

Article

Fault Detection and Normal Operating Condition in Power Transformers via Pattern Recognition Artificial Neural Network

André Gifalli ^{1,*} , Alfredo Bonini Neto ² , André Nunes de Souza ¹, Renan Pinal de Mello ¹, Marco Akio Ikeshoji ³ , Enio Garbelini ² and Floriano Torres Neto ¹ 

- ¹ School of Engineering, São Paulo State University (UNESP), Bauru 17033-360, SP, Brazil; andre.souza@unesp.br (A.N.d.S.); renan.p.mello@unesp.br (R.P.d.M.); floriano.torres@unesp.br (F.T.N.)
- ² School of Sciences and Engineering, São Paulo State University (UNESP), Tupã 17602-496, SP, Brazil; alfredo.bonini@unesp.br (A.B.N.); enio.garb@gmail.com (E.G.)
- ³ Federal Institute of Education, Science and Technology (IFSP), Birigui 16201-407, SP, Brazil; ikeshoji.akio@ifsp.edu.br
- * Correspondence: andre.gifalli@unesp.br

Abstract: Aging, degradation, or damage to internal insulation materials often contribute to transformer failures. Furthermore, combustible gases can be produced when these insulation materials experience thermal or electrical stresses. This paper presents an artificial neural network for pattern recognition (PRN) to classify the operating conditions of power transformers (normal, thermal faults, and electrical faults) depending on the combustible gases present in them. Two network configurations were presented, one with five and the other with ten neurons in the hidden layer. The main advantage of applying this model through artificial neural networks is its ability to capture the nonlinear characteristics of the samples under study, thus avoiding the need for iterative procedures. The effectiveness and applicability of the proposed methodology were evaluated on 815 real data samples. Based on the results, the PRN performed well in both training and validation (for samples that were not part of the training), with a mean squared error (MSE) close to expected (0.001). The network was able to classify the samples with a 98% accuracy rate of the 815 samples presented and with 100% accuracy in validation, showing that the methodology developed is capable of acting as a tool for diagnosing the operability of power transformers.

Keywords: artificial intelligence; classification; power transformers; dissolved gas analysis (DGA)



Citation: Gifalli, A.; Bonini Neto, A.; de Souza, A.N.; de Mello, R.P.; Ikeshoji, M.A.; Garbelini, E.; Neto, F.T. Fault Detection and Normal Operating Condition in Power Transformers via Pattern Recognition Artificial Neural Network. *Appl. Syst. Innov.* **2024**, *7*, 41. <https://doi.org/10.3390/asi7030041>

Academic Editor: Emmanuel Karapidakis

Received: 29 April 2024
Accepted: 16 May 2024
Published: 24 May 2024



Copyright: © 2024 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

The socioeconomic development of a country directly influences the in demand for energy, resulting in the expansion of the infrastructure that supports it. This is reflected in the implementation of new substations, transmission, and distribution lines, and improvements to existing electrical networks, in which transformers play a crucial role. Oil power transformers represent assets with high added value, essential for the efficient operation of electrical energy transmission and distribution infrastructure [1].

In the event of faults in these transformers, negative socioeconomic impacts occur, such as fines, legal proceedings, interruption in production, safety concerns, and environmental damage, among others. These impacts affect both energy suppliers and consumers [2].

These elements have significantly impacted research and the development of methodologies with predictive and preventive maintenance approaches, aiming to mitigate the causes and consequences of unscheduled service stops. Therefore, understanding the health condition of the transformer and identifying possible initial failures has become an area of intense interest for both researchers and companies in the electrical sector, as evidenced by [3]. The main objective of this focus is to reinforce security and continuity in energy supply.

Faults in transformers can originate from several sources, and, consequently, there are several assays and tests available for their detection. Among the possible causes, thermal and electrical stresses triggered by certain events, which result in the degradation of the insulating system, composed of oil and paper, can be diagnosed in the early stages [3–5].

The dissolved gas analysis (DGA) can be measured using either the more popular method of extracting periodic samples and analyzing them in a laboratory or on-site using portable analytical equipment or continuous gas monitoring equipment installed in the transformer (high value added). Since significant gas concentration growth over a short period of time is a strong indicator of internal evolution failure, in this last example, the time interval between the periodic analyses varies according to the evolution of the gas concentrations detected between the previous and the new measurements. Under normal circumstances, this is usually reviewed once or twice a year; if failure is suspected, this period is decreased to months, weeks, or days, according to the degree of severity [6,7].

In this context, several approaches and algorithms have been developed with the support of computational intelligence, seeking to provide an accurate and effective diagnosis of the condition of oil transformers [3–5]. However, the immediate adoption of these techniques by energy utilities is not yet common. This is due to the deep-rooted traditional use of standardized analytical methods, such as gas ratios, key gas, and relative percentage graphs (triangle, pentagon), among others, which are widely used in the evaluation of transformers. These procedures have encouraged companies to invest in tools aimed at optimizing the performance of power transformers (PT). Artificial neural networks (ANN) stand out as one of these tools.

Convolutional neural network (CNN) based approach is presented for classifying six types of discharge faults in power transformers in reference [8]. The results obtained in the experiments demonstrate that the proposed method significantly outperforms conventional algorithms, such as linear and nonlinear support vector machines. In reference [9], the potential of employing deep neural networks to uncover concealed patterns within vibration time series for early-stage prediction of transformer under-excitation, over-excitation, and interturn fault progression is explored.

The developed network designed for excitation voltage prediction demonstrates outstanding performance, achieving a relative absolute error of 0.56%. However, predicting interturn faults poses a more intricate challenge, with the constructed RNN for this task exhibiting a relative absolute error of 17.58%.

Artificial neural network (ANN) is employed to improve the precision of the Rogers ratio method in reference [10]. However, it is important to acknowledge that the intricacy of an ANN requires substantial storage and computational resources. To tackle this challenge, an optimization approach is utilized with the aim of maximizing accuracy while minimizing the architectural complexity of the ANN. Consequently, post-optimization, the implemented ANN demonstrated a notable level of accuracy, reaching up to 90.7%.

A novel intelligent system utilizing dissolved gas analysis (DGA) with a dual purpose: to address the limitations of conventional methods and to enhance transformer diagnosis efficiency through the application of artificial intelligence techniques. The obtained area under the ROC curve and sensitivity average percentages of 98.78–95.19% (p -value < 0.001), respectively, underscore the impressive performance of the proposed system, offering a fresh perspective on DGA analysis is presented in [11].

In reference [12], a multimodal mutual neural network is introduced for assessing the health of power transformers. The experimental findings demonstrate that the proposed approach achieves a high level of classification accuracy and provides precise health assessments for power transformers. Topics related to the application of artificial neural networks (ANNs) in the analysis of the operating conditions of power transformers and electrical energy systems have currently received attention [13–15]. In recent years, a series of studies and research have been published in this field, demonstrating the growing interest and relevance of these approaches [16–19].

In this context, this study presents an innovative approach using a pattern recognition artificial neural network (PRN), aiming to diagnose the operating conditions of power transformers (normal operation, thermal faults, and electrical faults) based on combustible gases (H_2 (hydrogen), CH_4 (methane), C_2H_2 (acetylene), C_2H_4 (ethylene), and C_2H_6 (ethane)) present inside over time of use. The difference between this methodology and existing approaches in the literature lies in its remarkable accuracy, achieving a 98% accuracy rate in classification, with model validation reaching 100% accuracy and with a training time of just 2 s for the network with 5 neurons in the hidden layer and with 10 neurons, just 10 s of processing.

In contrast to conventional approaches that are limited to identifying a certain class of transformer failure, this research proposes a framework that allows for systematic testing with different configurations to optimize the performance of artificial neural networks. To put it simply, the method that was developed starts the training program over and again while adjusting the settings until the ideal weights are found and noted. Because the most effective parameters are stored, accuracy and efficiency are maximized throughout every training period. Selecting the best artificial neural networks (ANN) model architecture provides a more reliable methodology for practical applications, including pattern recognition skills that efficiently capture nonlinear interactions between input variables.

2. Dissolved Gas Analysis (DGA) Dataset

Three groupings of constituent gases are generated during the breakdown of transformer oil: H_2 , CH_4 , C_2H_2 , C_2H_4 , C_2H_6 , and carbon oxides (CO and CO_2). The type and concentration of gas created because of the breakdown of insulating oil give clues about the potential transformer incipient fault type, as the creation of these fault gases is dependent on temperature [5,6].

Table 1 lists the gases that can originate because of insulation failure and potential transformer early failure [6].

In thermal faults below 300 °C, ethane points to overheating of the paper or mineral oil, while methane suggests degradation of the insulating materials. At higher temperatures, between 300 °C and 700 °C, methane and ethylene indicate more serious faults, such as carbonization of the paper. Above 700 °C, the predominance of ethylene signals extreme conditions that can cause severe damage to the insulating system.

Electrical faults, on the other hand, are characterized by the presence of hydrogen in low-energy discharges and acetylene in high-energy arcs, indicating more intense and destructive electrical events. Regular analysis of dissolved gases is crucial for the early detection of these faults, allowing quick and effective interventions to prevent further damage and guarantee the continued safe operation of power transformers.

Table 1. Gas importance by faults in power transformers (adapted of [6]).

Cause of Gas Generation	Main Gas Type	Medium Gas Type	Minor Gas Type
Partial discharge—Corona	H_2	-	-
Stray gassing— $T < 200$ °C	H_2	CH_4	CH_4 , C_2H_6
Thermal fault— $T < 300$ °C	C_2H_6	CH_4	C_2H_4 , C_2H_6
Overheating of paper ¹ or mineral oil	C_2H_6	CH_4	H_2 , C_2H_4
Carbonization of paper ¹	CH_4	C_2H_4 , C_2H_6	H_2 , C_2H_4
Thermal fault— 300 °C $< T < 700$ °C	CH_4 , C_2H_4	H_2 , C_2H_6	H_2 ,
Thermal fault— $T > 700$ °C	C_2H_4	H_2 , CH_4	-
Discharge low energy or sparking	H_2	C_2H_2 , C_2H_4	C_2H_2 , C_2H_6
Discharge high energy (arcing)	H_2 , C_2H_2	C_2H_4 , C_2H_4	CH_4

¹ Paper overheating, carbonization or aging can produce CO and CO_2 .

Reference [20] is the standard for analyzing gases generated and dissolved in transformer-insulating oil. It covers everything from recommended practices for monitoring, analysis, diagnosis, and maintenance to the theory of gas creation and how it relates to faults. A totally dissolved combustible gas study examines the evolution rates of the gases (by kind and total value) and connects them to criteria for maintenance and monitoring during the transformer's operational life. Under typical working conditions, the following are the limit values for gas concentrations specified in the standard [20], expressed in $\mu\text{L}/\text{L}$ (or ppm): $\text{H}_2 = 100$, $\text{CH}_4 = 120$, $\text{C}_2\text{H}_2 = 1$, $\text{C}_2\text{H}_4 = 50$, and $\text{C}_2\text{H}_6 = 65$.

The dataset is composed of 815 samples used for training and validation phases obtained from reference [19–25]. The data is distributed as 691 normal operation samples, 52 thermal fault samples, and 72 electrical fault samples for five types of combustible gases (H_2 , CH_4 , C_2H_2 , C_2H_4 , and C_2H_6). The CO and CO_2 data were not used due to the absence of such data in some of the sources and because they were related to paper degradation [20].

The histograms of the 691 normal concentration samples for a particular type of gas are shown in Figure 1. The concentration limit for each gas at the transformer's typical operating condition is shown by the dotted red vertical line. A condition that requires care may be indicated by values that exceed this limit.

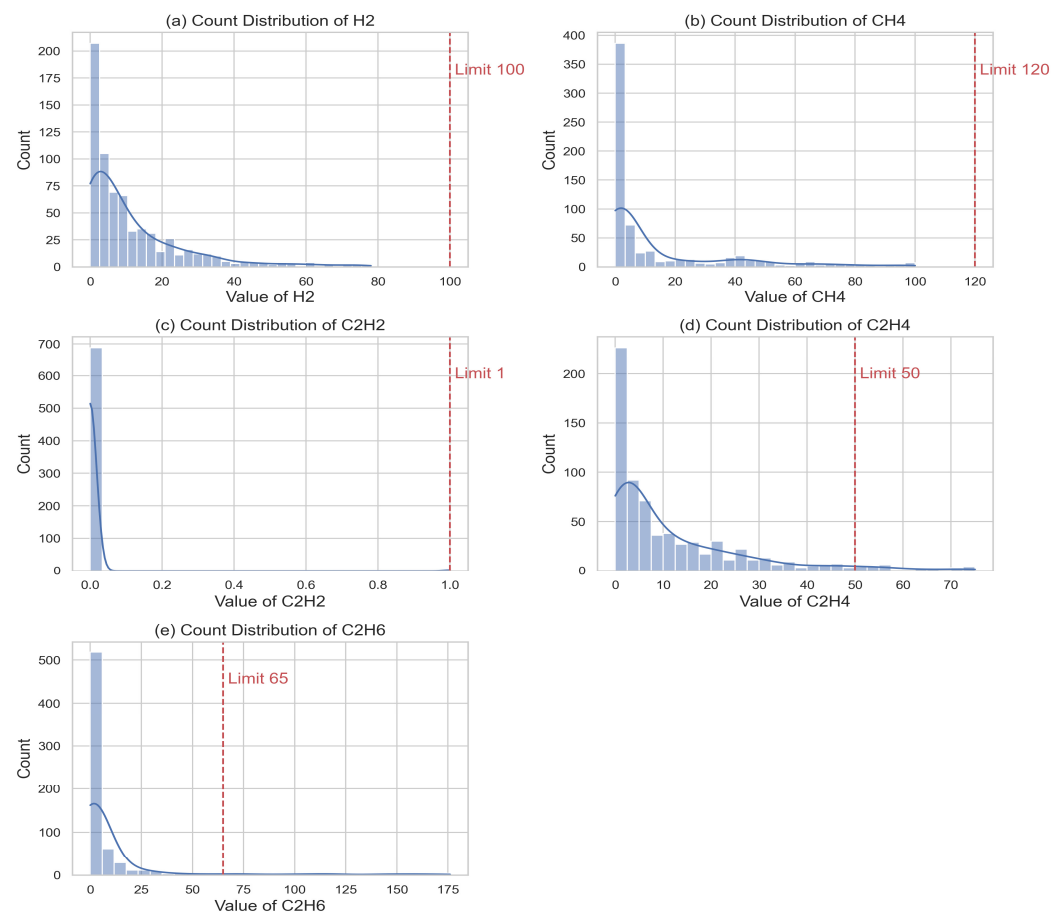


Figure 1. Distribution of the 691 samples classified as normal: (a) samples of H_2 ; (b) samples of CH_4 ; (c) samples of C_2H_2 ; (d) samples of C_2H_4 ; and (e) samples of C_2H_6 .

In general, analysis of gases dissolved in insulating oil shows that power transformers function normally. Most of the hydrogen (H_2) and methane (CH_4) concentrations are below the 100 ppm and 120 ppm standards, respectively; therefore, they do not indicate impending electrical failures. Because the concentration of acetylene (C_2H_2) is less than the threshold limit of 1 ppm, there are no potentially harmful electrical arcs present. It is evident that certain samples of ethylene (C_2H_4) and ethane (C_2H_6) surpassed 50 ppm and 65 ppm in that order. However, the limits shown are references for evaluation, but the final decision on the condition of the transformer must consider the wider context, including age, loading, and rated power of the equipment. The analysis should be complemented with an expert assessment that considers all these factors.

Figures 2 and 3 display the histograms for the 52 thermal fault samples and the 72 electrical fault samples, respectively.

Except for a few acetylene samples of electrical failures, all the gas concentration samples are shown to be above the standard limits.

The characteristics and nonlinear evolution of the samples are categorized as being in normal condition, electrical fault, and thermal fault, which are frequently different. It should be noted that a variety of factors could influence how early transformer insulation system problems manifest, including age, construction details, cooling methods, nominal power and voltage, and other factors.

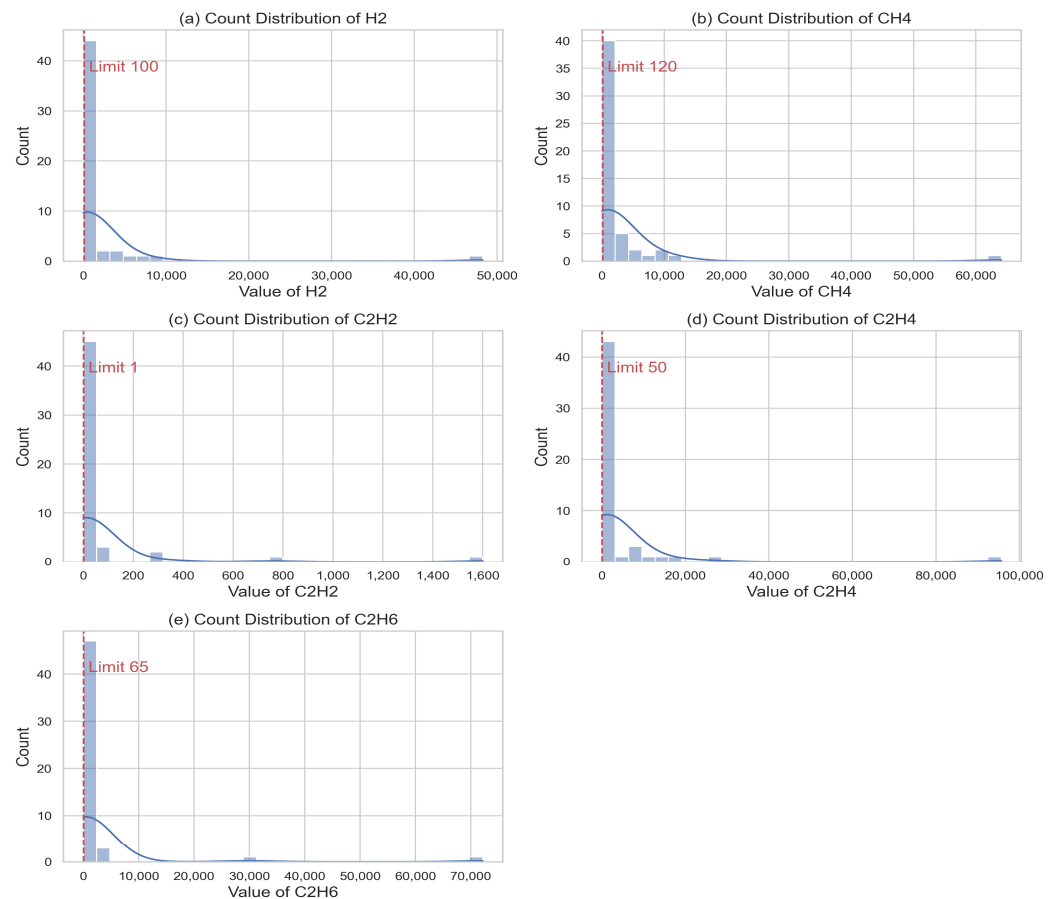


Figure 2. Distribution of the 52 samples classified as thermal failure: (a) samples of H_2 ; (b) samples of CH_4 ; (c) samples of C_2H_2 ; (d) samples of C_2H_4 ; and (e) samples of C_2H_6 .

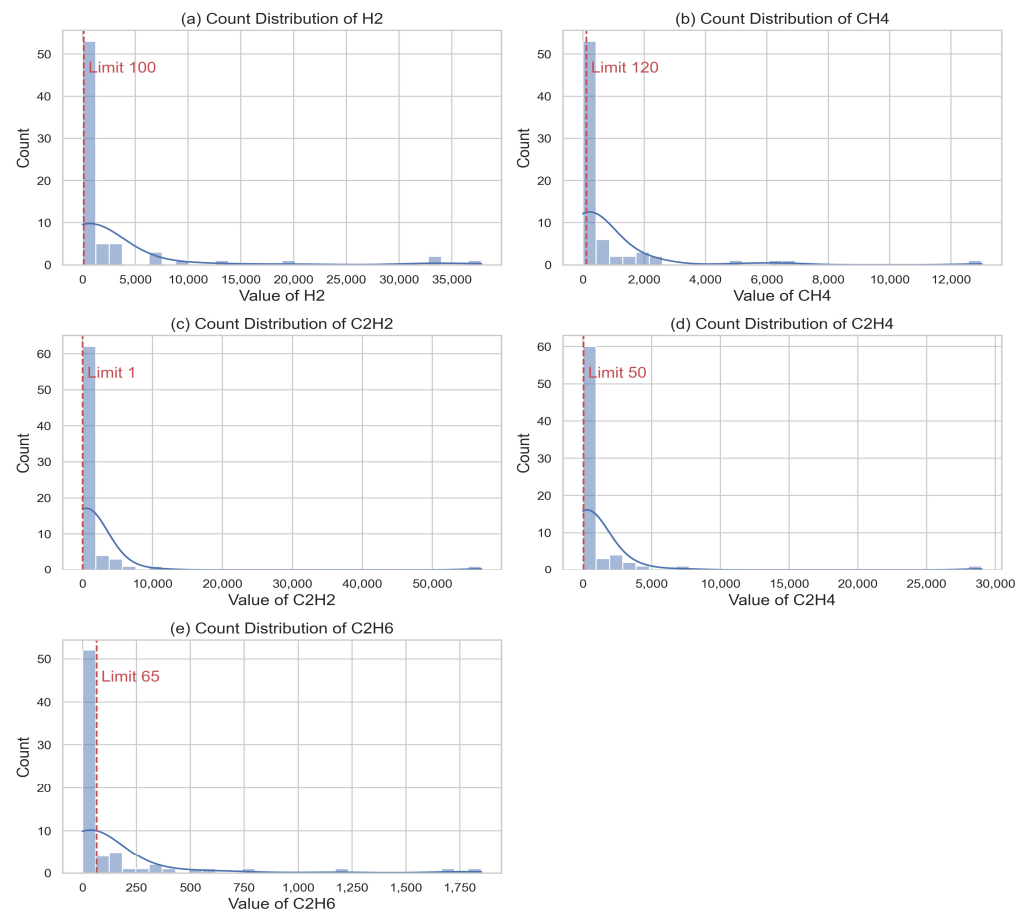


Figure 3. Distribution of the 72 samples classified as electrical failure: (a) samples of H_2 ; (b) samples of CH_4 ; (c) samples of C_2H_2 ; (d) samples of C_2H_4 ; and (e) samples of C_2H_6 .

3. Materials and Methods

The artificial neural network (ANN) used in this study, as shown in Figure 4, was the pattern recognition artificial neural network (PRN) composed of 5 neurons in the input layer (represented by the production of the combustible gases H_2 (hydrogen), CH_4 (methane), C_2H_2 (acetylene), C_2H_4 (ethylene) and C_2H_6 (ethane)), 5 or 10 neurons in the hidden layer (a comparison) and 3 neurons in the output layer, represented by 1—transformer normal operating condition, 2—thermal faults, and 3—electrical faults.

After an extensive series of tests, it was concluded that it would be essential to allocate 10 neurons in the hidden layer instead of 5. This decision was motivated by the striking similarity between the samples, mainly the samples that produced the fault outputs, 2—thermal and 3—electrical, resulting from the common characteristics they presented, which justified the need to increase the number of neurons in the hidden layer. From 10 neurons on in the hidden layer, training started to get worse.

Pattern recognition artificial neural networks (PRNs) are feedforward networks designed to classify inputs into predefined target classes. In pattern recognition networks, the target data typically comprises vectors containing all zeros except for a 1 in the element corresponding to the class it represents [26]. In pattern recognition problems, it is desired that a neural network classifies inputs into a set of target categories. For example, classify the operating condition of a power transformer as normal, thermal faults, or electrical faults based on combustible gases concentrated inside them. There are two classification methods in pattern recognition: supervised and unsupervised. To apply supervised pattern recognition, a large set of labeled data is required. If these are not available, an unsupervised approach can be applied. This work presents a supervised approach.

The scaled conjugate gradient backpropagation (SCG) is a training function for neural networks that iteratively updates weight and bias values. It utilizes the scaled conjugate gradient method, which is applicable to any network, provided its weight, net input, and transfer functions possess derivative functions. Backpropagation is employed to compute the derivatives of performance concerning the weight and bias variables [26]. SCG was the training function used in this work.

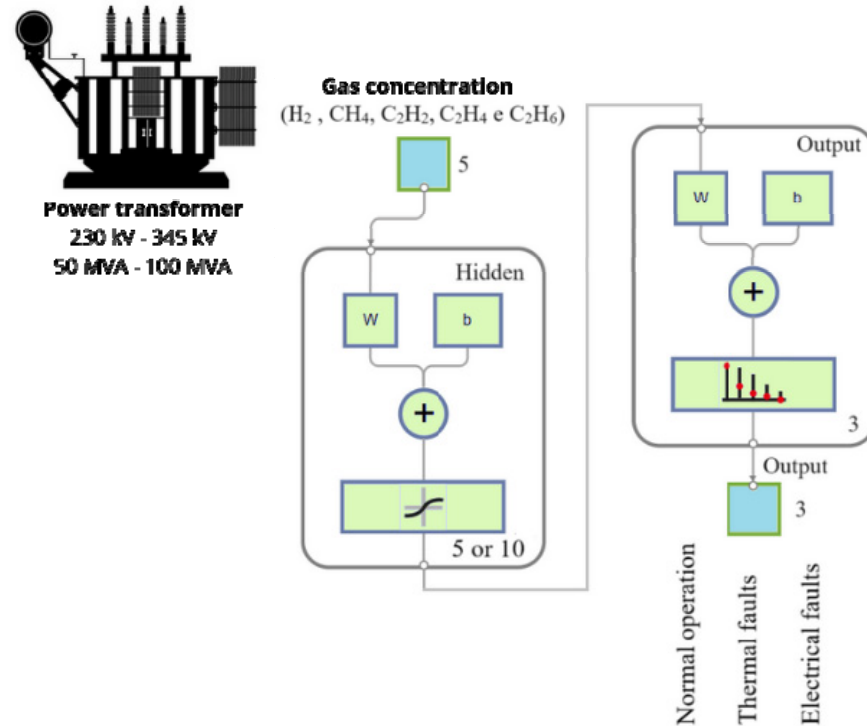


Figure 4. ANN used in this work (classified by pattern recognition network —PRN). Weights (*w*): Weights are real values assigned to each input/feature to indicate the importance of that specific characteristic in classifying the final output. Bias (*b*): Bias is used to move the activation function to the left or right. The bias determines when the activation function is called and so effects the network’s overall behavior.

The PRN network employed the hyperbolic tangent function as activation in the hidden layer, defined by Equation (1). On the other hand, for the output layer, the softmax function was adopted, represented by Equation (2).

$$f(u) = \frac{1 - e^{-tu}}{1 + e^{-tu}} \tag{1}$$

where *t* is an arbitrary constant, corresponding to the slope of the curve.

$$f(u) = \frac{e^u}{\sum e^u} \tag{2}$$

The softmax function accepts a vector *u* containing *K* real numbers as input and transforms it into a probability distribution comprising *K* probabilities. These probabilities are proportional to the exponentials of the input numbers, ensuring normalization. In other words, before applying the softmax, some components of the vector may be negative or greater than one and may not sum to 1. However, after applying the softmax, each component will be in the range (0, 1), and the components will sum to 1 so that they can be interpreted as probabilities. Additionally, higher input components will correspond to higher probabilities [21].

The mean square error (MSE) vector of the neural networks is calculated using (3):

$$\text{MSE} = \frac{1}{p} \sum_{i=1}^n (\mathbf{Y}_{\text{ob}} - \mathbf{Y}_{\text{des}})^2 \tag{3}$$

where \mathbf{Y}_{ob} and \mathbf{Y}_{des} are the obtained and desired outputs of the artificial neural network (PRN), compared during the network training, and p is the number of samples.

Neural networks employing the backpropagation algorithm, such as SCG, along with various other types of artificial neural networks, are often perceived as ‘black boxes.’ This is because it is largely unclear why these networks produce specific outcomes, as they lack explicit justifications for their predictions. Recognizing this limitation, numerous studies have focused on extracting knowledge from artificial neural networks and developing explanatory techniques to provide insights into the network’s behavior in particular situations [15,27]. Hence, it should be observed that each time the network undergoes retraining, a distinct value will be obtained [13,27]. Figure 5 presents the flowchart of the PRN network used to classify the operating conditions of a power transformer.

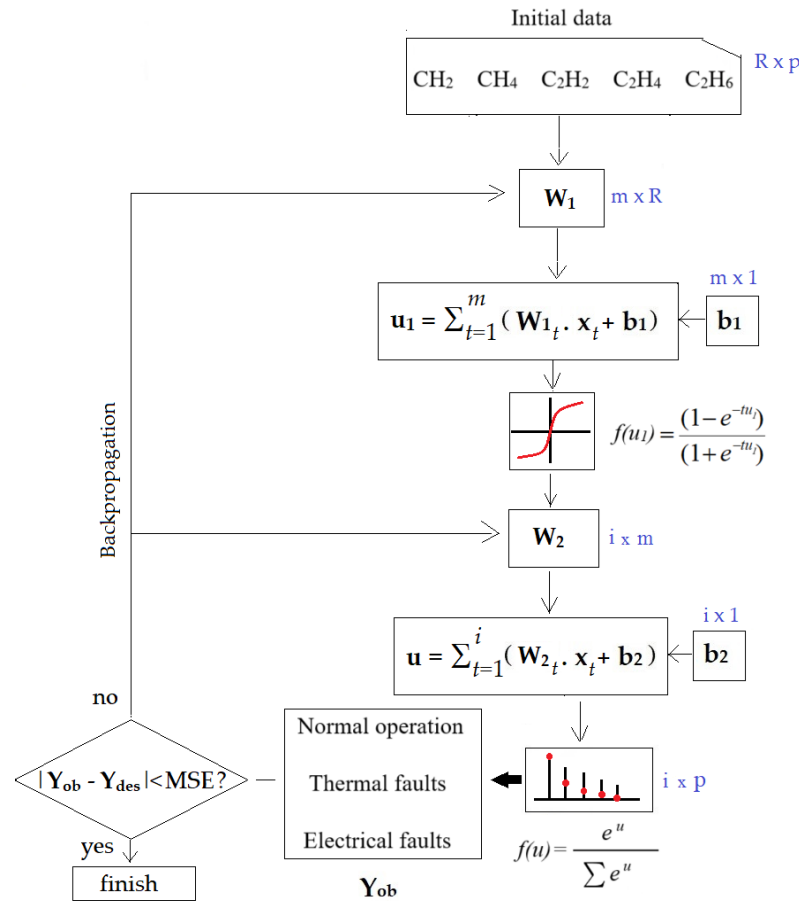


Figure 5. Flowchart of the PRN used in this work.

During training using the backpropagation algorithm (SCG), the network follows a two-step process. Initially, a pattern is introduced to the network’s input layer. The resulting activity propagates through the network, layer by layer, until the output layer generates a response. In the second step, this output is compared to the desired output for that specific pattern. If they do not match, the error is computed. This error is then propagated backward from the output layer to the input layer, and the connection weights of the internal layer units are adjusted accordingly. This process underscores the potential of the PRN application, which can function both as a classification and prediction tool. Based on this, a procedure was developed to initialize the training program multiple

times, using different configurations, both for the number of hidden layers and the number of neurons (varying in increments of 1), as well as for the proportions of training and validation samples (varying in increments of 5%), as presented in the flowchart in Figure 6. After repeating the process n times, the best result obtained was stored, corresponding to the most effective configuration (optimal number of hidden layers and neurons) that resulted in the highest accuracy percentage in the validation phase. In the study in question, the most successful configuration consisted of one hidden layer with 10 neurons, with 20% of the samples reserved for validation.

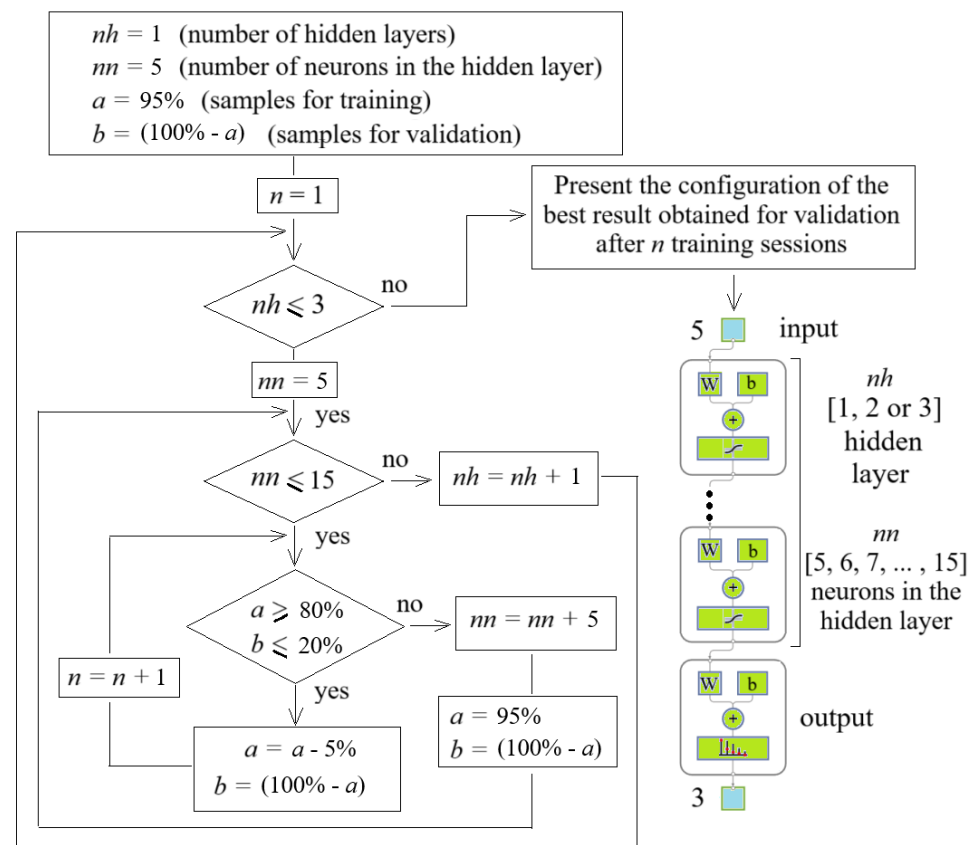


Figure 6. Flowchart of starting the training program several times (n) to choose the best configuration.

In the present situation, the most effective configuration for artificial neural network (ANN) after $n = 132$ training was the following: $nh = 1$ intermediate layer composed of $nm = 10$ neurons, with $a = 90\%$ of the samples intended for training and $b = 10\%$ for the validation phase. The second-best result found was with 5 neurons in the hidden layer.

4. Test Results and Discussion

Of the 815 samples randomly, 90% were used for training (733 samples) and 10% for validation (82 samples). The samples allocated for validation (82) are those that were not part of the initial training process. In other words, after the training, these samples (input data—gases) were introduced to the network to simulate which of the outputs (1, 2, or 3) the network classified. Two configurations were used for the hidden layer, one with 5 and the other with 10 neurons.

Figure 7 shows the training and validation performances of the artificial neural networks (PRN) used in this study. Figure 7a shows the MSE for each iteration of the PRN with five neurons in the hidden layer; the value achieved for the MSE was 0.0456 for training and 0.0595 for validation. Figure 7b shows the histogram of the error (obtained output Y_{ob} relative to the desired output Y_{des}), with 20 intervals for the 2445 data in the training and validation related to Figure 7a. The errors were around zero for most of the data. Better

results were found for the PRN with 10 neurons in the hidden layer, **MSE** at 0.0179 for training and 0.0057 for validation, which is close to what was expected (0.001), as shown in Figure 7c, with higher data accumulation and with errors around zero for the histogram, as shown in Figure 7d. Table 1 proves these results. The training and validation parameters were iterations, time, performance, and correlation. For the PRN with five neurons in the hidden layer, the values achieved were 10 s, with 113 iterations (10 validation checks achieved with 103 iterations), a correlation between the desired and obtained outputs of 0.8769 and 0.7599 to training and validation, respectively. Similar results are presented for the PRN with 10 neurons in the hidden layer, which are better results but with greater training time and number of iterations.

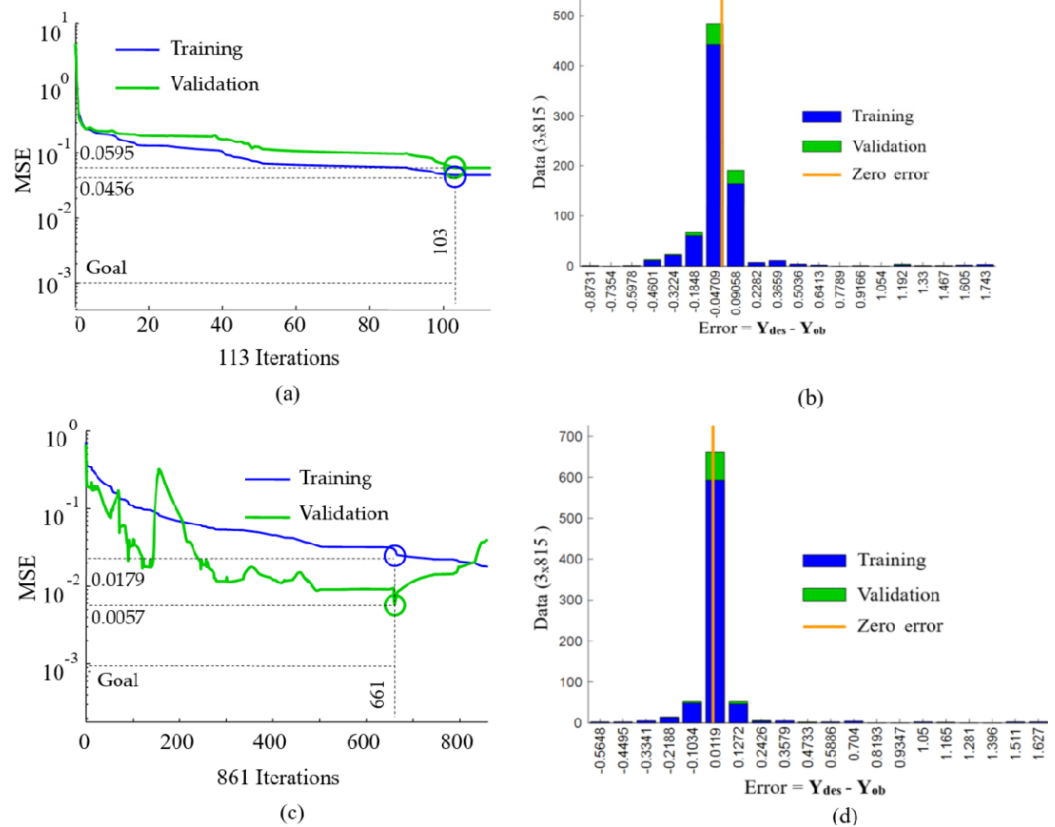


Figure 7. PRN performance: (a) training and validation performance (**MSE**) for PRN with 5 neurons in hidden layer, (b) error histogram ($Y_{des} - Y_{ob}$) for PRN with 5 neurons in hidden layer and 20 intervals for the 815 output samples, (c) training and validation performance (**MSE**) for PRN with 10 neurons in hidden layer, (d) error histogram ($Y_{des} - Y_{ob}$) for PRN with 10 neurons in hidden layer and 20 intervals for the 815 output samples.

Figure 8a displays the results with five neurons in the hidden layer, comprising a comparison between the outputs obtained by the ANN (Y_{ob}) and the desired outputs (Y_{des}) derived from experiments during the training phase (90%, 733 samples). A notable resemblance between the Y_{ob} and Y_{des} outputs is observed, indicating effective network training, as illustrated in Figure 7 and described in Table 2. Consequently, the ANN is now capable of estimating data (such as composition) for samples that were not part of the training process. An automated model has been developed to estimate these parameters based on new sets of input data (combustible gases H_2 (hydrogen), CH_4 (methane), C_2H_2 (acetylene), C_2H_4 (ethylene) and C_2H_6 (ethane)). Figure 8b presents the results of the network validation phase for 82 samples (10%) of the input data that were not part of the training, with the desired and obtained output values. Again, there is a similarity between the outputs, proving the effectiveness of the model created via ANN. The **MSE** between outputs for this phase was 0.0595, similar to the specified value.

Table 2. Values specified and achieved in training and validation of PRN with 5 and 10 neurons in the hidden layer compared to Y_{des} output.

PRN (5, 5, 3)	Specified Values	Achieved Values
Iterations	1000	113
Time (s)	20	2
Performance (MSE) training	0.001	0.0456
Correlation (R^2)	1.0	0.8769
Performance (MSE) validation	0.001	0.0595
Correlation (R^2) validation	1.0	0.7599
Validation checks	10	10 *
PRN (5, 10, 3)	Specified Values	Achieved Values
Iterations	1000	861
Time (s)	20	10
Performance (MSE) training	0.001	0.0179
Correlation (R^2)	1.0	0.9252
Performance (MSE) validation	0.001	0.0057
Correlation (R^2) validation	1.0	0.9718
Validation checks	10	10 *

* Achieved criterion.

Figure 8c represents the results for 100% of the samples (815 in total), encompassing both the training and validation phases simultaneously. There were a total of 18 errors, with 15 errors during training and only 3 errors during validation, demonstrating the effectiveness of the model.

As a result, the following confusion matrix, Figure 9, was obtained. On the confusion matrix plot, the rows correspond to the predicted class (Y_{ob} —obtained via ANN), and the columns correspond to the true class (Y_{des} —Target). The cells along the diagonal represent instances that are accurately classified, while those of the diagonal denote misclassified observations. Each cell displays both the count and the percentage of observations relative to the total.

The rightmost column provides the percentages of all predicted examples for each class that are correctly and incorrectly classified. These metrics are commonly known as precision (or positive predictive value) and false discovery rate, respectively [20].

Similarly, the bottom row displays the percentages of all examples belonging to each class that are correctly and incorrectly classified. These metrics are often referred to as recall (or true positive rate) and false negative rate, respectively. Finally, the cell in the bottom right corner of the plot indicates the overall accuracy [26].

In Figure 9, the first three diagonal cells show the number and percentage of correct classifications after the training (Figure 9a) and validation (Figure 9b) of the network, respectively. For example, in Figure 9a, 616 samples are correctly classified as class 1 (normal operation). This corresponds to 84% of all 733 samples. Similarly, 46 cases are correctly classified as class 2 (thermal faults). This corresponds to 6.3% of all samples. Finally, 56 samples were classified as class 3 (electrical faults), corresponding to 7.6% of all samples.

Overall, 97.9% of the predictions are correct, and 2.1% are wrong for training. Similar results are presented for validation in Figure 9b: it is observed that 96.3% of the predictions are correct and 3.7% are wrong. In both phases, the network achieved an accuracy of 97.8%, hitting 797 samples and an error rate of just 2.2%, as shown in Figure 9c.

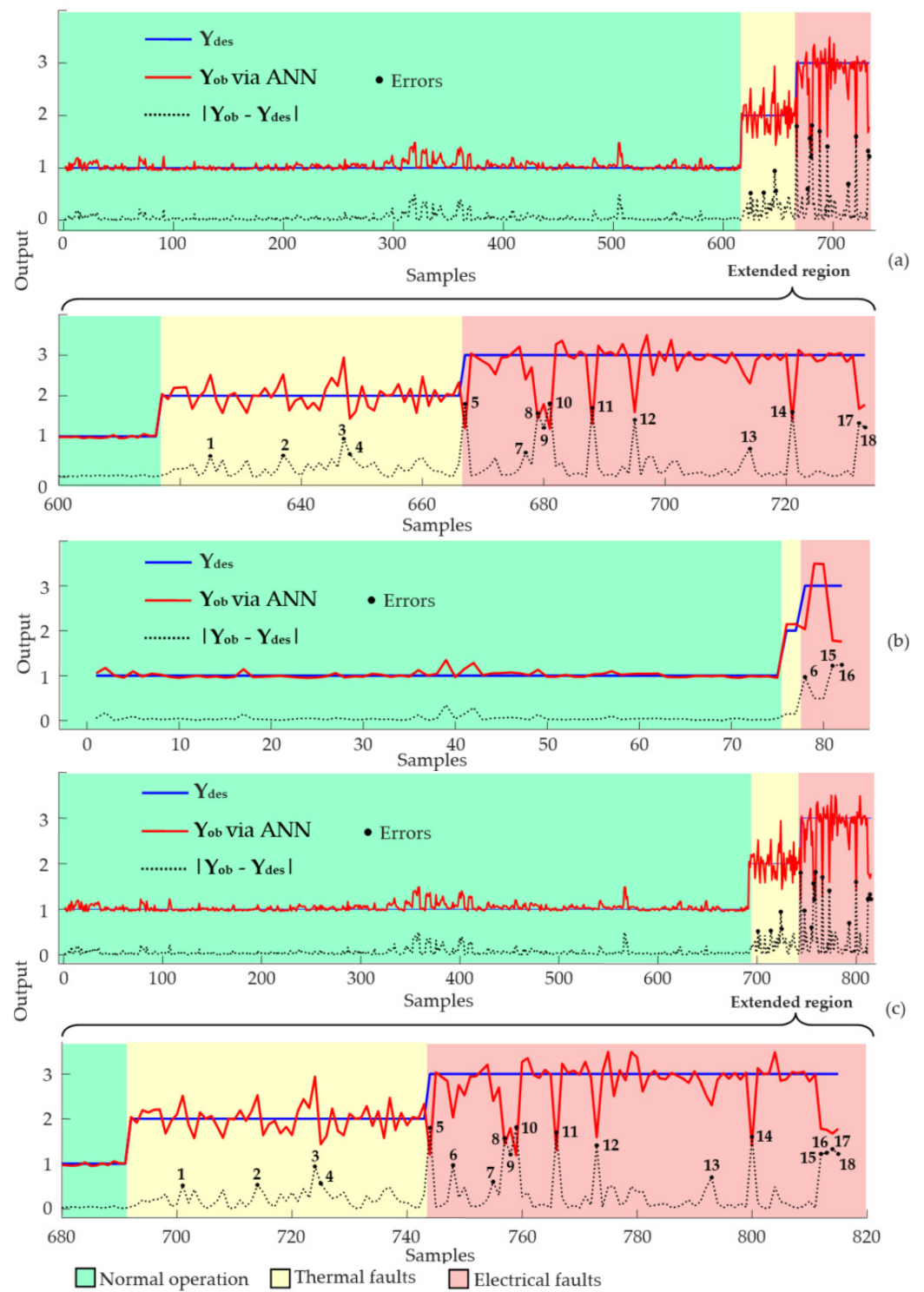


Figure 8. MSE performance for PRN with 5 neurons in hidden layer, (a) training phase with 90% of samples (733), (b) validation phase with 10% of samples (82), (c) 100% of samples (815).

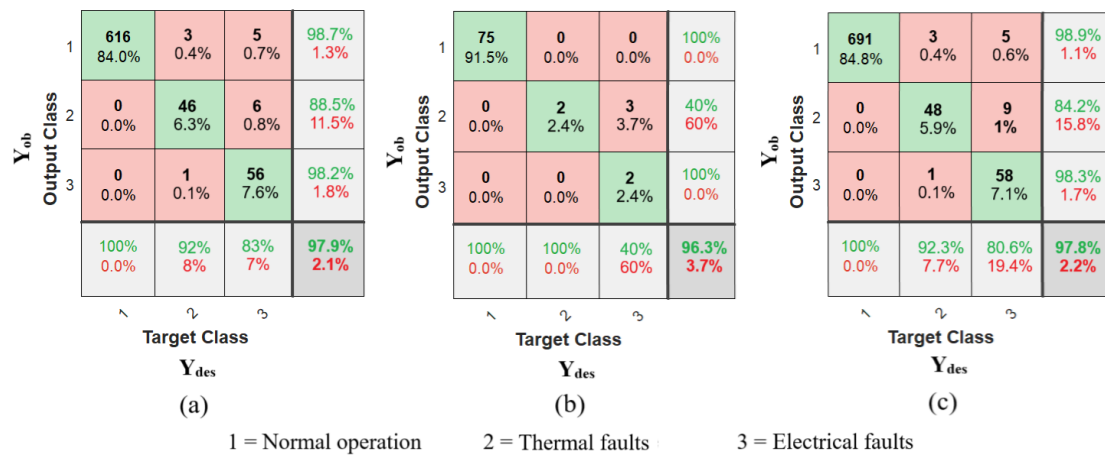


Figure 9. Confusion matrix for classifying the three classes (1—normal operating condition, 2—thermal faults, 3—electrical faults) obtained by PRN with 5 neurons in the hidden layer, (a) training phase with 90% of samples (733), (b) validation phase with 10% of samples (82), (c) 100% of samples (815).

Figure 10 presents results for the PRN network with 10 neurons in the hidden layer. Better outcomes were achieved for both training and validation, as shown in Figure 10a,b. During network training, there were a total of 16 errors, resulting in an accuracy rate of 97.8%, as illustrated in the confusion matrix presented in Figure 11a. During the model validation, the ANN (PRN) classified with 100% accuracy the samples that were not included in the training. Of the 82 classified samples, all were correctly assigned to their corresponding outputs (75 samples for normal operation—class 1, 4 samples for thermal failure—class 2, and 3 samples for electrical failure—class 3), as illustrated in Figure 10b and in the confusion matrix presented in Figure 11b.

The results regarding the training and validation of the PRN network applied to 100% of the samples are shown in Figure 10c, as well as in the confusion matrix presented in Figure 11c. It is observed that, when using 10 neurons in the hidden layer, the network recorded only 14 errors, occurring exclusively in the training phase, resulting in a success rate of 97.8%. Overall, the accuracy rate achieved was 98%, slightly surpassing the 97.8% achieved with five neurons in the hidden layer.

Both configurations proved to be effective as automatic models for classifying the operating conditions of power transformers based on the gases present inside them.

Tables 3–5 present the weights of the connections between the input and hidden layers, from the hidden layer to the output, and the respective bias weights of the hidden and output layers for the network with 10 neurons in the hidden layer.

Table 3. Weights of the connections between the neurons of the input layer and the hidden layer (W_{Rm}).

		Neurons Hidden Layer									
		m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}
Neurons input layer	R_1	0.0322	−0.3067	0.5928	1.3855	−0.4574	−1.2926	−0.9642	0.6219	5.0458	1.1364
	R_2	−1.2657	0.5935	0.3621	−0.3653	0.1491	1.3524	0.4590	0.9758	1.1697	−0.8334
	R_3	1.4485	−0.4754	3.0243	−1.4813	−1.6088	−0.6485	1.2424	1.9435	0.7289	−1.2262
	R_4	0.2304	1.2848	0.1588	0.7570	−1.5067	3.0534	1.4099	0.3570	1.6211	−0.2550
	R_5	1.0501	−1.5996	−1.4915	0.8923	0.2894	−0.2318	−0.0893	−0.5393	−0.6183	1.2611

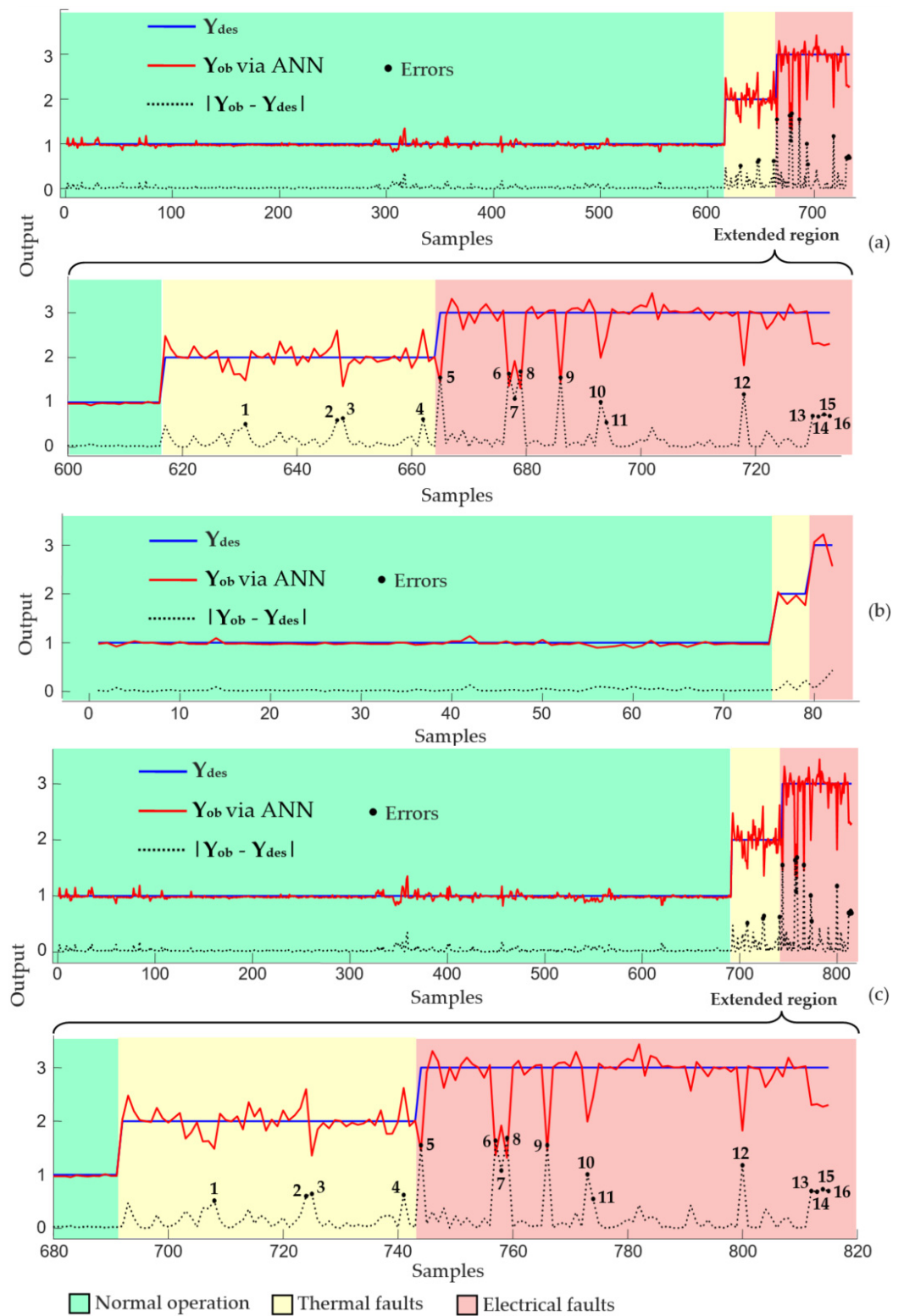


Figure 10. MSE performance for PRN with 10 neurons in hidden layer, (a) training phase with 90% of samples (733), (b) validation phase with 10% of samples (82), (c) 100% of samples (815).



Figure 11. Confusion matrix for classifying the three classes (1—normal operating condition, 2—thermal faults, 3—electrical faults) obtained by PRN with 10 neurons in the hidden layer, (a) training phase with 90% of samples (733), (b) validation phase with 10% of samples (82), (c) 100% of samples (815).

Table 4. Weights of the connections between the neurons of the hidden layer and the output layer (W_{mi}).

		Neurons Output Layer		
		i_1	i_2	i_3
Neurons hidden layer	m_1	−0.1062	0.8985	0.1317
	m_2	−0.1235	−0.3878	−0.8237
	m_3	−2.2868	0.5225	2.9425
	m_4	0.2307	1.4119	−0.8567
	m_5	−0.3885	0.2586	−0.9823
	m_6	−2.1606	2.8833	−1.0345
	m_7	−0.7728	0.0791	0.2054
	m_8	−0.5972	−0.3729	−0.3257
	m_9	−8.1766	3.4204	3.5715
	m_{10}	−0.5427	0.4388	−1.2068

Table 5. Bias of the neurons of each layer (hidden and output).

Neurons Hidden Layer (m_x1)		Neurons Output Layer (i_x1)	
1	−2.2544	1	−0.6331
2	1.8252		
3	2.4682		
4	−0.2260		
5	0.3246	2	0.9741
6	1.8809		
7	−0.6293		
8	1.0534		
9	7.8957	3	0.7542
10	2.1513		

5. Conclusions

This study presented a methodology via PRN artificial neural networks to obtain the operating conditions of power transformers (normal, thermal faults, or electrical faults) as a function of the combustible gases (H_2 (hydrogen), CH_4 (methane), C_2H_2 (acetylene), C_2H_4 (ethylene), and C_2H_6 (ethane)) presented therein.

In the two configurations presented (5 and 10 neurons in the hidden layer), the network had good training, a little better when using 10 neurons, with 10 s of processing and **MSE** at 0.0175. In the validation phase (data that was not part of the training), the **MSE** was 0.0057, and a correlation between the obtained output and the desired output was 0.9718, showing the effectiveness of the model.

In the best results, the network was able to classify the samples for both training and validation with low error. The network presented only 14 errors out of 815 samples, all in the training phase. In the validation phase, the network presented 100% accuracy. In total, for 815 samples in both phases (training and validation), the hit rate was 98%.

As a result, the methodology presented via ANN (PRN) proved to be efficient in classifying the operating conditions of power transformers (normal—class 1, thermal faults—class 2, and electrical faults—class 3) depending on the combustible gases created inside of the same (H_2 , CH_4 , C_2H_2 , C_2H_4 , and C_2H_6).

Energy utilities can improve their annual predictive and preventive maintenance planning process by implementing this proposed method. In Brazil, dissolved gas analysis (DGA) tests are often carried out annually or every six months. However, depending on the severity of the DGA results, the interval between these tests can be reduced, which can promote a change in maintenance planning, preventing further damage to the equipment and ensuring continuity of service.

Author Contributions: A.G. and A.B.N.: conceptualization; A.G. and A.B.N.: writing—original draft; A.N.d.S., R.P.d.M., M.A.I., E.G. and F.T.N.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by CAPES and CNPq under grant number 88887.704285/2022-00.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the financial support from São Paulo State University—UNESP—Prope.

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

Nomenclature

ANN	Artificial neural network
PRN	Pattern recognition artificial neural network
H_2	Hydrogen
CH_4	Methane
C_2H_2	Acetylene
C_2H_4	Ethylene
C_2H_6	Ethane
SCG	Scaled conjugate gradient
DGA	Dissolved gas analysis
PT	Power transformers
Y_{ob}	Obtained outputs
Y_{des}	Desired outputs

References

1. Bustamante, S.; Manana, M.; Arroyo, A.; Castro, P.; Laso, A.; Martinez, R. Dissolved Gas Analysis Equipment for Online Monitoring of Transformer Oil: A Review. *Sensors* **2019**, *19*, 4057. [CrossRef] [PubMed]
2. Lima, S.L.; Saavedra, O.R.; Miranda, V. A two-level framework to fault diagnosis and decision making for power transformers. *IEEE Trans. Power Deliv.* **2015**, *30*, 497–504. [CrossRef]
3. Cheng, L.; Yu, T. Dissolved Gas Analysis Principle-Based Intelligent Approaches to Fault Diagnosis and Decision Making for Large Oil-Immersed Power Transformers: A Survey. *Energies* **2018**, *11*, 913. [CrossRef]
4. PC57.104/D6.2; IEEE Approved Draft Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers. IEEE: New York, NY, USA, 2019. Available online: <https://ieeexplore.ieee.org/document/8666950/metrics> (accessed on 11 January 2024).
5. Yang, M.T.; Hu, L.S. Intelligent fault types diagnostic system for dissolved gas analysis of oil-immersed power transformer. *IEEE Trans. Dielectr. Electr. Insul.* **2013**, *20*, 2317–2324. [CrossRef]
6. IEC 60599; Mineral Oil-Filled Electrical Equipment in Service—Guidance on the Interpretation of Dissolved and Free Gases Analysis (3rd ed.). International Electrotechnical Commission: Geneva, Switzerland, 2015.
7. IEEE C57.104-2019; IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers. IEEE: New York, NY, USA, 2019.
8. Do, T.D.; Tuyet-Doan, V.-N.; Cho, Y.-S.; Sun, J.-H.; Kim, Y.-H. Convolutional-Neural-Network-Based Partial Discharge Diagnosis for Power Transformer Using UHF Sensor. *IEEE Access* **2020**, *8*, 207377–207388. [CrossRef]
9. Zollanvari, A.; Kunanbayev, K.; Bitaghsir, S.A.; Bagheri, M. Transformer Fault Prognosis Using Deep Recurrent Neural Network Over Vibration Signals. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2502011. [CrossRef]
10. Rokani, V.; Kaminaris, S.D.; Karaisas, P.; Kaminaris, D. Power Transformer Fault Diagnosis Using Neural Network Optimization Techniques. *Mathematics* **2023**, *11*, 4693. [CrossRef]
11. Hendel, M.; Meghnefi, F.; Senoussaoui, M.E.A.; Fofana, I.; Brahami, M. Using Generic Direct M-SVM Model Improved by Kohonen Map and Dempster–Shafer Theory to Enhance Power Transformers Diagnostic. *Sustainability* **2023**, *15*, 15453. [CrossRef]
12. Xing, Z.; He, Y. Multimodal Mutual Neural Network for Health Assessment of Power Transformer. *IEEE Syst. J.* **2023**, *17*, 2664–2673. [CrossRef]
13. Odinaev, I.; Pazderin, A.; Safaraliev, M.; Kamalov, F.; Senyuk, M.; Gubin, P.Y. Detection of Current Transformer Saturation Based on Machine Learning. *Mathematics* **2024**, *12*, 389. [CrossRef]
14. Beura, C.P.; Wolters, J.; Tenbohlen, S. Application of Pathfinding Algorithms in Partial Discharge Localization in Power Transformers. *Sensors* **2024**, *24*, 685. [CrossRef] [PubMed]
15. Bonini Neto, A.; Alves, D.A.; Minussi, C.R. Artificial Neural Networks: Multilayer Perceptron and Radial Basis to Obtain Post-Contingency Loading Margin in Electrical Power Systems. *Energies* **2022**, *15*, 7939. [CrossRef]
16. Cichoń, A.; Włodarz, M. OLTC Fault detection Based on Acoustic Emission and Supported by Machine Learning. *Energies* **2024**, *17*, 220. [CrossRef]
17. Das, S.; Paramane, A.; Chatterjee, S.; Rao, U.M. Sensing Incipient Faults in Power Transformers Using Bi-Directional Long Short-Term Memory Network. *IEEE Sens. Lett.* **2023**, *7*, 7000304. [CrossRef]
18. Rana, K.; Kishor, N.; Negi, R.; Biswal, M. Fault Detection and VSC-HVDC Network Dynamics Analysis for the Faults in Its Host AC Networks. *Appl. Sci.* **2024**, *14*, 2378. [CrossRef]
19. Thango, B.A. On the Application of Artificial Neural Network for Classification of Incipient Faults in Dissolved Gas Analysis of Power Transformers. *Mach. Learn. Knowl. Extr.* **2022**, *4*, 839–851. [CrossRef]
20. C57.110/2008; IEEE Recommended Practice for Establishing Liquid-Immersed and Dry-Type Power and Distribution Transformer Capability When Supplying Nonsinusoidal Load Currents. IEEE: New York, NY, USA, 2018.
21. Eqbal, D.; Khan, S.A.; Islam, T. Transformer incipient fault diagnosis on the basis of energy-weighted DGA using an artificial neural network. *Turk. J. Electr. Eng. Comput. Sci.* **2018**, *26*, 77–88. [CrossRef]
22. Ghoneim, S.; Ward, S.A. Dissolved Gas Analysis as a Diagnostic Tools for Early Detection of Transformer Faults. *Adv. Electr. Eng. Syst.* **2012**, *1*, 152.
23. Filho, G.L. Comparison between Diagnostic Criteria by Chromatographic Analysis of Gases Dissolved in Power Transformer Insulating Oil. Master’s Dissertation, São Carlos School of Engineering, University of São Paulo, São Carlos, Brazil, 2012.
24. Li, E. Dissolved gas data in transformer oil—Fault Diagnosis of Power Transformers with Membership Degree. *IEEE Access* **2019**, *7*, 28791–28798. [CrossRef]
25. Malik, H.; Mishra, S. Extreme Learning Machine Based Fault Diagnosis of Power Transformer Using IEC TC10 and Its Related Data. In Proceedings of the 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 17–20 December 2016.
26. Mathworks. Available online: <http://www.mathworks.com> (accessed on 20 January 2024).
27. De Souza, A.V.; Bonini Neto, A.; Piazentin, J.C.; Dainese, B.J., Jr.; Gomes, E.P.; Bonini, C.S.B.; Putti, F.F. Artificial neural network modelling in the prediction of bananas’ harvest. *Sci. Hortic.* **2019**, *257*, 108724. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.