Review

Artificial Neural Network Application for Partial Discharge Recognition: Survey and Future Directions

Abdullahi Abubakar Mas'ud [1,*], Ricardo Albarracín [2], Jorge Alfredo Ardila-Rey [3], Firdaus Muhammad-Sukki [4], Hazlee Azil Illias [5], Nurul Aini Bani [6] and Abu Bakar Munir [7,8]

[1] Department of Electrical and Electronics Engineering, Jubail Industrial College, Jubail 10099, Saudi Arabia
[2] Department of Electrical, Electronics and Automation Engineering and Applied Physics,
     Universidad Politécnica de Madrid, Ronda de Valencia 3, Madrid 28012, Spain; rasbarracin@gmail.com
[3] Department of Electrical Engineering, Federico Santa María Technical University,
     Santiago de Chile 8940000, Chile; jorge.ardila@usm.cl
[4] School of Engineering, Faculty of Design and Technology, Robert Gordon University,
     Aberdeen AB10 7GJ, UK; f.b.muhammad-sukki@rgu.ac.uk
[5] Department of Electrical Engineering, Faculty of Engineering, University of Malaya,
     Kuala Lumpur 50603, Malaysia; h.illias@um.edu.my
[6] UTM Razak School of Engineering and Advanced Technology, Universiti Teknologi Malaysia,
     Kuala Lumpur 54100, Malaysia; nurulaini.kl@utm.my
[7] Faculty of Law, University of Malaya, Kuala Lumpur 50603, Malaysia; abmunir@um.edu.my
[8] University of Malaya Malaysian Centre of Regulatory Studies (UMCoRS), University of Malaya,
     Jalan Pantai Baru, Kuala Lumpur 59990, Malaysia
* Correspondence: masud_a@jic.edu.sa; Tel.: +966-538138814

Abstract: In order to investigate how artificial neural networks (ANNs) have been applied for partial discharge (PD) pattern recognition, this paper reviews recent progress made on ANN development for PD classification by a literature survey. Contributions from several authors have been presented and discussed. High recognition rate has been recorded for several PD faults, but there are still many factors that hinder correct recognition of PD by the ANN, such as high-amplitude noise or wide spectral content typical from industrial environments, trial and error approaches in determining an optimum ANN, multiple PD sources acting simultaneously, lack of comprehensive and up to date databank of PD faults, and the appropriate selection of the characteristics that allow a correct recognition of the type of source which are currently being addressed by researchers. Several suggestions for improvement are proposed by the authors include: (1) determining the optimum weights in training the ANN; (2) using PD data captured over long stressing period in training the ANN; (3) ANN recognizing different PD degradation levels; (4) using the same resolution sizes of the PD patterns when training and testing the ANN with different PD dataset; (5) understanding the characteristics of multiple concurrent PD faults and effectively recognizing them; and (6) developing techniques in order to shorten the training time for the ANN as applied for PD recognition Finally, this paper critically assesses the suitability of ANNs for both online and offline PD detections outlining the advantages to the practitioners in the field. It is possible for the ANNs to determine the stage of degradation of the PD, thereby giving an indication of the seriousness of the fault.

1. Introduction

Over the years, partial discharge (PD) recognition has been a topic of interest for a number of reasons, in particular the need to distinguish between different PD fault sources within the insulation systems of power apparatus and discriminate them from extraneous interference events considered as noise [1–5]. PDs are the electrical discharges that occur within or outside the insulation of a high-voltage (HV) system under electric stress [6,7]. It is essential to recognize these faults at an early stage before they lead to disastrous conditions of the equipment with serious financial and safety implications. Therefore, developing techniques to characterize and classify PD has become of profound importance to condition monitoring (CM) engineers [8]. The nature, form and characteristics of PD have been widely investigated and, in many ways, established [9–11]. Despite that, a step forward must be given to determine novel techniques that can effectively classify PD patterns and give reliable assessment on the nature of the PD faults. To carry out the pattern recognition task, four main techniques have been recognized [12]. They are the template matching, statistical approach, syntactic approach and the intelligence systems:

(1) In template matching, a sample of the patterns to be recognized is readily available and correlated with a stored template. Examples of this technique are the distance classifiers, e.g., the minimum distance classifier [13].

(2) In statistical approach, each pattern is characterized by some measured features and represented as a point in multi-dimensional space [14]. The objective of this second technique is to choose those features that allow pattern fingerprints belonging to various categories to occupy separate regions in a multi-dimensional feature space.

(3) The syntactic approach is another technique for recognizing complex patterns. In this case, hierarchical observation is adopted where a pattern is regarded as being composed of subpatterns, which are individually less complex sub patterns [15]. The main complex pattern is a function of interrelationships between these smaller sub patterns.

(4) The intelligence system techniques one example is the artificial neural network (ANN).

Based on the aforementioned pattern recognition techniques, distance classifiers, statistical classifiers and artificial intelligence classifiers have been applied to recognize PD. Some examples of the distance classifiers which have been applied are the minimum distance classifier [13] and nearest neighbour classifiers [16]. Statistical classifiers employed are the Bayes classifiers [17] and the recognition rate classifiers [18], while the intelligence classifiers include ANNs [7,10,19–21], fuzzy logic controllers [17], hidden Markov models [22,23], support vector machines [24,25], genetic algorithms [26] and data mining techniques [27]. Among them, the ANN is one of the most successful pattern recognition techniques because of its capability to learn input-output relationships from a few examples. Several ANN techniques applied for PD pattern recognition include the feed-forward neural network using the back-propagation (BP) [10,16,28], the Kohonen self-organizing map (KOH) and learning vector quantization (LVQ) [16,29], adaptive resonance theory [30], counter propagation neural network [31], probabilistic neural network (PNN) [18,32], cellular neural network [33], modular neural network (MNN) [34,35], extension neural network [36], fuzzy neural networks [37], and most recently the ensemble neural network (ENN) [7]. These techniques yield encouraging results, with recognition rates reaching as high as 90% in some instances, when testing was done with unknown PD fingerprints. When applied for PD pattern recognition, the template matching approach (e.g., minimum distance classifier) and the intelligent technique (e.g., ANN) recorded up to 100% recognition rate for some PD fault examples [10,38]. However, the statistical approach (e.g., principal component analysis) is commonly applied as feature extraction technique of PD data in order to determine the most suitable parameters for classification. The Syntactic approach has never been applied to classify PD fingerprints.

In this paper, a comprehensive survey on the performance of several ANN models applied for PD recognition has been carried out, together with their strengths and limitations. ANN is chosen for this study because of its wider adoption for PD recognition by several researchers. The main advantage of the ANN over all other techniques is its ability to learn complex nonlinear input-output relationships and apply sequential training procedures in order to adapt themselves to the data to be recognized. Suggestions for improvement are made and the impact of the ANN research on real time PD location and recognition is critically analysed.

Section 1 is the introduction, while Section 2 evaluates the impact of ANN research on practical PD recognition. Section 3 describes the BP algorithm topologies. In Section 4, the ANN models applied for PD recognition are discussed. Section 5 presents the previous research works on ANNs for PD recognition. Afterwards, Section 6 provides some discussions and strength of the ANN when applied for PD recognition. Section 7 presents the limitations and suggestions for future improvement. Finally, the conclusions are presented in Section 8.

2. Impact of the Artificial Neural Networks Research on Practical Partial Discharge Recognition

CM has become a vital technique in HV equipment maintenance and is increasingly attracting attention globally [39,40]. The need to minimize fault alarms and allowing planned maintenance bring a series of advantages to the power industry. There are needs for lower maintenance cost, reducing the severity of damages, minimizing accidents and ensuring the safety of personnel. Due to these challenges, the development of new techniques for identifying PD sources has become the main challenge of many experts interested in improving the procedures currently used in condition-based maintenance (CBM).

This paper proposed that the ANN could be a potential tool for future CM equipped with improved sensitivity, reliability, intelligence, and cost savings. The question now is how the ANN can be applied to improve CM and its assessment. It is obvious that the ANN on its own cannot perform all the CM functions but it can be used in conjunction with the existing techniques to provide a robust CM tool. Due to many advantages of the online CM, the capability of the ANN to be applied to online monitoring can be examined. Reviewing the literature [41] shows that online CM systems (e.g., PD measurement) have four main parts, i.e., sensors, data acquisition, fault detection and diagnosis. The sensors usually detect the fault and convert a physical quantity to an electrical signal. Then, data acquisition systems process this information from the sensors usually using microcomputers. Finally, the fault detection and diagnosis systems determine the nature of the fault and clear indication for maintenance.

Current fault detection techniques involve the application of frequency and time-domain signal processing techniques to obtain signatures for a fault or normal condition. At this moment, fault diagnosis is being carried out by experts with the aid of computers and advanced techniques, such as in the one demonstrated by Álvarez et al. [42]. The ANN can be very attractive for both online and offline fault detections and diagnoses and can reduce the reliance on experts for fault interpretation, thereby reducing cost and visual implementation work. The ANN can be trained offline with all possible fault data. If sufficient data is not available in terms of scope (i.e., data must be available from wide range of operating conditions ideally including fault, unusual and undesirable conditions), then the developed ANN may not possess the adequate accuracy or functionality for intended applications. Therefore, the data used for training the ANN can come from the actual HV plant either in service or from factory test. If such data is not available then the possibility of using simulated data from laboratory experiment may be investigated. After sufficient training, the ANN can now act as an experienced evaluator. The moment the fault data is fed into the developed ANN, it can simply indicate the fault within seconds. Through training and testing with the known fault, the ANN can also track the degradation level and indicate the urgency for fault correction. However, ANN has some limitations. These include excessive training involved and lack of sufficient PD fault examples currently occurring in the field and in some cases one fault may lead to another fault. Multiple concurrent faults and the

problems associated with the presence of noise sources of different nature may also hinder the process of identification.

A diagram showing the overall structure of the proposed CM technique encompassing the ANN with post-processing elements is shown in Figure 1. It is obvious that the ANN can be applied in two stages. First, the online ANN for detecting the fault and second is the offline ANN for tracking the level of degradation of the insulation thereby offering significant potential for improving plant CM functions. The PD degradation assessment will be done offline because of repeated training and testing schemes of the ANN involved. The ANN will have significant influence on the overall maintenance cost plus reliability, by greatly reducing the time and increasing the accuracy of fault diagnosis.
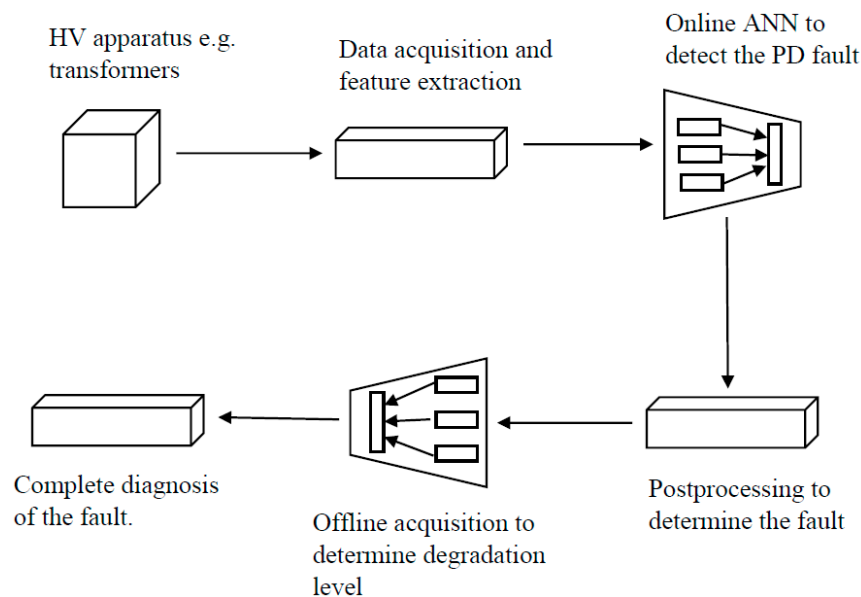
Figure 1. A proposed online/offline partial discharge (PD) based condition monitoring (CM) using the artificial neural network (ANN).

3. The Back Propagation Algorithm

The BP algorithm is a form of supervised learning for the feed-forward ANN [10,43] and it occurs in two steps, namely forward and backward learning. It learns through recurring presentation of the input and output examples and each time back-propagating the error and updating the weights and biases until the error is minimized to the desired value [7].

Figure 2 shows how the BP algorithm is implemented together with the neurons (also known as the processing elements). The inputs $(X_{P1}, \ldots , X_{PNo})$ are initially propagated through the network and the output is computed. Then, the error at the output $(T_{P1}, \ldots , T_{PNM})$ is back-propagated through the network and the weights are updated according to gradient descent algorithm [34]. This process continues until the mean square value at the output reaches the minimum acceptable value. Some of the major drawbacks of the BP are longer convergence time and susceptibility to training failure [44]. One of the improvements made to address these problems includes adding a momentum term for faster training but at the cost of extra memory space [22,44]. Despite all these issues, the BP has been widely applied for PD recognition because of its easier implementation and ability to provide better PD recognition result as compared to other ANN algorithms. Gulski and Krivda [10] applied different ANN algorithms for PD recognition and the results show that BP provides better recognition result.
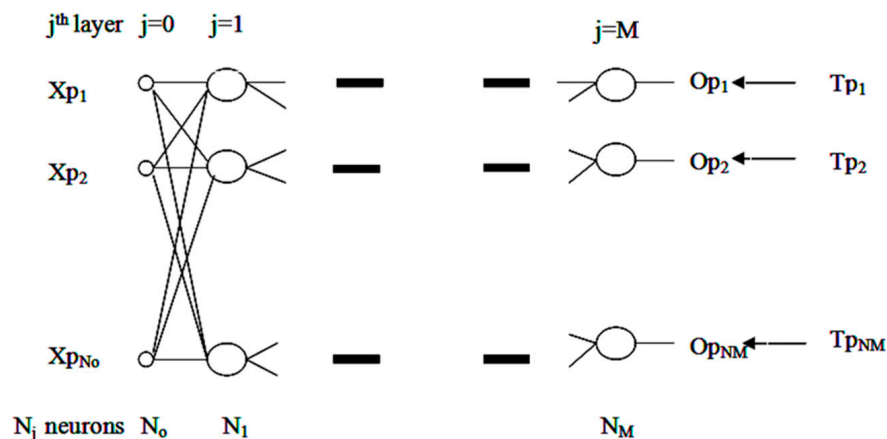
Figure 2. The back propagation (BP) algorithm, adapted from [7].

4. Artificial Neural Network Models Applied for Partial Discharge Pattern Recognition

4.1. Modular Neural Network

Several implementations of the MNN exist. The feature decomposition based on MNN was employed for PD recognition [34,35]. In this case, bulk-training fingerprints are divided into several subsets, with each subset comprising values of a specific parameter, as illustrated in Figure 3. Each ANN can be trained independently using a subset of the data by the BP algorithm. To determine the output of the modular network, majority-voting technique is employed, in order to combine the output of these constituent ANNs and get the final decision. The MNN therefore recognizes a particular input parameter belonging to a particular group if the majority of the sub networks assign this input to this particular group [35].
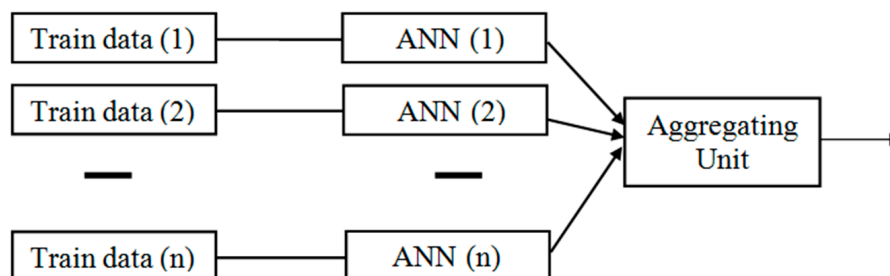

Figure 3. The modular neural network (MNN), adapted from [34].

4.2. The Ensemble Neural Network

An ENN, as presented in Figure 4, is a method of training several BP ANN topologies and combining their component predictions [45]. The inspiration for this method lies on the fact that by combining the component ANN predictions, it is expected that there would be a considerable improvement on the generalization performance of the ANN. The literature [46] proves that this is only possible if the constituent neural networks forming the ensemble are concurrently diverse and accurate. Several techniques evolved for training the ENN [45], but bagging (bootstrapping) is seen to be the most effective. In bootstrapping, a number of training fingerprints are generated by bootstrap resampling of the original fingerprint. Several training samples are repeated while others are simply ignored. The bootstrapping prevents over fitting associated to NNs and provides correct values of the bias and variance [43].
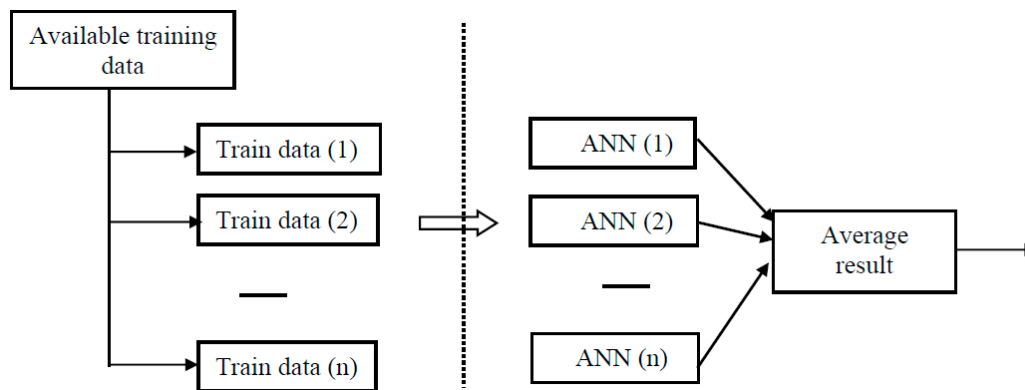
Figure 4. The ensemble neural network (ENN) topology, adapted from [46].

4.3. The Probabilistic Neural Network

The PNN is a technique based on competitive learning procedure based on the Parzen window concept of multivariate probability approximation [18] (Figure 5). The PNN obtains the probability density function (PDF) based on Bayes' decision making approach [18,28]. The PNN is made up of input layer, hidden layer and output layer [27]. The hidden layer consists of the exemplar and summation layers. Input parameters are fed into the network through the input layers. The exemplar layer is made up of Gaussian functions formed using a specified set of data points representing the centres [47]. The summation or class layer performs the summing operation of the outputs coming from the second layer for each class [47]. The decision layer then performs the voting—choosing the highest value [28]. Then, the related class label is obtained.
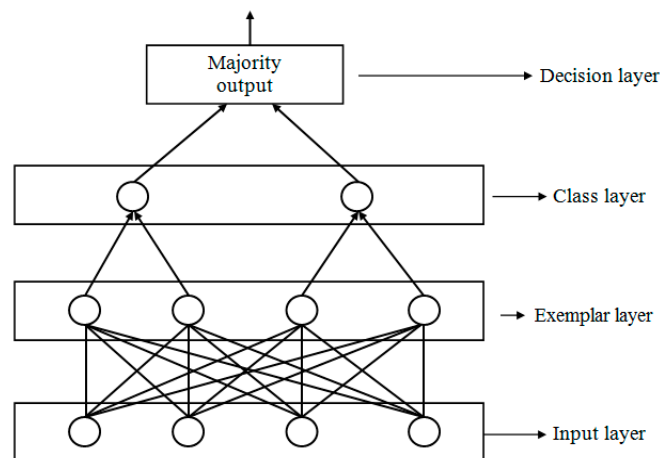
Figure 5. The probabilistic neural network (PNN) topology, adapted from [13].

4.4. The Radial Basis Function Network

The radial basis function network (RBFN) is another ANN model mostly applied to solve interpolation problems and consists of two layers [43], as shown in Figure 6. The neurons in the first layer do not give the weighted sum of inputs through the sigmoid function. The middle layer consists of the basis functions (φ_i), mostly made up of Gaussian functions. The centre of the basis function and the network input give the output of the first layer neurons. When the input moves away from a given centre, the neurons output drops off quickly to zero. The second layer of the RBFN network possesses receptive fields because they only respond to the inputs that are closer to their centres [43]. The RBFN provides quicker training and has unsupervised learning characteristics compared to the feed-forward network, but requires many neurons for high-dimensional input spaces.
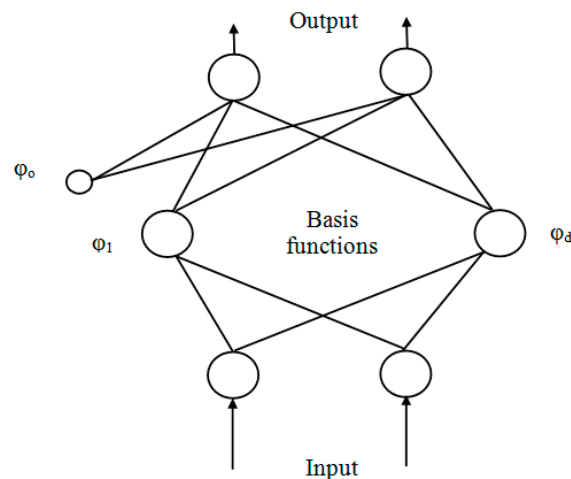
Figure 6. The radial basis function network (RBFN) topology.

5. Relevant Previous Research Works on Artificial Neural Network for Partial Discharge Recognition

Previous research has been undertaken by many authors on the application of NNs for PD pattern recognition. Research on the application of NNs for PD pattern recognition seems to have started in the early nineties. Because of the advantages of the phase-amplitude-number (φ-q-n) patterns (e.g., its visible discriminating capability) [48] in evaluating PD defects, earlier research started by extracting useful information from these distributions. Since different types of PD faults generate different φ-q-n patterns (see example, Figure 7), the ANN was able to discriminate these faults even with slight pattern variations. The initial stage in pattern recognition was the choice of appropriate fingerprints that can be applied as training and testing parameters for the ANN.
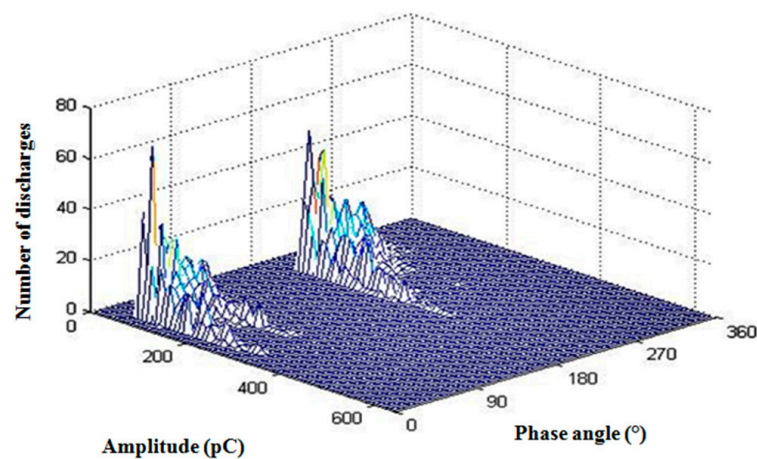

Figure 7. A typical φ-q-n pattern.

Earlier research work by Suzuki and Endoh [49] showed how the φ-q-n patterns from a needle-type defect in cross-linked polyethylene (XLPE) cable are transformed into smaller patterns by reducing the number of pixels, thereby minimizing the number of amplitude and phase resolutions. This is to ensure reduction of the input data to the ANN. A pixel corresponds to a specific phase angle range and a specific discharge magnitude. The paper applied the BP algorithm and the results showed that the correct response reaches 100% detection probability and converges rapidly for the smaller distributions as compared to larger distributions. This result clearly indicates that smaller numbers of pixels in the φ-q-n distributions are better PD recognition parameters for the ANN.

The technique of choosing learning fingerprints by Suzuki and Endoh [49] was adopted by Hozumi et al. [50] and Phung et al. [51]. They also applied the BP algorithm and the result shows that the ANN learns and updates faster with high recognition rate above 90%. Gulski and Krivda [10] evaluated the performance of different ANN algorithms, though they used a different approach in determining the input pattern when compared to that of Suzuki and Endoh [49].

Gulski and Krivda [10] studied the application of three types of ANN algorithms for classifying two-electrode PD models. These are models of artificial defects of industrial objects in 400 kV gas insulated substation (GIS) compartments. The work derived the $H_n(\varphi)+$, $H_n(\varphi)-$, $H_{qn}(\varphi)+$, $H_{qn}(\varphi)-$, $H_n(q)+$ and $H_n(q)-$ plots during 20 min of testing at 20% above PD inception voltage and the patterns were evaluated using 15 sets of statistical fingerprints. These included the Skewness (*sk*) and Kurtosis (*ku*) of the positive and negative half cycles of the $H_n(\varphi)$ and $H_{qn}(\varphi)$ histograms, as well as the cross-correlation (*cc*), discharge factor (*Q*) and the number of peaks. Definition of these statistical parameters can be found from the literature [10]. These statistical tools form the bulk of training and testing data for ANN models and encouraging performance, up to 100% rate was recorded for trained PD fingerprints. Recognition efficiency of 100% was obtained for the BP as compared to others, which had efficiency of 70%. Despite the success of this scheme, a number of PD fault misclassifications were recorded. For each algorithm, approximately 8 out of 12 PD defects were misclassified as belonging to others, but Gulski and Krivda [10] did not come up with logical conclusion regarding this observation.

Along with Gulski and Krivda [10], further literature has adopted the use of statistical fingerprints derived from *φ-q-n* patterns. For example, Candela et al. [52] developed a PD recognition system where statistical Weibull analysis was applied to the 3-D for feature extraction. The paper considered three artificial PD geometries, i.e., dielectric surface discharges in air, a metallic dielectric parallel air gap and a dielectric bounded spherical cavity. The *sk*, *ku*, α and β values formed the input parameters to the ANN. Based on the application of these parameters, success rates up to 98% were recorded. Further, Mirelli and Schifani [29] evaluated the application of statistical fingerprints as input variables to the ANN. They evaluated two separate parameters from each of the three PD patterns, i.e., $H_n(q)$, $H_n(\varphi)$ and $H_{qn}(\varphi)$. From the $H_n(q)$ histograms, α and β were determined, while from the $H_n(\varphi)$, *sk* and *ku* were evaluated. The lacunarity (measure of denseness of the fracture surface) and dimensionality (quantification of the surface roughness of the *φ-q-n* plot) are other factors derived from the *φ-q-n* plots. These six parameters were applied as inputs for training the ANN using the BP algorithm. The discriminating capability of the system was tested with high recognition rate of 92% in 20 kV insulators. In another work, Karthikeyan et al. [28] investigated the effectiveness of the BP algorithm for recognition of PD defects in voids, corona and surface discharges, using various statistical measures in order to obtain the fingerprints for the ANN. Some statistical parameters considered as inputs to the ANN include: (1) maximum and minimum values of a specific parameter; (2) measures of dispersion i.e., range, mean deviation, quartile deviation; and (3) measures of the central tendency. Their results show how the BP algorithm was able to show a good recognition rate up to 100% for some testing examples, though a number of misclassifications were recorded The paper concludes that the BP algorithm does not possess the ability to be utilized for online training because of the excessive training time needed to obtain the MSE at the output of the ANN. The learning rate also plays an important role in the convergence of the system as high values of learning rates yield less training time with high convergence error. However, with corrective measures as proposed by the author of this paper in the preceding section these issues can be addressed.

Other investigations employ a large matrix of parameters derived from the phased-resolved patterns as training and testing fingerprints, comprising the number of discharges, the amplitude and the phase angle. These have demonstrated good classification potential even with untrained data. For example, Badent et al. [53] developed a novel PD diagnosis system using artificial ANNs. The training data consisted of a (125 × 125) matrix derived from the phase resolved patterns, where the row index is the apparent charge and the column index is the phase resolution. The value in each index corresponds to the number of discharges. The system recorded excellent performance with

100% recognition rate for trained patterns and approximately 90% for new patterns. Trained patterns represent patterns already used for training while new patterns are patterns not used for training.

Hong et al. [34] in their paper investigated the application of feature decomposition based MNN for classifying PD patterns. The training data is comprised of fingerprints derived from the φ-q-n PD patterns. These included pulse count, average PD magnitude and maximum PD magnitude. The bulk training parameters were partitioned into three subsets with each subset comprising values of a specific parameter. Three ANNs were independently trained by the BP algorithm by using a subset of the data. To determine the output of the MNN, majority-voting technique combined the output of this ANN to give the final decision. The MNN learns faster than the single network and has been shown to perform better especially when discriminating unknown data.

Apart from the work of Badent et al. [53], Yamazaki et al. [54] investigated the application of ANNs to categorize PD patterns from voids with or without ultraviolet radiation. Partial discharge inception voltages (PDIV) were used as the training fingerprints. The performances of the ANNs were evaluated based on their mean square error (MSE). The network with the highest number of layers has the least MSE.

Mazroua et al. [16] did not adopt the usual feature extraction technique using φ-q-n PD patterns but rather used pulse train patterns for classifying PD from electrical trees and voids. Some specific parameters chosen included peak amplitude, rise time, fall time, duration and the area covered by the PD pulse. The ANN was also able to recognize some changes as a result of ageing. The ANN was only able to give good recognition on single discharge sources but the paper considered future work into designing ANNs for multiple defects.

In another work, Tian et al. [44] also applied Fourier transform for PD identification but with spectral density of acoustic signals for choosing the learning fingerprints and so did not utilize the φ-q-n patterns as training sets for the ANN. This technique reduces the number of input neurons and training data. The ANN algorithms also showed satisfactory results of up to 90% success rate.

Because of the misclassification problem inherent to the BP algorithm, Hoof et al. [55] developed a new ANN classification technique using the guarded neural classifier (GNC) that performs better than the usual BP algorithm when both are applied to the multi-layer perceptron neural network (MLPN). The GNC applies the nearest neighbour classification. The main advantage of the GNC network is the way it handles misclassification problem inherent to the BP algorithm, i.e., strange inputs are treated separately during learning and uncertain classifications. In order to classify these strange input vectors during training, the network output obtained during the training session is supervised and evaluated by an independent unit known as the guarded independent neuron (GI-neuron). Overall performance shows its ability to recognize new patterns without forgetting previous ones.

More recent research focused on: (1) determining the best technique of choosing initial weights and biases for training the ANNs; (2) noise elimination ability; and (3) short training and convergence times.

The noise elimination capability of ANNs was investigated by Chang et al. [56]. Four experimental PD models of PD in cast-resin current transformers (CT)s, with some insulating defects were used as training sets for the ANN. The insulating defects were a perfect CT, corona discharge, low voltage coil PD and HV coil PD. From the experiments, φ-q-n patterns were derived from this model and about 120 matrices were obtained. Each matrix was of (M × N) dimension, where the x-axis covers $0°$–$360°$ with each phase segment of size ($360°/M$), and the q-axis having a range of 0–400 pC with each division equivalent to ($400/N$). In order to ascertain the recognition efficiency, different levels of random noise were created to distort the original measurement. Their result shows that there is about 80% recognition rate for PD measurements up to a 20% noise level.

A potential technique to choose the optimal set of initial weights for training an ANN was determined by Kuo [57]. Particle swarm optimization (PSO) was applied to provide the optimal set of initial weights and biases for the ANN model. PSO is the first optimization technique to get the best initial weights and biases for the BP ANN and is achieved by optimizing certain objective functions [57]. The paper shows the efficiency of the scheme in identifying the insulation aging status

of cast-resin transformers for both noisy and noiseless environments. There is high recognition rate of 94% without noise and with around 30% noise level; the recognition rate was still around 84%.

The work undertaken by Chen et al. [58] produced a faster learning algorithm for training the ANN. In this paper, the BP network was applied to transformer insulation diagnosis, which include analysing: (1) low-voltage coil PD; (2) coil PD; (3) corona discharge; and (4) a healthy transformer. A faster learning algorithm known as resilient propagation was adopted as the learning rule because of its high convergence speed. The algorithm showed high recognition precision and high noise elimination capability. With 30% added random noise, recognition as high as 80% for some defects was recorded. Despite the faster convergence rate implemented by Chen et al. [58], the recognition rate was lower than that obtained with the same BP ANN, but implementing PSO as it was implemented by Kuo [57]. In order to improve PD recognition, there is need to investigate novel BP ANN optimization technique with resilient propagation to obtain high PD recognition in the presence of noise.

Recently, attention has been paid to the application of PNN [18,32], and RBFN [59] to categorize PD fault geometries, i.e., corona and surface discharges in air and oil. Evagorou et al. [60] applied the PNN to categorize some PD fault geometries, i.e., corona and surface PD in air and oil. After training the PNN algorithm, the input vectors containing the features for classification were then applied to calculate the PDF of each category and collectively by assigning the cost for misclassification; the result minimizes the likely risk taken. Maximum likelihood training was applied here and encouraging recognition probabilities of 99% were recorded for corona, while lower rates recorded for floating and internal discharges. In other research, Karthikeyan et al. [18] also applied the PNN to categorize single source PD patterns and a recognition performance of 100% was obtained for some input PD classes, though misclassification still persisted. This indicates that misclassification is still an issue with the BP ANN, where PD faults are misclassified as others and certain techniques to eliminate this issue must be investigated. Recently, Venkatesh and Gopal [32] focused on recognizing complex multiple source PD patterns using a composite version of the PNN, with a recognition efficiency of 97% being recorded. Chang [59] applied the RBFN to classify insulation defects such as external discharge, internal discharge and corona, etc.

The results indicate that the RBFN has the potential for PD recognition and is very effective for clustering PD defects of insulators with less complex features, which greatly reduced the size of the PD fault database. A summary of relevant ANN implementations is shown in Table 1.

Table 1. Summary of relevant ANNs that have been used for PD recognition.

Reference	ANN Type	PD Faults Tested	Input Parameters	Recognition Rate
Suzuki and Endoh [49]	BP	Needle-type defect in XLPE cable	No of discharges, amplitude and phase obtained from the φ-q-n patterns	100%
Gulski and Krivda [10]	BP, KOH, and LVQ	Corona in air, corona in oil, dielectric voids, surface discharges in oil and air, air bubbles in oil and background noise	Statistical parameters from the φ-q-n patterns	100% recorded for trained PD faults
Candela et al. [52]	BP	Surface discharge in air, metallic dielectric parallel air gap	Statistical parameters from the φ-q-n patterns	98% recorded
Mirelli and Schifani [29]	BP	Artificial and natural PD defects in practice, e.g., surface discharges, corona and voids	Statistical parameters from the φ-q-n patterns	92% in 20 kV insulators
Karthikeyan et al. [30]	BP	PD defects in voids, surface discharges and corona	Statistical parameters from the φ-q-n patterns	100% for some testing examples with misclassification recorded
Badent et al. [53]	BP	16 fault types of naturally occurring PD faults	Phase resolving PD matrix derived from the φ-q-n patterns	100% for trained PD faults and 90% for new patterns
Hong and Fang [35]; Hong et al. [56]	MNN using the BP	PD defects in 11 kV bushings, HV wire to surface and corona	Pulse count, average PD magnitude and maximum PD magnitude derived from the φ-q-n patterns	98% recorded
Mazroua et al. [16]	BP	Electrical trees and voids	PD pulse train patterns, e.g., peak amplitude, rise time, fall time and duration	94.5% for trained set and 94.0% for test set
Tian et al. [44]	BP	Different types of internal and bounded voids	Spectral density of acoustic signals	90%
Hoof et al. [31]	GNC using BP	Electrode bounded cavity, corona in air, surface discharges, electrical treeing, and noise sequences	Statistical parameters from the phase resolved patterns	100% recognition rate recorded for some defects
Chang et al. [56]	BP	PD in cast resin transformers, e.g., corona, and surface discharge	120 matrices of data derived from the φ-q-n patterns	80% recognition efficiency with 20% noise level
Kuo [57]	BP, and BP with PSO	Insulation ageing data such as significant and failure imminent	Statistical parameters from the PD signal	94% recognition rate even with 30% noise
Chen et al. [58]	BP	Low-voltage coil PD, HV coil PD, HV corona discharge and healthy transformer	Statistical parameters from the φ-q-n patterns	80% with 30% noise
Evagorou et al. [60]	PNN	Corona and surface discharges in air and oil	Moments of probability function of the wavelet coefficient	99% recorded for some defects
Karthikeyan et al. [18]	PNN	Corona and void discharges in air and oil	Statistical parameters from the φ-q-n patterns	100% recognition rate for some PD defects

6. Strength of the Artificial Neural Networks Applied to Partial Discharge Recognition

From the review carried out, PD pattern recognition using ANNs covers the following aspects:

- Selection of faults to be investigated.
- PD detection, measurement and quantification.
- Selection of appropriate fingerprints, which can be used to train and test the ANN.
- Achievement of high recognition rate targets.

Accordingly, several PD defects from two-electrode models as well as other models of artificial defects have been investigated and PD patterns have been captured and established. ANNs have utilized these artificial PD for pattern classification tasks. These defects include cavities or voids at various positions in insulation, corona and surface discharges in air and oil, electrical trees and floating parts in insulation. Mechanisms of cavities of different sizes and positions within the HV insulation systems such as polyethylene-terephthalate and epoxy resin have been established, including treeing pattern development and void in transformer oil. Corona discharges in air and oil have been widely investigated. Corona discharge in air is studied using a point-plane arrangement with different gap distances. Corona in liquid is studied by sharp needlepoint placed on pressboard in oil. The mechanism of corona in air and oil has already been established. Surface PD activity in oil-pressboard interface is a well-known phenomenon that has been identified as a critical effect in HV apparatus. Repeatable surface discharge measurements have been investigated from oil-pressboard interface using point-plane configuration or any other sharp point. These faults represent some of the most common defects found in transformers, underground cables and electrical machines.

Among the variety of PD parameters fed into the ANN, pulse-height ($H_{qn}(\varphi)$) and phase analysis ($H_n(\varphi)$) fingerprints appears to be the most common among researchers tackling PD classification problems. These phase and pulse-height distributions are obtained from the φ-q-n patterns, which are complex to analyse mathematically. They are then further converted into a 2D, i.e., $H_{qn}(\varphi)$ and $H_n(\varphi)$, distribution to simplify the analysis. However, several fingerprints have been determined from these distributions. These include discharge numbers, amplitude, phase and statistical operators. Statistical PD operators have been the most widely utilized and have been found to give good recognition probability, since they allow to identify uniquely parameters associated for each type of PD without being affected by the experimental set-up used or the applied voltage level during measurement. These include the *sk* and *ku* of the $H_n(\varphi)+$, $H_n(\varphi)-$, $H_{qn}(\varphi)+$, $H_{qn}(\varphi)-$ distributions, *Q*, *cc* and the *mcc*. However, since the work of Gulski and Krivda [10], there appears to be no research carried out to investigate further parameters best suited for PD recognition by the ANN. On the other hand, the BP network is the most widely used training model for the ANN as applied to solve PD classification problems. This is due to its ease of implementation and track record of classifying complex data in other field of applications [61–63]. High recognition rates above 90% were recorded with BP, which is a major success of this scheme. The MNN using the BP has also shown improved classification results with recognition reaching as high as 90%. Nowadays, attention has been paid to the application of PNN and classification results up to 100% have been recorded for some geometry. Based on this achievement, ongoing research by Venkatesh and Gopal [32] has determined the robustness of PNNs with regard to multiple concurrent PD sources, which are difficult to be recognized effectively.

7. Limitations and Suggestions for Improvement

After critical analysis of the literature, some of the limitations have been identified and suggestion for future improvements of the ANN made:

- The BP depends on a trial and error approach to determine the optimum topology, which is time consuming [64]. Though the PSO technique [58] has been implemented to obtain the optimum weights and biases of the ANN, a simpler approach has not yet been determined to ensure shorter training time of the ANN.

- Very few works have considered the application of ANNs for discriminating PD source positioning [49]. Discharges from similar PD sources (e.g., corona, void, surface discharges) at various positions of the HV insulation may vary in characteristics depending on whether it is single or multiple sources. ANN topologies for these scenarios should be further developed.

- Most of the PD fingerprints applied for training and testing the ANN, have been captured over a short period of time. It is necessary to capture data over long stressing periods because some fault PD patterns change over time scales of hours (e.g., voids), which can produce significant changes in parametric statistics. The authors of this paper made an attempt to discriminate different degradation levels of the pressboard subjected to sustained oil surface discharges [7].

- Despite great success recorded by Gulski in applying certain statistical operators as input to the ANN, there is need to investigate novel PD parameters and those better suited for ANN.

- Also, as stated previously, feature selection from the φ-q-n pattern using statistical tools has been the most extensively utilized parameter extraction to select inputs for the ANN. However, PD data has been captured and examined using fixed phase resolutions and amplitude bins by all users. It seems that no work reported in the literature that investigates the statistical operator's sensitivity to different φ-q-n resolution sizes and how this can potentially affect the recognition rates of the ANNs. This is important because different equipment instruments may have different settings for the φ-q-n patterns and this may likely provide inaccurate PD classification result. Although, in recent work [65], has been investigated the effect of φ-q-n resolution sizes on ANN recognition result.

- Since PD fault research is still an on-going activity, to date, there appears to be no comprehensive and up to date databank of PD faults, making the recognition task more challenging.

- To date, multiple concurrent PD sources have become increasingly complex due to the overlapping of discharges and several attempted classification techniques do not yield a reasonably degree of success. There is need to investigate and understand the mechanism of these fault scenarios in order to effectively recognize them. However, several papers reported the correct separation of simultaneous PD sources but not with the ANN (e.g., as demonstrated by Albarracín et al. [66]). The previous application of these separation techniques will significantly improve the PD sources recognition process by the ANN.

- The electrical and radiated noise may be considered as a serious challenge for PD classification research, because as it is shown in Figure 8, these disturbances can be coupled with the PD signals completely altering its spectral content. Moreover, some periodic-pulsing noise from thyristors operation for example, can hide the presence of PD. Uncertainty, regarding the level and proportion of noise in PD patterns may lead to inaccurate classification. To this end, the influence of types and levels of noise on PD source classification needs to be investigated and well understood. Recently, the research carried out by Carvalho et al. [67], and Álvarez et al. [68] shows that noise can be eliminated from PD by using wavelet signal denoising that in combination with ANN techniques resulting in improved classification result.

As a summary, if these weaknesses are adequately addressed in the future work, it is possible to realize a robust and more reliable PD pattern recognition tool.
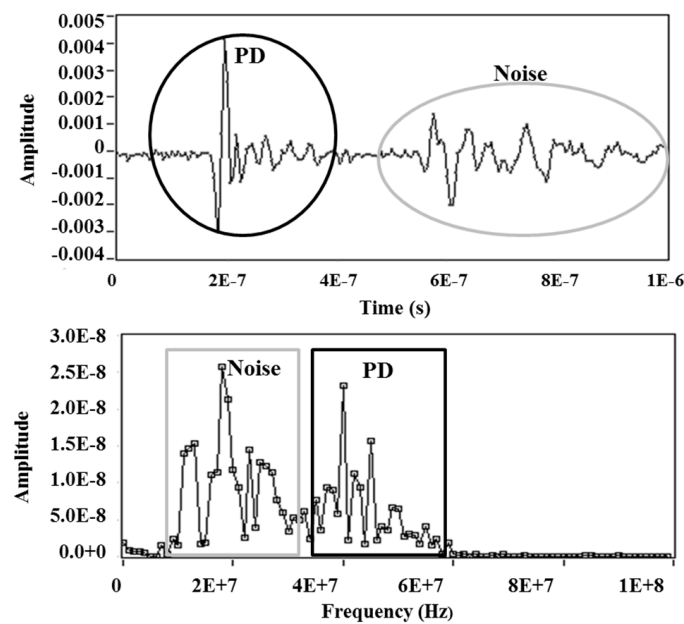
Figure 8. Example of a signal formed by components of PD and noise.

8. Conclusions

This paper has reviewed the recent progress made on the application of ANN for PD recognition as well as proposing some suggestions for future improvement. Tremendous success has been achieved in discriminating a number of PD fault examples, with recognition rate reaching as high as 90%. Different proportion and levels of noise in PD patterns hinders recognition task. There is the issue of long training time of the ANN using the conventional trial and error approach. There appear to be no reliable databank for PD faults as novel PD faults are being investigated, making the recognition task challenging. The mechanisms of multiple defects are not yet understood and effectively classified by the ANN, although this has been achieved with other techniques. There is also the issue of PD patterns variation over different time periods and degradation levels, which has not been well addressed and established. PD data has been captured and recognized using fixed phase resolutions and amplitude bins by all users and different PD testing apparatus may have different resolution settings that may provide an incorrect recognition result. In order to improve PD recognition using ANN certain suggestions were proposed by the authors. These include investigating new optimization techniques for PD recognition using the ANN, using PD data captured over long stressing period for training and testing the ANN, fully understanding the mechanism of multiple faults and using identical φ-q-n resolution sizes in training the ANN.

As a further contribution of this paper, the suitability of the ANN for practical PD recognition has been assessed, and benefits to the practitioners outlined. The ANN can give information regarding seriousness of the PD and the urgency of the need to rectify the fault.

Author Contributions: Abdullahi Abubakar Mas'ud wrote the review paper and provides the analysis of the results. Firdaus Muhammad-Sukki helps in the analysis of the review results and provides suggestions for improvement. Ricardo Albarracín and Jorge Alfredo Ardila-Rey proposed additional chapter and analysis. Hazlee Azil Illias helps in the analysis of the review results. Nurul Aini Bani and Abu Bakar Munir carried out a proofreading on the article and provide suggestions for improvement.

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

Abbreviations

ANN	Artificial neural network
BP	Back-propagation
CM	Conditioning monitoring
CT	Current transformer
ENN	Ensemble neural network
GI-neuron	Guarded independent neuron
GIS	Gas insulated substation
GNC	Guarded neural classifier
HV	High-voltage
KOH	Kohonen self-organizing map
LVQ	Learning vector quantization
MLPN	Multi-layer perceptron neural network
MNN	Modular neural network
MSE	Mean square error
PD	Partial discharge
PDF	Probability density function
PNN	Probabilistic neural network
PSO	Particle swarm optimization
XLPE	Cross-linked polyethylene
α	Scale parameter
β	Shape factor
$\varphi\text{-}q\text{-}n$	Phase-amplitude-number
cc	Cross-correlation
$H_n(\varphi)$	Pulse amplitude distribution
$H_{qn}(\varphi)$	Mean pulse-height distribution
$H_n(\varphi)+$	Pulse-count distribution (+ve half cycle) in phase
$H_n(\varphi)-$	Pulse-count distribution (−ve half cycle) in phase
$H_{qn}(\varphi)+$	Mean pulse-height distribution (+ve half cycle) in phase
$H_{qn}(\varphi)-$	Mean pulse-height distribution (−ve half cycle) in phase
$H_n(q)+$	Pulse amplitude distribution (+ve half cycle) in amplitude
$H_n(q)-$	Pulse amplitude distribution (−ve half cycle) in amplitude
ku	Kurtosis
Q	Discharge factor
TP	Error at the output

References

1. CIGRE Working Group. Report 21.03—Recognition of discharges. *Electra* **1969**, *11*, 61–98.
2. Gulski, E. *Computer-Aided Recognition of Partial Discharges Using Statistical Tools*; Delft University: Delft, The Netherlands, 1991.
3. Hirata, A.; Nakata, S.; Kawasaki, Z.-I. Toward Automatic Classification of Partial Discharge Sources with Neural Networks. *IEEE Trans. Power Deliv.* **2006**, *21*, 526–527. [CrossRef]
4. *High Voltage Test Techniques—Partial Discharge Measurements*; IEC 60270 (2000); International Electrotechnical Commission (IEC): London, UK, 2000.
5. Garnacho, F.; Sánchez-Urán, M.A.; Ortego, J.; Álvarez, F.; González, A. Control of Insulation Condition of Smart Grids by Means of Continuous PD Monitoring. In Proceedings of the 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), Stockholm, Sweden, 10–13 June 2013.
6. Kreuger, F.H. *Partial Discharge Detection in High-Voltage Equipment*; Butterworth-Heinemann: Woburn, MA, USA, 1989.
7. Mas'ud, A.A.; Stewart, B.G.; McMeekin, S.G. Application of an ensemble neural network for classifying partial discharge patterns. *Electr. Power Syst. Res.* **2014**, *110*, 154–162. [CrossRef]
8. Kane, C.; Lease, B.; Golubev, A. Practical experiences of on-line partial discharge measurements on a variety of medium-voltage electrical equipment. *IEEE Trans. Ind. Appl.* **1999**, *35*, 1238–1246. [CrossRef]
9. Morshuis, P.H.F. Partial discharge mechanisms: Mechanisms leading to breakdown, analyzed by fast electrical and optical measurements. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1993.
10. Gulski, E.; Krivda, A. Neural networks as a tool for recognition of partial discharges. *IEEE Trans. Electr. Insul.* **1993**, *28*, 984–1001. [CrossRef]

11. Sarathi, R.; Nandini, A.; Tanaka, T. Understanding electrical treeing phenomena in XLPE cable insulation under harmonic AC voltages adopting UHF technique. *IEEE Trans. Dielectr. Electr. Insul.* **2012**, *19*, 903–909. [CrossRef]

12. Jain, A.K.; Duin, R.P.W.; Mao, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 4–37. [CrossRef]

13. Salama, M.A.; Bartnikas, R. Determination of neural-network topology for partial discharge pulse pattern recognition. *IEEE Trans. Neural Netw.* **2002**, *13*, 446–456. [CrossRef] [PubMed]

14. Bishop, C.M. *Neural Network for Pattern Recognition*; Clarendon Press: Oxford, UK, 1995.

15. Fu, K.-S. *Syntactic Pattern Recognition and Applications*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1982.

16. Mazroua, A.A.; Bartnikas, R.; Salama, M.M.A. Neural network system using the multi-layer perceptron technique for the recognition of PD pulse shapes due to cavities and electrical trees. *IEEE Trans. Power Deliv.* **1995**, *10*, 92–96. [CrossRef]

17. Cavallini, A.; Montanari, G.C.; Contin, A.; Pulletti, F. A new approach to the diagnosis of solid insulation systems based on PD signal inference. *IEEE Electr. Insul. Mag.* **2003**, *19*, 23–30. [CrossRef]

18. Karthikeyan, B.; Gopal, S.; Venkatesh, S. Partial discharge pattern classification using composite versions of probabilistic neural network inference engine. *Expert Syst. Appl.* **2008**, *34*, 1938–1947. [CrossRef]

19. Gulski, E.; Kreuger, F.H. Computer-aided analysis of discharge patterns. *J. Phys. D Appl. Phys.* **1990**, *23*, 1569–1575. [CrossRef]

20. Satish, L.; Zaengl, W.S. Artificial neural networks for recognition of 3-D partial discharge patterns. *IEEE Trans. Dielectr. Electr. Insul.* **1994**, *1*, 265–275. [CrossRef]

21. Majidi, M.; Fadali, M.S.; Etezadi-Amoli, M.; Oskuoee, M. Partial discharge pattern recognition via sparse representation and ANN. *IEEE Trans. Dielectr. Electr. Insul.* **2015**, *22*, 1061–1070. [CrossRef]

22. Satish, L.; Gururaj, B.I. Use of hidden Markov models for partial discharge pattern classification. *IEEE Trans. Electr. Insul.* **1993**, *28*, 172–182. [CrossRef]

23. Abdel-Galil, T.K.; Hegazy, Y.G.; Salama, M.M.A.; Bartnikas, R. Partial discharge pulse pattern recognition using hidden Markov models. *IEEE Trans. Dielectr. Electr. Insul.* **2004**, *11*, 715–723. [CrossRef]

24. Hao, L.; Lewin, P.L.; Dodd, S.J. Comparison of Support Vector Machine Based Partial Discharge Identification Parameters. In Proceedings of the Conference Record of the 2006 IEEE International Symposium on Electrical Insulation, Toronto, ON, Canada, 11–14 June 2006; pp. 110–113.

25. Robles, G.; Parrado-Hernández, E.; Ardila-Rey, J.; Martínez-Tarifa, J.M. Multiple partial discharge source discrimination with multiclass support vector machines. *Expert Syst. Appl.* **2016**, *55*, 417–428. [CrossRef]

26. Ziomek, W.; Reformat, M.; Kuffel, E. Application of genetic algorithms to pattern recognition of defects in GIS. *IEEE Trans. Dielectr. Electr. Insul.* **2000**, *7*, 161–168. [CrossRef]

27. Lai, K.; Phung, B.; Blackburn, T. Application of data mining on partial discharge part I: Predictive modelling classification. *IEEE Trans. Dielectr. Electr. Insul.* **2010**, *17*, 846–854. [CrossRef]

28. Karthikeyan, B.; Gopal, S.; Srinivasan, P.S.; Venkatesh, S. Efficacy of Back Propagation Neural Network Based on Various Statistical Measures for Pd Pattern Classification Task. In Proceedings of the 2006 IEEE 8th International Conference on Properties and Applications of Dielectric Materials, Bali, Indonesia, 26–30 June 2006; pp. 40–43.

29. Mirelli, G.; Schifani, R. A Novel Method for the Recognition of PD Patterns by Neural Network. In Proceedings of the 1999 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, Austin, TX, USA, 17–20 October 1999; pp. 206–209.

30. Karthikeyan, B.; Gopal, S.; Venkatesh, S. ART 2—An unsupervised neural network for PD pattern recognition and classification. *Expert Syst. Appl.* **2006**, *31*, 345–350. [CrossRef]

31. Hoof, M.; Freisleben, B.; Patsch, R. PD source identification with novel discharge parameters using counterpropagation neural networks. *IEEE Trans. Dielectr. Electr. Insul.* **1997**, *4*, 17–32. [CrossRef]

32. Venkatesh, S.; Gopal, S. Robust Heteroscedastic Probabilistic Neural Network for multiple source partial discharge pattern recognition—Significance of outliers on classification capability. *Expert Syst. Appl.* **2011**, *38*, 11501–11514. [CrossRef]

33. Khan, J.A.; Ravichandran, S.; Mallikarjunappa, K. Partial discharge object recognition using cellular neural networks on digital signal processor. *Int. J. Comput. Sci. Netw. Secur.* **2008**, *8*, 115–120.

34. Hong, T.; Fang, M.T.C.; Hilder, D. PD classification by a modular neural network based on task decomposition. *IEEE Trans. Dielectr. Electr. Insul.* **1996**, *3*, 207–212. [CrossRef]

35. Hong, T.; Fang, M.T.C. Detection and classification of partial discharge using a feature decomposition-based modular neural network. *IEEE Trans. Instrum. Meas.* **2001**, *50*, 1349–1354. [CrossRef]

36. Chen, H.-C.; Gu, F.-C.; Wang, M.-H. A novel extension neural network based partial discharge pattern recognition method for high-voltage power apparatus. *Expert Syst. Appl.* **2012**, *39*, 3423–3431. [CrossRef]

37. Zhao, H.; Lin, Z. Fault Diagnosis of Partial Discharge in the Transformers Based on the Fuzzy Neural Networks. In Proceedings of the 2010 International Conference on Computational and Information Sciences (ICCIS), Chengdu, China, 17–19 December 2010; pp. 1253–1256.

38. Tho, N.T.N.; Chakrabarty, C.K.; Siah, Y.K.; Ghani, A.B.A. Implementation of Minimum Distance Classifier for PD Pulse Classification on FPGA. In Proceedings of the Global Engineering, Science and Technology Conference, Dubai, UAE, 1–2 April 2013; pp. 1–13.

39. Han, Y.; Song, Y.H. Condition monitoring techniques for electrical equipment—A literature survey. *IEEE Trans. Power Deliv.* **2003**, *18*, 4–13. [CrossRef]

40. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. [CrossRef]

41. James, R.; Su, Q. *Condition Assessment of High Voltage Insulation in Power System Equipment—IET Power and Energy Series*; The Institution of Engineering and Technology: Stevenage, UK, 2008.

42. Álvarez, F.; Garnacho, F.; Ortego, J.; Sánchez-Urán, M.Á. Application of HFCT and UHF sensors in on-line partial discharge measurements for insulation diagnosis of high voltage equipment. *Sensors* **2015**, *15*, 7360–7387. [CrossRef] [PubMed]

43. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice-Hall: Upper Saddle River, NJ, USA, 1998.

44. Tian, Y.; Lewin, P.L.; Davies, A.E.; Richardson, Z. PD Pattern Identification Using Acoustic Emission Measurement and Neural Networks. In Proceedings of the Eleventh International Symposium on High-Voltage Engineering (ISH 99), London, UK, 23–27 August 1999; Volume 5, pp. 41–44.

45. Mas'ud, A.A.; Stewart, B.G.; McMeekin, S.G.; Nesbitt, A. Partial Discharge Pattern Classification for an Oil-Pressboard Interface. In Proceedings of the IEEE International Symposium on Electrical Insulation, San Juan, Puerto Rico, 10–13 June 2012; pp. 122–126.

46. Mas'ud, A.A.; Stewart, B.G.; McMeekin, S.G.; Nesbitt, A. An Ensemble Neural Network for Recognizing PD Patterns. In Proceedings of the 2010 45th International Universities Power Engineering Conference (UPEC), Cardiff, UK, 31 August–3 September 2010; pp. 1–6.

47. University of Reading. Probabilistic Neural Network (PNN). Available online: http://www.personal.reading.ac.uk/~sis01xh/teaching/CY2D2/Pattern3.pdf (accessed on 6 October 2014).

48. Krivda, A. Automated recognition of partial discharges. *IEEE Trans. Dielectr. Electr. Insul.* **1995**, *2*, 796–821. [CrossRef]

49. Suzuki, H.; Endoh, T. Pattern Recognition of Partial Discharge in XLPE Cables Using a Neural Network. In Proceedings of the 3rd International Conference on Properties and Applications of Dielectric Materials, Tokyo, Japan, 8–12 July 1991; pp. 43–46.

50. Hozumi, N.; Okamoto, T.; Imajo, T. Discrimination of partial discharge patterns using a neural network. *IEEE Trans. Electr. Insul.* **1992**, *27*, 550–556. [CrossRef]

51. Phung, B.T.; Blackburn, T.R.; James, R.E. The Use of Artificial Neural Networks in Discriminating Partial Discharge Patterns. In Proceedings of the Sixth International Conference on Dielectric Materials, Measurements and Applications IET, Manchester, UK, 7–10 September 1992; pp. 25–28.

52. Candela, R.; Mirelli, G.; Schifani, R. PD recognition by means of statistical and fractal parameters and a neural network. *IEEE Trans. Dielectr. Electr. Insul.* **2000**, *7*, 87–94. [CrossRef]

53. Badent, R.; Kist, K.; Lewald, N.; Schwab, A.J. Partial-Discharge Diagnosis with Artificial Neural Networks. In Proceedings of the 4th International Conference on Properties and Applications of Dielectric Materials (ICPADM), Brisbane, Australia, 3–8 July 1994; pp. 638–641.

54. Yamazaki, A.; Tsutsumi, Y.; Yonekura, T. Partial Discharge Recognition Using a Neural Network. In Proceedings of the 4th International Conference on Properties and Applications of Dielectric Materials (ICPADM), Brisbane, Australia, 3–8 July 1994; pp. 642–645.

55. Hoof, M.; Patsch, R.; Freisleben, B. GNC-network: A New Tool for Partial Discharge Pattern Classification. In Proceedings of the Electrical Insulation Conference and Electrical Manufacturing and Coil Winding Conference, Cincinnati, OH, USA, 26–28 October 1999; pp. 511–515.

56. Chang, H.-C.; Kuo, Y.-P.; Lee, C.-Y.; Lin, H.-W. A Partial Discharge Based Defect-Diagnosis System for Cast-Resin Current Transformers. In Proceedings of the 39th International Universities Power Engineering Conference (UPEC 2004), Bristol, UK, 6–8 September 2004; Volume 1, pp. 233–237.

57. Kuo, C.-C. Artificial identification system for transformer insulation aging. *Expert Syst. Appl.* **2010**, *37*, 4190–4197. [CrossRef]

58. Chen, P.-H.; Chen, H.-C.; Liu, A.; Chen, L.-M. Pattern Recognition for Partial Discharge Diagnosis of Power Transformer. In Proceedings of the 2010 International Conference on Machine Learning and Cybernetics, Qingdao, China, 11–14 July 2010; pp. 2996–3001.

59. Chang, W.-Y. Partial Discharge Pattern Recognition Using Radial Basis Function Neural Network. In Proceedings of the 2010 Asia-Pacific Power and Energy Engineering Conference, Chengdu, China, 28–31 March 2010; pp. 1–4.

60. Evagorou, D.; Kyprianou, A.; Lewin, P.L.; Stavrou, A.; Efthymiou, V.; Metaxas, A.C.; Georghiou, G.E. Feature extraction of partial discharge signals using the wavelet packet transform and classification with a probabilistic neural network. *IET Sci. Meas. Technol.* **2010**, *4*, 177–192. [CrossRef]

61. Zhang, J.; Chau, K. Multilayer ensemble pruning via novel multisubswarm particle swarm optimization. *J. Univ. Comput. Sci.* **2009**, *15*, 840–858.

62. Wang, W.; Chau, K.; Xu, D.; Chen, X. Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resour. Manag.* **2015**, *29*, 2655–2675. [CrossRef]

63. Zhang, S.; Chau, K. Dimension Reduction Using Semi-Supervised Locally Linear Embedding for Plant Leaf Classification. In Proceedings of the 5th International Conference on Intelligent Computing (ICIC 2009), Ulsan, Korea, 16–19 September 2009; pp. 948–995.

64. Kolen, J.F.; Pollack, J.B. Back propagation is sensitive to initial conditions. *Complex Syst.* **1990**, *4*, 269–280.

65. Mas'ud, A.A.; Stewart, B.G.; McMeekin, S.G. An investigative study into the sensitivity of different partial discharge φ-q-n pattern resolution sizes on statistical neural network pattern classification. *Measurement* **2016**, *92*, 497–507. [CrossRef]

66. Albarracín, R.; Robles, G.; Martínez-Tarifa, J.M.; Ardila-Rey, J. Separation of sources in radiofrequency measurements of partial discharges using time-power ratio maps. *ISA Trans.* **2015**, *58*, 389–397. [CrossRef] [PubMed]

67. Carvalho, A.T.; Lima, A.C.S.; Cunha, C.F.F.C.; Petraglia, M. Identification of partial discharges immersed in noise in large hydro-generators based on improved wavelet selection methods. *Measurement* **2015**, *75*, 122–133. [CrossRef]

68. Álvarez, F.; Ortego, J.; Garnacho, F.; Sánchez-Urán, M.Á. A clustering technique for partial discharge and noise sources identification in power cables by means of wavelet transform. *IEEE Trans. Dielectr. Electr. Insul.* **2016**, *23*, 469–481. [CrossRef]