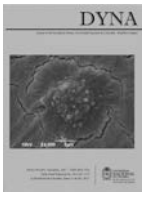
UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS



DYNA
http://dyna.medellin.unal.edu.co/

Wavelet denoising of partial discharge signals and their pattern classification using artificial neural networks and support vector machines

Ian Carlo Guzmán, José Luis Oslinger & Rubén Darío Nieto

Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia. ian.guzman@correounivalle.edu.co, jose.oslinger@correounivalle.edu.co, ruben.nieto@correounivalle.edu.co

Received: March 31th, 2017. Received in revised form: August 11th, 2017. Accepted: October 20th, 2017.

Abstract
This paper presents two pattern recognition approaches using Partial Discharges fingerprints as input features to classify PD patterns. A multi-layer perceptron (MLP) backpropagation neural network and a support vector machine (SVM) were trained to recognize three types of PD patterns. Experimental results showed that the algorithms can achieve high recognition rates. Moreover, the Discrete wavelet transform (DWT) was used to denoise PD signals as a prior stage to the classification process. Different mother wavelets were tested for different levels of decomposition in order to find appropriate wavelet parameters for better signal to noise ratio (SNR) and less distortion after the denoising process.

Keywords: Partial Discharge (PD), Discrete Wavelet Transform (DWT), Artificial Neural Network (ANN), Support Vector Machine (SVM)

Filtrado wavelet de descargas parciales y su clasificación de patrones usando redes neuronales artificiales y máquinas de soporte vectorial

Resumen
Este artículo presenta dos enfoques de reconocimiento de patrones usando huellas dactilares de descargas parciales como características de entrada para llevar a cabo la clasificación de patrones de DP. Un perceptrón multicapa (MLP) basado en el algoritmo de propagación hacia atrás y una máquina de soporte vectorial fueron entrenados para reconocer tres tipos de patrones de DP. Los resultados experimentales demostraron que los algoritmos pueden arrojar altas tasas de reconocimiento. De otra parte, la trasformada wavelet discreta (DWT) fue utilizada para eliminar el nivel de ruido presente en las DP como una etapa previa al proceso de clasificación. Diferentes wavelets madre fueron probadas a diferentes niveles de descomposición con el objeto de encontrar parámetros wavelet apropiados para obtener una mejor relación señal-ruido (SNR) y menos distorsión después del proceso de filtrado.

Palabras clave: Descarga parcial (DP), transformada wavelet discreta, red neuronal artificial, máquina de soporte vectorial.

1. Introduction

A Partial Discharge (PD) is a localized electrical pulse that occurs in the insulation system of high voltage (HV) equipment. Partial discharges appear as short duration pulses having a duration much less than 1μs due to electrical stress concentrations in the insulation systems or on the surface of the insulation [1]. PD ocurrence may cause damage leading to a breakdown of the insulation system.

Therefore, early detection of insulation defects in high voltage equipment is a major concern to be adressed. Extensive research related to PD phenomena has been

How to cite: Guzmán, I.C., Oslinger, J.L. and Nieto, R.D., Wavelet denoising of partial discharge signals and their pattern classification using artificial neural networks and support vector machines DYNA, 84(203), pp. 240-248, December, 2017.

conducted and including: its detection and acquisition techniques, chemical and physical processes, denoising methods, feature extraction and PD pattern classification algorithms [2]. PD denoising is an important task mainly in PD online monitoring systems [3]. Several studies have reported methods based on the wavelet transform for PD denoising of white noise [3-8]. However, the signal to be filtered through the wavelet transform may be distorted. The distortion of a de-noised signal is closely related to the selected mother wavelet and to the threshold selection [9].

Among research areas related to PD phenomena, feature extraction and PD pattern classification are fundamental processes in order to implement an insulation condition monitoring system that can automatically identify the type of PD sources [10-12].

Since 1990, researchers have shown a strong interest in performing PD classification using artificial intelligence techniques [13-22]. Studies like [2,10,13,19,23-33] have reported the use of Multilayer neural networks for PD classification, in which high recognition rates have been found for different types of PD defects. The advantage in using neural networks over other types of classifiers is the possibility they offer to learn from examples [34]. But as early as 2004, researchers have shown a growing interest in DP classification using support vector machines [35-46]. In general, DP classification results using support vector machines have yielded better classification results compared to other classifiers (including neural networks) [2,47]. However, there are still important challenges related to apply appropriate AI techniques for automatic PD source classification such as: feature extraction, application of suitable pattern recognition algorithms and recognizing multiple PD sources that may occur simultaneously in (HV) equipment.

In this paper, we propose a multi-layer perceptron (MLP) backpropagation neural network and a binary support vector machine architecture as artificial intelligence algorithms to automatically identify PD patterns using statistical operators as input features, and compare their performance based on classification results. Additionally, multilevel wavelet decomposition was used to denoise PD signals prior to the classification process. Different mother wavelets for different levels were tested to find better signal to noise ratio (SNR) with less distortion after the denoising process.

This paper is organized as follows: In section 2, the practical setup used in this work is presented. A high level overview of the developed real-time application is explained in section 3. In section 4, we give a brief introduction to Discrete wavelet transform (DWT) and results of PD denoising are presented. The feature extraction process is explained in section 5. The artificial intelligence algorithms and the results of PD pattern classification are presented in section 6. Finally, conclusions are stated in section 7.

2. Experimental setup

The practical setup for PD measurement was implemented according to IEC60270 recommendations and is depicted in Fig. 1.

A voltage level coming from the secondary of the transformer is applied across the test object to generate PD signals.

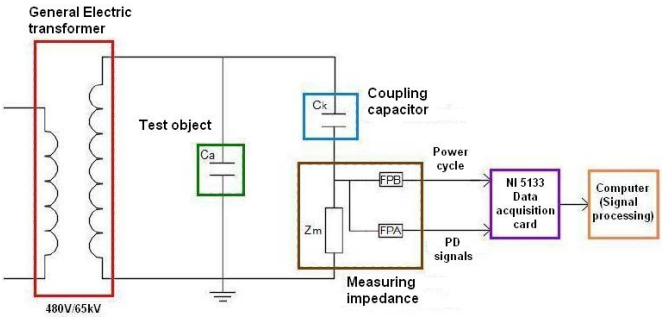


Figure 1. PD measurement setup.
Source: The authors

A quadrupole which contains the measuring impedance includes a high-pass and a low-pass filter in order to separate the DP signals from the power cycle, respectively. The data acquisition card captures the power cycle and the DP signals in different channels to send them to the computer for signal processing.

Artificial PD models such as a tip-hemispherical electrode; a rod-plane electrode and a stator coil were used to simulate corona, surface and internal partial discharges, respectively. The data acquisition card used in this work was the National Instruments NI-5133, which provides two simultaneously sampled channels, a sample frequency up to 100MSPS, 50MHz bandwidth, 8-bit resolution, etc.

3. GUI application

In order to process the PD signals, a Labview application was developed which consists of several tabs that allow: view the acquired signals in scope mode, plot the phase resolved partial discharge diagram, display statistical parameters associated to PD patterns and classify the corona, surface and internal PD patterns using artificial intelligence-based algorithms. Fig. 2 shows the application main window where the user can interact with the acquired PD signals in real time and enter different values depending on the acquisition needs. Once the acquisition and processing are finished, all the results are available to be consulted in their associated tabs.

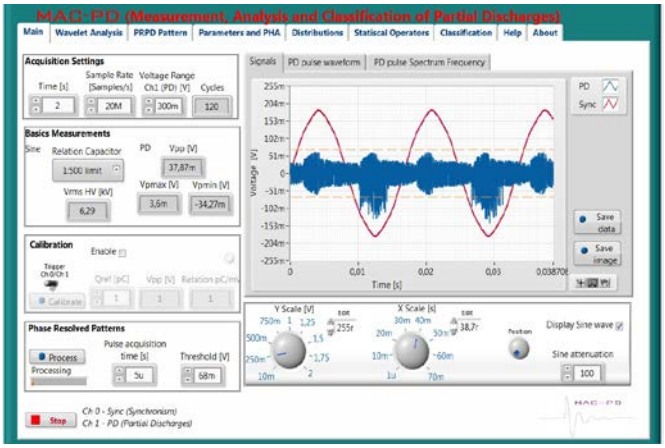


Figure 2. Typical PD patterns
Source: The authors

A 256x256 matrix was designed to record all the informationrelated to PD pulses, where rows represent peak values and columns represent phase angles, as shown in Fig. 3. Each coordinate was associated with a specific PD pulse of a certain peak value and phase angle, therefore, a coordinate is incremented in one every time its associated PD pulse is detected.

The 256x256 matrix facilitated the statistical parameters computation and allowed to plot the PRPD diagram.

Fig. 4 shows the three typical PRPD patterns obtained for each type of PD.

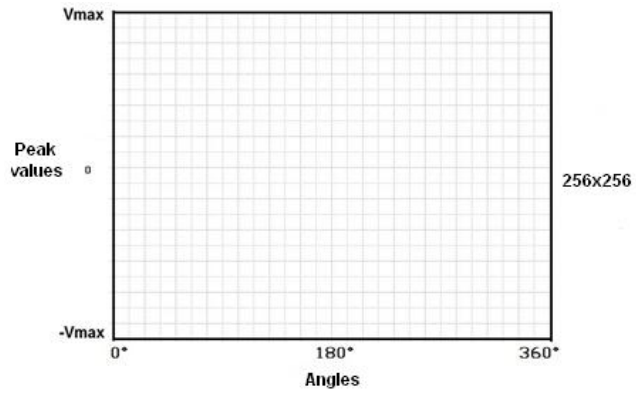
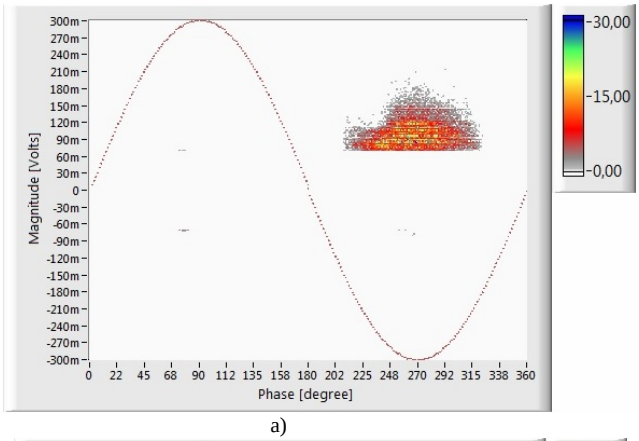
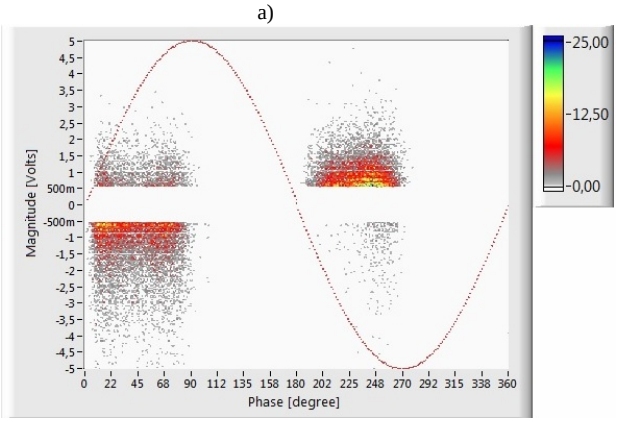


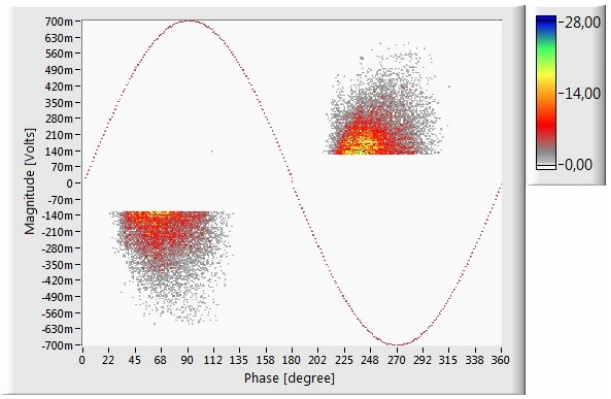
Figure 3. 256x256 Matrix
Source: The authors



a)



b)



c)

Figure 4. Typical PD patterns of: a) Corona discharge, b) Surface discharge, c) Internal discharge.
Source: The authors

4. Discrete Wavelet Transform (DWT)

4.1. A brief introduction to discrete wavelet transform

A wavelet is a short duration small wave that has zero mean value; it increases in amplitude and then decreases back to zero quickly. It satisfies:

$$\int_{-\infty}^{\infty} \psi(t)dt \tag{1}$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \tag{2}$$

Where $\psi(t)$ is the mother wavelet. A family of wavelet functions associated with $\psi(t)$ can be denoted as:

$$\psi(at) = a^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \tag{3}$$

Where a is a scaling factor to determine the amplitude and duration of the mother wavelet. The translation factor b is used to shift the mother wavelet along the time axis. The continuous wavelet transform (CWT) of a signal $f(t)$ is defined as [48-50]:

$$W_\psi f(a,b) = |\, a \,|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt \tag{4}$$

The discrete wavelet transform (DWT) of a signal is obtained by means of a filtering scheme called quadrature mirror filters (QMF) which is a digital filter bank structure. QMF allows for signals to be decomposed into several frequency coefficients and then reconstruct the original signal by using the inverse discrete wavelet transform (IDWT). The original signal is fed through a number of complementary low-pass (L) and high-pass (H) filters and down-sampled by two. The low and high frequency content, also known as "approximation" and "detail" coefficients are given by the low-pass and high-pass filters, respectively [51-55]. A three-level discrete wavelet transform for a signal $s(k)$ is shown in Fig. 5.

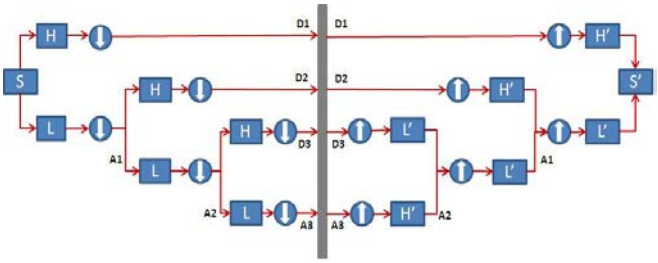


Figure 5. Three-level discrete wavelet transform for a signal "S"
Source: The authors

4.2. Wavelet denoising of partial discharges

There is still no universal or precise method regarding to wavelet type, threshold or level selection since every approach depends on a specific problem [53]. Therefore, the selection of the best wavelet parameters has to be made by means of a trial and error process.

In order to compare the denoising performance with different mother wavelets for different levels of decomposition, the signal to noise ratio (SNR), the reduction in noise level and the root mean square error (RMSE) were considered.

4.2.1. Signal to Noise Ratio (SNR)

It is a measure that indicates the ratio of the power of a signal to the level of noise, usually expressed in decibels (dB). The signal to noise ratio is defined as [53]:

$$SNR(dB) = 10 * log \frac{\sum_{i=1}^{N} y(i)^2}{\sum_{i=1}^{N} \big(x(i) - y(i)\big)^2} \quad (5)$$

Where, $x(i)$ is the signal of reference, $y(i)$ is the denoised signal and N is the number of sample points. A positive value of SNR means the power of signal is greater than the power of noise and vice versa for a negative value of SNR.

4.2.2. Reduction in noise level

In practical measurements, there is no signal of reference, therefore, only the reduction of noise level can be computed, which is the amount of suppressed noise. The normalized reduction of noise level is computed as [54]:

$$RNL\ (dB) = 10 * log \sum_{i=1}^{N} \frac{1}{N} (z(i) - y(i))^2 \quad (6)$$

Where $z(i)$ is the noisy signal, $y(i)$ is the denoised signal and N is the number of samples.

4.2.3. Root Mean Square Error (RMSE)

RMSE is an indicator of the signal distortion after the filtering process. The smaller the RMSE, the more similar is the denoised signal to the signal of reference and less the distortion after filtering. The RMSE is defined as [55]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y(i) - x(i))^2} \quad (7)$$

Where, $x(i)$ is the original signal, $y(i)$ is the denoised signal and N is the number of sample points.

4.3. Denoising results

Many tests were made to find the most appropiate wavelet parameters for better denoising results. It was found that PD signal denoising using a soft-threshold function yields a greater value of RMSE compared to a hard-threshold function. Therefore, all the denoising tests in this work were carried out using the latter.

In order to analyze PD signals with higher level of noise, Gaussian white noise was simulated and added to PD signals. Fig. 6 shows a typical internal PD signal used as a reference signal. Fig. 7 shows the reference signal with simulated white noise added such that SNR=-13dB. Results of denoising in the form of SNR and RMSE values are shown in Table 1.

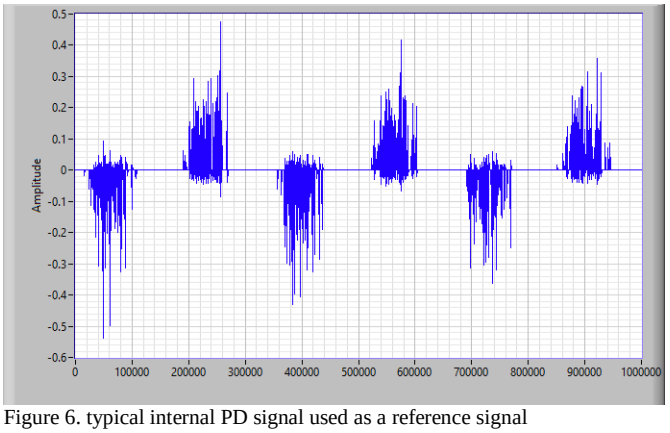


Figure 6. typical internal PD signal used as a reference signal
Source: The authors

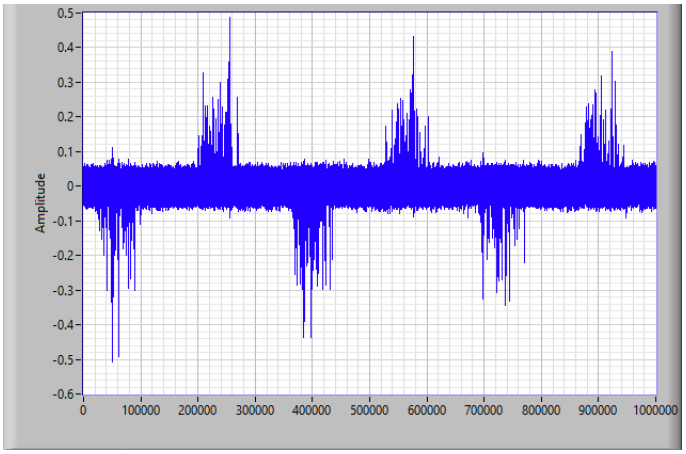


Figure 7. Reference signal with simulated white noise added such that SNR=-13dB.
Source: The authors

Table 1.
Results of denoising a typical PD signal with white noise added

	Decomposition Levels									
Mother wavelet	**1**		**2**		**3**		**4**		**5**	
	SNR(dB)	**RMSE(%)**	**SNR(dB)**	**RMSE(%)**	**SNR(dB)**	**RMSE(%)**	**SNR(dB)**	**RMSE(%)**	**SNR(dB)**	**RMSE(%)**
Db02	18.7804	0.192354	16.9436	0.228477	15.9350	0.249433	15.0973	0.267487	14.7070	0.275679
Db03	18.5311	0.196676	16.8399	0.230734	15.3854	0.261317	14.569	0.278867	14.1272	0.288442
Db04	18.4446	0.198342	16.6296	0.235006	15.0816	0.267047	14.1055	0.288071	13.6588	0.297163
Db05	18.5987	0.195602	16.2031	0.243501	14.7930	0.272898	13.7681	0.294066	13.2726	0.303759
Db06	18.5027	0.197734	15.7586	0.252825	14.3080	0.283637	13.4371	0.301464	13.104	0.307934
Db07	18.2584	0.202324	15.6372	0.255444	14.0520	0.288419	13.1259	0.307225	12.7664	0.314648
Db08	18.1173	0.204887	15.3415	0.262257	13.7274	0.295796	12.7466	0.316293	12.2738	0.325291
Db09	18.0710	0.205806	15.1400	0.266005	13.4883	0.300856	12.5192	0.320433	12.1925	0.329654
Db10	18.0612	0.206055	14.9801	0.268408	13.0349	0.309614	12.1240	0.327537	12.5418	0.335091

Source: The authors

From Table 1, it is observed that the highest values of reduction in noise level and signal to noise ratio are obtained for a single level of DWT decomposition. Also, it can be noted from Table 1 that the RMSE value is smaller for one level compared to higher levels of DWT, which implies that the distortion of the signal is less for a single level. Similar results are reported in [56, 57].

5. Feature extraction

Feature extraction is carried out as a prior stage for pattern recognition. In this work, the feature extraction process was implemented using statistical operators, which can be obtained from statistical distributions. The shape of statistical distributions provide information related to the type of PD signals, therefore, statistical operators can be used as discriminatory parameters for classification purposes. The statistical distributions for PD signals are defined as [16,34]:

$\boldsymbol{Hn^{\pm}(\varphi)}$: The number of discharges per each window as a function of the angle φ.

$\boldsymbol{Hqn^{\pm}(\varphi)}$:The average pulse amplitude per each window as a function of the angle φ.

$\boldsymbol{Hqmax^{\pm}(\varphi)}$:The maximum pulse amplitude per each window as a function of the angle φ.

A set of statistical operators obtained from the distributions described above can be also called PD-figerprints. In this work, PD-fingerprints include:

Skewness (Sk): Measures the degree of asymmetry of a distribution. Sk=0 means the distribution is fully symmetric. A positive skewness means the distribution is skewed to the left. A negative skewness means the distribution is skewed to the right. The skewness is defined as:

$$\boldsymbol{S_k} = \frac{N}{(N-1)*(N-2)}\sum_{i=1}^{N}\frac{(x_i-\mu)^3}{\sigma^3} \tag{8}$$

Where x_i is the i-th variable, μ is the mean value, σ is the variance and N is the number of data points.

Kurtosis: it is an indicator of the sharpness of a distribution. Ku=0 means a normal distribution. A

positive kurtosis means a sharp distribution. A negative kurtosis means a flat distribution. Kurtosis is defined as:

$$\mathrm{K_u} = \sum_{\mathrm{i=1}}^{\mathrm{N}}\frac{(\mathrm{x_i}-\mu)^4}{\mathrm{N}\sigma^4} - 3 \tag{9}$$

Mean: it is the average for each semi-cycle of each distribution.

Cross-correlation: it describes the difference in distribution shape between the positive and negative semi-cycle of a distribution. c.c=0 means total shape asymmetry. c.c=1 means total shape symmetry. Cross-correlation is defined as:

$$\mathrm{c.c} = \frac{\sum \mathrm{x_i} * \mathrm{y_i} - \sum \mathrm{x_i}\sum \mathrm{y_i}/\mathrm{N}}{\sqrt{\left(\sum \mathrm{x_i^2} - \frac{(\sum \mathrm{x_i})^2}{\mathrm{N}}\right)\left(\sum \mathrm{y_i^2} - \frac{(\sum \mathrm{y_i})^2}{\mathrm{N}}\right)}} \tag{10}$$

6. Pattern recognition algorithms

6.1. Classification using ANN

There are different ANN-based algorithms to solve pattern recognition problems [34].

The neural network implemented in this work is a Multilayer Perceptron (MLP) based on the backpropagation algorithm [58]. The MLP structure used in this work is made up of three layers as illustrated in Figure 8: an input layer, a hidden layer and an output layer. Layers are made up of a number of neurons in which each neuron has a sigmoid activation function; the output layer is made up of three neurons, each for one of the three types of PD patterns to be classified (internal, corona and surface). PD fingerprints were used to train the neural network using measurements files so that a new feature input vector can be classified into one of the PD patterns. The feature set consists of 20 statistical operators as shown in Figure 8, including skewness, kurtosis, mean, number of peaks and cross-correlation, which are obtained from the statistical distributions$Hn^{\pm}(\varphi)$, $Hqn^{\pm}(\varphi)$ and $Hqmax^{\pm}(\varphi)$.

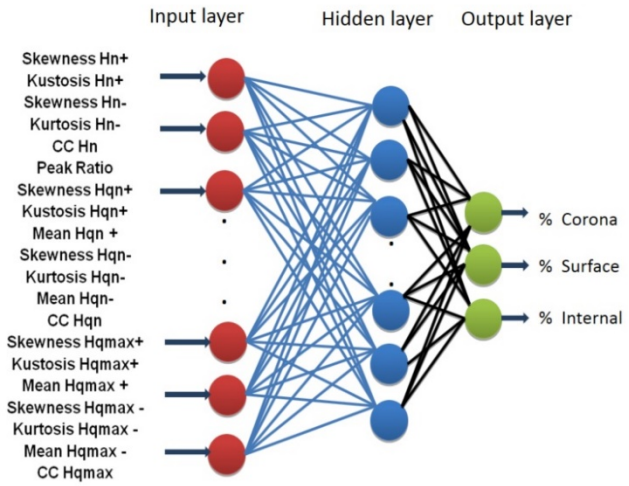


Figure 8. MLP structure
Source: The authors

6.2. Classification using SVM

A support vector machine (SVM) is a binary classifier and a supervised learning model that takes the data points belonging to a spatial domain into another spatial domain of a major dimension which is defined by a kernel function that separates the two classes by means of an hyperplane (also called support vector). This facilitates the classification of the input data into one of the two classes. A support vector machine always looks for the hyperplane that maximizes the margin between the two spaces of points.

Mathematically, an SVM is a linear classifier according to equation 11, which maximizes the margins of classification. The margin of classification is the distance between the classification boundary and the closest point to each class. Equation 12 describes the separation hyperplane. Margin maximization (which minimizes the risk of errors in classifying data) gives the SVM an excellent generalization capability.

$$f(x) = sgn(w^T \bullet x + b) \tag{11}$$

$$w^T \bullet x + b = 0 \tag{12}$$

Where w^T is a vector orthogonal to the hyperplane and "•" is a dot product.

For non-linear separable data, a kernel function is introduced in order to allow the data to be brought into a larger space called a feature space, so that equation 11 can be rewritten as in equation 13 [35]:

$$f(x) = sgn(w^T.\Phi(x) + b) \tag{13}$$

Where $\Phi(x)$ is the kernel function.

A polynomial Kernel was used in this work and is defined as:

$$K(x, y) = (x^T y + C)^n \tag{14}$$

Where x and y are vectors in the input space, n is the degree of the polynomial. The degree of the polynomial Kernel used in this work was quadratic ($n = 2$).

Support vector machines are binary classifiers, that is, they can only classify two classes. However, it is possible to classify more than two classes using several units in parallel, where the number of units to be used is calculated by equation 15.

$$N = \frac{J(J-1)}{2} \tag{15}$$

Where J is the number of classes to be classified.

A "voting process" must be held where each class can obtain up to $J-1$ possible votes. The entry pattern will be assigned to the most voted class [36].

The SVM architecture implemented in this work is shown in Figure 9. According to equation 15, three units in parallel were required to classify the three types of PD patterns (internal, corona and surface).

Each unit consists of a learning and testing stage, therefore, each unit has to be trained with the corresponding dataset in order to process new input patterns. A new input pattern is going to be classified into one of two PD patterns at the outputs of a testing stage; afterwards, the voting class algorithm receives the classification results of each testing unit in order to perform a count and classify the input pattern into one of the three PD patterns.

6.3. Classification results

To construct the training dataset, a preprocessing stage was carried out in order to acquire all the set of statistical operators explained in section 5 for each type of PD. After the training process was finished, the classifiers were tested with new input patterns and they could achieve high recognition rates. Tables 2 and 3 show the results of MLP and SVM of four tests for each type of PD, in order to obtain recognition rates for input features not included in the training dataset, respectively. From Table 2, it can be observed that the highest average recognition rate is for corona discharge with a 99.9% rate, 98.31% for internal discharge and 95.31% for surface discharge. Given that a SVM is a binary classifier, it can be noted from Table 3 that recognition rates provided by the SVM are 0% or 100%, where 100% means an identified pattern. Therefore, input PD patterns are well identified by the SVM as well.

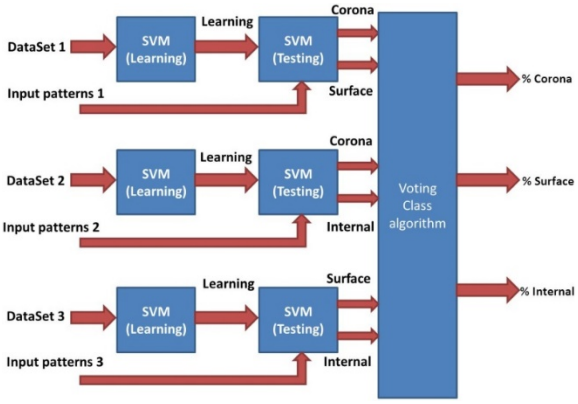


Figure 9. SVM architecture
Source: The authors

Table 2.
Recognition rates of MLP

Recognition rate					
Type of PD	**Internal (%)**	**Corona (%)**	**Surface (%)**	**Average (%)**	**Accuracy (%)**
Internal	99.99	0.005	0.005	98.31	100
	98.5549	0.7225	0.72256		
	98.0844	1.4366	0.47889		
	96.6435	1.6782	1.67823		
Corona	0.05	99.99	0.05	99.99	100
	0.05	99.99	0.05		
	0.05	99.99	0.05		
	0.05	99.99	0.05		
Surface	0.05	0.05	99.99	95.42	100
	1.791	1.9422	96.2668		
	4.06089	1.1489	94.7902		
	7.91322	1.4481	90.6386		

Source: The authors

Table 3.
Recognition rates of SVM

Recognition rate					
Type of PD	**Internal (%)**	**Corona (%)**	**Surface (%)**	**Average (%)**	**Accuracy (%)**
Internal	100	0	0	100	100
	100	0	0		
	100	0	0		
	100	0	0		
Corona	0	100	0	100	100
	0	100	0		
	0	100	0		
	0	100	0		
Surface	0	0	100	100	100
	0	0	100		
	0	0	100		
	0	0	100		

Source: The authors

7. Conclusions

Implementation and results related to PD denoising using the Discrete wavelet transform (DWT) and PD pattern classification using artificial intelligence-based algorithms have been presented in this paper.

According to the results of filtering PD signals through the Discrete Wavelet Transform, first level of decomposition was enough to remove white noise with greater signal to noise ratio (SNR) than to higher levels of decomposition. Also, it was noted that as the number of levels increases, the distortion of the signal increases as well and, therefore, the SNR decreases.

The multilayer perceptron (MLP) neural network and the support vector machine (SVM) implemented in this work were trained and tested using PD-fingerprints as input features and both achieved high recognition rates. Since the SVM is a binary classifier, recognition rates at its outputs were provided as 0% and 100%, where 100% means an identified PD pattern. Recognition rates of more than 90% were achieved by the MLP neural network. Both algorithms were executed simultaneously in the Labview real-time application developed in this work.

Future research will consider multiple PD source classification.

References

[1] IEC, IEC 60270 High-voltage test techniques: Partial discharge measurements, 2000.

[2] Ma, H., Chan, J.C., Saha, T.K. and Ekanayake, C., Pattern recognition techniques and their applications for automatic classification of artificial partial discharge sources. IEEE Transactions on Dielectrics and Electrical Insulation, 20(2), pp. 468-478, 2013. DOI: 10.1109/TDEI.2013.6508749

[3] Florkowska, M.F.B., Phase-resolved rise-time-based discrimination of partial discharges. IET Generation, Transmission and Distribution, 3(1), pp. 115-124, 2009.

[4] Lozano, B., Diseño de software off-line para el análisis estadístico de descargas parciales, Tesis, Universidad Carlos III de Madrid, Madrid, España, 2013.

[5] van Veen, L.W., Comparison of measurement methods for partial discharge measurements in power cables, Thesis, Delft University of Technology, Delft, Holand, 2014.

[6] Seghir, T., Mahi, D., Lebey, T. and Malec, D. Analysis of the Electric field and the potential distribution in cavities inside solid insulating electrical materials. Proceedings of the COMSOL User Conference, Electr. Eng., pp. 2-5, 2006.

[7] Ahmed, N. and Srinivas, N., PD types and their detection possibilities by PD test methods, Annual Report Conference on Electrical Insulation and Dielectric Phenomena, Kitchener, Ont., 2001, pp. 307-310. DOI: 10.1109/CEIDP.2001.963545

[8] Bodega, R., Morshuis, P.H.F., Lazzaroni, M. and Wester, F.J., PD recurrence in cavities at different energizing methods, in IEEE Transactions on Instrumentation and Measurement, 53(2), pp. 251-258, 2004.

[9] Gutfleisch, F. and Niemeyer, L., Measurement and simulation of PD in epoxy voids, in IEEE Transactions on Dielectrics and Electrical Insulation, 2(5), pp. 729-743, 1995. DOI: 10.1109/94.469970

[10] Turner, M. and Gulski, E., (2010). Pattern recognition for partial discharge measurement. Tettex Instruments Division, Haefely Test AG,[online]. (July), pp. 1-18. Available from: http://www.haefely.com/pdf/solutions/DSWpaper.pdf

[11] Contin, A. and Pastore, S., Classification and separation of partial discharge signals by means of their auto-correlation function evaluation, in: IEEE Transactions on Dielectrics and Electrical Insulation, 16(6), pp. 1609-1622, 2009. DOI: 10.1109/TDEI.2009.5361581

[12] Hao, L. et al., Discrimination of multiple PD sources using wavelet decomposition and principal component analysis, in: IEEE Transactions on Dielectrics and Electrical Insulation, 18(5), pp. 1702-1711, 2011. DOI: 10.1109/TDEI.2011.6032842

[13] Gulski, E. and Krivda, A., Neural networks as a tool for recognition of partial discharges. IEEE Transactions on Electrical Insulation, 28(6), pp. 984-1001,1993. DOI: 10.1109/14.249372

[14] Gulski, E., Computer-aided recognition of partial discharges using statistical tools, Delft: Delft University Press, 1991.

[15] Kreuger, F.H., Gulski, E. and Krivda, A., Classification of Partial Discharges. IEEE Transactions on Electrical Insulation, 28(6), pp. 917-931, 1993. DOI: 10.1109/14.249365

[16] Gulski, E. and Kreuger, F.H., Diagnostics of insulating systems using statistical tools. Conference record of the 1992 IEEE International Symposium on Electrical Insulation, pp. 393-396, 1992. DOI: 10.1109/ELINSL.1992.246981

[17] Gulski, E. and Kreuger, F.H. Computer-aided Analysis of discharge patterns, J. Phys, D: Appl. Phys. 23, pp. 1569-1575, 1990.

[18] Krivda, A., Recognition of discharges: Discrimination and classification. Electra, 1995.

[19] Mazroua, A.A., Salama, M.M.A. and Bartnikas, R., PD pattern recognition with neural networks using the multilayer perceptron technique, in: IEEE Transactions on Electrical Insulation, 28(6), pp. 1082-1089, 1993. DOI: 10.1109/14.249382

[20] Kranz, H.G., Diagnosis of partial discharge signals using neural networks and minimum distance classification, in: IEEE Transactions on Electrical Insulation, 28(6), pp. 1016-1024, 1993. DOI: 10.1109/14.249375

[21] James, R.E. and Phung, B.T., Development of computer-based measurements and their application to PD pattern analysis, in: IEEE Transactions on Dielectrics and Electrical Insulation, 2(5), pp. 838-856, 1995.

[22] Chen, P. and Chen, H. (n.d.). Application of back-propagation neural network to power transformer insulation diagnosis, pp. 26-34.

[23] Chang, C.S., Jin, J., Chang, C., Hoshino, T., Hanai, M. and Kobayashi, N., Separation of corona using wavelet packet transform and neural network for detection of partial discharge in gas-insulated substations, in: IEEE Transactions on Power Delivery, 20(2), pp. 1363-1369, 2005. DOI: 10.1109/TPWRD.2004.839187

[24] Li, J., Sun, C., Wang, Y., Yang, J. and Du, L., PD pattern recognition using combined features, Conference Record of the 2004 IEEE International Symposium on Electrical Insulation, 2004, pp. 139-142. DOI: 10.1109/ELINSL.2004.1380490

[25] Evagorou, D., Kyprianou, A., Lewin, P.L., Stavrou, A., Efthymiou, V. and Georghiou, G.E., Classification of partial discharge signals using probabilistic neural network, in: IEEE International Conference on Solid Dielectrics, Winchester, UK, 2007, pp. 609-615. DOI: 10.1109/ICSD.2007.4290887

[26] Boczar, T., Borucki, S., Cichon, A. and Zmarzly, D., Application possibilities of artificial neural networks for recognizing partial discharges measured by the acoustic emission method, in: IEEE Transactions on Dielectrics and Electrical Insulation, 16(1), pp. 214-223, 2009. DOI: 10.1109/TDEI.2009.4784570

[27] Shurrab, I.Y., El-Hag, A., Assaleh, K. and Ghunem, R., Partial discharge on-line monitoring of outdoor insulators. Conference Record of IEEE International Symposium on Electrical Insulation, pp. 391-394, 2012. DOI: 10.1109/ELINSL.2012.6251496

[28] Nasional, U.T., FPGA implementation of neural network classifier for partial discharge time resolved data from magnetic probe, 2011, pp. 451-455.

[29] Macedo, E.C.T., Villanueva, J.M., Da Costa, E.G., Freire, R.C.S., Araujo, D.B., De Souza Neto, J.M.R. and Glover, I.A. Assessment of dielectric degradation by measurement, processing and classification of Partial Discharges, Proceedings of the 2012 IEEE International Power Modulator and High Voltage Conference, IPMHVC 2012, 2, pp. 587-590. DOI: 10.1109/IPMHVC.2012.6518812

[30] Gaoyang, Li, Mingzhe, R., Xiaohua, W., Xi, L. and Yunjia, L., Partial discharge patterns recognition with deep Convolutional Neural Networks, 2016 International Conference on Condition Monitoring and Diagnosis (CMD), Xi'an, 2016, pp. 324-327. DOI: 10.1109/CMD.2016.7757816

[31] Suprayogi, B. and Khayam, U., Design of partial discharge location identifier software for high voltage generator using artificial neural network, in: Proceedings of the Joint International Conference on Electric Vehicular Technology and Industrial, Mechanical, Electrical and Chemical Engineering (ICEVT & IMECE), Surakarta, 2015, pp. 82-87. DOI: 10.1109/ICEVTIMECE.2015.7496664

[32] Pattanadech, N., Nimsanong, P., Potivejkul, S., Yuthagowith, P. and Polmai, S., Partial discharge classification using probabilistic neural network model, in: 18th International Conference on Electrical Machines and Systems (ICEMS), Pattaya, 2015, pp. 1176-1180. DOI: 10.1109/ICEMS.2015.7385217

[33] Catterson, V.M. and Sheng, B., Deep neural networks for understanding and diagnosing partial discharge data, in: IEEE Electrical Insulation Conference (EIC), Seattle, WA, 2015, pp. 218-221. DOI: 10.1109/ICACACT.2014.7223616

[34] Sahoo, N.C., Salama, M.M.A. and Bartnikas, R., Trends in partial discharge pattern classification: A survey. IEEE Transactions on Dielectrics and Electrical Insulation, 12(2), pp. 248-264, 2005, DOI: 10.1109/TDEI.2005.1430395

[35] Hao, L., Lewin, P.L., Tian, Y. and Dodd, S.J., Partial discharge identification using a support vector machine, in: CEIDP '05. Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 2005, pp. 414-417. DOI: 10.1109/CEIDP.2005.1560708

[36] Robles, G., Parrado-Hernández, E., Ardila-Rey, J. and Martínez-Tarifa, J.M., Multiple partial discharge source discrimination with multiclass support vector machines. Expert Systems with Applications, 55, pp. 417-428, 2016. DOI: 10.1016/j.eswa.2016.02.014

[37] Poyhonen, S., Conti, M., Cavallini, A., Montanari, G.C. and Filippetti, F., Insulation defect localization through partial discharge measurements and numerical classification, IEEE International Symposium on Industrial Electronics, 1, pp. 417-422, 2004. DOI: 10.1109/ISIE.2004.1571844

[38] Hao, L., Lewin, P.L., Tian, Y. and Dodd, S.J., Partial discharge identification using a support vector machine, in: CEIDP '05. Annual Report Conference on Electrical Insulation and Dielectric Phenomena, 2005, pp. 414-417. DOI: 10.1109/CEIDP.2005.1560708

[39] Aziz, N.F.A., Hao, L. and Lewin, P.L., Analysis of partial discharge measurement data using a support vector machine, in: 5th Student Conference on Research and Development, Selangor, Malaysia, 2007, pp. 1-6. DOI: 10.1109/SCORED.2007.4451430

[40] Hao, L., Lewin, P.L. and Swingler, S.G., Identification of multiple partial discharge sources, in: International Conference on Condition Monitoring and Diagnosis, Beijing, 2008, pp. 118-121. DOI: 10.1109/CMD.2008.4580244

[41] Hao, L. and Lewin, P.L., Partial discharge source discrimination using a support vector machine, in: IEEE Transactions on Dielectrics and Electrical Insulation, 17(1), pp. 189-197, 2010. DOI: 10.1109/TDEI.2010.5412017

[42] Tang, J., Liu, F., Zhang, X., Liang, X. and Fan, Q., Partial discharge recognition based on SF_6 decomposition products and support vector machine, in: IET Science, Measurement & Technology, 6(4), pp. 198-204, 2012. DOI: 10.1049/iet-smt.2011.0163

[43] Ma, H., Saha, T.K. and Ekanayake, C., Statistical learning techniques and their applications for condition assessment of power transformer, in: IEEE Transactions on Dielectrics and Electrical Insulation, 19(2), pp. 481-489, 2012. DOI: 10.1109/TDEI.2012.6180241

[44] Sarathi, R., Merin-Sheema, I.P. and Abirami, R., Partial discharge source classification by support vector machine, in: IEEE 1st International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), Kolkata, 2013, pp. 255-258. DOI: 10.1109/CATCON.2013.6737508

[45] Xu, J., Niu, H. and Hu, R., The feature extraction and pattern recognition of partial discharge type using energy percentage of wavelet packet coefficients and support vector machines, in: 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), Changsha, 2015, pp. 1776-1779. DOI: 10.1109/DRPT.2015.7432530
[46] Zhi-Qin, M., Xian, Y., Chun-Yao, L., Dan, Z., Hong-Ming, H. and Jian-Guang, C., Partial discharge developing stages identification based on cluster-hierarchical decision SVM in oil-paper insulation, in: IEEE International Conference on High Voltage Engineering and Application (ICHVE), Chengdu, 2016, pp. 1-4. DOI: 10.1109/ICHVE.2016.7800654
[47] Hunter, J.A., Hao, L., Lewin, P.L., Evagorou, D., Kyprianou, A. and Georghiou, G.E., Comparison of two partial discharge classification methods. Conference Record of IEEE International Symposium on Electrical Insulation. 2010. DOI: 10.1109/ELINSL.2010.5549736
[48] Ma, X., Zhou, C. and Kemp, I.J., Interpretation of wavelet analysis and its application in partial discharge detection. IEEE Transactions on Dielectrics and Electrical Insulation, 9(3), pp. 446-457. DOI: 10.1109/TDEI.2002.1007709
[49] Mallat, S.G., A wavelet tour of signal processing, Academic Press, 1998.
[50] S.G. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), pp. 674-693, 1989. DOI: 10.1109/34.192463
[51] Zhang, H., Blackburn, T.R., Phung, B.T. and Sen, D., A novel wavelet transform technique for on-line partial discharge measurements part 1: WT de-noising algorithm. IEEE Transactions on Dielectrics and Electrical Insulation, 14(1), pp. 3-14, 2007. DOI: 10.1109/TDEI.2007.302864
[52] Zhang, H., Blackburn, T.R., Phung, B.T. and Liu, Z., Signal processing of online partial discharges measurements in HV power cables, Australasian Universities Power Engineering Conf., Brisbane, Australia, Paper ID 161, 2004.
[53] Skea, J., Anderson, D., Green, T., Gross, R., Heptonstall, P. and Leach M., Intermittent renewable generation and the cost of maintaining power system reliability. Generation, Transmission & Distribution, IET, 1(2), pp. 324-325, 2007. DOI: 10.1049/iet-gtd:20070023
[54] Satish, L. and Nazneen, B., Wavelet-based denoising of partial discharge signals buried in excessive noise and interference, in: IEEE Transactions on Dielectrics and Electrical Insulation, 10(2), pp. 354-367, 2003. DOI: 10.1109/TDEI.2003.1194122
[55] He, H., Tan, Y. and Wang, Y., Optimal base wavelet selection for ECG noise reduction using a comprehensive entropy criterion. Entropy, 17(9), pp. 6093-6109, 2015. DOI: 10.3390/e17096093
[56] Arivazhagan, S., Deivalakshmi, S. and Kannan, K., Performance analysis of image denoising system for different levels of wavelet decomposition. International Journal of Imaging Science and Engineering, 1(3), pp. 104-107, 2007.
[57] Sharmila, G., Maheswari, R.V. and Subburaj, P., Partial discharge signal denoising using wavelet techniques-on site measurements, in: International Conference on Circuits, Power and Computing Technologies (ICCPCT), Nagercoil, 2013, pp. 673-678. DOI: 10.1109/ICCPCT.2013.6528974
[58] Bishop, C.M., Neural Networks for pattern recognition, Oxford, Clarendon Press, USA, 1995.

I.C. Guzmán-Velásquez, received his BSc. degree in Electronics Eng., from Universidad del Valle, Cali, Colombia in 2013. He received his MSc. degree in Electrical Engineering at Universidad del Valle, Cali, Colombia in 2017. His research interests include: signal processing and machine learning.
ORCID: 0000-0002-4532-8814

J.L. Oslinger-Gutierrez, received his BSc. degree in Electrical Eng., from Universidad del Valle, Cali, Colombia in 1996. He received his PhD degree from Universidad del Valle, Cali, Colombia in 2007. Currently, he is a full time professor at Universidad del Valle in the electrical and electronics school. His research interests include: PD phenomena and design, diagnostics and simulation of rotating electric machines.
ORCID: 0000-0001-9004-3562

R.D. Nieto-Londoño, received his BSc. degree in Electrical Eng., from Universidad del Valle, Cali, Colombia in 1995. He received his MSc. degree with an emphasis in electronics engineering from Universidad del Valle, Cali, Colombia in 2001 and received his PhD degree from Universidad del Valle, Cali, Colombia in 2009. Currently, he is a professor at Universidad del Valle in the electrical and electronics school. His research interests include: digital circuits design, low-power digital design and computer architecture.
ORCID: 0000-0002-1113-3269