Therefore, this paper uses QPSO algorithm to optimize SAE and RVM respectively. Taking RVM as an example, the optimization steps of QPSO algorithm are as follows:

The steps for QPSO to optimize RVM are as follows:

Step1: Initialize the QPSO algorithm, where the population number and the maximum number of iterations are set to 40 and 50, respectively, and the population particle position is limited to between 0.01 and 100, and then you map the kernel width to the position of the particle. The fitness function is the mean square error of the RVM output value and the true value:

$$fit = \frac{\sum_{i=1}^{r} [\tilde{z}(i) - z(i)]^2}{T} \tag{18}$$

Step2: Calculating the width of the RVM kernel function and the initial fitness function value;

Step3: Update the global optimal position of the particle according to formulas (15)-(17), and recalculate the new RVM kernel functions width and fitness function value;

Step4: Compare with the previous fitness function value. If it is smaller, replace the fitness function value and the kernel function width. Otherwise, the fitness function value and the kernel function width are reserved;

Step5: Repeat Step2-Step4 until the conditions at the end of the iteration are met;

Step6: Finally, the optimal result is output, and the global optimal position of the particle is the width of the kernel function.

IV. FAULT DIAGNOSIS OF IGBT DRIVE CIRCUIT

In this paper, the incipient fault diagnosis method of IGBT drive circuit based on deep learning is presented. The method of SAE signal is used for feature extraction, and the RVM is used to diagnose the data. The principle is shown in fig.6:

A. STEPS FOR IGBT FAULT DIAGNOSIS

1) The simulation circuit diagram of IGBT driving circuit is drawn by Pspice, The normal state and fault state of the circuit are simulated separately in Pspice to obtain the initial signal, and then the original signal is processed, which is used as the training sample needed in this paper;
2) Determining faulty components of the IGBT drive circuit;
3) Determining the tolerance range of each faulty component;
4) Monte Carlo is used to simulate the drive circuit to collect priory sample data. The excitation voltage of the IGBT drive circuit is a pulse voltage, and voltage signals in different fault modes are obtained at the test node;
5) The obtained voltage signals are normalized as data samples, At the same time, the category label is

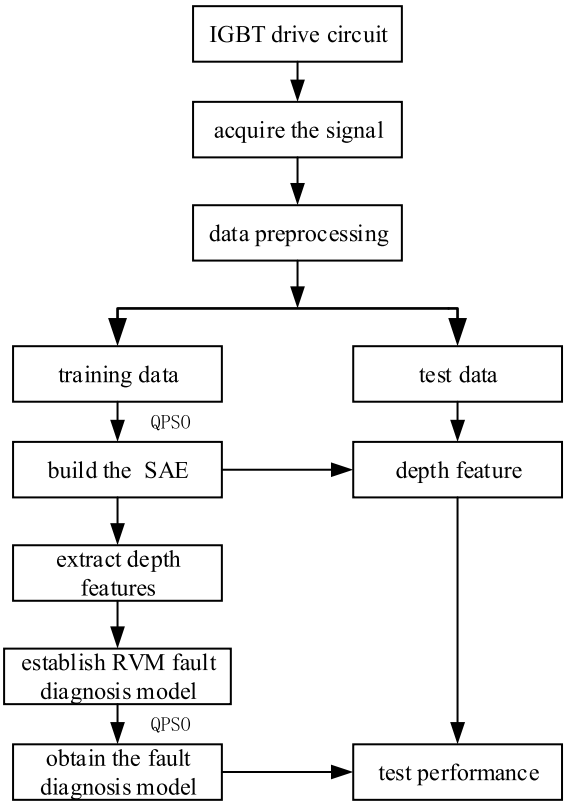


FIGURE 6. The schematic diagram of fault diagnosis method for IGBT drive circuit.

identified in the interval [0, 1], thereby completing the pre-processing of the data;

6) When setting up the network initially, you need to set the number of training cycles, the number of layers of SAE, and the dimensions of feature vectors;
7) Input the pre-processed data samples to pre-train the stacked auto-encoder;
8) The output eigenvalues and classification labels are regarded as the basis for network micro-adjustment, Then, according to the back propagation algorithm to train the whole network, thereby completing the micro adjustment of the whole network;
9) Input the real-time monitoring of the collected data set, use the trained network to complete the fault diagnosis, and output the fault mode category.

B. IGBT DRIVE CIRCUIT SIMULATION AND RESULTS ANALYSIS

The fault of the IGBT drive circuit is mainly single fault, generally the method applied to the fault of the single component can also be applied to the fault of the multi-component, so this paper studies the failure of the single component of the IGBT drive circuit. Fig.8 shows the simulation circuit of the IGBT drive circuit. It consists of 11 resistors, 5 capacitors, 5 triodes, 2 zener diode, 1 optocoupler PS2501, 1 IGBT, etc. The nominal values of the components are marked on the graph. Taking this circuit as an example, the entire flow of the diagnostic

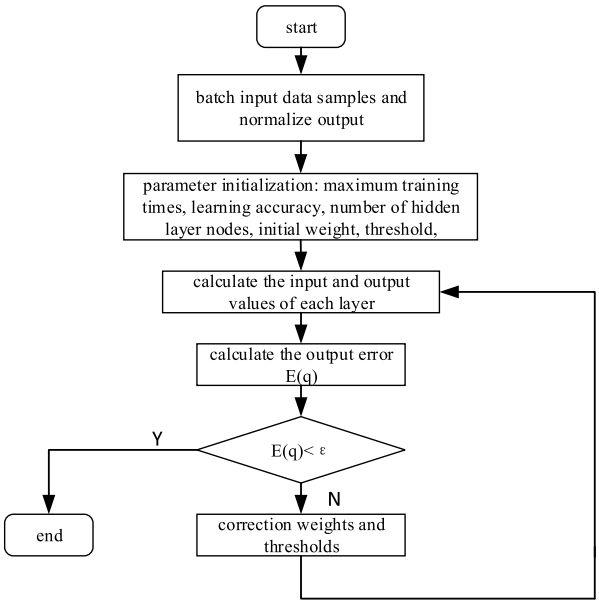


FIGURE 7. Establishment of SAE process.

method proposed in this paper is demonstrated. The excitation source uses pulse wave, and the specific parameters are: V1 = 6V, V2 = 0V, TD = 0, TR = 0, TF = 0, PW = 40us, PCR = 50us.

The voltage response wave at the IGBT output node is collected. The fault times domain response signal is obtained at the output of the circuit. The tolerances of the resistors and capacitors are set to 10% and 5% respectively. Through the sensitivity test, C5, R2, R3, R5, R9 and R11 were selected as test objects for fault diagnosis experiment of circuit components.

Table 1 show the fault categories, nominal values, and fault values of each component of the circuit, where ↑ and ↓ indicate higher and lower than the nominal value respectively. The fault categories of C5, R2, R3, R5, R9 and R11 components are listed below: C5↑, C5↓, R2↑, R2↓, R3↑, R3↓, R5↑, R5↓, R9↑, R9↓, R11 ↑, R11↓, form 12 failure modes. When the component parameter value deviates from the nominal value by 50%, the component can generally be considered to be faulty. When it deviates from the nominal value by 25%, it can be considered as an incipient fault. Each type of failure mode samples 200 sets of data, of which 100 sets of data are used as training sample sets, and the remaining samples are used as test data samples.

C. ANALYSIS OF EXPERIMENTAL RESULTS

Take the optimization process of kernel function width as an example, and fig.9 is the iterative process of QPSO algorithm. It can be seen from the figure that the QPSO algorithm avoids the disadvantage of local optimization. Therefore, the parameters obtained by the optimization of QPSO algorithm are the best. The optimal results were found at 7, 17 and 39 respectively.

When using SAE to identify the fault state of IGBT driving circuit, the hidden layer structure of SAE should be

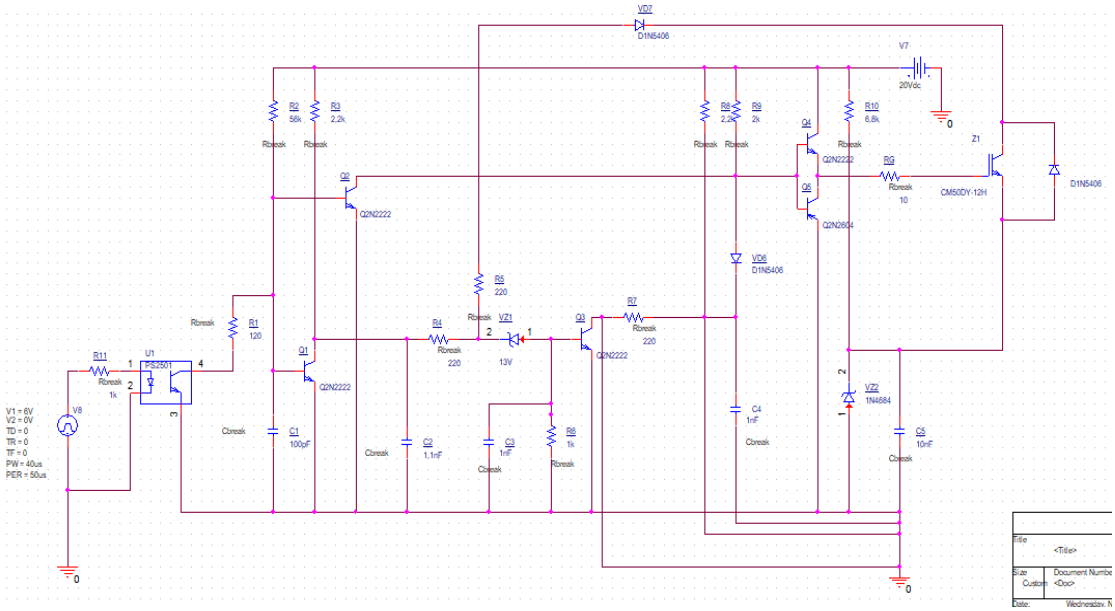


FIGURE 8. The diagram of IGBT drive circuit.

TABLE 1. Fault modes and corresponding fault setting of IGBT drive circuit.

fault mode f_N	*fault mode*	*nominal*	*tolerance*	*Corresponding fault value*
f_1	*normal*	———	———	———
f_2	C5↓	10nF	5%	7.5 nF
f_3	C5↑	10nF	5%	12.5 nF
f_4	R2↓	56k	10%	42k
f_5	R2↑	56k	10%	70k
f_6	R3↓	2.2k	10%	1.65k
f_7	R3↑	2.2k	10%	2.75k
f_8	R5↓	220	10%	165
f_9	R5↑	220	10%	275
f_{10}	R9↓	2k	10%	1.5k
f_{11}	R9↑	2k	10%	2.5k
f_{12}	R11↓	6.8k	10%	5.1k
f_{13}	R11↑	6.8k	10%	8.5k

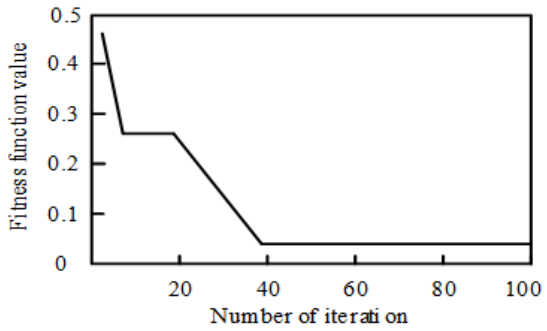


FIGURE 9. The iterative process of QPSO algorithm.

fully considered. According to the number of SAE hidden layers set in this paper, $N = 3$, optimized by QPSO, the number of iteration steps is set to 100, the optimal node

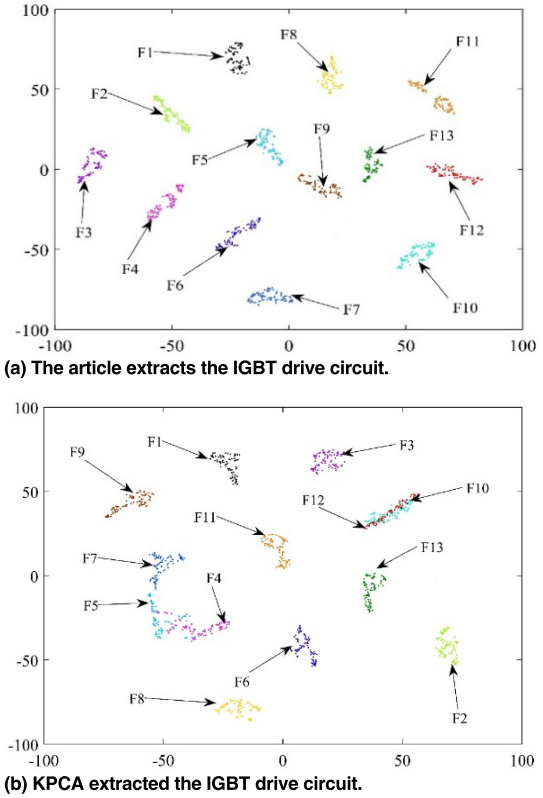
TABLE 2. SAE hidden layer structure based on optimization algorithms.

algorithm	Not optimized	QPSO
Hidden layer structure	300-200-100	325-108-247
Learning rate	0.1	0.2128
Kernel function	0.1	0.3357
Training accuracy	100%	78.5%
Test accuracy	100%	100%

structure of hidden layer of SAE network is obtained as shown in Table 2.

In this paper, SAE and wavelet packet decomposition and Kernel Principal Component Analysis (KPCA) algorithm are used to extract the fault signals of IGBT drive circuit. The purpose is to illustrate the superiority of the SAE algorithm proposed in this paper in feature extraction. The radial basis function is used as the kernel function of KPCA, the wavelet basis function is DB4, and the number of layers of wavelet decomposition is set to 4. In this paper, the Monte Carlo method is used to simulate the IGBT drive circuit, and 200 sets of circuit data are obtained. Among them, 100 groups are the health status data of the IGBT drive circuit, and the other 100 groups are the data in the fault state. Fig.10 shows the results of two-dimensional feature distribution extracted by two algorithms.

It can be seen from the above figure that comparing the two experimental results, both methods can extract the fault features of the IGBT drive circuit, and can also perform fault diagnosis of the drive circuit. However, the two-dimensional features extracted by wavelet packet decomposition and KPCA overlap each other. As can be seen from fig.10(b), the fault signals of F4 and F5, F5 and F7, F10 and F12 all overlap, indicating that the method proposed in this paper is the best for the IGBT driver circuit. Finally, it concludes that the following conclusions can be seen from the



(a) The article extracts the IGBT drive circuit.

(b) KPCA extracted the IGBT drive circuit.

FIGURE 10. Based on different methods to extract IGBT drive circuit two-dimensional fault feature distribution.

comparison of the KPCA method, as shown in fig.10(a), The two-dimensional features extracted in this paper have a high degree of separation, and the distribution among the same kind is also very uniform. The distribution of the same kind of features are also very close. And there is no coincidence of features, indicating that this method has strong classification ability.

SVM is a common algorithm at fault diagnosis. In order to verify the superiority of the method proposed in this paper, SVM and RVM are respectively adopted for fault diagnosis of the above two groups of feature extraction data. We use the QPSO algorithm to optimize the kernel function width of SVM and RVM. Then, the parameters of SVM and RVM are optimized without QPSO, and then the results of SAE feature extraction are diagnosed by SVM and RVM respectively. Table 3 shows the results of various fault diagnosis. Among them, KPCA indicates feature extraction using wavelet packet and KPCA, and SAE indicates feature extraction using the method of this paper. QRVM represents feature extraction using QPSO-optimized RVM algorithm, and QSVM represents feature extraction using QPSO-optimized SVM algorithm.

According to the data in the table, the best results can be obtained by using SAE feature extraction and using QPSO optimized RVM for fault diagnosis, and the test time is also the shortest. And RVM is more spars than SVM. Completing the training process ahead of the test has no effect on the overall troubleshooting time.

TABLE 3. Comparison of fault diagnosis results.

Algorithm	Training	Test accuracy	Testing time
KPCA-QRVM	83.65%	73.1%	3.17s
KPCA-QSVM	84.11%	74.3%	4.65s
SAE-RVM	100%	91.6%	0.59s
SAE-SVM	100%	90.7%	1.33s
SAE-QRVM	100%	100%	0.64s
SAE-QSVM	100%	100%	1.48s

V. CONCLUSION

At present, in view of the fact that the existing analog circuit fault diagnosis method cannot effectively diagnose the incipient faults of the IGBT drive circuit, an incipient fault diagnosis method for the IGBT drive circuit based on deep learning technology is proposed. The method overcomes the drawback that the traditional fault diagnosis method cannot effectively diagnose the early fault of the IGBT drive circuit. And it can be effectively identified when the fault is in the incubation period, thereby reducing the economic loss caused by the failure. In this paper, deep level features of early fault signals are extracted from collected data by SAE method, and a fault diagnosis model based on multi-classification RVM is established. Considering that the parameter settings in SAE and multi-classification RVM have a significant impact on the final diagnostic effect, Therefore, this paper adopts QPSO algorithm to optimize parameters of RVM and SAE to achieve the best fault diagnosis effect. By comparing the proposed method with traditional feature extraction and fault classification methods, the superiority and advancement of the proposed method are proved. Experimental results show that the proposed multi-classification RVM for early fault of IGBT driver circuit can achieve 100% diagnostic accuracy.

REFERENCES

[1] 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.

[2] 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.

[3] Y. Xiao and Y. He, "A novel approach for analog fault diagnosis based on neural networks and improved kernel PCA," *Neurocomputing*, vol. 74, no. 7, pp. 1102–1115, Mar. 2011.

[4] H. Han, H. Wang, S. Tian, and N. Zhang, "A new analog circuit fault diagnosis method based on improved Mahalanobis distance," *J. Electron. Test.*, vol. 29, no. 1, pp. 95–102, Feb. 2013.

[5] J. Shi, Y. Deng, Z. Wang, and Q. He, "A combined method for analog circuit fault diagnosis based on dependence matrices and intelligent classifiers," *IEEE Trans. Instrum. Meas.*, to be published.

[6] A. Zhang and C. Chen, "Fault diagnosis based semi-supervised global LSSVM for analog circuit," in *Proc. Int. Conf. Mechatronics Control*, Jinzhou, China, 2014, pp. 744–748.

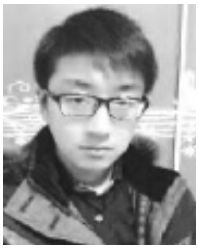
[7] S. Chen, M. Du, Z. Peng, Q. He, W. Zhang, and M. Liang, "High-accuracy fault feature extraction for rolling bearings under time-varying speed conditions using an iterative envelope-tracking filter," *J. Sound Vib.*, vol. 448, pp. 211–229, May 2019.

[8] S. Siuly and Y. Li, "Designing a robust feature extraction method based on optimum allocation and principal component analysis for epileptic EEG signal classification," *Comput. Methods Programs Biomed.*, vol. 119, no. 1, pp. 29–42, Apr. 2015.

[9] Z.-G. Zhao and F. Liu, "On-line nonlinear process monitoring using kernel principal component analysis and neural network," in *Proc. Adv. Neural Netw. (ISNN)*, vol. 3973, 2006, pp. 945–950.
[10] G. D. Pelegrina, L. T. Duarte, and C. Jutten, "Blind source separation and feature extraction in concurrent control charts pattern recognition: Novel analyses and a comparison of different methods," *Comput. Ind. Eng.*, vol. 92, pp. 105–114, Feb. 2016.
[11] P. Li, F. Kong, Q. He, and Y. Liu, "Multiscale slope feature extraction for rotating machinery fault diagnosis using wavelet analysis," *Measurement*, vol. 46, no. 1, pp. 497–505, Jan. 2013.
[12] F. S. Abousaleh, T. Lim, W.-H. Cheng, N.-H. Yu, M. A. Hossain, and M. F. Alhamid, "A novel comparative deep learning framework for facial age estimation," *EURASIP J. Image Video Process.*, vol. 2016, no. 1, p. 47, Dec. 2016.
[13] S. M. Siniscalchi, D. Yu, L. Deng, and C.-H. Lee, "Exploiting deep neural networks for detection-based speech recognition," *Neurocomputing*, vol. 106, pp. 148–157, Apr. 2013.
[14] T. Ma, H. Li, H. Yang, X. Lv, P. Li, T. Liu, D. Yao, and P. Xu, "The extraction of motion-onset VEP BCI features based on deep learning and compressed sensing," *J. Neurosci. Methods*, vol. 275, pp. 80–92, Jan. 2017.
[15] X. Cao *et al.*, "Ship recognition method combined with image segmentation and deep learning feature extraction in video surveillance," *Multimedia Tools Appl.*, 2019. doi: 10.1007/s11042-018-7138-3.
[16] Q. Zhang, L. T. Yang, and Z. Chen, "Deep computation model for unsupervised feature learning on big data," *IEEE Trans. Services Comput.*, vol. 9, no. 1, pp. 161–171, Jan./Feb. 2016.
[17] H. Yin, X. Jiao, Y. Chai, and B. Fang, "Scene classification based on single-layer SAE and SVM," *Expert Syst. Appl.*, vol. 42, no. 7, pp. 3368–3380, 2015.
[18] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.
[19] J. Qu, Z. Zhang, and T. Gong, "A novel intelligent method for mechanical fault diagnosis based on dual-tree complex wavelet packet transform and multiple classifier fusion," *Neurocomputing*, vol. 171, pp. 837–853, Jan. 2016.
[20] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, and F. Herrera, "An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes," *Pattern Recognit.*, vol. 44, no. 8, pp. 1761–1776, Aug. 2011.
[21] M. M. M. Islam and J.-M. Kim, "Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector machines," *Rel. Eng. Syst. Saf.*, vol. 184, pp. 55–66, Apr. 2019.
[22] L. Gonzalez-Abril, F. Velasco, C. Angulo, and J. A. Ortega, "A study on output normalization in multiclass SVMs," *Pattern Recognit. Lett.*, vol. 34, no. 3, pp. 344–348, Feb. 2013.
[23] P. Lingras and C. Butz, "Rough set based 1-v-1 and 1-v-r approaches to support vector machine multi-classification," *Inf. Sci.*, vol. 177, no. 18, pp. 3782–3798, Sep. 2007.
[24] J. C. Platt, N. Cristianini, and J. Shawe-Taylor, "Large margin DAGs for multiclass classification," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 1999, vol. 2000, no. 12, pp. 547–553.
[25] H. Wu, G. Liu, and Z. Pu, "Multi-class image recognition based on relevance vector machine," in *Proc. Int. Workshop Intell. Syst. Appl.*, WuHan, China, 2009, pp. 1–4.
[26] B. Kijsirikul and N. Ussivakul, "Multiclass support vector machines using adaptive directed acyclic graph," in *Proc. Int. Joint Conf. Neural Netw.*, Honolulu, HI, USA, 2002, pp. 980–985.
[27] H. Yi, X. Song, and B. Jiang, "Structure selection for DAG-SVM based on misclassification cost minimization," *Int. J. Innov. Comput. Inf. Control*, vol. 7, no. 9, pp. 5133–5143, Sep. 2011.
[28] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Hum. Sci.*, Nagoya, Japan, 1995, pp. 39–43.
[29] J. Sun, W. Fang, X. Wu, V. Palade, and W. Xu, "Quantum-behaved particle swarm optimization: Analysis of individual particle behavior and parameter selection," *Evol. Comput.*, vol. 20, no. 3, pp. 349–393, 2012.
[30] E. Salahshour, M. Malekzadeh, F. Gordillo, and J. Ghasemi, "Quantum neural network-based intelligent controller design for CSTR using modified particle swarm optimization algorithm," *Trans. Inst. Meas. Control*, vol. 41, no. 2, pp. 392–404, Jan. 2019.
[31] S. L. Sabat, L. dos Santos Coelho, and A. Abraham, "MESFET DC model parameter extraction using quantum particle swarm optimization," *Microelectron. Rel.*, vol. 49, no. 6, pp. 660–666, 2009.

YIGANG HE (M'17) received the M.Sc. degree in electrical engineering from Hunan University, Changsha, China, in 1992, and the Ph.D. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1996. In 1990, he joined the College of Electrical and Information Engineering, Hunan University, where he was promoted to Associate Professor in 1996, and a Professor in 1999. From 2006 to 2011, he was the Director of the Institute of Testing Technology for Circuits and Systems, Hunan University. He is currently a Professor with the Hefei University of Technology. His teaching and research interests include the areas of circuit theory and its applications, testing and fault diagnosis of analog and mixed-signal circuits, electrical signal detection, smart grid, radio frequency identification technology, and intelligent signal processing.

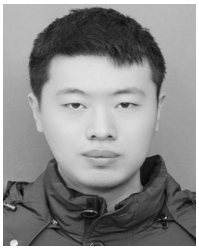
CHENCHEN LI received the B.Sc. degree from the Liaoning University of Science and Technology, in 2015. She is currently pursuing the master's degree with the Hefei University of Technology. Her current research interest includes the reliability research of power electronics devices.



TAO WANG received the B.S. degree in electrical engineering from the Hefei University of Technology, Hefei, China, in 2014, where he is currently pursuing the Ph.D. degree. His current interests include the research of intelligent and real-time information processing, smart electronics devices, and the diagnosis and prognosis of high voltage equipment.

TIANCHENG SHI was born in 1990. He received the bachelor's degree in engineering from the Hefei University of Technology, in 2013, where he is currently pursuing the Ph.D. degree. His current research interests include condition monitoring, reliability analysis, fault diagnosis, and fault tolerance techniques for PWM power electronics converter systems.

LIN TAO was born in 1994. She received the B.Sc. degree from North China Electric Power University, in 2015. She is currently pursuing the master's degree with the Hefei University of Technology. Her current research interests include fault diagnosis and the prediction of power equipment.



WEIBO YUAN received the B.Sc. degree from the Hefei University of Technology, in 2017, where he is currently pursuing the Ph.D. degree. His current research interests include multi-physics modeling and reliability of power semiconductor devices, high voltage, and insulation.