

# Industrial Data-driven Monitoring based on Incremental Learning applied to the Detection of Novel Faults

**Abstract**— The detection of uncharacterized events during electromechanical systems operation represents one of the most critical data challenges dealing with condition-based monitoring under the Industry 4.0 framework. Thus, the detection of novelty conditions and the learning of new patterns are considered as mandatory competencies in modern industrial applications. In this regard, this study proposes a novel multi-fault detection and identification scheme, based on machine learning, information data-fusion, novelty-detection and incremental learning. First, statistical time-domain features estimated from multiple physical magnitudes acquired from the electrical motor under inspection are fused under a feature-fusion level scheme. Second, a self-organizing map structure is proposed to construct a data-based model of the available conditions of operation. Third, the incremental learning of the condition-based monitoring scheme is performed adding self-organizing structures and optimizing their projections through a linear discriminant analysis. The performance of the proposed scheme is validated under a complete set of experimental scenarios and results compared with a classical approach.

**Index Terms**—Condition monitoring, Fault detection, Feature extraction, Incremental learning, Novelty detection, Machine learning.

## I. INTRODUCTION

The promotion of a sustainable industrialization represents one of the main targets considered in the sustainable development goals foreseen for the next decade [1]. In this regard, the deployment of an effective Industry 4.0 is critical to enable the regeneration of the industrial sector towards more efficient, reliable, resilient and competitive infrastructures. Indeed, the Industry 4.0 relies in the convergence between Information Technology (IT) and Operation Technology (OT), in order to reach industrial plants where the interconnectivity between virtual and physical worlds. Such interconnectivity allows the exploitation of data and the enhancement of advanced services as the condition monitoring of critical systems to detect and identify problems before they even occur, preventing downtimes and even planning future actions [2].

A great deal of industrial applications is supported by the transformation of electrical to mechanical energy through electric machines, which, nowadays, are particularly oriented to permanent-magnet synchronous motors (PMSM). The PMSM are considered a clear future choice due to their high-speed operation and precise torque control, high torque to current ratio as well as a high-power density and high efficiency [3-5]. However, despite its characteristics, the appearance of unexpected faults may occur at any time due to wear and tear of different parts, causing undesired breakdowns that affect productivity and, in consequence, economic effectiveness. Thus, the condition monitoring approaches based on data-driven methods are highly demanded to increase the availability and efficiency of industrial equipment [6-7]. Yet, although a great deal of data-driven diagnosis methods have been proposed, most of them are focused on the characterization of a set of single faulty patterns previously defined.

Indeed, the detection and identification of electromechanical system conditions present new challenges in the modern industry framework. Such requirement represents a serious issue, thereby, new fusion information strategies, based on the combination of different physical magnitudes, and the exploitation of databases by means of reliable data-driven methodologies, represent the most promising approaches to describe unequivocal patterns related to the different conditions [8-9]. Certainly, the detection of operating novelties, novelty detection, is considered as a required competency for the next generation of condition-based monitoring (CBM) schemes. Most of the available CBM approaches dealing with novelty detection are limited by [10-12]: (i) the incorporation of new scenarios to the initial models is not considered, (ii) the processing stage is focused on the detection of specific faults and, (iii) the models require a large amount of data for characterization. In this regard, classical incremental learning approaches are based on the re-training of the machine learning based solution once a new operating condition (i.e. fault) is detected. The most extended solution lies on the consideration of a linear feature reduction technique, as Principal Component Analysis (PCA), and Support Vector Machine (SVM), as domain based classifier and novelty detector [13-14]. This approach presents two main disadvantages: representative data of each scenario must be stored, and algorithms hyper-parameters must be updated during each iteration.

Thereby, the contribution of this work lies in the proposal of a data-driven monitoring methodology based on adaptive and incremental learning. The proposed methodology overcomes the need of representative data storage as well as the re-design of new classification and novelty detection structures each time new patterns appear. Thus, facing one of the current data science challenges in Industry 4.0: the detection, identification and learning of novel conditions of operations in PMSM based electromechanical systems. Indeed, the high-dimensional set of features estimated from the characterization of the considered physical magnitudes, stator currents and vibration signals, is exploited to design the digital model approach for novel faults detection and identification. Originality of this work includes the consideration of an incremental learning stage supported by a re-training procedure based on LDA projection, and the validation of the proposed fault detection and identification structure supported by Self-Organizing Maps (SOM) for data modeling, diagnosis and novelty assessment. It should be noticed that dealing with industrial applicability requirements, two additional benefits arise from the proposed methodology. First, the quantification of the reliability degree related to the diagnosis outcome, by means of a similarity analysis between the measurements under assessment and the training data base. Second, the identification of deviations over the nominal condition, which represents an important capability since unexpected faults not initially characterized may appear during the useful life of the electromechanical system.

## II. PROPOSED METHOD

Most of condition-based monitoring proposals work under the premise that information about fault scenarios of the machine under monitoring is available. This consideration reduces, significantly, the reproducibility of the method for being implemented to industrial applications. Thus, the main considerations to be faced are: (i) the premise that only information of the healthy condition is initially available and, (ii) the adaptation of the fault detection and identification scheme in order to incorporate new conditions. In this regard, such challenges are addressed in this work by means of the proposed fault detection and identification methodology depicted in Fig. 1. The general structure of the methodology is intended to serve as a guide to extend such a condition monitoring scheme to other electrical machine-based systems presenting the same aforementioned circumstances; yet, the modules to process the data and the feature calculation have been proposed to obtain relevant information dealing with an electromechanical system.

The proposed methodology is divided in four parts: (i) the feature set estimation, (ii) the initial data modeling stage, (iii) the on-line assessment, (iv) the incremental learning.

### A. Feature estimation

During the continuous monitoring stage, vibration and stator current signals are acquired from the electromechanical system under inspection. Then, a set of ten statistical time-domain features is estimated from each physical magnitude following classical approaches. The set of statistical features is composed by: mean, RMS, standard deviation, variance, shape factor, crest factor, latitude factor, impulse factor, skewness, and kurtosis. The mathematical equations of the proposed set of features can be found in [15]. Although features from frequency domain and time-frequency domain could be also considered in order to enhance the characterization procedure, it has been demonstrated that the proposed set of statistical time-domain features perform properly dealing with condition assessment in electromechanical systems [16-17].

### B. Training

From a practical point of view, most of applications start with healthy behavior data. Thus, the proposed method considers the

modeling of the available healthy data distribution by means of SOM technique in order to detect eventual novelty patterns. In fact, the topological characteristics of the data distribution are highly connected with the fault effects over the considered physical magnitudes and estimated statistical time-domain features. Thus, SOM is a neuron grid network which uses a learning rule based on the preservation of the topological properties of a  $d$ -dimensional manifold (i.e. initially, the data set corresponding to the healthy behavior), which allows a deeper characterization for eventual faults dealing with the novel occurrence of malfunction conditions. In fact, considering only the healthy condition, a one-class problem arises, that is, the *known* class corresponding to the healthy condition and *unknown* class corresponding to any other behavior, that is, novel conditions.

The SOM neuron grid consists of a predefined number of neurons known as Matching Unit (MU), which are randomly initialized and successively adapted in order to retain as much as possible the topological properties of the original data distribution during the training process. Later, new data samples are evaluated over the resulting grid, topologically representative of the original data manifold. The Euclidean distance to each MU is estimated, and the nearest neuron is activated as Best Matching Unit (BMU), inheriting all the properties (i.e., class and membership degree) [18].

### C. Operation

As the monitoring of the machine under inspection progresses, eventual novel patterns will be identified. In case of novel condition, the incremental learning stage is triggered to incorporate the new pattern into the novelty detection and fault identification structure. In this sense, the quantization error,  $E_q$ , represents the distance of the data vector, that is, the measurements under assessment, to its BMU. Accordingly, the  $E_q$  resulting from a measurement assessment reveals the amount of knowledge that the corresponding SOM model has over such data point, that is, a similarity degree in regard with the original data used during the corresponding SOM model training. Indeed, during the training stage, the mapping of the data distribution provides a mean quantization error,  $\bar{E}_q$ , describing the average distance error between the data set and the corresponding MU set. Then, during the assessment stage, quantization errors,  $E_q$ , within the  $\bar{E}_q$  obtained during the

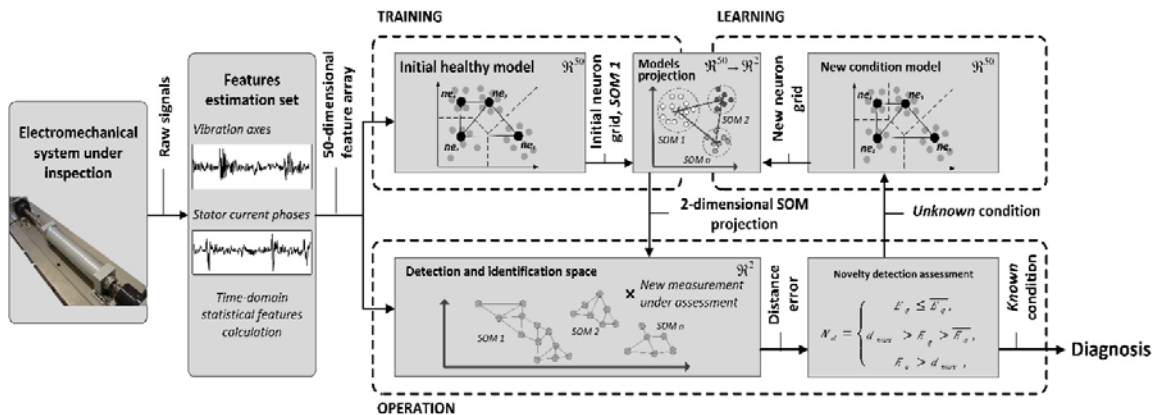


Fig. 1. Flow chart of the proposed multi-fault detection and identification scheme based on adaptive and incremental learning. Training stage: modeling of the healthy condition. Operation stage: distance estimation between the measurements under assessment and all neurons of the models to determine the diagnosis or novelty degree of the measurements. Learning stage: new model generation and optimized projection to maximize separation among models.

training,  $\bar{E}_{q-training}$ , represent the lowest novelty degree,  $N_d=0$ , since the measurement is understood as *known*. An  $E_q$  between the  $\bar{E}_q$  obtained during the training,  $\bar{E}_{q-training}$ , and the maximum distance obtained during the training,  $d_{max-training}$ , represents a proportional novelty degree from 0 to 1. Finally, an  $E_q$  bigger than the maximum distance obtained during the training,  $d_{max}$ , results in the highest novelty degree,  $N_d=1$ , since the measurement is understood as *unknown*. The  $E_q$  is, then, used to provide information regarding the novelty degree of the measurement under analysis following the next equation:

$$N_d = \begin{cases} E_q \leq \bar{E}_{q-training}, & 0 \\ d_{max} > E_q > \bar{E}_{q-training}, & 1 - \frac{E_q - \bar{E}_{q-training}}{d_{max-training} - \bar{E}_{q-training}} \\ E_q > d_{max-training}, & 1 \end{cases} \quad (11)$$

#### D. Learning

The incremental learning procedure is performed by considering the following steps:

*Step 1.* The incremental learning is only carried out whether a novelty detection occurs, that is, novelty degree  $N_d=1$ .

*Step 2.* A new SOM based model is trained considering the data related to the new class distribution.

*Step 3.* The available SOM based models are subjected to a dimensionality reduction procedure by means of LDA projection resulting into a two-dimensional projection maximizing classes separation and defining the *know* information.

Through these three steps, the incremental learning procedure is carried out. It should be noticed that this approach allows including new information to the proposed monitoring system without need of access to the measurements that were used initially for the training of the SOM models. The training of the new model is supported by the selection of the configuration hyper-parameters, specifically, the number of neurons to be used and the initial shape of the structure. The SOM structure depends on the number of available data points and its complexity in terms of distribution. However, as most of works dealing with SOM algorithm suggest, sheet-type structures composed by a number of neurons following a 1:10 ratio in regard to the available data represents a good tradeoff between structure complexity and performance, and this is the rule proposed to be followed in this work [19-20].

### III. EXPERIMENTAL SETUP

In order to validate the performance of the proposed methodology, different operating patterns have been considered in a laboratory scale test bench. The experimental test bench is composed of two identical electric motors, a gearbox, and a screw shaft. One of the electric motors is the motor under study which is used as the driving motor, this motor drives the input shaft of the gearbox and the output shaft of the gearbox drives, in turn, the screw shaft that performs a displacement of a movable part. The electric motors belong to the model ABB-SPMSMs and have 3 pairs of poles, a rated torque of 3.6 Nm, 230 Vac, and a rated rotational speed of 6000 rpm. To drive and control these motors an ABB power converters model ACSM1 is used. The mechanical vibrations and the stator currents are continuously acquired by means of a triaxial accelerometer and current probes, respectively. The acquisition is performed through a PXIe 1062 data acquisition system from National

Instruments with a sampling frequency of 20kS/s to acquire the measured signals at one second of sampling time. A flowchart of the electromechanical system and the data acquisition is shown in Fig. 2.

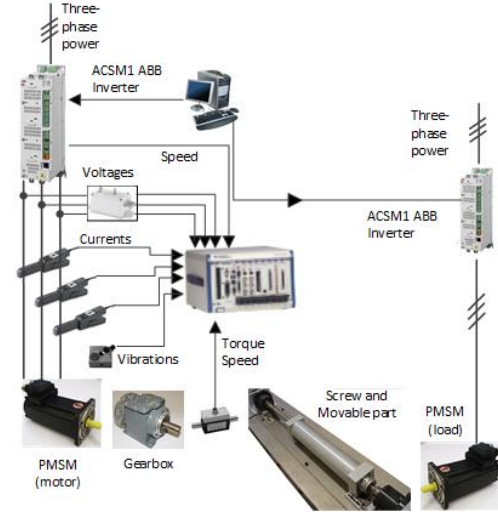


Fig. 2. Experimental test bench considered to evaluate the effectiveness of the proposed incremental learning approach for the novelty fault identification.

Four different conditions have been experimentally tested. The first condition is the healthy condition (HLT). The second condition is degraded bearing (DB), where the inner and the outer raceways of the non-end bearing have been uniformly scraped aiming to produce a generalized roughness defect. The third condition is partial demagnetization of the electric motor (DEM). Such condition considers around 50% of nominal flux reduction in one pair of poles. Finally, the fourth condition, a static eccentricity (ECC) has been produced by means of attaching a bolt in the output shaft of the gearbox. All these conditions have been evaluated under different experiments, thus, two-speed setpoints combined with two different torque patterns setpoints have been considered, 1500 rpm, 3000 rpm, and, 0 % and 50 % of the nominal rated torque, respectively. Consequently, each considered condition has been experimentally evaluated under four different operating regimes; fifty measurements were performed for each condition. It must be noticed that the data imbalance represents a critical issue in most data science challenges and, in particular dealing with electromechanical systems, since the normal operating condition is the most represented usually [21-22]. However, the proposed methodology avoids the non-desired effects of such situation since the adaptive nature of the self-organizing maps technique considered in the proposed methodology characterizes topological aspects of the data manifold as data density, independently of the data number.

### IV. RESULTS AND DISCUSSION

Each physical magnitude acquired during the operation of the electromechanical system is characterized through the estimation of the proposed set of ten numerical statistical time-domain features. Thus, the initial available condition (i.e. HLT) is characterized by a high-dimensional set of fifty statistical features. Then, the first fault detection and identification model is trained by mapping such high-dimensional set of features corresponding to the HLT condition through a first SOM

neuron grid model. Thus, the first data modeling is performed and represented by the  $SOM_1$  grid, in which only information related to the HLT condition is considered during the training procedure. A qualitative representation of the resulting 50-dimensional data manifold of this initial HLT condition by means of a 2-dimensional PCA projection is shown in Fig. 3(a); and, the resulting 2-dimensional PCA projection of the predefined 10 x 10  $SOM_1$  neuron grid is shown in Fig. 3(b), which represents the  $SOM_1$  model. From both figures, it can be observed that HLT topological data distribution and  $SOM_1$  neuron distribution match properly.

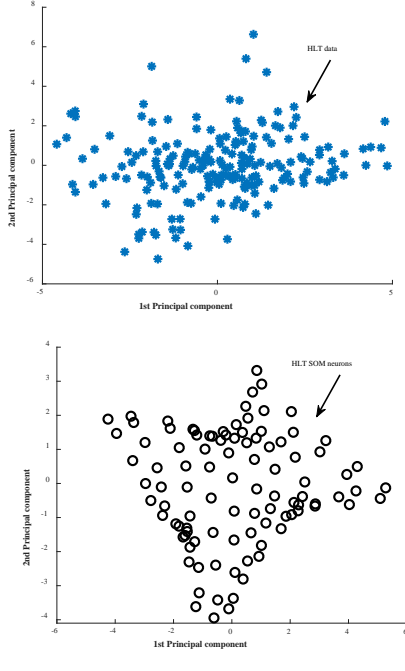


Fig. 3. Qualitative representation of the resulting 50-dimensional data manifold of this initial HLT condition. (a) PCA projection of available measurements. (b) PCA projection of resulting  $SOM_1$  neurons.

The  $SOM$  model performance is quantitatively defined by its quantization error. That is, the mean distance from each available measurement to its BMU, is represented by the average quantization error,  $\bar{E}_q$ . The resulting mean quantization error during the training process of this  $SOM_1$  is 5.23. In order to evaluate the performance of this first model, a first test set, test set 1, is evaluated. In Table I is specified the content of each test set used during the experimental validation, including *known* and *unknown* class conditions. Specifically, in this test set 1, the *known* class corresponds to the HLT condition, while the *unknown* class corresponds to the bearing defect (BD) condition. The evaluation of this novel faulty condition (i.e. BD), over the  $SOM_1$  results in a  $\bar{E}_q$  of 167.03. Indeed, an increase in the mean quantization error is due to the detection of a different condition in regard to the *known* data distribution. That is, such increase is produced because the topological characteristics of the database used during the training (composed by the HLT condition), are different from the topological properties of the database used during the evaluation (composed by the BD condition). In Fig. 4 is shown a qualitative representation, obtained by means of a PCA projection, of the resulting  $SOM_1$  model together with the resulting 50-dimensional data manifold of the BD condition.

TABLE I  
DETAIL OF TRAINING AND TEST SETS

Test case	Training set	Test set	
		<i>Known</i> class	<i>Unknown</i> class
1	HLT	HLT	BD
2	HLT, BD	HLT, BD	DEM
3	HLT, BD, DEM	HLT, BD, DEM	ECC

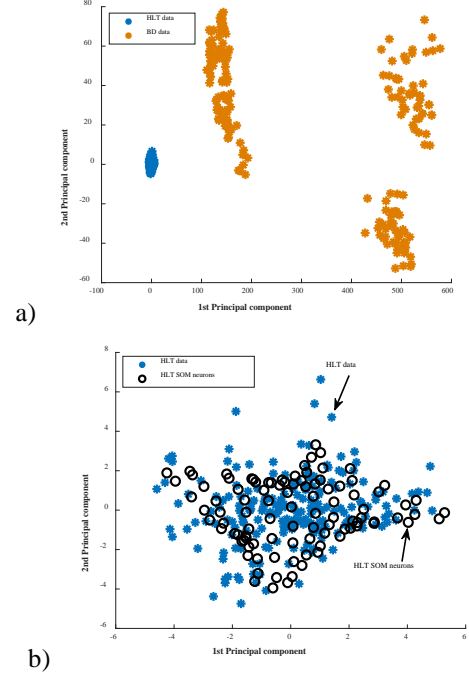


Fig. 4. 2-dimensional PCA projection resulting from the evaluation of  $SOM_1$ . (a) HLT and BD data projection. (b) zZoom over the HLT data with the  $SOM_1$  neurons.

The  $E_q$  values obtained during the assessment of both conditions over the  $SOM_1$  are shown in Fig. 5, where the first 200 test measurements, corresponding to the HLT condition, exhibit  $E_q$  values within the range of the  $\bar{E}_q$ , 5.23, while, the  $E_q$  values of the next 200 measurements, corresponding to the *unknown* class condition (i.e. BD), exhibit values over 100. Such increase in the  $E_q$  value reveals a novelty detection. It must be noted that, for both conditions, HLT and BD, the four different operating conditions were considered, thus, leading to variation in the obtained  $E_q$  values within the same condition.

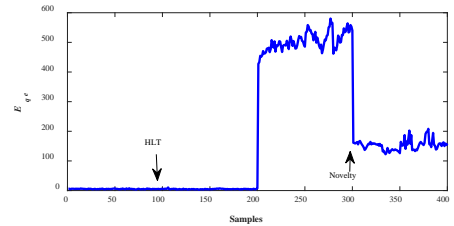


Fig. 5. Resulting quantization error during the assessment of the *known* class HLT and *unknown* class BD over the first model,  $SOM_1$ .

Once the *unknown* condition is detected, the knowledge of the novelty detection structure must be updated to include the pattern related to the new detected condition. Thus, such data modeling is carried out by means of a second novelty detection model,  $SOM_2$ , represented by a new 10 x 10 neuron grid.



Thereby, during the model training of such BD condition, an averaged quantization error,  $\bar{E}_q$ , of 18.96 is obtained. Thus, the new novelty detection structure is composed of two models,  $SOM_1$  and  $SOM_2$ , that represent the data distribution of the two *known* class conditions HLT and BD, respectively. A qualitative representation of the resulting  $SOM_1$  and  $SOM_2$  models by means of a 2-dimensional PCA projection is shown in Fig. 6(a), and the resulting  $SOM_2$  model represented by the U-Matrix is shown in Fig. 6(b). Certainly, the unified distance matrix, the U-matrix, applied over the resulting  $SOM_2$  trained structure, allows the visualization of distances between the weight vector adjacent neuron units in the defined two-dimensional map. Specifically, the U-matrix represents the distances between neurons of such trained SOM where darkest areas depict those areas in the map where the data concentrations are presented (there exist smallest distances between neurons), and lightest areas depict boundaries between clusters (there exist bigger distances between neurons). It is shown in Fig. 6(b) that three main large areas can be distinguished, which means that the BD data behaves as neighbor clusters.

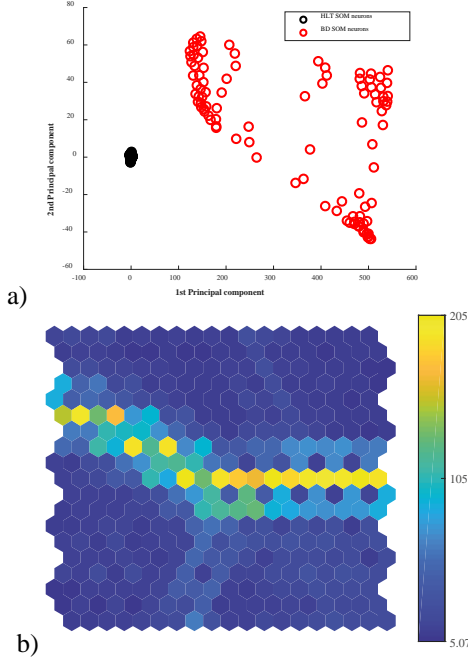


Fig. 6. Updated novelty detection structure. (a) Qualitative representation of the resulting 50-dimensional data manifold of the HLT and BD known conditions through the  $SOM_1$  and  $SOM_2$  neurons. (b) Detail of the resulting U-matrix resulting from the  $SOM_2$  model. The lighter the color between two neurons is, the larger is the relative distance between them.

In this regard, if a new *unknown* condition takes place, the current novelty detection structure will be assessed to perform its detection and, if proceed, to update the current knowledge and increase the novelty detection structure again. In order to enhance the novelty detection capabilities of the proposed novelty detection structure, it is proposed the application of a linear projection supported by LDA. That is, through LDA projection, the 50-dimensional space containing  $SOM_1$  and  $SOM_2$  is reduced into 2-dimensional space called detection and identification space, where the separation between SOM models is augmented due to the discriminant capabilities of the

LDA. Consequently, the new 2-dimensional space allows obtaining a visual representation of the 50-dimensional spaces represented by the neuron grids of the involved models.

Considering the experimental case under analysis, the LDA projection of the two SOM neuron grids,  $SOM_1$  and  $SOM_2$ , results in the 2-dimensional projection shown in Fig. 7. Thus, the HLT and BD conditions are represented by the obtained projection in which both conditions appear separated by different clusters.

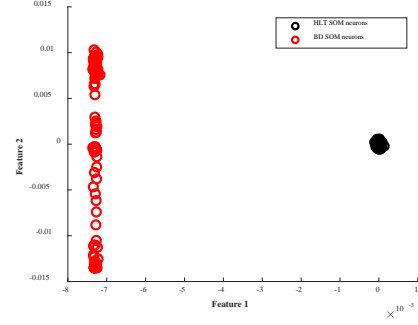


Fig. 7. 2-dimensional projection performed by the LDA over the two available model,  $SOM_1$  and  $SOM_2$ .

Consequently, as aforementioned, the appearance of a new *unknown* condition has to be assessed and detected by the current novelty detection structure composed by  $SOM_1$  and  $SOM_2$ . Then, in case of novel condition, such *unknown* patterns would be characterized by another SOM based model and, then, a new projection would be also obtained by carrying out a dimensionality reduction of all SOM grids through the LDA. In this sense, aiming to highlight the effectiveness of the proposed method, another test case, test case 2, with *known* and *unknown* classes is evaluated through the current novelty detection structure composed by  $SOM_1$  and  $SOM_2$ . The evaluation of such test case 2 results in an averaged quantization error,  $\bar{E}_q$ , of 152.59 and 72.97 from  $SOM_1$  and  $SOM_2$  respectively. Hence, the novelty detection is performed by comparing these values with those obtained during the training of the corresponding novelty models,  $SOM_1$  and  $SOM_2$ . This detection capability is quantitative represented in Fig. 8(a) and Fig. 8(b) where the  $E_q$  values obtained during the test case 2 evaluation is shown.

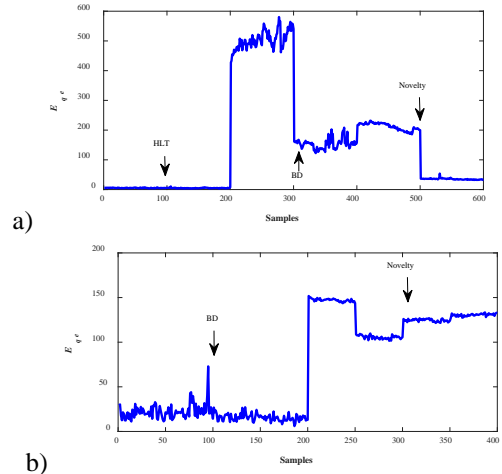


Fig. 8. Quantization error achieved during the assessment of the *known* classes HLT and BD and *unknown* class DEM over available SOM's. (a) Assessment over  $SOM_1$  model. (b) Assessment over  $SOM_2$  model.

Because a novel condition has been detected, the knowledge of the novelty detection structure is updated and the dataset of such *unknown* condition has to be modeled by a new SOM grid. Specifically, the novel detected condition belongs to an eccentricity condition (ECC). Once the data modeling is performed, the ECC pattern is represented by the third novelty detection model, a  $10 \times 10$  neuron grid,  $SOM_3$ , with an averaged quantization error of 6.52 obtained during its training. As a result, the new novelty detection structure is composed of three models:  $SOM_1$  representing the HLT condition,  $SOM_2$  representing the BD condition and  $SOM_3$  representing the ECC condition. Then, the three SOM models are again subjected to a dimensionality reduction procedure by means of the LDA projection. The resulting 2-dimensional projection is shown in Fig. 9. The three already *known* conditions are represented by different clusters that are perfectly separated from each other.

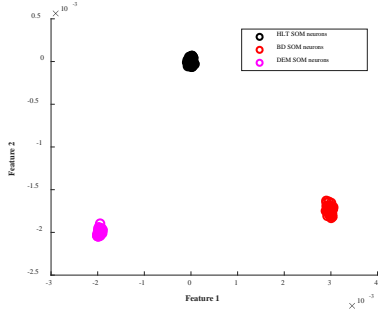


Fig. 9. 2-dimensional LDA projection over the three SOM grids:  $SOM_1$ ,  $SOM_2$  and  $SOM_3$ .

Finally, a third test set, test set 3, is evaluated over the current novelty detection structure composed by  $SOM_1$ ,  $SOM_2$  and  $SOM_3$ . Accordingly, the averaged quantization errors obtained during the evaluation of the *unknown* class of such third test case through  $SOM_1$ ,  $SOM_2$  and  $SOM_3$  are 133.53, 95.72 and 46.62, respectively. The graphical representation of the  $E_q$  value performed by the three models is shown in Fig. 10.

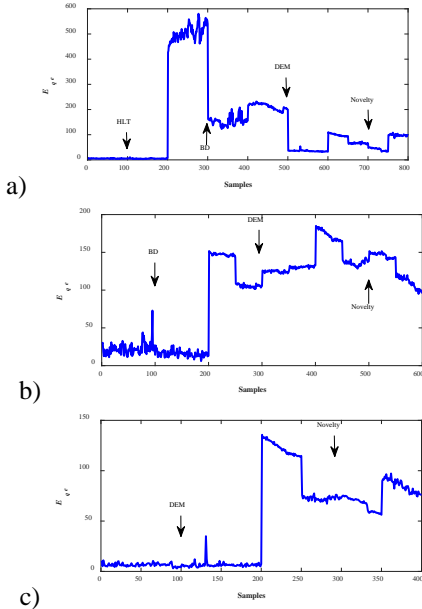


Fig. 10. Quantization error achieved during the assessment of the *known* HLT, BD and ECC conditions, and the *unknown* DEM condition. (a)  $SOM_1$  response. (b)  $SOM_2$  response. (c)  $SOM_3$  response.

The *unknown* pattern corresponds in this case to a demagnetization condition (DEM). Afterwards, the data modeling of such novel condition detected is carried out by means of a fourth model,  $SOM_4$ , which exhibited during the training process an averaged quantization error of 7.27. After updating the knowledge of the novelty detection structure,  $SOM_1$ ,  $SOM_2$ ,  $SOM_3$  and  $SOM_4$  are subjected to the proposed dimensionality reduction procedure through the LDA projection. The resulting 2-dimensional LDA projection that represent the novelty detection structure is shown in Fig. 11, where four different clusters appear completely separated one from each other.

In Table II, the resulting ratios obtained for the diagnosis and novelty detection are summarized. It should be highlighted that all evaluated samples have been correctly diagnosed and identified to its corresponding class with membership probability higher than 96% in all cases.

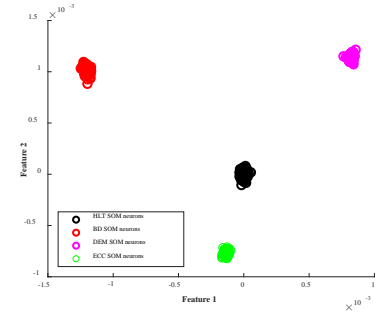


Fig. 11. 2-dimensional LDA projection over the four available SOM grids, that is,  $SOM_1$ ,  $SOM_2$ ,  $SOM_3$  and  $SOM_4$ .

TABLE II RESULTING DIAGNOSTIC AND NOVELTY DETECTION MEMBERSHIP RATIOS		
Test case	Test set	
	Known class (%)	Unknown class (%)
1	HLT (100%)	BD (100%)
2	HLT (100%), BD (100%)	DEM (>96%)
3	HLT (100%), BD (100%), DEM (100%)	ECC (>96%)

The obtained results shown the effectiveness of the proposed data-driven monitoring method for detection, identification and learning of multiple faulty conditions in an electromechanical system. Aiming to emphasize and demonstrate the effectiveness of proposed dimensionality reduction through LDA projection, an additional test case is presented. Fig. 12 shows the resulting 2-dimensional PCA projection of HLT, BD, ECC and DEM conditions, where a significant overlapping between the DEM and ECC classes takes place. In this sense, it should be mentioned that, in addition to the novelty detection performed by the proposed SOM models structure, their projection by means of LDA leads to avoid overlapping problems that may produce misclassifications. It must be noticed that this projection can be carried out iteratively trough the learning stages due to the neuron grids resulting from the self-organizing map, since each model act as a representative set of neurons in regard with the original data distribution. Thus, providing superior data processing capabilities compared with classical SVM based solution as shown in Fig 13, where classes overlapping problem reduce novelty detection performance till 12% dealing with the characterization of multiple patterns.

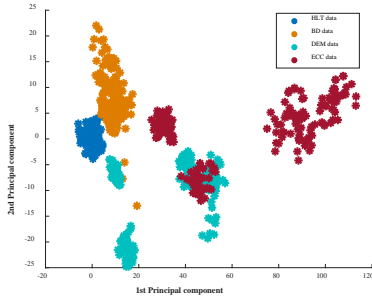


Fig. 12. 2-dimensional PCA projection performed over all considered classes, that is, HLT, BD, ECC and DEM.

Furthermore, although the proposed multi fault identification scheme manages as many sub-models as identified faults, each sub-model is focused on the characterization of a specific pattern of operation leading to a high capability of response for detecting novel conditions. Such capability of response can be quantified in terms of the computational burden, thus, over an Intel Core i7-4770K @3.50GHz CPU, the execution of the proposed algorithm in Matlab 2019a takes less than 200 msec in all considered conditions. In this regard, the inertia of most of the electromechanical faults takes several orders above such value. It should be also mentioned that the initial training, as well as the incremental learning procedures, have been analyzed in terms of required time resulting in less than 1 second for all considered conditions. Therefore, it can be stated that the proposed diagnosis methodology can be implemented as a diagnostic tool for real-time operation.

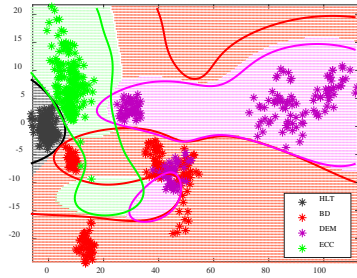


Fig. 13. Decision regions and projected boundaries of considered conditions by means of a 2-dimensional PCA projection and SVM-based novelty detection scheme.

## V. CONCLUSIONS

Modern industrial production is characterized by the consideration of machine learning data based models to support the main aspects of the manufacturing process. In this regard, two main data science challenges related with condition monitoring of electromechanical assets in the industry 4.0 framework are: (i) the premise that only information of the healthy condition is initially available and, (ii) the adaptation of the fault detection and identification scheme in order to incorporate new operating conditions. Thus, this paper proposes a new methodology for multi-fault detection and identification based on incremental learning applied to novel faults detection on electromechanical systems by analyzing vibrations and stator current signatures of the electric motor drive.

The new method has four key features: (i) It considers collaborative self-organizing maps as adaptive data patterns

characterization; the SOM-based approach allows class topology modeling by means of a reduced set of neurons starting from the healthy operation of the electromechanical system. (ii) It uses a novelty detection structure based on quantization error, which successfully detected all novel scenarios considered. (iii) The LDA based learning procedure enables novel scenarios to be incorporated into the models to upgrade the knowledge available without need of original information and enhancing classification performance. (iv) The same SOM-LDA structure is used for detection and identification, that is, novelty assessment and classification of the measurement under analysis.

Twelve different experimental scenarios have been considered including faults and operating conditions, representing a significant set of situations. Under each of these conditions, the proposed methodology provides satisfactory diagnosis results. The novelty detection is carried out in the original feature space, that is, the 50-dimensional space, by means the assessment of the resulting quantization error,  $Eq$ , for each measurement under inspection. This is due to the challenge that the novelty detection framework must address. When dealing with *unknown* fault conditions, it is critical to use multiple feature analysis approaches, as proposed with the novelty detection model structure, since the lower the number of features considered, the higher the risk of misidentification.

A dimensionality reduction procedure is later applied by means of the LDA. Thus, by using the LDA, a better separation of the considered conditions is achieved, moreover, such 2-dimensional reduction allows the implementation of visual procedures for better interpretations. Indeed, the user supervision is necessary after a novel scenario detection in order to confirm and track the corresponding root-cause. Thus, the proposed methodology is constrained to two-dimensional representations, where the underlying physical phenomena of the electromechanical system can be visualized. However, the incremental learning capabilities of the proposed method avoid model retraining and data storage. It must be noted that the proposed methodology shows a diagnostic accuracy of 100%, and 99% in novelty detection accuracy, which, compared to classical approaches based on SVM, represent high-performance ratios. It must be clarified that the detection of *unknown* operating conditions is not limited to single faulty modes, since the occurrence of multiple or combined faulty modes would lead to novel patterns that would be detected, modeled and incorporated to the knowledge of the monitoring system for posterior identification.

Indeed, from an industrial viewpoint, the development of the proposed multi fault identification structure may be extended aiming to be applied as a diagnostic tool for real-time operations. The model complexity is low in terms of computational burden, since considering a regular on-line iteration, just a change of basis must be applied over the measurement under analysis using the LDA projection matrix, together with Euclidian distance calculations between the projected measurement and the neurons of each available SOM neuron grid for membership assessment. Certainly, the proposed structure is suitable for being implemented in embedded systems such as field-programmable gate arrays (FPGA).

Finally, it should be noticed that the proposed diagnosis scheme, including SOM under an incremental learning framework supported by LDA projection has not been previously studied in multi-fault diagnosis and the obtained results reveal a its suitability for other rotating mechanical component faults and operating conditions since the proposed method has the capability to adapt to different sets of available physical magnitudes.

## REFERENCES

- [1] Marco S. Reis and Geert Gins, "Industrial process monitoring in the big data/industry 4.0 era: from detection, to diagnosis, to prognosis," *Processes*, vol. 5, pp. 1-12, 2017.
- [2] C. Jing and J. Hou, "SVM and PCA based fault classification approaches for complicated industrial process," *Neurocomputing*, vol. 167, pp. 636-642, 2015.
- [3] S. Sahoo, P. Rodriguez, and M. Sulowicz, "Evaluation of different monitoring parameters for synchronous machine fault diagnostics," *Electrical Engineering*, vol. 99, issue 2, pp. 551-560, 2017.
- [4] Kao, W. Wang, Y. Lai and J. Perng, "Analysis of permanent magnet synchronous motor fault diagnosis based on learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 2, pp. 310-324, 2019.
- [5] N. H. Obeid, T. Boileau, and B. Nahid-Mobarakkeh, "Modeling and diagnostic f incipient interturn faults for a three-phase magnet synchronous motor," *IEEE Transactions on Industry Applications*, vol. 51, issue 5, 2016.
- [6] N. R. Devi1, D. V. S. S. S. Sarma, and P. V. R. Rao "Detection of stator incipient faults and identification of faulty phase in three-phase induction motor-simulation and experimental verification," *IET Electric Power Applications*, vol. 9, issue 8, pp. 540-548, 2015.
- [7] M. Kang, M. R. Islam, J. Kim, J.-M. Kim, and M. Pecht, "A hybrid feature selection scheme for reducing diagnostic performance deterioration caused by outliers in data-driven diagnostics," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3299-3310, May 2016.
- [8] R. Loiselle, Z. Xu and I. Voloh, "Essential Motor Health Monitoring: Making Informed Decisions About Motor Maintenance Before a Failure Occurs," in *IEEE Industry Applications Magazine*, vol. 24, no. 6, pp. 8-13, Nov.-Dec. 2018. doi: 10.1109/MIAS.2017.2740465
- [9] M. Kang, Md. R. Islam, J. Kim, J. M. Kim, and M. Pecht, "Hybrid Feature Selection Scheme for Reducing Diagnostic Performance Deterioration Caused by Outliers in Data-Driven Diagnostics," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3299-3310, 2016.
- [10] Wang S, Yub J, Lapirac E, Lee J. A modified support vector data description based novelty detection approach for machinery components. *Appl Soft Comput J* 2013;13:1193–205. doi:10.1016/j.asoc.2012.11.005.
- [11] Costa BSJ, Angelov PP, Guedes LA. Fully unsupervised fault detection and identification based on recursive density estimation and self-evolving cloud-based classifier. *Neurocomputing* 2015;150, Part:289–303. doi:10.1016/j.neucom.2014.05.086.
- [12] Filev DP, Chinnam RB, Tseng F, Baruah P. An industrial strength novelty detection framework for autonomous equipment monitoring and diagnostics. *IEEE Trans Ind Informatics* 2010;6:767–79. doi:10.1109/TII.2010.2060732.
- [13] J.A.Cariño, M. Delgado-Prieto, D. Zurita, A. Picot, J. A.Ortega and R. J. Romero-Troncoso, "Incremental novelty detection and fault identification scheme applied to a kinematic chain under non-stationary operation," *ISA Transactions*, In Press, July 2019. <https://doi.org/10.1016/j.isatra.2019.07.025>
- [14] X. Zhang, W. Chei, B. Wang, and X. Chen, "Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization," *Neurocomputing*, vol. 167, pp. 260-279, 2015.
- [15] J.J. Saucedo-Dorantes, M. Delgado-Prieto, R. A. Osornio-Rios and R. J. Romero-Troncoso, "Multifault Diagnosis Method Applied to an Electric Machine Based on High-Dimensional Feature Reduction," *IEEE Transactions on Industry Applications*, vol. 53, issue 3, pp. 3086-3097, 2017.
- [16] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3398–3407, 2013.
- [17] P. Gangsar and R. Tiwari, "Comparative investigation of vibration and current monitoring for prediction of mechanical and electrical faults in induction motor based on multiclass-support vector machine algorithms," *Mechanical Systems and Signal Processing*, vol. 94, pp. 464-481, 2017.
- [18] W. Li, S. Zhang, and G. He, "Semisupervised distance-preserving self-organizing map for machine-defect detection and classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5, pp. 869-879.
- [19] J. M. Bossio, C. H. Angelo, G. R. Bossio, "Self-organizing map approach for classification of mechanical and rotor faults on induction motors," *Neural Computing and Applications*, vol. 22, issue 1, pp. 41-51, 2013.
- [20] M. Skowron, M. Wolkiewicz, T. Orłowska-Kowalska, and C. T. Kowalski, "Application of self-organizing neural networks to electrical fault classification in induction motors," *Applied Sciences*, vol. 9, issue 4, pp. 1-22, 2019.
- [21] Lai, C.S., Tao, Y., Xu, F., Ng, W.W., Jia, Y., Yuan, H., Huang, C., Lai, L.L., Xu, Z. and Locatelli, G., "A robust correlation analysis framework for imbalanced and dichotomous data with uncertainty," *Information Sciences*, 470, pp.58-77, 2019.
- [22] Wing, W. Y. Ng., Zhang, J., Lai, C.S., Pedrycz, W., Lai, L.L. and Wang, X., "Cost-Sensitive Weighting and Imbalance-Reversed Bagging for Streaming Imbalanced and Concept Drifting in Electricity Pricing Classification" *IEEE Transactions on Industrial Informatics*, vol. 15, issue 3, pp.1588-1597, 2019.