Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 00 (2016) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd INNS Conference on Big Data and Deep Learning 2018

Unintrusive Monitoring of Induction Motors Bearings via Deep Learning on Stator Currents

Francesca Cipollini[a], Luca Oneto[a,*], Andrea Coraddu[b], Stefano Savio[c], Davide Anguita[a]

[a]*DIBRIS - University of Genova, Via Opera Pia 13, I-16145 Genova, Italy*
[b]*NAOME - Naval Architecture, Ocean & Marine Engineering, Strathclyde University, Glasgow G1 1XW, UK*
[c]*DITEN - University of Genova, Via Opera Pia 11a, I-16145 Genova, Italy*

Abstract

Induction motors are fundamental components of several modern automation system, and they are one of the central pivot of the developing e-mobility era. The most vulnerable parts of an induction motor are the bearings, the stator winding and the rotor bar. Consequently, monitoring and maintaining them during operations is vital. In this work, authors propose an induction motor bearings monitoring tool which leverages on stator currents signals processed with a Deep Learning architecture. Differently from the state-of-the-art approaches which exploit vibration signals, collected by easily damageable and intrusive vibration probes, the stator currents signals are already commonly available, or easily and unintrusively collectable. Moreover, instead of using now-classical data-driven models, authors exploit a Deep Learning architecture able to extract from the stator current signal a compact and expressive representation of the bearings state, ultimately providing a bearing fault detection system. In order to estimate the effectiveness of the proposal, authors collected a series of data from an inverter-fed motor mounting different artificially damaged bearings. Results show that the proposed approach provides a promising and effective yet simple bearing fault detection system.
© 2016 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of KES International.

Keywords: Deep Learning, Monitoring, Induction Motors, Bearings, Stator Currents.

1. Introduction

Induction Motors (IMs) are ubiquitous in many industrial systems, such as modern automation systems, and e-cars [1,2]. In particular, IMs are cheap, characterized by a reasonably high efficiency, and require low maintenance activities [2,1]. Hence, IMs are a perfect industrial workhorse [3].

However, IMs are subject to different types of undesirable faults which cause additional costs and losses in production time [3]. Moreover, decayed components inside the IMs often result in a higher power consumption with respect to properly maintained ones, thus requiring additional costs in energy supply [1]. As a result, IMs maintenance is a critical

* Luca Oneto. Tel.: +39-010-353-2192; fax: +39-010-353-2897.
E-mail address: luca.oneto@unige.it

1877-0509 © 2016 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of KES International.

problem which requires the optimization of both costs and performance[4,5]. Hence, new methods for assessing the status of the IM components are becoming vital in order to maximize availability and performance[6].

As far as IMs are concerned, the most vulnerable parts are the bearings, the stator winding, the rotor bar, and the shaft[2]. In particular, bearings play a primary role in the reliability and performance of an IM because they are subject to continuous mechanical stress and because they produce undesirable vibrations when degraded[7]. Results on various studies show that bearing decay account for the 41% of all IMs failures[8]. Stator winding and rotor bar faults are responsible for respectively 37% and 10% of the total IMs faults. The remaining 12% of IMs faults are associated with other components (e.g. the shaft).

According to the literature on this subject[9,10], a common approach for monitoring the IMs bearing is to monitor the vibrations. By installing vibration sensors it is possible to easily analyse fault signatures and salient fault features[11,12]. Unfortunately, this procedure is not free from technical and economic downsides: placing sensors on the IM might not be effective nor economical. In fact, vibration sensor are not cheap, are prone to faults, are hard or impossible to install on many systems, and are sensible to corrosive and dusty environments[13].

An alternative way, with respect to study the vibration signals, is to consider the stator currents[14]. This approach has many advantages with respect to the previous one since it does not require the installation of any additional sensor and no direct access to the device is needed[15]. In fact, the stator currents signals are already commonly available, or easily and unintrusively collectable[14]. Among the different techniques found in literature, developed to predict bearings damages from the stator current signal, two main approaches exist[4]. The first one is based on the analysis of the spectrum of the current and thrust while the second one is based on Data Analytics (DA) tools applied to the raw signals.

In literature, a general review of the different frequency domain techniques applied to the analysis of motor currents can be found[16], and recently other approaches have been proposed[17,1,6]. However, all these methods rely on a simple idea: collecting raw data, filter them and then apply frequency analysis to study the variations in the frequency spectrum before and after the damage is injected. The drawback of all these methods is the high dependence on the motor specific characteristics, whose parameters need to be known in advance, to determine a reference frequency spectrum function describing the nominal bearings state[16]. Moreover, the frequency spectrum vibration analysis tend to be inaccurate if there are some slight variations on the load and current, due for example to different operative conditions or noise[18,19].

In this context, many DA tools have been developed to overcome the limitation of the frequency-based approaches[16]. In fact, DA tools do not need any a-priory information about the IM and are robust to noise[20]. Among the DA tools developed for bearings fault detection purposes different works have been proposed in literature[20,21,22] but these methods do not exploit state-of-the-art approaches developed in the last years like the Deep Neural Network (DNN). In fact, before learning, feature learning is required and conducted in many applications to achieve a satisfactory accuracy[23,24]. DNN represents a state-of-the-art choice in this context[25,26]. The deep architecture extracts features through a multilayered feature representation framework where the higher layers represent more abstract information than those from the lower ones.

In this paper authors propose to use a DNN for detecting the state of decay of the IM bearings able to extract from the stator current signal a compact and expressive representation of the bearings state, ultimately providing a bearing fault detection system. In order to estimate the effectiveness of the proposed procedure, authors collected a series of data from an inverter-fed motor with different damaged bearings. As endurance tests would have resulted in the bearings break-down, thus compromising the approach test settings, the authors decided to artificially induce different kinds of damage to the bearing to trace the evolution of their degradation state. The test was repeated with different motor loads for each damaged condition of the bearings, to asses the developed DNN in different operative conditions. Results show that the proposed approach provides a promising and effective yet straightforward bearing fault detection system.

2. Induction Motor Bearings

Three-phase IMs are asynchronous speed machines, their low cost, ability to operate in hostile environments, good dynamic performance, and wide speed range operation make them a perfect candidate for industrial applications. IMs are subjected to different modes of failures. The most common failure mode is the bearing failure, followed by stator

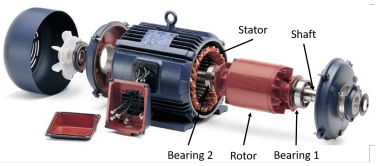

Fig. 1: Induction Motor.

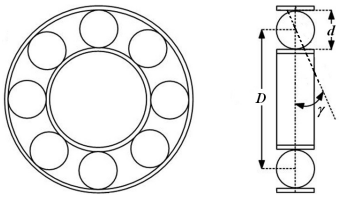

Fig. 2: Bearing characteristic parameters.

Table 1: Ball bearing parameters detail.

Bearing Parameter	Value
Outer diameter	52.0 mm
Pitch diameter D	38.5 mm
Roller diameter d	7.9 mm
Rollers number n	8
Contact angle γ	0°

winding and rotor bar failures. In Figure 1 an exploded view of an IM is reported. In Figure 2, the geometry of a rolling-element bearings (REB) is reported. It consists of two rings (one inner and the other outer), where a set of rolling elements placed in the raceways rotates inside these rings [27]. Under normal operating conditions, considering a balanced load and a proper alignment, fatigue failure begins with small fissures, positioned below the surfaces of the raceway and rolling elements and slowly propagates to the surface producing detectable vibrations and increasing noise levels.

From a mechanical point of view, local defects inside an IM bearing cause periodic impulses in vibration signals [27]. Period and amplitude of these impulses are in a relationship with the fault position, the shaft rotational speed and the bearing geometry [27]. REBs related defects can be categorized as outer bearing race defects, inner bearing race defects, ball defects, and train defects [28]. The vibration frequencies to detect these faults can be described by the following relationships [29]:

$$f_{od}=\frac{n\cdot f_s}{2d}\left(1-\frac{d}{D}\cos(\gamma)\right),\ f_{id}=\frac{n\cdot f_s}{2}\left(1+\frac{d}{D}\cos(\gamma)\right),\ f_{bd}=\frac{D\cdot f_s}{d}\left(1-\frac{d^2}{D^2}\cos^2(\gamma)\right),\ f_{tr}=\frac{f_s}{2}\left(1-\frac{d}{D}\cos(\gamma)\right), \quad (1)$$

where f_{od} is the outer race defect frequency, f_{id} is the inner race defect frequency, f_{bd} is the ball defect frequency and f_{tr} is the train defect frequency. In these expressions, f_s is the shaft rotation frequency, n the rollers number, d and D are the roller and the pitch diameter of the bearing respectively and γ the contact angle, as depicted in Figure 2.

The knowledge of the bearing characteristic frequencies can be used to detect the cause of the defect. Based on this information it is possible to implement a condition monitoring system of the REB using vibration spectrum analysis to find the location, the cause, and the severity of defects. Traditionally, IMs condition monitoring was developed considering measurements such as vibration and temperature. The implementation of these systems could be expensive and is economical in the case of large motors or for critical applications [1]. Recently, most of the recent research [1,30] has been directed towards the monitoring of the IMs inspecting the phase current. In fact, the use of quantities that are already measured for command and control purposes in an IM, such as the machine's stator current, is favoured because allow the realization of cheaper, noninvasive and more reliable monitoring and diagnostic systems.

Moreover, since defects generate components in both vibration and current signals and the impact vibration generated by a bearing faults in their early stages have relatively low energy, it is difficult to identify the bearing fault in the spectra using conventional frequency-based approaches. The authors proposed to use a DNN for detecting the state of decay of the IM bearings able to extract from the stator current signal a compact and expressive representation of the bearings state, ultimately providing a bearing fault detection system.

To estimate the effectiveness of the proposed approach, authors carried out an experimental campaign collecting the stator current signal from an inverter-fed IM mounting different artificially damaged bearings.

The bench set has been designed within the University of Genoa, using a classic Motor-Transmission-User (MTU) system with a configuration that facilitates the bearings replacement procedure. The experimental setup consists of a three-phase IM directly connected to a brushless motor, acting as an electrical brake. An inverter controls the driving IM. In Table 2 the IM plate data are reported, while Figure 3 shows the experimental setup for collecting healthy and faulty stator current waveforms.

The stator current signals have been collected from the current sensors connected to the power inverter. This configuration ensured the maximum distance between the current sensor and the motor, reducing the noise generated by the inverter magnetic field. Since it was not possible to carry out endurance tests, which would have led to the

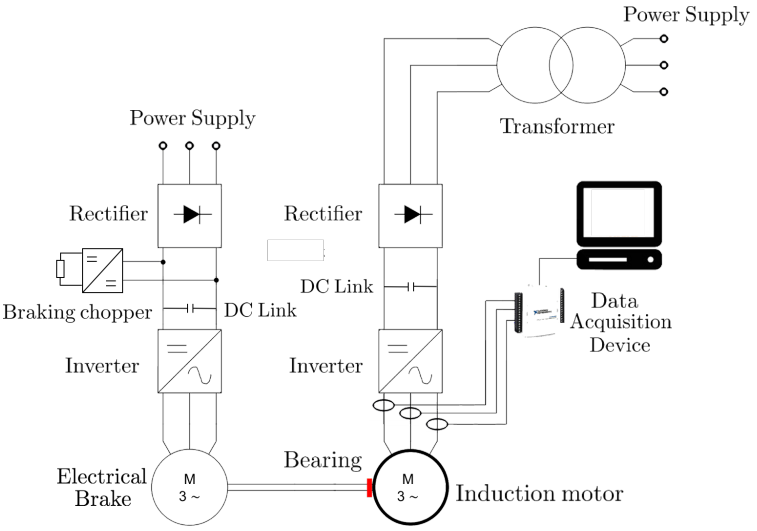

Fig. 3: Principle scheme of the experimental setup.

Table 2: IM plate data (delta-connected windings).

	Value	Unit
Rated mechanical power P_n	1.5	[kW]
Rated line-to-line voltage V_n	230	[V]
Rated line current I_n	5.90	[A]
Polar pairs p	2	
Base frequency f_n	50	[Hz]

Fig. 4: Bearings Condition: New (H0), 1.6 mm hole (H1) and 5 mm hole (H2).

breakdown of the bearings, the authors decided to introduce artificially damage into the bearing, to trace the evolution of the degradation state. The damage taken into account is located on the outer track of the bearing and can be easily inspected at the end of the tests to verify the integrity of the component. The parameters detail of the bearing used are reported in Table 1. Three identical bearings have been used covering the following bearing fault scenarios: no damages, size-1 artificial induced hole (1.6 mm), size-2 artificial induced hole (5 mm). These damages will be lately respectively referenced as H0, H1, and H2. As a first step, the experiments were carried out for the healthy bearing condition to establish the base-line data. In Figure 4 the artificially damaged bearings condition is reported. For each bearing damaged condition, four different mechanical conditions have been investigated, applying to the motor shaft different resistive torques at the same rotational speed. In the following, each mechanical condition will be identified by the corresponding stator current: 25%, 50%, 75% and 100% of the rated line current I_n, and will be lately referenced as L1, L2, L3 and L4. For each experiment, once the steady-state conditions have been reached, the stator currents have been acquired for 30 seconds. Experiments have been repeated many times in different conditions in order to build a large enough set of experiments.

3. Data Analytic Techniques

In this section, authors report the adopted workflow for the purpose of monitoring the bearings decay status. In particular authors will focus their attention on the following phases: (i) raw data collection and cleaning, (ii) data segmentation, (iii) initial simple feature mapping, (iv) advance feature mapping via deep unsupervised learning, and (v) bearings decay status prediction via deep supervised learning. These phases are depicted in Figure 5.

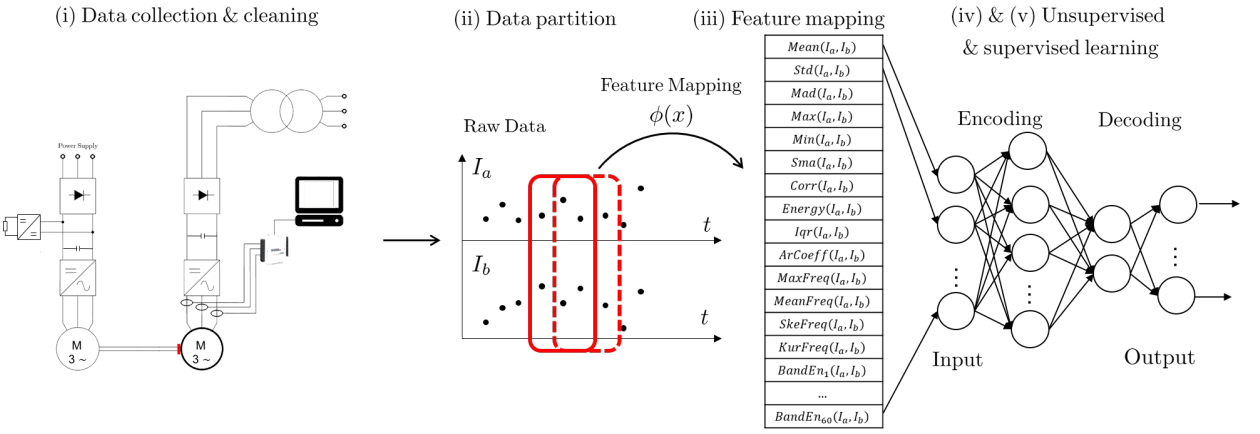

Fig. 5: Bearings decay status monitoring system workflow.

The first phase, described in details in Section 2, consists in the process of collecting the raw data with an analog to digital converter device cleaned from the higher noise frequencies. The result of this process is a time series reporting, with a sampling frequency of 20 KHz, the value of two of the three stator currents, I_a and I_b. The third one is not informative, since it can be retrieved with a simple linear combination of these two, in fact $I_a + I_b + I_c = 0$.

The second phase consists in segmenting these raw data in overlapping sliding time windows of 24 s. This quantity has been selected considering the peculiar characteristic of the studied IM, so to have a window large enough to capture the dynamics of the IM.

In the third phase, authors extracted from the windowed raw data a series of simple yet informative features, which have been chosen based on previous studies on similar context[31]. The list of these features is reported in Table 3. The result of this feature mapping is a sample $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^d$ with $d = 155$ (see Table 3) with associated its label

Table 3: Simple feature set extracted from the windowed raw data.

Time Domain		Frequency Domain
Signal magnitude area of I_a and I_b	Mean value of I_a and I_b	Largest frequency component of I_a and I_b
Correlation coefficient between I_a and I_b	Standard deviation of I_a and I_b	Frequency signal average of I_a and I_b
Average sum of the squares of I_a and I_b	Median absolute value of I_a and I_b	Frequency signal Skeewness of I_a and I_b
Interquartile range of I_a and I_b	Largest values of I_a and I_b	Frequency signal Kurtosis of I_a and I_b
Signal Entropy of I_a and I_b	Smallest value of I_a and I_b	Energy at 60 different band frequencies of I_a and I_b
Autoregression coefficients of I_a and I_b		

$\boldsymbol{y} \in \mathcal{Y}$, where $y_1 \in \mathcal{Y}_1 = \{1, 2, 3\}$ represents the decay state of the bearing (see Figure 4) and $y_2 \in \mathcal{Y}_2 \subseteq \mathbb{R}$ represents the load level (see Section 2). For each experiment and each window, a different sample which composes the data $\mathcal{D}_n = \{(\boldsymbol{x}_1, \boldsymbol{y}_1), \cdots, (\boldsymbol{x}_n, \boldsymbol{y}_n)\}$ is then provided. Based on the experiments described in Section 2 a total of $n = 1400$ samples have been collected. Consequently, authors deal with a multioutput (two labels) and multitasks problem (one label brings to a classification task while the other to a regression one)[32].

Unfortunately, even if the simple feature mapping of Table 3 is quite informative, it is characterized by some drawbacks. The first one is that this feature mapping is quite high dimensional and consequently hard to interpret for a human operator. The second one is that due to this high dimensional space and the low number of experiments (see Section 2), and consequently low number of samples, the risk is to overfit the available data instead of learning some meaningful information out of them.

To overcome these issues, an unsupervised dimensionality reduction approach must be applied to reduce the space and do not overfit the data. A simple approach is to use the Principal Component Analysis (PCA)[33]. PCA assumes that data lie in a low dimensional informative space, which have been roto-translated in a higher dimensional space. PCA can be thought of as fitting an n-dimensional ellipsoid over the data. Each axis of the ellipsoid represents a

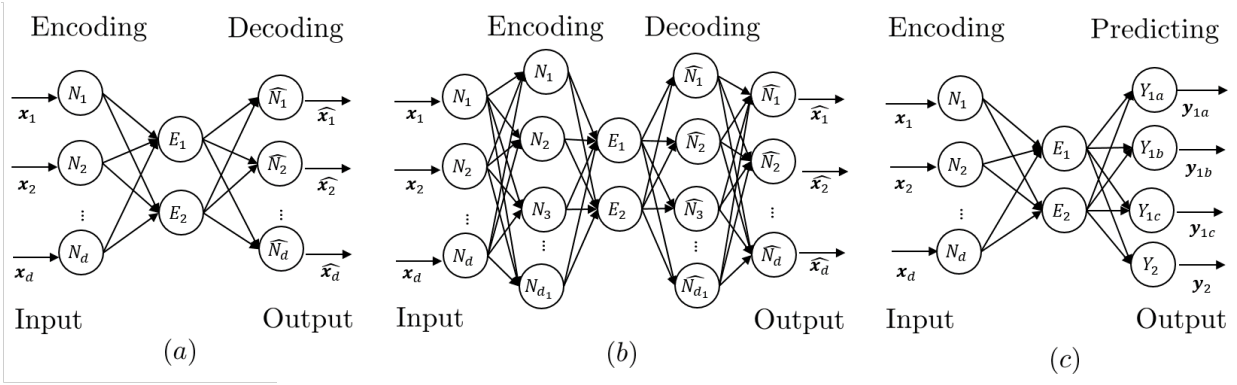

Fig. 6: NN Architectures.

new component. The larger is the axis, the higher is the variance of the data along that dimension and, consequently, the more relevant is that component as it varies more. In other words, the components with low variance are less informative. Unfortunately, as reported in Section 4, this approach is too naive, and also scaling-dependent. Hence, PCA did not allow the authors to obtain an informative low dimensional representation of the data.

For this reason, in this paper authors propose to exploit the DNN as an unsupervised dimensionality reduction approach. DNN, contrarily to the PCA, does not make any assumption on the distribution of the data and is not so sensitive to scaling. DNNs are a subfamily of the Neural Networks (NNs). NNs are a DA technique which aims at emulating the components of the brain (the neurons) with a simple mathematical abstraction, the perceptron[34], and their interaction by connecting more perceptrons together[35]. The neurons are organized in stacked layers connected together, their parameters are learned based on the available data via backpropagation[35]. If the architecture of the NN consists of a single hidden layer it is called Shallow NN (SNN), while if it is composed by multiple layers stacked together, the network is defined as DNN[26,25]. Recently, many advances have been made in this field or research by developing new neurons[36], new activation functions[37], new optimization techniques[38], new regularization methods in order to reduce the overfit tendency of complex and deep networks[39]. These advances allowed the researcher to successfully apply these methods on increasingly different and difficult real world problems. In particular, DNN have been shown to be an extremely powerful tool for feature selection and extraction purposes, both in the supervised and unsupervised context[25,26,24].

To perform the unsupervised learning feature selection process, DNNs, instead of learning the relationship between the input space $\mathcal{X}$ and the output space $\mathcal{Y}$, try to perform an often lower[24] (but sometimes higher[26]) dimensional feature mapping, which is able to explain, in a more informative way, the point sampled from $\mathcal{X}$. This architecture is called autoencoder[24] and is depicted in Figure 6.A. By staking many autoencoders it is possible to obtain a deep autoencoder[24], depicted in Figure 6.B. The stacked autoencoders can be learned incrementally by adding and learning one simple autoencoder at a time in order to avoid the gradient vanishing effect[24]. Once the representation has been learned, it is possible to use the latter in order to learn the relation between $\mathcal{X}$ and $\mathcal{Y}$ with the final DNN architecture depicted in Figure 6.C[26]. The final architecture uses the unsupervised learned representation as starting point, which is then fine-tuned based on the desired target tasks.

DNNs, as one can easily understand from the previous description, are characterized by many hyperparameters, which deeply influence the representation and generalization abilities of the architecture. In particular, the hyperparameters of a DNN are[40,24]:

- the activation function (e.g. sigmoidal, hyperbolic tangent, and rectified linear);
- the number of layers;
- the number of neurons for each layer;
- the type of regularizes and magnitude of regularization (e.g. norm of the weights, dropout, and early stopping);
- the loss function (e.g. quadratic and linear);
- the optimizer and optimization time (e.g. stochastic gradient descent and mini-batch gradient descent).

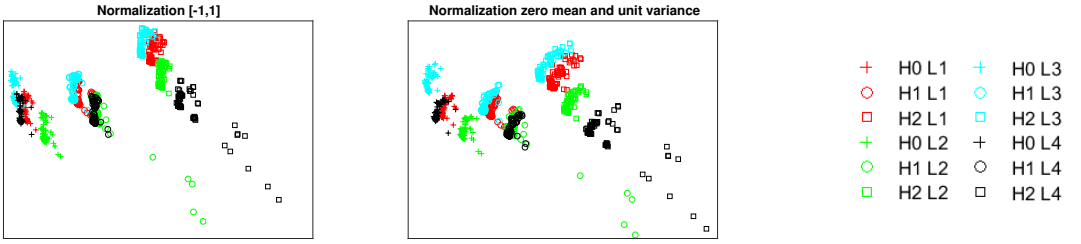

Fig. 7: Projected test point in the two-dimensional space defined by the two most informative PCA.

Moreover, after fixing the above mentioned hyperparameters the resulting architecture depends also on the initialization of the weights of the network[41]. Note that, in a DNN, this last variability is not so pronounced if the hyperparameters have been appropriately set.

Consequently, to tune these hyperparameters, it is necessary to adopt a reliable model selection strategy[42]. A common approach for tuning the hyperparameters of a leaning algorithm is to build a grid of possible configuration of hyperparameters, or to randomly select a subset of possible configurations[43], and then select the best one according to the K-Fold Cross Validation[42] of the Bootstrap[44]. Unfortunately, in a DNN, the number of hyperparameters is too high to perform this task effectively and, for this reason, the configurations are often chosen based on the experience of the data scientist and with a bit of data snooping[45]. This approach, even if commonly exploited, may lead to significant biases in estimation[46]. For this reason it is always necessary to keep apart a set of unused data, the test set, that can be seen just once, in order to test the validity of the applied procedure and selected architecture in order to report unbiased results[24,45,46].

4. Results

In this section, authors will report the results obtained by adopting the DA workflow described in Section 3, over the data collected during the trial described in Section 2, to obtain an IM Bearings fault detection system.

It is worth recalling the different campaigns of experiments performed for collecting the stator current data for a better understanding of the results. In particular, authors tested three different configurations of bearing damages: no damages, hole of 1.6 mm, and hole of 5.0 mm, respectively H0, H1, and H2. Moreover, authors tested different load conditions: 25%, 50%, 75% and 100% of the rated line current I_n, respectively L1, L2, L3, L4. Each experiment has been repeated for each condition and for a long period.

The main purpose of this study is to obtain, from the initial simple feature mapping (see Section 3 and Table 3), a compact and expressive representation of the bearings state.

For this purpose, the first approach that authors adopted was to use the PCA for reducing the dimensionality of the original data to check whether, in a lower dimensional space that can be interpreted by an operator, it is possible to give a meaningful low dimensional representation of the IMs bearing fault phenomena. Since PCA is scaling-sensitive, authors tested many normalization methods. Figure 7 reports the results obtained via PCA with different normalization methods. From Figure 7 it possible to note that PCA it is not able to represent the phenomena in a two-dimensional space since the data cloud of the different operational conditions (faults H0, H1, and H2 and loads L1, L2, L3, and L4) are overlapped.

Based on the result obtained with the PCA, authors decided to exploit an SNN and then a DNN trying to compress all the information in the two neurons in the second-last layer. Moreover, to simulate in a more realistic scenario, authors assume not to have a significant amount of labeled samples. For this reason authors exploited just $n \in \{100, 150, 200\}$ labeled samples to train the network and in Figure 8 the projected test points in the two-dimensional space, defined be the different networks, is depicted. Note that the training and the validation phases have been performed according to what described in Section 3. From Figure 8 it is possible to observe that:

- DNN and partially SNN, differently from PCA, are able to find a compact and expressive representations of the bearings damage status, by grouping the data in separate clusters based on load and damage conditions;

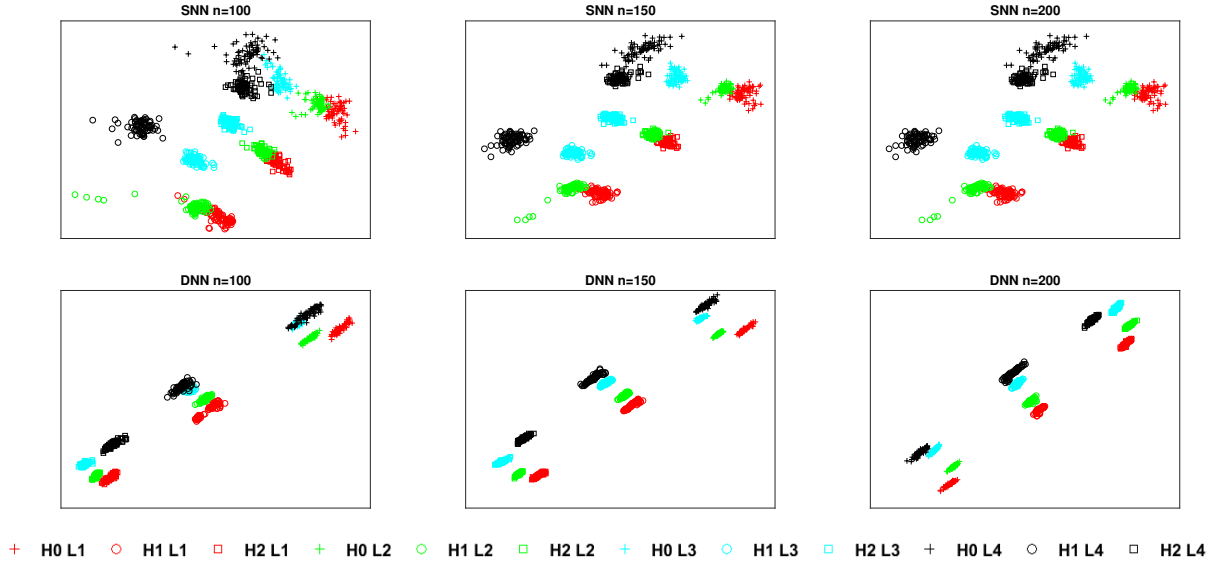

Fig. 8: Projected test point in the two-dimensional space defined by the different networks.

- both in SNN and DNN learned representation groups are ordered by load and entity of the damage;
- DNN provide clearer and more defined clusters with respect to SNN ones, showing higher classification performances even when the number of training samples is extremely limited.

In conclusion, based on the reported results, it is possible to state that the DNNs are able to extract from the stator current signal a compact and expressive representation of the bearings state, ultimately providing a bearing fault detection system.

5. Conclusions

In this paper authors dealt with the problem of assessing the performance of the induction motor bearings. Induction motors are fundamental components of any modern automation system and bearings are their first cause of faults, followed by stator winding and the rotor bar. To detect faults in the bearings, contrarily to the state-of-the-art approaches exploiting vibration signals, collected by easily damageable and intrusive vibration probes, in this paper authors exploit the stator currents signals, which are already commonly available, or easily and unintrusively collectable. Moreover, authors showed that using state-of-the-art deep neural network, instead of the now classic techniques like the PCA, it is possible to extract from the stator current signal a compact and expressive representation of the bearings state, ultimately providing a bearing fault detection system. By exploiting a series of real-data collected from an inverter-fed motor mounting different artificially damaged bearings, authors showed the effectiveness of their proposal.

References

1. M. Blodt, P. Granjon, B. Raison, G. Rostaing, Models for bearing damage detection in induction motors using stator current monitoring, IEEE transactions on industrial electronics 55 (4) (2008) 1813–1822.
2. S. Karmakar, S. Chattopadhyay, M. Mitra, S. Sengupta, Induction Motor Fault Diagnosis: Approach Through Current Signature Analysis, Springer, 2016.
3. R. Dekker, Applications of maintenance optimization models: a review and analysis, Reliability engineering & system safety 51 (3) (1996) 229–240.
4. S. Nandi, H. A. Toliyat, X. Li, Condition monitoring and fault diagnosis of electrical motors?a review, IEEE transactions on energy conversion 20 (4) (2005) 719–729.

5. A. Coraddu, L. Oneto, A. Ghio, S. Savio, D. Anguita, M. Figari, Machine learning approaches for improving condition-based maintenance of naval propulsion plants, Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment 230 (1) (2016) 136–153.
6. W. Zhou, B. Lu, T. G. Habetler, R. G. Harley, Incipient bearing fault detection via motor stator current noise cancellation using wiener filter, IEEE Transactions on Industry Applications 45 (4) (2009) 1309–1317.
7. I. Y. Önel, K. B. Dalci, I. Senol, Detection of bearing defects in three-phase induction motors using park's transform and radial basis function neural networks., Sadhana 31 (3).
8. R. R. Schoen, T. G. Habetler, F. Kamran, R. G. Bartfield, Motor bearing damage detection using stator current monitoring, IEEE transactions on industry applications 31 (6) (1995) 1274–1279.
9. F. Immovilli, M. Cocconcelli, A. Bellini, R. Rubini, Detection of generalized-roughness bearing fault by spectral-kurtosis energy of vibration or current signals, IEEE Transactions on Industrial Electronics 56 (11) (2009) 4710–4717.
10. B. Samanta, K. R. Al-Balushi, S. A. Al-Araimi, Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection, Engineering Applications of Artificial Intelligence 16 (7) (2003) 657–665.
11. N. Gebraeel, M. Lawley, R. Liu, V. Parmeshwaran, Residual life predictions from vibration-based degradation signals: a neural network approach, IEEE Transactions on industrial electronics 51 (3) (2004) 694–700.
12. V. Sugumaran, V. Muralidharan, K. I. Ramachandran, Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing, Mechanical systems and signal processing 21 (2) (2007) 930–942.
13. M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, H. Henao, Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks, IEEE Transactions on Industrial Electronics 60 (8) (2013) 3398–3407.
14. B. Yazici, G. B. Kliman, W. J. Premerlani, R. A. Koegl, G. B. Robinson, A. Abdel-Malek, An adaptive, on-line, statistical method for bearing fault detection using stator current, in: Conference Record of the 1997 IEEE, Industry Applications Conference, Thirty-Second IAS Annual Meeting Industry Applications, 1997.
15. G. B. Kliman, J. Stein, Methods of motor current signature analysis, Electric Machines and power systems 20 (5) (1992) 463–474.
16. M. E. H. Benbouzid, G. B. Kliman, What stator current processing-based technique to use for induction motor rotor faults diagnosis?, IEEE Transactions on Energy Conversion 18 (2) (2003) 238–244.
17. L. Frosini, E. Bassi, Stator current and motor efficiency as indicators for different types of bearing faults in induction motors, IEEE Transactions on Industrial electronics 57 (1) (2010) 244–251.
18. H. Su, K. T. Chong, Induction machine condition monitoring using neural network modeling, IEEE Transactions on Industrial Electronics 54 (1) (2007) 241–249.
19. B. Ayhan, M.-Y. Chow, M.-H. Song, Multiple discriminant analysis and neural-network-based monolith and partition fault-detection schemes for broken rotor bar in induction motors, IEEE Transactions on Industrial Electronics 53 (4) (2006) 1298–1308.
20. R. R. Schoen, B. K. Lin, T. G. Habetler, J. H. Schlag, S. Farag, An unsupervised, on-line system for induction motor fault detection using stator current monitoring, IEEE Transactions on Industry Applications 31 (6) (1995) 1280–1286.
21. C. T. Kowalski, T. Orlowska-Kowalska, Neural networks application for induction motor faults diagnosis, Mathematics and Computers in Simulation 63 (3) (2003) 435–448.
22. F. Filippetti, G. Franceschini, C. Tassoni, P. Vas, Recent developments of induction motor drives fault diagnosis using ai techniques, IEEE transactions on industrial electronics 47 (5) (2000) 994–1004.
23. I. Guyon, A. Elisseeff, An introduction to variable and feature selection, Journal of machine learning research 3 (Mar) (2003) 1157–1182.
24. J. Schmidhuber, Deep learning in neural networks: An overview, Neural networks 61 (2015) 85–117.
25. Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE transactions on pattern analysis and machine intelligence 35 (8) (2013) 1798–1828.
26. G. E. Hinton, S. Osindero, Y. W. Teh, A fast learning algorithm for deep belief nets, Neural computation 18 (7) (2006) 1527–1554.
27. G. B. Kliman, J. Stein, Induction motor fault detection via passive current monitoring-a brief survey, in: Meeting of the Mechanical Failures Prevention Group, 1990, pp. 49–65.
28. P. Vas, Parameter estimation, condition monitoring, and diagnosis of electrical machines, Oxford: Clarendon Press; New York: Oxford University Press, 1993.
29. H. Saruhan, S. Sandemir, A. Çiçek, I. Uygur, Vibration analysis of rolling element bearings defects, Journal of applied research and technology 12 (3) (2014) 384–395.
30. Z. Wang, C. S. Chang, X. German, W. W. Tan, Online fault detection of induction motors using independent component analysis and fuzzy neural network, in: International Conference on Advances in Power System Control, Operation and Management, 2009.
31. D. Anguita, A. Ghio, L. Oneto, X. Parra, J. L. Reyes-Ortiz, A public domain dataset for human activity recognition using smartphones., in: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2013.
32. J. Shawe-Taylor, N. Cristianini, Kernel methods for pattern analysis, Cambridge university press, 2004.
33. K. Pearson, Principal components analysis, The London, Edinburgh and Dublin Philosophical Magazine and Journal 6 (2) (1901) 566.
34. F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain., Psychological review 65 (6) (1958) 386.
35. D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors, Cognitive modeling 5 (3) (1988) 1.
36. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86 (11) (1998) 2278–2324.
37. R. H. R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, H. S. Seung, Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit, Nature 405 (6789) (2000) 947–951.
38. J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, Q. V. Le, A. Y. Ng, On optimization methods for deep learning, in: International conference on machine learning, 2011.
39. N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting., Journal of machine learning research 15 (1) (2014) 1929–1958.

40. F. Chollet, et al., Keras, `https://keras.io` (2015).
41. A. García-Gamboa, N. Hernández-Gress, M. González-Mendoza, R. Ibarra-Orozco, J. Mora-Vargas, A comparison of different initialization strategies to reduce the training time of support vector machines, in: International Conference on Artificial Neural Networks, 2005.
42. D. Anguita, A. Ghio, L. Oneto, S. Ridella, In-sample and out-of-sample model selection and error estimation for support vector machines, IEEE Transactions on Neural Networks and Learning Systems 23 (9) (2012) 1390–1406.
43. J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, Journal of Machine Learning Research 13 (Feb) (2012) 281–305.
44. B. Efron, R. J. Tibshirani, An introduction to the bootstrap, CRC press, 1994.
45. H. White, A reality check for data snooping, Econometrica 68 (5) (2000) 1097–1126.
46. M. Hardt, J. Ullman, Preventing false discovery in interactive data analysis is hard, in: Annual Symposium on Foundations of Computer Science, 2014.