

Article

Exploring the Impact of Parallel Architecture on Improving Adaptable Neuro-Fuzzy Inference Systems for Gas-Insulated Switch Defect Recognition

I-Hua Chung ^{1,*}  and Yu-Hsun Lin ²

¹ Master's Program in Offshore Wind Energy Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 807, Taiwan

² Department of Electrical Engineering, National Penghu University of Science and Technology, Magong 880, Taiwan; yuhsunlin@gms.npu.edu.tw

* Correspondence: steven.chung@nkust.edu.tw

Abstract: Gas-insulated switchgear malfunctions during power system operation may occur due to electrical, thermal, or human errors in the manufacturing process. The leading causes of insulation deterioration of gas-insulated switchgear are discharging along the surface caused by dirt on the insulating material, internal discharge caused by impurities and cavities in the insulating material, corona discharge caused by poor assembly or construction at the site, and electric tree channel discharge caused by the intense internal discharge. Since different defects produce different partial discharge characteristics, the operating power equipment can be analyzed using measurement instruments to detect partial discharge for preventive equipment fault diagnosis, avoiding unnecessary power outages and losses; therefore, evaluating the defects in gas-insulated switchgear is essential. In this study, three gas-insulated switchgears were prefabricated with different defects before encapsulation, and the partial discharge data of each defect were measured by applying different test voltages. The adaptive neuro-fuzzy inference system (ANFIS) input data were used to evaluate the recognition effect, showing that the average recognition rate of the core for all defects was over 90%. The proposed system architecture can continuously accumulate the defect measurement database of gas-insulated switchgear and be used as a reference for constructing electrical equipment defect recognition systems.

Keywords: partial discharge; gas-insulated switches; adaptive class neuro-fuzzy inference systems



Citation: Chung, I.-H.; Lin, Y.-H. Exploring the Impact of Parallel Architecture on Improving Adaptable Neuro-Fuzzy Inference Systems for Gas-Insulated Switch Defect Recognition. *Energies* **2022**, *15*, 3940. <https://doi.org/10.3390/en15113940>

Academic Editor: Xiao-Hua Wang

Received: 25 April 2022

Accepted: 23 May 2022

Published: 26 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Increasing the transmission voltage to transmit power increases transmission efficiency and reduces transmission losses but also subjects the insulation materials of gas-insulated switches (GIS) to higher electrical stresses. The leading causes of insulation degradation are manufacturing, mechanical effects, external environment, and the possibility of heat build-up and partial discharge under high electrical stress, resulting in oxidation, decomposition, and corrosion of the material and, thereby, arcing and leakage of electrical equipment [1]. Overall, although GIS is reliable with controlled uniformity of insulation materials and impurity components, inevitably, there may be air gaps in the solid insulation material, and metal particle defects may occur inside the equipment due to frequent switching operations. In addition, the partial discharge phenomenon generated under the action of high voltage leads to accelerated deterioration of insulation materials, which also affects the effectiveness of equipment operation [2].

Statistics from the International Council on Large Electric Systems (CIGRE), which investigates the causes of GIS failures, show that the failure rate increases year by year as GIS is installed in the power system after 25 years of operation.

Insulation damage accounts for more than half of the GIS failure factors, and the rest include mechanical failure, gas leakage, and others [3].

Although GIS has a high degree of reliability in manufacturing and installation, inevitably, due to human negligence and GIS internal manufacturing defects, such as metal particles in the box and insulation materials in the air gap, these minor defects in the GIS operation process may produce an internal discharge phenomenon. Although the energy released by partial discharge is minimal, it can trigger chemical and physical reactions in the insulation. The insulation material gradually deteriorates, eventually leading to equipment insulation breakdown, indirectly destroying the relevant equipment and endangering the operators on site. Therefore, long-term online monitoring of partial discharges in power equipment plays an essential role in maintaining high-voltage equipment [4,5].

Since the service life of equipment often depends on the strength of insulation materials and insulation structure design. With the end-user requirements for equipment reliability, domestic and foreign equipment or instrumentation manufacturers and academic researchers are gradually paying attention to partial discharge analysis of power equipment and other related research topics [6]. The main research direction is to observe the partial discharge phenomena generated during the degradation of gas-insulated switch insulation caused by defects [7]. For example, Qi et al. used the step-up method to conduct long-term tests on partial discharges caused by high-voltage electrode faults on the insulator surface and obtained various statistical profiles at different test stages by observing the initial discharge to the development of flashover along the surface as a basis for diagnosing and assessing the severity of partial discharges caused by high-voltage electrode faults on the insulator surface of GIS equipment [8].

The types of insulation defects commonly found in GIS are as follows:

- (a) Fixed metal protrusion on high-voltage conductors (conductor);
- (b) Fixed metal protrusion inside the enclosure;
- (c) Floating potential formed by the gap between the insulating support and the conductor;
- (d) Free metal particles inside the enclosure;
- (e) Dirt on the surface of the insulation support.

As each type of defective partial discharge has unique phase characteristics, the characteristics of the partial discharge pattern shown are different, and the degree of damage caused by GIS may also vary. The following are further descriptions of the characteristics of the partial discharge signals generated by the above common types of defects in GIS research [9–11]:

- (a) Fixed metal protrusions on high-voltage conductors

The discharge generated on the high-voltage conductor is mainly caused by small metal chips attached to the conductor, as in the case of corona discharge generated by sharp metal protrusions on the shell. As the protrusions form the tip part of the electric field, concentration intensity is high. As the voltage changes to the peak, the nearby gas is ionized, and its positive ions leave a tip under the action of the electric field. As the SF₆ gas molecules have strong electronegativity to adsorb electrons to form electron clusters, the electric field around the tip increases, forming a “corona discharge”. There are several discharge characteristics, including: (1) when the initial discharge occurs, only in the negative half-cycle near the discharge signal; (2) the discharge signal appears in the positive half-cycle as the test voltage increases; (3) The discharge pulse in the partial discharge inception voltage (PDIV) of the positive and negative half-cycle is more asymmetric and, as voltage increases, the positive half-cycle discharge signal amplitude is stable, and the negative half-cycle discharge volume and number of discharges are more significant than the positive half-cycle [12]. In addition, Gao’s study on partial discharge detection using ultra-high frequency (UHF) reported that the discharge volume differs depending on the protrusions’ cross-sectional diameter, resulting in different discharge volumes and distribution situations. The maximum discharge volume increased from 10 mV to 60 mV as

the diameter length increased, with a significant spread in the distribution of the discharge volume amplitude [13].

(b) Fixed metal protrusions inside the box

The discharge phenomenon of fixed metal protrusions located inside the box is similar to that of fixed metal protrusions on high-voltage conductors, mainly concentrated at the peak of the power supply. The discharge characteristics are as follows: (1) the partial discharge is mainly concentrated at the positive peak of the supply voltage and produces a very high discharge amplitude. Unlike the discharge formed by the fixed metal protrusions on the high-voltage conductor, the discharge signal at the negative peak has a lower repetition rate and smaller amplitude; (2) the discharge near the peak of the negative half-period does not have a significant increase in amplitude with increasing voltage. The overall discharge volume increases according to the longer cross-sectional diameter of the protrusions. When the cross-sectional diameter of the protrusions is 0.4 mm, 0.8 mm, and 1.3 mm, the distribution range of the discharge volume widens with the increase in cross-sectional diameter, except for a significant increase in the highest amplitude of the discharge volume [13]. In addition, the results from the observation of the shape of the discharge volume distribution show that each forms a particular distribution profile at different cross-sectional diameters.

(c) The floating potential formed by the gap between the insulator and the conductor

From the GIS, high-voltage conductor cracks or insulation support (insulator) and the conductor electrode gap (electrode gap) form a floating potential, which is different from the gas ionization generated by the tip of the discharge phenomenon; floating potential defects caused by the discharge have the characteristics of local insulation breakdown. When the voltage at both ends of the equivalent capacitor reaches its withstand voltage, the insulation gap between the insulation support and the conductor breaks down, generating a partial discharge. When the voltage at both ends of the equivalent capacitor reaches its withstand voltage, the discharge stops. Due to the continuous external voltage application, the equivalent capacitor is discharged and recharged, and the partial discharge phenomenon is repeatedly triggered.

The discharge characteristics are summarized as follows: (1) the discharge phenomenon is similar to a discharge capacitor, which generates a high discharge amplitude; (2) the partial discharge area is mainly concentrated at the leading edge of quadrant I and quadrant III [12]; (3) as the applied voltage increases, several discharge pulses occur sporadically in the I and III quadrants of the voltage cycle near the peak voltage. In a study by Gao et al. to detect partial discharges in UHF, the air gap diameter lengths in the control insulation material were 0.2 mm, 0.5 mm, and 1 mm, respectively, and the discharge volume and amplitude distribution were observed by applying the test voltage until the discharge stabilized [13]. The maximum discharge decreases from more than 100 mV to less than 60 mV as the diameter increases. When comparing the air gap diameters of 1 mm and 0.5 mm, there is a tendency for the discharge amplitude distribution to spread slightly, whereas the discharge amplitude spreads significantly for air gap diameters of 0.2 mm and 1 mm.

(d) Free particles inside the box

In the GIS box, there is a partial discharge when moving particles in the high electric field collide with the box, therefore the discharge pulse appears randomly without any phase correlation, and the contour of the discharge signal after accumulating several cycles is similar to the test voltage waveform [12,14].

(e) Insulation supports surface dirt

When the free metal particles are immobilized on the surface of the insulator, this causes the local electric field to be concentrated, resulting in a significant reduction in the flashover voltage on the insulator surface. Compared to the discharge amplitude generated by fixed particles and floating electrodes, the highest amplitude occurs at the positive and

negative half of the voltage near the peak, and the discharge phase is mainly concentrated near zero [15].

In addition, this defect's discharge volume and distribution in SF₆ and air, respectively, and the discharge volume in SF₆ was more significant than that in air as measured by a bandpass partial discharge detection instrument [13].

Therefore, if different detection frequency instruments measure the partial discharge signal, the discharge volume may have a different distribution trend.

Furthermore, the forms of partial discharge include surface discharge caused by surface dirt, internal discharge caused by impurities and cavities in the insulation material, corona discharge caused by the tip of poor assembly or construction on-site, and electric tree channel discharge caused by intense internal discharge [16]. Therefore, if online monitoring technology can issue early warning signals before faults, this could avoid substantial economic losses caused by equipment failure outages. After determining possible defects, the test data can be used as a reference for on-site maintenance or essential equipment part preparation, which can help improve the reliability of the power equipment [17].

2. Research Methodology and Measurement Planning

2.1. Research Methodology

Insulation material deterioration is prone to equipment destruction due to flashover caused by the high-voltage surge. Once the switchgear is out of service due to partial discharge, it not only interrupts regular production but also may require human and material resources for emergency support [18–20], as well as being extremely costly [21]. Current international research on partial discharge recognition is generally based on the condition of a single defective source. For example, Satish and Zaengl proposed to establish a database using PD mapping of known defect sources [22] applying fractal (fractal) theory to the mapping analysis [23]. Sahoo et al. also proposed multi-source recognition using two-dimensional wavelet transformed vertical, horizontal, and diagonal coefficients of variation [24], statistical, class neural network path, fuzzy logic, and time series methods in partiality discharge mapping analysis [22–25]. Montanari et al. used the time and frequency center of gravity of each signal and plotted the pulsed signals in different regions of the T–W plane to find signals with different characteristics such as noise or different types of partial discharge signals, showing that signals with different characteristics, such as noise or different types of partial discharge signals, form clusters for defect type recognition [26–28]. Tang et al. used the concentration ratio of SF₆ decomposed into several gases during partial discharge as a characteristic parameter and the decision tree theory to identify types of partial discharge defects [29]. Since SF₆ GIS has been tested before leaving the factory, it is not possible to perform power failure inspection in-field operation, so online judgment of switchgear status has become a pivotal issue to prevent accidents, and long time detection of discharge behavior and defect type analysis is an essential procedure [30,31].

Thus, domestic industry experts and academic researchers plan a series of studies on maintenance methods and detection and analysis techniques for power equipment to avoid unnecessary industrial safety accidents and equipment losses caused by insulation deterioration. In this paper, we propose a process for the construction of a database of equipment defects, as shown in Figure 1.

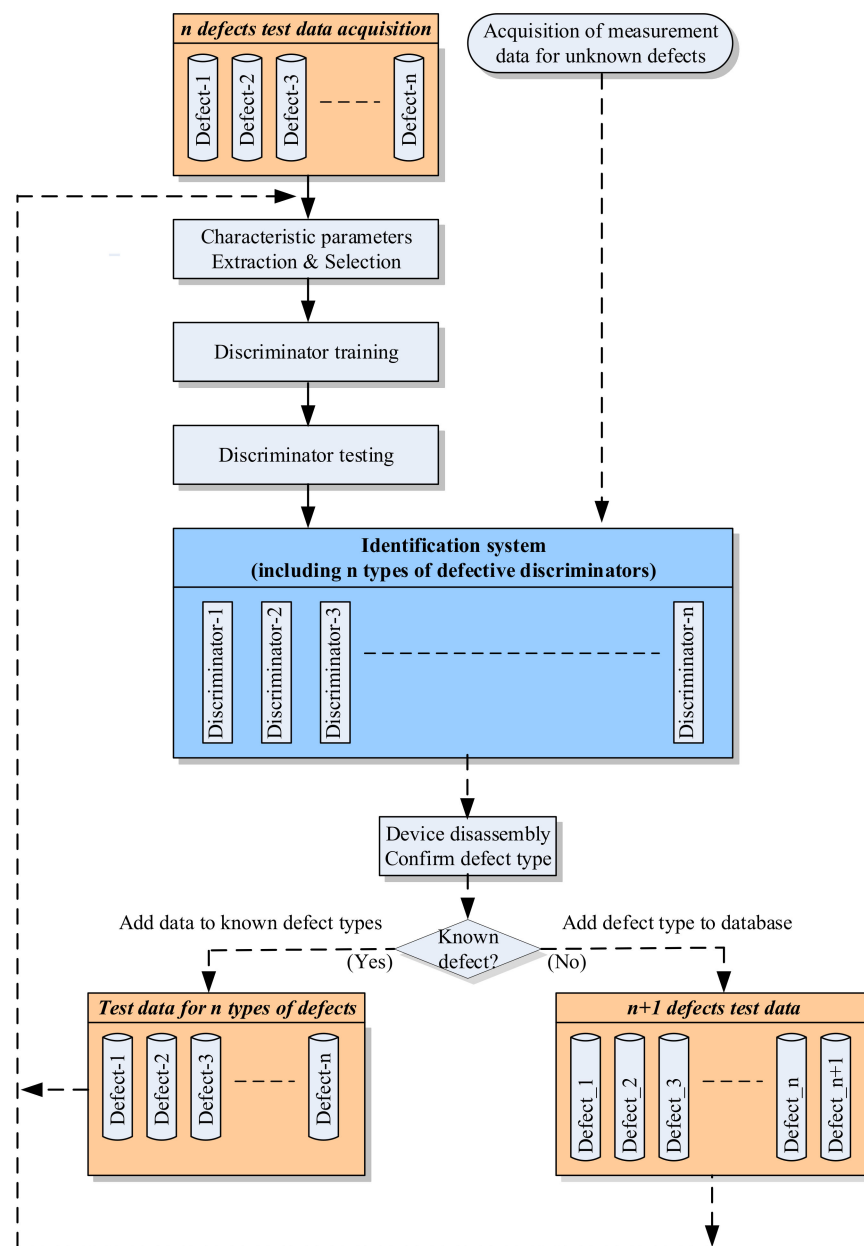


Figure 1. Schematic diagram of the construction of a defect measurement database.

Monitoring the partial discharge phenomenon of SF6 insulation equipment for an extended period can also be used to develop a database of various types of defects for follow-up analysis by domestic researchers and scholars. It is expected to improve the national equipment diagnosis technology and accurately identify the types of defects during insulation failure, which can be used as a reference for routine maintenance.

Three sets of simple 15 kV gas-insulated load start/stop switches manufactured by a domestic electrical company were used in the test. Generally, if a GIS has internal insulation defects, it is easy to cause equipment failure and power outage. Common internal defects of GIS include metal protrusions on the conductors or inside the enclosure, free-moving metal particles, floating potential due to poor contact, surface stains, and defects in the solid insulation medium. Therefore, as shown in Figure 2, a single defect was initially created inside each switch case during the factory assembly process using the test object:

Defect 1: Oil stain added to the inner layer of the porcelain sleeve on the power side.

Defect 2: Metal particles sprinkled on the bottom of the box.

Defect 3: Metal projection welded to the bottom of the connecting shaft inside the box (to create a tip discharge effect).

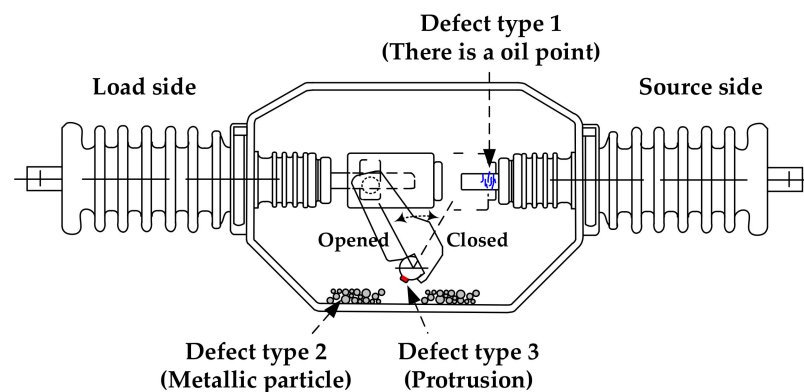


Figure 2. Diagram of the internal defects of the switch.

As discussed, general GIS internal insulation defects very easily cause equipment failure and cause power outages, therefore, different defects were created during the assembly of the GIS before analysis in the high-voltage test laboratory to identify and diagnose the defects. The experimental plan involved three sections: planning the test procedure, converting the characteristic parameters, and exploring the identifier's training and test data structure.

1. Planning the test flow

Figure 3 shows the planning concept of the accelerated insulation degradation test, starting from the measurement of voltage V_{st} , planning the equal interval of the test time ΔT , and then increasing the test voltage ΔV of equal magnitude to the end measurement voltage V_{end} each time to explore the recognition ability of the discriminator under all levels of test voltage. It is considered a complete test cycle when all defective test items are tested for accelerated insulation deterioration. The test pieces are transferred to each unit for GIS-related research and then transported back to the high-pressure laboratory for the second pressurization test. As a result of the repeated cycle of the above process, the following test measurement data can be used to verify the previous research results.

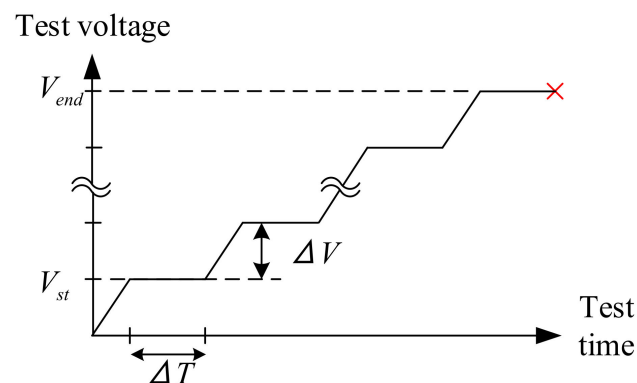
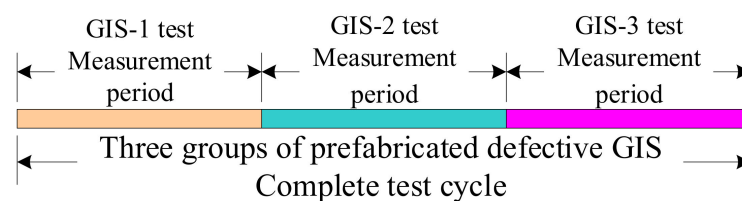


Figure 3. Experimental planning concept.

2. Conversion of features

After pre-processing the original localized measurement data, a series of seventeen features are extracted from the data using several statistical methods, as shown in Figure 4. Some of the feature parameters with good recognition rates were selected to improve the performance of the recognition system.

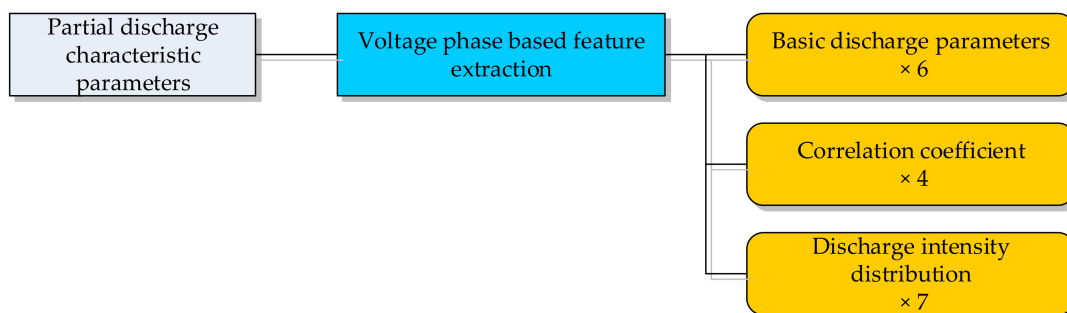


Figure 4. Features used in this study.

3. To explore the training and testing data structure of the discriminator

Since partial discharge phenomena containing many forms and different defect types produce different partial discharge characteristics, partial discharge and its analysis can be performed to assess the defects present in the GIS [24]. Statistical methods are used to extract several characteristic parameters from the partial discharge measurements as input data for training and testing of the discriminator or directly from phase mapping for image recognition after analysis [32]. In the past 15–20 years, many partial discharge recognition methods have been proposed and tested, including statistical tools, signal processing tools, image processing techniques, fuzzy logic, and artificial neural networks.

In addition to the properties and advantages of general neural networks, the adaptive neuro-fuzzy inference system (ANFIS) uses the fuzzy inference mechanism, which makes the computation of processing units easier and faster and increases the system's fault tolerance. Therefore, ANFIS was used to identify the partial discharge data of each defect, and seventeen partial discharge characteristic parameters were used to identify the defect inside the specimen. The ANFIS analysis is similar to that of a general supervised neural network, in which the object of study is first identified, and the input data and output categories are given. The input data must have information that can express the differences between different categories, so the features extracted from the data are often used as the input data. The ANFIS training process adjusts the internal parameters to find the relationship between input features and output categories.

Initializing the fuzzy inference system (FIS) model of ANFIS is more complex, involving the selection of inputs [33] and the selection of FIS models, the design of if-then rules, and membership functions within the FIS model. After the initial FIS model is designed, training data can be input to train the ANFIS before initial validation is performed by inputting test data. Finally, by analyzing the recognition rate of each defect by each feature, several modules of features can be selected to represent the characteristics of a particular defect, as shown in Figure 5, and the set of feature modules of each defect forms a parallel recognition system, as shown in Figure 6.

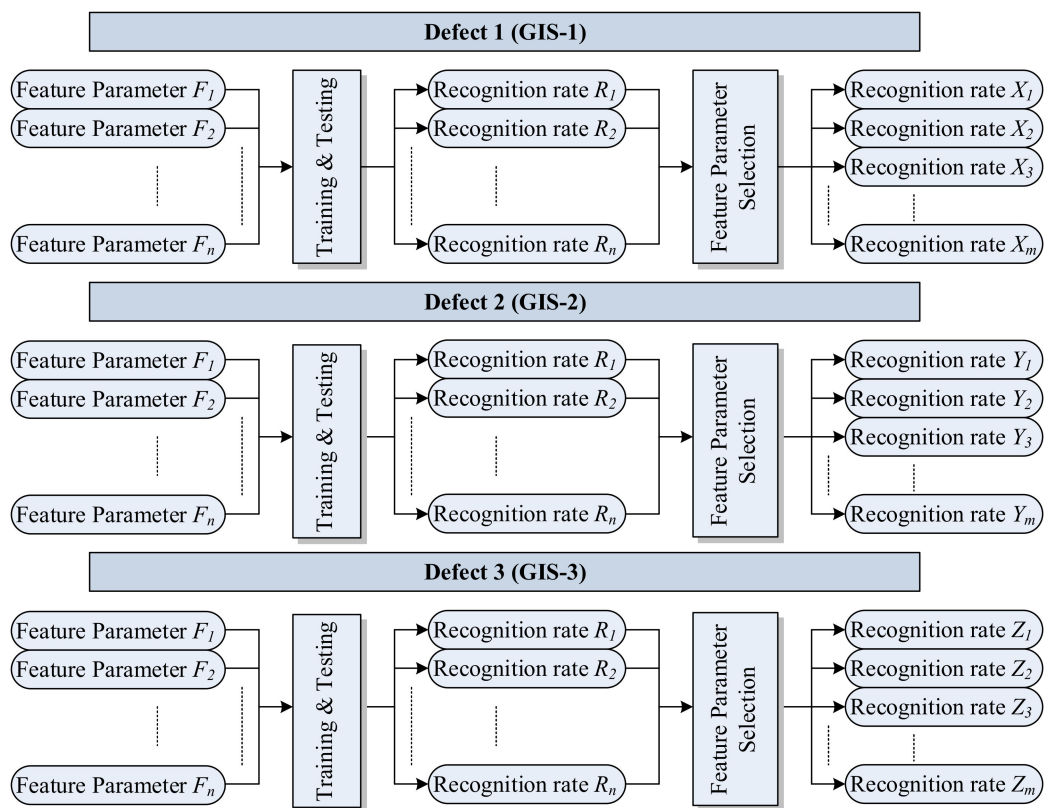


Figure 5. Selection of features for specific defects.

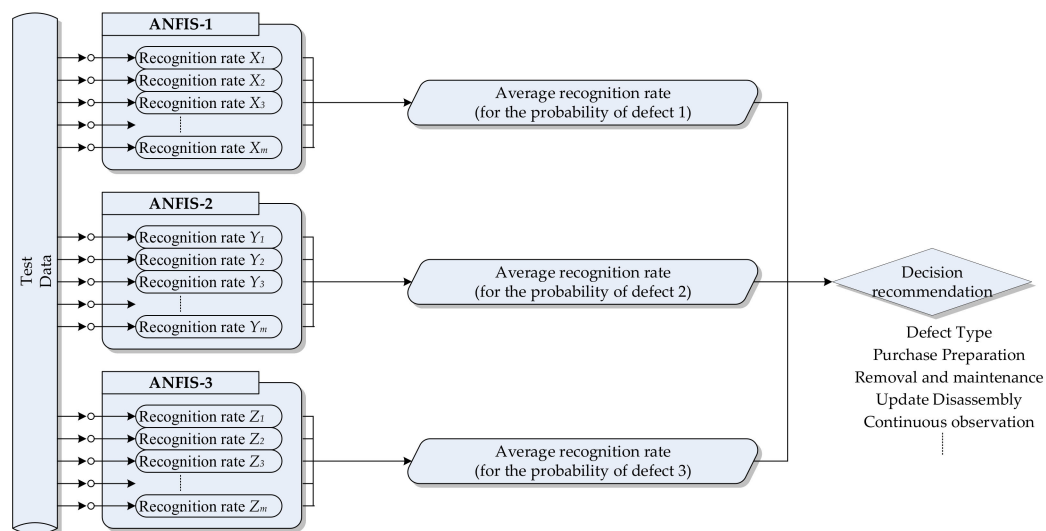


Figure 6. Schematic diagram of the parallel recognition system.

The feature extraction program extracts the test data input into the recognition system in Figure 6 and inputs it to ANFIS according to the required features of each defect identifier for recognition. The average recognition rate of each feature analysis is output from each defect identifier, which is used as a reference basis for analysts' decisions. In addition, over time, an appropriate weight can be given to each defect identifier in the FIS according to the experts' suggestions, strengthening the recognition system's capability.

2.2. Measurement Planning

The high-voltage laboratory consists of pressurized equipment and partial discharge measuring equipment (measuring device, MD) [34]. The pressurized equipment is adjusted by the inductive voltage regulator on the control panel, and the maximum output test voltage is 200 kV after stepping up through the primary and secondary step-up transformers [35]. The partial discharge phenomenon occurs when the electrical insulation of equipment components is weak. When the test voltage exceeds the starting voltage of partial discharge of the defective equipment, the MD senses the partial discharge pulse signal flowing through the object to be tested and the coupling capacitor. The partial discharge signal generated by the insulation deterioration is measured through the MD, then transmitted to the personal computer for storage via the circuit protection device.

In addition to using partial discharge measurement instruments to measure the current at the moment of discharge, the test can be conducted by a capacitive divider after voltage division, and the currently applied test voltage can be connected to the output of the self-designed attenuation circuit to observe the discharge pulse and phase-related information. Finally, a personal computer equipped with an A/D conversion adapter is used to record the voltage signal and discharge pulse signal. The developed program is then used to process the original partial discharge signal data in real-time to display the relevant graphs of the discharge signal.

The insulation deterioration test was accelerated by increasing the test voltage at equal intervals with equal magnitude each time, as shown in Figure 3, where V_{st} is the starting test voltage, V_{end} is the ending test voltage, ΔT is the duration interval of the voltage increase, and ΔV is the test voltage for each increment. The partial discharge data measurement records a partial discharge signal of 40 electrical cycles every 4 min.

In the study, the test voltage was increased from the initial discharge voltage to 50 kV. From the starting test voltage V_{st} to 40 kV, the voltage was increased by 2 kV every 8 h for the voltage increase and measurement. Based on the experience of GIS insulation damage, the voltage was increased by 1 kV every 8 h after 40 kV for voltage increase and measurement. After 40 kV, the test voltage was increased by 1 kV every 8 h for voltage boosting and measurement. The partial discharge test was completed when the test voltage reached V_{end} .

Assuming the starting test voltage is 10 kV, the recording device was set to capture one partial discharge data every 4 min during pressurization, and it is expected that each voltage level can be measured $(8 \text{ h} \times 60 \text{ min}) \times (1 \text{ stroke} 4 \text{ min}) = 120$ data. The total number of voltage levels is expected to be $26 \times 120 = 3120$.

In addition, the partial discharges are significantly different at the same test voltage waveform. Subsequently, the raw measurement data were simplified before extracting the characteristic parameters using statistical methods as the input data for the analysis of the discriminator.

3. Measurement Data Processing and Analysis

3.1. Data Simplification Method

During the test, the test voltage and discharge pulse signal were obtained by an A/D capture card. The starting test voltage was assumed to be 10 kV, and the test voltage was continuously increased to the end voltage of 50 kV, with 26 sets of voltage levels. The sampling period (cycle) was set to 42 cycles, and the sampling rate was 20 MHz, so the number of data points that could be obtained during the sampling cycle was:

$$\frac{20 \text{ MHz}}{60 \text{ Hz}} \times 42 \text{ cycles} = 14,000,000 \text{ data points}$$

The measurement data were stored in an array format of $1 \times 14,000,000$. After the first and last cycles (which may contain incomplete data) are eliminated, the raw data are taken from 40 cycles, then subjected to a simplification and filtering procedure. To observe the partial discharge data captured by the measurement device, the current signal

of the raw data was split into j phase windows based on the test voltage signal period. After filtering by the wavelet method [36,37], the maximum discharge of the PD current signal was stored. In this study, 60 Hz was used as the reference frequency, and each cycle was divided into 600 phase windows. The data were stored after being reduced to a 40×600 array using a written Matlab program for subsequent analysis of the discharge profiles, called $q - t - \varphi$ plots or phase evolution plots. Overall, the number of discharge cycles increases with more measurements, so the characteristics of the discharge frequency can be further explored, which is called $n - q - \varphi$ diagram or phase analysis diagram. Since the discharge quantity q and phase φ are functions of the discharge number n , they are commonly written as $H_n(\varphi, q)$.

Since this experiment was conducted without the discharge conversion parameter test, the partial discharge quantities are shown by the measured pulse amplitude v . The phase evolution and resolution patterns of the three defects at 26 kV are shown in Figures 7 and 8, respectively.

The above-streamlined data process can significantly improve the data storage problem, but there is still a shortage of information on the application of 3-D mapping data to analyze partial discharge. Since the partial discharge phenomenon is random, the following section describes the method of converting data patterns using statistical methods.

3.2. 2-D Distribution Mapping of Partial Discharges

Although the original partial discharge data are reduced to phase-resolved partial discharge (PRPD) mapping by the method above, there is a significant improvement in data storage. However, to facilitate statistical analysis, this section converts the 3-D map into a 2-D distribution map by using the concept of projection onto a plane.

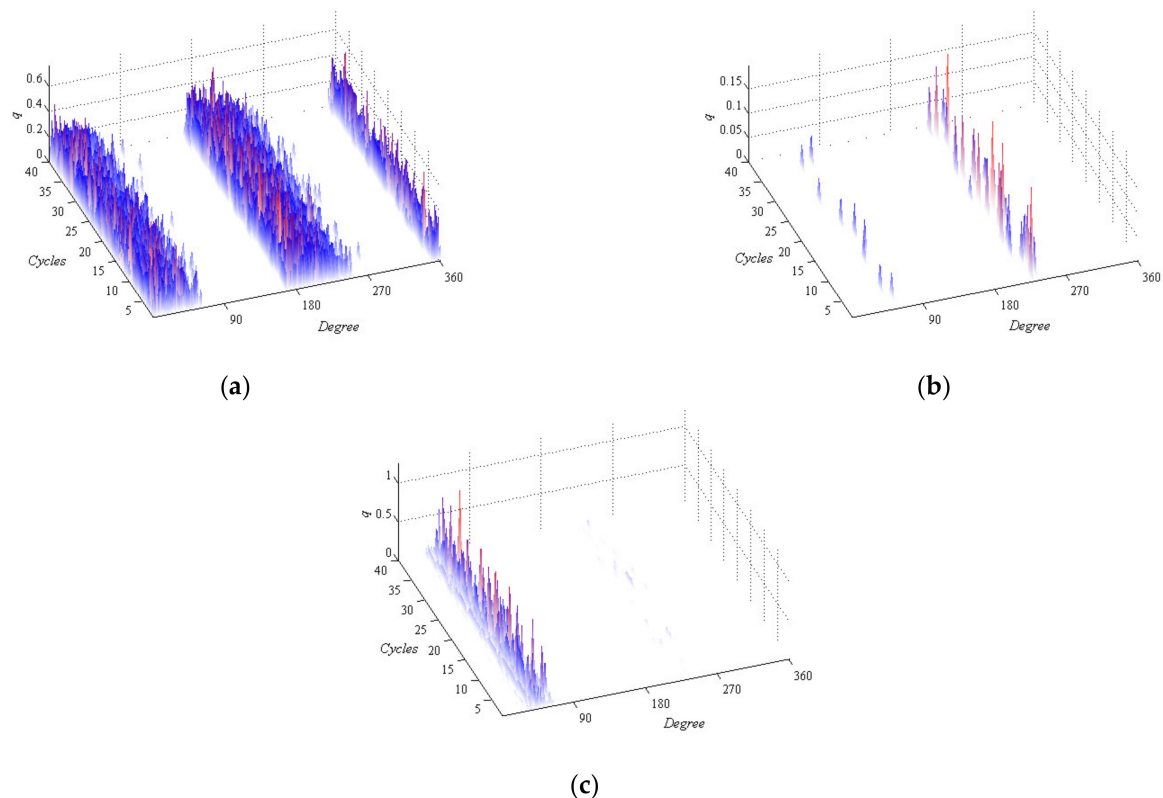


Figure 7. 26 kV $q - t - \varphi$ patterns: (a) GIS-1; (b) GIS-2; (c) GIS-3.

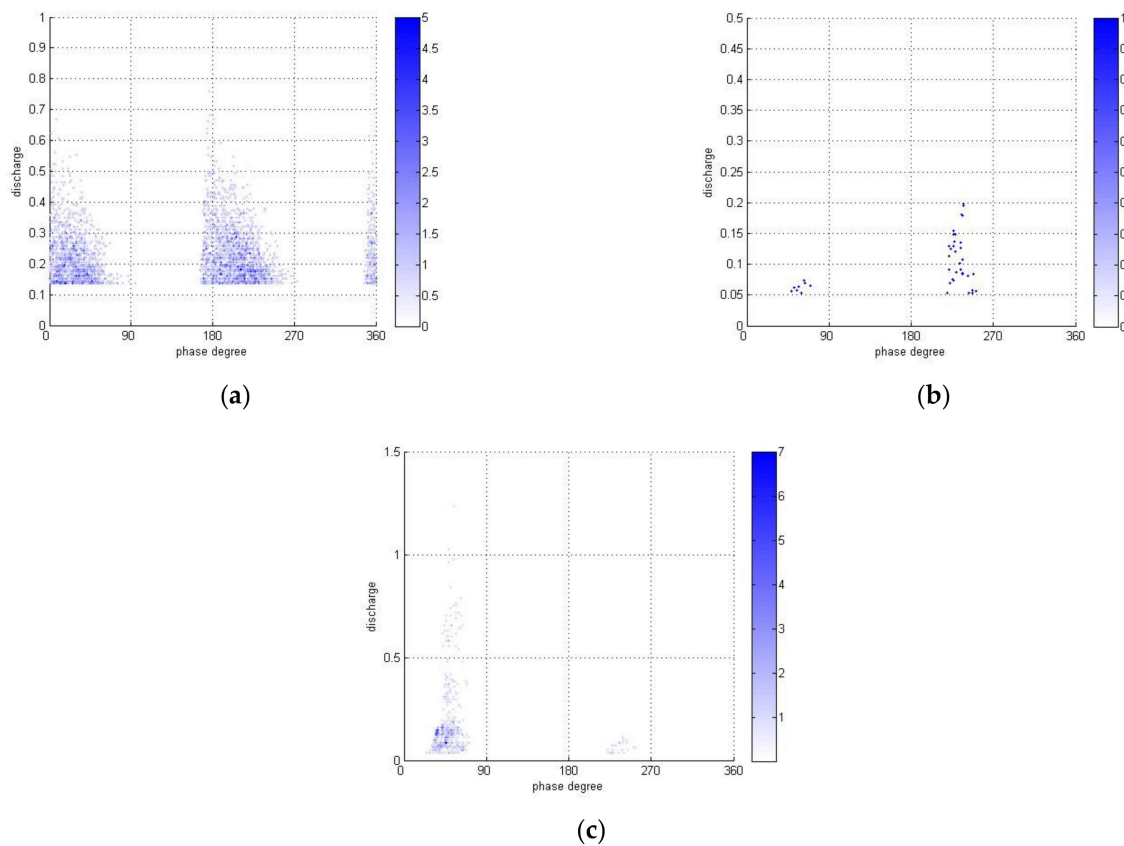


Figure 8. 26 kV $n - q - \varphi$ patterns: (a) GIS 1; (b) GIS 2; (c) GIS 3.

The 2-D distribution conversion data used are in the form of “phase angle φ —sum of discharges q_{sum} ”, “phase angle φ —maximum discharge q_{max} ”, “phase angle φ —number of discharges q_{num} ”, and “phase angle φ —average discharge q_{ave} ”, where q_{sum} , q_{max} , q_{num} and q_{ave} are called the fundamental derived quantities of discharges, and their mathematical definitions can be written as (1) to (4) [38].

- Total discharge: q_{sum}

$$q_{sum}(m, \varphi) = \sum_{t=1}^T q(m, t, \varphi) \quad (1)$$

where $q_{sum}(m, \varphi)$: the sum of discharges of phase window φ in the m_{th} measurement data, m : the m_{th} measurement data, φ : phase window (divided into 600 phase windows in this study), T : the number of electrical cycles captured (40 electrical cycles are taken in this study), $q(m, t, \varphi)$: m_{th} measurement data, the discharge volume of the φ_{th} phase window in the t_{th} electrical cycle.

- The number of discharges: q_{num}

$$q_{num}(m, \varphi) = \sum_{t=1}^T n(m, t, \varphi) \quad (2)$$

where $q_{num}(m, \varphi)$: the number of discharges in the m_{th} measurement data, the φ_{th} phase window, $n(m, t, \varphi)$: the number of discharges of the φ_{th} phase window in the t_{th} electrical cycle for the m_{th} measurement data,

$$n(m, t, \varphi): \begin{cases} 1, & \text{if } |q(m, t, \varphi)| \neq 0 \\ 0, & \text{else} \end{cases}, \quad \begin{matrix} t = 1, 2, \dots, T \\ \varphi = 1, 2, \dots, \Phi \end{matrix}$$

- Average discharge volume: q_{ave}

$$q_{ave}(m, \varphi) = \frac{q_{sum}(m, \varphi)}{q_{num}(m, \varphi)} \quad (3)$$

- Maximum discharge: q_{max}

$$q_{max}(m, \varphi) = \text{Max}\{q(m, t, \varphi) \mid t = 1, \dots, T; \varphi = 1, \dots, \Phi\} \quad (4)$$

where $q_{max}(m, \varphi)$: the maximum discharge of the φ_{th} phase window in the m_{th} measurement data.

3.3. Extraction of Characteristic Parameters of Voltage Phase

3.3.1. Basic Discharge Parameters

According to the discharge evolution spectrum ($q - \varphi - t$), the 2-D phase spectrum can be extended, and the fundamental discharge quantities can be handled as follows [39]:

- The sum of full-cycle discharge quantities: Q_{sum}

$$Q_{sum}(m) = \sum_{\varphi=1}^{\Phi} \sum_{t=1}^T q(m, t, \varphi) \quad (5)$$

where m : the m_{th} measurement data, Φ : the total number of phase windows, a total of 600 phase windows, T : total number of electrical cycles, 40 electrical cycles in total, $q(m, t, \varphi)$: the amount of discharge at “ t_{th} electrical cycle and φ_{th} phase window” in the m_{th} measurement data.

- The number of full-cycle discharges: Q_{num}

$$Q_{num}(m) = \sum_{\varphi=1}^{\Phi} \sum_{t=1}^T n(m, t, \varphi) \quad (6)$$

where $n(m, t, \varphi)$: $\begin{cases} 1, & \text{if } |q(m, t, \varphi)| \neq 0 \\ 0, & \text{else} \end{cases}$, $t = 1, 2, \dots, T$.

- Full-cycle average discharge: Q_{avg}

$$Q_{avg}(m) = \frac{Q_{sum}(m)}{Q_{num}(m)} \quad (7)$$

- Full-cycle maximum discharge: Q_{max}

$$Q_{max}(m) = \text{Max}\{q(m, t, \varphi) \mid t = 1, 2, \dots, T, \varphi = 1, 2, \dots, \Phi\} \quad (8)$$

In addition to these four basic discharge parameters, the discharge amount of the m th data is sorted from the smallest to the largest, and the discharge amount of the middle net is called the “median of the full-cycle discharge amount: Q_{med} ”. The value with the highest number of discharges in the m th data is also counted and called the “mode of full-cycle discharge: Q_{mod} ”. The relevant characteristics in this section are compiled in Table 1.

3.3.2. The Modified Cross-Correlation Factor

The correlation coefficient (9) is used to observe the linear relationship between the positive and negative half-cycles and is expressed as a correlation coefficient between $[-1, +1]$, also known as the Pearson correlation coefficient. The closer the magnitude is to

1, the higher the correlation, where a positive value represents a positive correlation and a negative value represents a negative correlation [39].

$$CC(m) = \frac{\sum_{\varphi=1}^{\Phi} (q^+(m, \varphi) - Q_{avg}^+(m)) (q^-(m, \varphi) - Q_{avg}^-(m))}{\sqrt{\sum_{\varphi=1}^{\Phi} (q^+(m, \varphi) - Q_{avg}^+(m))^2} \sqrt{\sum_{\varphi=1}^{\Phi} (q^-(m, \varphi) - Q_{avg}^-(m))^2}} \quad (9)$$

where $q^+(m, \varphi)$: the m_{th} measurement data, the positive half-period of the φ_{th} phase window value, $Q_{avg}^+(m)$: the average value of the positive half-period phase sequence of the m_{th} measurement data, $q^-(m, \varphi)$: the m_{th} measurement data, the negative half-period of the φ_{th} phase window value, $Q_{avg}^-(m)$: the average value of the negative half-cycle phase sequence of the m_{th} measurement data, Φ : the number of phase windows of the electrical cycle.

Table 1. Partial discharge characteristic parameters: basic discharge parameters.

No.	Program Code	Description
1	Q-All Cycle-Quantity Height-Sum	Total number of discharges in the whole cycle
2	Q-All Cycle-Quantity Height-Num	The number of discharges in the whole cycle
3	Q-All Cycle-Quantity Height-Ave	Average discharge volume of the whole cycle
4	Q-All Cycle-Quantity Height-Max	Maximum discharge volume of the whole cycle
5	Q-All Cycle-Quantity Height-Med	Median discharge volume of the whole cycle
6	Q-All Cycle-Quantity Height-Mod	Plurality of discharges for the whole cycle

Since the correlation coefficients describe the similarity of shapes, they cannot express the difference in numerical ratios; therefore, the modified cross-correlation factor (mCC) can be obtained by combining the discharge asymmetry (Q_{asy}) and phase asymmetry (Φ_{asy}) [24].

$$mCC = CC \times Q_{asy} \times \Phi_{asy} \quad (10)$$

where

$$Q_{asy}(m) = \frac{Q_{sum}^-(m) / Q_{num}^-(m)}{Q_{sum}^+(m) / Q_{num}^+(m)} \quad (11)$$

$$\Phi_{asy}(m) = \frac{\varphi^-(m)}{\varphi^+(m)} \quad (12)$$

where $\varphi^+(m)$ and $\varphi^-(m)$ denote the initial phase of the discharge signal for the m th measurement data average discharge distribution in the positive and negative half-periods, respectively. According to (10), the four base discharge parameters in Section 3.3.1 are calculated separately, and the extracted characteristic parameters are organized in Table 2.

Table 2. Partial discharge characteristic parameters: modified cross-correlation factor.

No.	Program Code	Description
93	Q- All Cycle-CC-Sum	Correlation coefficient between the sum of positive and negative half-cycle discharges and phase
94	Q- All Cycle-CC-Num	Number of discharges in positive and negative half-weeks - phase correlation coefficient
95	Q- All Cycle-CC-Ave	Average discharge volume of positive and negative half-weeks - phase correlation coefficient
96	Q- All Cycle-CC-Max	Maximum discharge volume of positive and negative half-weeks - phase correlation coefficient

3.3.3. Discharge Intensity Distribution

After calculating the number of discharges in the q_{ith} discharge period from the 3-D phase-resolved mapping, the PD intensity distribution can be obtained in Figure 9 [40]. In addition to the visual observation of the shape of the discharge intensity distribution to observe the change with increasing measurements, statistical methods such as mean, standard deviation, skewness, kurtosis [39–42], Weber scale parameter, Weber shape parameter, and Weber calibration value [43] were used to extract the characteristic parameters.

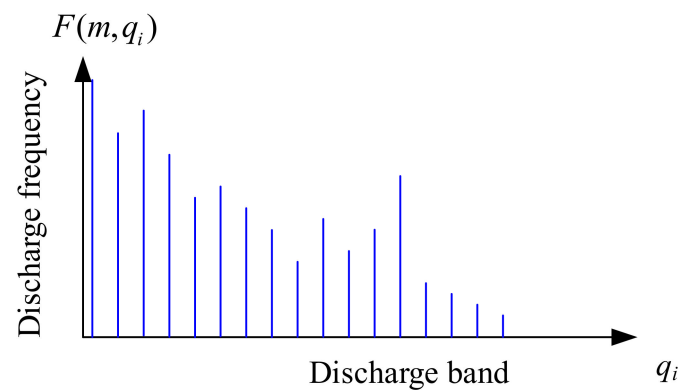


Figure 9. Schematic diagram of the discharge intensity distribution.

- Mean value: μ_F

It is used to describe the mean value of the data in the intensity distribution:

$$\mu_F(m) = \frac{1}{N_m} \times \sum_{i=1}^{N_m} F(m, q_i) \quad (13)$$

where N_m : the number of discharge zones (split into 100 discharge zones in the study), $F(m, q_i)$: the number of discharges of the q_{ith} discharges in the m_{th} measurement data.

- Standard deviation: σ_F

It is used to describe the width of the intensity distribution of the data in units of time.

$$\sigma_F(m) = \sqrt{\frac{1}{S_h(m)} \times \sum_{i=1}^{N_m} (F(m, q_i) - \mu_F(m))^2} \quad (14)$$

where

$$S_h(m) = \sum_{i=1}^{N_m} F(m, q_i)$$

- Skewness: Sk_F

It is used to describe the degree of skewness of the discharge in the shape of the intensity distribution. If the distribution shows a normal distribution, the skewed value is zero; if the tail points to the right, the skewed value is positive; if the tail points to the left, the skewed value is negative, as shown in Figure 10.

$$Sk_F(m) = \frac{\frac{1}{S_h} \times \sum_{i=1}^{N_m} (F(m, q_i) - \mu_F(m))^3}{\sigma^3} \quad (15)$$

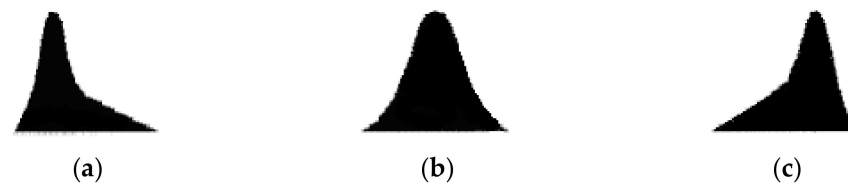


Figure 10. The shape of the skewed distribution: (a) negative value; (b) equal to zero; (c) positive value.

- Kurtosis: Ku_F

The kurtosis is used to describe the degree of the sharp and flat shape of the discharge distribution. If the distribution shows a normal distribution, the kurtosis value is zero; if it is flat, the kurtosis value is negative; if it is sharp, the kurtosis value is positive, as shown in Figure 11.

$$Ku_F(m) = \frac{\frac{1}{S_h(m)} \times \sum_{i=1}^{N_m} (F(m, q_i) - \mu_F(m))^4}{\sigma^4} - 3 \quad (16)$$

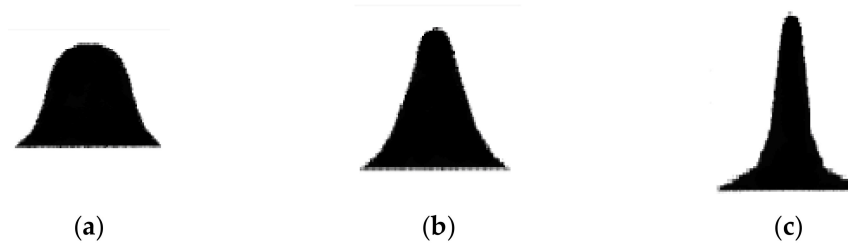


Figure 11. The shape of the kurtosis distribution: (a) negative value; (b) equal to zero; (c) positive value.

- Weber distribution chance parameters: λ_F, k_F

Assuming a Weber distribution of the number of discharges $F(m, q_i)$ with parameters (k_F, λ_F) and that the number of discharges is ordered from smallest to largest as $\{x_1, x_2, \dots, x_n\}$, the Weber cumulative distribution function can be written as Equation (17):

$$Wb(x; k_F, \lambda_F) = 1 - e^{-(x/\lambda_F)^{k_F}} \quad (17)$$

where λ_F : Weber scale parameter, k_F : Weber shape parameter, x : the difference between the highest number of discharges x_n and the lowest number of discharges x_1 , i.e., $x = x_n - x_1$.

The derivation process of the numerical solution of the above parameters can be described in the literature [44].

- Weber K-S check: H_F

H_F is used to indicate whether the intensity distribution of the check discharge is Weberian or not and is used as a characteristic parameter in this study. H_F is the assumption established by the Kolmogorov–Smirnov check method (KS-test, for short), expressed as follows:

$$H_F(m) = \begin{cases} 0 & , \text{The data is verified as Weber distribution} \\ 1 & , \text{The data is verified as non – Weberian distribution} \end{cases} \quad (18)$$

This method mainly uses the difference between the Weberian theoretical cumulative distribution function value and the actual data cumulative distribution to analyze the fitness of each measurement data. The detailed verification process of the KS-test is described in the literature [45]. Finally, the characteristic parameters extracted in this section are organized in Table 3.

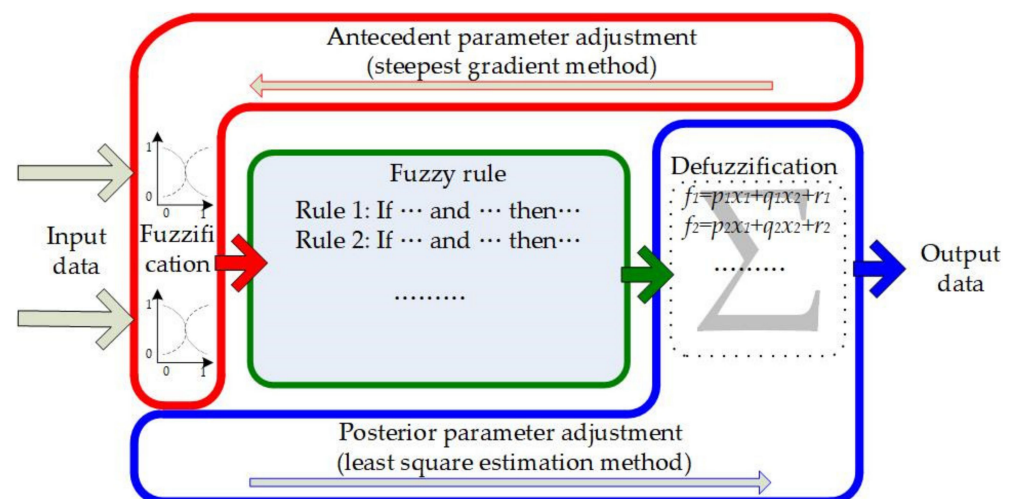
Table 3. Partial discharge characteristic parameters: discharge intensity distribution.

No.	Program Code	Description
97	Q- All Cycle-Intensity freq_mu	Average value of discharge intensity
98	Q- All Cycle-Intensity freq_std	Standard Deviation of Discharge Intensity
99	Q- All Cycle-Intensity freq_sk	Bias of discharge intensity
100	Q- All Cycle-Intensity freq_ku	Peak state of discharge intensity
101	Q- All Cycle-Intensity freq_WblScale	Weber Scale of Discharge Intensity
102	Q- All Cycle-Intensity freq_WblShape	Weber shape parameter of discharge intensity
103	Q- All Cycle-Intensity freq_KStest	Weber calibration value of discharge intensity

4. An Adaptive Class Neuro-Fuzzy Inference System

4.1. The Architecture of the Adaptive Neural-Like Fuzzy Inference System

To address the advantages and disadvantages of FISs and neural-like networks, Jang and Sun developed ANFIS in 1993 [46,47]. In addition to enhancing semantic data processing, ANFIS combines the hybrid learning procedure of FISs and neural network-like algorithms, making the network training results more perfect, as shown in Figure 12.

**Figure 12.** Schematic diagram of the composite learning process.

The primary approach is to structure the if–then rules of the FIS on a multilayer neural network-like architecture [46,47], that is, self-learning and organizing, thereby increasing the system’s ability to handle uncertainty and imprecision. It can adapt to the appropriate membership function to satisfy the relationship between input and output of the FIS, thus solving the problem that the parameters of the membership function in the FIS must be adjusted repeatedly by human thinking [48].

The fuzzy system combined with the neural-like network forms the ANFIS system with the same function as the Sugeno fuzzy system, which has the NN-like node and link structure, where the adaptive node is used to represent each corresponding fuzzy rule model in the fuzzy system. Figure 13 shows the basic architecture of ANFIS with five layers of neural nodes: input layer, rule layer, regularization layer, inference layer, and output layer, all nodes in the same layer use the same attribution function. The inference model of Sugeno is used in the fourth layer.

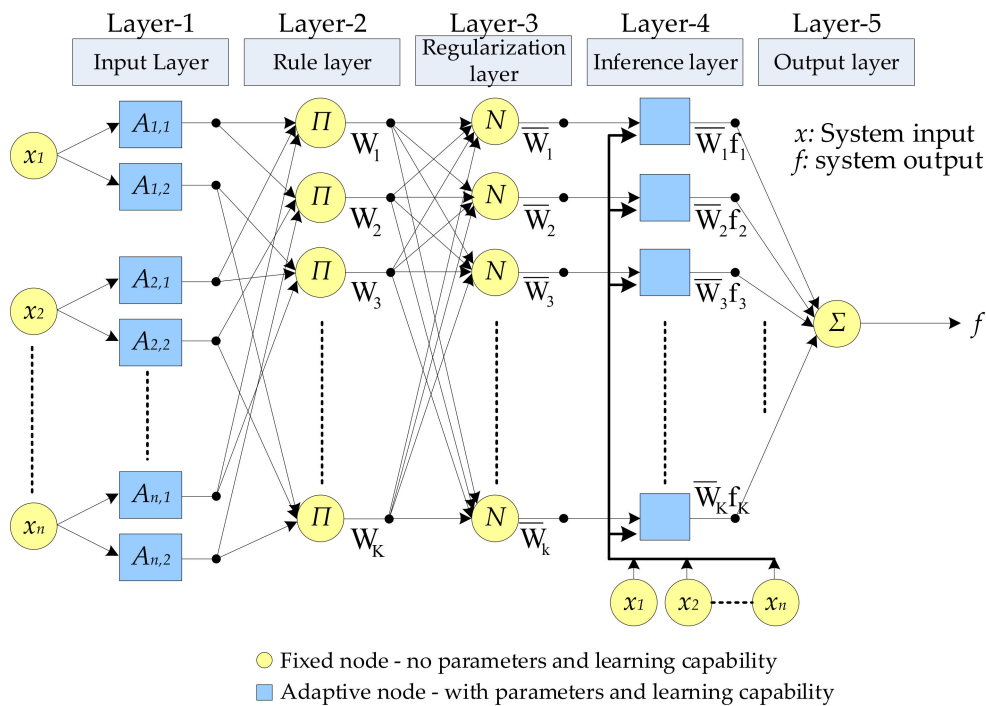


Figure 13. Schematic of ANFIS architecture (n input, 1 output).

As shown in Figure 13, the process of ANFIS inference is to fuzzify the input data and structure the attribution function of the input data. The initialization parameters of the attribution function are adjusted by the learning function of the analog neural network, and finally, the fuzzy evaluation results are obtained through the output layer. The functions of each layer are described below:

- Layer-1 input layer (adaptive nodes)

Each node of this layer is used to represent the membership function of the input independent variable (e.g., x_i). The output O_1 of the node, the function in (19) indicates the attribution degree (e.g., $\mu_{1_1}(x_1), \mu_{1_2}(x_1), \dots, \mu_{M_i}(x_1)$) of the input independent variable (e.g., x_1) imaged to the set of fuzzy functions (e.g., $A_{1,1}, A_{1,2}, \dots, A_{1,M_i}$) to which it belongs. (e.g., $\mu_{1_1}(x_1), \mu_{1_2}(x_1), \dots, \mu_{M_i}(x_1)$).

$$O_{1,j_i} = \mu_{j_i}(x_i), \text{ for } i = 1, 2, \dots, n \text{ and } j_i = 1, 2, \dots, M_i \quad (19)$$

Depending on the situation or analysis results, the attribution function can be modified to any suitable parametric function, such as triangular, trapezoidal, Gaussian, and generalized bell fuzzy functions, as in (20):

$$f_{gbellmf}(x; a_{j_i}, b_{j_i}, c_{j_i}) = \frac{1}{1 + \left| \frac{x - c_{j_i}}{a_{j_i}} \right|^{2b_{j_i}}} \quad (20)$$

The parameters a_{j_i} , b_{j_i} , and c_{j_i} are the parameters of the bell-shaped attribution function (the initialized values of the parameters are $a = 1$, $b = 1$, and c = the average of the input training data), and these “premise parameters” in ANFIS are updated by the backward transfer learning process.

- Layer-2 rule layer (fixed node)

Each node of this layer is marked by $\prod$, which means that the incoming signal of the upper layer is arithmetically multiplied (product) in each node.

$$O_{2,k} = W_k = \prod_{i=1}^n \mu_{j_i}(x_i), \quad \text{for } j_i = 1, 2, \dots, M_i \text{ and } k = 1, 2, \dots, K \quad (21)$$

where W_k denotes each rule's output triggering strength (firing strength), which can be regarded as the weight value.

Supposing there are s input variables, then s of the M_i fuzzy sets are selected for a permutation to obtain $K = (M_i)^s$ pairings, and the result O_2 is output from the nodes as (21). For example, when two input variables are imaged to the attribution function with the fuzzy set number of 3, then $3^2 = 9$ rules (rule nodes) can be composed, as shown in Figure 14. The result is output from the nodes as (21).

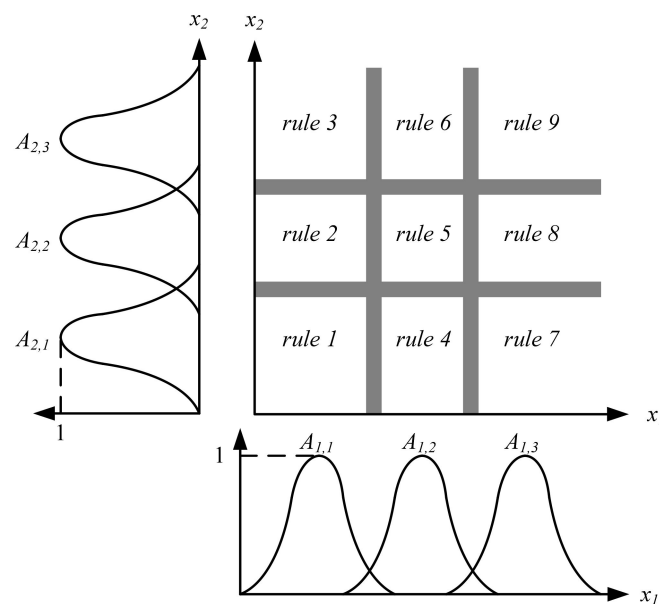


Figure 14. The input space divided into nine fuzzy sets.

- Layer-3 Regularization layer (fixed node)

Each node of this layer is labeled with N . The output is calculated as the ratio of the k^{th} rule input weight value to the trigger strength of all rules so that the output value O_3 is between 0 and 1, called normalized firing strength, as shown in (22).

$$O_{3,k} = \bar{W}_k = \frac{W_k}{\sum_{k=1}^K W_k}, \quad \text{for } k = 1, 2, \dots, K \quad (22)$$

- Layer-4 Inference layer (adaptive node)

As shown in (23), the output O_4 of each adaptive node of this layer indicates that the output of the regularization layer is multiplied by the fuzzy rule.

$$O_{4,j} = \bar{W}_k f_k = \bar{W}_k \sum_{i=0}^n r_{k_i} x_i, \quad x_0 = 1 \quad (23)$$

where $x_0 = 1$ and r_{k_i} is the parameter of the first-order Sugeno fuzzy rule (parameter initialization value is 0), called “consequent parameters”, updated in the ANFIS forward learning process.

In the compound learning process, the least square estimation method is used to find the error rate of the posterior parameters when learning in the forward direction, then the error rate is passed back to update the antecedent parameters a_{j_i} , b_{j_i} , and c_{j_i} by inverse

transfer's steepest slope descent method. The process of adjusting the parameters is thus iteratively trained to present an appropriate graph of the attribution function.

- Layer-5 Output layer (fixed node)

This layer is mainly used to convert the output generated by fuzzy inference into a clear value, so it is also called the defuzzification layer. This layer has only a single node, marked as Σ . The result is the sum of all the inputs of the previous layer as the system's overall output. By fixing the previous parameters, the overall output can be considered a linear combination of the latter parameters, and its output can be expressed by the weighted average method (24).

$$O_{5,i} = f = \sum_{k=1}^K \bar{W}_k f_k = \frac{\sum_{k=1}^K W_k f_k}{\sum_{k=1}^K W_k} \quad (24)$$

ANFIS uses a composite learning process in which, in the forward pass, the input data are passed to the Layer-4 inference layer by generating output values at each layer node, and the posterior parameters are adjusted by the least square estimation method. In the backward transfer, the error signal is backwardly transferred layer by layer to the Layer-1 input layer, and the previous parameters are updated by the steepest slope method [49]. After the supervised learning of each parameter by the neural network-like method, the FIS can be adjusted by the least-measure method to correct the parameter values of the attribution function of the fuzzy system.

4.2. Adaptive Class Neuro-Fuzzy Inference System Learning Algorithm

In the previous section, the nonlinear antecedent parameters of ANFIS were adjusted simultaneously with the posterior linear parameters to solve the nonlinear optimization problem. The nonlinear problem can also be complicated if more rules are established or more variables are input. In the error backward transfer algorithm, in addition to slowing down the convergence by reducing the learning efficiency, the problem of falling into local solutions due to poor convergence may also occur. Therefore, a compound learning algorithm was used to adjust the linear antecedent parameter group of each rule and the nonlinear posterior parameter group of the attribution function.

Table 4 shows the compound learning method and the summary of parameter correction adopted in ANFIS. After fixing the posterior parameter values and backward transferring them to the first layer, the steepest slope method is used to correct the posterior parameter values, i.e., update the first layer's attribution function. From the process of inputting fuzzy independent variables in the first layer to the process of integrating all the independent variables in the fifth layer to obtain the results, it can be seen that the primary purpose of ANFIS is to determine the fuzzy attribution function used for each independent variable (e.g., the fuzzy functions used for the independent variable x_1 are $A_{1,1}$, $A_{1,2}$), and the process of changing the parameters through repeated training finally presents various appropriate attribution function graphs.

Table 4. The compound learning program approach.

	Forward Pass	Backward Pass
premise parameters	fixed	gradient descent
consequent parameters	least-square estimator	fixed
signals	node output	error signal

5. Analysis of the Measured Data

5.1. Evaluation of Discriminator Parameters

5.1.1. Training and Testing Samples

Figure 15 shows the proportion of training data and test data in this paper, and the three data architectures used are as follows:

- Architecture 1: two of the three elements of the feature covariates are selected as training data, and the other is test data.
- Architecture 2: four out of six elements of the feature covariates are selected as training data and two as test data.
- Architecture 3: six of the nine elements of the feature covariates are selected as training data and three as test data.

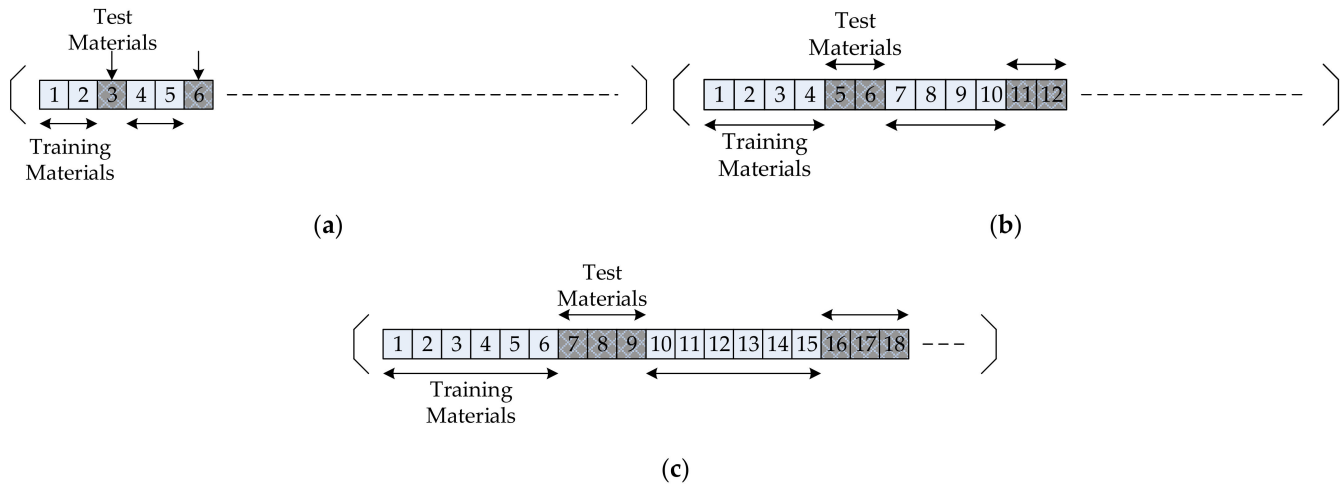


Figure 15. Schematic diagram of training data and test data architecture: (a) Architecture I: 2/3 are training materials, and 1/3 are test materials; (b) Architecture II: 4/6 for training and 2/6 for testing; (c) Architecture III: 6/9 for training data and 3/9 for testing data.

Although the partial discharge pulses are presented randomly under short time observation, for the overall insulation cracking trend, the initial discharge to insulation collapse can be regarded as a complete insulation degradation cycle, i.e., each time point is closely related to the discharge data. Therefore, the data arrangement of the three architectures is slightly different regarding the number of selected samples. When dealing with the arrangement of the elements of each characteristic parameter, they are still arranged into the training data and test data in order of measurement time.

In this study, the partial discharge data measured at 26 kV test voltage are discussed and the test statistics of the bell-type attribution function are shown in Figure 16. The average recognition rate increased only slightly with the increase in training data from architectures 1 to 3. In addition, the maximum training error value only changed slightly in architecture 3, which should not lead to overfitting and thus cannot effectively analyze the new data source. The average of the recognition rate obtained from the three data architectures is considered to make the recognition rate output by the discriminator more objective.

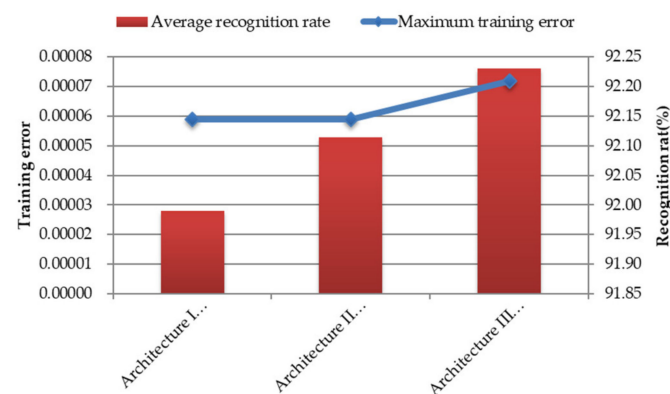


Figure 16. Analysis of training and test data structure.

5.1.2. Selecting the Type of Membership Function

In addition to the commonly used membership function determination in [50], four other built-in Matlab attribution functions, such as hyperbolic attribution function, Gaussian hybrid curve attribution function, and Π -type attribution function hyperbolic product curve attribution function, were also tested. The types of attribution functions selected are organized as follows:

- Triangular-shaped membership function;
- Trapezoidal-shaped membership function;
- Gaussian curve membership function;
- Generalized bell-shaped membership function;
- Double sigmoidal membership function;
- Gaussian combination membership function;
- Π -shaped membership function;
- Product of two sigmoidally shaped membership functions.

A feature parameter of defect type 1 (defect-1) is assumed to be used as the training data, and the trained discriminator is given the code ANFIS-1. Then the test data of defect type 1, defect type 2, and defect type 3 (from GIS-1, GIS-2, and GIS-3) are input to the discriminator ANFIS-1 for analysis, and the results of the discriminator's recognition rate of the test data are output, as shown in Figure 17.

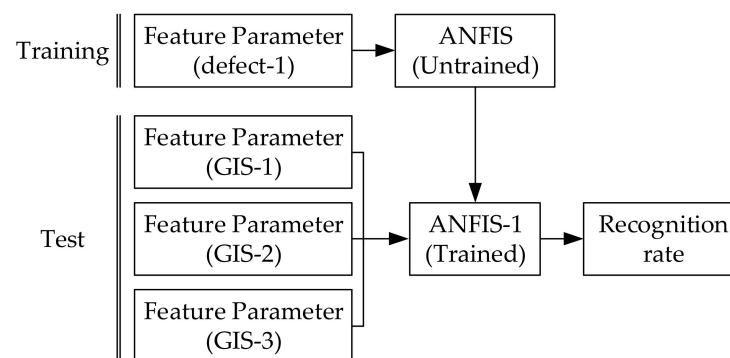


Figure 17. Discriminator training flow.

After analyzing each defect type, nine recognition rates were obtained for a single feature, as shown in Figure 18. Take ANFIS-1 analysis results as an example; “Success (%)” indicates the recognition rate of “defect type 1” when inputting GIS-1 data, and “Failure (%)” indicates the recognition rate of “defect type 1” when inputting GIS-2 or GIS-3 data.

Therefore, after executing seventeen features using the above test structure, $17 \times 9 = 153$ tests can be obtained for a single attribution function. It is not meant to list the success rate of all tests directly because it contains invalid data and a relatively high failure rate. As the data processing flow is shown in Figure 19, this section initially uses an 85% recognition rate as the threshold between “successful recognition” and “failed recognition.” When the failure rate is greater than 85%, the number of valid tests for a specific feature is ignored in the calculation and “NaN” is removed for invalid information after analysis by the discriminator for success rate information.

After the initial data processing according to the screening process in Figure 19, the success rate results below 85% are still retained for the objective evaluation of the attribution function. In addition, considering that a particular feature is not correlated with the recognition rate of each defect, a specific feature cannot be directly counted on all defects, and a single recognition rate can be used as a reference for the attribution function.

Therefore, after using (25) to calculate the effective data percentage of each attribution function, the effective data percentage histogram of each attribution function shown in Figure 20 can be obtained.

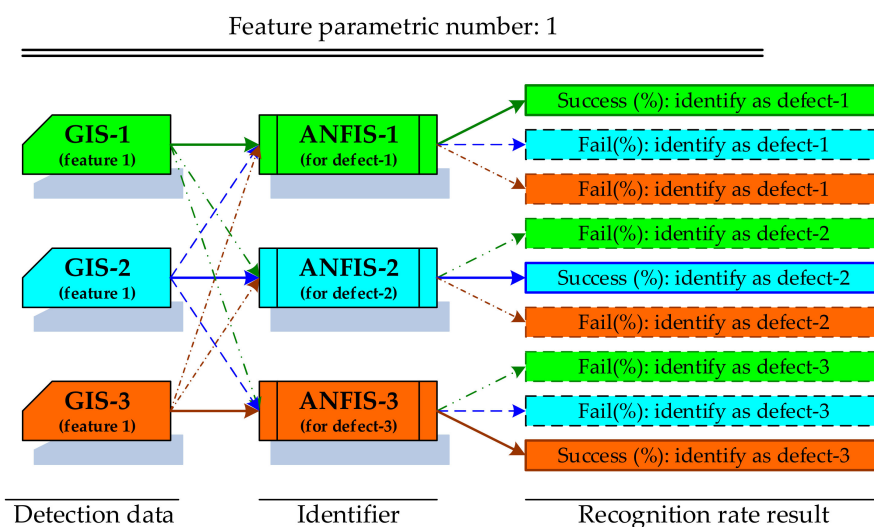


Figure 18. System architecture for testing attributable function types.

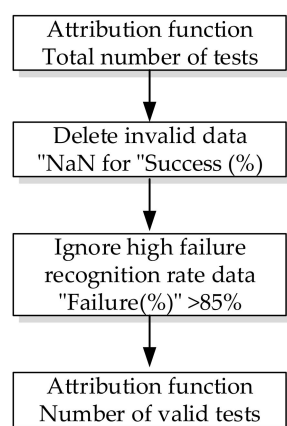


Figure 19. Data filtering process for attribution functions.

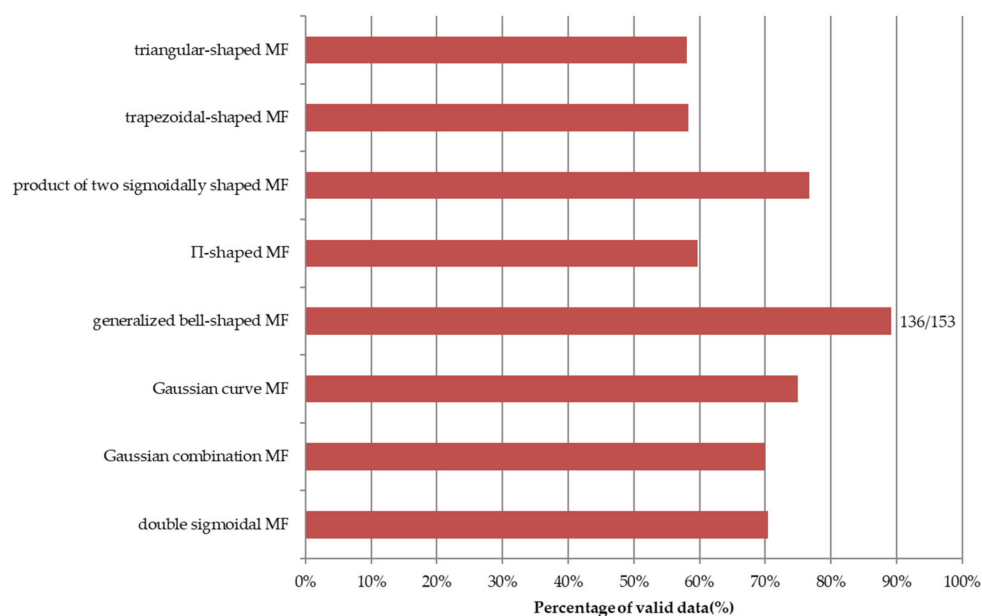


Figure 20. Percentage of valid data for each attribution function.

The percentage of valid data for the bell-type attribution function is the highest, therefore, the bell-type attribution function is used to calculate the attribution degree of the input discriminator features:

$$\text{Percentage of valid data(\%)} = \frac{\text{Number of valid tests}}{\text{total number of tests}} \times 100 \quad (25)$$

5.1.3. Number of Attribution Functions

In this section, the number of attribution functions numMF is adjusted to 3, 6, and 12 to observe the training error and training time for each feature covariate. The number of different attribution functions in the ANFIS architecture is shown in Figure 21, and the primary setup data are shown in Table 5.

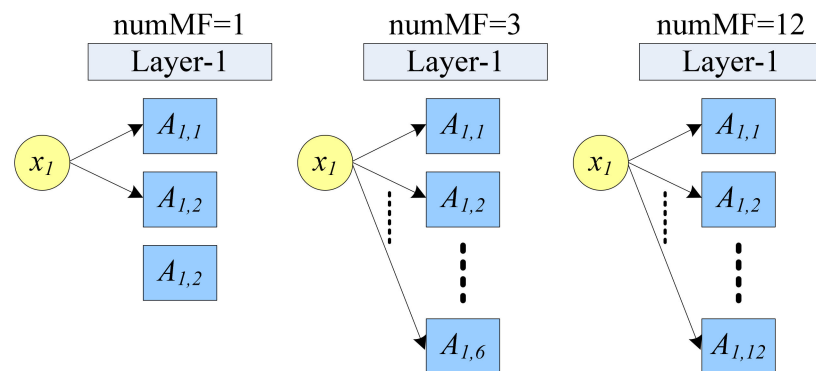


Figure 21. Diagram of different attribution functions in the ANFIS architecture.

Table 5. Basic ANFIS setup information.

	Test 1	Test 2	Test 3
MF(Layer-1):	generalized bell	generalized bell	generalized bell
numMF:	3	6	12
T-norm(Layer-2)	product	product	product
Output(Layer-5)	Weighted average	Weighted average	Weighted average

The discriminator Epoch is set to 1000, and all the data are used of each feature under 26 kV test voltage by continuous training three times, and the three training error values for each time and average are shown in Figure 22.

The above figure shows that each training error almost overlaps with the average training, showing that the system is stable. In addition, the results of the average training error values for different numbers of attribution functions are collated in Figure 23, and the training error slightly decreases as the number of MFs increases. The results of the average training time for different numbers of attribution functions are shown in Figure 24.

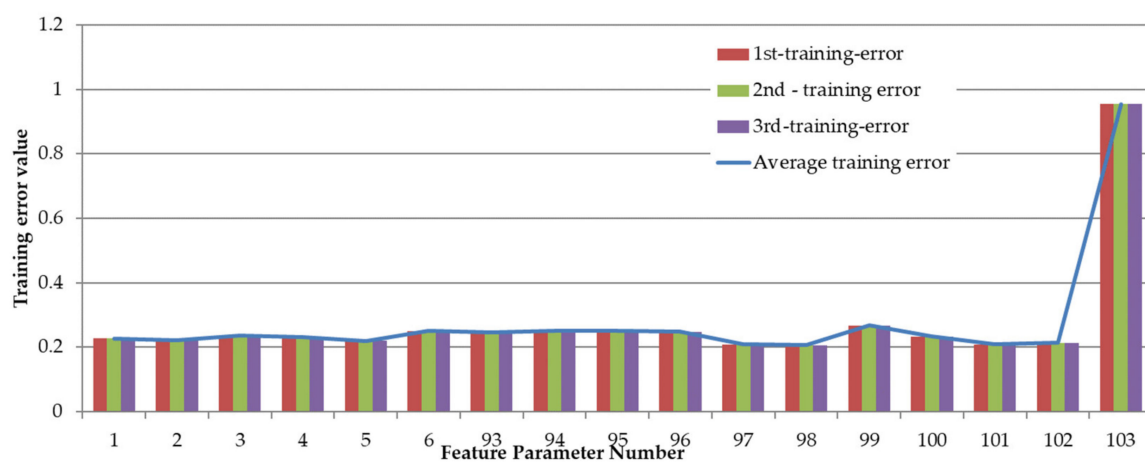
It should be noted that too-small training error values may result from overfitting, i.e., if new feature values of the local sampling data are input, they may be unrecognizable. There are two common reasons for overfitting in general training networks.

- Insufficient training samples

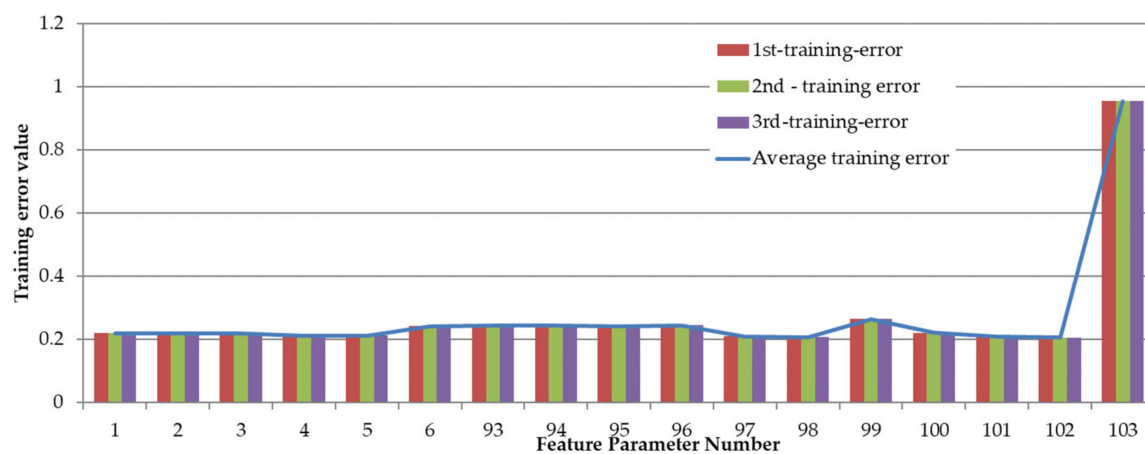
Even though the error may reach a minimum value after training, it may generate a considerable training error value when another new set of samples is provided to the network, that is, an overfitted trained network cannot classify the new source data effectively, thus creating a system that does not perform well.

- Too many network parameters

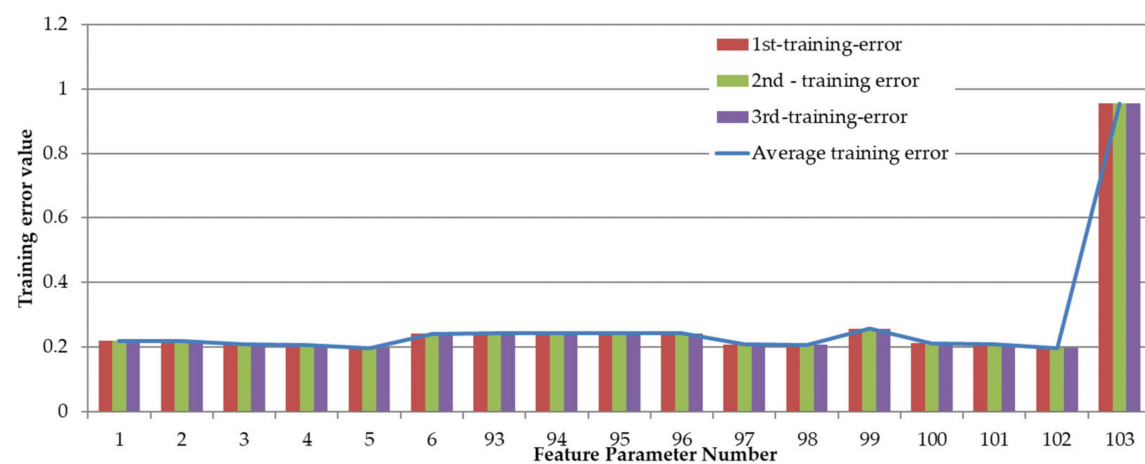
Overfitting may occur when the network is too extensive, and the wasted time (cost) due to excessive training time caused by too many parameters can also be considered a factor in evaluating the system.



(a)



(b)



(c)

Figure 22. Training error for different MFs: (a) Test 1: numMF = 3; (b) Test 2: numMF = 6; (c) Test 3: numMF = 12.

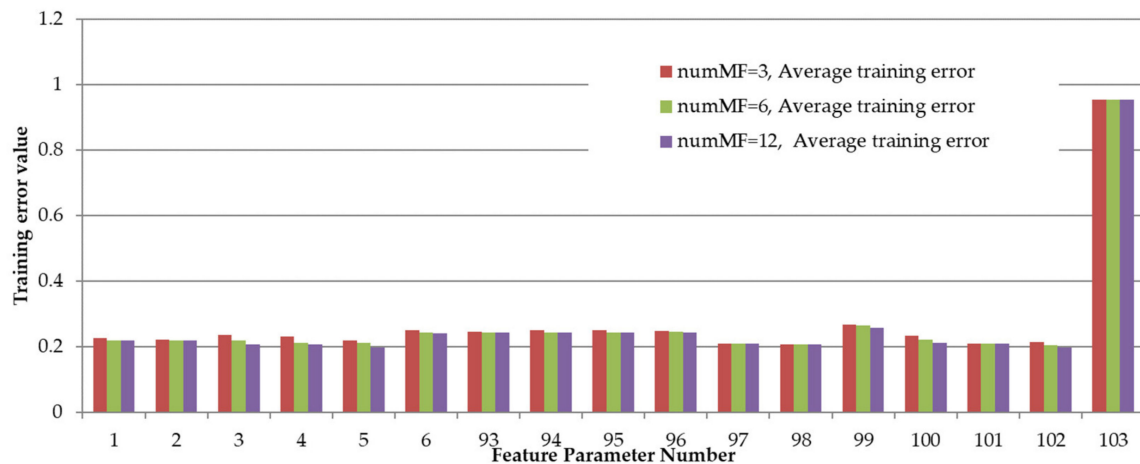


Figure 23. Average training error for different MFs.

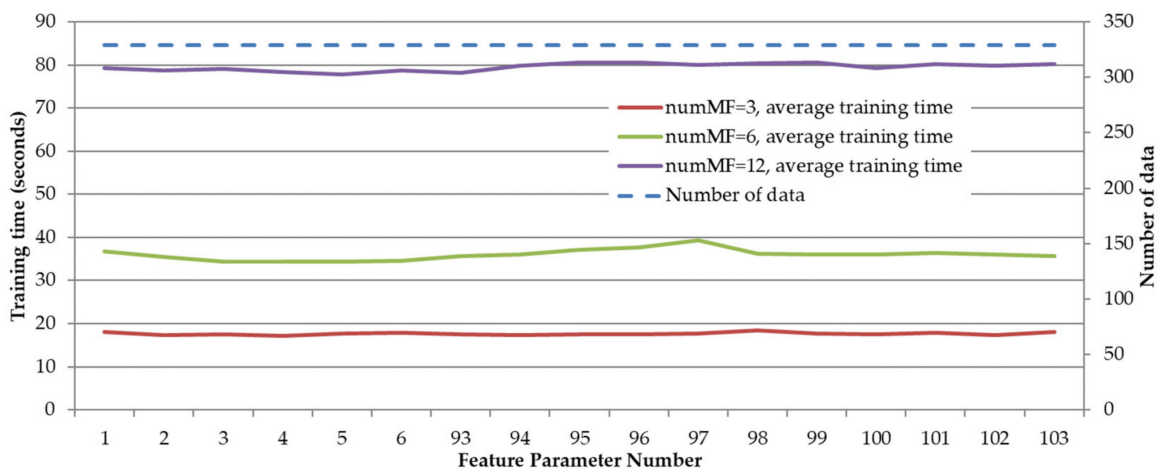


Figure 24. Average training time for different MFs.

Figure 24 shows that the number of attribution functions increases proportionally to the analysis time. Since there is no specific guideline for the size of the training error after convergence, the training efficiency of the discriminator is affected after considering the accumulation of massive data (big data) in the subsequent cycle of measurement under the acceptable recognition rate result. Therefore, the number of attribution functions $\text{numMF}=3$ is proposed to be used in the initial stage to build the recognition system.

5.1.4. Recognition Threshold

In this study, the training sample data are used as the check data to set the recognition threshold, and the mean value and standard deviation are used to set the threshold for successful recognition. The standard deviation is mainly used in statistics to measure the average distance between the observed value and the mean and is often used for statistical dispersion. In other words, the standard deviation represents the degree of dispersion centered on the mean. When all observations are the same, the standard deviation equals zero; otherwise, the standard deviation is always greater than zero. The standard deviation increases when the observed values are scattered and further away from the mean.

Assuming a set of data with a normal distribution, let the mean of this set of data be the center, and the standard deviation is used to describe the degree of dispersion of the data. Figure 25 shows the rule of chance distribution of normal distribution, and the range of chance distribution is indicated by percentage (%). Figure 25 shows that a range of plus or minus one standard deviation from the mean means that it contains more than 68% of

the data, a range of plus or minus two standard deviations from the mean covers more than 95% of the data, while a range of three standard deviations from the mean covers more than 99% of the data, i.e., it is close to covering all the data.

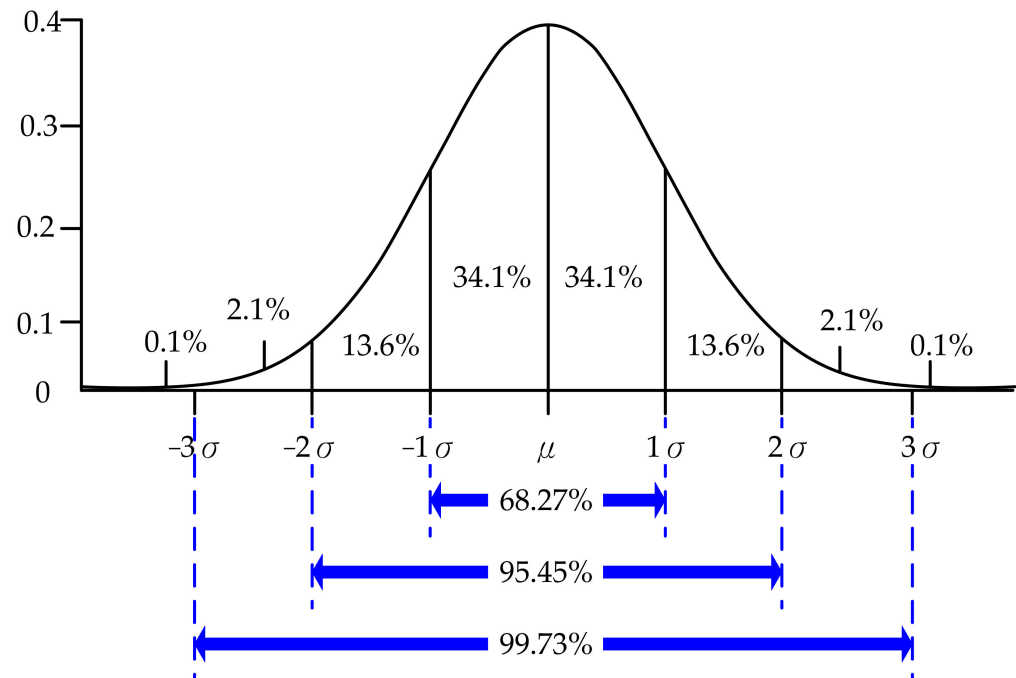


Figure 25. Rules for the mean and standard deviation in normally distributed data.

In this section, we observe the change in recognition by varying the standard deviation multiplier, as shown in Figure 26.

The graph shows that the change in standard deviation is positively correlated with the recognition rate; therefore, in addition to the completeness of the data, to make the evaluation of the recognition rate more rigorous, we adopted two times the standard deviation as the range of data variation, i.e., the average value of the recognition result of the test data plus two times the standard deviation as the threshold value of the upper and lower recognition rate.

5.2. Feature Selection Analysis and Discussion

The primary purpose of feature selection is to:

- Improve the recognition rate;
- Simplify the calculation of the classifier;
- Understand the causal relationship between the classification of features.

In this section, the analysis was performed using the professional software Matlab on a personal computer (Intel 1.17GHz CPU) with a Windows XP operating system, using the adaptive class neuro-FIS as the analysis tool for the recognition of insulation defects. Before using ANFIS as a discriminator, the initial FIS model must be established, so the arranged training data are input into the discriminator for training to evaluate whether the performance can meet the requirements. The ANFIS training settings and architecture are shown in Table 6 and Figure 27.

After the three data architectures, training data and test data are selected for the training and testing of the discriminator to achieve an objective evaluation of the initial selection of feature parameters. Finally, the recognition rate results obtained from the three architectures are the average successful recognition rate, as indicated by the black dashes in Figure 28.

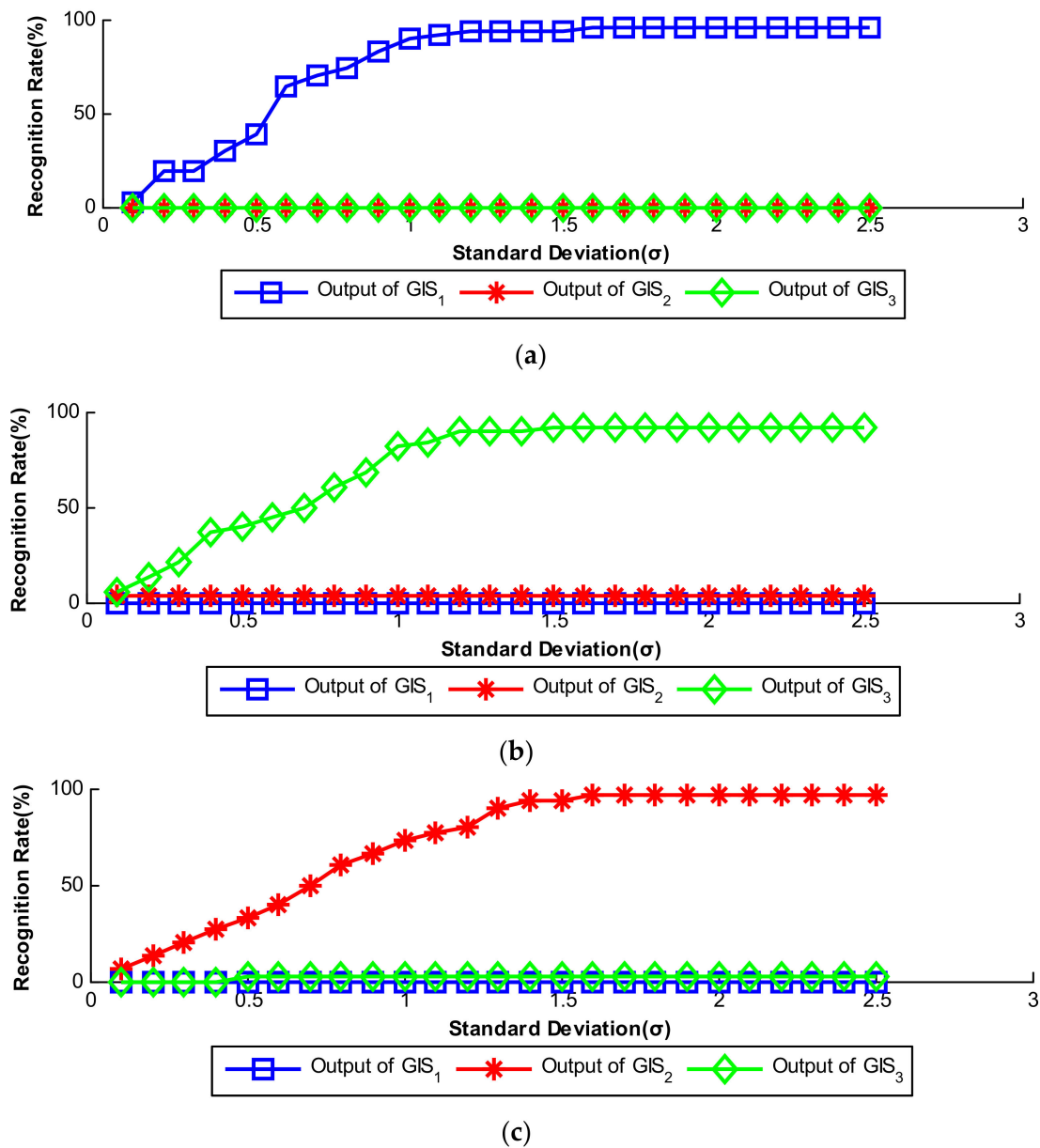


Figure 26. Standard deviation–recognition rate curve for the recognition threshold. (a) Defect type 1 (take Feature-1 as an example). (b) Defect type 2 (take Feature-93 as an example). (c) Defect type 3 (take Feature-97 as an example).

Table 6. ANFIS training setup items.

Name	Item
Number of training samples	360 (26 kV)
Number of training epochs	1000 (epochs)
Operation method measurement	T-normMax-product
Optimization method	Hybrid
Attribution function MF	generalized bell
Number of functions numMF	3

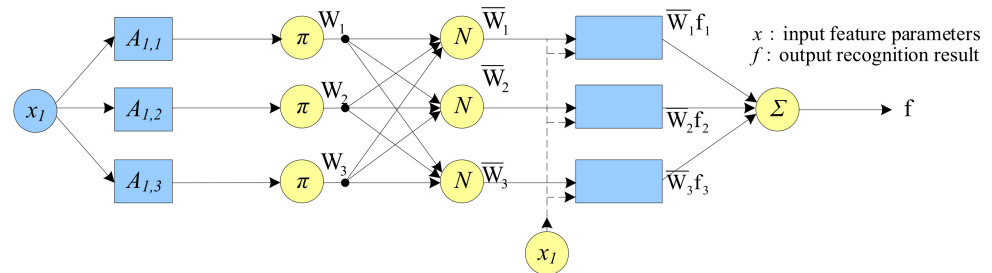


Figure 27. Schematic diagram of ANFIS architecture for partial discharge defect classification.

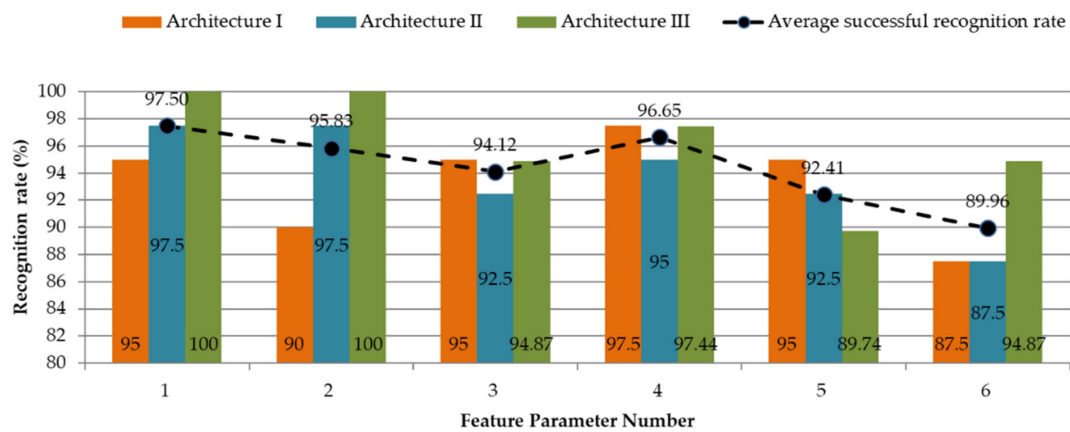


Figure 28. The successful recognition rate of each architecture (GIS-2).

Considering the practical application of defect recognition, we do not know the actual defect type of the initial input identifier data. In this section, to critically evaluate the feature selection process, we define the active average recognition rate as (26) after considering the maximum of the two average failure recognition rates:

$$REC_{EA}(i) = REC_{AS}(i) \times \{1 - \text{Max}[REC_F(i)]\} \quad (26)$$

where $REC_{EA}(i)$: The effective average recognition rate of a feature for defect type i (%), $REC_{AS}(i)$: The average successful recognition rate of a feature in defect type i (%), and $REC_F(i)$: The average failure rate of a feature for defect type i (%).

The parameters in (26) are plotted in Figure 29, where the average successful recognition rate of the feature number 10 is higher than 89.96%. However, because of its maximum average failure rate recognition rate of 44.47%, its effective average recognition rate is 49.96% after the calculation using (26).

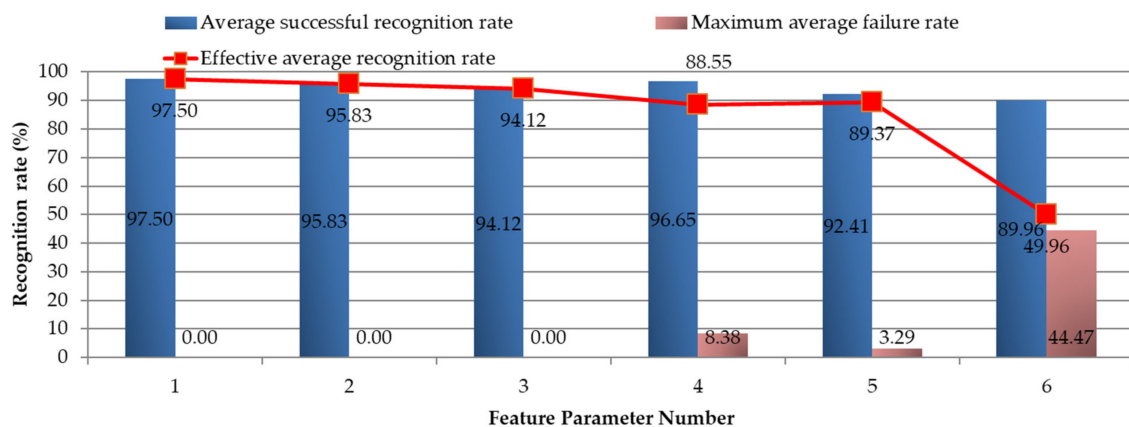


Figure 29. The effective average recognition rate.

In the subsequent sections, to save space, the successful recognition rate is not shown in each structure, but rather only the average successful recognition rate. In addition to distinguishing the effective average recognition from the average success recognition rate and the maximum average failure recognition rate by a line graph, only the data labels of the effective average recognition rate are shown to avoid confusion.

5.2.1. Feature Selection for Defect Type 1 (GIS-1)

The characteristic parameters used as partial discharges in this paper can be divided into three groups:

1. Essential discharge parametric recognition rate;
2. Cross-correlation coefficient recognition rate;
3. Discharge intensity distribution recognition rate.

A discussion of the results follows to select the feature parameters with the best recognition rate in each group.

- Partial discharge characteristic parameter group 1: basic discharge parameter recognition rate

The highest active average recognition rate in Figure 30 is 92.6% for Feature No. 1, followed by Feature No. 2, and the parameters are shown in Table 7.

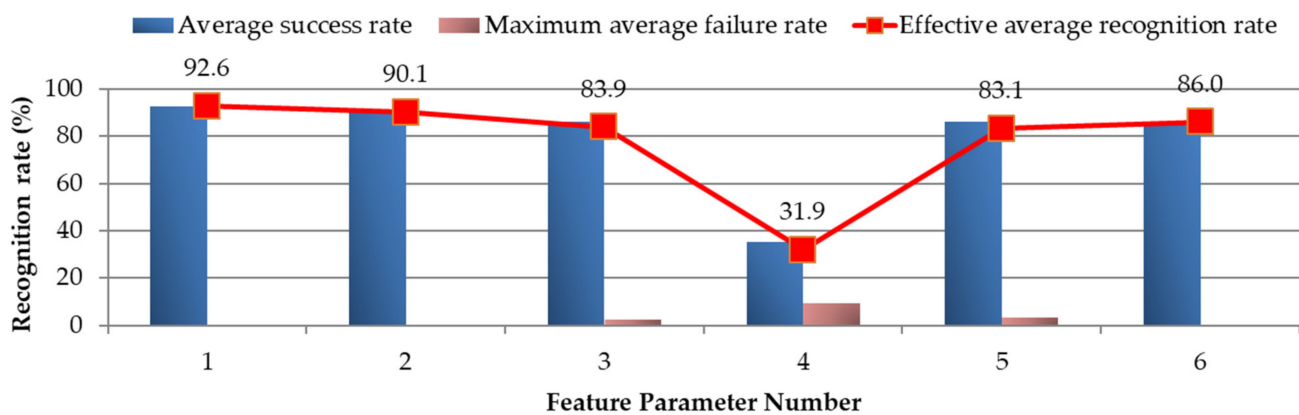


Figure 30. GIS-1 partial discharge feature: basic discharge parameter recognition rate.

Table 7. GIS-1 better recognition rate of characteristic parameters: basic discharge parameters.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
1	Total discharge of the whole cycle	92.6	0.0	92.6

- Partial discharge parameter group 2: cross-correlation coefficient recognition rate

As shown in Figure 31, the partial discharge characteristic parameters—cross-correlation coefficient recognition rate, the characteristic parameters with better recognition rate, are compiled in Table 8.

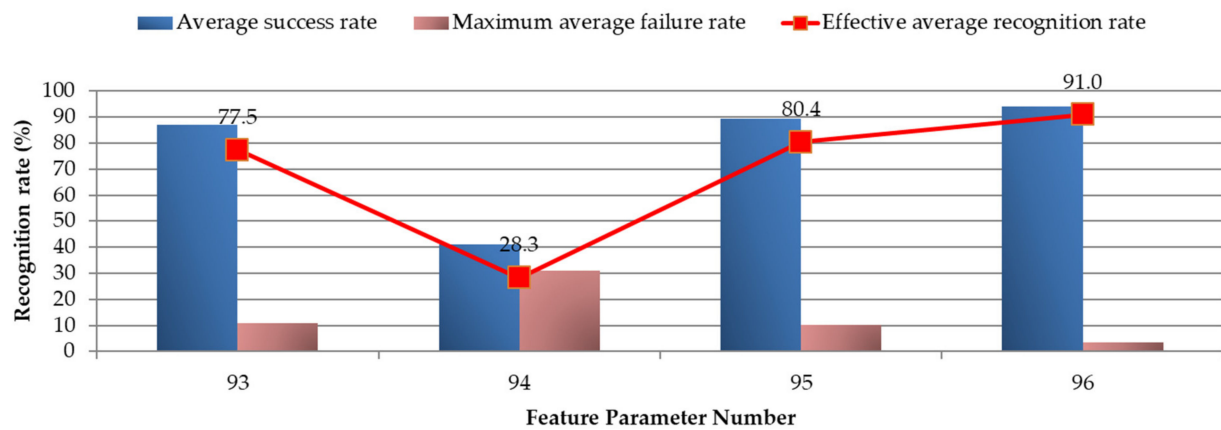


Figure 31. GIS-1 partial discharge parameter: cross-correlation coefficient recognition rate.

Table 8. GIS-1 better recognition rate of features: cross-correlation coefficients.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
96	Correlation coefficient of maximum discharge-phase of positive and negative half-cycle	94.2	3.4	91.0

- The recognition rate of the partial discharge characteristic parametric group 3: discharge intensity distribution

Figure 32 shows the recognition results of the discharge intensity distribution of the features. The features with an average recognition rate of less than 85% are No. 99, No. 100, No. 102, and No. 103, and the features with a better recognition rate of the discharge intensity distribution are shown in Table 9.

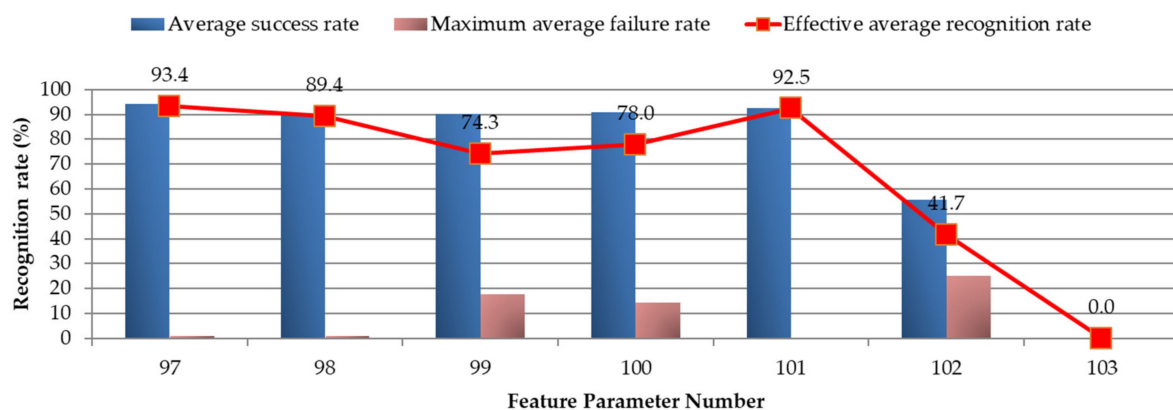


Figure 32. GIS-1 partial discharge characteristics: discharge strength distribution recognition rate.

Table 9. GIS-1 best recognition rate of characteristic parameters: discharge intensity distribution.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
97	Average value of discharge intensity	94.2	0.9	93.4

5.2.2. Feature Selection for Defect Type 2 (GIS-2)

This section selects the characteristic parameters for each group from Figures 33–35 for defect type 2, as shown in Tables 10–12.

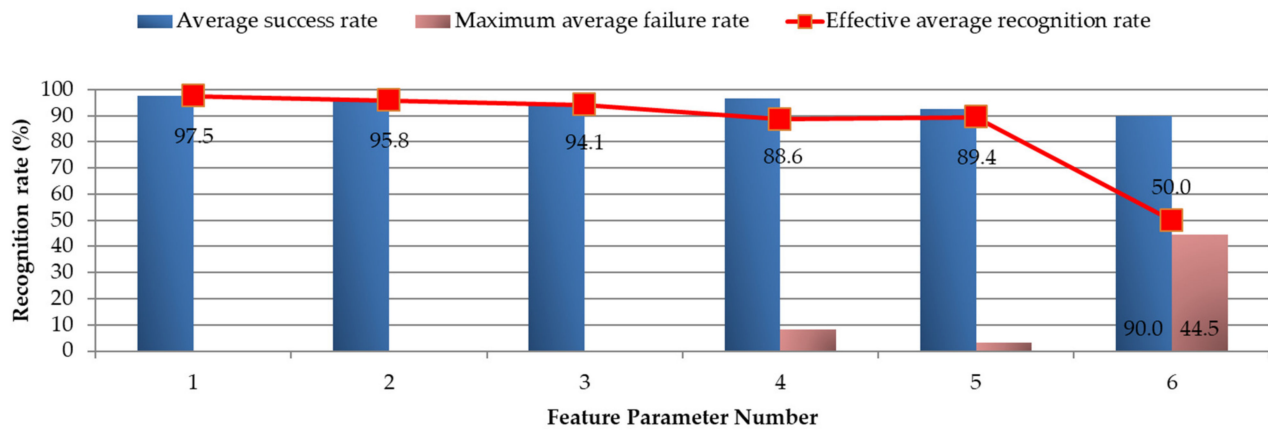


Figure 33. GIS-2 partial discharge parameter: basic discharge parameter recognition rate.

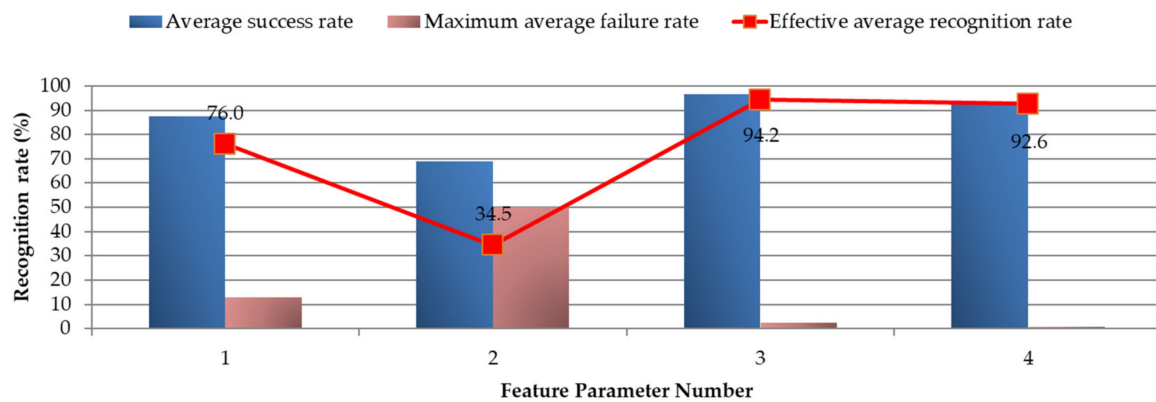


Figure 34. GIS-2 Partial discharge parametric: cross-correlation coefficient recognition rate.

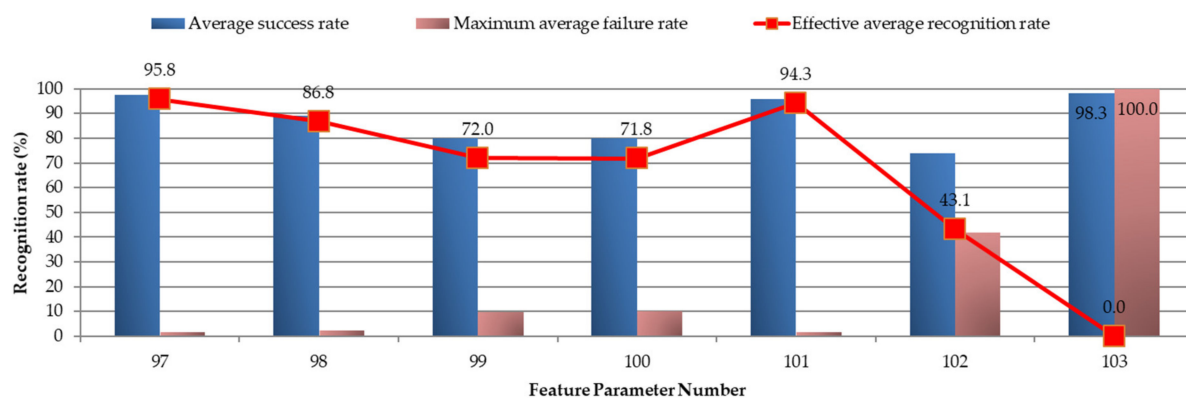


Figure 35. GIS-2 partial discharge characteristic parameter: discharge intensity distribution recognition rate.

Table 10. GIS-2 better recognition rate: basic discharge parameters.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
1	Total discharge of the whole cycle	97.5	0	97.5

Table 11. GIS-2 best recognition rate of the characteristic parametric: cross-correlation coefficients.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
95	Average discharge volume and phase correlation coefficient of positive and negative half-cycle	96.6	2.5	94.2

Table 12. GIS-2 better recognition rate of the characteristic parameters: discharge intensity distribution.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
97	Average value of discharge intensity	97.4	1.7	95.8

- Partial discharge parameter group 1: basic discharge parameter recognition rate

The recognition rate of basic discharge parameters for defect type 2 is shown in Figure 33.

The average success rate of feature number 6 in Figure 33 is up to 90%, but the maximum average failure rate is 44.5%, which means that the trained discriminator is less discriminative than other features. The features with better recognition rates for the basic discharge parameters are compiled in Table 10.

- Partial discharge characteristic parameter group 2: cross-correlation coefficient recognition rate

As shown in Figure 34, the cross-correlation coefficients of the four basic discharge parameters are shown in the cross-correlation coefficients of positive and negative half-periods. The characteristic parameters with better recognition rates of the basic discharge parameters are compiled in Table 11.

- The recognition rate of the partial discharge characteristic parameter group 3: discharge intensity distribution

Figure 35 shows the recognition results of the discharge intensity distribution of the features, with the features with less than 85% recognition rate being No. 99, No. 100, No. 102, and No. 103. The average successful recognition rate of feature number 103 is 98.3%, but the maximum average failure rate is 100%, which means that the trained discriminator has no discriminative power against foreign defects compared with other features. The features with a better recognition rate of discharge intensity distribution are shown in Table 12.

5.2.3. Feature Selection for Defect Type 3 (GIS-3)

This section selects the features for each group for defect type 2 of Figures 36–38, as shown in Tables 13–15. The selection method is similar to Sections 5.2.1 and 5.2.2. Each

group first evaluates the selection, and then the selection is judged by the result of each group's selection, whether the selection is made from the whole.

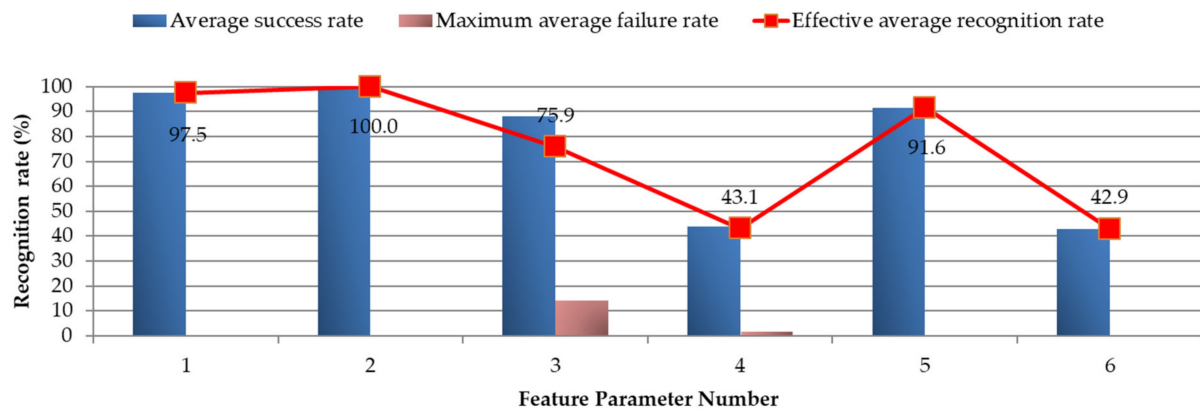


Figure 36. GIS-3 partial discharge parameter: basic discharge parameter recognition rate.

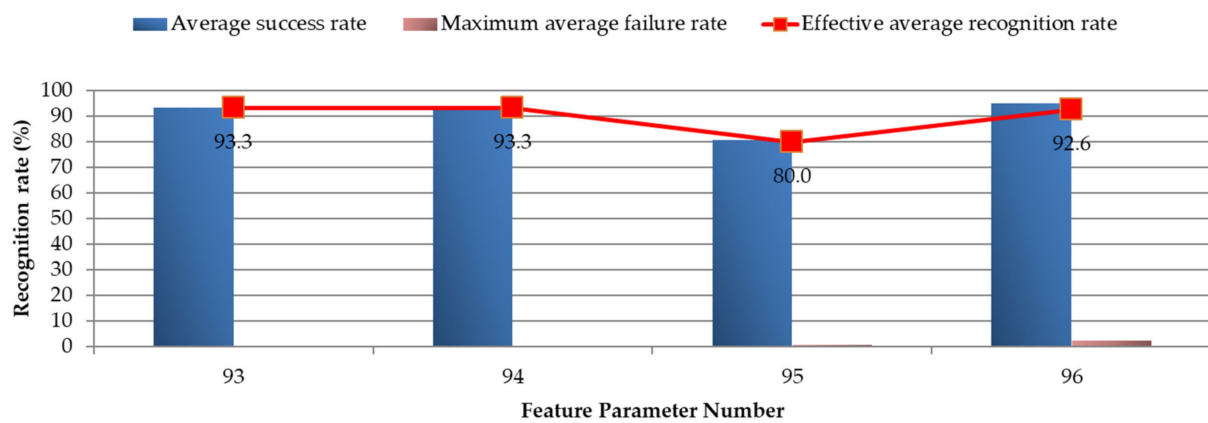


Figure 37. GIS-3 partial discharge parametric: cross-correlation coefficient recognition rate.

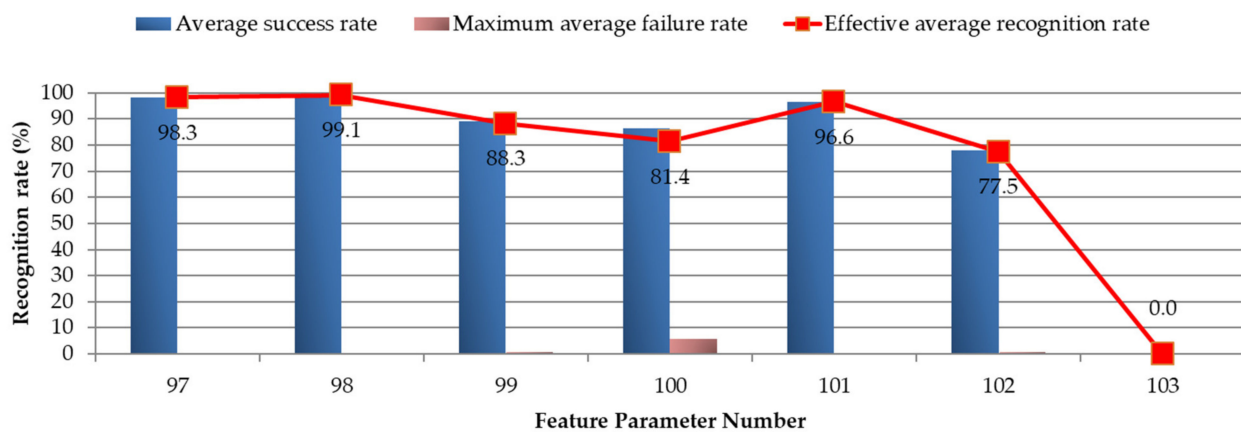


Figure 38. GIS-3 partial discharge characteristic parameters: identification rate of the discharge intensity distribution.

Table 13. GIS-3 best recognition rate: basic discharge parameters.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
2	Number of discharges in the whole cycle	100	0	100

Table 14. GIS-3 best recognition rate of the characteristic parameters: cross-correlation coefficients.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
93	Correlation coefficient of sum-phase of positive and negative half-cycle discharges	93.3	0.0	93.3

Table 15. GIS-3 better recognition rate of the characteristic parameters: discharge intensity distribution.

Feature Parameters		REC _{AS}	REC _F	REC _{EA}
No.	Name	Average Success Rate	Maximum Average Failure Rate	Effective Average Recognition Rate
98	Standard deviation of discharge intensity	99.1	0.0	99.1

- Partial discharge characteristics parameter group 1: basic discharge parameter recognition rate

The recognition rate of the basic discharge parameters for defect type 3 is shown in Figure 36.

The highest effective average recognition rate in Figure 36 is 100% for feature number 2, followed by feature number 1. The features with better recognition rates for basic discharge parameters are shown in Table 13.

- Partial discharge parametric group 2: cross-correlation coefficient recognition rate

The recognition rate of GIS-3 cross-correlation coefficients is shown in Figure 37, with a recognition rate higher than 90%, except for feature number 95, which is lower than 85%.

- Partial discharge parameter group 3: discharge strength distribution recognition rate

Figure 38 shows the recognition results of the discharge intensity distribution of the characteristic parameters. The effective average recognition rates of the characteristic parameters less than 85% are No. 100, No. 102, and No. 103, and the characteristic parameters with a better recognition rate of the discharge intensity distribution are also compiled in Table 15.

5.3. Parallel Recognition System Test Results and Discussion

The characteristic parameter numbers of each defect type are organized in Table 16, where each identifier ANFIS-1, ANFIS-2, and ANFIS-3 was used to identify the defect types that may exist inside the GIS, as follows.

ANFIS-1: The inner layer of the porcelain casing on the power supply side has oil stains.

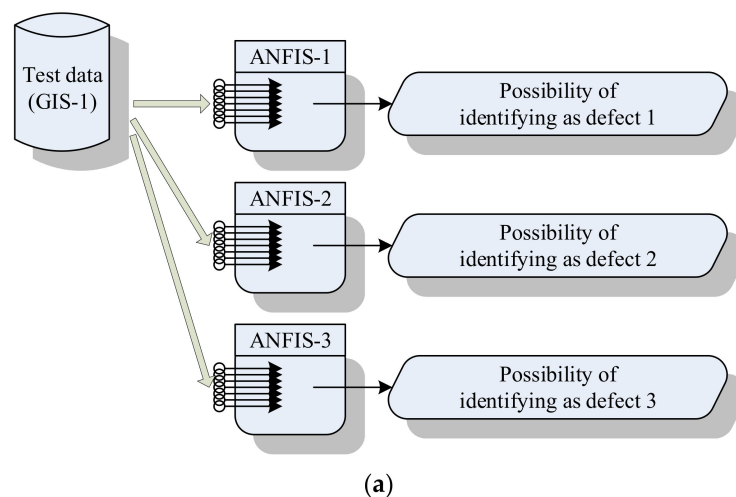
ANFIS-2: There are metal particles in the bottom layer inside the box.

ANFIS-3: There are bulges in the box linkage mechanism or conductors.

Table 16. Parallel recognition system feature numbers.

Description \ Group	I	II	III
ANFIS-1	1	96	97
ANFIS-2	1	95	97
ANFIS-3	2	93	98

From the test planning concept, the first accelerated insulation degradation test was conducted for each defective test article, with the data measured during the complete test cycle considered the initial characteristic parameter selection for each identifier. After the first measurement, each test item was transported to the required test site by a private institution commissioned by the unit or by the planned research unit. Due to the transportation process, the equipment is inevitably shaken and collides, or, according to the research plan, the SF6 gas leakage procedure needs to be conducted before the pressurization test. Therefore, after the equipment is returned to the unit through the cyclic transfer process, the data from the second accelerated insulation degradation integrity test can be used as test data for evaluating the recognition system. As shown in Figure 39a, the data obtained from the second test of GIS-1 are used as test data and input to the discriminator ANFIS-1~3 for recognition.

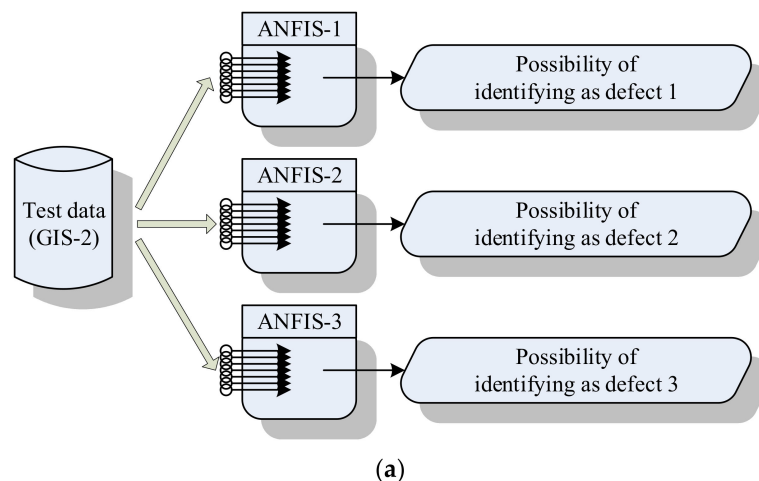


Discriminators	Feature Parameter	Recognition rate	Discriminators	Feature Parameter	Recognition rate	Discriminators	Feature Parameter	Recognition rate
ANFIS-1	1	82.5	ANFIS-2	1	0	ANFIS-3	2	0
	96	90		95	0		93	0
	97	92.5		97	2.5		98	0
	104	95		104	2.5		104	0
Average recognition rate		90.00	Average recognition rate		1.25	Average recognition rate		0.00

(b)**Figure 39.** GIS-1 in parallel recognition system: (a) test architecture; (b) test results.

The average recognition rate in Figure 39b shows that 90% are defect type 1, 1.25% are defect type 2, and 0% are defect type 3. Therefore, it is assumed that the above analysis results can be used to determine the existence of defects in the equipment if the defects are unknown.

Similarly, the second GIS-2 and GIS-3 tests' data were used as the identifier test data and input into the ANFIS-1 to ANFIS-3 identifiers. The average recognition rate was more than 90%, as shown in Figures 40 and 41, respectively, which shows that the designed recognition system can correctly identify the defects tested.

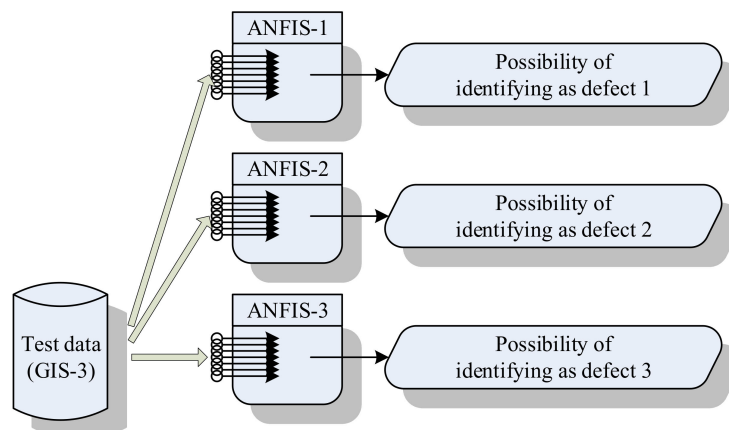


(a)

Discriminators	Feature Parameter	Recognition rate	Discriminators	Feature Parameter	Recognition rate	Discriminators	Feature Parameter	Recognition rate
ANFIS-1	1	0	ANFIS-2	1	95	ANFIS-3	2	0
	96	0		95	95		93	0
	97	2.6		97	92.3		98	0
	104	17.5		104	80		104	0
Average recognition rate		5.03	Average recognition rate		90.58	Average recognition rate		0.00

(b)

Figure 40. GIS-2 in parallel recognition system: (a) test architecture; (b) test results.



(a)

Discriminators	Feature Parameter	Recognition rate	Discriminators	Feature Parameter	Recognition rate	Discriminators	Feature Parameter	Recognition rate
ANFIS-1	1	0	ANFIS-2	1	0	ANFIS-3	2	100
	96	5		95	5		93	87.5
	97	0		97	0		98	97.4
	104	0		104	0		104	97.4
Average recognition rate		1.25	Average recognition rate		1.25	Average recognition rate		95.58

(b)

Figure 41. GIS-3 in parallel recognition system: (a) test architecture; (b) test results.

6. Conclusions

In this study, ANFIS was used as the core discriminator in a defect recognition system, with seventeen partial discharge characteristic parameters extended by several statistical

methods based on four basic discharge derivatives from the measured data as the training and testing data for the discriminator. The proposed recognition system and the selected features performed well in the recognition of three types of internal GIS defects, with recognition rate above 90%, and thus can be used as an essential reference for defect type recognition.

In addition to the continuous accumulation of laboratory test data, to implement an expert system based on partial discharge insulation diagnosis in the future, the researcher also needs to monitor actual equipment failures in the field. The partial discharge database can be continuously added to using measurement data of artificially defective specimens in the laboratory or measurements of equipment failure damage in the field. According to the proposed parallel architecture in this paper, continuous analysis and discriminative training procedures can make the recognition system more stable, producing more accurate results for practical applications.

Author Contributions: Conceptualization, I.-H.C. and Y.-H.L.; methodology, I.-H.C.; software, I.-H.C.; validation, I.-H.C. and Y.-H.L.; formal analysis, I.-H.C.; investigation, I.-H.C.; resources, I.-H.C. and Y.-H.L.; data curation, I.-H.C.; writing—original draft preparation, I.-H.C.; writing—review and editing, I.-H.C.; visualization, I.-H.C. supervision, I.-H.C.; project administration, I.-H.C.; funding acquisition, I.-H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. Tanaka, T. Aging of polymeric and composite insulating materials. Aspects of interfacial performance in aging. *IEEE Trans. Dielectr. Electr. Insul.* **2002**, *9*, 704–716. [\[CrossRef\]](#)
2. Moreno, V.M.; Gorur, R.S. AC and DC performance of polymeric housing materials for HV outdoor insulators. *IEEE Trans. Dielectr. Electr. Insul.* **1999**, *6*, 342–350. [\[CrossRef\]](#)
3. Jones, C.J. CIGRE working group 13.09-monitoring and diagnostic techniques for switch equipment. *IEEE/PES Transm. Distrib. Conf. Expo.* **2001**, *2*, 1083–1087.
4. Baker, P.C.; Judd, M.D.; McArthur, S.D.J. A frequency-based RF partial discharge detector for low-power wireless sensing. *IEEE Trans. Dielectr. Electr. Insul.* **2010**, *17*, 133–140. [\[CrossRef\]](#)
5. Sabot, A.; Petit, A.; Taillebois, J.P. GIS insulation co-ordination: On-site tests and diagnostic techniques. A utility point of view. *IEEE Trans. Power Deliv.* **1996**, *11*, 1309–1316. [\[CrossRef\]](#)
6. Cheng, Y.-H. *Electrical Equipment Insulation Inspection and Diagnosis*; China Electric Power Press: Beijing, China, 2001.
7. Suzuki, K.; Mizoguchi, H.; Ozaki, Y.; Yanabu, S. Investigation of interruption performance of newly developed 300 kV 3-phase-in-one-tank-type GCB and its application to a reduced size GIS. *IEEE Trans. Power Deliv.* **1989**, *4*, 362–367. [\[CrossRef\]](#)
8. Qi, B.; Li, C.; Geng, B.; Hao, Z. Severity diagnosis and assessment of the partial discharge provoked by high-voltage electrode protrusion on GIS insulator surface. *IEEE Trans. Power Deliv.* **2011**, *26*, 2363–2369. [\[CrossRef\]](#)
9. Tang, J.; Zhou, Q.; Tang, M. Study on mathematical model for VHF partial discharge of typical insulated defects in GIS. *IEEE Trans. Dielectr. Electr. Insul.* **2007**, *14*, 30–38. [\[CrossRef\]](#)
10. Gao, W.; Ding, D.; Liu, W. Research on the typical partial discharge using the UHF detection method for GIS. *IEEE Trans. Power Deliv.* **2011**, *26*, 2621–2629. [\[CrossRef\]](#)
11. Zhang, X.; Ren, J.; Tang, J.; Sun, C. Kernel statistical uncorrelated optimum discriminant vectors algorithm for GIS PD recognition. *IEEE Trans. Dielectr. Electr. Insul.* **2009**, *16*, 206–213. [\[CrossRef\]](#)
12. Pearson, J.S.; Farish, O.; Hampton, B.F.; Hampton, M.D.; Templeton, D.; Pryor, B.M.; Welch, I.M. Partial discharge diagnostics for gas insulated substations. *IEEE Trans. Dielectr. Electr. Insul.* **1995**, *2*, 893–905. [\[CrossRef\]](#)
13. Gao, W.; Ding, D.; Liu, W.; Huang, X. Analysis of the intrinsic characteristics of the partial discharge induced by typical defects in GIS. *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 782–790.
14. Hara, M.; Negara, Y.; Setoguchi, M.; Kurihara, T.; Suehiro, J. Particle-triggered pre-breakdown phenomena in atmospheric air gap under ac voltage. *IEEE Trans. Dielectr. Electr. Insul.* **2005**, *12*, 1071–1081. [\[CrossRef\]](#)
15. Mansour, D.-E.A.; Nishizawa, K.; Kojima, H.; Hayakawa, N.; Endo, F.; Okubo, H. Charge accumulation effects on time transition of partial discharge activity at GIS spacer defects. *IEEE Trans. Dielectr. Electr. Insul.* **2010**, *17*, 247–255. [\[CrossRef\]](#)
16. Kuffel, E.; Zaengl, W.S.; Kuffel, J. *High Voltage Engineering: Fundamentals*, 2nd ed.; Butterworth-Heinemann: Oxford, UK, 2000; pp. 421–456.

17. Ueta, G.; Wada, J.; Okabe, S.; Miyashita, M.; Nishida, C.; Kamei, M. Insulation performance of three types of micro-defects in inner epoxy insulators. *IEEE Trans. Dielectr. Electr. Insul.* **2012**, *19*, 947–954. [[CrossRef](#)]
18. Boggs, S.A.; Petterson, K.; Diessner, A.; Fujimoto, N.; Leibold, A.A. Partial discharge testing of gas insulated substations. *IEEE Trans. Power Deliv.* **1992**, *7*, 499–506.
19. Fruth, B.; Niemeyer, L. The importance of statistical characteristics of partial discharge data. *IEEE Trans. Electr. Insul.* **1992**, *27*, 60–69. [[CrossRef](#)]
20. Portugues, I.E.; Moore, P.J.; Glover, I.A.; Johnstone, C.; Mckosky, R.H.; Goff, M.B.; van der Zel, L. RF-based partial discharge early warning system for air-insulated substations. *IEEE Trans. Power Deliv.* **2009**, *24*, 20–29. [[CrossRef](#)]
21. Kusumoto, S.; Itoh, S.; Tsuchiya, Y.; Mukae, H.; Matsuda, S.; Takahashi, K. Diagnostic technique of gas insulated substations by partial discharge detection. *IEEE Trans. Power Appar. Syst.* **1980**, *99*, 1456–1465. [[CrossRef](#)]
22. Krivda, A. Automated recognition of partial discharges. *IEEE Trans. Dielectr. Electr. Insul.* **1995**, *2*, 796–821. [[CrossRef](#)]
23. Satish, L.; Zaengl, W.S. Can fractal features be used for recognizing 3-D partial discharge patterns. *IEEE Trans. Dielectr. Electr. Insul.* **1995**, *2*, 352–359. [[CrossRef](#)]
24. Sahoo, N.C.; Salama, M.M.A.; Bartnikas, R. Trends in partial discharge pattern classification: A survey. *IEEE Trans. Dielectr. Electr. Insul.* **2005**, *12*, 248–264. [[CrossRef](#)]
25. Du, B.X.; Liu, Y. Pattern analysis of discharge characteristics for hydrophobicity evaluation of polymer insulator. *IEEE Trans. Dielectr. Electr. Insul.* **2011**, *18*, 114–121. [[CrossRef](#)]
26. Cavallini, A.; Contin, A.; Montanari, G.C.; Puletti, F. Advanced PD inference in on-field measurements. I. Noise rejection. *IEEE Trans. Dielectr. Electr. Insul.* **2003**, *10*, 216–224. [[CrossRef](#)]
27. Cavallini, A.; Conti, M.; Contin, A.; Montanari, G.C. Advanced PD inference in on-field measurements. II. Identification of defects in solid insulation systems. *IEEE Trans. Dielectr. Electr. Insul.* **2003**, *10*, 528–538. [[CrossRef](#)]
28. Wang, L.; Cavallini, A.; Montanari, G.C.; Testa, L. Evolution of PD patterns in polyethylene insulation cavities under AC voltage. *IEEE Trans. Dielectr. Electr. Insul.* **2012**, *19*, 533–542. [[CrossRef](#)]
29. Tang, J.; Liu, F.; Meng, Q.; Zhan, X.; Tao, J. Partial discharge recognition through an analysis of SF₆ decomposition products part 2: Feature extraction and decision tree-based pattern recognition. *IEEE Trans. Dielectr. Electr. Insul.* **2012**, *19*, 37–44. [[CrossRef](#)]
30. Ren, M.; Dong, M.; Ren, Z.; Peng, H.-D.; Qiu, A.-C. Transient earth voltage measurement in PD detection of artificial defect models in SF₆. *IEEE Trans. Plasma Sci.* **2012**, *40*, 2002–2008. [[CrossRef](#)]
31. Chang, C.; Chang, C.S.; Jin, J.; Hoshino, T.; Hanai, M.; Kobayashi, N. Source classification of partial discharge for gas insulated substation using wave shape pattern recognition. *IEEE Trans. Dielectr. Electr. Insul.* **2005**, *12*, 374–386. [[CrossRef](#)]
32. Chang, C.-H. The Application of Partial Discharge Wavelet Transform Low-Frequency Components in The Gas Insulated Switchgear Defect Identification. Master's Thesis, National Taiwan University of Science and Technology, Taipei, Taiwan, 2012.
33. Wu, R.N.; Chung, I.H.; Chang, C.K. Classification of Partial Discharge Patterns in GIS Using Adaptive Neuro Fuzzy Inference System. *J. Chin. Inst. Eng.* **2014**, *37*, 1043–1054. [[CrossRef](#)]
34. Wu, R.N.; Chang, C.K. The use of partial discharges as an online monitoring system for underground cable joints. *IEEE Trans. Power Deliv.* **2011**, *26*, 1585–1591. [[CrossRef](#)]
35. Wu, R.-N.; Chung, I.-H.; Chang, C.-K.; Chang, C.-H.; Guo, W.-C. The Application of Partial Discharge Wavelet Transform Low-Frequency Components in The Gas Insulated Switchgear Defect Identification. In Proceedings of the Taiwan 33rd Power Engineering Conference, Taipei, Taiwan, 7–8 December 2012; pp. 2661–2734.
36. Lin, Y.H.; Wu, R.N.; Chung, I.H. Novel trend of “ ℓ ” shape in PD pattern. *IEEE Trans. Dielectr. Electr. Insul.* **2008**, *15*, 292–301. [[CrossRef](#)]
37. Zhongrong, X.; Ju, T.; Caixin, S. Application of complex wavelet transform to suppress white noise in GIS UHF PD signals. *IEEE Trans. Power Deliv.* **2007**, *22*, 1498–1504. [[CrossRef](#)]
38. Wu, R.N.; Lin, Y.H.; Chung, I.H.; Chang, C.K. Recognition of insulation status of high voltage cast-resin current transformers. *Int. J. Electr. Eng.* **2008**, *15*, 49–56.
39. Gulski, E.; Kreuger, F.H. Computer-aided recognition of discharge sources. *IEEE Trans. Electr. Insul.* **1992**, *27*, 82–92. [[CrossRef](#)]
40. Hirose, H.; Hikita, M.; Ohtsuka, S.; Tsuru, S.; Ichimaru, J. Diagnosis of electric power apparatus using the decision tree method. *IEEE Trans. Dielectr. Electr. Insul.* **2008**, *15*, 1252–1260. [[CrossRef](#)]
41. Gulski, E. Digital analysis of partial discharges. *IEEE Trans. Dielectr. Electr. Insul.* **1995**, *2*, 822–837. [[CrossRef](#)]
42. Contin, A.; Montanari, G.C.; Ferraro, C. PD source recognition by Weibull processing of pulse height distributions. *IEEE Trans. Dielectr. Electr. Insul.* **2000**, *7*, 48–58. [[CrossRef](#)]
43. Cacciari, M.; Contin, A.; Montanari, G.C. Use of a mixed Weibull distribution for the identification of PD phenomena. *IEEE Trans. Dielectr. Electr. Insul.* **1995**, *2*, 1166–1179. [[CrossRef](#)]
44. Shen, J.-Y. Estimation in Parametric Model under Interval Data. Master's Thesis, National Cheng Kung University, Tainan, Taiwan, 2002.
45. Lin, C. Real-Time Insulation Status Assessment of Underground Cable Joints Based on Statistical Theory. Master's Thesis, National Taiwan University of Science and Technology, Taipei, Taiwan, 2010.
46. Jang, J.S.R. ANFIS: Adaptive-network-based fuzzy inference systems. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]
47. Jang, J.S.R.; Sun, C.T. Neuro-fuzzy modeling and control. *Proc. IEEE* **1995**, *83*, 378–406. [[CrossRef](#)]
48. Jang, J.S.R.; Sun, C.T.; Mizutani, E. *Neural-Fuzzy and Soft Computing*; Pearson Hall: Hoboken, NJ, USA, 2004; Chapters 7–11.

-
49. Passino, K.M.; Yurkovich, S. *Fuzzy Control*; Addison Wesley: San Francisco, CA, USA, 1998; pp. 235–256.
 50. Lotfi, A.Z.; Berkely, C. *Fuzzy Logic Toolbox: User's Guide*; The MathWorks Incorporated: Natick, MA, USA, 2008; Chapters 2–7.