Journal of *Marine Science and Engineering*

MDPI

Article

A Layering Linear Discriminant Analysis-Based Fault Diagnosis Method for Grid-Connected Inverter

Guangfeng Jin [1], Tianzhen Wang [1,*], Yassine Amirat [2,*], Zhibin Zhou [2] and Tao Xie [1]

[1] Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China; 201930210074@stu.shmtu.edu.cn (G.J.); 201840210002@stu.shmtu.edu.cn (T.X.)
[2] ISEN Yncréa Ouest, L@bISEN, 20, Rue Cuirassé Bretagne, 29200 Brest, France; zhibin.zhou@isen-ouest.yncrea.fr
* Correspondence: tzwang@shmtu.edu.cn (T.W.); yassine.amirat@isen-ouest.yncrea.fr (Y.A.); Tel.: +86-021-38282640 (T.W.)

Abstract: Grid-connected inverters are the core equipment for connecting marine energy power generation systems to the public electric utility. The variation of current sensor fault severity will make fault samples multimodal. However, linear discriminant analysis assumes that the same fault is independent and identically distributed. To solve this problem, this paper proposes a layering linear discriminant analysis method based on traditional linear discriminant analysis. The proposed method divides the historical fault data based on the sensor fault severity layer-by-layer until the distribution of the same fault category in each subset is very close. Linear discriminant analysis is used to analyze historical fault data in each subgroup, and the kappa coefficient is applied as the basis for ending the training process. A BP neural network is employed to estimate the fault severity during the testing process, and the fault diagnosis sub-model is selected. The proposed method enables the accurate diagnosis of faults with different distributions in the same category and provides an accurate estimate of the sensor's fault severity degree. The estimated value of the sensor's fault degree can provide critical information for the maintenance of the equipment and can be used to correct the sensor's output.

Keywords: grid-connected inverter; open-circuit fault; current sensor fault; linear discriminant analysis; layering fault diagnosis

Citation: Jin, G.; Wang, T.; Amirat, Y.; Zhou, Z.; Xie, T. A Layering Linear Discriminant Analysis-Based Fault Diagnosis Method for Grid-Connected Inverter. *J. Mar. Sci. Eng.* **2022**, *10*, 939. https://doi.org/10.3390/jmse10070939

Academic Editor: Eugen Rusu

Received: 3 June 2022
Accepted: 5 July 2022
Published: 8 July 2022

1. Introduction

In recent years, ocean energy has attracted more attention due to its reproducibility, cleanness, and vast reserves [1,2]. Ocean energy includes wave energy, tidal energy, and ocean current energy [3,4]. Unfortunately, ocean energy conversion systems, as either immersed or floating systems, are affected by severe constraints due to their geographical location and weather conditions. These constraints lead to ocean energy systems potentially suffering from a high failure rate, mainly related to the mechanical parts, the electric generator, and even the power converters. In this context, the cascaded H-bridge inverter is favored for the electrical energy transmission of ocean energy fields due to its characteristics of easy modularization and convenient expansion of level numbers [5]. Due to the unstable output of ocean power generation, the power transmitted by the inverter fluctuates, aggravating the material's temperature cycle in the power semiconductor switches. Temperature cycle, mechanical vibration, and electromagnetic interference can damage the power semiconductor switch and current sensor [6,7]. The failure of power switches causes distortions in the inverter's output voltage and deteriorates the safety of other healthy components in the system [8]. The measurement errors caused by the current sensor fault propagate in the control loop through feedback [9]. These errors continue to accumulate and act on the driving module of the switch through the controller's signals, disrupting the normal operation of the power switch and then causing compound failure or even leading

to secondary accidents. Therefore, it is of great significance to study the open-circuit and current sensor faults in the grid-connected inverter to improve the reliability and stability of the whole system [10,11].

The analytical model-based method uses state variables to infer other variables that characterize the system's operation status [12]. This method compares the estimated and actual values to locate the fault [13]. The signal-based fault diagnosis method depends on the signal distortion caused by the fault [14]. The distinguishing fault features are extracted by signal processing methods such as Fourier transform [15–17], wavelet decomposition [18,19], and variational modal decomposition [20,21]. The extracted features are often used as input for an intelligent algorithm, and the intelligent algorithm finally identifies the fault types. Compared with model-based and signal processing-based fault diagnosis methods, data-driven fault diagnosis methods are more suitable for applications with complex system structures. Data-driven fault diagnosis methods can be divided into fault feature extraction and identification based on intelligent algorithms [22]. The feature extraction method extracts the features that represent the system's operating state from the monitoring signals. An intelligent algorithm locates the fault by analyzing the fault characteristics, such as BP network, extreme learning machine, support vector machine, etc. A BP neural network is a multilayer network with back-propagation errors that adjust weights and thresholds to learn input-to-output mappings. BP networks are theoretically capable of learning arbitrary nonlinear mapping relationships. Many different optimization algorithms have been proposed to improve the effectiveness of BP networks by optimizing the initial weights and thresholds [23,24].

Because of their effectiveness and easy implementation, scholars frequently make use of multivariate statistical analysis methods in feature extraction, such as principal component analysis, independent component analysis, and linear discriminant analysis [25]. These methods reduce the data dimension as much as possible on the premise of retaining sufficient fault information, thereby reducing the complexity of the model. Cai et al. [26] used principal component analysis (PCA) to extract the main feature components. The extracted data only reflect the projections that maximize the variance in the traditional PCA method. Wang et al. [27] used robust principal component analysis (RPCA) to extract fault features, adding weight factors to change the original PCA projection matrix so that the projected variables can better distinguish similar faults. The literature [28] shows that variable selection based on linear discriminant analysis can improve diagnostic accuracy while reducing the complexity of the model. With the continuous development of equipment, the complexity of sample data is also increasing, and the performance of traditional linear discriminant analysis methods has declined. The system's nonlinear, non-Gaussian, and multimodal properties are indexed as the main reasons for the performance degradation of these methods. The nonlinear characteristics of the system make the linear method unable to complete high-precision fault diagnosis. Reference [29] first performs a nonlinear transformation on the sample, then performs principal component analysis, and finally realizes the nonlinear principal component analysis of the original sample. This method is based on the assumption that the principal components are Gaussian distributed. In another study, a new monitoring statistic was constructed by introducing a local anomaly factor, which can find anomalous outliers in the given sample under any data distribution [30]. Regarding the non-Gaussian nature of samples, Reference [31] divides the non-Gaussian sample data into several Gaussian distribution subsets, and then each subset is analyzed. Tests on synthetic and natural datasets show that the method can significantly reduce the influence of non-Gaussian data distribution.

Affected by equipment operating states or different fault degrees, fault data belonging to the same fault category may exhibit multimodal characteristics [32]. One study from the literature [33] considers the global information represented by the between-class and within-class matrices. It uses the reverse nearest neighbor to mine the local details in the sample data, enabling the method to better describe faults containing multimodal sample

data. Reference [34] takes the neighborhood of the sample as the smallest subset through the reverse nearest neighbor, which can explain multimodal data in more detail.

In the event of a current sensor fault, multimodal data also appear. Under the influence of closed-loop control, the change in the sensor fault degree significantly impacts the fault signal. Moreover, the information on the fault degree of the sensor has important guiding significance for the overhaul and maintenance of the equipment.

Few methods can simultaneously diagnose open-circuit and current sensor faults. Gou et al. [35] consider these two types of fault in the three-phase inverter of an asynchronous motor, ignoring the fault severity degree variation. Variations in the fault severity degree can lead to significant differences in the dispersion of data for different faults, which reduces the diagnostic accuracy. This paper proposes a fault diagnosis method for grid-connected inverters based on layering linear discriminant analysis (LLDA) to solve these problems. The basic idea of this paper is to divide the historical fault data into several subsets according to the current sensor's fault degree in a layering manner in the offline modeling process and to use linear discriminant analysis to extract the features of these subsets. As the number of layers increases, the number of subsets will also increase, and the difference in the fault degree in each subset will be smaller. We use the kappa coefficient as the evaluation index to build multiple fault diagnosis sub-models gradually. A BP neural network estimates the fault severity degree of the sample, and the fault diagnosis sub-model to which it belongs is identified based on the estimation result.

The rest of this paper is organized as follows. Section 2 describes the effect of the multimodality of the sample data. Section 3 details the composition of the proposed method. Section 4 presents the experimental platform and results, and Section 5 gives the conclusions.

2. Problem Description

Figure 1 shows a typical three-phase, three-wire, voltage source grid-connected inverter. Each phase consists of two H-bridges, and each H-bridge is supplied by a DC source. Three current sensors are placed on the AC side to measure the currents for the closed-loop control. The controller adjusts the IGBT drive signal according to the feedback information of the current sensor and controls the output voltage of the inverter.

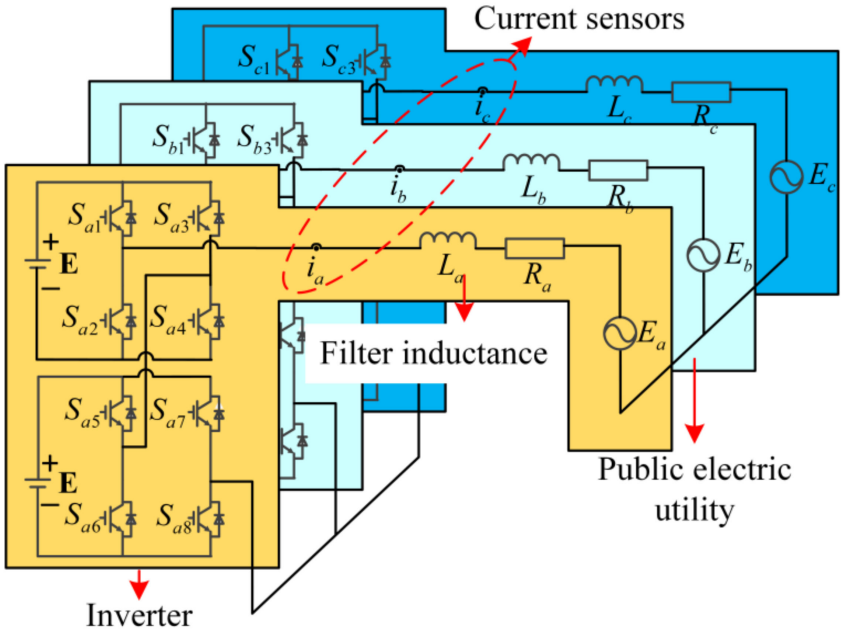

Figure 1. Block diagram of three-phase grid-connected inverter.

Sensor failure is manifested as a difference between the measured value and the actual operating state of the system. In a closed-loop system, measurement errors can cause the control signal to deviate from its expected value through a feedback loop. Under the effect of the abnormal control signal, abnormal signs can also appear in other healthy sensors, and this poses a challenge for fault diagnosis and isolation of sensor faults. There are two common fault types in the current sensor: gain-variation fault and zero-offset fault, which are defined in Equations (1) and (2), respectively.

$$I_{x\_\text{gain}} = (1 + \varepsilon_x) i_x \tag{1}$$

$$I_{x\_\text{offset}} = i_x + D_x \tag{2}$$

where I_{x_gain} is the output of the current sensor in gain-variation fault, I_{x_offset} is the output of the current sensor in zero-offset fault, i_x is the actual current in the circuit, ε_x is the variation of the gain, D_x is the sensor's offset, and x represents the phase. Under the normal case, ε_x and D_x are both zero, and the measured value of the sensor is a true reflection of the current in the circuit. While a zero-offset fault causes significant DC components in the three-phase currents, a gain-variation fault changes the amplitudes of the three-phase currents.

In the d-q coordinate system, two PI controllers are used to perform feedback current control with decoupling terms for producing voltage references, as shown in Figure 2. i_d^* and i_q^* are the reference currents on the d-q coordinates, respectively, ω is the angular frequency of the grid, and E_d and E_q are the grid voltages in the d-q coordinate system, respectively. In the sensor fault, the current fed back to the controller changes from i_d, i_q to I_d, I_q. A current sensor failure makes the three-phase current measurements no longer symmetrical, so I_d and I_q can have AC components. Under the adjustment of the PI controller, u_d and u_q also become AC components. u_d and u_q cause fluctuations in i_d and i_q, and the grid-connected inverter will simultaneously transmit fluctuating active and reactive power into the public electric utility, threatening the public electric utility.

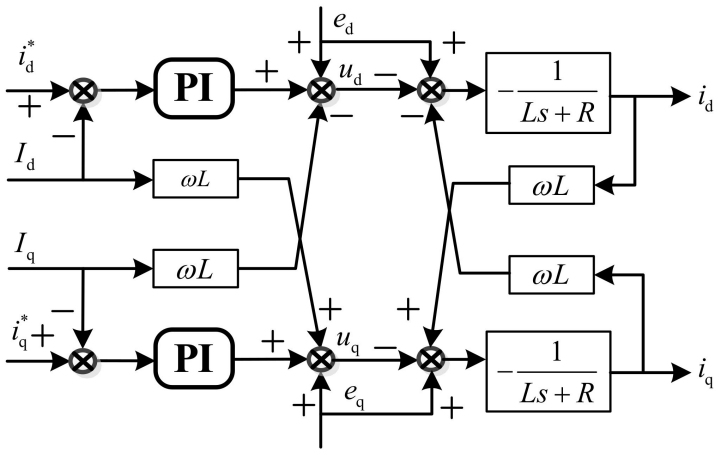

Figure 2. Schematic diagram of current decoupling control.

The severity degree of the sensor fault can be considered constant over a short period. However, the severity degree can change due to nonlinear effects such as the sensor usage time, device temperature drift in the power supply, and signal conditioning circuits. The variation of the fault degree will make the fault samples more dispersed. The average and standard deviation reflect the distribution of the samples [36]. The current power frequency and DC components are strongly influenced by gain-variation fault and zero-offset fault, respectively, so these are chosen to calculate the range and standard deviation separately. The samples were normalized based on z-score before calculating the average and standard deviation of the samples. Table 1 shows the range and standard deviation under different

fault severity degrees. The range and standard deviation of the samples are more significant when the fault severity varies over a broader range, implying a higher dispersion degree of the samples.

Table 1. Time-domain indicators under different fault severity degrees.

	$\varepsilon_a = 0$	$-0.1 \leq \varepsilon_a \leq 0.1$	$-0.15 \leq \varepsilon_a \leq 0.15$	$-0.2 \leq \varepsilon_a \leq 0.2$
Average	0.7373	0.7336	0.7290	0.7238
Standard Deviation	0.0046	0.0839	0.1171	0.1487
	$D_a = 0$	$-1 \leq D_a \leq 1$	$-1.5 \leq D_a \leq 1.5$	$-2 \leq D_a \leq 2$
Average	0.4988	0.5062	0.5076	0.5082
Standard Deviation	0.0016	0.0934	0.1311	0.1654

Figure 3 shows the performance of LDA under multimodal data. In this figure, the variance of the first type of faults is considerable in the second-dimensional feature. Due to the high dispersion of fault type 1, there is a significant error in using the mean value of each variable to inscribe the center of the category. In the traditional LDA method, the projection is in the direction perpendicular to the second-dimensional part to obtain a minor within-class distance. However, there is a partial overlap region between the second and third types of faults after projection. When the samples of the same fault are not independent and identically distributed, it is difficult to obtain the minimum within-class distance while guaranteeing the maximum between-class distance. Traditional LDA ignores the difference in dispersion caused by the variation of fault severity degrees. Because the main influences on fault data distribution are gain-variation and offset, this paper proposes dividing the current sensor fault data into layers according to the values of ε_x and D_x. Finally, the dispersion of feature data for different categories is relatively close within each subset.

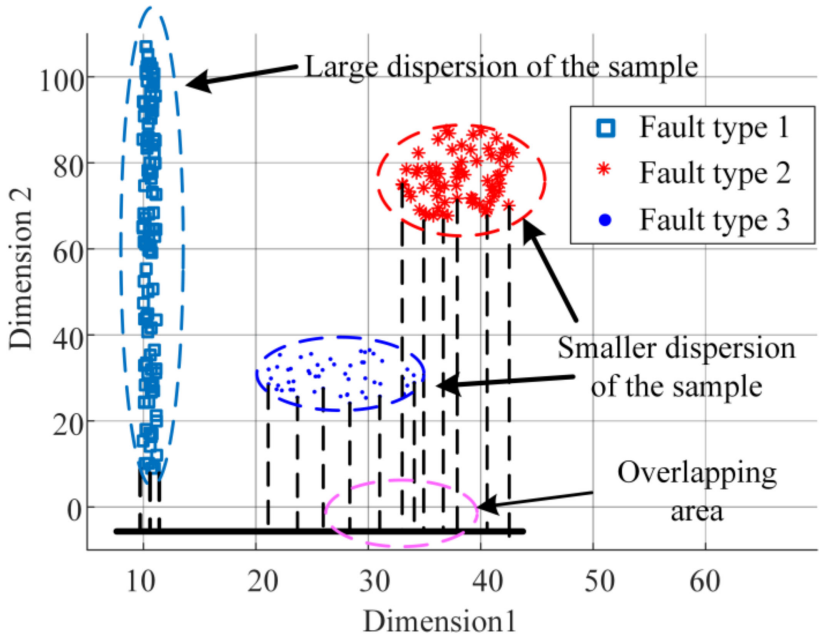

Figure 3. Classification of samples with different dispersion degrees.

3. The LLDA-Based Fault Diagnosis Method for Grid-Connected Inverter

The proposed fault diagnosis method includes the establishment and evaluation of layering linear discriminant analysis, followed by fault severity degree estimation based on a BP neural network.

3.1. Basic Principles of LDA

LDA is a classical dimension reduction method [36], which expects data in the same category to be as close as possible and can better reflect the differences between data. The combination of LDA and probability mode can realize fault classification. Multi-class LDA can be achieved by optimizing the following objectives:

$$\operatorname{argmax} J(\mathbf{W}) = \frac{\prod\limits_{diag} \mathbf{W}^{\mathrm{T}} \mathbf{S}_{\mathrm{b}} \mathbf{W}}{\prod\limits_{diag} \mathbf{W}^{\mathrm{T}} \mathbf{S}_{\mathrm{W}} \mathbf{W}} \tag{3}$$

where $\prod\limits_{diag} A$ is the product of the diagonal elements of A, $\mathbf{W}$ is a matrix of $n \times d$, $\mathbf{S}_{\mathrm{b}}$ is the between-class scatter matrix, and $\mathbf{S}_{\mathrm{W}}$ is the within-class scatter matrix. m_i is the number of samples for the i-th class fault, μ_i is the average vector of the i-th class, and μ is the mean vector of the total samples.

$$\mathbf{S}_{\mathrm{b}} = \sum_{i=1}^{N} m_i (\mu_i - \mu)(\mu_i - \mu)^{\mathrm{T}} \tag{4}$$

$$\mathbf{S}_{\mathrm{W}} = \sum_{i=1}^{N} \sum_{x \in X_i} (x - \mu_i)(x - \mu_i)^{\mathrm{T}} \tag{5}$$

After obtaining the projection data for each fault type, the maximum likelihood estimation can be used to calculate the mean and variance of each fault.

$$\ln p(X \Big| \mu, \sigma^2) = -\frac{1}{2\sigma^2} \sum_{i=1}^{K} (X_i - \mu)^2 - \frac{K}{2} \ln \sigma^2 - \frac{K}{2} \ln(2\pi) \tag{6}$$

where X_i is the i-th sample in each fault, and K represents the number of each fault.

When a new fault sample appears, each fault's distribution probability density function calculates the probability that the current fault sample belongs to that fault type. The class corresponding to the maximum probability is the prediction class. The probability density function is shown below.

$$\mathrm{f}(X) = \frac{1}{(2\pi)^{d/2} |\sum|^{1/2}} \exp\left[-\frac{1}{2} (X - \mu)^T \sum^{-1} (X - \mu) \right] \tag{7}$$

where d represents the dimensions of fault features, and $\sum$ is the covariance matrix of training samples for each class.

3.2. Establishment and Evaluation of Layering Diagnostic Model Based on Linear Discriminant Analysis

The fault diagnosis model is trained in layers based on LDA in this part. The proposed method builds a fault diagnosis model hierarchically and progressively. The fault data are firstly divided based on the current sensor's fault severity degree, and then the diagnostic model's kappa statistic is used as the layering termination criterion.

3.2.1. Division of Historical Fault Data

Both ε_x and D_x are considered to be zero in the case of healthy and open-circuit fault. The inverter output voltage possesses more significant fault characteristics under

open-circuit fault. At the same time, the current signal better reflects the current sensor fault; therefore, the inverter output voltage and current signals are selected to diagnose the open-circuit fault and sensor fault, respectively.

When building a layering fault diagnosis model based on linear discriminant analysis, historical fault data need to be divided. Figure 4 depicts the division process of historical fault data in a 4-layer structure. This uses ε_x and D_x as the division basis to build a layering fault diagnosis model. The sensor fault data are first separated from other faults in the dividing process. After each layer finishes dividing the data, a part of the data in each subset is extracted to train the LDA. Another part of the data is used as test data to verify the performance of the diagnostic model. In the proposed method, the division process of historical fault data is repeated until the diagnostic model meets the conditions. There will be lower dispersion of fault data in each fault subset with more layers, resulting in higher diagnostic accuracy. In the division of sensor fault data, ε_x and D_x are used to separate gain and offset historical fault data. The farther the values of ε_x and D_x deviate from zero, the more severe the failure. In decomposing from 2-layer to 3-layer, III.2 contains relatively severe gain-variation and zero-offset faults, and III.1 includes samples of relatively minor severity.

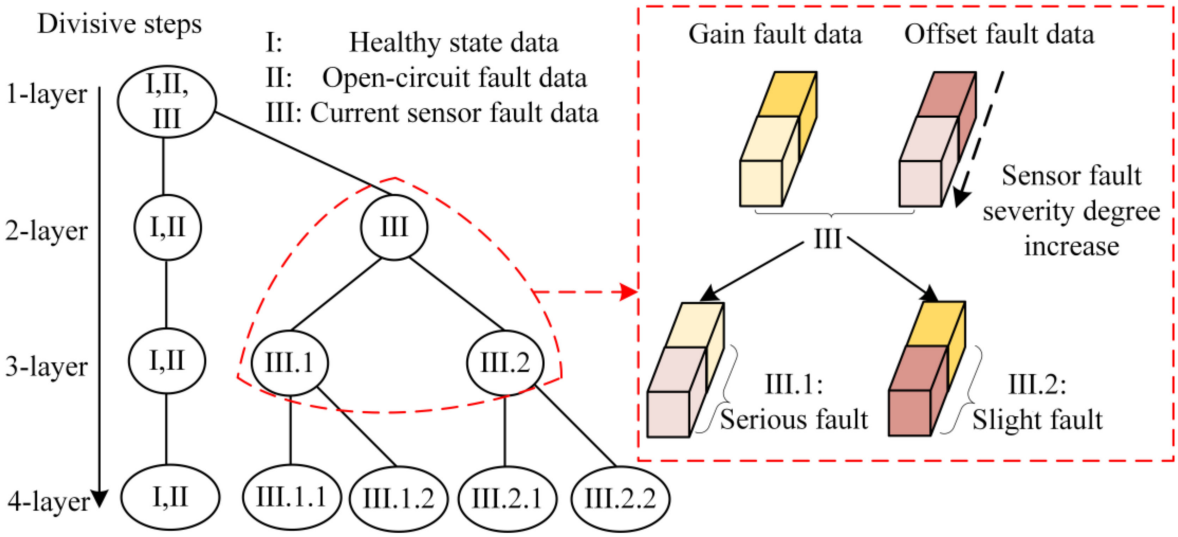

Figure 4. Schematic diagram of the historical fault data division.

3.2.2. Evaluation of Diagnostic Models Based on Kappa Statistic

In dividing the fault data and building a layering diagnosis model, the evaluation of the diagnosis model is crucial. For this purpose, we apply the kappa statistic to evaluate the model. The kappa coefficient was proposed by Cohen et al. and is a statistic that reflects the agreement between the measured and actual variables [37]. In fault classification, it can be used to determine whether the output results of the diagnostic model are consistent with the actual results. The higher k is, the higher the agreement between the two results is, with values in the range of 0–0.2 indicating low agreement and values in the range of 0.8–1 indicating almost complete agreement. The calculation of the kappa coefficient is shown below.

$$k = \frac{p_o - p_e}{1 - p_e} \tag{8}$$

where p_o is the overall diagnostic accuracy, and p_e can be calculated through Equation (9). N is the number of samples, m is the number of fault types, x_i denotes the total number of

samples whose predicted outcome is the i-th fault, and X_i is the total number of the i-th fault in the actual results.

$$p_e = \frac{\sum_{i=1}^{m} x_i X_i}{N \times N} \tag{9}$$

3.3. Fault Severity Degree Estimation Based on BP Neural Network

Since the division of fault data and the creation of the diagnostic model are based on the fault severity degree, the values of ε_x and D_x need to be estimated when a new fault signal appears. There is a nonlinear relationship between the magnitudes and average values of the three-phase currents with ε_x and D_x. Since the BP neural network method has a powerful nonlinear fitting ability, it is applied in this work to estimate the fault severity degree.

As shown in Figure 5, the neural network has three input neurons and one output neuron. When the command current of the system is stable, When the command current of the system is stable, under the action of closed-loop control, the average values of the three-phase currents are zero. In the zero-offset fault, DC components in the three-phase currents are unavoidable. So, in estimating the offset, the input to the neural network is the average values of the three-phase currents, and the output is the estimated value of the offset. $\overline{I}_x$ is calculated through Equation (10), and I_x is the output of the current sensor. N is the number of sampling points in a grid cycle. During the system's real-time operation, to obtain a faster diagnosis speed, the calculation of $\overline{I}_x$ is also continuously performed. To calculate $\overline{I}_x$ in a sliding window manner, the length of the window is the fundamental period of the grid, and the sliding step of the window is one control period.

$$\overline{I}_x = \frac{\sum_{i=1}^{N} I_x(i)}{N} \tag{10}$$

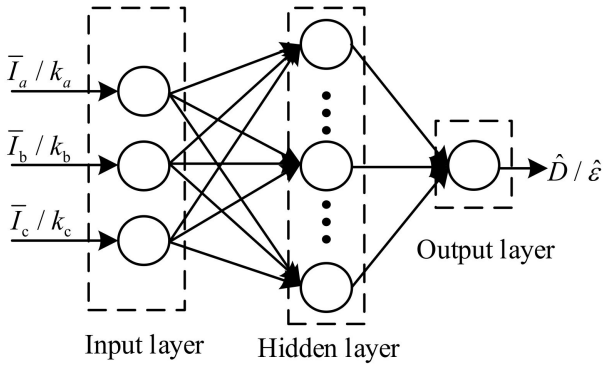

Figure 5. Estimation of ε_x, D_x based on BP neural network.

The gain-variation fault and the variation of command current can both change the amplitudes of the three-phase currents, and it is difficult to extract the amplitudes due to measurement noise. The integral of the absolute current value in (11) can also reflect the current's amplitude. Unlike load-variation, the gain-variation fault causes the amplitudes to be different. The three-phase currents will remain symmetrical under the closed-loop control when the load changes, so k_x is close to 1. However, under the gain fault, k_x is no longer close to 1 and has a close relationship with the degree of gain-variation. The variable k_x shown in Equation (11) is finally chosen to estimate the gain-variation to eliminate the effect of command current variation.

$$\begin{cases} L_x = \int_{(i-1)T_s}^{iT_s} |I_x| dt \\ k_a = \frac{L_b}{L_c}; k_b = \frac{L_c}{L_a}; k_c = \frac{L_a}{L_b} \end{cases} \tag{11}$$

3.4. Flow Chart of the LLDA-Based Fault Diagnosis Method

Figure 6 shows the flowchart of the proposed fault diagnosis method based on LLDA. The starting point of the proposed method is to divide the fault data into layers, so that the dispersion of the fault data in each layer is relatively low. When the degree of dispersion of different faults is relatively low, the error caused by using the mean value of each variable to represent the cluster center is also relatively reduced.

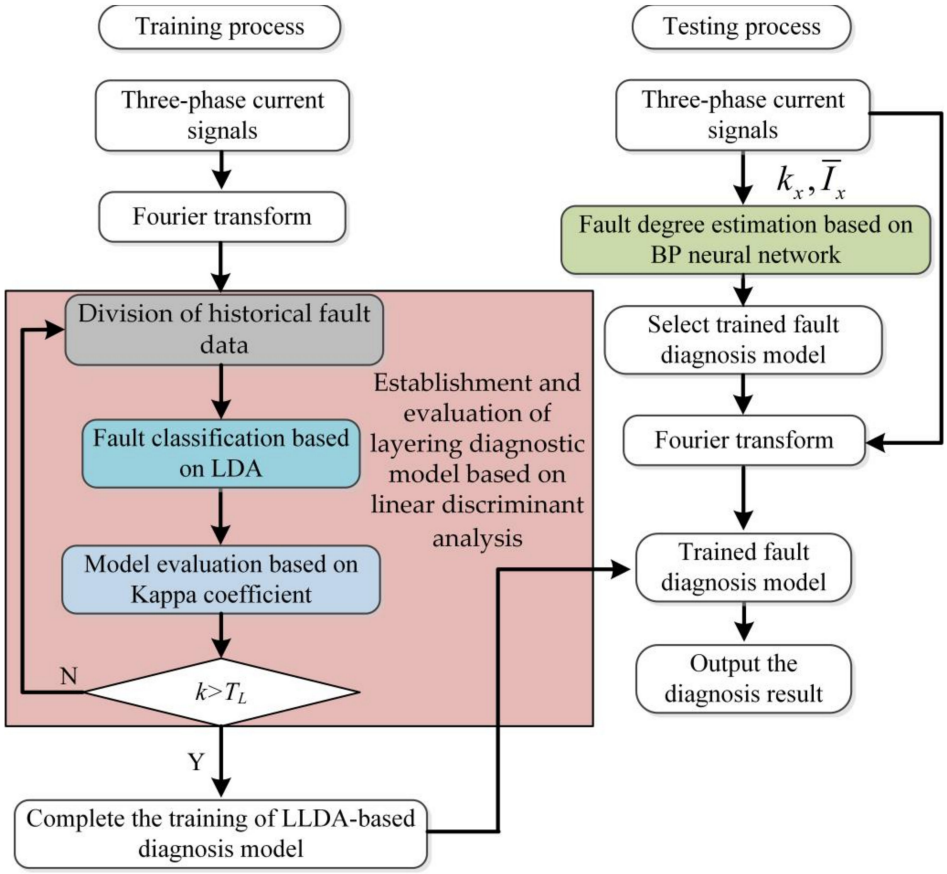

Figure 6. Flow chart of the LLDA-based fault diagnosis method for the grid-connected inverter.

The proposed method consists of a training process and a testing process. In the training process, the Fourier transform is used to obtain the fault characteristics implied in the three-phase currents. After the historical fault data have been divided into layers, a part of the data is used as training data to train the LDA. The building of the layering diagnosis model continues until the stopping condition is met. In the testing process, the BP neural network is used to estimate the fault severity degree of the sensor. Based on the estimation results, a diagnostic sub-model is selected, and the trained model can output the final diagnostic result.

After locating the fault and estimating the fault severity, the estimated value can be used to correct the output of the faulty sensor. Gain-variation and zero-offset faults are processed using (12) and (13) to eliminate the effects of ε_x and D_x, respectively. Fault-tolerant control can be realized by replacing the output signal of the faulty sensor with $\hat{I}_{x\_\text{gain}}$ and $\hat{I}_{x\_\text{offset}}$.

$$\hat{I}_{x\_\text{gain}} = \frac{I_{x\_\text{gain}}}{(1+\hat{\varepsilon}_x)} \tag{12}$$

$$\hat{I}_{x\_\text{offset}} = I_{x\_\text{offset}} - \hat{D}_x \tag{13}$$

where $\hat{I}_{x\_\text{gain}}$ and $\hat{I}_{x\_\text{offset}}$ are the estimation values of the actual current under gain-variation and zero-offset faults, respectively.

4. Experimental Validation

4.1. Data Acquisition

A three-phase grid-connected experimental platform was built to validate the proposed method. The experimental platform is composed of several main parts, as shown in Figure 7. Table 2 lists the parameters of the experimental system. The open-circuit fault is simulated by lowering the driving signal. This experimental platform can be explained as follows: The open-circuit fault setting unit connects the controller and the drive circuit of the switch, which utilizes the cooperation of the NOR gate logic circuit and the self-locking button. When the self-locking button is pressed, 5 V voltage is input to the NOR gate logic circuit with the drive signal output by the controller, and the control signal transmitted to the switch is set to a low level. The switch remains open under the continuous action of the low-level control signal, and this state is consistent with the open-circuit fault. In the measurement of voltage and current signals, the signal conditioning circuit adjusts the sensor output signal to a range suitable for sampling by the controller. Since it is impossible to make the current sensor fail at a specific moment, the sensor fault is simulated by changing the measured value in the controller. The results sampled by the ADC need to be scaled up to reflect the actual status precisely. In this paper, the gain-variation fault of the current sensor is simulated by changing this scaling, and the zero-offset fault is simulated by superimposing a constant term onto the normal scaling result. For example, the zero-offset fault is simulated by superimposing a constant onto the measured signal. Figure 7 illustrated the experimental test bench. The controller used in the experiment is dSPACE's DS1102.

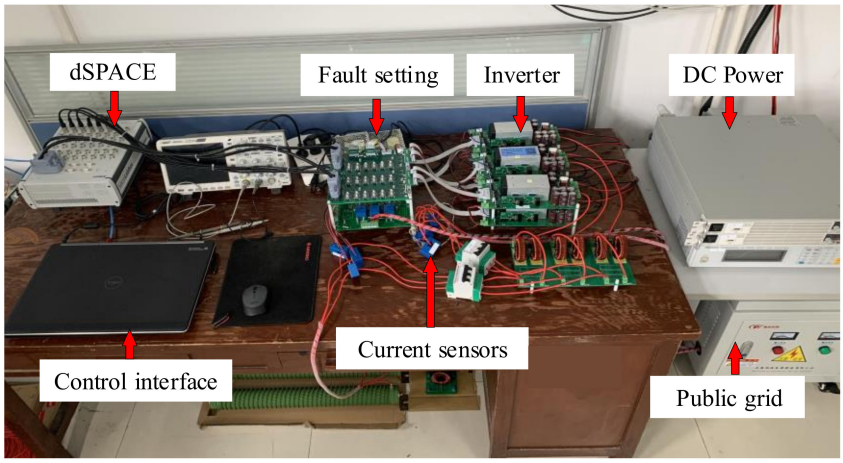

Figure 7. Three-phase grid-connected inverter experimental platform.

Table 2. System parameters.

Parameter	Value
DC-link voltage	100 V
Filter inductance	6 mH
Switching frequency	5 kHz
Grid voltage (RMS)	220 V
Grid frequency	50 Hz
d-axis reference current	4 A
q-axis reference current	0 A
Isolation transformers	1:2

The types of faults studied in this paper belong to three categories: open-circuit fault of IGBT, gain-variation fault, and zero-offset fault of the current sensor. Among them, the open-circuit fault of the IGBT consists of eight types according to the fault location. There

are six common fault types in the current sensor: gain-variation and zero-offset fault in A, B, and C-phase, respectively. The specific fault classification is shown in Table 3.

Table 3. Labeling of the faults.

Fault Type	Fault Location	Label
Healthy state	-	1
Gain-variation fault of current sensor	A-phase current sensor	2
	B-phase current sensor	3
	C-phase current sensor	4
Zero-offset fault of current sensor	A-phase current sensor	5
	B-phase current sensor	6
	C-phase current sensor	7
Open-circuit fault	S_{a1}	8
	S_{a2}	9
	S_{a3}	10
	S_{a4}	11
	S_{a5}	12
	S_{a6}	13
	S_{a7}	14
	S_{a8}	15

Fault data are collected under various operating conditions to examine the performance and robustness of the proposed method. The reference current on the d-axis is 4 A, the switching frequency is 5 kHz, and the sampling frequency is 20 kHz. The gain-variation varies from -0.3 to 0.3 in an interval of 0.05, and the zero-offset varies from -3 A to 3A in an interval of 0.5 A.

4.2. Layering Performance under Variation of Fault Severity Degree

Figure 8 shows the distribution of the studied faults under the DC and power frequency components in a 1-layer structure. The change in the severity degree of the sensor fault significantly affects the data distribution. It can be seen from the figure that as the gain-variation varies from -0.3 to 0.3, the power frequency component of the gain-variation fault changes from 4.2 to 5.3, while the difference in the DC component is not apparent. Due to the high dispersion of sensor fault samples, there is a significant error in using the mean value of each variable to inscribe the center of the category. Traditional LDA finds the most extensive interclass matrix based on inaccurate category centers, increasing the difficulty of finding a projection matrix that distinguishes all fault samples.

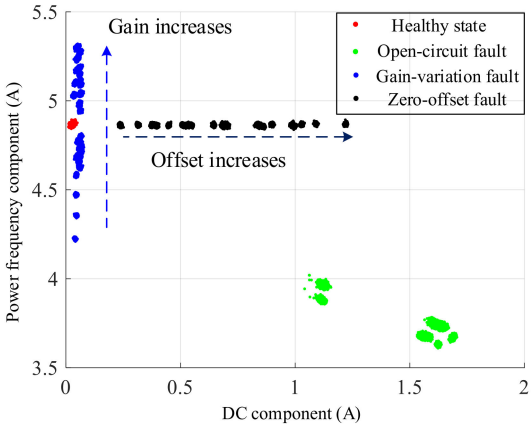

Figure 8. The distribution of faults under the DC and power frequency components.

The t-distributed stochastic neighbor embedding (t-SNE) method is a nonlinear dimensionality reduction technique suitable for visualization by embedding high-dimensional data into a two-dimensional format. Figure 9 shows the two-dimensional feature distribution produced using t-SNE, which does not have any real physical significance. The numbers in the legend represent the fault labels in Table 3. In the 1-layer structure, LDA is used for feature extraction of all types and severity degrees. The visualization result in Figure 9a shows overlapping between faults 2–7, which are six types of current sensor faults. Meanwhile, there are clear boundaries between open-circuit faults. Figure 9b shows that the characteristics of various faults have a high degree of aggregation, and the distance between different fault categories is noticeable. Figure 9c,d shows the visualization results of the current sensor fault features in 2-layer and 3-layer, respectively. The overlap between the sensor fault samples decreases as the number of layers increases. This phenomenon shows that layering effectively improves the separability of current sensor faults. The reduction in overlap area is mainly due to the decreasing dispersion of the sensor fault samples within each subset as the number of layers increases. The lower the dispersion is, the lower the error is in using the mean of each variable as the category center, and the LDA can find a projection matrix that distinguishes all faults.

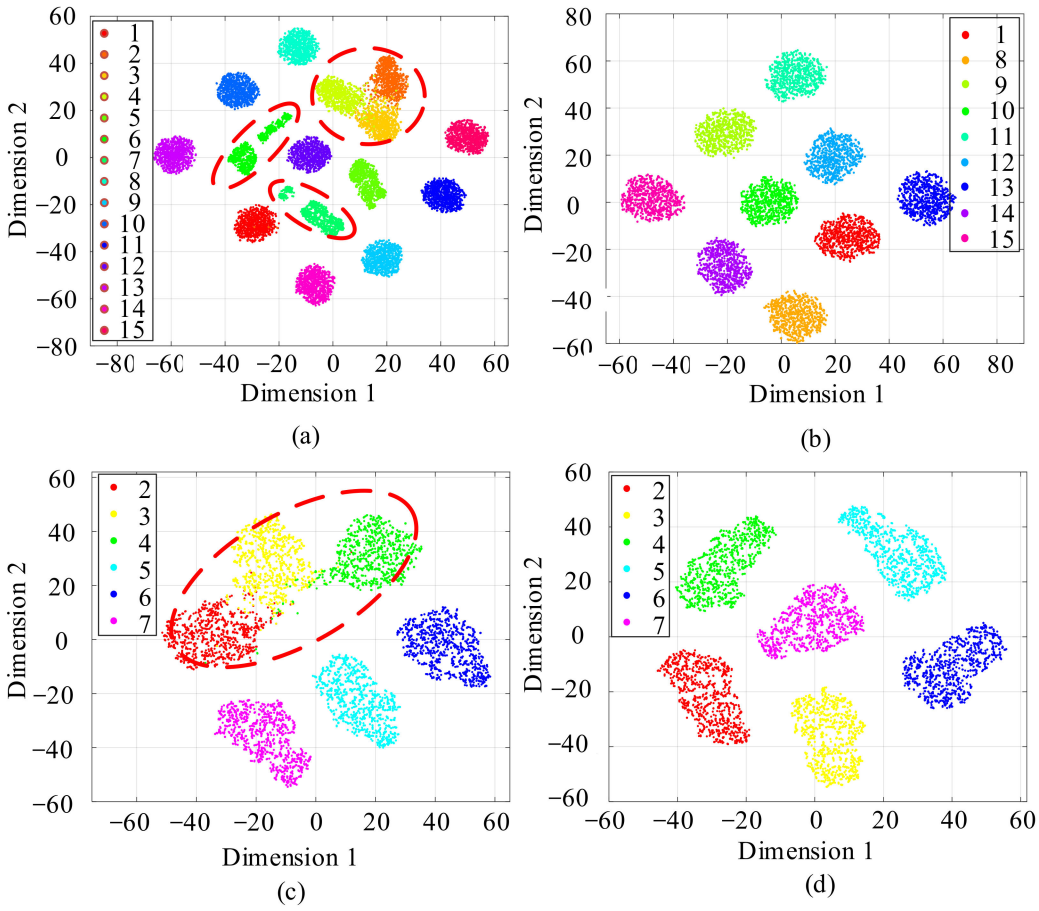

Figure 9. Visualization results of the features extracted under different layers based on t-SNE: (**a**) result of all faults under 1-layer; (**b**) result of open-circuit fault and healthy state under 2-layer; (**c**) result of current sensor fault under 2-layer; (**d**) result of current sensor fault under 3-layer.

In the layering establishment of the fault diagnosis model, the kappa coefficient is used as an indication for iteration termination. Table 4 shows the variation of the kappa coefficient with the increase in layers, based on the experimental data. In 1-layer, the kappa coefficient is only 0.8393, indicating that the model has a poor diagnostic performance on the test set. When the layer number becomes 2, the kappa coefficient rapidly increases to

0.9815. When the number of layers is increased to 3, the kappa coefficient grows further, but the increase is not significant. For application scenarios that require incredibly high diagnostic accuracy, further increasing the number of layers could make sense. In this application scenario, when the number of layers is two, there is a favorable diagnostic performance.

Table 4. Kappa coefficient under different layers.

Layers	Kappa
1	0.8393
2	0.9815
3	0.9907

4.3. Fault Severity Degree Estimation Based on BP Neural Network

The estimation of ε_x and D_x is performed using a BP neural network, with the number of neurons in the input, hidden, and output layers being 3, 10, and 1, respectively. The activation function in the hidden layer is the sigmoid function. Figure 10 shows the estimated results of the gain-variation and offset of the A-phase current sensor. The results demonstrate that when the fault severity degree changes, the network can control the estimation error of the gain-variation in the range [−0.005–0.005], and the estimation error of the offset does not exceed 5×10^{-4}, indicating that the neural network has an excellent estimating performance.

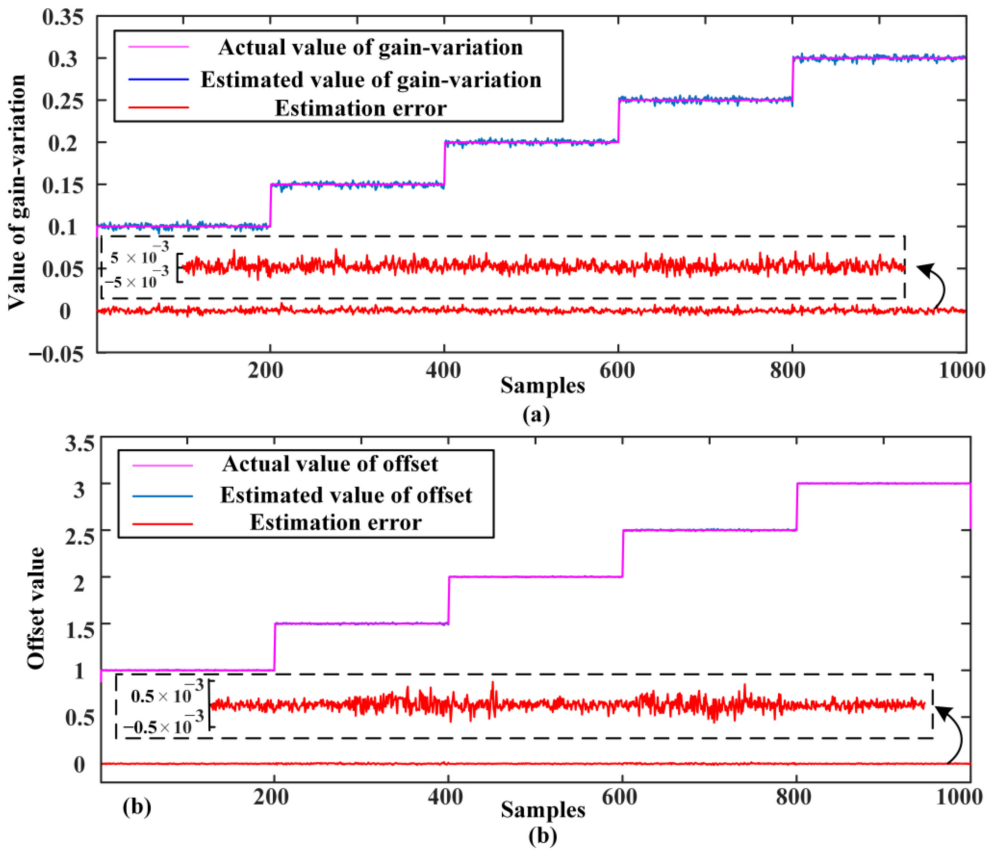

Figure 10. Estimation results of sensor faults: (**a**) estimation results of gain-variation fault; (**b**) estimation results of zero-offset fault.

Figure 10 shows the estimated error distributions for the sensor gain-variation and zero-offset faults. The blue bar represents the number of samples that fall within the

current estimation error. The red line represents the cumulative sample size as a proportion of the total sample, with higher slopes indicating a relatively high concentration of estimation error. The cumulative probabilities in Figure 11 suggest that the estimation error of gain-variation concentrates on $\left[-5\times10^{-3}, 5\times10^{-3}\right]$, and the estimation error of zero-offset concentrates on $\left[-5\times10^{-4}, 5\times10^{-4}\right]$. The estimation errors for offset are mainly concentrated around zero compared with the estimation errors for gain-variation.

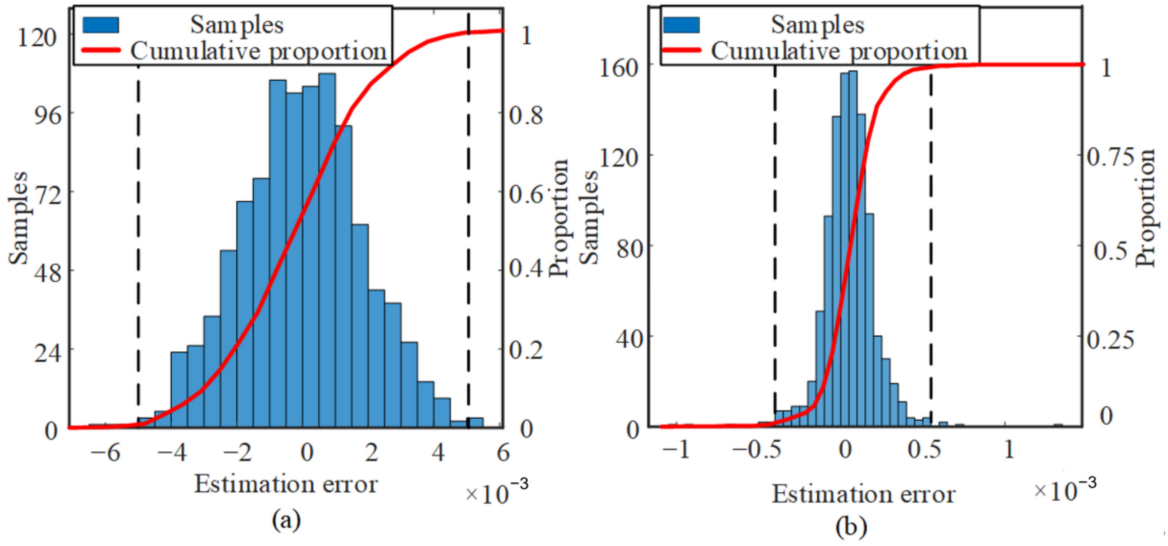

Figure 11. Error distribution of sensor fault severity degree estimation: (**a**) error distribution of gain-variation fault; (**b**) error distribution of zero-offset fault.

The proposed method is compared with the diagnostic methods in [26,35], and Table 5 shows their diagnostic accuracies. The first method in the table uses PCA for feature extraction, replacing more original variables with fewer new variables while retaining as much information about the original variables as possible. To obtain more information, PCA needs the variance of the projected data to be as large as possible in the projection direction. Compared with the open-circuit fault, the fault characteristics of the current sensor fault are not significant. Therefore, the features of sensor failures are submerged in the feature extraction process, leading to this method's low diagnostic accuracy. ReliefF is a feature selection algorithm that assigns different weights to features based on their ability to distinguish close-range samples. The variation of the current sensor's fault severity degree increases the distance between the same fault category. Therefore, the ReliefF algorithm reduces the weight of the feature, failing to select the optimal feature subset.

Table 5. Accuracy of different diagnosis methods.

Method	Accuracy (%)
FFT-PCA-BN [26]	90.40
FFT-ReliefF-RVFL [35]	88.87
FFT-LLDA	99.19

5. Conclusions

Variations in the current sensor's fault degree cause the fault sample data to exhibit significant multimodal characteristics. Therefore, this paper uses the current sensor's fault degree to divide the historical fault data layer-by-layer. In this way, the multimodal data are divided into multiple equally distributed subsets, enabling the linear discriminant analysis to perform well in these subsets. The final experimental results also show that LLDA can

effectively reduce the influence of multimodality on fault diagnosis. The estimated value of failure degree is important information for equipment operation and maintenance.

In this work, a BP neural network is applied to select the average value and the ratio of three-phase currents' integral value to estimate the fault severity degree. The network achieves good estimation results under different fault degrees, with the maximum estimation error being less than 0.005. The estimated value of the sensor fault severity degree can also be used to correct the faulty sensor's output signal and to achieve fault-tolerant control. This information will also be useful for guiding subsequent equipment maintenance.

It should be noted that in the proposed method, the number of layers is a critical variable. Theoretically, as the number of layers increases, the effectiveness of fault diagnosis will improve, and the kappa coefficient in the experimental results also verifies this. The selection of the number of layers in this paper is based on the kappa coefficient. The selection of layers needs to be further improved in future research.

Author Contributions: Conceptualization, G.J., Y.A., T.X. and T.W.; methodology, G.J. and Y.A.; software, G.J., Y.A. and T.W.; validation, G.J., Y.A., Z.Z. and T.W.; formal analysis, G.J. and Y.A.; investigation, G.J., Y.A., Z.Z. and T.W.; resources, G.J., Y.A., Z.Z. and T.W.; data curation, G.J. and T.W.; writing—original draft preparation, G.J., T.W., Y.A., Z.Z. and T.X.; writing—review and editing, G.J., T.W., Y.A., Z.Z. and T.X.; visualization, G.J., T.W., Y.A., Z.Z. and T.X; supervision, T.W. and Y.A.; project administration, T.W. and Y.A.; funding acquisition, T.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant no.s 62073213).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, L.; Li, W.; Li, J.; Fu, Q.; Wang, T. Evolution trend research of global ocean power generation based on a 45-year scientometric analysis. *J. Mar. Sci. Eng.* **2021**, *9*, 218. [CrossRef]
2. Draycott, S.; Sellar, B.; Davey, T.; Noble, D.R.; Venugopal, V.; Ingram, D.M. Capture and simulation of the ocean environment for offshore renewable energy. *Renew. Sustain. Energy Rev.* **2019**, *104*, 15–29. [CrossRef]
3. Li, M.; Luo, H.; Zhou, S.; Kumar, G.M.S.; Guo, X.; Law, T.C.; Cao, S. State-of-the-art review of the flexibility and feasibility of emerging offshore and coastal ocean energy technologies in East and Southeast Asia. *Renew. Sustain. Energy Rev.* **2022**, *162*, 112404. [CrossRef]
4. Clemente, D.; Rosa-Santos, P.; Taveira-Pinto, F. On the potential synergies and applications of wave energy converters: A review. *Renew. Sustain. Energy Rev.* **2021**, *135*, 110162. [CrossRef]
5. Ekstroem, R.; Leijon, M. Grid Connection of Wave Power Farm Using an N-Level Cascaded H-Bridge Multilevel Inverter. *J. Electr. Comput. Eng.* **2013**, *2013*, 26.1–26.9. [CrossRef]
6. Ji, B.; Song, X.; Sciberras, E.; Cao, W.; Hu, Y.; Pickert, V. Multiobjective design optimization of IGBT power modules considering power cycling and thermal cycling. *IEEE Trans. Power Electron.* **2014**, *30*, 2493–2504. [CrossRef]
7. Huang, H.; Mawby, P.A. A lifetime estimation technique for voltage source inverters. *IEEE Trans. Power Electron.* **2012**, *28*, 4113–4119. [CrossRef]
8. Choi, U.M.; Blaabjerg, F.; Lee, K.B. Study and handling methods of power IGBT module failures in power electronic converter systems. *IEEE Trans. Power Electron.* **2014**, *30*, 2517–2533. [CrossRef]
9. Li, D.; Wang, Y.; Wang, J.; Wang, C.; Duan, Y. Recent advances in sensor fault diagnosis: A review. *Sens. Actuators A Phys.* **2020**, *309*, 111990. [CrossRef]
10. Gao, Z.; Cecati, C.; Ding, S.X. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [CrossRef]
11. Chen, W.; Zhang, L.; Pattipati, K.; Bazzi, A.M.; Joshi, S.; Dede, E.M. Data-driven approach for fault prognosis of SiC MOSFETs. *IEEE Trans. Power Electron.* **2019**, *35*, 4048–4062. [CrossRef]
12. Gong, Z.; Huang, D.; Ma, L.; Qin, N.; Jadoon, H.U.K. Grid voltage sensor fault-tolerant control for single-phase two-level PWM rectifier. *IET Electr. Power Appl.* **2021**, *16*, 776–788. [CrossRef]

13. Poon, J.; Jain, P.; Spanos, C.; Panda, S.K.; Sanders, S.R. Fault prognosis for power electronics systems using adaptive parameter identification. *IEEE Trans. Ind. Appl.* **2017**, *53*, 2862–2870. [CrossRef]
14. Zhao, H.; Cheng, L. Open-switch fault-diagnostic method for back-to-back converters of a doubly fed wind power generation system. *IEEE Trans. Power Electron.* **2017**, *33*, 3452–3461. [CrossRef]
15. Shabbir, M.N.S.K.; Liang, X.; Chakrabarti, S. An ANOVA-based fault diagnosis approach for variable frequency drive-fed induction motors. *IEEE Trans. Energy Convers.* **2020**, *36*, 500–512. [CrossRef]
16. Reyes-Malanche, J.A.; Villalobos-Pina, F.J.; Cabal-Yepez, E.; Alvarez-Salas, R.; Rodriguez-Donate, C. Open-circuit fault diagnosis in power inverters through currents analysis in time domain. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3517512. [CrossRef]
17. Wu, F.; Sun, J.; Zhou, D.; Liu, Y.; Geng, T.; Zhao, J. Simplified Fourier series based transistor open-circuit fault location method in voltage-source inverter fed induction motor. *IEEE Access* **2020**, *8*, 83953–83964. [CrossRef]
18. Veerendra, A.S.; Mohamed, M.R.; Punya Sekhar, C. A novel fault-detection methodology of proposed reduced switch MLI fed induction motor drive using discrete wavelet transforms. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12820. [CrossRef]
19. Wang, Z.; Huang, Z.; Song, C.; Zhang, H. Multiscale adaptive fault diagnosis based on signal symmetry reconstitution preprocessing for microgrid inverter under changing load condition. *IEEE Trans. Smart Grid* **2016**, *9*, 797–806. [CrossRef]
20. Abdelkader, R.; Cherif, B.D.E.; Bendiabdellah, A.; Kaddour, A. An Open-Circuit Faults Diagnosis Approach for Three-Phase Inverters Based on an Improved Variational Mode Decomposition, Correlation Coefficients, and Statistical Indicators. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3510109. [CrossRef]
21. Zhang, J.; Sun, H.; Sun, Z.; Dong, Y.; Dong, W. Open-circuit fault diagnosis of wind power converter using variational mode decomposition, trend feature analysis and deep belief network. *Appl. Sci.* **2020**, *10*, 2146. [CrossRef]
22. Cen, J.; Yang, Z.; Liu, X.; Xiong, J.; Chen, H. A Review of Data-Driven Machinery Fault Diagnosis Using Machine Learning Algorithms. *J. Vib. Eng. Technol.* **2022**, 1–27. [CrossRef]
23. Liu, Y.; Zhu, X.; Yang, J. Fault diagnosis of PV array based on optimised BP neural network by improved adaptive genetic algorithm. *J. Eng.* **2017**, *2017*, 1427–1431. [CrossRef]
24. Zhang, L.; Gao, T.; Cai, G.; Hai, K.L. Research on electric vehicle charging safety warning model based on back propagation neural network optimized by improved gray wolf algorithm. *J. Energy Storage* **2022**, *49*, 104092. [CrossRef]
25. Xu, X.; Yan, X.; Yang, K.; Zhao, J.; Sheng, C.; Yuan, C. Review of condition monitoring and fault diagnosis for marine power systems. *Transp. Saf. Environ.* **2021**, *3*, 85–102. [CrossRef]
26. Cai, B.; Zhao, Y.; Liu, H.; Xie, M. A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems. *IEEE Trans. Power Electron.* **2016**, *32*, 5590–5600. [CrossRef]
27. Wang, T.; Qi, J.; Xu, H.; Wang, Y.; Liu, L.; Gao, D. Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter. *ISA Trans.* **2016**, *60*, 156–163. [CrossRef]
28. Chiang, L.H.; Kotanchek, M.E.; Kordon, A.K. Fault diagnosis based on Fisher discriminant analysis and support vector machines. *Comput. Chem. Eng.* **2004**, *28*, 1389–1401. [CrossRef]
29. Schölkopf, B.; Smola, A.; Müller, K.R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **1998**, *10*, 1299–1319. [CrossRef]
30. Deng, X.; Wang, L. Modified kernel principal component analysis using double-weighted local outlier factor and its application to nonlinear process monitoring. *ISA Trans.* **2018**, *72*, 218–228. [CrossRef]
31. Nie, F.; Zhao, X.; Wang, R.; Li, X. Fast Locality Discriminant Analysis with Adaptive Manifold Embedding. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**. [CrossRef] [PubMed]
32. Yu, J. Process monitoring through manifold regularization-based GMM with global/local information. *J. Process Control* **2016**, *45*, 84–99. [CrossRef]
33. Guo, Y.R.; Bai, Y.Q.; Li, C.N.; Shao, Y.H.; Ye, Y.F.; Jiang, C.Z. Reverse nearest neighbors Bhattacharyya bound linear discriminant analysis for multimodal classification. *Eng. Appl. Artif. Intell.* **2021**, *97*, 104033. [CrossRef]
34. Zhu, F.; Gao, J.; Yang, J.; Ye, N. Neighborhood linear discriminant analysis. *Pattern Recognit.* **2022**, *123*, 108422. [CrossRef]
35. Gou, B.; Xu, Y.; Xia, Y.; Deng, Q.; Ge, X. An online data-driven method for simultaneous diagnosis of IGBT and current sensor fault of three-phase PWM inverter in induction motor drives. *IEEE Trans. Power Electron.* **2020**, *35*, 13281–13294. [CrossRef]
36. Rao, C.R. The utilization of multiple measurements in problems of biological classification. *J. R. Stat. Soc.* **1948**, *10*, 159–203. [CrossRef]
37. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Med.* **2012**, *22*, 276–282. [CrossRef]