A Constrained Neural Classifier for Pulsed Eddy Current based Flaw Detection in Industrial Pipes

Gilvan F. da Silva, Edmar E. P. de Souza, Paulo C. M. A. Farias and Eduardo F. Simas Filho
Digital Systems Laboratory, Electrical Engineering Program, Federal University of Bahia, Salvador, Brazil
gilvan.farias.silva@gmail.com, {edmar.egidio, eduardo.simas, paulo.farias}@ufba.br
Maria C. S. Albuquerque, Ivan C. da Silva and Cláudia T. Farias
Federal Institute of Bahia, Salvador, Brazil
{cleaalbuquerque, ivan.silva, c.farias}@ifba.edu.br

***Abstract*—Decision support systems are important to improve the efficiency of nondestructive evaluation, specially for industrial equipment. Pulsed eddy-current is a magnetic method used for evaluation of metallic equipment. In this paper, is proposed the combination of pulsed eddy current evaluation, digital signal processing, and neural networks to detect flaws in industrial pipes. A novel method using particle swarm optimization is proposed for imposing performance constraints during neural classifier training process. Results obtained for experimental signals acquired from composite-insulated metallic industrial pipes presenting internal and external corrosion areas are used to validate the proposed method. A comparison to neural networks trained from the traditional back-propagation algorithm was presented.**

***Index Terms*—Artificial Neural Networks, Particle Swarm Optimization, pulsed-eddy current evaluation, signal processing.**

I. Introduction

Pulsed Eddy Current (PEC) is a nondestructive testing technique based on electromagnetic induction that can be applied for detecting flaws in metallic equipment [1]. PEC has advantages over the traditional Eddy Current testing, such as deeper penetration depth and more accurate information about flaws [2] [3].

An application of PEC is testing the corrosion level of industrial metallic pipes. Some pipes present an external thermal insulation coating of composite material and due to the infiltration of water or humidity these pipes can present corrosion under isolation (CUI), that, in extreme cases, may cause the pipe rupture [4]. Using time-domain information from PEC signals it is possible to detect pipe corrosion, but it is difficult to efficiently determine where it occurs, in internal or external pipe walls, which is an important information for predictive maintenance.

A way to classify pipe defects from PEC signals is through the application of Artificial Neural Networks (ANN) [5] [6] [7]. Back-propagation [8] is a classic method to train ANN and is based on minimization of Mean Square Error (MSE). For nondestructive evaluation on industrial equipment and parts, the characterization of the defect type is very important for definition of a proper maintenance procedure. In some industrial activities, specific types of defects may present more potential danger than others. Thus, information on defect criticality when considered during the training of the multiclass classifier may lead to improvements in the detection of defects considered most important. However This issue cannot be accomplished by the back-propagation training method, as it is based on MSE minimization, which disregards the distinct cost of error between classes.

In this paper is proposed an ANN constrained training method based on Particle Swarm Optimization (PSO) [9] for multiclass classification problems. The method allows including constraints to prioritize a defect class over other. The PSO algorithm was chosen because of its simplicity and the possibility of including performance constraints.

The data obtained from PEC testing has large dimension, which means that the ANN training complexity increases, hampering the ANN learning. To improve the classification efficiency and avoid the curse of dimensionality, the PEC signal was pre-processed, reducing its dimension [5]. In this paper two processing techniques were employed, Discrete Fourier Transform (DFT) [10] and Discrete Wavelet Transform (DWT) [11]. Both methods were used in order to enable posteriorly comparison. For selecting the most important features and reduce the data dimensionality, in order to improve the classification efficiency, Principal Components Analysis (PCA) [12] was applied.

II. Proposed Method

The proposed processing sequence comprises three stages: (i) feature extraction; (ii) feature selection; (iii) classification. In first stage were applied DFT and DWT, in second stage was applied PCA and in third stage an ANN was trained by PSO to accomplish discrimination performance constraints. Fig. 1 shows the proposed signal processing diagram.

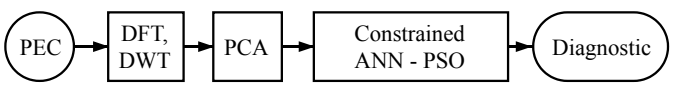

Fig. 1: Proposed signal processing diagram.

The testing piece for this experiment was a carbon steel pipe coated with composite material. This pipe presents corrosion on both, inside and outside walls, as illustrated in Fig. 2. The dimensions for the area with corrosion are shown in

Table I [13]. Three classes of interest were defined for this experiment, the no-defect (ND) class for region without corrosion, the external defect (ED) and internal defect (ID).

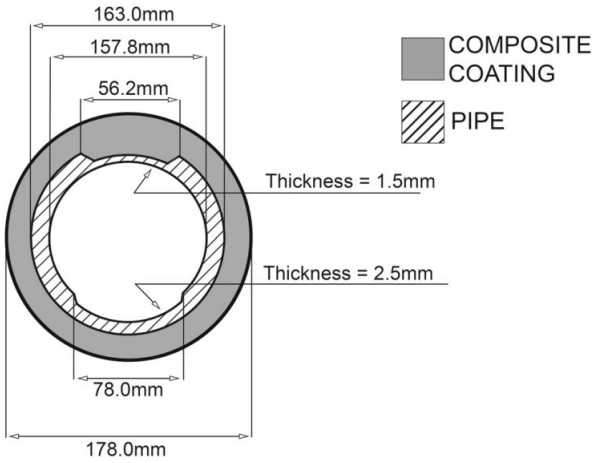

Fig. 2: Cut view of testing piece [13].

TABLE I: Testing piece dimensions [13].

Defect	Length (mm)	Width (mm)	Thickness (mm)
ED	86,8	56,2	1,5
ID	96,0	78,0	2,5

In pulsed eddy current nondestructive evaluation, a coil energized by an alternating current generates a primary alternating magnetic field, which induces eddy currents in the metallic object under test. A secondary magnetic field is generated in opposition to the primary field [14], [15]. The presence of flaws in testing piece, like cracks or corrosion, changes the eddy current characteristics, hence the secondary magnetic field changes. Fig. 3 illustrates the pulsed eddy current principle.

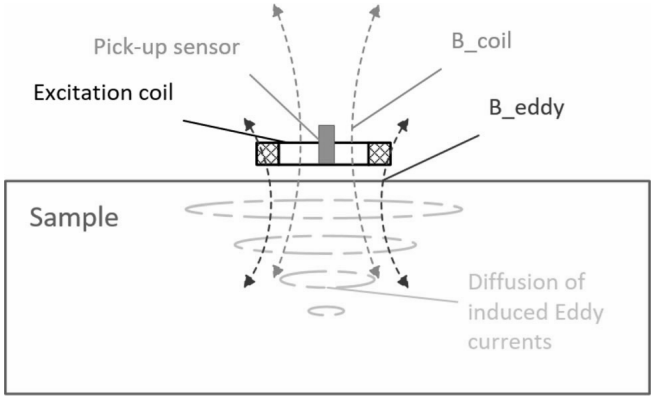

Fig. 3: Illustration of pulsed eddy current principle [15]

For the PEC test, was employed a coil with the following features: outside diameter 33 mm, inside diameter 17 mm, height 20 mm, wire diameter 0,27 mm and 500 turns. There are two Giant Magneto-Resistance (GMR) sensors. One sensor is employed for reference signal measurement.

During the test, the coil was positioned at the three zones of interest (ND, ED and ID). The coil was excited from a square wave with amplitude $4V$, frequency 1 KHz and Duty Cycle 60%. For each zone the PEC test was applied 100 times in different positions resulting on 100 observations with 200 time-domain samples for each observation. The sampling interval was 5×10^{-6} seconds.

A. Signal Pre-processing

Signal feature extraction was performed by both, discrete Fourier transform (DFT) and discrete wavelet transform (DWT). The DFT transforms a discrete-time signal $x[n]$ into discrete frequency domain components $X[k]$ [10]. The DWT may be performed through the multi-resolution decomposition algorithm, where the input signal $x[k]$ is separated into high and low frequencies using quadrature mirror filters, as shown in Fig. 4. The low-pass $l[n]$ filters are associated with approximation coefficients $C_A[n]$ Equation 1, and the high-pass filters $h[n]$ are associated with detail coefficients $C_D[n]$ Equation 2 [11].

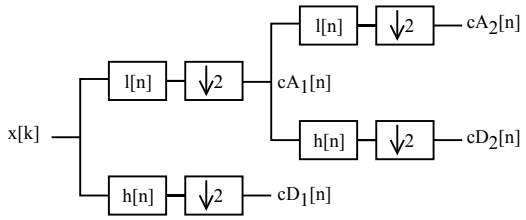

Fig. 4: Illustration of filter bank decomposition for DWT computing.

$$C_A[n] = \sum_k l[k-2n]x[k] \tag{1}$$

$$C_D[n] = \sum_k h[k-2n]x[k] \tag{2}$$

For features selection, Principal Components Analysis (PCA) was applied. PCA is a linear statistical signal processing method that transforms the multivariate input variables into orthogonal components. Besides that, the principal components are ordered according to his energy (variance), which makes PCA an optimum transformation for information compaction (in the mean square error sense) [12].

B. Artificial Neural Network

Artificial Neural Networks are algorithms inspired on the brain structure, which are composed of a set of artificial neurons [5]. A neuron can be mathematically represented by

$$y = \phi\left(\sum_{j=1}^{D} w_j x_j + b\right) \tag{3}$$

where x_j are the inputs, D is the dimension of the input signals, w_j are the synapses weights, b is the bias, $\phi(\cdot)$ is the activation function and y is the neuron output. The neurons can be grouped on layers resulting in the Multilayer Perceptron (MLP) structure. A classic method to train an ANN is the back-propagation algorithm, which uses gradient descent optimization to minimize the mean square error between target and output values.

For classification problems, back-propagation training considers all classes in the same way and is not possible to consider a specific class as priority. For some classification problems, including nondestructive evaluation, errors for different classes may have different costs. Considering the flaw detection in industrial pipes, incorrectly classifying a defect as no-defect represents a potential danger to the equipment and employees.

C. Particle Swarm Optimization

Particle Swarm Optimization (PSO) [9] is an algorithm based on the behavior of social animals. It initializes a population of M particles (candidate solutions) that are moving in a D-dimensional search space, by means adjusting its position $\mathbf{p}(h)$ and velocity $\boldsymbol{v}(h)$. The particle position update is performed by the Eq. 4:

$$\mathbf{p_m}(h+1) = \mathbf{p_m}(h) + \boldsymbol{v}_m(h+1), \tag{4}$$

where $\mathbf{p_m}(h)$ is the m-th particle position vector for iteration h and $\boldsymbol{v}_m(h+1)$ is the updated particle velocity, which may be computed from Eq. 5:

$$\begin{aligned}\boldsymbol{v}_m(h+1) = w(h) \times \boldsymbol{v}_m(h) + c_1 \times r_1 \times (\mathbf{p_{best}}(h) \\ -\mathbf{p_m}(h)) + c_2 \times r_2 \times (\mathbf{g_{best}}(h) - \mathbf{p_m}(h)),\end{aligned} \tag{5}$$

where $\boldsymbol{v}_m$ is the m-th particle speed, w is the inertia weight, c_1 and c_2 are the social and cognitive parameters, respectively, r_1 and r_2 are random numbers between 0 and 1, $\mathbf{p_{best}}$ is the best position the particle has been and $\mathbf{g_{best}}$ is the position of the best swarm particle.

D. Proposed Method for ANN Training

The proposed method for ANN training based on constrained PSO considers each particle as being composed from weights of an MLP with a single hidden layer, as illustrated in Fig. 5, where $x_i, i = 1, ..., j$ are the inputs, $\mathbf{W^{(1)}}$ and $\mathbf{W^{(2)}}$ are the synaptic weights matrices and y is the network output. The weights are randomly initialized as proposed by [16] and the hyperbolic tangent is used as activation function for all neurons. The used PSO parameters are shown in Table II.

TABLE II: Implemented PSO parameters

Topology	Global
Cognitive constant	2
Social constant	1
Swarm size	200
Acceleration constants	Random [0; 1]
Inertia limits	[0,4; 0,9]

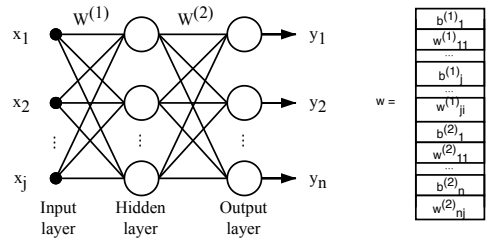

Fig. 5: MLP vector representation to be encoded as a PSO candidate solution.

At each PSO algorithm iteration the particles are evaluated by the corresponding MLP classification results. The constraint was defined to increase the probability of detection (P_D) for class ND while minimizing the errors for other classes. In this work, the PSO algorithm minimizes the fitness function (f_T), that evaluates for class c the classification error (PF_c) and constraints the probability of detection (PD_{ND}) for class ND to a minimum value λ:

$$\begin{aligned}\min f_T = \sum_{c=1}^{3} P_{Fc}(y), \\ \text{subject to} \begin{cases} PD_{ND} \geq \lambda, \end{cases}\end{aligned} \tag{6}$$

The proposed algorithm refines the candidate solutions by minimizing the fitness function, restricting the particles to satisfy the constraints, as shown in Fig. 6.

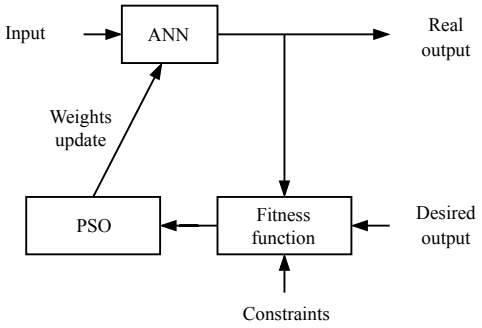

Fig. 6: Proposed constrained MLP PSO training flow chart.

III. Experimental Results

A way to summarize classification results is by using the confusion matrix, which is a square ($NC \times NC$) matrix $\mathbf{C} = \{c_{ij}\}$, where the columns represent the real classes and the rows represent the predicted classes [17], and NC is the number of classes of interest. In this work, for classifiers comparison using a single index, it was used the Efficiencies Product (EP), which is calculated by geometric mean [18] of all elements from main diagonal of confusion matrix as a performance measure.

$$EP = \left(c_{11} \times c_{22} \times c_{33}\right)^{1/3} \tag{7}$$

where c_{11}, c_{22} and c_{33} are elements from $\mathbf{C}$ main diagonal.

The DFT was applied to signal resulting on 100 frequency-domain coefficients, in this work it was considered the amplitude spectrum. For the DWT, it was employed function Daubechies 4 mother-wavelet function and three decomposition levels resulting on 31 approximation coefficients. This configuration was chosen experimentally after exhaustive testing different mother-wavelet functions and decomposition levels.

The typical time-domain signals for the classes of interest are shown in Fig. 7. Fig. 8 illustrates frequency spectrum obtained with DFT. The approximation coefficients obtained with DWT are shown in Fig. 9. It is important to note that most of the acquired signals present significant variation from those typical patterns.

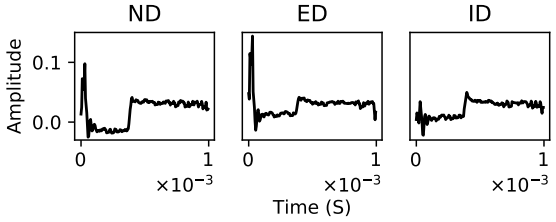

Fig. 7: Typical time-domain signals for the classes of interest.

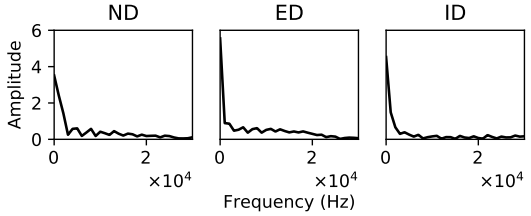

Fig. 8: Typical frequency spectrum signals for the classes of interest obtained by DFT.

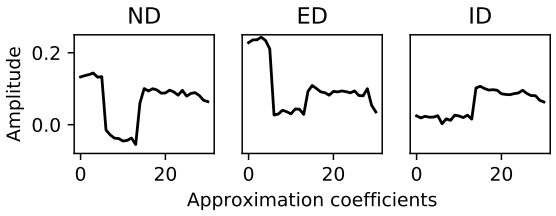

Fig. 9: Typical approximation coefficients obtained by DWT for the classes of interest.

At the second pre-processing stage, PCA compaction was applied to retain 90% of total variance. From the PCA over DFT coefficients were selected 13 principal components, from the PCA applied to DWT were selected 14 principal components.

The ANN used in this work has a single hidden layer. The input layer size match with the input vector size. The output layer has three neurons matching with the classes number. In order to select the hidden layer size the hold-out cross-validation method was employer.

Cross-validation is a widely used methods to model selection of a neural classifier. For the classifiers training the dataset was randomly split according to the hold-out cross-validation method [19] into 50% for training set and 50% for testing set. The training process was restarted for 10 different random selection of training and testing sets. Ten hidden layer sizes were valued, 2, 4, 6, 8, 10, 12, 14, 16, 18 and 20 neurons.

The Fig. 10 shows the mean and standard deviation of scores for the network trained with data without pre-processing (time-domain information). The same process was employed for all pre-processing cases, concluding that the best hidden layer size is 10 neurons for all cases.

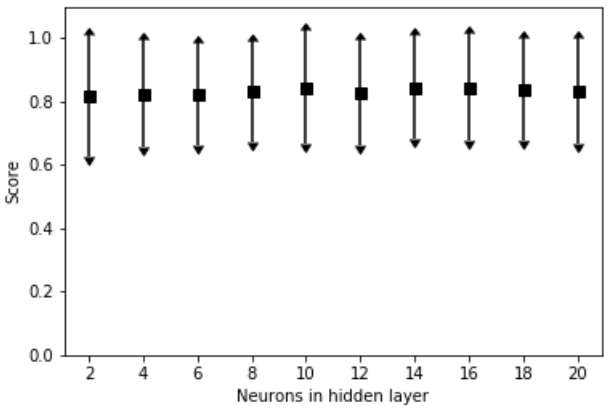

Fig. 10: Error bar graphic of scores for the network trained with data without pre-processing (time-domain information). The horizontal axis denotes the number of neurons in hidden layer.

Then the classifiers were trained with 10 neurons in the hidden layer. For the classifiers training, the dataset was randomly split according to the hold-out cross-validation method [19] into 50% for training set and 50% for testing set. The training process was restarted 10 times, for 10 different random selection of training and testing sets (resulting in 100 initializations).

After the MLP neural networks training process, the best classifiers were selected. From Table III, confusion matrix for network trained with time-domain information, it may be observed that the proposed method achieved higher value of EP and also higher detection efficiency for the no-defect case.

The proposed method presented higher EP values when compared to the traditional back-propagation algorithm for networks trained with pre-processed data by both methods DWT (Table IV) and DFT (Table V).

After PCA compaction, the classifiers trained from the constrained PSO method presented also higher EP values for both case, DWT+PCA (Table VI) and DFT+PCA (Table VII).

As a summary of the obtained results, it was observed that all classifiers trained from PSO presented EP higher than

TABLE III: Confusion matrix (in %) for data without pre-processing (time-domain information).

		ND	ED	ID
MLP-BP	ND	90	2	2
	ED	4	66	18
EP = 78.0	ID	6	32	80
MLP-PSO	ND	98	4	8
	ED	2	84	4
EP = 89.8	ID	0	12	88

TABLE IV: Confusion matrix (in %) for data pre-processed by DWT.

		ND	ED	ID
MLP-BP	ND	96	12	2
	ED	2	76	8
EP = 86.9	ID	2	12	90
MLP-PSO	ND	92	4	4
	ED	6	88	6
EP = 90.0	ID	2	8	90

TABLE V: Confusion matrix (in %) for data pre-processed by DFT.

		ND	ED	ID
MLP-BP	ND	98	8	16
	ED	0	74	4
EP = 83.4	ID	2	18	80
MLP-PSO	ND	96	2	6
	ED	0	92	8
EP = 91.2	ID	4	6	86

TABLE VI: Confusion matrix (in %) for data pre-processed by DWT+PCA.

		ND	ED	ID
MLP-BP	ND	80	6	0
	ED	6	68	20
EP = 75.8	ID	14	26	80
MLP-PSO	ND	88	4	8
	ED	2	90	20
EP = 82.9	ID	10	6	72

TABLE VII: Confusion matrix (in %) for data pre-processed by DFT+PCA.

		ND	ED	ID
MLP-BP	ND	84	2	2
	ED	8	86	14
EP = 84.7	ID	8	12	84
MLP-PSO	ND	90	6	6
	ED	4	88	6
EP = 88.7	ID	6	6	88

equivalent ones trained from back-propagation. The biggest difference was achieved for MLP trained from data without pre-processing, where the constrained PSO classifier presented EP 11.8 percentage points higher than the network trained by back-propagation. The smallest difference occurred for networks trained for data pre-processed from DWT (EP 3.1 percentage points higher).

The network which presented the highest EP was trained by PSO for data pre-processed from DFT (Table V), resulting on EP$= 91.2\%$ and a probability of detection of $PD_{ND} = 96\%$ for no-defect signatures. For this case, the total confusion by classifying defects as no-defect $ConfND$, that is the sum of elements $c_{1,2}$ and $c_{1,3}$ from confusion matrix, was 8%. For the back-propagation trained classifier with equal PD_{ND} (DWT pre-processing case) (Table IV), 14% of such confusion was observed.

IV. Conclusions

Multiclass classification problems may present different classification error costs for the classes of interest. This is generally the case for non-destructive evaluation decision support systems, in which not detecting a defect is more serious than incorrectly classifying a non-defective equipment as if it presents some kind of defect. In this work, a novel MLP training method was proposed, using PSO to deal with operational constraints during the classifier design stage. The proposed constrained PSO method was effective for training neural networks for classification of pulsed eddy-current evaluation data. Different signal pre-processing chains were considered in order to obtain higher discrimination efficiencies. It was observed that the combination of constrained PSO training and signal pre-processing from DFT provided high-discrimination efficiency for no-defect signatures, while minimizing the classification errors for defects.

Acknowledgements

This study was financed in part by the Coordination of Superior Level Staff Improvement - Brazil (CAPES) - Finance Code 001. The authors are also thankful to CNPq and FAPESB for the financial support.

References

[1] K. S. Rao, S. Mahadevan, B. P. C. Rao, and S. Thirunavukkarasu, "A New Approach to Increase the Subsurface Flaw Detection Capability of Pulsed Eddy Current Technique," *Measurement*, vol. 128, pp. 516–526, 2018.

[2] G. Y. Tian, Y. He, I. Adewale, and A. Simm, "Research on Spectral Response of Pulsed Eddy Current and nde Applications," *Sensors and Actuators A: Physical*, vol. 189, pp. 313–320, 2013.

[3] L. Xie, B. Gao, G. Tian, J. Tan, B. Feng, and Y. Yin, "Coupling pulse eddy current sensor for deeper defects ndt," *Sensors and Actuators A: Physical*, vol. 293, pp. 189 – 199, 2019.

[4] S. Winnik, *Corrosion Under Insulation (CUI) Guidelines: Revised*. Woodhead Publishing, 2015.

[5] S. O. Haykin *et al.*, *Neural Networks and Learning Machines*. New York: Prentice Hall,, 2009.

[6] X. Fu, C. Zhang, X. Peng, L. Jian, and Z. Liu, "Towards end-to-end pulsed eddy current classification and regression with cnn," in *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–5, May 2019.

[7] A. Habibalahi, M. Dashtbani Moghari, K. Samadian, S. S. Mousavi, and M. S. Safizadeh, "Improving pulse eddy current and ultrasonic testing stress measurement accuracy using neural network data fusion," *IET Science, Measurement Technology*, vol. 9, no. 4, pp. 514–521, 2015.

[8] A. L. Paul and P. C. Byrne, "An efficient learning algorithm for the backpropagation artificial neural network," in *IEEE Proceedings on Southeastcon*, pp. 61–63, April 1990.

[9] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," in *IEEE international conference on neural networks*, pp. 1942–1948, IEEE, 1995.

[10] B. P. Lathi, *Linear Systems and Signals*. Oxford University Press, Inc, 2 ed., 2005.

[11] A. Ukil and A. Barlocher, "Implementation of Discrete Wavelet Transform for Embedded Applications Using tms320vc5510," in *2007 International Symposium on Industrial Embedded Systems*, pp. 357–360, IEEE, 2007.

[12] A. Tharwat, "Principal component analysis-a tutorial," *International Journal of Applied Pattern Recognition*, vol. 3, no. 3, pp. 197–240, 2016.

[13] C. B. Larocca, C. T. Farias, E. F. Simas Filho, and I. C. Silva, "Wall Thinning Characterization of Composite Reinforced Steel Tube Using Frequency-Domain PEC Technique and Neural Networks," *Journal of Nondestructive Evaluation*, vol. 37, no. 3, p. 44, 2018.

[14] J. García-Martín, J. Gómez-Gil, and E. Vázquez-Sánchez, "Non-Destructive Techniques Based on Eddy Current Testing," *Sensors*, vol. 11, no. 3, pp. 2525–2565, 2011.

[15] A. Sophian, G. Tian, and M. Fan, "Pulsed Eddy Current Non-Destructive Testing and Evaluation: A Review," *Chinese Journal of Mechanical Engineering*, vol. 30, no. 3, p. 500, 2017.

[16] D. Nguyen and B. Widrow, "Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights," in *IJCNN International Joint Conference on Neural Networks*, pp. 21–26, IEEE, 1990.

[17] K. M. Ting, *Confusion Matrix*, pp. 260–260. Boston, MA: Springer US, 2017.

[18] S. Lahmiri, D. A. Dawson, and A. Shmuel, "Performance of Machine Learning Methods in Diagnosing Parkinson's Disease Based on Dysphonia Measures," *Biomedical engineering letters*, vol. 8, no. 1, pp. 29–39, 2018.

[19] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, vol. 2. John Wiley and Sons, 2014.