Detection of Stator Fault in Synchronous Reluctance Machines Using Shallow Neural Networks

Siwan Narayan
School of Information Technology, Engineering and Physics
The University of the South Pacific
Suva, Fiji
siwan.narayan@usp.ac.fj

Rahul R Kumar
School of Information Technology, Engineering and Physics
The University of the South Pacific
Suva, Fiji
rahul.kumar@usp.ac.fj

Giansalvo Cirrincione
Department of Electrical Engineering, Lab. LTI
University of Picardie Jules Verne
Amiens, France
giansalvo.cirrincione@u-picardie.fr

Maurizio Cirrincione
School of Information Technology, Engineering and Physics
The University of the South Pacific
Suva, Fiji
maurizio.cirrincione@usp.ac.fj

2021 IEEE Energy Conversion Congress and Exposition (ECCE) | 978-1-7281-5135-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ECCE47101.2021.9595518

***Abstract*—Fault detection in electrical drives can be really challenging, especially when the input data is collected from an operational electrical machine. In order to prevent machine damages and downtimes, it is really important to detect pre-fault conditions. This paper presents the detection of stator inter-turn fault for Synchronous Reluctance Motor (SynRM) with a severity as low as 1.3%. After the transformation of the three-phase currents using Extended Park Vector (EPV) approach, the temporal features were calculated. Thereafter, the geometry of the features has been studied by using the Principal Component Analysis (PCA) and the Curvilinear Component Analysis (CCA) to estimate the best intrinsic dimensionality and extract the most significant features. Finally, a variety of classifiers have been trained with this feature-set (FS) and the shallow neural network has proved to give the best performance.**

Keywords—synchronous reluctance motors, diagnostics, shallow neural networks, data, classification, Park's transformation

I. Introduction

Over the recent years, a wide research interest in Synchronous Reluctance Motors (SynRM) has been ever increasing. Indeed, SynRMs have drawn immense attention due to high demand in high power and high speed applications. For the past few decades, the SynRM has been considered as valid alternative to the Permanent Magnet Synchronous Motor (PMSM) due to its structural simplicity, lower prices, and mechanical robustness [1-3]. The high dynamic performance, the existence of cold rotor, the capability of working in a wide speed range, and the ability to operate under deep flux weakening are all properties that make the SynRM one of the most interesting electrical machines [4].

Although the SynRM has regained a widespread interest in research, this has been generally limited to either motor design or motor control techniques. On the other hand, there has been a number of research on the detection and classification of faults in electrical machines, but this has not been as extensive when it comes to SynRMs. Actually, SynRMs are utilized ever increasingly in motor drive applications nowadays and it is consequently important to detect and classify SynRM fault conditions. This paper wants to give a contribution with this respect and focusses on one class of motor faults: stator faults in SynRM. The rationale of this choice is given in the next paragraph. Internal Faults in electrical machines can be classified into two major categories; mechanical and electrical faults, which are sub-divided into other specific fault types as shown in Fig. 1:

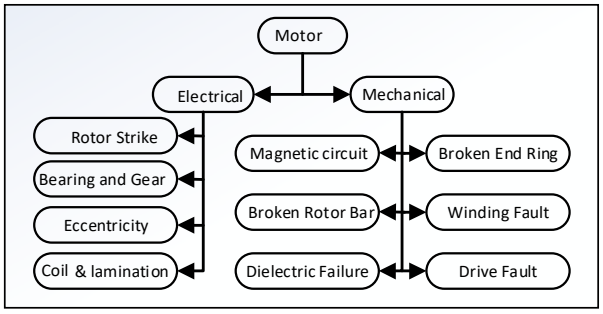


Fig. 1. Classification of Faults in Electrical Machines [5]

According to the statistics available from EPRI for motor faults [6], 41% of overall synchronous motor faults are bearing and gear related, 37% are stator winding related, and 10% are rotor related. As stated by [7], a number of surveys relating to electrical machine failure indicate that in general, almost 37% of any motor failure are related to stator winding faults. Early identification and precise determination of machine faults will not only help in quick maintenance, but reduce downtime as well as prevent financial losses and permanent damages to the machines. It is most important for SynRM drives being used in dangerous applications to have reliable fault detection systems together with an appropriate remedial action to allow the drive to maintain uninterrupted operation. In order to have an accurate and reliable classification and detection of motor faults, it is fundamental to employ an analysis method which can use minimum amount of data extracted from the machine to classify and detect motor faults at its early stages without interrupting the operation.

This paper presents the detection and classification of the stator inter-turn fault in SynRMs. Using the Extended Park

Vector (EPV) transformation of the stator current, the temporal [8] features are calculated. Afterwards, with the use of the Principal Component Analysis (PCA) and the Curvilinear Component Analysis (CCA), the data geometry and topology are studied to estimate the intrinsic dimensionality of the Feature Set (FS). After reducing the dimension of the FS, neural and non-neural based classifiers are trained to determine the best classification model. Finally, this approach has been assessed experimentally.

II. METHODOLOGY

Figure 2 below gives an overview of the approach and the different steps are described in the following.

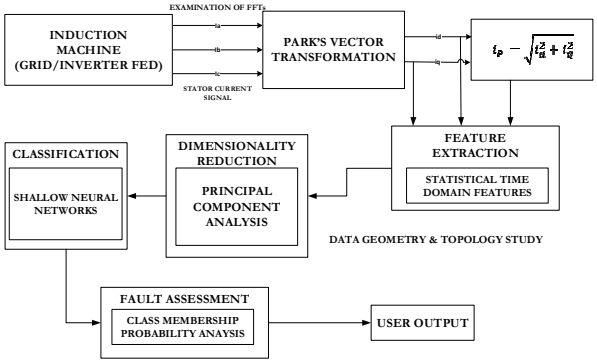


Fig. 2. Flowchart of the proposed methodology

A. Feature Extraction

In order to extract the features from the signal, the most reliable approach would be to use the Extended Park Vector (EPV) transformation of the stator current where the direct and quadrature currents (i_d, i_q) are acquired from the three-phase stator currents (i_{sa}, i_{sb}, i_{sc}) of the SynRM.

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

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

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

Using the EPV current i_p, 15 temporal features have been generated. The set formed by all of these 15-component vectors is the feature set FS. These features are adapted from [8] and listed as follows: mean, maximum value, root mean square (RMS), square root mean (SRM), standard deviation, variance, RMS shape factor, SRM shape factor, crest factor, latitude factor, impulse factor, skewness, kurtosis, 5th moment and 6th moment. After this step, the PCA technique is used in two ways: (1) to study the geometry of the data and explore the differences between healthy and faulty clusters to infer appropriate intrinsic dimensionality and, (2) to reduce the dimensionality of the features on the basis of variability of the data, so that they can be ready for classification.

B. Data Geometry and Topology

After generation of the features, the next step is the interpretation of the geometry and topology of the data [9]. While many techniques exist for the so-called Dimensionality Reduction (DR), a rule of thumb is to first explore the features using linear techniques for their speed and reliability, and then confirm it by using non-linear approaches. To reduce the input dimension, it is essential to estimate the intrinsic dimensionality of the FS; however, prior to that, the FS should be normalized, so that each attribute has zero mean and variance equal to one.

In this paper, to study geometry of the data, the authors use PCA [10], the Pareto charts and the data projected along the first three PCs. The Pareto chart gives details on the percentage variability of each PC together with the cumulative values as the PCs accumulate. The data projected along the first three PCs shows how the class clusters are positioned inside the 3-D space. This phase allows the user to observe either the separation or the overlapping of class clusters. In this way, the intrinsic dimensionality is estimated.

An additional analysis of the FS is also carried out by the CCA (Curvilinear Component Analysis - [11, 12]), another non-linear DR (Dimension Reduction) technique. The target of the CCA is to confirm the deductions made by the PCA (confirm the intrinsic dimensionality value deduced by PCA analysis). The CCA is one of the most powerful nonlinear DR techniques. It is derived from Sammon's mapping, but it proves its superiority in terms of data unfolding and extrapolation [11]. Essentially, it performs the quantization of the input space (training dataset) and projects it nonlinearly into a latent space (reduced dimension of the training dataset). For further details regarding CCA projection (and applications) see [8, 9, 11, 13].

CCA can show how good the DR is by using the *dy-dx* diagram to study the mapping between the input data space and the latent space with the chosen intrinsic dimensionality value. This diagram is the plot of the distances of samples in the latent space (*dy*) versus the distances of corresponding samples in the input data space (*dx*). In this scenario, it acts as a tool for the detection and analysis of nonlinearities. In most cases, the output space is lower than that of the input. A *"good mapping"* is achieved when the unfolding occurs for large values (points lie on the *dy>dx* of the diagonal) and a projection arises for small values (when the points lie on the *dy<dx* side of the diagonal). If the points lie well onto the diagonal (where *dy* = *dx*), it is possible to lower the dimension; otherwise, a "*thicker*" data cloud implies that the chosen dimensionality is very small and in order to attain a good mapping, the dimensionality value needs to be increased.

Thus, as a final remark on the study of the data geometry and topology, PCA is initially applied to the FS to have a generic idea on the intrinsic dimensionality and its corresponding linear projection. Then, the CCA analysis is carried out (using *dy-dx* diagrams) only to confirm on the intrinsic dimensionality value deduced by PCA. Through these verifications, appropriate intrinsic dimensionality for the FS is deduced and in this paper, the authors utilize PCA for DR as the data manifold was found to be linear according to the results presented in section III.

C. Fault Detection and Classification

In recent years, Neural Networks (NN) have been widely used in many practical applications. Actually, it is quite commonly used in classification, prediction and also system identification. In this paper, a multilayer perceptron (MLP) NN classifier has been developed to detect the stator-based fault and also yield the fault severity. The reduced set of features obtained from the PCA acts as the input to the NN and the class membership together with its probabilities acts as the output of the classifier. Also, for the purpose of comparison, non-neural based classifiers have been trained and tested using the same FS (partitioned into train and test sets).

III. Experimental Results and Discussion

A. Experimental Test Rig

The experimental test rig, shown in Fig. 3, has been developed to measure the three-phase stator currents from a 2.2kW SynRM with parameters shown in Table I.

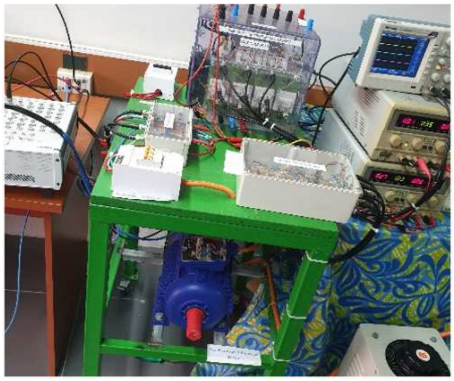


Fig. 3. Experimental Test Rig

The experimental test rig and its component layout, as shown in Figs. 3 and 4, have been developed to extract the three-phase stator currents from the 2.2kW SynRM. The three-phase power supply is connected to the 12kVA SEMIKRON IGBT voltage source inverter, which drives the SynRM. The three-phase stator current signals are acquired using the LEM current transducers (LA 55-P) that are connected to dSPACE MLBX (DS1202) using the BNC cable.

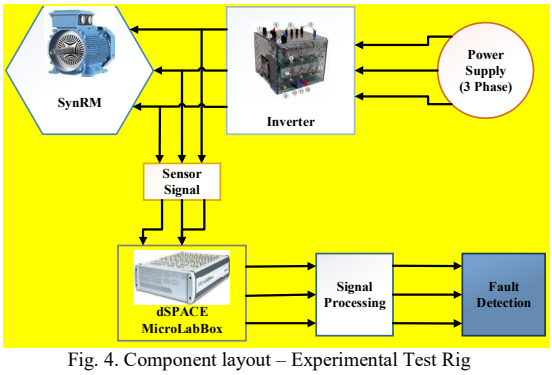


Fig. 4. Component layout – Experimental Test Rig

Table I below lists the parameters of the SynRM that have been used to obtain the current signatures for healthy and faulty conditions.

TABLE I. Parameters of SynRM

Fig. 5. 2.2 kW SynRM

Power Rating	2.2 kW
Number of Poles	8
Frequency	50 Hz
Rated Speed	750 rpm
Rated Voltage	365Y
Rated Current	5.7A
Rated Torque	14Nm
Weight	24 Kg

B. Stator inter-turn fault emulation – Hardware

This section describes the acquisition of the three-phase stator current signals (non-invasively) of the healthy and faulty SynRM. In particular, the low severity levels of the stator inter-turn faults have been acquired under no load conditions. The faulty motor data is taken at 3 different severity levels, that is when stator winding is short-circuited at 1.3%, 1.5%, and at 2.07% (pre-faults), as explained further below. The authors have chosen these low severity faults at no-load, because detection of faults under these circumstances is extremely difficult especially when the no-load conditions is at low speeds. For the acquisition of the healthy phase current signature, the motor is operated in its optimal condition and the three-phase currents are retrieved with a Hall-effect sensor through the dSPACE MLBX platform.

For the implementation of stator inter-turn fault, the winding coil short circuit is characterized by an appreciable shunt resistance R_{shunt}, and the number of shorted winding turns is emulated by varying the R_{shunt} according to the fault severity. For this purpose, the fault in the SynRM described in Fig. 3 is emulated in phase A windings with the R_{shunt} values of 0.75Ω, 1.1Ω and 2.2Ω for 1.3%, 1.5%, and 2.07%, fault severities, respectively. The phase currents are obtained with a sampling frequency of 1 kHz.

C. Feature Generation and Data Partition

After the computation of the extended Park's Vector, the previously mentioned 15 temporal features of the i_P current have been calculated using a moving rectangular window with width of 100 samples. The heathy data have consisted of 5710 samples and the faulty data have consisted of 19128 samples. Subsequently, the data corresponding to the 1.3%, 1.5% and 2.07% severities have been assigned to separate classes. With this aim, the stator fault with 1.3% severity consisted of 5303 samples, 1.5% was made up of 8115 samples and 2.07% comprised 5710 samples.

For classification (neural and non-neural based), the following data division is considered: 50% for training and validation; 50% for the test set. This ratio of data partitioning has been selected to critically observe how each classifier behaves and measure their learning in terms of accuracy. For classification problems, the accuracy [14] is defined as:

$$\%\ Classification\ Accuracy = \frac{TP + TN}{FP + FN} \times 100 \qquad (4)$$

where: TP-True Positive, TN-True Negative, FP-False Positive and FN-False Negative

D. Geometrical Analysis of Data

After the feature generation and data partitioning, an essential step prior to classification is the study of the geometry of the data by linear techniques. Under these circumstances, the PCA was utilized and through Pareto chart (Fig. 6), the intrinsic dimensionality was deduced as 4, stating that only 4 Principal Components (PCs) were sufficient to explain over 95% of variability in the dataset. Using this "*explained variability*" metric, the cumulative variance of 100% was attained with only 4 PCs. Thus, projecting from a dimension of 15 (initial dimensionality of FS) to a dimension of 4, linearly (via PCA) and retaining the cumulative variance, implies that the intrinsic dimensionality of the FS should be around 4. Confirmation of this value is done via CCA analysis in the following sub-section.

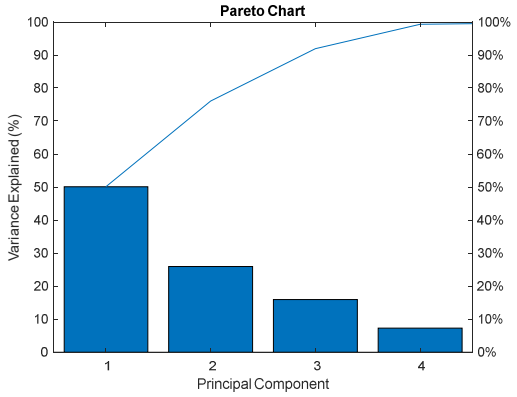
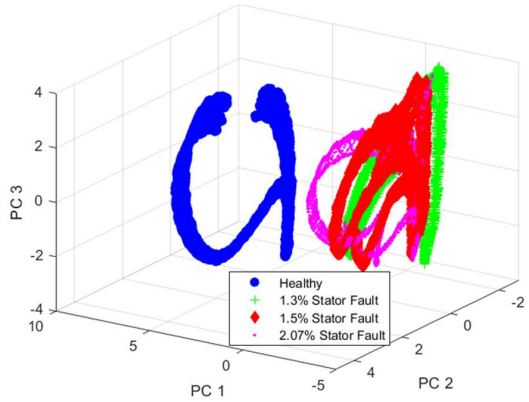
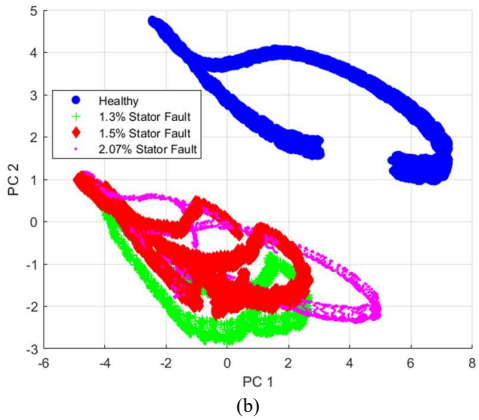


Fig. 6. Pareto Chart

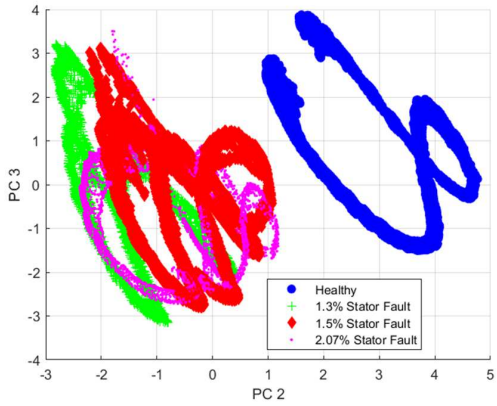
After application of PCA, data was projected in its first three PC's as illustrated in Fig. 7, which also denotes the healthy and faulty conditions. It should be noted from Fig. 7a-c that the healthy cluster appears to be well separated from the faulty ones; however; there is a certain level of overlapping among the faulty class clusters. This means that in terms of fault detection (healthy versus faulty), there are fewer chances for misclassification and the probability of the class output upon detection should be above the 50% confidence level. On the other hand, there is a high degree of overlapping among the faulty class clusters and therefore linear based classification techniques cannot be used. As a result, for classification, various neural and non-neural nonlinear classifiers have been trained and tested to determine the best model. This is portrayed in section IV.



(a)



(b)

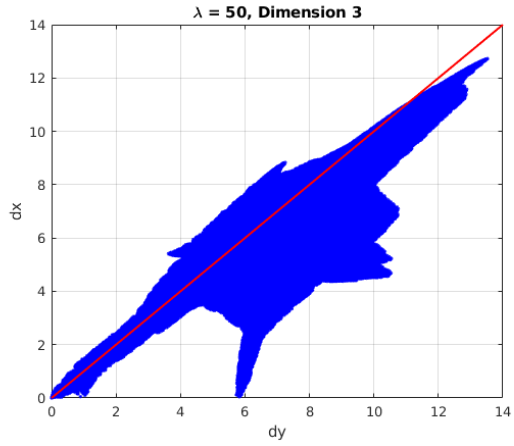


(c)

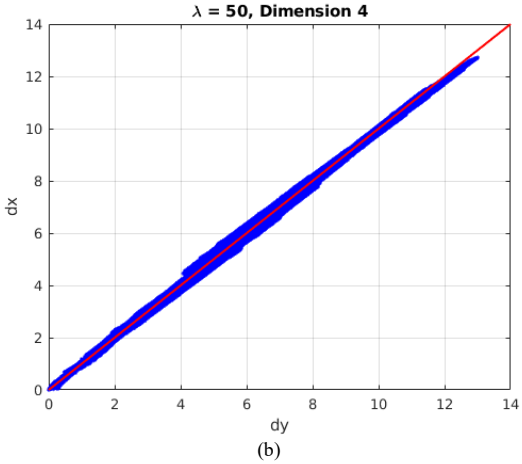
Fig. 7. PCA Projection Plots: (a) 3D view with the first 3 PCs, (b) 2D view with PC 1 and PC 2, (c) 2D view with PC 2 and PC 3

E. *CCA Analysis*

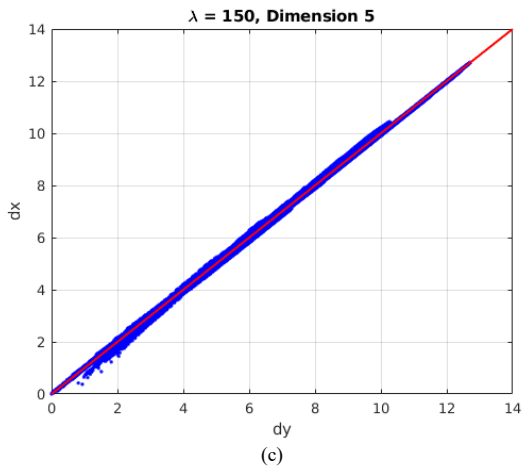
While the inferred intrinsic dimensionality deduced via PCA (Pareto chart) is 4, to confirm this, the 15 dimensional FS has also been analyzed using CCA. With this aim, starting from the dimension of 4, the *dy-dx* diagrams for dimensions 3 and 5 (4 ± 1 , with 4 being the inferred intrinsic dimensionality from PCA) were also explored to check the input-output mapping. Figures 8a-c represents the *dy-dx* diagrams for dimensions 3-4, respectively.



(a)



(b)



(c)

Fig. 8. Input-output mapping using CCA for: (a) Dimension 3, (b) Dimension 4, and (c) Dimension 5

While in Fig. 8a (dimension of 3) the scattering of the data indicates a poor representation, the input-output mapping is well presented for dimensions 4-5 (Figs. 8b-c). This results from the fact that the data cloud in dimensions 4-5 (Figs. 8b-c) lies along the bisector ($dy = dx$), showing that the topology is preserved with these dimensionality values. As for the dimension of 3, the majority of the data cloud is not well aligned to the bisector.

Thus, through these justifications, it can be concluded that the data manifold is linear (for dimensions 4 and 5) and the intrinsic dimensionality lies between the dimensions 4-5. Hence, for DR, the intrinsic dimensionality of 4 has been chosen.

IV. Classification

For the purpose of the classification, various families of classifiers have been trained and tested. The classification accuracy (test set) for all the classifiers have been calculated using equation (4) and are summarized in Table II. It can be concluded that the "*shallow dense neural based classifier*" performs better than all the others. In particular, it has a test accuracy of as much as 99.99%. Unlike the other 14 classifiers (listed in Table II), the neural based classifier is able to output the class together with its membership probabilities without any further calculations. This information is essential when it comes to subsequent risk analysis. In addition, it has the lowest computational complexity. This results from the fact that during the recall phase (i.e. during the classification phase) the computational complexity is proportional to the numbers of samples fed to the network. In this study, for the shallow neural network classifier, the computational complexity is linear .i.e. $O(n)$.

To assess the validity of the proposed shallow neural network based classifier, the confusion matrix in Fig. 9 has been plotted. It shows the correct and incorrect predictions where:

- Class 1- Healthy,
- Class 2 – 1.3% Stator fault severity,
- Class 3 – 1.5% Stator fault severity,
- Class 4 – 2.07% Stator fault severity.

Also, the training, validation and test performance (using the cross-entropy error function) is visualized in Fig. 10 on the basis of different numbers of neurons in the hidden layer.

In terms of low computational complexity and number of hidden layers, the best neural network has 9 neurons in its hidden layer. This number has been determined after examining the accuracy graph (Fig. 11) vs. number of neurons.

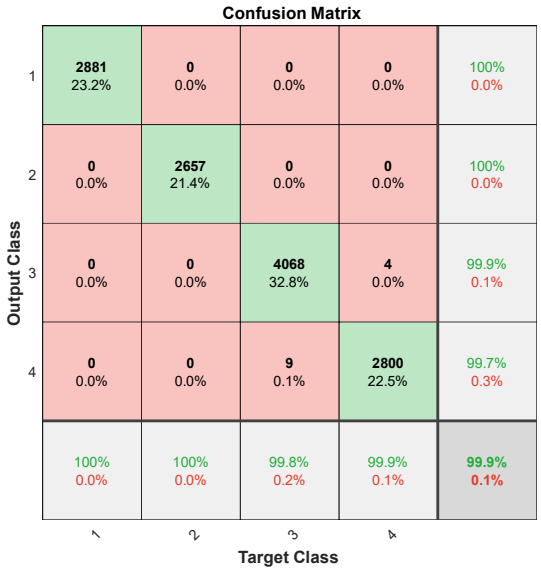


Fig. 9. Confusion Matrix

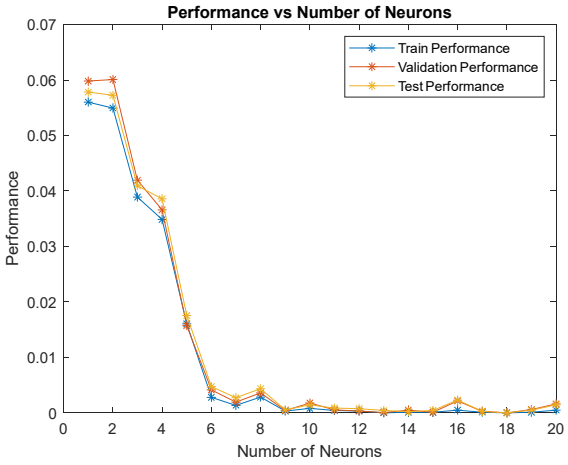


Fig. 10. Performance of the shallow dense neural network

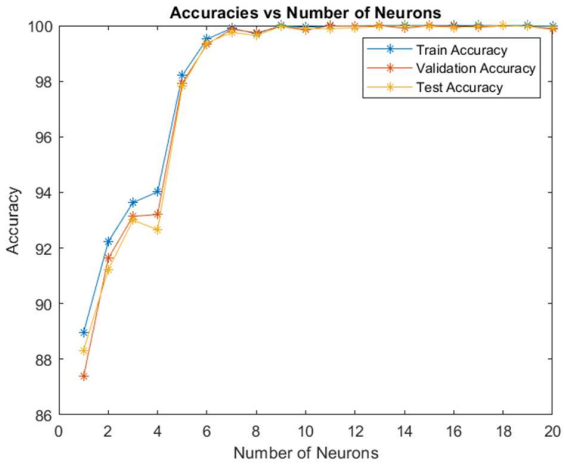


Fig. 11. Accuracy vs. number of neurons

V. CONCLUSIONS

This paper explores the possibility to detect stator inter-turn faults in SynRMs at low severity levels. From the developed neural and non-neural based approaches, the shallow dense neural network with few neurons in the hidden layer has shown the best performance not only for its lower computational complexity, but also for the ability to output probabilities together with the class memberships. The most essential step prior to classification is the generation of temporal features and their geometrical and topological analyses. In particular, the DR techniques: PCA and CCA are firstly used to determine whether the data manifold is linear or non-linear. The PCA is applied first to have generic idea on the intrinsic dimensionality value, and then it is confirmed by the CCA through its *dy-dx* map. The CCA can be also used to check other DR techniques and their input-output mappings. In this study, the use of CCA has confirmed the linearity of the data together with its intrinsic dimensionality value that was deduced initially by PCA. Thus, on this basis, the authors use PCA for DR of the FS and using this reduced-FS train neural and non-neural based classifiers. After a thorough comparison among the classifiers listed in Table II, the final classification accuracy (of the test set) using a shallow neural network was 99.99% with $O(n)$ as the computational complexity.

Future work will focus on the application of other neural based topologies for the detection of other faults in SynRM along with the corresponding sensitivity analysis.

TABLE II. CLASSIFICATION RESULTS ON THE TEST SET

Classifier	Classification Accuracy (%) Test Set	Comments
Shallow Dense NN	**99.99**	**Activation: TANH activation, 0% dropout rate, Architecture: $4_{IN}\|9_{FC}\|4_{OUT}$, No. Parameters = 85, Preprocessor: PCA, Ouput Activation: Softmax**
Fine Tree	94.46	Max. Number of Splits = 100, Split Criterion: Gini's Diversity index
Medium Tree	66.69	Max. Number of Splits = 20, Split Criterion: Gini's Diversity index
Course Tree	75.26	Max. Number of Splits = 4, Split Criterion: Gini's Diversity index
Linear Discriminant	82.16	Full Covariance Structure
Quadratic Discriminant	87.28	Full Covariance Structure
Linear SVM	88.35	Kernel Function: Linear
Quadratic SVM	83.54	Kernel Function: Quadratic
Cubic SVM	89.68	Kernel Function: Cubic
Gaussian SVM	90.22	Kernel Function: Gaussian, Kernel Scale: 13.0
Ensemble (Boosted Trees)	82.91	Ensemble Method: AdaBoost, Learner: Decision Tree, Max. Splits: 20, Number of Learners: 30, LR: 0.1
Ensemble (Bagged Trees)	92.44	Ensemble Method: Bagging, Learner: Decision Tree, Number of Learners: 30
Ensemble (Subspace Discriminant)	77.74	Ensemble Method: Subspace, Learner: Discriminant, Number of Learners: 30, Subspace Dimension: 5
Ensemble (Subspace kNN)	97.52	Ensemble Method: Subspace, Learner: Nearest Neighbours, Number of Learners: 30, Subspace Dimension: 5
Ensemble (RUSBoost)	79.24	Ensemble Method: RUSBoost, Max. Splits: 20, Number of Learners: 30, LR: 0.1

REFERENCES

[1] L. Xu, X. Xu, T. A. Lipo, and D. W. Novotny, "Vector control of a synchronous reluctance motor including saturation and iron loss," IEEE transactions on industry applications, vol. 27, no. 5, pp. 977-985, 1991.

[2] R. E. Betz, R. Lagerquist, M. Jovanovic, T. J. Miller, and R. H. Middleton, "Control of synchronous reluctance machines," IEEE Transactions on Industry Applications, vol. 29, no. 6, pp. 1110-1122, 1993.

[3] P. Vas. "Sensorless vector and direct torque control." Oxford University Press. (accessed.

[4] A. Accetta, M. Cirrincione, M. Pucci, and A. Sferlazza, "A nonlinear control of synchronous reluctance motors (SynRM) based on feedback linearization considering the self and cross-saturation effects," in 2019 IEEE Energy Conversion Congress and Exposition (ECCE), 2019: IEEE, pp. 1804-1809.

[5] Y. Zhongming and W. Bin, "A review on induction motor online fault diagnosis," in Proceedings IPEMC 2000. Third International Power Electronics and Motion Control Conference (IEEE Cat. No. 00EX435), 2000, vol. 3: IEEE, pp. 1353-1358.

[6] G. Singh, "Induction machine drive condition monitoring and diagnostic research—a survey," Electric Power Systems Research, vol. 64, no. 2, pp. 145-158, 2003.

[7] I. Bouchareb, A. Bentounsi, and A. Lebaroud, "Performance evaluation and remedial strategy of synchronous reluctance motor drive under winding turn short circuit fault," in 2016 XXII International Conference on Electrical Machines (ICEM), 2016: IEEE, pp. 2313-2320.

[8] R. R. Kumar et al., "Induction Machine Stator Fault Tracking Using the Growing Curvilinear Component Analysis," IEEE Access, vol. 9, pp. 2201-2212, 2020.

[9] R. R. Kumar, G. Cirrincione, M. Cirrincione, A. Tortella, and M. Andriollo, "A Topological Neural Based Scheme for Classification of Faults in Induction Machines," IEEE Transactions on Industry Applications, pp. 1-1, 2020, doi: 10.1109/TIA.2020.3032944.

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

[11] 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, no. 1, pp. 148-154, 1997.

[12] J. W. Sammon, "A nonlinear mapping for data structure analysis," IEEE Transactions on computers, vol. 100, no. 5, pp. 401-409, 1969.

[13] G. Cirrincione, V. Randazzo, and E. Pasero, "The Growing Curvilinear Component Analysis (GCCA) neural network," Neural Networks, vol. 103, pp. 108-117, 2018.

[14] C. M. Bishop, Pattern recognition and machine learning. springer, 2006.