VWM-DCRNN: A Method of Combining Signal Processing With Deep Learning for Fault Diagnosis in Single-Phase PWM Rectifier

Na Qin, *Member, IEEE*, Tianwei Wang, Deqing Huang, *Senior Member, IEEE*, Yiting You, and Yiming Zhang

***Abstract*—Based on variational mode decomposition and dual model of convolutional recurrent neural network (VMD-DCRNN), a novel diagnosis method is proposed to discover the faults of insulated gate bipolar transistors (IGBTs), diodes, and series resonant circuit in a single-phase pulsewidth modulation ac–dc rectifier, an important part of traction power supply system of high-speed railway. By virtue of the combination of signal processing and deep learning schemes, the multiscale feature information for fault diagnosis is extracted by regarding the input current of the rectifier and the voltage of dc side as the original signal. More clearly, VMD is adopted to decompose the original signal to a series of intrinsic mode function components (IMFs). Then, the submodel of current is designed to detect the fault types of IGBTs and diodes. Meanwhile, the submodel of voltage is established to identify the fault types of series resonant circuit. These IMFs are fed into DCRNN consisting of the two submodels for training and testing. Finally, the features extracted from the two submodels are integrated and converted to label space to complete fault diagnosis. Experimental results show that the proposed VMD-DCRNN algorithm is superior to the single-channel models, reaching 96.27%.**

***Index Terms*—Convolutional recurrent neural network (CRNN), fault diagnosis, rectifier, variational mode decomposition (VMD).**

I. Introduction

THE ac–dc–ac traction drive system is widely used in the field of high-speed railway traction systems, and the structure of traction drive system is shown in Fig. 1. Power electronic transformer (PET) plays a role in power conversion in traction drive system, mainly including high-frequency transformers and power electronic converters. PET is always with complex structure, a large number of components, and frequent hidden faults. The fault of main devices of PET often leads to abnormal fluctuation of voltage and current, which endangers the driving safety of high-speed trains. Most of the existing fault diagnosis methods for PET are model based, where the basic idea is to use a mathematical model to calculate fault residuals and conduct fault diagnosis through residual analysis. However, in practical engineering applications, the nonlinearity and discreteness of switching devices limit the precision of the analytical model. Therefore, in the fault diagnosis of complex nonlinear systems, it is urgent to study the fault diagnosis method with less dependence on the model. Along this way, those data-based fault diagnosis methods prefer making full use of the rich device state information contained in the system detection data, and exploring the internal relation between the data and the fault modes.

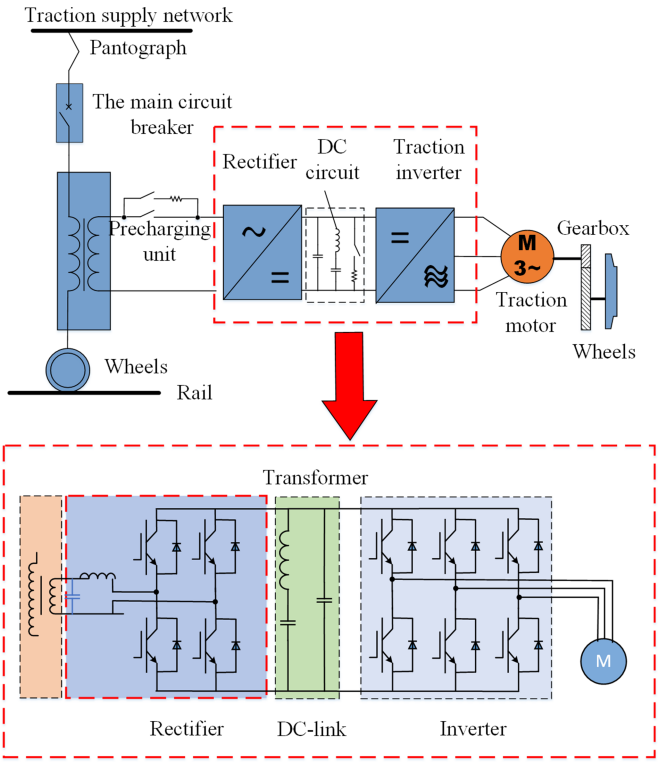

Fig. 1. Structure of traction drive system. Among all the components of the system, the rectifier and the DC-link are of the main concern of this article.

As an important part of traction drive system, the performance of the single-phase pulsewidth modulation (PWM) rectifier will directly affect that of the drive system, where the power semiconductor device is the most vulnerable [1]. In consequence,

Manuscript received 15 November 2022; revised 2 March 2023; accepted 23 March 2023. Date of publication 30 March 2023; date of current version 19 May 2023. This work was supported in part by the Natural Science Foundation of China under Grants 62173279 and U1934221, and in part by the Sichuan Science and Technology Program under Grants 2022YFG0247, 2021JDJQ0012, and 2020YFQ0057. Recommended for publication by Associate Editor F. W. Fuchs. *(Corresponding author: Deqing Huang.)*

The authors are with the Institute of Systems Science and Technology, Southwest Jiaotong University, Chengdu 610031, China (e-mail: qinna@swjtu.edu.cn; wtw@my.swjtu.edu.cn; elehd@home.swjtu.edu.cn; youyit@my.swjtu.edu.cn; zhangym@my.swjtu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2023.3263226.

Digital Object Identifier 10.1109/TPEL.2023.3263226

open-circuit and short-circuit faults of insulated gate bipolar transistors (IGBTs) have become the common faults of the grid-side rectifier. In general, when a short-circuit fault occurs, the short-circuit fault evolves rapidly and has serious consequences, making it difficult to perform fault diagnosis. However, the short-circuit fault can be converted into an open-circuit fault by adding a fast fuse and other hardware protection circuits in the converter [2]. Therefore, analysis and diagnosis of open-circuit faults of IGBTs are playing a central role in the fault diagnosis of the rectifier. Besides, the abnormal voltage stress caused by the diode open-circuit fault is much higher than the blocking voltage of the IGBT, which may result in the IGBTs overvoltage failure in a short time [3]. For the intermediate dc circuit, there are few research works on the fault diagnosis method for the series resonant circuit.

Model-based methods are the most widely used methods for the diagnosis of open-circuit faults of IGBTs [4], [5]. More detailedly, a generalized model-based method is proposed in [4] for fault detection and identification, where the main idea is to make full use of the dynamic characteristics of error residual. In addition, abundant state-observer-based and state-estimator-based methods are also introduced for open-circuit fault diagnosis [6], [7], [8], [9], [10]. It is worth pointing out that the diagnosis performance of the model-based methods is highly dependent on the model accuracy, which, however, is hard to be guaranteed, especially when considering the complexity of the traction system as well as the effect of current direction change on the operation of antiparallel diodes.

On the other hand, signal-processing-based [11], [12] and machine-learning-based [12], [13] methods have become more and more popular in the field of power system fault diagnosis. Considering that the harmonics near the switching frequency would be different under different conduction conditions, a frequency-domain-analysis-based method is introduced for the open-circuit fault diagnosis of IGBTs in [11]. Further, some relatively intelligent methods, such as vector machine and fuzzy network [12], are brought into open-circuit fault diagnosis. In order to improve the diagnosis efficiency of the neural network (NN) on open-circuit fault, an improved adaptive hierarchical genetic algorithm is used to optimize the neuron weights [13]. It is important to select an appropriate signal for fault diagnosis, which can directly affect the results. Current-based methods are favored for converter fault diagnosis since no extra sensors are needed and they are independent of converter parameters [14]. In [15], a real-time fault diagnosis based on absolute normalized current and adaptive threshold for converter open-circuit (OC) faults is proposed in doubly fed induction generator (DFIG)-based wind turbine systems, and this method can detect multiple open-circuit faults, but it is prone to false alarms when the load changes suddenly. In [16], a current residual-based method is proposed, and converter faults are diagnosed by comparing with adaptive threshold. When a fault occurs in different converter switches, converter voltages present different characteristics. Therefore, it is possible to diagnose converter switch faults by analyzing voltage signals [14]. In [17], the switched bridge voltage and corresponding duration are measured to detect inverter open-circuit faults, and this method has a short detection time, but requires extra circuit, which increases the implementation cost. The dc-link voltage can also be used for fault diagnosis [18]. Jung et al.[19] used grid voltage phase angle and dc-link voltage to diagnose faults. The time–frequency analysis of dc-link voltage is used to obtain converter fault states [20].

Signal-based and machine-learning-based fault diagnosis methods are different from model-based methods in that accurate modeling may not be compulsory, and the fault is diagnosed by analyzing the amplitude–frequency features of voltage and current. As such, the effectiveness of data-driven methods is heavily affected by the performance of feature extracting, where the typical ways include fast Fourier transform (FFT) [11], short-time Fourier transform (STFT) [21], and wavelet transform (WT) [22], [23]. However, STFT has a single resolution for a given window function because only its position can be moved and its shape cannot be changed. Alternatively, the wavelet function of WT is designed to not only move in the time domain but also scale and shift the shape of the transformation window. Noting that STFT and WT are subjective since both of the window functions need to be preset, empirical mode decomposition (EMD) [24], [25] and two improved versions of EMD, ensemble empirical mode decomposition (EEMD) [26] and complementary ensemble empirical mode decomposition (CEEMD) [27], are proposed to further degrade the subjectivity in analyzing nonlinear nonstationary timing signal. In [28], the measured signal that has been decomposed via the variational mode decomposition (VMD) is used to extract the feature parameters according to the properties of multiple parameter faults in power electronic circuits.

Overall, the aforementioned feature extraction methods are basically applicable for shallow feature extraction, where the features include amplitude information, frequency information, energy information, etc. By analyzing the abovementioned characteristic information, the corresponding relationship between the fault signal and the fault mode could be established. However, it is often difficult to establish such a relationship directly through the characteristics of the data itself. In order to extract those deeper features, deep learning is introduced for fault diagnosis [29], [30], [31], [32], [33], [34], [35], [36], [37], [38]. For instance, based on Gaussian–Bernoulli deep belief network, a fault diagnosis method is proposed to reduce the loss of feature information [30]. In [31], deep neural network (DNN) is used to extract fault information of high-speed train bogie adaptively, which performs well in bogie fault diagnosis. In [32], WT and DNN are combined to achieve fault diagnosis of microgrids.

The two most important parameters associated with a signal are its frequency and amplitude. Thus, the faults of the rectifier can be located by analyzing the changes in the abovementioned two parameters. Considering that VMD is able to decompose mixed signals with complex frequencies into a series of stable signals, this article uses VMD to decompose fault signals. Then, given that CRNN can establish the mapping relationship between fault features and fault modes through autonomous learning, the effective features extracted by VMD are used as input signals of CRNN to train the fault classification model for fault diagnosis of the rectifier. Furthermore, current-based

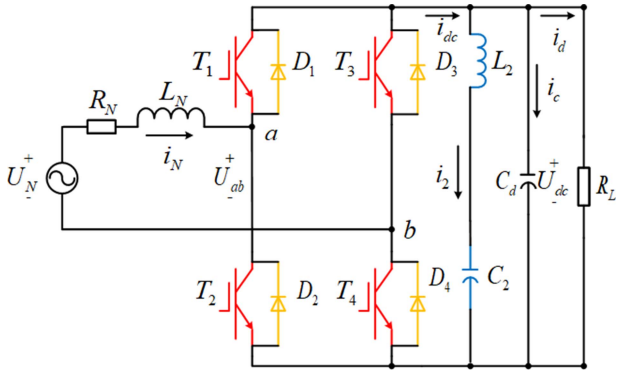

Fig. 2. Topology of a single-phase rectifier.

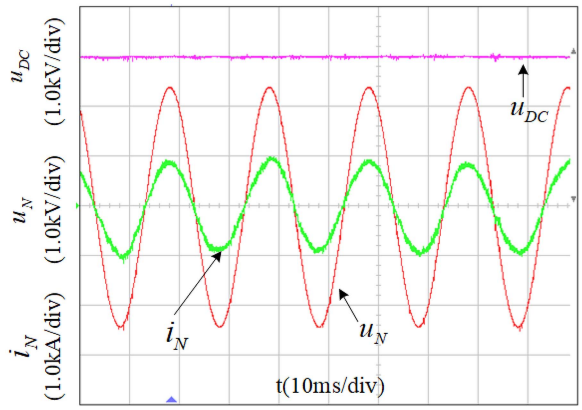

Fig. 3. Waveforms of DC-side voltage U_{DC}, grid-side voltage U_N, and grid-side current i_N in a steady state.

methods and voltage-based methods are quite common in fault diagnosis but few combine them. In this article, the voltage and current submodels are established separately to fix the problem that the characteristics of IGBTs open-circuit, diodes open-circuit, series resonant circuit, and open and short circuit are similar. Overall, the signal-analysis-based fault diagnosis method is combined with an additional deep learning scheme, which not only enhances the effectiveness of the extracted features but also makes the fault diagnosis model more intelligent.

II. Fault Modeling and Analysis of Rectifier

The topology of the single-phase rectifier is shown in Fig. 2. It consists of two bridge arms, each of which contains two IGBTs and two diodes. Further, L_2 and C_2 form a resonant circuit, C_d is the supporting capacitance of the dc side, and R_N and L_N denote the winding leakage inductance and resistance, respectively. Moreover, U_N and i_N represent the grid-side voltage and current, respectively, and U_{ab} is the input voltage of the rectifier. In this article, the series resonant circuit is considered to be part of the rectifier. Since short-circuits fault of IGBT can be converted into open-circuit fault by adding a fast fuse and other hardware protection circuits, only open-circuit faults of IGBTs and diodes along with open and short-circuit faults of series resonant circuit are considered here. In addition, the proposed fault diagnosis scheme is based on the implicit assumption that the load and inductance values of the system are kept invariant during the process of fault diagnosis. When the circuit is normal, the grid-side current i_N would be a sinusoidal signal and the voltage of dc side U_{dc} has slight pulsation but close to a straight line, as shown in Fig. 3.

A. Analysis of IGBT Open-Circuit Fault

The direction in which the grid-side current i_N flows into the rectifier is defined as the reference direction. When T_1 fails, if the grid-side current i_N is positive, the current can flow continuously through diode D_1 without affecting the normal operation of the rectifier. If i_N is negative, the diode D_1 cannot continue to flow, resulting in the change of the topology, as shown in Fig. 4. More specifically, it can be divided into the following two conditions when an open-circuit fault occurs in T_1.

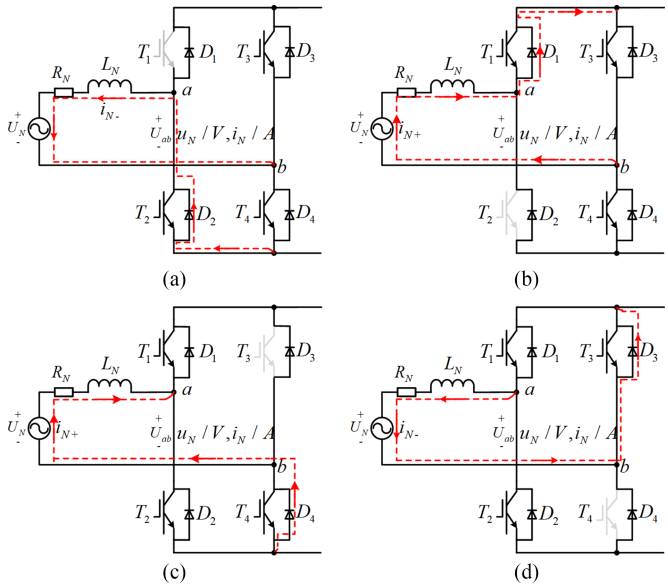

Fig. 4. Schematic diagram of grid-side current flow when open-circuit faults occur in IGBTS. (a) T1. (b) T4. (c) T2. (d) T3.

1) The grid-side current flows through diodes D_3 and D_2. Traction winding U_N and inductor L_N charge the capacitor of dc side C_d simultaneously, making i_N decrease and U_{dc} increase.
2) The grid-side current flows through T_4 and D_2. Capacitor C_d supplies power to the load R_L and dc voltage U_{dc} releases energy by forming a loop through the load to release energy, leading U_{dc} to drop. On the other hand, Only U_N is applied to inductor L_N. Therefore, compared with the normal state, the amplitude of i_N decreases and the inductive energy storage decreases.

To sum up, when open-circuit faults occur in T_1, the positive part of i_N is not affected while the negative part of i_N and the amplitude of U_{dc} decrease, as shown in Fig. 5(a). Fig. 5(b) shows that the waveform of i_N under T_2 failure is completely opposite to that under T_1 failure. Since T_1 and T_4 and T_2 and T_3 belong to one bridge arm, respectively, all the previous explanations can be extended to the remaining switches, T_2, T_3, and T_4.

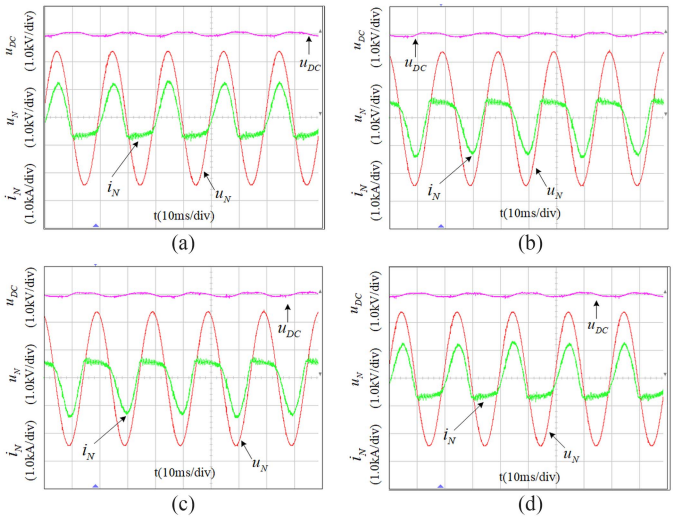

Fig. 5. Waveforms of DC-side voltage U_{DC}, grid-side voltage U_N, and grid-side current i_N: (a) under T_1 failure, (b) under T_2 failure, (c) under T_3 failure, and (d) under T_4 failure.

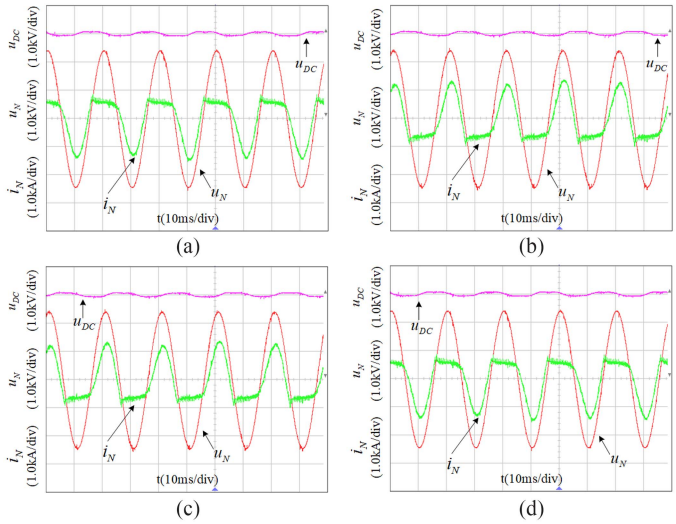

Fig. 6. Waveforms of DC-side voltage U_{DC}, grid-side voltage U_N, and grid-side current i_N: (a) under D_1 failure, (b) under D_2 failure, (c) under D_3 failure, and (d) under D_4 failure.

B. Analysis of Diode Open-Circuit Fault

The direction in which the grid-side current i_N flows into the rectifier is defined as the reference direction. When D_1 fails, if the grid-side current i_N is negative, i_N flows through T_1 and D_2, leaving D_1 in an inoperative state. Hence, the open-circuit fault will not affect the normal operation of the rectifier. If the grid-side current i_N is positive, the current has no way to flow through unless the switch T_2 is turned ON. Since inductor current cannot change immediately, the input voltage of the rectifier U_{ab} rises sharply based on Kirchhoff law, which may lead to tremendous voltage stress on T_1 and T_2. Therefore, the abnormal voltage stress caused by the diode open-circuit fault is much higher than the blocking voltage of the IGBT, which may result in the IGBTs overvoltage failure in a short time. The waveform of the rectifier under diodes failure is shown in Fig. 6.

C. Analysis of Series Resonant Circuit Fault

The series resonant circuit consists of a capacitor and an inductor. They both have the potential for open-circuit and short-circuit failures. The topologies in the event of failures are shown in Fig. 7.

When an open-circuit fault occurs on L_s, the voltage of C_s is discharged through the discharge resistor R_s. The intermediate dc circuit loses the function of filtering out the 100 Hz pulse voltage, resulting in increased voltage fluctuations on the dc side. The open-circuit fault of the inductor will not immediately cause severe fluctuations on U_{dc}, but the input of the resonant voltage to the traction motor will increase the harmonic content of the motor stator current and increase the electromagnetic torque ripple, which is likely to cause secondary faults in the traction drive system. The increase of the stator current harmonics and torque ripple of the traction motor will further aggravate the fluctuations on U_{dc}, which will affect the control of the rectifier, increase the grid-side current harmonic content, and eventually lead

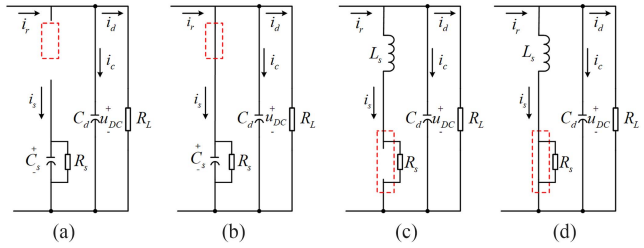

Fig. 7. Topology of rectifier when open-circuit faults occur in series resonant circuit. (a) Inductor open circuit. (b) Inductor short circuit. (c) Capacitor open circuit. (d) Capacitor short circuit.

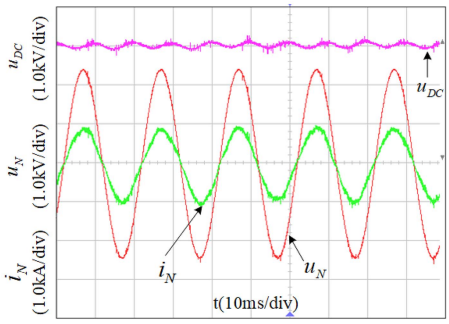

Fig. 8. Waveforms of DC-side voltage U_{DC}, grid-side voltage U_N, and grid-side current i_N under series resonant open circuit.

to the collapse of the traction drive system. Fig. 8 demonstrates the waveforms of the main parameters.

When L_s experiences a short-circuit fault, C_s is connected in parallel with the support capacitor C_d. The intermediate dc circuit loses the function of filtering out the 100 Hz pulse voltage, resulting in increased voltage fluctuations on the dc side. The impact of the resonant inductor short circuit is similar to that of the open-circuit fault. However, since the resonant capacitor value (4.56 mF) is larger than the support capacitor value (3 mF), it is equivalent to increasing C_d. Therefore, when the

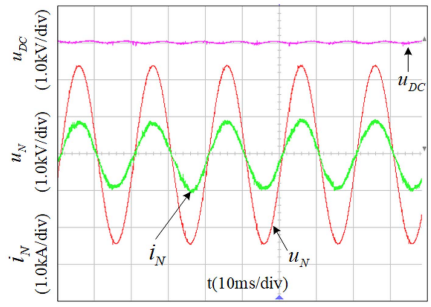

Fig. 9. Waveforms of DC-side voltage U_{DC}, grid-side voltage U_N, and grid-side current i_N under short circuit of L_s.

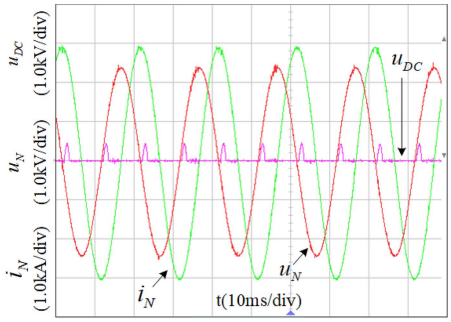

Fig. 10. Waveforms of DC-side voltage U_{DC}, grid-side voltage U_N, and grid-side current i_N under short circuit of C_s.

resonant inductor has a short-circuit fault, the voltage fluctuation of U_{dc} is smaller than that of U_{dc} when open-circuit faults occur. Waveforms of the main parameters are shown in Fig. 9.

When an open-circuit fault occurs on C_s, this condition is similar to the open-circuit fault of the inductor. Thus, these two conditions are unified as the series resonant open-circuit fault.

When C_s experiences a short-circuit fault, L_s is connected in parallel with C_d, which may bring about the LC parallel resonance, leading to a sharp increase in the voltage fluctuation of U_{dc}. Meanwhile, when C_s is short circuited, the voltage across the capacitor is rapidly discharged, which can easily lead to the destruction of the capacitor. Fig. 10 shows the waveforms of the main parameters.

D. Preliminary Summary

The similarity of T_1 and T_4 (T_2 and T_3) fault signals makes it difficult to correctly locate the failed IGBT. For the same reason, the fault diagnosis of D_1 and $D_4(D_2$ and $D_3)$ is also hard. Besides, diodes and IGBTs can sometimes misidentify each other due to the similarity of their waveform. From Figs. 8 and 9, it can be seen that the grid side current i_N is not affected by open circuit and short circuit of the inductor, the waveform of which is exactly the same as the normal case, making it troublesome to distinguish them by i_N alone. Due to all the abovementioned problems, additional means are required for comprehensive fault diagnosis.

III. Theoretical Background

A. Principle of VMD

VMD assumes that any signal is composed of a series of subsignals with a specific center frequency and limited bandwidth [i.e., intrinsic modal component (IMF)] [39]. The center frequency and bandwidth limit are obtained by solving the variational problem, and the corresponding effective components of each center frequency in the frequency domain are found to obtain the mode function.

Different from the concept of IMF proposed by Huang et al. [24], the VMD algorithm redefines the limited-bandwidth IMF, which is defined as the component mode of AM–FM signals with the expression given as follows with stricter constraints:

$$u_k(t) = A_k(t)\cos(\varphi_k(t)) \tag{1}$$

where $\varphi_k(t)$ is the signal phase and $A_k(t)$ is the instantaneous amplitude.

Utilizing the constrained variational model, K IMFs with specific sparsity are sought, such that the sum of each component's estimated bandwidth is minimized, and the constraint condition is that the sum of each component is equal to the original signal. In order to obtain the K IMFs with limited bandwidth, the single marginal spectrum of each IMF component $u_k(t)$ is achieved first through Hilbert transform, and then, the center frequency ω_k of each IMF is estimated, multiplying its exponential signal $e^{-j\omega_k t}$ to modulate the modal spectrum to the corresponding baseband. In the end, the square norm of the analytic signal gradient L^2 is calculated, constructing the variational mode as follows:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left(\sum_{k=1}^{K} \left\|\partial_t\left[\left(\delta_t + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t}\right\|_2^2\right) \\ \text{s.t.} \sum_{k=1}^{K} u_k(t) = f(t) \end{cases} \tag{2}$$

where K is the number of mode decompositions, δ_t represents the unit impulse function, j is an imaginary unit, $*$ represents convolution operation, and ∂_t represents partial derivative operation. $\{u_k\} = \{u_1, u_2, u_3, \ldots, u_K\}$ denotes the K IMFs and $\{\omega_k\} = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_K\}$ denotes the center frequency of each component.

The penalty factor α and the Lagrange multiplier λ are introduced to transform the constrained variational problem into an unconstrained variational problem to solve the abovementioned variational problem, and the augmented Lagrange expression is given as follows:

$$\begin{aligned} &L(\{u_k\},\{\omega_k\},\lambda) \\ &= \alpha \sum_{k=1}^{K} \left\|\partial_t\left[\left(\delta_t + \frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t}\right\|_2^2 \\ &\quad + \left\|f(t) - \sum_{k=1}^{K} u_k(t)\right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t)\right\rangle. \end{aligned} \tag{3}$$

The alternate direction multiplier algorithm is used to update and iteratively solve the saddle point of (3) to obtain the optimal

Algorithm 1: Variational Mode Decomposition.

Initialize $\{\hat{u}_k^1\}, \{\omega_k^1\}, \hat{\lambda}_k^1, n \leftarrow 0$

Repeat

$n \leftarrow n + 1$

for $k = 1 : K$ **do**

Update $\hat{u}_k$ for all $\omega \geq 0$:

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) - \sum_{i=1}^{k-1} \hat{u}_i^{n+1}(\omega) - \sum_{i=k+1}^{K} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}_i(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{4}$$

end for

for $k = 1 : K$

Update ω_k:

$$\omega_k^{n+1}(\omega) \leftarrow \frac{\int_0^\infty \omega \left|\hat{u}_k^{n+1}(\omega)\right|^2 \mathrm{d}\omega}{\int_0^\infty \left|\hat{u}_k^{n+1}(\omega)\right|^2 \mathrm{d}\omega} \tag{5}$$

end for

Dual ascent for all $\omega \geq 0$:

$$\hat{\lambda}^{n+1}(w) \leftarrow \hat{\lambda}^n(w) + \tau \left(\hat{f}(w) - \sum_{k=1}^{K} \hat{u}_k^{n+1}(\omega) \right) \tag{6}$$

until convergence:

$$\sum_{k=1}^{K} \left(\|\hat{u}_k^{n+1}(\omega) - \hat{u}_k^n(\omega)\|_2^2 / \|\hat{u}_k^n(\omega)\|_2^2 \right) < \varepsilon \tag{7}$$

solution that can decompose the original signal into K IMFs. Algorithm 1 reveals the implementation process. An update of the corresponding variable is expressed by (4)–(6), where τ is the fidelity coefficient, $\wedge$ stands for the Fourier transform, and n is the number of iterations.

When it comes to solving the variational model, the center frequency and bandwidth of the IMF are continuously updated until the iterative stop condition is met, and the expression can be seen in (7), where ε is the discriminant accuracy, which is taken as 10^{e-6} here.

When the iteration stops, the frequency domain characteristics of the signal have been adaptively separated. The modulated signal $\hat{u}_k(\omega)$ is transformed into a time-domain IMF by inverse Fourier transform.

VMD is an adaptive, fully nonrecursive method for modal variation and signal processing, which gives VMD many unique advantages.

1) The desired number of modes can be assigned.
2) The IMFs decomposed by VMD method have independent center frequencies, showing the characteristics of sparsity in the frequency domain, which have the quality of sparse research.

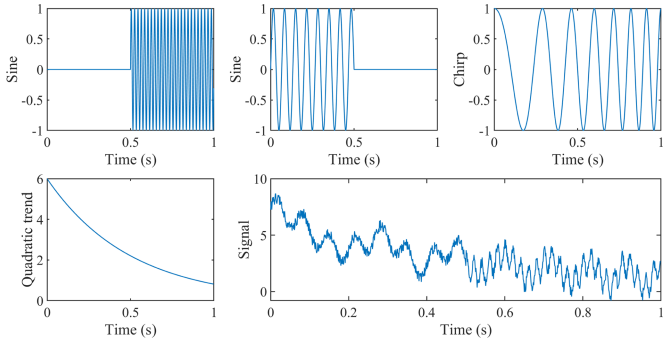

Fig. 11. Four signal components and the final synthesized signal.

3) In the process of solving the IMF, the end effect similar to the EMD method is avoided by means of mirror extension.
4) Mode mixing can be avoided effectively (if the value of K is selected properly).

B. Analysis of Decomposition Results of Simulated Signal

The principle of VMD is introduced in detail in the previous section. Next, the performance of the decomposition methods will be shown with the signal in Fig. 11 as the original decomposition signal. This signal is composed of quadratic trend term, chirp signal, and piecewise sine, the sampling frequency of which is 1000 Hz. Besides, an additional random noise signal is added to simulate the effect of noise in actual situations. This level of composite signal can be considered quite complicated.

The effect of VMD decomposition is mainly affected by the selected value of the mode number K. If the value of K is small, since the VMD algorithm is equivalent to an adaptive filter bank, some important information in the original signal will be filtered, affecting the accuracy of subsequent classification. Nevertheless, when the selected value of K is large, the center frequencies of adjacent IMF will be close to each other, resulting in mode mixing or additional noise. The main difference between IMFs is that the center frequency of each IMF varies from each other and corresponds to the FFT result. Therefore, appropriate K is selected by observing the distribution of center frequency with the corresponding spectrum.

Fig. 12(a) and 12(b) shows the VMD results when $K = 4$ and $K = 5$, respectively. When $K = 5$, the center frequency corresponding to IMF1 cannot be found in the spectrum of the original signal, which means this IMF is useless. In order to show the superiority of the VMD method, the commonly used CEEMD, which has made great improvements on the basis of EMD and EEMD, is selected here for comparison. It can be seen that both VMD and CEEMD perform well. Specifically, combined with the frequency spectrum, it can be seen that many spurious and useless components, such as IMF1, IMF2, IMF6, IMF7, and IMF8, are generated after CEEMD, making it hard to choose the right and useful IMF. On the other hand, benefitting from the VMD's advantage that the number of decomposed modes can be selected, the value of K can be chosen as four according to the FFT result. Compared with that of CEEMD,

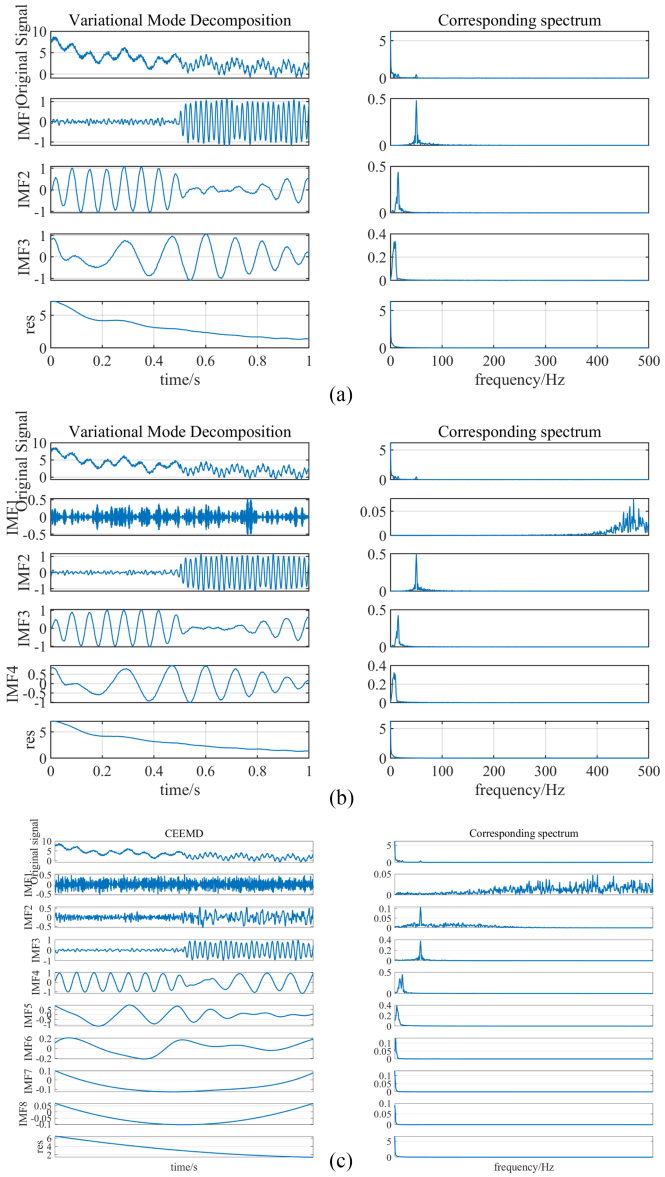

Fig. 12. Decomposition results of mixed signal. (a) Decomposition results of VMD ($K = 4$). (b) Decomposition results of VMD ($K = 5$). (c) Decomposition results of CEEMD.

the decomposed result of VMD is in perfect agreement with the composition of the synthesized signal. In such sense, VMD is superior to CEEMD in decomposing complex fixed signal.

Fig. 13 shows the VMD results of grid-side current when $T1$ fails. K is chosen as 3 according to the FFT result. It can be seen that the center frequency of each IMF can find a good correspondence in the spectrum of the original signal, indicating that 3 is the appropriate value of K. Other fault conditions are the same.

C. Building of Fault Classification Network CRNN

Considering that the faults of IGBTs and diodes are related to current while the faults of series resonant circuit are more related to voltage, two independent feature extraction submodels are

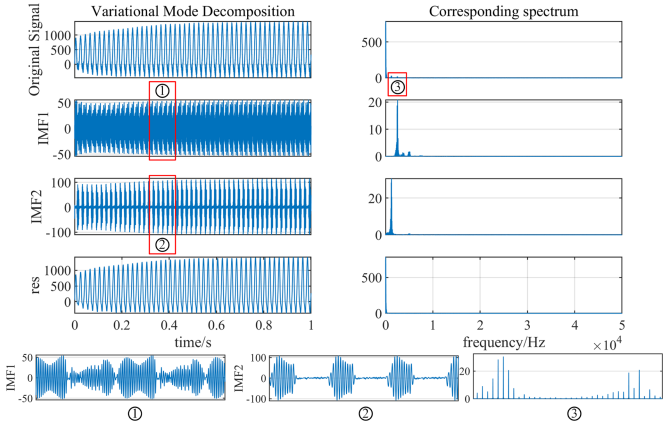

Fig. 13. VMD results of grid-side current when $T1$ fails ($K = 3$).

adopted to process the voltage and current signals, respectively. Since the features of the signal are not obvious and the signal processing method cannot extract the signal features entirely, CRNN is used as the model for the fault diagnosis. Fig. 14 shows the structure of CRNN. The framework contains the following five parts.

1) After the output of VMD is generated, it is used as an input for a 1-D convolutional block. This block is composed of 1-D convolutional layers and 1-D pooling layers in an alternating pattern.
2) The feature maps output by the last convolutional layer are unstacked over the time axis.
3) The feature maps of the unstack layer are passed to the recurrent layer.
4) The tanh activation function helps the fully connected layers to understand the outputs of the last recurrent layer so that the useful information is retained.
5) The output layer with the softmax function figures out the probability of samples for each class.

D. 1D-CNN Block

The 1D-CNN block generally consists of two parts: 1) the first part performs the convolution operation to extract features; 2) the second part then adjusts the output of the convolutional layers through the pooling operation. The sparsity of CNN weights enables the detection of small but significant features through the use of convolutional filters that are smaller than the input. This implies that CNN requires fewer parameters to be saved and greatly enhances the effectiveness of feature extraction.

In a convolutional layer, the feature maps x_j^l are obtained through convolution operation for lth convolutional layer ($l \in l_c$), which is elaborated in the following:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{8}$$

where k_{ij} and b_j^l represent the weight and bias of jth convolutional filter, respectively, M_j is the number of input feature maps, and $*$ represents convolution operation.

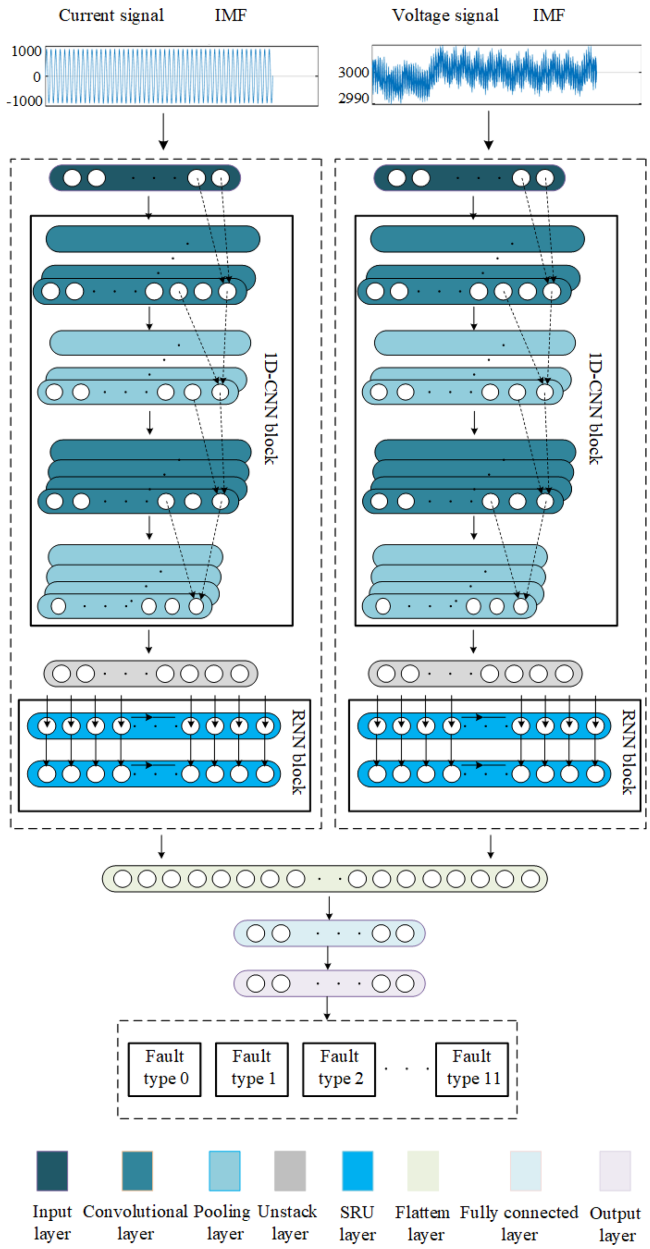

Fig. 14. Structure of DCRNN.

Commonly used activation functions are rectified linear unit (ReLU), Tanh, etc. The ReLU function overcomes the problem of gradient disappearance in the positive range, having the advantages of fast calculation and fast convergence speed. When the input is negative, ReLU is not activated at all. Although it enhances the sparsity of the network and alleviates overfitting, it loses negative value features in the feature map, suitable for stable and symmetrical waveform data. Therefore, the current submodel uses the ReLU activation function for the current waveform which is relatively symmetrical. The output range of the Tanh function is $(-1, 1)$, the waveform of which is centrosymmetric, and the nonlinear monotonic relationship between the input and output is maintained. As a result, the Tanh function can handle both positive and negative eigenvalues in the feature map. Despite the problems Tanh function may

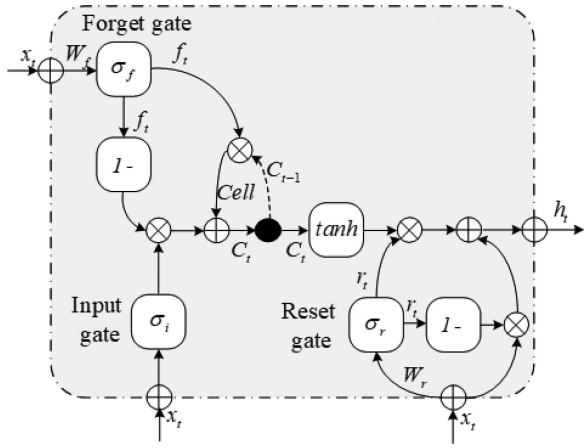

Fig. 15. Structure of SRU.

have, such as gradient disappearance in the saturated region, it is still suitable for processing asymmetric signals. In light of the characteristics of voltage, the Tanh activation function is applied to the voltage submodel. The pooling layer is mainly used for feature dimensionality reduction and data compression to avoid overfitting and ensure the generalization ability of the network. Max pooling and average pooling are the two most common pooling methods. When it comes to processing texture information, max pooling is superior to average pooling. In practice, the voltage and current data contain a small amount of harmonic information, of which the max pooling method can reduce the influence. The max pooling is described as follows.

$$\hat{x}_j^l = f\left(\beta_j^l \text{down}(x_j^l) + b_j^l\right) \tag{9}$$

where β_j^l and b_j^l represent the weight and bias of max pooling, and down() means the max pooling function. After performing the operation in l_c layers, the output of the CNN block is a tensor with deep features, which contains the most effective information in a small dimension.

E. RNN Block

RNN is applicable to tasks, such as natural language processing and time series modeling. Among the RNNs, long short-term memory (LSTM), and simple recurrent units (SRU) stand out. LSTM's ability to extract time series features makes it better able to use historical data to characterize fault signals. However, the long-term dependence of LSTM on time series makes parallel computation difficult. SRU preprocesses the input data relatively independently in each recursive extraction of time series features, and then performs recursive feature extraction through relatively lightweight parallel operations. Fig. 15 shows the structure of SRU. The output lossy summary h_t of SRU can be calculated as follows:

$$\hat{x}_t = W \cdot x_t \tag{10}$$

$$f_t = \sigma_f(W_f \cdot x_t) + b_f \tag{11}$$

$$r_t = \sigma_r(W_r \cdot x_t) + b_r \tag{12}$$

$$c_t = f_t \odot c_{t-1} + (1 - f_t) \odot \hat{x}_t \tag{13}$$

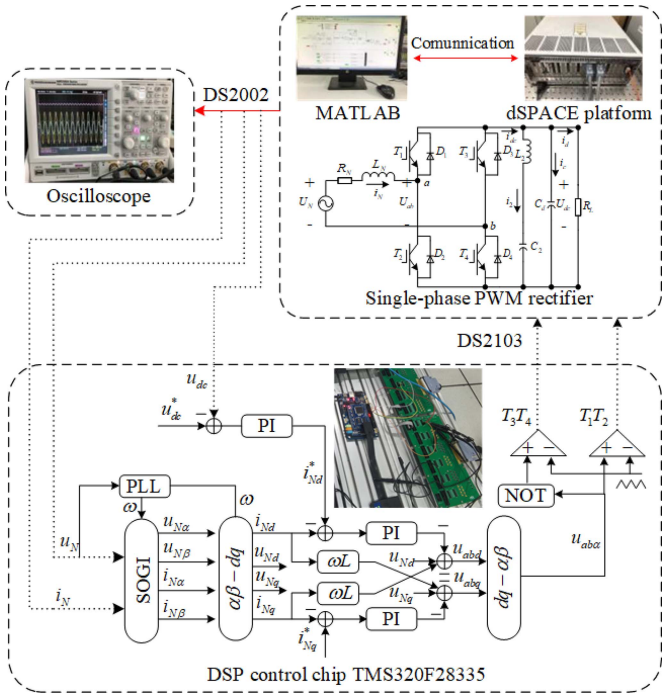

Fig. 16. dSPACE experimental setup.

$$h_t = r_t \odot \tanh(c_t) + (1 - r_t) \odot \hat{x}_t \tag{14}$$

where σ_f, σ_r and b_f, b_r are the sigmoid functions and bias term of forget gate f_t and reset gate r_t, respectively. Besides, W, W_f, and W_r denote the weight of linear mapping $\hat{x}_t$, forget gate, and reset gate, respectively. $\odot$ indicates that the corresponding elements are multiplied. The status of gates $(\hat{x}_t, f_t, r_t)$ no longer relies on historical iteration state h_{t-1}, and reset gate is introduced to improve time sequence feature extraction. In such a way, SRU realizes parallel acceleration from the aspects of structure and operation optimization.

IV. Experimental Results

A. Collection of Raw Data

In this section, the fault data based on dSPACE real-time simulator will be presented to verify the effectiveness of the proposed method. The dSPACE can control virtual objects with a real controller to achieve extreme simulation testing of the real environment. Thus, using dSPACE-based data as measured data have a better verification effect on the effectiveness of the proposed method. The data acquisition setup based on dSPACE is shown in Fig. 16. It consists of a digital signal processor (DSP), as the executor of the control algorithm, of which the control chip is TMS320F28335, a dSPACE simulator, and a host PC as the control interface. PWM pulse signal is input to dSPACE by the DS2103 board. The grid voltage U_N, the input current of rectifier circuit i_N, and dc-side voltage $U_{\rm dc}$ of the main circuit are output to the AD sampling module of the DSP by the DS2002 board and displayed by the oscilloscope. The DQ current decoupling control strategy is widely used in the field of electric locomotives due to its excellent steady-state performance. In Fig. 16, PLL is a phase-locked loop, $u^*_{\rm dc}$ is the given dc voltage, $u_{ab\alpha}$ is the voltage modulation signal, and $\alpha\beta/dq$ and $dq/\alpha\beta$ represent the rotation coordinate relationship. Main circuit parameters of the hardware-in-the-loop experiment system are given in Table I.

TABLE I
System Parameters of Single-Phase PWM Rectifier

Electrical parameters	Value
Voltage of grid side $u_{\rm sm}$/V	2200
Reference voltage of dc side $u^*_{\rm dc}$/V	3000
Support capacitor of dc side $C_{\rm d}$/μF	3000
Equivalent resistance of grid side R/Ω	0.068
Equivalent inductance of grid side L/mH	2.2
Equivalent load of dc side $R_{\rm L}/\Omega$	10
Secondary filter capacitor C_2/μF	4560
Secondary filter inductance L_2/mH	0.55
Switching frequency $f_{\rm pwm}$/Hz	1250
Proportion parameter of outer voltage loop	0.6
Integral parameter of outer voltage loop	20
Proportion parameter of inner current loop	0.7
Integral parameter of inner current loop	38

TABLE II
Twelve Kinds of Working Conditions of the Rectifier

Label	Fault category	Label	Fault category
0	Normal	6	$D2$ Open-circuit
1	$T1$ Open-circuit	7	$D3$ Open-circuit
2	$T2$ Open-circuit	8	$D4$ Open-circuit
3	$T3$ Open-circuit	9	Series resonant open-circuit
4	$T4$ Open-circuit	10	Inductor short-circuit
5	$D1$ Open-circuit	11	Capacitor short-circuit

The conduction situations of the rectifier have been analyzed in Section II. Furthermore, two working conditions converter overloaded and voltage spike are added. A sudden change in load resistance from 10 to 5 Ω is used to simulate the rectifier overloaded condition. DC-side reference voltage $U_{\rm dc}$ spikes from 3000 to 3500 V to simulate voltage spike. The faults are set to occur 0.1 s after rectifier overloaded and voltage spike. Taking half-cycle data as a sample, 250 samples are collected for each of the 12 fault types, the sample size of which is 1000×1. As for the training/test split, the percentage of the data that should be held over for testing is 0.2. Therefore, the dataset has 9000 samples, of which the training set contains 7200 and the validation set contains 1800. Table II tabulates the fault types of the rectifier.

B. Fault Identification With Comparative Methods

To verify the effectiveness of using dual channels, experiments are performed first on the single channel of CNN, CRNN, and VMD-CRNN, respectively. For the convolutional module in both 1D-CNN and CRNN model, the $\text{Layer}_{\text{1D-CNN}}$ is set to 2. With regard to the method combining VMD with CRNN, the number of $\text{Layer}_{\text{1D-CNN}}$ is correspondingly increased to three because the sample size of the input is 1000*3. Each

TABLE III
RESULTS OF THE COMPARATIVE METHODS

Index	1D-CNN layers(cells)	SRU layers(cells)	Accuracy(%)
CNN-I	2(12/24)	-	54.5
SRU-I	-	2(50*2)	55.4
CRNN-I	2(12/24)	2(50*2)	58.3
VMD-CRNN-I	3(12/24/48)	2(50*2)	66.3
CNN-U	2(12/24)	-	33.7
SRU-U	-	2(50*2)	29.9
CRNN-U	2(12/24)	2(50*2)	37.9
VMD-CRNN-U	3(12/24/48)	2(50*2)	43.3

layer consists of convolutional and max-pooling layers. Detailed parameter settings and results are given in Table III. It can be seen that whether it is a single-channel model of current or voltage, the accuracy is sequentially improved. Furthermore, Fig. 17(a) illustrates the testing accuracies through 2000 iterations for each model. VMD-CRNN outperforms CNN, SRU, and CRNN to a certain degree. The CNN network is the first to reach convergence, and the highest accuracy is 54.5%. Hence, without any preprocessing, the features extracted directly from the data of the same conduction interval are quite similar, so the classification performance is hard to be further improved. Fig. 17(b) reveals the testing loss of each model through 2000 iterations. From Fig. 17(c) and 17(d), a similar conclusion can be drawn: VMD-CRNN outperforms CNN and CRNN. On the other hand, compared with CNN, CRNN has improvement both in accuracy and loss value, especially in loss value. After comprehensive consideration of accuracy and robustness, VMD-CRNN is selected as the model for subsequent experiments.

More clearly, the confusion matrix evaluates the accuracy of classifier classification. The size of the confusion matrix is $n_{\text{classes}} \times n_{\text{classes}}$, where each column represents the class prediction of the sample by the classifier, and each row represents the real class to which the sample belongs.

The confusion matrix of VMD-CRNN (under current dataset) is given in Fig. 18(a), from which it can be seen that labels $y_n \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ can be correctly classified in a certain degree while labels $y_n \in \{0, 9, 10, 11\}$ are misidentified with each other. The reason is that the fault characteristics of IGBT and diode are mostly reflected in current but the fault characteristics of series resonant circuit are not sensitive to current. Fig. 18(b) shows the confusion matrix of VMD-CRNN (under voltage dataset), which reveals the exact opposite of VMD-CRNN (current dataset) because the fault characteristics of series resonant circuit are mostly reflected in voltage. From Table III, it can be seen that the maximum accuracy of all comparative methods is 66.3%, which means, either way, it is not appropriate to perform complete fault diagnosis with a single-channel model. From Fig. 18, it is obvious that the VMD-CRNN (under voltage dataset) and VMD-CRNN (current dataset) exhibit complementary characteristics. Therefore, the dual-channel model is applied in the following experiment.

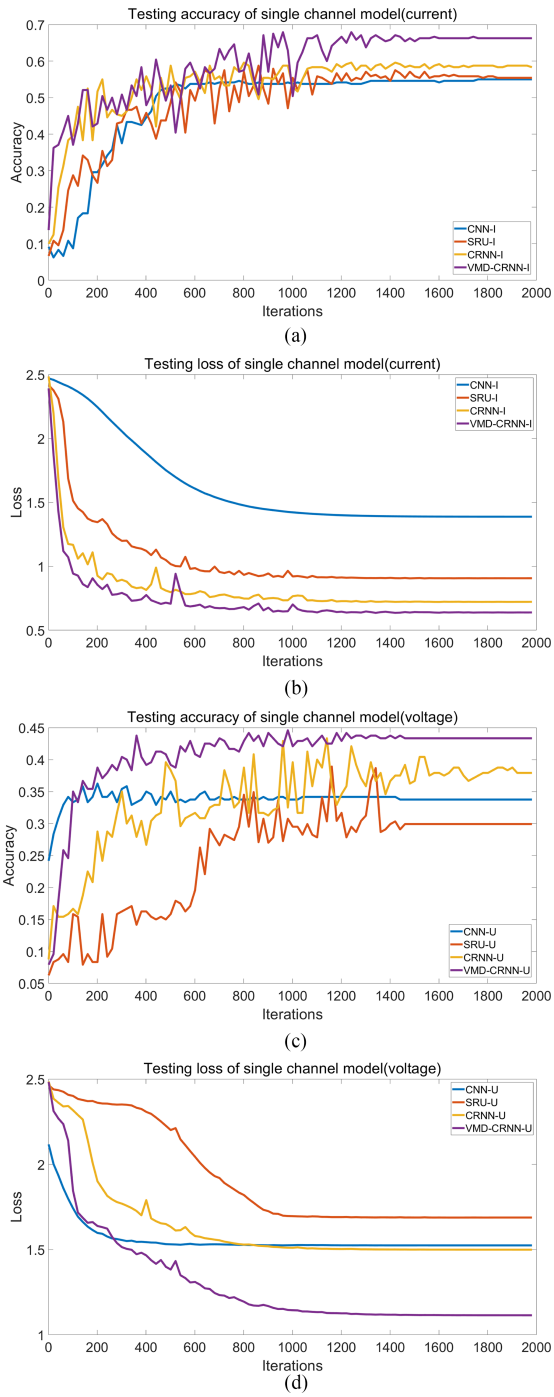

Fig. 17. Testing results of the comparative methods through 2000 iterations. (a) Testing accuracies of single-channel model (current). (b) Testing losses of single-channel model (current). (c) Testing accuracies of single-channel model (voltage). (d) Testing losses of single-channel model (voltage).

C. Fault Identification With VMD-DCRNN

The model establishment of fault type identification is discussed in Section III. The CRNN framework used for fault identification combines three 1D-CNN layers with two SRU layers. This framework not only has 96.27% of fault type identification but also achieves the best robustness. These results demonstrate that, compared with the single-channel models and models using

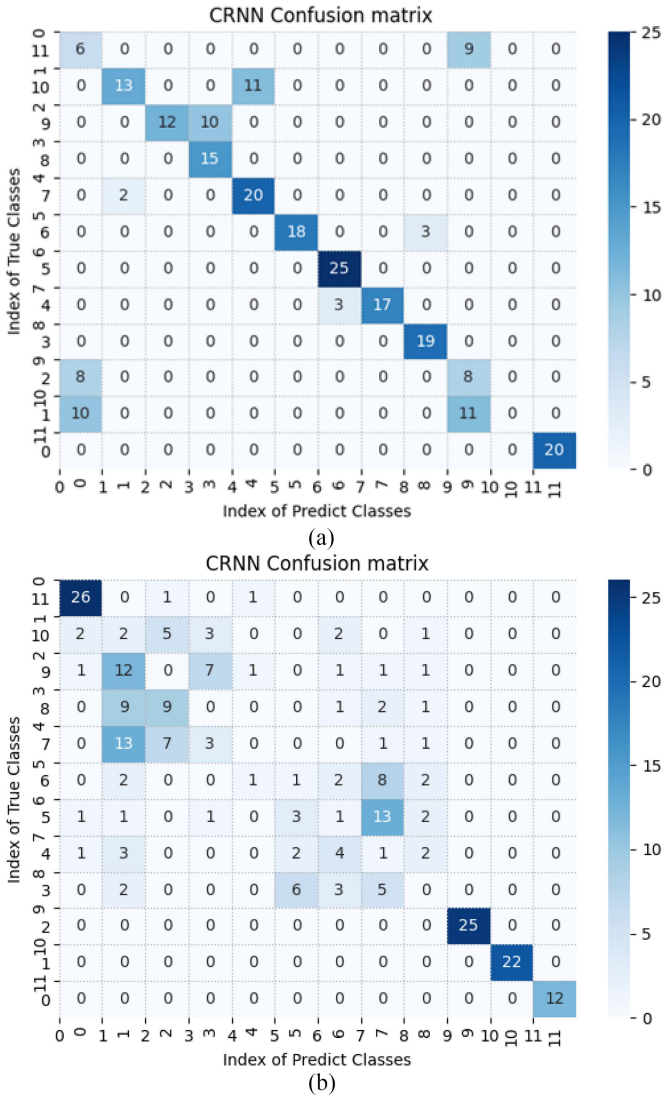

Fig. 18. Confusion matrix of VMD-CRNN. (a) VMD-CRNN for identifying the fault types (under current dataset). (b) VMD-CRNN for identifying the fault types (under voltage dataset).

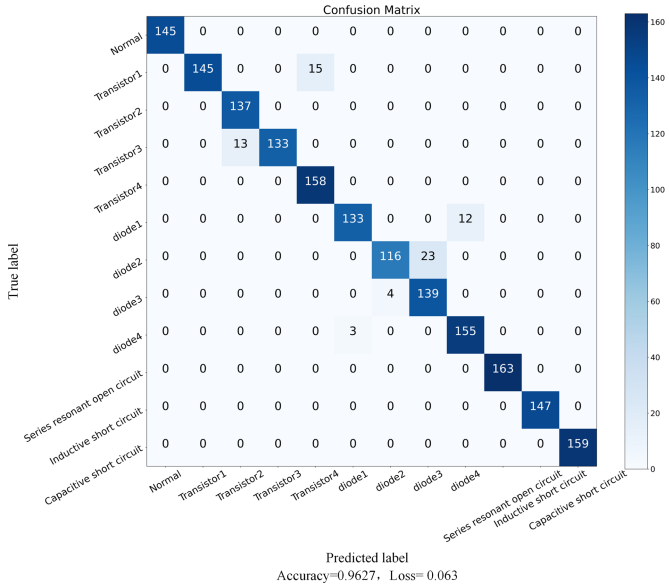

Fig. 19. Confusion matrix of VMD-DCRNN.

TABLE IV
TEST RESULTS AND TIME

Working conditons	Fault category	Test result	Time (s)
Steady state	Normal(0)	0:40	1.64
	$T_1(1)$	(1:27)(4:13)	1.61
	$T_2(2)$	(2:34)(3:4)(9:2)	1.59
	$T_3(3)$	(3:26)(9:11)(2:3)	1.59
	$T_4(4)$	4:40	1.55
	$D_1(5)$	(5:12)(2:8)(8:8)(7:8)(4:4)	1.77
	$D_2(6)$	(6:18)(7:15)(5:7)	1.65
	$D_3(7)$	(7:15)(6:11)(3:11)(2:3)	1.63
	$D_4(8)$	(8:23)(5:11)(6:4)(2:2)	1.63
	L OC(9)	9:40	1.57
	L SC(10)	10:40	1.57
	C SC(11)	(11:36)(3:4)	1.59

data without signal processing as input, the proposed method provides clear advantages by combining signal processing with deep learning. The 1D-CNN layer processes data in the time domain, yielding that the features containing most of the fault information can be considered as the equivariant representations in the time domain. In addition, SRU has the ability to mine hidden timing information used to model time series.

Specifically, the confusion matrix of VMD-DCRNN is shown in Fig. 19, from which it can be seen that the VMD-DCRNN is sensitive to the difference of each fault type. Among the 1800 test samples, 1733 samples are correctly classified. In total, 13 samples' fault type, which should belong to T_3 failure, is misidentified as the T_2 failure. The reason is that T_2 and T_3 have the similar fault characteristics, as discussed in Section II. As such, the features they extract will be similar, yielding classification errors. Other sample classification errors are also for this reason. However, most of the samples are correctly classified, which indicates that the VMD-DCRNN is effective for classifying the fault type of the rectifier.

Table IV gives the results and time of the test using HIL data. The number in the fault category indicates the fault label, which corresponds to Table II in the manuscript. The number of test results represents the number of each category. For example, for T_1 faults under steady state, the test results are (1:27) and (4:13), which means that 27 samples are identified as T_1 faults and 13 samples as T_4 faults. In this way, it is considered that T1 open-circuit fault occurs in the rectifier. The duration of the test is roughly between 1.5 and 1.8 s, and all the faults are correctly categorized.

D. Robustness Verification

Since the condition of voltage spike for training is that the voltage surges to 3500 V, the condition of rectifier overloaded for training is that the load jumps to 5 Ω, and more experiments are done to explore the boundaries of this model. For the voltage

TABLE V
RESULTS OF THE ROBUSTNESS VERIFICATION

Working conditons	Situation	Test result	Loss
Voltage spike	3400V	96.33%	0.1092
	3750V	86.5%	0.3927
	4000V	75.67%	0.6983
Rectifier overloaded	7 Ω	90.67%	0.2995
	4Ω	85.5%	0.4116
	3Ω	46.5%	2.3137

spike condition, the data of U_{dc} surging to 3400, 3750, and 4000 V are selected and put into the trained model. From Table V, the test accuracy under the condition of voltage surging to 3400 V is 96.33%, which means that the rectifier can diagnose faults perfectly when the voltage changes between 3000 and 3500 V. The test accuracy under the condition of voltage surging to 4000 V is 75.67% and some errors cannot be recognized. It can be seen that the model can still complete the fault diagnosis normally when the voltage surges to 3750 V. Therefore, it can be considered that as long as the voltage surge does not exceed 3750 V, this model is competent for the task of fault diagnosis. For the rectifier overloaded condition, the data of R_L decreasing to 7, 4, and 3ω are selected and put into the trained model. In the same way, as can be seen in Table V, it can be considered that as long as the load is not less than 4 Ω, this model is competent for the task of fault diagnosis. To sum up, this model has good robustness.

V. CONCLUSION

This article proposed a fault diagnosis method of rectifier based on VMD-DCRNN, where VMD is used for feature decomposition and DCRNN is used for feature extraction and fault identification. First, the fault current and voltage signal have been decomposed into a series of stable IMFs to simplify the signal adaptively. Next, the selected IMFs of current and voltage have been taken as the input eigenvectors of CRNN, respectively, to further extract the deeper characteristics. Finally, the VMD-DCRNN model has been trained iteratively, and the experimental results demonstrate the efficacy of the proposed method in fault diagnosis.

The proposed method makes full use of the data and characteristics of the grid-side current and the dc-side voltage along with the combination of signal processing and deep learning, which solves the problem of fault diagnosis for almost all components in the single-phase rectifier that are prone to failure.

Future work includes online fault diagnosis and network visualization, so as to better explain the feature extracting process and further guide signal-channel sequencing. In addition, multisensor fault diagnosis of single-phase rectifier is a place where deep learning can flex its muscles. There should be a lot of work that can be done in this direction.

REFERENCES

[1] S. Yang, A. Bryant, P. Mawby, D. Xiang, L. Ran, and P. Tavner, "An industry-based survey of reliability in power electronic converters," *IEEE Trans. Ind. Appl.*, vol. 47, no. 3, pp. 1441–1451, May/Jun. 2011.

[2] M. Alavi, D. Wang, and M. Luo, "Short-circuit fault diagnosis for three-phase inverters based on voltage-space patterns," *IEEE Trans. Ind. Electron.*, vol. 61, no. 10, pp. 5558–5569, Oct. 2014.

[3] B. Gou, X. Ge, S. Wang, X. Feng, J. B. Kuo, and T. G. Habetler, "An open-switch fault diagnosis method for single-phase PWM rectifier using a model-based approach in high-speed railway electrical traction drive system," *IEEE Trans. Power Electron.*, vol. 31, no. 5, pp. 3816–3826, May 2015.

[4] J. Poon, P. Jain, I. C. Konstantakopoulos, C. Spanos, S. K. Panda, and S. R. Sanders, "Model-based fault detection and identification for switching power converters," *IEEE Trans. Power Electron.*, vol. 32, no. 2, pp. 1419–1430, Feb. 2017.

[5] X. Ge, J. Pu, B. Gou, and Y. 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.

[6] S. Tong, B. Huo, and Y. Li, "Observer-based adaptive decentralized fuzzy fault-tolerant control of nonlinear large-scale systems with actuator failures," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 1, pp. 1–15, Feb. 2014.

[7] S. Tong, T. Wang, and Y. Li, "Fuzzy adaptive actuator failure compensation control of uncertain stochastic nonlinear systems with unmodeled dynamics," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 3, pp. 563–574, Jun. 2014.

[8] D. Dolz, I. Pearrocha, and R. Sanchis, "Performance tradeoffs for networked jump observer-based fault diagnosis," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2692–2703, May 2015.

[9] L. Guo, Y.. Zhang, H. Wang, and J. Fang, "Observer-based optimal fault detection and diagnosis using conditional probability distributions," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3712–3719, Oct. 2006.

[10] Y. Liu, Q. Zeng, S. Tong, C. L. P. Chen, and L. Liu, "Actuator failure compensation-based adaptive control of active suspension systems with prescribed performance," *IEEE Trans. Ind. Electon.*, vol. 67, no. 8, pp. 7044–7053, Aug. 2020.

[11] Z. Tian and X. Ge, "An on-line fault diagnostic method based on frequency-domain analysis for IGBTs in traction PWM rectifiers," in *Proc. IEEE 8th Int. Power Electron. Motion Control Conf.*, 2016, pp. 3403–3407.

[12] H. Guo, J. A. Crossman, Y. L. Murphey, and M. Coleman, "Automotive signal diagnostics using wavelets and machine learning," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1650–1662, Sep. 2000.

[13] B. Fan, J. Niu, and J. Zhao, "Three-phase full-controlled rectifier circuit fault diagnosis based on optimized neural networks," in *Proc. 2nd Int. Conf. Artif. Intell. Manage. Scie. Electron. Commerce*, 2011, pp. 6048–6051.

[14] J. Liang, K. Zhang, A. Al-Durra, S. Muyeen, and D. Zhou, "A state-of-the-art review on wind power converter fault diagnosis," *Energy Rep.*, vol. 8, pp. 5341–5369, 2022.

[15] H. Zhao and L. Cheng, "Open-circuit faults diagnosis in back-to-back converters of DF wind turbine," *IET Renewable Power Gener.*, vol. 11, no. 4, pp. 417–424, 2017.

[16] I. Jlassi and A. J. M. Cardoso, "IGBTs and current sensors fault diagnosis in direct-drive PSMG wind turbine systems using adaptive thresholds," in *Proc. 43rd Annu. Conf. IEEE Ind. Electron. Soc.*, 2017, pp. 5072–5077.

[17] D. Chen, Y. Liu, and S. Zhang, "Open-circuit fault diagnosis method for the t-type inverter based on analysis of the switched bridge voltage," *IET Power Electron.*, vol. 12, no. 2, pp. 295–302, 2019.

[18] M. Sen, M. Alaraj, and J.-D. Park, "Open circuit fault detection and localization in modular multilevel converter," in *Proc. North Amer. Power Symp.*, 2016, pp. 1–6.

[19] J.-H. Jung, H.-K. Ku, Y.-D. Son, and J.-M. Kim, "Open-switch fault diagnosis algorithm and tolerant control method of the three-phase three-level NPC active rectifier," *Energies*, vol. 12, no. 13, 2019, Art. no. 2495.

[20] A. Ismail, L. Saidi, and M. Sayadi, "Wind turbine power converter fault diagnosis using dc-link voltage time–frequency analysis," *Wind Eng.*, vol. 43, no. 4, pp. 329–343, 2019.

[21] J. Burriel-Valencia, R. Puche-Panadero, J. Martinez-Roman, A. Sapena-Bano, and M. Pineda-Sanchez, "Short-frequency fourier transform for fault diagnosis of induction machines working in transient regime," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 3, pp. 432–440, Mar. 2017.

[22] I. Bayram and I. W. Selesnick, "Frequency-domain design of overcomplete rational-dilation wavelet transforms," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 2957–2972, Aug. 2009.

[23] G. D. Gonzalez, R. Paut, A. Cipriano, D. R. Miranda, and G. E. Ceballos, "Fault detection and isolation using concatenated wavelet transform variances and discriminant analysis," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1727–1736, May 2006.

[24] N. E. Huang, Z. Shen, and S. R. Lon, "The empirical mode decomposition and hilbert spectrum for nonlinear and nonstationary time series analysis," *Proc. Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.
[25] N. ur Rehman and D. P. Mandic, "Empirical mode decomposition for trivariate signals," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1059–1068, Mar. 2010.
[26] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data. Anal.*, vol. 1, no. 1, pp. 1–41, 2009.
[27] J. R. Y. J. S. Shieh and N. E. Huang, "Complimentary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," *Adv. Adapt. Data. Anal.*, vol. 2, no. 2, pp. 135–156, 2010.
[28] Y. Jiang, L. Chen, W. Zeng, and Y. Xin, "Adaptive weighted VMD-WPEE method of power-electronic-circuit multiple-parameter-fault diagnosis," *IEEE Trans. Emerg. Sel. Topics Power Electron.*, vol. 8, no. 4, pp. 3878–3890, Dec. 2020.
[29] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, "Next-generation Big Data analytics: State of the art, challenges, and future research topics," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1891–1899, Aug. 2017.
[30] Z. Liu, Z. Jia, C. Vong, S. Bu, J. Han, and X. Tang, "Capturing high-discriminative fault features for electronics-rich analog system via deep learning," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1213–1226, Jun. 2017.
[31] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, "Intelligent fault diagnosis of the high-speed train with Big Data based on deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2106–2116, Aug. 2017.
[32] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1694–1703, Mar. 2019.
[33] Y. Liu, Q. Zeng, S. Tong, C. L. P. Chen, and L. Liu, "Adaptive neural network control for active suspension systems with time-varying vertical displacement and speed constraints," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9458–9466, Dec. 2019.
[34] K. Liang, N. Qin, D. Huang, and Y. Fu, "Convolutional recurrent neural network for fault diagnosis of high-speed train bogie," *Complexity*, vol. 5, pp. 1–13, Oct. 2018.
[35] L. Su, L. Ma, N. Qin, D. Huang, and A. H. Kemp, "Fault diagnosis of high-speed train bogie by residual-squeeze net," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 3856–3863, Jul. 2019.
[36] D. Huang, Y. Fu, N. Qin, and S. Gao, "Fault diagnosis of high-speed train bogie based on LSTM neural network," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 119–203, 2021.
[37] J. Yang et al., "An improved deep network for intelligent diagnosis of machinery faults with similar features," *IEEJ Trans. Elect. Electron. Eng.*, vol. 14, no. 12, pp. 1851–1864, Oct. 2019.
[38] A. Murray and J. Penman, "Extracting useful higher order features for condition monitoring using artificial neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2821–2828, Nov. 1997.
[39] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

Na Qin (Member, IEEE) received the B.S. degree in electrical technology from the School of Electrical Engineering, Zhengzhou University, Zhengzhou, China, in 2000, and the M.S. degree in electric power system and automation and the Ph.D. degree in electrical engineering from the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China, in 2003 and 2014, respectively.

She is currently an Associate Professor with the School of Electrical Engineering, Southwest Jiaotong University. Her research interests include intelligent information processing, fault diagnosis, pattern recognition, and intelligent systems.

Tianwei Wang received the B.Eng. degree in electronic information engineering in 2021 from Southwest Jiaotong University, Chengdu, China, where he is currently working toward the M.S. degree in control science and engineering.

His research interests include fault diagnosis, fault tolerant control of rectifier, and intelligent information processing.

Deqing Huang (Senior Member, IEEE) received the B.S. and Ph.D. degrees in applied mathematics from Mathematical College, Sichuan University, Chengdu, China, in 2002 and 2007, respectively, and the second Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 2011.

From 2013 to 2016, he was a Research Associate with the Department of Aeronautics, Imperial College London, London, U.K. In January 2016, he joined the Department of Electronic and Information Engineering, Southwest Jiaotong University, Chengdu, as a Professor and the Department Head. His research interests include modern control theory, fluid analysis and control, convex optimization, and robotics.

Yiting You received the B.S. degree in electronic information engineering from Southwest Petroleum University, Chengdu, China, in 2021. She is currently working toward the M.S. degree in control science and engineering with the Southwest Jiaotong University, Chengdu.

Her research interests include intelligent information processing, fault diagnosis, and deep learning.

Yiming Zhang received the B.S. degree from the Lanzhou Jiaotong University, Lanzhou, China, in 2019. He is currently working toward the Ph.D. degree in electrical engineering with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China.

His research interests include fault diagnosis and deep learning.