Computers and Electrical Engineering 71 (2018) 440–451

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A reliable approach for detection of incipient faults of short-circuits in induction generators using machine learning☆

Pedro Pedrosa Rebouças Filho*[,a], Navar M.M. Nascimento[a], Igor R. Sousa[a], Cláudio M.S. Medeiros[a], Victor Hugo C. de Albuquerque[b]

[a] Laboratório de Processamento de Imagens, Sinais e Computação Aplicada, Instituto Federal do Ceará, Av. Treze de Maio, 2081, Benfica, Fortaleza, Ceará 60040-531, Brazil

[b] Programa de Pós-Graduação em Informática Aplicada, Universidade de Fortaleza, Fortaleza, Ceará, Brazil

ARTICLE INFO

Keywords:
Fault detection
Induction generator
Machine learning
Neural networks
Reliability
Short-circuit
Wind turbine

ABSTRACT

This research contributes with a reliable approach to detect incipient stator winding inter-turn short-circuits in induction generators applied in wind turbines. Using a wind turbine test-bench, we inserted different types of short-circuit in the generator. The electrical current is acquired to build a fault database. We propose the use of four feature extraction techniques with three classifiers. The MLP identified 100% of the generator's Normal conditions with less than 1% false positives and negatives. Using different topologies of MLP, it was possible to identify incipient short-circuits in 1.41% turns with 99.33% accuracy. The combination Fourier-MLP is more useful for fault detection, since it obtained 84.48% of accuracy, with 99.98% of Normal conditions correctly classified.

1. Introduction

Among the renewable energy sources, wind energy has become the most effective and accepted solution for electricity generation worldwide, contributing 486.7 GW to global demand [1]. This energy production represents only 3% of the world's energy demand, but it is estimated that in 2030 the wind power will be able to supply 17–19% of the global demand [1].

Despite the growing exploitation of this energy source at 17% per year [1] the technologies are not consolidated, there are still engineering and science challenges to be solved to support this expansion.

Operational problems directly impact on the cost of energy, according to Polinder et al. [2] only with a reliable and available wind turbine system that the cost of energy can be mitigated. Moreover, in the light of reliability are the maintenance operating costs, which account for up to 30% of the cost of energy [2].

Hahn, Durstewitz and Michael have already evidenced the concern with maintenance in wind farms [3]. They exhibited records of fault types in wind turbines. The compilation of data from a set of wind farms installed in Europe has shown that the most costly faulty component for the wind farm is the electric generator.

Among the areas of study surrounding wind turbines, the present text focuses on electric generators, especially the Squirrel Cage Induction Generator (SCIG). The importance of the SCIG is based on its robustness, consolidated technology and future trends. Also, Yaramasu et al. [4] forecasts this generator will dominate the market of wind turbine in the next years.

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. G. R. Gonzalez.

* Corresponding author.

E-mail addresses: navarmedeiros@ppger.ifce.edu.br (N.M.M. Nascimento), igor.sousa@ppger.ifce.edu.br (I.R. Sousa), claudiosa@ifce.edu.br (C.M.S. Medeiros), victor.albuquerque@unifor.br (V.H.C. de Albuquerque).

https://doi.org/10.1016/j.compeleceng.2018.07.046
Received 3 April 2018; Received in revised form 9 July 2018; Accepted 31 July 2018
0045-7906/

Despite the versatility of SCIG, it is not immune to faults, and has limitations. In general terms, these faults are associated with: overheating, electrical', dynamic and mechanical effects [5]. Bonnet and Soukup [5] emphasizes that the short-circuit between turns is the most incipient fault and the most difficult to detect.

So, this work focus on detecting incipient short-circuits in SCIG, to provide a means to improve wind turbine reliability. Moreover, this could be potentially good for giant wind turbines. Because, detecting an incipient short-circuit before total degradation could guide corrective maintenance to make repairs on site, without having to remove the generator from the wind turbine and decrease the machine availability.

Since wind turbines are a sensor-based equipment, monitored and controlled by a Supervisory Control and Data Acquisition (SCADA) system, the incipient short-circuit detector could be deployed inside the hardware responsible for data acquisition. So, one could feasibly replicate the methodology of this paper in real wind turbine systems

This paper is organized as follows: in Section 2 state of the art in fault detection for reliability improvement in electrical machines is presented; in Section 3 the feature extractions methods used in this work are explained; in Section 4 the procedures to emulate the fault emulation, acquire data and create a dataset for pattern recognition are presented; in Section 5 the machine learning methods used in this work are described, as well as its configurations and metrics for evaluations; in Section 6 our results are analyzed and conclusions presented in Section 7.

2. State of the art in short-circuit faults in electrical machines

Stator winding faults are usually associated with a variety of factors, such as, overheating, electrical overload, electrical discharges and mechanical stress [5]. The environment where the SCIG is installed is harsh, so a monitoring tool is required

The Current Signature Analysis (MCSA) technique was identified as a promising mean to describe induction generator operational conditions. It consists of a set of consolidated methods, which developed between the years of 1975 and 1985 in surveys around the world that had in common the analysis of the current and its spectrum to characterize induction machines [6].

In [7], a neural classifier is trained with the motor current, the supply voltage and the rotor speed. The authors use a single-phase, 0.5 hp motor running at 50 Hz unloaded. They were considered satisfactory. However, the problem is limited to only one motor configuration (single phase, no load, operating at 50 Hz) and requires a speed sensor along with the already installed current sensors.

Bouzid and Champenois [8] worked on a computational model of a 1.1 kW induction motor under stator winding failure. The authors used neural networks with current analysis and were possible to identify 100% of the modeled problems. However, the simulating contemplated the machine operating only under 50 Hz.

Palácios et al. [9] acquired current signals in a one *hp* motor, with multiple load conditions and varying the percentage of turn under short-circuit (3%, 5% and 10%). The authors achieved results from 75% to 98% of accuracy for binary classifications between normal or fault conditions.

Oliveira et al. [10] use a single current sensor, without speed sensor, and 1.4% of turns under short-circuit, in an electric induction machine, can be identified with 67% of accuracy using neural networks. The authors proposed the use of the MCSA technique, combined with Fourier, from the frequency spectrum theory of Thomson and Fenger [6]. Vieira et al. [11] continued these studies and was able to develop an embedded system to detect these failures, based on the Multi-layer Perceptron (MLP).

The members, of the same research group where this work was developed, have experience on detection of incipient short-circuits in induction motors. This was consolidated during years and by theirs publications: Oliveira et al. [10], Coelho et al. [12], Vieira et al. [11].

The literature review exhibited most of the researches were concerned in identifying failure under certain operation conditions, either by frequency or load and without much description of the fault emulations. The approach proposed in our research was built upon the need for a more generalist method, wherein an expert system could detect failure under different operational points. Moreover, also, triggered by the fast expansion of wind farms in the region of Ceará, Brazil, this present work addresses the problem to detect incipient faults in generators, since it is unprecedented for both research group and region.

3. Feature extraction methods used for induction machines

The Fourier Transform can be used to express time-domain signals into a frequency domain. Oliveira et al. [10] investigated the harmonics in electrical current signal from electrical motor under several inter-turn short-circuits faults and concluded the most important frequencies are $0.5f_n$, $1.5f_n$, $2.5f_n$, $3.0f_n$, $5.0f_n$ and $7f_n$, wherein f_n is the stator's fundamental frequency. Oliveira et al. [10], Coelho et al. [12] and Vieira et al. [11], Vieira et al. [11] also used the same frequencies to purpose different approaches for fault detection in induction motors. The studies considered the electrical machine operating as a motor, but we believe the same frequencies will be representative while the machine works as a generator. Thus, the values of these frequencies, normalized by its fundamental, are going to be used as features.

The Higher-Order Statistics is a promising technique for describing time domain signals [13]. The usage of Kurtosis was first proposed by Dwyer [14] as a statistical tool to indicate non-Gaussian components in signals. This theory is under the High-Order Statistics basis and described it as the fourth statistical moment. However, Antoni [15] highlights the capacity and efficiency of characterizing non-stationary signals. The advantage of HOS for signals indicated by Mendel [13], and it relies on the robustness (i.e., filtering) to Gaussian noise when using moment higher than second order. The features extracted from HOS are going to be skewness and kurtosis, alongside with variance and rms values.

The Structural Co-occurrence Matrix consists of a method based on co-occurrence statistics and is directed to the structural

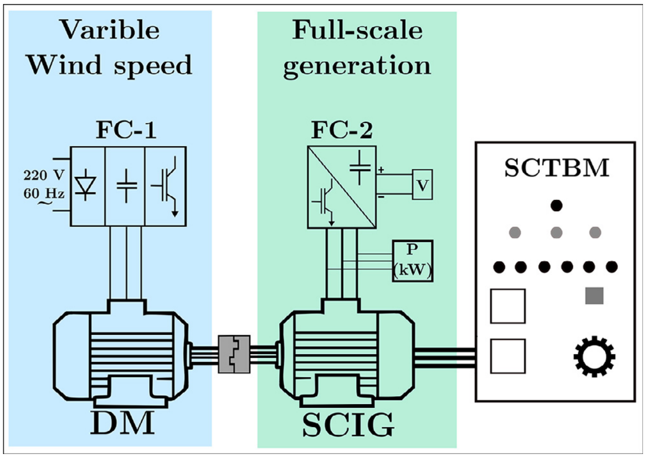


Fig. 1. The diagram contains the wind turbine test bench, based on the Full-scale e Full-Variable-Speed.

analysis of discrete signals, admitting existing connections between low-level structures of two discrete signals in n-dimensions [16]. Its main characteristic is to introduce a prior knowledge about the analyzed signals, emphasizing the detection of details. Its output is a two-dimensional histogram, wherein SCM provides the co-occurrences between structures of the input signals [16]. The authors proposed 6 features to be extracted from any signal and its calculations are shown in its paper. Among the feature extraction methods used in this paper, the SCM is the most recent one and has demonstrated being able to recognize patterns in details from signals [17].

4. Experimental setup and dataset creation

The wind turbine emulation is based on the configuration described by Yaramasu et al. [4], which consists of a squirrel cage induction generator, of the type Full-scale e Full-Variable-Speed, which means that the electric machine can generate electrical energy throughout the entire speed range. A test bench based on this wind turbine system was developed for this work and is shown in Fig. 1.

The Drive Machine (DM) emulates the transformation of kinetic energy from the wind in rotating motion to the shaft of the generator, in allusion to the blades of the wind turbine. The frequency converter FC-1 emulates the variable wind speed that powers the DM. FC-2 only feeds the generator's stator. When the frequency set in FC-1 is higher than in FC-2, the kinetic energy is converted into electrical energy, owing through the SCIG to the dc-bus of FC-2.

To perform the experiments we used a 4-pole induction generator, with mechanical power of 1 hp, electrically connected in delta configuration for power supply voltage of 220 V, and its rated current is 3 A. Both generator and drive machine are powered by a WEG CFW-08 frequency converter. The DM is mechanically coupled to the generator, and we used an induction motor with the same characteristics and wiring.

A Short-circuit Test Board conducts the experiments in Machines (SCTBM). There is an electrical circuit, which function is to emulate connections for the short-circuit between turns of one coil stator. It also has an acquisition and data sending module to acquire the electrical current of the electric machine.

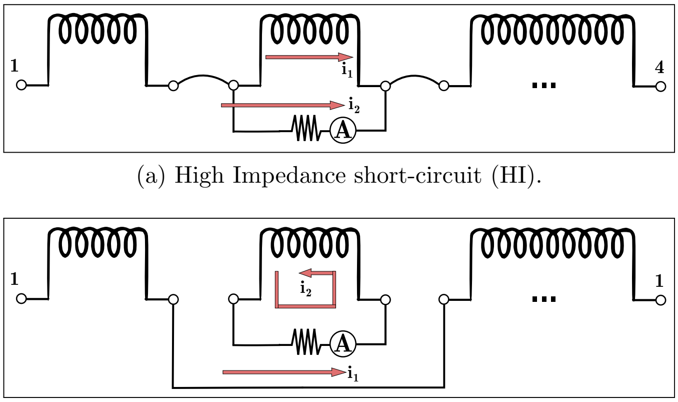
4.1. Modifications on generator for fault emulation

The inter-turn short-circuit is the most primary fault state in stator winding [5]. It is mostly neglected, since it usually does not trigger instruments or circuit protections, but it might reach phase-to-phase or phase-to-ground short-circuit, which is usually destructive to the electrical machine [5].

The electrical machine used to acquire operational data is the same as used by Oliveira et al. [10] and Vieira et al. [11]. This machine was properly prepared to enable emulation of stator winding inter-turn short-circuits.

A diagram of the types of short-circuits is exhibited in Fig. 2. The high impedance (HI), Fig. 2a, is a parallel path created for the electric current to flow. This condition emulates the most incipient short-circuit, wherein the insulation is starting to decay until, reaches the low impedance (LI) case, Fig. 2b. The LI is a condition before total degradation, wherein an amount of turns are under short-circuits. This is emulated by removing coils turns from the circuit, but keeping them immersed in the electromagnetic field. In both cases, the short-circuit current is limited to its rated value, through external resistor, to protect our generator during experiments. So, all of our emulated faults being incipient, even the LI short-circuit, because in a real case scenario the current might exceed the rated value.

There is always a question: how much time does it take for the inter-turn short-circuit degrade insulation until the total failure? This question has been addressed for researches as, Wang and Butler [18], Neti and Wilhelm [19], Tshiloz et al. [20] over the years, and there still not a general sense or answer. However, a prevailing opinion of manufacturers and users is low voltage machines ($\leq$ 1000 V) resist for more extended periods on short-circuit than higher voltage machines [21].



(a) High Impedance short-circuit (HI).

(b) Low Impedance short-circuit (LI).

Fig. 2. Electrical diagram of the modifications for stator inter-turn short-circuit emulation.

The evolution of the incipient short-circuit might take dozens of seconds, minutes or hours. This decay rate indeed is a function of the number of turns under failure. So, we purpose a methodology to focus on the first occurrences of the failure, where there is enough time for the expert system detects before complete degradation of the generator.

All short-circuits are applied in just one of the phases during experiments, in the following amount of total's stator winding: 1.41%, 4.81%, and 9.26%. Combining intensities and types of short-circuiting, we arrived at six different types of fault. In the following text, HI might be referred to as HI-1, HI-2, and HI-3, that means HI with 1.41%, with 4.81% and with 9.26% of short-circuited winding respectively and the same is valid to LI faults. With this arrange it is possible to emulate the evolution of short-circuiting from its most incipient stage (LI-1) until a more severe condition for the insulation (HI-3), but also incipient. This was based on a study made by Bonnet and Soukup [5].

4.2. Experiments

The experimental setup is shown in Fig. 3. A data acquisition system of National Instruments, composed by the NI-USB6009 module programmed to read a 10 s signal, sampled at 5 kHz, with 14-bits resolution and a LabVIEW interface running in a microcomputer was used to acquire the signal, from the three current sensors. This process resulted in 1356 acquisitions, being 248 Normal and 1108 Fault conditions among 6 different types of faults. In all experiments, the values of the frequency applied into the generator's stator (f_g) are registered, as well as the voltage in the dc-bus of the frequency converter in the generator side (V_{dc}) and power generated (P). The f_g ranges from 43.65 to 59.27 Hz, V_{dc} from 210 to 380 V.

The next step is to apply feature extraction methods to retrieve useful information from the electrical current from phase R of the generator.

4.3. Dataset creation

To compose our datasets we used three different feature extraction methods: one time-domain method (i) the Higher-Order Statistics (HOS) and two frequency-domain methods (ii) the Fourier Transform and (iii) the Structural Co-occurrence Matrix (SCM). The features used by each one of the methods are described in Section 3. The number of features summarized in Table 1. Additionally to each one of the methods we purpose using the measurements of V_{dc}, f_g and I_{Rrms} as features, as well. In HOS one of the features is the rms value of the current signal, so in this method the only additional features are dc-bus and generator frequency.

All datasets[1] have seven classes, being one Normal and six Faults. The number of samples are also the same, 1356, and they split into classes as follows: Normal conditions has 248 samples, Fault HI-1 has 203, HI-2 has 179, HI-3 has 183, LI-1 has 177, LI-2 has 208 and LI-3 has 158 samples.

5. Machine learning methods adopted

With the datasets in our hand, the trained classifiers are Multi-layer Perceptron, Support Vector Machines, and Bayesian. First, preliminary studies to tune hyperparameters in each model are made, to the performance discussion be conclusive and valid.

The Bayesian classifier is considered a statistical technique, being applied in the classification of samples according to the probability density function of each one of these samples belonging to a particular class. Its learning is supervised and based on Bayes Decision Theory. Bayesian Classifier distributes the samples according to the value of the posterior probability, calculated from values of the conditional densities and prior probabilities [22]. The inspiration for use this classifiers relies on the assumption that variables

[1] Databases are available at https://github.com/navarmn/Wind_turbine_failure_prediction.

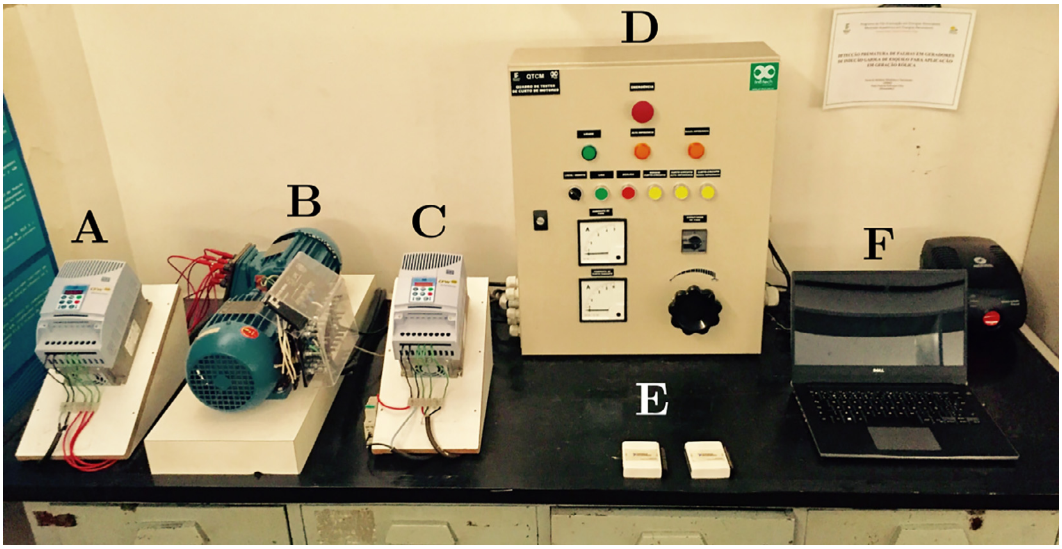


Fig. 3. The experimental setup. **A** is the frequency converter in wind size, **B** is the coupled DM and SCIG, **C** is the frequency converter in generator side, **D** is the SCTBM, **E** is the NI-USB6009 and **F** is the microcomputer which runs LabView.

Table 1
Summary of datasets composed by the feature extraction methods.

Feature extraction	Number of features	Additional features	Total of features
Fourier	6	V_{dc}, f_g and I_{Rrms}	9
HOS	4	V_{dc} and f_g	6
SCM	6	V_{dc}, f_g and I_{Rrms}	9

are random and could be modeled by a Gaussian probability density function, we even consider non-random signals as randomly behavior. The number of samples in all datasets is much higher than its dimensionality, so we used the independent estimations of the covariance matrix for each class and, this gives us a quadratic Bayesian classifier.

The Multi-layer Perceptron is a biologically inspired neural network, which consists in a sensitive input layer, a processing hidden layer and an output layer, representing a non-linear mapping in an input vector and an output vector [23]. Its basic unit, commonly called neurons, are in fact Perceptrons, a non-linear mathematical and computational model for a neuron, which is all interconnected layer by layer through numerical weights. The learning process occurs in a training phase, where weights are adjusted recursively based on the backpropagation of the error. Non-linear functions are imposed on hidden neurons as a mean to solve non-linearly separable problems [23]. The premise for using neural networks is different from the Bayesian classifiers, and it is based on the capability to perform a non-linear mapping between the input/output pair pattern.

In all databases, The MLP was trained using gradient descent algorithm, and the following hyperparameters are defined through a grid search: number of hidden neurons varying from 1 to 50; initial learning rate from 0.001 to 0.5; momentum rate from 0.1 to 0.5 and number of epochs from 100 to 5000. It was used one hidden layer and hyperbolic tangent was chosen to be the activation function of all neurons. After a preliminary investigation the number of hidden units of the network were 10 at the database formed by Fourier, 13 at HOS and 9 at SCM.

Support Vector Machines (SVM) are based on the Statistical Learning Theory created by Cortes and Vapnik [24]. This technique has as primary objective to determine classes with boundaries that increase the distance between them. Originally, SVM was designed to solve binary problems, however, this technique becomes complicated when used in multi-class problems and, therefore, approaches such as one-versus-one e one-versus-all are examples of variations of this method for this purpose. The kernel machines are versatile since it combines statical theory with non-linear projection to a new feature space where classes could be easier classified. It is a different paradigm than Bayesian and MLP classifiers. The SVM configuration is found automatically using the polynomial kernel.

The hold-out validation on the datasets was performed, but since the number of samples per classes is unbalanced, a few precautions were taken care. The class with fewer samples was used as a reference, and 80% of its samples were used to establish the number of samples per classes in the training set. We arrived in a training set with 882 samples, being 126 per class (65% of total samples). This processed should avoid bias during training. The test set has 474 samples, unbalanced between classes.

The metrics used to access performance of all classifiers are Accuracy (Acc), Specificity (Spe), Sensitivity (Sen) and F-Score (Fsc). The average confusion matrix will be presented and discussed to identify where the misclassifications occur. To determinate, the classifier robustness the false positive and Negative rates are going to be analyzed, since they represent Normal conditions classified as Fault and Fault conditions classified as Normal, respectively.

Table 2
Results from average metrics of overall classes: Accuracy (Acc), Sensitivity (Sen), Specificity (Spc), F-score (Fsc) obtained from a extractors Fourier, HOS and SCM, combined with classifiers MLP, SVM and Bayes.

Classifier	Acc (%)	Sen (%)	Spc (%)	Fsc (%)
Fourier				
MLP	**84.48 ± 2.65**	84.48 ± 2.65	97.01 ± 0.59	84.48 ± 2.65
SVM	83.94 ± 1.06	51.31 ± 3.58	91.42 ± 0.59	55.04 ± 2.57
Bayes	77.30 ± 0.63	77.47 ± 1.13	95.33 ± 0.16	77.39 ± 0.81
HOS				
MLP	83.54 ± 1.42	83.54 ± 2.31	96.82 ± 0.34	83.54 ± 1.57
SVM	**89.60 ± 0.52**	76.79 ± 1.35	95.43 ± 0.34	77.08 ± 1.24
Bayes	78.96 ± 0.64	78.96 ± 0.64	95.75 ± 0.16	78.96 ± 0.64
SCM				
MLP	40.93 ± 5.58	40.93 ± 7.97	80.61 ± 3.56	40.93 ± 4.75
SVM	**89.14 ± 0.74**	64.53 ± 2.52	93.13 ± 0.53	62.87 ± 3.32
Bayes	54.97 ± 1.00	55.03 ± 1.05	88.00 ± 0.41	55.03 ± 1.00

6. Results and discussions

In this section is discussed the results from all combinations feature extraction methods versus classifiers, by its end the better configuration is shown and the second part concerns different propositions to achieve both accuracy and robustness of the classifier to identify incipient inter-turn stator winding short-circuit.

All results presented are from the test set. The results on training set are not exhibited since they are very similar to the test set. The programming routines of feature extraction, training and classifications are implemented on MATLAB®. Note that all experiments reported here used a PC Intel i7 running at 3.1 GHz and 8Gb of RAM on a Linux Ubuntu operating system installed in a solid-state drive.

6.1. Comparison of feature extraction methods and classifiers

In Table 2 it is exhibited the average metrics overall classes from all combinations feature extraction method versus classifiers. Using Fourier as a feature extractor, MLP achieved the average accuracy of 84.48% among all classes, which is better than Bayes and slightly than SVM. The sensitivity indicates MLP was far better than other methods while detecting true positive cases, which might indicate a capability to identify Normal operational conditions of the generator. The specificity also indicates MLP better than the other classifiers while making predictions of faulty conditions.

The results in Table 2 indicates that combining HOS with SVM provide better results than MLP and Bayes because achieved 89.6% of accuracy in all classes. However, the sensibility and specificity of MLP still higher than Bayes and SVM, this indicates MLP remain better while detecting Normal conditions among other classes. The dataset composed using SCM has shown results below 50% in MLP, which is no better than a random prediction. The same for Bayes, which is slightly better. Both methods might not be suitable for SCM extractor. However, SVM provided results much better than both, achieving 89.14% of accuracy.

The analysis indicates MLP alongside Fourier might be a more robust approach to detect stator winding inter-turn short-circuits in SCIG because its results are more stable than other methods as well as its false positive and negative rates. However, to affirm this, the results exhibited in Table 3 and the confusion matrix shown in Table 4 are analyzed.

The highlight of non-fault detection goes to the combination of MLP-Fourier and Bayes-Fourier, which achieved 99.8% and 99.14% respectively. Despite the better overall accuracy of the combination SVM-HOS, the Normal conditions were identified, with only 81.17% and it is not suitable for this work, which is to identify faulty conditions it is essential the classifier identify appropriately Normal operational conditions of the generator.

The most critical fault, LI-3 is properly identified by MLP-Fourier in all training. The other faults are correctly classified in the decreasing order: LI-2, HI-3, HI-1, HI-2 and LI-1. In the confusion matrix, Table 4, it is possible to see fault HI-1 is mostly misclassified as LI-1, because the low amount of turns in the stator winding under short-circuit, 1.41%, is not enough to provide significant differences in the current spectra between high and low impedance types of fault. The robustness comes with the 1% of samples classified as Normal conditions since the faulty HI-1 tend to be similar to the Normal operational condition. Although the labeling of high and low impedance short-circuits, all Faulty classes are incipient, as explained in Section 4.1, so a real-case scenario is going to be easily detected.

Similar patterns occur in other groups of faults, HI-2 is mostly confused with LI-1, HI-1, and LI-3. LI-1 misclassified as HI-1, HI-2, HI-3, LI-2, and LI-3. However, faults HI-3, LI-2, and LI-3 have a higher accuracy rate, because the number of turns under short-circuit grows alongside its discrepancy from normal conditions and other classes. Only 4% of the samples in faulty conditions were misclassified as Normal, which indicates the misclassifications by MLP tend to be between faulty classes and not with Normal conditions.

Regarding time analysis our results indicated Fourier took 2.48 ms to extract feature from the current signal and MLP takes an average of 0.034 ms to classify one sample. Both feature extraction time and classification time are around 2.5 ms, and it can be unconsidered for our system, since most of the time is spent to acquire the electrical current. For our experiments we used as standard 10 s acquisition at 5 kHz to guarantee a 0.1 Hz/sample of resolution in the frequency spectrum. The total time for our system to

Table 3
Results of an average accuracy rate comparison between different generator operational states (Classes), over the features extraction methods Fourier, HOS, and SCM.

Class	Fourier		
	MLP	SVM	Bayes
Normal	**99.98 ± 0.11**	80.87 ± 2.27	99.14 ± 0.33
HI-1	73.51 ± 12.26	76.34 ± 1.82	62.94 ± 3.27
HI-2	66.00 ± 8.74	80.26 ± 0.75	54.11 ± 3.28
HI-3	94.05 ± 3.82	81.18 ± 1.88	86.71 ± 1.31
LI-1	58.94 ± 12.36	83.36 ± 2.99	39.49 ± 5.91
LI-2	98.89 ± 1.12	87.66 ± 1.29	92.10 ± 1.21
LI-3	100.00 ± 0.00	97.96 ± 0.17	100.00 ± 0.00
Class	HOS		
	MLP	SVM	Bayes
Normal	63.93 ± 1.24	81.17 ± 3.30	59.24 ± 4.49
HI-1	74.03 ± 2.60	72.04 ± 0.93	47.46 ± 5.40
HI-2	98.11 ± 1.39	78.10 ± 0.90	96.24 ± 0.77
HI-3	98.25 ± 1.22	88.71 ± 0.27	99.22 ± 0.52
LI-1	80.39 ± 1.71	84.16 ± 1.03	63.23 ± 2.54
LI-2	98.78 ± 0.55	98.20 ± 0.47	100.00 ± 0.00
LI-3	96.88 ± 0.33	99.20 ± 0.53	97.37 ± 0.67
Class	SCM		
	MLP	SVM	Bayes
Normal	51.59 ± 0.57	77.66 ± 2.12	49.03 ± 3.97
HI-1	31.75 ± 1.70	81.31 ± 1.92	73.51 ± 3.68
HI-2	50.00 ± 0.80	92.63 ± 1.55	10.38 ± .29
HI-3	46.03 ± 1.09	97.30 ± 0.63	50.12 ± 2.10
LI-1	69.84 ± 2.01	83.14 ± 2.68	17.03 ± 7.14
LI-2	64.29 ± 0.95	97.66 ± 0.96	90.35 ± 1.20
LI-3	90.48 ± 0.44	99.34 ± 0.35	92.87 ± 1.06

Table 4
Percentage average confusion matrix for **Fourier-MLP**. The highlights correspond to correct classifications and are also the accuracy rate of each class, while the other number are the misclassifications rates. The sum in each row correspond to 100%.

Labels	Predictions						
	Normal	HI-1	HI-2	HI-3	LI-1	LI-2	LI-3
Normal	**99.98%**	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%
HI-1	0.93%	**73.50%**	4.41%	0.49%	19.25%	0.85%	0.53%
HI-2	1.84%	7.44%	**66%**	6.38%	14.12%	3.25%	0.95%
HI-3	0.12%	0.82%	3.26%	**94.04%**	1.00%	0.68%	0.04%
LI-1	1.09%	28.11%	8.84%	1.26%	**58.93%**	1.04%	0.69%
LI-2	0.02%	0.05%	0.25%	0.49%	0.14%	**98.88%**	0.15%
LI-3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**100%**

analyze one sample is 10.0025 s.

The research conducted here is investigatory, in embedded applications hardware and software optimizations are required, as a trade-off with accuracy. However, Thomson and Fenger [25] affirms using MCSA we are capable of detecting incipient short-circuits. Thomson and Fenger [25] also performed several tests of short-circuits on the induction machine operating under rated voltage. Moreover, their results exhibit the lead time until the failure is an order of minutes since in our approach took 10.002 s to classify one sample, this is enough time for the expert system detect the failure in initial conditions, within the time before total degradation.

We believed the classifier robustness relies not only on its accuracy but on its capacity for identifying and reject samples that provide confusion, like false positives and false negative. Thus, to increase the classifier reliability, by reducing the number of faulty conditions misclassified as Normal, it is necessary to induce the classifier to learn differences between faulty and normal patterns. Also, those results imply faults should be better classified in groups, so the following experiments are performed by training the MLP with faults grouped by its type or/and intensity since this classifier obtained the best results.

Table 5
Results obtained by MLP after grouping faults by two types, HI and LI. It is exhibited the overall accuracy among all classes, as well as its independent rates. The confusion table is also shown.

Overall classes			
Acc (%)	95.15 ± 0.94		
Stratified by classes			
Class	Acc (%)		
Normal		100 ± 0.00	
HI		96.73 ± 2.71	
LI		88.71 ± 2.74	
		Confusion table	
Labels	Predictions		
	Normal	HI	LI
Normal	**100%**	0%	0%
HI	0.03%	**96.73%**	3.23%
LI	0.13%	11.15%	**88.71%**

6.2. Classifier reliability improvement

In Table 5 the average accuracy of three-class training is shown: normal motor, high impedance fault, and low impedance fault. In this experiment, the HI and LI classes do not have all the data, since it would unbalance the data set in relation to the normal class. The Fourier-MLP combination is also used as the feature extraction-classifier approach. The MLP achieved 95.15% overall set-up. It is interesting to note that the Normal class obtained 100% hit rate, there being no false positives. As the confusion matrix shows, the LI class is confused 11.5% of times with the class HI. One explanation is that HI class mix data from HI-1, HI-2 and HI-3 and as seen in Table 4 HI-1 and LI-1 are extremely confused. So the fact of the short-circuit current is limited at the generator rated current in order not to damage the machine, making the emulated LI short-circuits not much different from HI faults.

New training are performed, this time between the Normal classes, short-circuit in 1.41% of turns (LVL-1), in 4.81% (LVL-2) and 9.26% (LVL-3). The result is shown in Table 6, where the average accuracy and the confusion matrix are shown. Again, there is suppressed data to avoid unbalance of data relative to the Normal class, as in the case of the class of levels 1, 2 and 3. The overall accuracy increased to 97.12%, and it is notable that once again, the Normal class is classified correctly 100% of the time. It is induced to presume the MLP accuracy rate increase following the number of turns under short-circuit, because critical failures tend to show significant differences in the current spectra. However, the results indicate a different behavior by the classifier. Since the concerning

Table 6
Results after grouping faults by intensities, 1.41%, 4.81%, and 9.26%. It is exhibited the overall accuracy among all classes, as well as its independent rates. The confusion table is also shown.

Overall classes				
Acc (%)	**97.12 ± 0.97**			
Stratified by classes				
Class	Acc (%)			
Normal		100 ± 0.00		
LVL-1		99.33 ± 0.70		
LVL-2		91.22 ± 3.77		
LVL-3		97.92 ± 1.43		
Confusion table				
Labels	Predictions			
	Normal	LVL-1	LVL-2	LVL-3
Normal	**100%**	0%	0%	0%
LVL-1	0.11%	**99.33%**	0.48%	0.07%
LVL-2	1.56%	4.49%	**91.22%**	2.71%
LVL-3	0.06%	0.18%	1.83%	**97.91%**

Table 7
Summary of results, containing accuracy after grouping all faults in one class. It is exhibited the overall accuracy among all classes, as well as its independent rates. The confusion table is also shown.

Overall classes		
Acc (%)	**100 ± 0.00**	
Stratified by classes		
Class	Acc (%)	
Normal		100 ± 0.00
Fault		100 ± 0.00
Confusion table		
Labels	Predictions	
	Normal	Fault
Normal	**100%**	0%
Fault	0%	**100%**

of this works is to detect incipient faults, identifying an LVL-1 short-circuits are our primary goal. LVL-1 short-circuits is the most challenging fault to identify because a low number of turns are under fault. However, the MLP classified them correctly 99.3% of the times. It is a considerable improvement since in 7 class classification, the faults HI-1 and LI-1 had lower hit rates of 73.50% and 58.93% respectively, as seen in Table 4.

The LVL-3 achieved high hit rates, around 97%, but the intermediary condition LVL-2 presented the lower hit rate of all classes, only 91.22% on average. By the confusion table, it is possible to see 4.49% of the LVL-2 faults were classified as LVL-1, that occurs because HI-2 faults tend to be similar to LI-1 and HI-1, as seen in Table 4. Some of the samples are classified as LVL-3, especially by the confusion between LI-2 and HI-3, and a few less are classified as Normal.

Final training are performed with the intention of a binary classification only the data between Normal and Faults. For this training, 80% of data were used for training. Although classes are unbalanced in the number of samples (887 Failure samples and 199 Normal samples), this helps the network learn more about fault data, where most of the knowledge of the data bank is found. Thus, one can avoid cases of false positives. The result is displayed in Table 7. With 100% average accuracy, there is no presence of false positive or false negative. This result was expected, since the Tables 4–6 already indicate values similar to the average accuracy of the Normal class.

In Fig. 4, the output $y_o(t)$ of the best MLP trained with the binary data set on the dataset is presented. It is possible to observe that although 100% accuracy has been obtained, some samples tend to be misclassified, such as the samples marked in the graphs. To

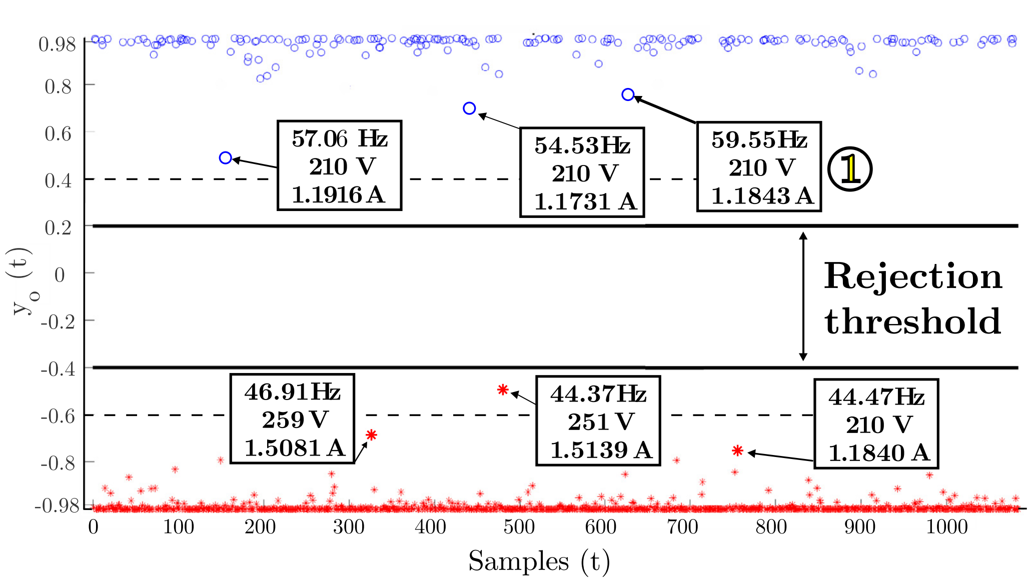
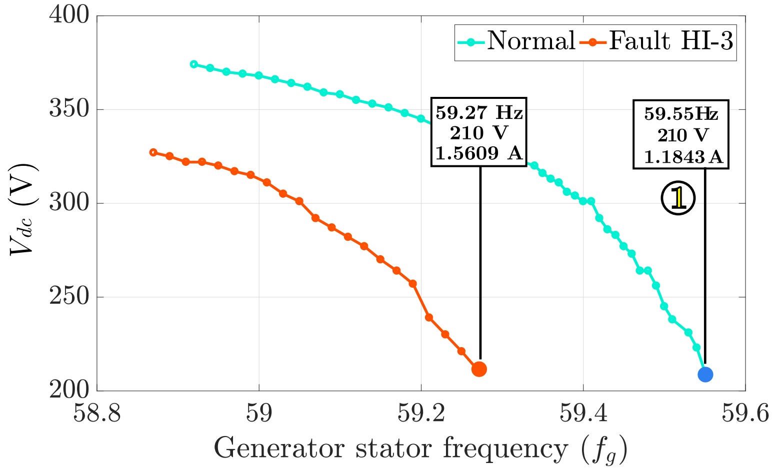
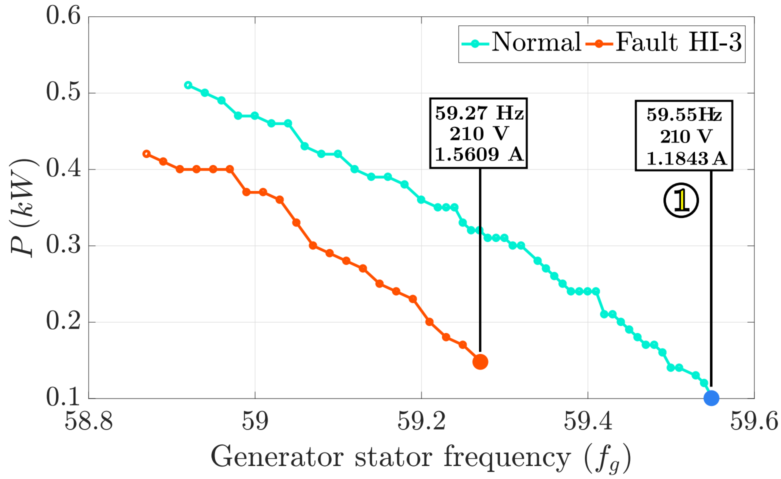


Fig. 4. The output from the MLP-Fourier fitted under the binary dataset. Normal conditions are circle, and Fault conditions are star.



(a)



(b)

Fig. 5. Curves dc-bus voltage in Fig. 5a and power in Fig. 5b.

avoid some samples could possibly be misclassified, we chose to introduce a rejection band at the output of the MLP. If any sample is within this range, the sample is rejected by the network and a new sample is analyzed. This procedure reduces false positive or false negative. In the specific case of this work, an asymmetric rejection range between + 0.2 and − 0.4 was implemented in order to highlight the decrease of false positives, since erroneous classification of the Normal condition as a Fault can lead to loss of classifier's reliability, since the machine should be stopped for later evaluation.

The analysis of the dataset in Fig. 4 indicates samples from normal conditions that are closer to the threshold have a low voltage on the dc-bus, around 210 V. Even in different generator frequencies, and in the faulty conditions as well.

That might occur because the imbalance produced by the short-circuit deprecate the generation capacity. TThis hypothesis is affirmed by the measured power and dc-bus voltage on the modified generator during the emulations in the test-bench. The sample highlighted by the number **1** in Figs. 4, 5 a and 5b are the same. When f_g was 59.55 Hz, this was much closer to the starting point of the generation curve, and this moment might be interpreted as a low wind-speed condition, as well. When the generation power capability, Fig. 5b, is too much under its nominal rate, so the V_{dc} in the frequency converter tend to be much less than its rated value, also might lead to misclassifications. This phenomenon incites the usage of a relative frequency between stator's, and rotor's generator is more appropriate to avoid misclassifications, but this requires a speed sensor coupled to the generator. However, we emphasize the reliability of the classifiers be kept since the rejection band prevent false positive misclassifications.

So, the comparison and propositions for the classifiers make possible to identify incipient short-circuits with a reliable classification, to meet necessities of wind farm maintenance routines. This classifier could prevent not also unwanted stops of the wind turbine, but provide alerts to direct repair in loco without having to remove generator from the site. Also, might potentially reduce downtime in wind farms.

6.3. Comparison of works related

In [9] the authors trained an MLP and were able to correctly identify 98.96% of the conditions using 30 features and 85.63% with seven features reduced with PCA, while we achieved 100% with nine features, as exhibited in Table 7. Despite varying the amount of turn under short-circuit, the author considered all faulty samples as faults and as we shown, the main difficult while detecting incipient short-circuits are the confusion between levels of short-circuits. We were able to identify 1.41% of turn under short-circuit with 99.93% of accuracy with less than 2% of false negatives and 0% of false positives. Palácios et al. [9] also acquires its signals with 25 kHz during 15 s, which might compromise memory storage in embedded applications, while we did 5 kHz in 10 s. However, the time analysis was not contemplated by their research.

We exhibited the fault detection is frequency dependent, especially in regions with lower voltage in dc-bus. So our methodology demonstrated to be more generalist, instead of the results of in [7,8], which uses the electrical machine running at only on frequency.

However, the studies made by previous works in our researches groups, like Oliveira et al. [10], Coelho et al. [12] and Vieira et al. [11], concluded one current sensor might be suitable to identify short-circuits in induction motors powered by frequency converters. Coelho et al. [12] achieved better results and achieved 97.5% of accuracy for binary classifications. In fact, this is also relevant to our approach, because using only one current sensor we obtained around 85% of accuracy rate overall classes and 100% in binary classifications.

7. Conclusions

The results from combinations of feature extraction methods and classifiers indicated that the combination Fourier and Multilayer Perceptron is a better tool to detect incipient stator winding inter-turn short-circuits in induction generator applied to wind turbines.

It was perceived in all training, Normal conditions were classified with accuracy over 99%, and by grouping all faults together, it was achieved 100% in a binary classification. Rejection thresholds are implemented to reduce false positive and negative rates, despite there were none of the samples rejected.

Analyzing the outputs from the binary neural networks it was possible to identify when dc-bus is far from its rated value (311 *V*), Normal samples tend to be misclassified as faulty, but the classifier reliability is kept by the rejection band, to avoid false positive misclassifications.

However, to identify incipient faults, it was observed that grouping faults in different kind of groups led to different interpretations. Therefore, to provide an early fault detection system a multi-classifier structure, containing the four, three and two classes neural networks should be used. It is important to emphasize that all short-circuit conditions evaluated are in fact incipient conditions, since the short-circuit current is limited to generators rated current. Thus, we believe real short-circuit conditions should be easily identified due to our classifier robustness.

The methodology used in this work demonstrated to be efficient and could be replicated in systems that are already installed in wind farms, and in a newer wind turbine as well. Since frequency converters power new generators, this solution could also be embedded in it, creating an integrated product responsible for powering, control and monitoring. This certainly would increase the reliability and availability of the wind farm.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.compeleceng.2018.07.046.

References

[1] GWEC. Global Wind Report 20162017; 76. http://files.gwec.net/files/GWR2016.pdf.
[2] Polinder H, Ferreira JA, Jensen BB, Abrahamsen AB, Atallah K, McMahon Ra. Trends in wind turbine generator systems. IEEE J Emerg Sel Top Power Electron 2013;1(3):174–85. https://doi.org/10.1109/JESTPE.2013.2280428.
[3] Hahn B, Durstewitz M, Rohrig K. Reliability of wind turbines. Wind Energy 2007:1–4. https://doi.org/10.1007/978-3-540-33866-6_62.
[4] Yaramasu V, Wu B, Sen PC, Kouro S, Narimani M. High-power wind energy conversion systems: state-of-the-Art and emerging technologies. Proc IEEE 2015;103(5).
[5] Bonnett AH, Soukup GC. Cause and analysis of stator and rotor failures in 3-phase squirrel cage induction motors. Pulp and paper industry technical conference, 1991, conference record of 1991 annual. 1991. p. 22–42.
[6] Thomson WT. Motor current signature analysis for incution motors. Ind Appl Case Hist 2017:359–73.
[7] Morsalin S, Mahmud K, Mohiuddin H, Halim MR, Saha P. Induction motor inter-turn fault detection using heuristic noninvasive approach by artificial neural network with Levenberg Marquardt algorithm. IEEE international conference on informatics, electronics & vision (ICIEV). 2014. p. 1–6.
[8] Bouzid MBK, Champenois G. Neural network based method for the automatic detection of the stator faults of the induction motor. Electrical engineering and software applications (ICEESA), 2013 international conference on. IEEE; 2013. p. 1–7.
[9] Palácios RHC, Silva IN, Goedtel A, Godoy WF, Lopes TD, Procópio C. Reconhecedor Neural de Defeitos no Estator em Motores de Indução Trifásicos Apoiado por Análise de Componentes Principais. In: XII Simpósio Brasileiro de Automação Inteligente, 2015, Natal/RN/Brasil. XII SBAI, 2015.
[10] Oliveira AG, Silva R, Medeiros CMS. Neural network used to stator winding interturn short-circuit fault detection in an induction motor driven by frequency converter. Proceedings - 1st BRICS countries congress on computational intelligence, BRICS-CCI 2013. 2013. p. 459–64.
[11] Vieira RG, Cunha RGC, Medeiros CMS, Silva ET. Embedding a neural classifier to detect faults in a three-phase induction motor. 2016 VI Brazilian symposium on computing systems engineering embedding. 2016. https://doi.org/10.1109/SBESC.2016.31.
[12] Coelho DN, Barreto GA, Medeiros CMS, Santos JDA. Performance comparison of classifiers in the detection of short circuit incipient fault in a three-phase induction motor. Computational intelligence for engineering solutions (CIES), 2014 IEEE symposium on. IEEE; 2014. p. 42–8.
[13] Mendel JM. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. Proc IEEE 1991;79(3):278–305.

[14] Dwyer R. Detection of non-Gaussian signals by frequency domain kurtosis estimation. ICASSP '83 IEEE international conference on acoustics, speech, and signal processing. 8. 1983. p. 607–10.
[15] Antoni J. The spectral kurtosis: a useful tool for characterising non-stationary signals. Mech Syst Signal Process 2006;20(2):282–307. https://doi.org/10.1016/j.ymssp.2004.09.001.
[16] Ramalho GLB, Ferreira DS, Filho PPR, de Medeiros FNS. Rotation-invariant feature extraction using a structural co-occurrence matrix. Measurement 2016;94:406–15.
[17] Rebouças Filho PP, Rebouças EdS, Marinho LB, Sarmento RM, Tavares JMR, de Albuquerque VHC. Analysis of human tissue densities: a new approach to extract features from medical images. Pattern Recognit Lett 2017;94:211–8.
[18] Wang H, Butler KL. Modeling transformers with internal incipient faults. IEEE Trans Power Delivery 2002;17(2):500–9.
[19] Neti P, Wilhelm B. Online health monitoring of motor insulation. General Electric Company; 2013.
[20] Tshiloz K, Smith A, Mohammed A, Djurović S, Feehally T. Real-time insulation lifetime monitoring for motor windings. Electrical machines (ICEM), 2016 XXII international conference on. IEEE; 2016. p. 2335–40.
[21] Thomson WT. On-line MCSA to diagnose shorted turns in low voltage stator windings of 3-phase induction motors prior to failure. Electric machines and drives conference, 2001. IEMDC 2001. IEEE international. IEEE; 2001. p. 891–8.
[22] Rodrigues MB, Da Nóbrega RVM, Alves SSA, Rebouças Filho PP, Duarte JBF, Sangaiah AK, et al. Health of things algorithms for malignancy level classification of lung nodules. IEEE Access 2018;6:18592–601.
[23] Rebouças Filho PP, Cortez PC, da Silva Barros AC, Albuquerque VHC, Tavares JMR. Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images. Med Image Anal 2017;35:503–16.
[24] Neto EC, Gomes SL, Filho PPR, de Albuquerque VHC. Brazilian vehicle identification using a new embedded plate recognition system. Measurement 2015;70:36–46.
[25] Thomson WT, Fenger M. Current signature analysis to detect induction motor faults. IEEE Ind Appl Mag 2001;7(4):26–34.

Pedro Pedrosa Rebouças Filho - Ph.D. degree in Teleinformatics Engineering from Federal University of Ceará, Brazil, in 2013, and is a professor at Federal Institute of Education, Science and Technology of Ceará (IFCE), Brazil. His current researches interest are applications in Computer Vision. For more information: https://professorpedrosa.com/

Navar de Medeiros Mendonça e Nascimento - Master of Science in Renewable Energies from Federal Institute of Education, Science and Technology of Ceará (IFCE) in 2018. Received the Bachelor degree in Mechatronics Engineering from Institute of Education, Science and Technology of Ceará, Brazil, in 2015. His current researches interest are applications in Pattern Recognition, Computer Vision, Electrical Machines and Renewable energies.

Igor Rocha de Sousa - Master degree student in Renewable Energies at Federal Institute of Education, Science and Technology of Ceará (IFCE), received the Bachelor degree in Mechatronics Engineering from Institute of Education, Science and Technology of Ceará (IFCE), Brazil, in 2017. His current researches interest are applications in Pattern Recognition, Computer Vision and Electrical Machines and Renewable energies.

Cláudio Marques de Sá Medeiros - Ph.D. degree in Teleinformatics Engineering from Federal University of Ceará, Fortaleza, /Brazil, in 2008, and is a professor at Federal Institute of Education, Science and Technology of Ceará (IFCE), Fortaleza, Brazil. His current researches interest is applications in Machine Learning, Electrical Machine Failure Detection and Renewable Energies.

Victor Hugo C. de Albuquerque - Ph.D. in Mechanical Engineering with emphasis on Materials from the Federal University of Paraiba (UFPB, 2010), an M.Sc. in Teleinformatics Engineering from the Federal University of Ceara (UFC, 2007). He is currently Assistant VI Professor of the Graduate Program in Applied Informatics at the University of Fortaleza (UNIFOR).