A Topological and Neural Based Technique for Classification of Faults in Induction Machines

R.R. Kumar[1,3*], G. Cirrincione[2,3], M. Cirrincione[3], A. Tortella[1], M. Andriollo[1]
[1]*Department of Electrical Engineering, University of Padova*, Padova, Italy
[2]*University of Picardie Jules Verne, Lab. LTI,* Amiens, France
[3]*School Engineering and Physics, University of the South Pacific,* Suva, Fiji
*E-mail: rahulranjeev.kumar@studenti.unipd.it

***Abstract*—This paper presents a data driven approach where at first the most significant features of the three phase current signal are analyzed and then a Curvilinear Component based analysis (CCA), which is a nonlinear manifold learning technique, is performed to compress and interpret the feature behaviour. Finally, a multi-layer perceptron network is used to develop a classifier. The effectiveness of the developed model is verified experimentally with data provided on-line and in real-time.**

***Keywords*—fault diagnosis, induction machine, neural networks, pattern recognition, dimensionality reduction**

I. Introduction

In recent years, research in the area of electrical machines and drives maintenance and diagnosis has experienced a breath-taking dynamism. This has been due to their incorporation into an endless number of industrial processes and applications. Indeed, extensive use of electrical motors and generators, whose possible failures may lead to serious repercussions in monetary terms (repair costs and shutdowns) as well as other less tangible factors have drawn the attention of industry to this topic.

Electric motors have been connected to our lives on a daily basis. Some examples of these are: manufacturing systems, air transportations, ground transportations, building air-conditioner systems, home energy conversion systems, various cooling systems in electrical devices, and even cell phone vibration systems. According to authors of [1-3], the quantity of working machines in the world was expected to be around 16.1 billion in 2011, with a rapid development of 50% in the preceding five years. Among these machines, induction motors (IM) are the most common ones and are widely used in the industry. They are reliable in operations, yet are liable to various sorts of undesirable faults.

With reference to the statistics available from IEEE and EPRI for motor faults [4-6], the majority of faults associated are from the IMs. It contributes to 80% of the failed components and also tops all the categories under the failed components. This also compares with the surveys included in [7-11] and the statistics from EPRI [6] (see Table I).

Correct determination and early identification of incipient faults results in quick maintenance and short downtime for the process under consideration. It also avoids harmful, yet devastating outcomes and reduces financial loss. A perfect analytic system must be able to extract minimum amount of measurement from a machine and by investigation, extricate a correct diagnosis, so that its condition can be inferred to give a clear indication of the incipient fault in a short amount of time.

Table I. Fault Comparison

IEEE Working Group [4]	EPRI [6]	References [7-11]
44% bearing	41% bearing	40% bearing
26% winding	37% winding/stator related	38% stator related
8% rotor/shaft/coupling	10% rotor related	10% rotor related

A brief survey of diagnostic systems starts with the prevalent motor current signature analysis (MCSA) taking into account the spectrum analysis of the stator current, which is effective for electrical machines working at both steady speed and rated load. Transient conditions are also essential, and a couple of methodologies have been proposed to deal with faults in this this situation [12-15]. In addition, various MCSA-based methodologies have been proposed to detect some types of interior mechanical failures like broken rotor bar faults, bearing faults, and mechanical unbalanced rotor faults. Identification of stator voltage unbalances and other phase impacts utilizing digital signal processing strategies haves been described in some MCSA-based methodologies.

Most of fault diagnosis schemes developed about three decades back, utilized the Fast Fourier Transform (FFT) as a base technique for the analysis of motor current or vibration signature. In any case, FFT presents a few weaknesses like masking of characteristic frequencies by supply frequency, inaccuracy for transient signals, and so on. To address these drawbacks, several new techniques can be utilized.

Over the last few years, there has been an ever increasing attention to new approaches in the area of FD and Condition Monitoring (CM). The majority of these new schemes adopt methods which are oriented towards Artificial Intelligence (AI). A generic AI scheme is based on a data driven approach which involves pre-historic or stored data, feature calculation and extraction by means of signal processing and dimensionality reduction techniques, respectively, as well as a classification system to define the healthy state of the machine under consideration.

In such cases, feature calculation and extraction play a major role in enhancing the overall system accuracy. Various characteristic features are used in FD and CM for IMs. Some of the important ones are: statistical time and frequency domain features [16], envelope of the signal using Hilbert transform

[17], harmonic retrieval using non-parametric and parametric methods [18], energy and kurtosis [19] of the signal, and cepstrum of the signal [18]. Some of these features can be used to denote abrupt changes in the signal. The significance of these features entirely depend on the nature of fault incurred in the IM. In general, it is not a good practice to use a high number of feature set because this often results in overfitting of the classifier and may cause serious errors upon identification of fault.

While it is difficult to point out exactly which feature is correctly able to demonstrate anomalies in the signal, a common practice is to make use of dimensionality reduction (DR) techniques. Principal Component Analysis (PCA) and its neural variants are commonly used for DR due to its speed and simplicity, however, they are limited by the fact that they are linear. Actually, the DR techniques can be divided into two groups which are: linear and non-linear. While the latter are in general slower, they are much more accurate than the former ones in real world applications [20]. The possibility of using DR technique working in real time is quite essential due to the fact that it not only allows a batch of data to be projected fast, but also permits non-stationary data to be tracked. This mechanism can be applied to the majority of real time pattern recognition applications, where feature reduction plays an important role.

In the above paragraphs, the non-linear DR techniques has been said to be unsuitable for online applications. While many efforts have been made in order to speed up these algorithms: which involves updating of graphs, new prediction of data and embedding updating, these (e.g. iterative LLE, [21]) have proved to be cumbersome from a computational point of view. Thus, these limitations exert a demanding burden on the processing unit which slows down the entire process and makes the algorithm useless for real time applications. However, Neural networks (NN) can be also used for data projection. In general, they are trained offline and used in real time (recall phase). In this case, they work only for stationary data and can be better considered as implicit models of the embedding. Example of such NN's are the Self-Organizing Maps (SOM) [22] and their variants [23, 24].

In this work, two major faults will be studied. These are: stator inter-turn fault (SITF) and broken rotor bar fault (BRBF). Both of these types of faults are critical and evolve over time which gives rise to the non-stationary signals. The FD is carried out by analysing only the stator current spectrum and observing changes with respect to the spectrum of a healthy IM. The proposed scheme (Fig. 1) presents a data driven approach where all possibilities of the SITF and BRBF for IM under consideration have been studied (possibilities such as varying load, speed, slip). A novel real time classification has been developed by means of a non-linear fast projection technique based on Curvilinear Component Analysis (CCA) [25] and a Multi-Layer Perceptron (MLP) network [26].

The rest of the paper is organized as follows: Fault diagnosis methodology according to the proposed scheme is discussed in Section II, which also outlines the diagnostic features based on statistical features of stator current and the key ideas for feature calculation and extraction using CCA. A thorough discussion of experimental result is carried out in Section III followed by concluding remarks in Section IV.

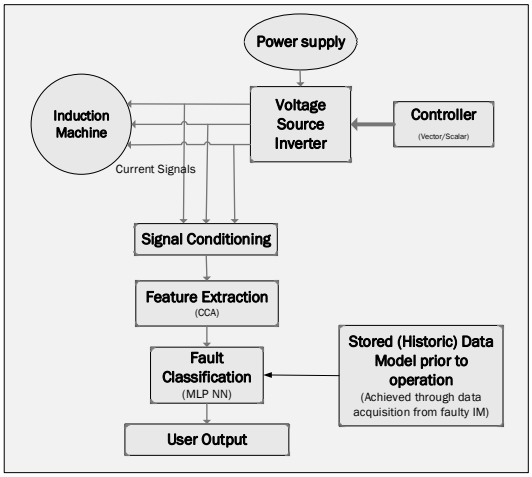

Fig.1. Proposed Scheme for fault classification

II. METHODOLOGY

A. Feature Set

After data acquisition of the 3-phase stator current signals of IM, significant features are extracted. While a healthy IM is designed to operate in symmetrical conditions, where equal resistances and inductances are present in each phase winding and rotor bar along with the air-gap, any element of non-uniformity may cause a specific characteristic which will persist at every rotation. More specifically, these asymmetry tends to be more visible in the frequency domain. Space Vectors [27] can be used as they are one of the most reliable approaches used in identifying faults in electrical drives. The goal is to acquire the direct and quadrature currents (i_d, i_q) from the 3-phase stator currents (i_{sa}, i_{sb}, i_{sc}) of the IM.

$$i_d = \sqrt{2/3}\, i_{sa} - \sqrt{1/6}\; i_{sb} - \sqrt{1/6} i_{sc} \quad (1)$$

$$i_q = \sqrt{1/2}\, i_{sb} - \sqrt{1/2}\, i_{sc} \quad (2)$$

$$i_P = \left|\left(i_d + j i_q\right)\right| \quad (3)$$

To this aim, the following features are obtained from i_P (Eq. 3), which is the norm of direct and quadrature currents (i_d, i_q) to build the feature-set (FS).

- Energy of the signal
- Kurtosis Value of the signal [19]
- Crest Factor
- Skewness of the signal
- Statistical features derived from time and frequency domains [16]
- Magnitudes of the prominent frequency harmonics

Thereafter, the Curvilinear Component Analysis (CCA) is used to reduce the dimensionality of the data to enhance generalization and avoid overfitting.

B. Dimensionality Reduction: The Curvilinear Component Analysis (CCA)

The CCA is one of the most powerful nonlinear DR technique. Actually it is derived from Sammon's mapping [28]

which proves its superiority in terms of data unfolding and extrapolation. It performs the quantization of the input space (training dataset) and projects it nonlinearly into a latent space (reduced dimension of the training dataset). Two weights are attached to each neuron. The first neuron has the dimensionality of the input space (dimension of training data, X) and the second one has a reduced dimension (dimensionality of the data in the latent space, Y).

The concept is as follows: for every pair of feature vectors in the original feature space (input space), an inter-point distance $D_{ij} = |x_i - x_j|$ is computed. The prime objective is to preserve these distances between the same points in the reduced feature space (latent space) which is given by $L_{ij} = |y_i - y_j|$, formed by reduced set of features.

To constraint the distance L_{ij} of the associated Y weights in the latent space to be equal to D_{ij}, the CCA defines a distance function (4), with a threshold λ, in order to determine short and long distances between feature vectors in the input space. In this way CCA allows matching for short distances, in a way respecting the local topology. The choice of λ, which determines the radius of the influence depends entirely on the data. In general, it is defined as three times the standard deviation of the of the feature set.

$$F_\lambda(L_{ij}) = \begin{cases} 0 \; if \; \lambda < L_{ij} \\ 1 \; if \; \lambda \geq L_{ij} \end{cases} \quad (4)$$

Equation 4 is a step function which limits only the under threshold inter-point distances L_{ij}. For each pair (i, j) of N neurons, the CCA cost function is given by:

$$E_{CCA} = \frac{1}{2}\sum_{i,j=1}^{N} E_{CCA}^{i} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(D_{ij} - L_{ij})^2 F_\lambda(L_{ij}) \quad (5)$$

Defining $y(j)$ as the weight of the j^{th} projecting neuron, the stochastic gradient algorithm for minimizing (5) is given by:

$$y(j) \leftarrow y(j) + \alpha(D_{ij} - L_{ij})F_\lambda(L_{ij})\,\frac{y(j) - y(i)}{L_{ij}} \quad (6)$$

where α is the learning rate. In this case, the longer distances are penalized and its asymmetry allows better unfolding of data. The global position of the projected map in the latent space will vary at every iteration whilst exhibiting the same topology throughout the course revealing that the CCA projection is not invariant.

Once the intrinsic dimensionality of the feature set has been correctly identified, it can be verified to what degree of mapping has occurred by using the *dy-dx* diagram. It is the plot of the distances of samples in the latent space (*dy*) versus the distances of corresponding samples in the data space (*dx*). In this scenario, it acts as a tool for the detection and analysis of nonlinearities. If the output and input space have the same dimension, then all the output inter-point distances, L_{ij} are equal to input inter-point distances, D_{ij}. Therefore, the joint distribution of input and output distances will lie along the bisector (*dy=dx*) as a consequence.

Generally, CCA is often used as a DR tool, in most cases, the output space is lower than that of the input. A "*good mapping*" is when: there is unfolding for large values (points lie on the *dy>dx* side of first diagonal) and a projection for small values (when points lie on the *dy<dx* side of the diagonal) [29]. If the distribution lies well on the diagonal, then it is possible to lower the dimension of the outer space, however, if the data cloud becomes thicker, then it implies the selected dimension size of the output space is too small to correctly represent the data. After obtaining a satisfactory output dimension value, a good practice is to vary the value of λ to align the data cloud along the bisector.

The CCA proves to be an appropriate method for representing non-linear data representation as is the case explored in this paper. It proves its superiority in terms of DR in comparison with PCA or other linear methods because of its revealing curvilinear views of more strongly folded structures. For further details regarding CCA projection see [25].

C. Classification using Neural Networks

Nowadays, Neural Networks (NN) have been used in many practical applications. In particular, they are more commonly used in system identification, prediction and classification. This is because they exhibit a property to learn complex non-linear models. The NN model is normally represented by the data itself. The drawback of the NNs is that it is prone to problems like overfitting. This should be avoided, because the NN model would have a good behaviours only for the data used under training with little generalization capability. A recommended practice is to partition the dataset into training, validation and test sets. Performing continuous validation checks upon training the NN model will ensure that the model does not overfit the training data. In this study, a multilayer perceptron (MLP) NN has been developed as a classifier to recognize the type of fault of the IM under consideration. The input of the NN classifier is the reduced set of features given by CCA and the output of the NN classifier is the class which gives the fault ID (Table II). Only four classes have been considered.

TABLE II. CLASS AND FAULT ID

Class	Fault ID
1	Healthy
2	SITF
3	BRBF
4	Combination of SITF and BRBF

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental Test Rig

The experimental test rig (Fig. 2) has been constructed to acquire stator current signals from three identical 3-phase squirrel cage IMs of 1.1kW. Data is acquired for two configurations: a) IM connected to the grid supply and is subjected to various loads for the following conditions: healthy, SITF, BRBF and combination of SITF and BRBF. b) IM is supplied by a SEMIKRON IGBT Voltage Source Inverter (VSI) of 12kVA, V/f method is used as a control scheme for the same conditions as in a). The stator currents are acquired using the LEM (LA 55-P) current transducers coupled to a DS1104 card (dSPACE). Parameters of the IMs are identical and are

shown in Table III (procedure adopted from [30]). The SITF is introduced by employing a variable power resistor connected in parallel to one of the stator phases. In this study, it occurs in phase C of the IM and the severity of the fault ranges from 0% to 10%. For BRBF, a CNC machine has been used to drill holes on the rotor bar.

TABLE III. PARAMETERS OF THE SQUIRREL CAGE IM

No. poles	4
Supply Frequency	50 Hz
Stator Resistance	3.6760 Ω
Rotor Resistance	3.8270 Ω
Stator Leakage Inductance	0.0268 H
Rotor Leakage Inductance	0.0400 H
Magnetizing Inductance	0.4490 H
Moment of Inertia	0.0059 kgm^{-2}

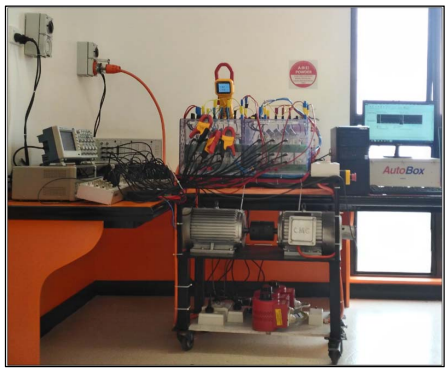

Fig. 2. Experimental Test Rig

B. *Feature Selection and Manifold Analysis*

In order to make up the FS for a four class problem, the 3-phase stator current signals were divided into equal segments. Then, using (1) – (3), Space Vector currents were obtained along with its magnitude. For every segment, features described in Part *A* of Section II are calculated. The FS comprised of 5345 observations with 33 features in total. For both healthy and faulty conditions, 70% of the data was used for training and validation, and 30% for the test set.

While too many features would cause overfitting or may prove difficulty in learning the data manifold when training the classifier, a pre-processing step becomes necessary to reduce dimensionality of the FS. The intrinsic dimensionality of the FS can be estimated using [31, 32]. From Table IV the estimated intrinsic dimensionality of the FS using a comparative approach, ranges in between 6 and 7.

TABLE IV. INTRINSIC DIMENSIONALITY OF THE FS

Method [31, 32]	**Intrinsic dimensionality: *(Healthy, SITF and BRBF)***	**Intrinsic dimensionality: *(Healthy, SITF and BRBF, Combined SITF & BRBF)***
Maximum Likelihood estimator (MLE)	6	6
PCA eigenvalues	6	7
Nearest Neighbour	-0.06	-0.06

As a result, all the above stated dimensions including 8 are studied to confirm the best dimensionality of the FS which consists data for healthy and all types of faults for the IM. The CCA approach is used in an original way, creating three different reduced FS of dimensions: 6,7 and 8. The analysis for each reduced FS is carried out by visualizing its corresponding *dy-dx* diagrams. It must be noted that prior to DR using CCA, the FS is normalized to retain an overall generalization property. Figures 3-5 represent the *dy-dx* diagrams for each reduced FS with their best CCA λ parameter.

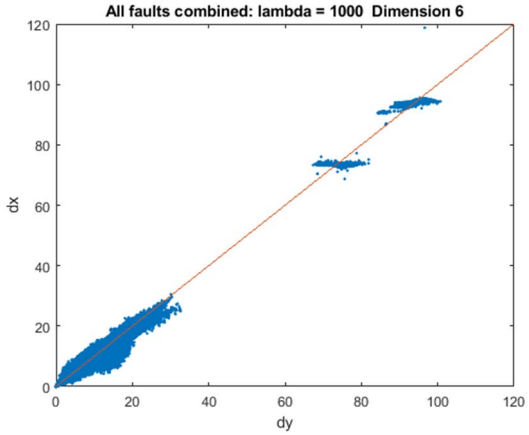

Fig. 3. *dy-dx* plot of reduced FS of dimension 6

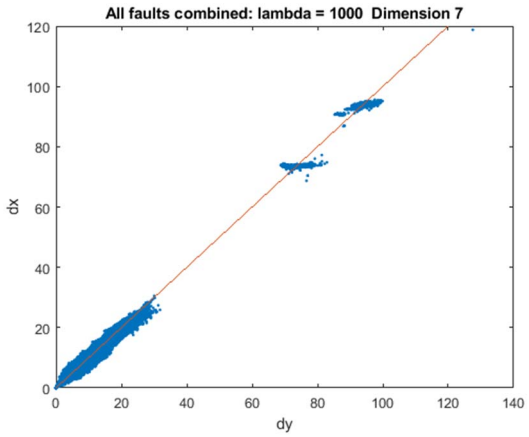

Fig. 4. *dy-dx* plot of reduced FS of dimension 7

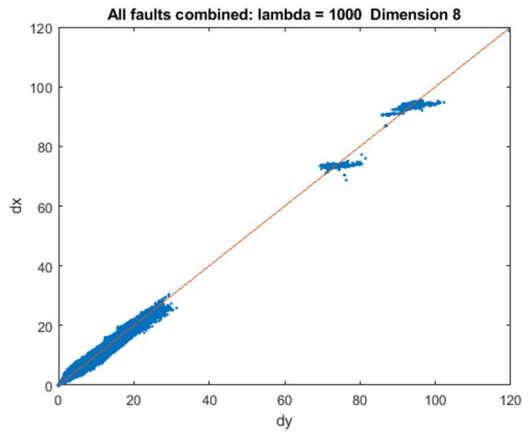

Fig. 5. *dy-dx* plot of reduced FS of dimension 8

Upon comparison of the plots above, it can be noticed that the dimension of 6 is ranked below others. This is due to the fact that it has a thicker cloud of data which moves out towards the *dx* line even at its best value of λ (see bending of data cloud in Fig. 3). Generally, the more the deviation of data cloud with respect to the bisector, the more nonlinear the manifold is, implying that the selected dimension is lower and it needs to be increased. Hence, it can be inferred that the intrinsic dimensionality of the original FS is between 7 and 8.

To select appropriate dimension of the original FS, the data cloud in the *dy*-*dx* diagram for dimensions 7 and 8 is zoomed and inspected to see if there are any major deviation from the bisector and any differences between the two plots (Figs. 6-7)

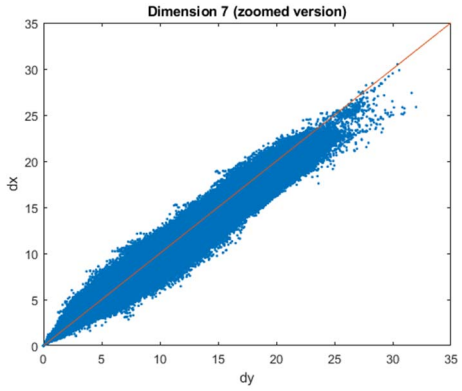

Fig. 6. *dy*-*dx* plot of reduced FS of dimension 7 (zoomed version)

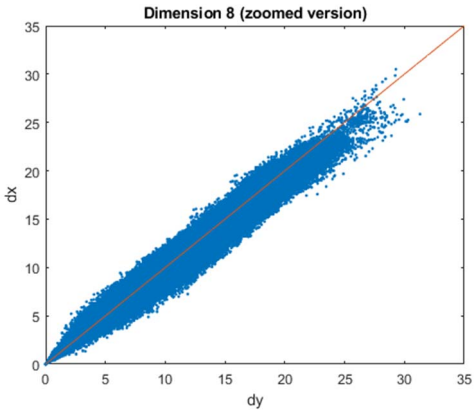

Fig. 7. *dy*-*dx* plot of reduced FS of dimension 8 (zoomed version)

Considering Table IV, Figs. 3-5 (intra and inter class distances) and the zoomed versions of *dy*-*dx* plots (Figs. 6-7), the best choice for dimensionality of the FS can be stated as 7. This is because the *dy*-*dx* plots for dimension of 7 and 8 are very much similar. Referring to Fig. 4, the *dy*-*dx* plot after projecting the training set to the reduced dimension space via CCA reveals that while the manifold is slightly nonlinear (proportional to the point thickness around the bisector near the origin), the cluster of points close to origin (intra-class distances) and the presence of two distinct clusters (inter-class distances) show the existence of at least three manifolds in the feature space, which are related to the classes. The same can be said for the reduced FS of dimension 8 (Fig. 5 in this case). Thus, this justifies the choice of selecting 7 as the final intrinsic dimension of the FS. Figure 8 gives a 3D interpretation of the first three CCA components of the reduced dimension space. Indeed, there's is not much overlap between the clusters.

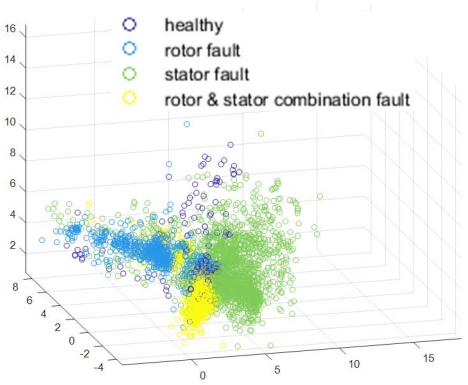

Fig. 8. Visualization of the first three CCA components of the 7 dimensional reduced space

C. *Classification using MLP NN*

After the DR step, a MLP NN is trained. As this is a pattern recognition problem, for classification of the type of fault, various choices of the number of neurons in the hidden layer have been tried. The final MLP model has 141 neurons in the hidden layer. The output of the network consists of *softmax* transfer functions [33] which output both class, new observations and the corresponding probability. The training algorithm for the network is based on the *scaled conjugate gradient technique* [34] and the performance of the network is consistently calculated by the *cross entropy* error function [26]. The results in form of the confusion matrix of the training data and the test data are given in Figs. 9 and 10, respectively. Note the class labels and its corresponding fault ID (Table II).

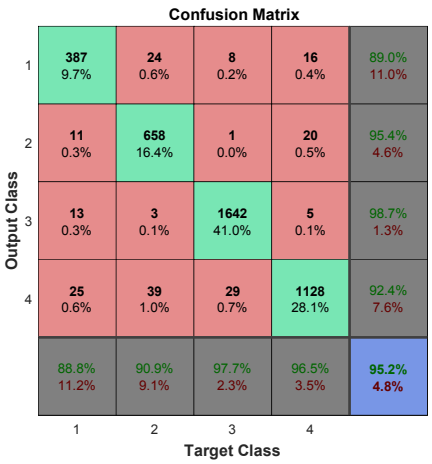

Fig. 9. Confusion Matrix for the training set

It must be noted that the above results have been obtained on an online basis. Starting from acquisition of the 3-phase stator current signal to calculation of features, the pre-trained CCA network is used to project each segmented data (dimensionality same as the original FS) to a lower dimension space to make the reduced FS classifier compliant. Classification is finally carried out by MLP NN.

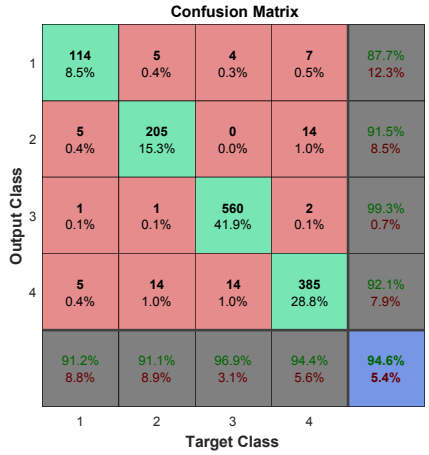

Fig. 10. Confusion matrix for the test set

IV. CONCLUSION

Using the proposed methodology, classification of stator and rotor related faults can be achieved at varying slips on an online basis. The 3-phase stator current signal reveals a vast amount of information on the performance of the IM. This is useful for detecting and classifying different kinds of faults incurred in the IM. Using Space Vectors, the 3-phase currents are reduced to direct and quadrature currents and finally combined using the modulus of both (i_P). Then, the features of the signals which refer to statistical properties of the time and frequency domains are calculated. Thereafter, CCA is used to project the original FS into a reduced dimension space (based on the intrinsic dimensionality of the original FS). Finally, MLP NNs are trained and the best configuration which is based on lower *cross entropy* error and higher validation accuracy is chosen. Following the scheme of Fig. 1, the whole process (starting from acquiring 3-phase current signals to classification) was carried out by using the pre-trained sets of CCA projection and the MLP based classifier. The overall classification accuracies show 95.2% and 94.6% for training and test sets, respectively.

REFERENCES

[1] H. Henao, G.-A. Capolino, M. Fernandez-Cabanas, F. Filippetti, C. Bruzzese, E. Strangas, *et al.*, "Trends in fault diagnosis for electrical machines: a review of diagnostic techniques," *IEEE industrial electronics magazine,* vol. 8, pp. 31-42, 2014.

[2] F. Filippetti, A. Bellini, and G.-A. Capolino, "Condition monitoring and diagnosis of rotor faults in induction machines: State of art and future perspectives," in *2013 IEEE Workshop on Electrical Machines Design, Control and Diagnosis (WEMDCD)*, 2013.

[3] H. A. Toliyat, S. Nandi, S. Choi, and H. Meshgin-Kelk, *Electric machines: modeling, condition monitoring, and fault diagnosis*: CRC press, 2012.

[4] "IEEE Recommended Practice for the Design of Reliable Industrial and Commercial Power Systems - Redline," *IEEE Std 493-2007 (Revision of IEEE Std 493-1997) - Redline,* pp. 1-426, 2007.

[5] S. Karmakar, S. Chattopadhyay, M. Mitra, and S. Sengupta, "Induction Motor Fault Diagnosis."

[6] G. Singh, "Induction machine drive condition monitoring and diagnostic research—a survey," *Electric Power Systems Research,* vol. 64, pp. 145-158, 2003.

[7] M. E. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE transactions on industrial electronics,* vol. 47, pp. 984-993, 2000.

[8] G. B. Kliman, R. A. Koegl, J. Stein, R. D. Endicott, and M. W. Madden, "Noninvasive detection of broken rotor bars in operating induction motors," *IEEE Transactions on Energy Conversion,* vol. 3, pp. 873-879, 1988.

[9] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors-a review," *IEEE transactions on energy conversion,* vol. 20, pp. 719-729, 2005.

[10] A. Siddique, G. Yadava, and B. Singh, "A review of stator fault monitoring techniques of induction motors," *IEEE transactions on energy conversion,* vol. 20, pp. 106-114, 2005.

[11] Y. Zhongming and W. Bin, "A review on induction motor online fault diagnosis," in *Power Electronics and Motion Control Conference, 2000. Proceedings. IPEMC 2000. The Third International*, 2000, pp. 1353-1358.

[12] Y. Gritli, L. Zarri, C. Rossi, F. Filippetti, G.-A. Capolino, and D. Casadei, "Advanced diagnosis of electrical faults in wound-rotor induction machines," *IEEE Transactions on Industrial Electronics,* vol. 60, pp. 4012-4024, 2013.

[13] S. H. Kia, H. Henao, and G.-A. Capolino, "Diagnosis of broken-bar fault in induction machines using discrete wavelet transform without slip estimation," *IEEE Transactions on Industry Applications,* vol. 45, pp. 1395-1404, 2009.

[14] M. Riera-Guasp, J. A. Antonino-Daviu, M. Pineda-Sanchez, R. Puche-Panadero, and J. Perez-Cruz, "A general approach for the transient detection of slip-dependent fault components based on the discrete wavelet transform," *IEEE Transactions on Industrial Electronics,* vol. 55, pp. 4167-4180, 2008.

[15] A. Stefani, A. Bellini, and F. Filippetti, "Diagnosis of induction machines' rotor faults in time-varying conditions," *IEEE Transactions on Industrial Electronics,* vol. 56, pp. 4548-4556, 2009.

[16] B. Li, P.-l. Zhang, Z.-j. Wang, S.-s. Mi, and D.-s. Liu, "A weighted multi-scale morphological gradient filter for rolling element bearing fault detection," *ISA transactions,* vol. 50, pp. 599-608, 2011.

[17] S. L. Hahn, *Hilbert transforms in signal processing* vol. 2: Artech House Boston, 1996.

[18] M. H. Hayes, *Statistical digital signal processing and modeling*: John Wiley & Sons, 2009.

[19] K. V. Mardia, "Measures of multivariate skewness and kurtosis with applications," *Biometrika,* vol. 57, pp. 519-530, 1970.

[20] L. Van Der Maaten, E. Postma, and J. Van den Herik, "Dimensionality reduction: a comparative," *J Mach Learn Res,* vol. 10, pp. 66-71, 2009.

[21] J. Weng, Y. Zhang, and W.-S. Hwang, "Candid covariance-free incremental principal component analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 25, pp. 1034-1040, 2003.

[22] X. Qiang, G. Cheng, and Z. Li, "A survey of some classic self-organizing maps with incremental learning," in *Signal Processing Systems (ICSPS), 2010 2nd International Conference on*, 2010, pp. V1-804-V1-809.

[23] B. Fritzke, "A growing neural gas network learns topologies," in *Advances in neural information processing systems*, 1995, pp. 625-632.

[24] T. Martinetz and K. Schulten, "A" neural-gas" network learns topologies," 1991.

[25] P. Demartines and J. Hérault, "Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets," *IEEE Transactions on neural networks,* vol. 8, pp. 148-154, 1997.

[26] C. M. Bishop, *Neural networks for pattern recognition*: Oxford university press, 1995.

[27] M. Cirrincione, M. Pucci, and G. Vitale, *Power converters and AC electrical drives with linear neural networks*: CRC Press, 2017.

[28] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on computers,* vol. 100, pp. 401-409, 1969.

[29] G. Cirrincione and A. Marvuglia, "A Novel Self-organizing Neural Technique for Wind Speed Mapping," in *Sustainability in Energy and Buildings*, ed: Springer, 2009, pp. 209-217.

[30] P. Krause, O. Wasynczuk, S. D. Sudhoff, and S. Pekarek, *Analysis of electric machinery and drive systems* vol. 75: John Wiley & Sons, 2013.

[31] J. A. Lee and M. Verleysen, *Nonlinear dimensionality reduction*: Springer Science & Business Media, 2007.

[32] E. Levina and P. J. Bickel, "Maximum likelihood estimation of intrinsic dimension," in *Advances in neural information processing systems*, 2005, pp. 777-784.

[33] W. Duch and N. Jankowski, "Survey of neural transfer functions," *Neural Computing Surveys,* vol. 2, pp. 163-212, 1999.

[34] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural networks,* vol. 6, pp. 525-533, 1993.