2018 26th Mediterranean Conference on Control and Automation (MED)
Zadar, Croatia, June 19-22, 2018

Motor Fault Detection and Diagnosis Using Fuzzy Cognitive Networks with Functional Weights

Georgios Karatzinis, Yiannis S. Boutalis, *Member, IEEE* and Yannis L. Karnavas, *Member, IEEE*

Abstract— **A Fuzzy Cognitive Network (FCN) is an operational extension of a Fuzzy Cognitive Map (FCM) which incorporates proven stability and guaranteed exponentially-fast error convergence to zero during its training and supports at the same time the continuous interaction with the system it describes. FCN is a suitable, reliable and very powerful FCM structure for system identification, control applications and adaptive decision making because of its convergence based operational architecture. In this paper we introduce the FCN based classifier and apply it on a motor fault detection problem. We use also an alternative form of the network that assumes functional interconnection weights which improves its performance. A benchmark data set is used for condition monitoring of rolling bearings, consisting of motor current and vibration signals. Statistical features are extracted from both current and vibration signals and are combined together to enhance the diagnosis performance of the model. The classification accuracy of the proposed approach shows that the FCN based classifier can be used as a very reliable diagnostic tool for motor fault detection and diagnosis.**

I. INTRODUCTION

Fuzzy Cognitive Maps (FCMs) are referred as human knowledge based inference networks, cyclic directed weighted graphs, cognition influence graphs and recurrent neural networks. These definitions-descriptions complete the graphical, computational and memory state representation of FCMs. They have been used in many scientific fields for various tasks such as modeling [1], control [2], pattern recognition applications [3],[4],[5],[6], decision making [7], forecasting [8]-[9] pointing out their capability, flexibility and effectiveness. FCMs have been introduced by Kosko [10], based on Axelrod's work on cognitive maps [11], in order to model complex behavioral systems using causal relations.

The graphical representation of the FCM, that is the cognitive graph, consists of nodes and weight interconnections. The nodes represent behavioral aspects of the system, that is, each one represents a system characteristic feature. Weight interconnections represent node interactions, that is, causal relationships between actions, goals, events, values and trends of the system. The concepts of the network represent the activation values that are simultaneously updated showing the iteratively interaction and the overall impact over the system. The activation level is implemented by a nonlinear function f which can be of various types such as bivalent, trivalent, logistic, and sigmoid [12]. The inference mechanism of an FCM [13] is as follows: first, the FCM is initialized. The activation level of each of the nodes of the system is set to specific values based on the belief of the expert regarding the current state. Then, the various concepts are free to interact. The activation of one node influences the nodes to which it is connected. This interaction continues until: 1) a fixed-point equilibrium is reached; 2) a limit cycle is reached; or 3) chaotic behavior is exhibited [14].

Kosko enhanced the power of cognitive maps considering fuzzy values for their nodes and fuzzy degrees of interrelationships between nodes. He also proposed the differential Hebian rule [15] to estimate the FCM weights expressing the fuzzy interrelationships between nodes based on acquired data. Other remarkable learning algorithms and some learning and architecture limitations of conventional FCMs are illustrated in [16]. Several FCM extensions have been proposed in [6],[17],[18],[19], which try to overcome various inherent limitations and architecture restrictions. Such an FCM extension that improves convergence issues, knowledge management and the cooperation with the system under investigation is the Fuzzy Cognitive Network.

Fuzzy Cognitive Networks (FCNs) [20] were initially introduced as a general computational and storage framework to facilitate the use of FCM in a strong interaction with the physical system they describe. The updating mechanism receives feedback from the real system, stores the acquired knowledge and imposes control values to it. FCNs and their storage mechanism assume that they reach equilibrium points, each one being associated with a specific operational instant of the underlying system. The concept of "steady" nodes which correspond to input values and influence but are not influenced by the other nodes of the FCM was also introduced in that work.

The complete general FCN framework [21] is given with the incorporation of certain weight interconnection conditions [22], the rigorous mathematical proofs for its stability and convergence properties and the adaptive parameter estimation algorithm [23], the implementation of linear and bilinear parametric models [24] and the automated fuzzy rule database [22,23]. Hence, the aforementioned conditions and proofs complete the FCN framework and the proper operation of the FCM-based models, allowing them to converge during their training and operate in cooperation with the system they describe. This way they become suitable for many applications. Also, they limit experts' contribution only to the initial graph by just defining the number of nodes

Georgios Karatzinis is with the Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece, 67100 GR (e-mail: gkaratzi@ee.duth.gr).

Yiannis S. Boutalis is with the Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece, 67100 GR (e-mail: ybout@ee.duth.gr).

Yannis L. Karnavas is with the Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece, 67100 GR (e-mail: karnavas@ee.duth.gr).

and the causal relationships irrespectively of their initial values. The associations between equilibrium points and network weights and concept values are stored in a fuzzy rule database to be recalled later on by using a fuzzy inference system.

In this paper, we introduce the use of the complete FCN framework in pattern classification applications and study its performance in a motor fault detection and diagnosis problem. The contribution of steady nodes is of great importance, because they allow the network to assign a set of input nodes to probably different output classes. That is achieved by using steady nodes for the input patterns and non-steady nodes for the output classes. The classifier undergoes supervised training and each specific class is assigned to a single output. So the number of the classifier outputs equals the number of each specific problem's classes.

Another important feature of the proposed approach is the use of an alternative form of storing the FCN acquired data knowledge. Instead of using large fuzzy rule data bases for storing that knowledge and fuzzy inference procedures for using it we introduce functional weights on the node interconnections. The functional forms of the weights are continuously updated during the training phase of the FCN framework. This way we alleviate the framework from large storage requirements and from the need of experts' judgment and knowledge for the issues related to the fuzzy database and fuzzy inference mechanism.

Similar to the use of a fuzzy rule database and fuzzy inference, the alternative storage approach allows, at each operational instant, the proper weight retrieval even in situations, where the equilibrium conditions had not been exactly met during the training phase. This is achieved by handling pure data from the FCN producing functional deterministic relationships between concepts and weight interconnection values using regression analysis. In this case, the weights represent the dependent variables and node values the independent ones, while the regression is achieved by multiple polynomials with proper order, based on the coefficient of determination (R squared) and Adjusted R squared.

The material of the paper is unfolded by giving first, in section II, the basics of FCN, its convergence and its weight updating. Next, in section III, the FCN with functional weights is introduced, presenting the way by which the proper functions are estimated. Section IV enhances the use of steady nodes in FCNs allowing their direct application in pattern classification. In section V, a brief description on motor bearing fault detection theory is presented, while in section VI simulation results are presented, which use benchmark data from real bearing damages and illustrate the excellent performance of the proposed FCN structure with functional weights in detecting and diagnosing motor faults. Conclusions and thoughts for future works are given in section VII.

II. Fuzzy Cognitive Networks

FCNs are operational extensions of FCMs to support the close interaction with the system they describe. They improve conventional FCMs by introducing: (a) the representation level of the cognitive graph incorporating the partial human knowledge in the model (human in the loop) (b) the convergence conditions of the model (c) the updating mechanism based on linear and bilinear parametric models improving its learning capabilities (d) the automated storage of the acquired knowledge throughout the operation.

A. Fuzzy Cognitive Networks Representation

A graphical representation of FCN is depicted in Fig. 1. Each concept represents a characteristic of the system, in general, it represents events, actions, goals, values, and trends of the system. Each concept is characterized by a number A_i that represents its value, and it results from the transformation of the real value of the systems variable, which is represented by this concept, either in the interval [0,1] or in the interval [-1,1]. All concept values from vector A are expressed as:

$$A = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix}^T$$

with $n=5$ being the number of the nodes. Causality between concepts allows degrees of causality and not the usual binary logic, so the weights of the interconnections may range in the interval [-1,1]. The existing knowledge on the behavior of the system is stored in the structure of nodes and interconnections of the map. The value of w_{ij} indicates how strongly concept C_j influences concept C_i. The sign of w_{ij} indicates whether the relationship between concepts C_j and C_i is direct or inverse. The nodes that influence but are not influenced by other nodes are called steady. For the FCN of Fig. 1 the weight interconnection matrix W is equal to

$$W = \begin{bmatrix} d_{11} & 0 & 0 & 0 & 0 \\ 0 & d_{22} & 0 & 0 & 0 \\ w_{31} & w_{32} & d_{33} & w_{34} & 0 \\ w_{41} & w_{42} & w_{43} & d_{44} & 0 \\ 0 & w_{52} & 0 & w_{54} & d_{55} \end{bmatrix}$$

The equation that calculates the values of concepts of FCNs is given by:

$$A_i\left(k\right) = f\left(d_{ii} A_i\left(k-1\right) + \sum_{\substack{j=1 \\ j \neq i}}^{n} w_{ij} A_j\left(k-1\right) \right) \quad (1)$$

where $A_i(k)$ is the value of concept C_i at discrete time k, $A_i(k-1)$ the value of concept C_i at discrete time $k-1$ and $A_j(k-1)$ is the value of concept C_j at discrete time $k-1$. w_{ij} is the weight of the interconnection from concept C_j to concept C_i and d_{ii} is a variable that takes on values in the interval [0,1] depending upon the existence of "strong" or "weak" self-feedback to node i. The concepts C_1 and C_2 are steady nodes and their diagonal elements are equal to zero.

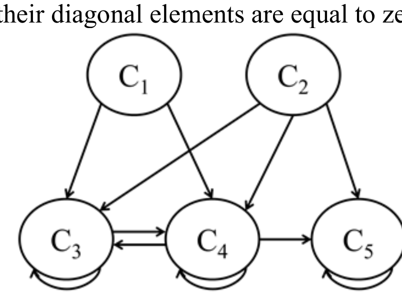

Figure 1. FCN with 5 nodes.

Most common function f is the squashing sigmoid function with saturation level 0 and 1. Sigmoid function f is a continuous and differentiable node function used both in FCMs and FCNs and squashes the result in the interval [0,1] and is expressed as

$$f = \frac{1}{1+e^{-c_l x}}$$

Where $c_l>0$ is used to adjust its inclination.

B. Adaptive Learning Algorithms of Fuzzy Cognitive Networks

In FCNs, proofs and conditions for the existence and uniqueness of solutions for (1) have been studied analytically in [22], [23]. The conditions are extended to take into account not only the weights of the map but also the inclination parameters of the involved sigmoid functions, increasing the structural flexibility of the network. Following the inverse procedure of (1) it would be helpful to find appropriate weight sets directly related to desired equilibrium points of the FCN. That leaded to the development of adaptive weight and sigmoid parameter estimation algorithms, which employs appropriate weight projection criteria to assure that the equilibrium is always achieved. The adaptive estimation algorithms are distinguished in linear and bilinear parametric model designs. Both approaches show great convergence and learning capabilities using appropriate projection methods which guarantee that the conditions of existence and uniqueness are always fulfilled. The linear approach shows inability to give always the appropriate parameter estimates. The bilinear approach takes into account both weight and inclination parameters of the nodes' sigmoid functions outperforming the linear one, allowing the use of smaller FCN structures [24].

C. Knowledge Storage and Recall using Fuzzy Rule Databases

The knowledge storage mechanism [23],[24] is applied on FCNs in order to store the acquired knowledge for weight interconnections and allow their retrieval for future use even in situations, where the equilibrium conditions have not been exactly met during the training phase. This is done by using a fuzzy if-then rule database, which associates in a fuzzy manner the various weights with the corresponding equilibrium node values. After storage has been completed, it can be used in place of the unknown system to predict its behavior as the FCN recalls the relative weight values.

III. Fuzzy Cognitive Networks with Functional Weights Representation

The obtained weight interconnections related to desired equilibrium points of the FCN and node values constitute the knowledge of the encountered operational situations. Instead of storing that knowledge in a fuzzy rule database, we assume an alternative scheme, where the weights are not plain values but non linear functions having a polynomial form

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 \qquad (2)$$

Which is non-linear in x, but it is linear in respect to the parameters a. In order to model the values of the weight interconnections, that is the dependent variables, in terms of the values of the independent (node) variables, we are using multivariable forms of (2) as follows:

$$\begin{gathered} W_1 = y_1 = f(x_1,\ldots,x_n) \\ W_2 = y_2 = f(x_1,\ldots,x_n) \\ \vdots \\ W_k = y_k = f(x_1,\ldots,x_n) \end{gathered}$$

where x_i, (i=1,2,…,n) are the node variables that influence the dependent variables (weights W_i ,i=1,2,…,k) and are used to fit weight interconnection data into a least squares linear regression model. It is conducive to use a polynomial generator to construct different n^{th} order polynomials. The repeatable increase of the order of the polynomials will change the best fit shape, that is the functional weights shape. The statistic measure "R squared" r^2 will lead to proper order that best fits the obtained data. This way the interconnection weights are constructed as high order functions. To estimate values of parameters linear least squares is needed. Linear least squares problems occur when solving overdetermined linear systems.

In our application the total of the pairs of node and weight interconnection values naturally give more equations than unknowns leading to overdetermined system. The solution of the linear least squares problem is a vector $x\in\mathbf{R}^n$ for which the norm of the residual $\boldsymbol{r}$ is minimized, that is

$$\|\boldsymbol{r}\|_2 = \|\boldsymbol{b} - M\boldsymbol{x}\|_2 \to \min$$

where $\boldsymbol{b}\in\mathbf{R}^m$ and in the full matrix $M\in\mathbf{R}^{m\times n}$ is overdetermined, that is $m>n$

$$M = \begin{bmatrix} A_1^1 & \cdots & A_n^1 & W_1^1 & \cdots & W_k^1 \\ & \vdots & & & \vdots & \\ A_1^m & \cdots & A_n^m & W_1^m & \cdots & W_k^m \end{bmatrix}$$

Where m is the number of equilibriums inserted as observations for least squares. Obtaining the normal equations

$$M^T M\boldsymbol{x} = M^T\boldsymbol{b}$$

the least squares solution is uniquely determined by

$$x_{LS} = M^+\boldsymbol{b} = \left(M^T M\right)^{-1} M^T\boldsymbol{b}$$

where M^+ is the pseudo-inverse of M. Instead of normal equations we use QR decomposition because it is numerically more stable [25].

Considering a matrix $M\in\mathbf{R}^{m\times n}$ with $m\geq n$ and rank equal to n there exists the Cholesky decomposition of

$$M^T M = R^T R$$

where R is an upper triangular matrix. Since R is non singular $(MR^{-1})^T(MR^{-1})=I$ and matrix $Q=MR^{-1}$ has orthogonal columns. The QR decomposition is

$$M = QR = Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

where R_1 is an $n\times n$ upper triangular matrix. Then the pseudo-inverse of $M\in\mathbf{R}^{m\times n}$, $m>n$ and rank equal to the minimum of m and n can be found by

$$M^+ = R^{-1}Q^T$$

Then the unique solution is given by

$$x_{LS} = R^{-1}Q^{T}\boldsymbol{b}$$

The method that is used for computing the QR decomposition is in our case the Householder transformations [26]. Hence, the weight interconnections assume the functional structure of (2) which is non-linear if $n\geq 2$ and $a_n\neq 0$.

IV. Fuzzy Cognitive Networks For Pattern Classification Applications

The effectiveness of the FCN framework in pattern recognition application is reflected in the following Lemma.

Lemma 1: *An FCN structure consisting of input and output concepts can lead to an efficient classifier if the input concepts assume steady nodes form and the output concepts non-steady ones*.

This Lemma arises as a direct consequence of the interpretation of FCN's convergence properties [23]. According to [23] in FCN with steady nodes the convergence fixed point (otherwise termed equilibrium point) does not depend solely on the weight and sigmoid inclination parameters, as in the case of FCN with non-steady nodes. It depends also on the values of the steady (input) nodes. This way, different external excitations on the input nodes will drive to different fixed-points. Moreover, the use of functional weights, instead of plain weight values, permits the mapping of different excitation input values to the same outputs classes, because a functional is actually the representative of many different plain values, each one representing a different input-output association. Needless to say that the conventional FCN framework with the fuzzy rule database has the same properties, but as pointed out in the previous sections, the functional weights approach replaces the need for storing the fuzzy rule database and using the required fuzzy inference procedure to recall the associations.

To demonstrate the above facts we compare the classification abilities of both conventional FCN framework and FCN framework with functional weights using the "Fertility" benchmark dataset[1] with 10 attributes, 100 instances and 2 classes. The experiments were performed using 5 fold cross-validation and were repeated 10 times.

The classification rates in Table I demonstrate the enhanced results using functional weights in the FCN framework. Both linear and bilinear FCN parametric models were tested, while the functional weights assume second order multivariate polynomial form.

TABLE I. Fertility-FCN Classifier Classification Rate

	Fuzzy Rule Database		***Functional Weights***	
Dataset	***Linear***	***Bilinear***	***Linear***	***Bilinear***
Fertility	0.8800	0.8850	0.9250	0.9390

[1] Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

V. Motor Bearing Fault Detection

The most common mechanical faults in industrial processes are related to the bearing damages [27], [28]. These are detected essentially by vibration analysis with different detection methods e.g. [29]. These monitoring techniques are expensive and require a challenging mechanical construction depending on the number of used accelerometers and the kind of bearing. For this reason, alternative detection approaches are studied e.g. [30], [31]. There, the stator current signal of a permanent magnet synchronous machine (PMSM) (as well as induction machine) is used to detect bearing damage in the drive train. The objective of this method is to utilize the embedded current sensors of the controller for faults detection as well. A bearing damage causes characteristic frequencies in the structural vibration spectrum, which depend on the mechanical rotational speed and the kind of bearing fault. Bearing damages lead to torque pulsations resulting in speed fluctuations, which induce oscillations in the motor current with the characteristic frequencies. These fault frequencies depend on the electric supply frequency, the vibration frequencies and load torque fluctuations. Afterwards, this procedure is applied to the stator current signal of a PMSM at different loadings. Vibrations are transferred also into the air gap of the machine related to the torque feedback of the bearing damage. Rolling bearings are usually used to guide and to support the shaft in rotating machines. They consist of an inner and outer race (ring), which are separated by rolling elements such as balls or cylindrical rollers. The rollers are integrated in a cage to avoid contact. Rolling bearings can be damaged by different causes so that local or distributed defects occur. Distributed defects due to generalized roughness affect the whole region of the bearing and can be recognized through particular frequencies related to local defects. This can be classified in function of the affected part of the bearing such as: inner ring fault, outer ring fault, cage fault and ball fault. These faults cause characteristic frequencies in the vibration spectrum, which depend on the geometry of the rolling elements and the mechanical rotational frequency f_{rm}. The characteristic frequencies of each fault type are then described in as follows:

Inner/Outer ring defect $$\begin{matrix} f_{inner} \\ f_{outer} \end{matrix} = \frac{N_{ball}}{2} f_{rm}\left(1 \pm \frac{D_{ball}}{D_{cage}}\cos\beta\right)$$

Ball defect $$f_{ball} = \frac{D_{cage}}{2D_{ball}} f_{rm}\left(1 - \frac{D_{ball}^2}{D_{cage}^2}\cos\beta^2\right)$$

Cage defect $$f_{cage} = \frac{1}{2} f_{rm}\left(1 - \frac{D_{ball}}{D_{cage}}\cos\beta\right)$$

whereas N_{ball} is the number of balls or cylindrical rollers, β the contact angle of the balls, D_{ball} the ball or roller diameter and D_{cage} the cage diameter, also known as the ball or roller pitch diameter.

VI. FAULT DETECTION RESULTS USING FUZZY COGNITIVE NETWORKS

The procedures presented in sections III and IV are implemented together in a motor fault detection and diagnosis problem as it is illustrated in Fig.2. For training and testing purposes, a benchmark data set of real bearing damages[2] is used which distinguish three classes: 1) healthy state, 2) outer ring damage and 3) inner ring damage. From this data set we are using three arrays of measurements, the first two are related to current signals and the third with the vibration signal. Statistical features are extracted from the raw data of each array in the preprocessing stage, including mean, variance, skewness and kurtosis, which are in the sequel normalized in the interval [0,1]. In the training phase each feature gives values to an input-steady node, while each class is associated with a non-steady node. Since we expect a 3-class classification, the number of non-steady output nodes is three. The supervised learning procedure produces proper weight matrices based on the given target label for each feature vector. The different plain weight values are "stored" in the Functional form of each weight. These trained functional weights are then provided to the diagnosis FCN (right part of Fig.2). For each test feature vector the Diagnosis FCN computes the equilibrium-output state corresponding to predicted class. Finally, the exact bearing condition is estimated using Euclidean distance between the target classes and the equilibrium-output states. In Table II the classification accuracy is given for three distinct case studies for the under investigation motor bearings.

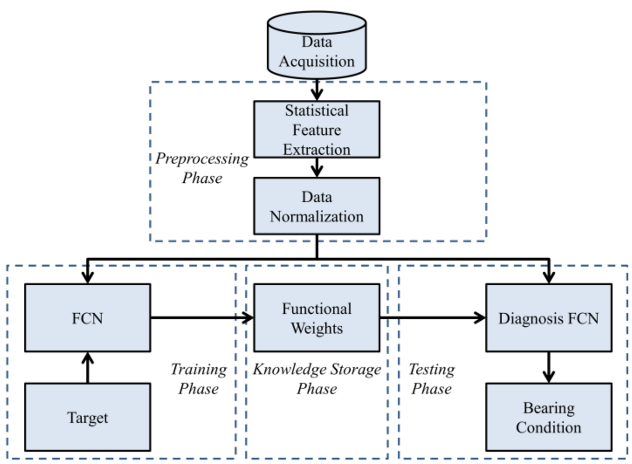

Figure 2. Motor fault detection and diagnosis scheme using FCN with functional weights

TABLE II. MOTOR FAULT DETECTION PERFORMANCE - FCN

	Motor Current Signals	Vibration Signal	Combination of all Signals
Linear	90.08	95.50	99.58
Bilinear	**91.42**	**95.83**	**99.83**
Number of Features	8 (total)	4	12

[2] Christian Lessmeier et al., KAt-DataCenter, Chair of Design and Drive Technology, Paderborn University, Germany: http://mb.uni-paderborn.de/kat/datacenter

TABLE III. MOTOR FAULT DETECTION PERFORMANCE – COMPARISON WITH WELL-KNOWN ML ALGORITHMS USING STATISTICAL FEATURES

	FCN-FW Bi-linear	Bagged Tree	AdaBoost	Fine Gaussian SVM	Weighted K Nearest Neighbor
1	91.42	**95.30**	87.90	94.80	87.20
2	95.83	**97.80**	**97.80**	94.80	96.40
3	**99.83**	99.80	98.50	98.90	97.80

Both linear and bilinear parametric models are used for the FCN training phase. It can be observed that the bilinear one improves slightly the correct classification rate. The first case study concerns motor current signals using 4 features for each of the two signal (8 features). The second one uses the 4 aforementioned statistical features computed solely on vibration signal demonstrating improved classification rate compared with the current signals case. The best performance of the final motor rolling bearing fault diagnosis model is summarized as a combination of the extracted statistical features from both raw motor current and vibration signals performing diagnosis accuracy of 99.83% for the bilinear parametric FCN model. Various comparisons were made regarding the order of the polynomial functional weights, which affects the generalization abilities of the classifier. The above best results were achieved using 3rd order polynomials and compared with well-known machine learning (ML) algorithms trained using the same type (statistical) and number of features. In Table III the classification rates for the three case studies: 1) motor current signals; 2) vibration signals; 3) their combination; are given demonstrating the classification abilities of the proposed model compared with powerful classifiers of the literature. Bagged Tree model performs the best performance of 95.30% on the first case study. The results have indicated that the ensemble ML algorithms Bagged Tree and AdaBoost outperform all models on the second case with 97.80% classification rate, while the bilinear FCN with functional weights and the Bagged Tree methods show great performance on the final combination of the extracted statistical features case. The achieved classification rates of the proposed model from the bilinear approach are compared with well-known classifiers using the same real damage dataset but with different type and number of features and are depicted in Table IV. In [32] frequency domain features were extracted using Fast Fourier Transform (FFT) and power spectral density (PSD) and after feature extraction and feature selection, a number of 18 features were emerged for both motor current signals and 15 features for the vibration signal. It can be seen that, although we use only 8 and 4 features respectively, the performance of the FCN classifier is quite comparable, while it is recalled from Table II that when combining all 12 features it outperforms all methods. The experiments were performed using 5 fold cross-validation and repeated 10 times and the experimental evaluation of Table III models is carried on in a FX-8150 3.6 GHz PC with 4GB memory.

VII. CONCLUSION

In this paper an FCN structure was proposed, which incorporates steady nodes for input concepts and non-steady

TABLE IV. MOTOR FAULT DETECTION PERFORMANCE - COMPARISON WITH OTHER MODELS TAKEN FROM [32]

Machine Learning Model	CART	RF	BT	NN	SVM-PSO	ELM	kNN	Ensemble	FCN with Functional Weights
Motor Current Signals	66.7	83.3	81.7	65.8	56.7	69.2	68.3	**93.3**	91.42
Vibration Signals	**98.3**	**98.3**	83.3	44.2	75.8	60.8	62.5	**98.3**	95.83

for the output ones and lead to a powerful FCN based classification tool. The new approach adopts functional weights as an alternative form of storing the FCN acquired data knowledge, avoiding the need for using fine-tuned fuzzy rule database and fuzzy inference procedures. Functional weights overcome limitations of conventional FCMs providing non-linear weight functions, permitting the storage of more than one relationship values. The performance of the FCN based classifier is tested on a real motor bearing fault detection problem. From the results obtained based on both linear and bilinear parametric training models, it is evident that the proposed model produce accurate predictions of motor faults. Future extensions of this work may include the use of multiple hierarchically ordered FCN to improve the classification performance.

REFERENCES

[1] D.M. Case, C.D. Stylios, "Fuzzy Cognitive Map to model project management problems," in *Proceedings of Annual Conference of the North American Fuzzy Information Processing Society* (NAFIPS), *IEEE*, Texas, USA, 2016, pp. 1-6.

[2] T. L. Kottas, Y. S. Boutalis, and A. D. Karlis, "New maximum power point tracker for PV arrays using fuzzy controller in close cooperation with fuzzy cognitive networks," *IEEE Transactions on Energy Conversion*, Vol. 21, pp. 793–803, 2006.

[3] W. Froelich, "Towards improving the efficiency of the fuzzy cognitive map classifier," *Neurocomputing,* Vol. 232, pp. 83 – 93, 2017.

[4] G.A. Papakostas , D.E. Koulouriotis, A.S. Polydoros, V.D. Tourassis, "Towards Hebbian learning of Fuzzy Cognitive Maps in pattern classification problems," *Expert Systems with Applications*, Vol. 39, pp. 10620-10629, 2012.

[5] G. Nápoles, I. Grau, R. Bello, R. Grau, "Two-steps learning of fuzzy cognitive maps for prediction and knowledge discovery on the hiv-1 drug resistance," *Expert Syst. with Appl.*, Vol. 41, pp. 821-830, 2014.

[6] H. J. Song, C. Y. Miao, R. Wuyts, Z. Q. Shen, M. D'Hondt, F. Catthoor, "An extension to fuzzy cognitive maps for classification and prediction," *IEEE Trans. on Fuzzy Syst.*, Vol.19, pp. 116–135, 2011.

[7] E. I. Papageorgiou, "A new methodology for Decisions in Medical Informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques," *Applied Soft Computing*, Vol. 11, No. 1, pp. 500–513, 2011.

[8] W Froelich, W Pedrycz, "Fuzzy cognitive maps in the modeling of granular time series," *Knowl.-Bas. Syst.*, Vol. 115, pp. 110-122, 2017.

[9] J.L. Salmeron, W. Froelich, E.I. Papageorgiou, "Application of fuzzy cognitive maps to the forecasting of daily water demand," In *Proceedings of 2015 ITISE*, Granada, Spain, 2015.

[10] B. Kosko, "Fuzzy cognitive maps," *International Journal of Man-Machine Studies,* Vol. 24, pp. 65–75, 1986.

[11] R. Axelrod, *Structure of Decision: The Cognitive Maps of Political Elites*, Princeton University Press, New Jersey, 1976.

[12] A.K. Tsadiras, "Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps," *Information Sciences*, Vol. 178, pp. 3880-3894, 2008.

[13] D. Yaman, S. Polat, "A fuzzy cognitive map approach for effect-based operations: An illustrative case," *Information Sciences*, Vol. 179, pp. 382-403, 2009.

[14] A.K. Tsadiras, K.G. Margaritis, "Cognitive mapping and certainty neuron fuzzy cognitive maps," *Information Sciences*, Vol. 101, pp. 109–130, 1997.

[15] B. Kosko, "Differential Hebbian Learning," in *Proceedings of AIP Conference*, pp. 277–282, 1986.

[16] E.I. Papageorgiou, J.L. Salmeron, "A review of fuzzy cognitive maps research during the last decade," *IEEE Trans. Fuzzy Syst*., Vol. 21, pp. 66-79, 2013.

[17] G. Nápoles, I. Grau, E. Papageorgiou, R. Bello, K. Vanoof, "Rough cognitive networks," *Knowl. Based Syst*. Vol 91 pp. 46–61, 2016.

[18] W. Pedrycz, A. Jastrzebska, W. Homenda, "Design of fuzzy cognitive maps for modeling time series," *IEEE Trans. Fuzzy Syst*., Vol. 24, pp. 120–130, 2016.

[19] G. Acampora, V. Loia, "On the temporal granularity in fuzzy cognitive maps," *IEEE Trans. Fuzzy Syst*., Vol. 19, pp. 1040–1057, 2011.

[20] T.L. Kottas, Y.S. Boutalis, M.A. Christodoulou, "Fuzzy cognitive networks: a general framework," *Intelligent Decision Technologies,* Vol. 1, pp. 183–196, 2007.

[21] Y. Boutalis, D. Theodoridis, T. Kottas, and M. A. Christodoulou, *System Identification and Adaptive Control: Theory and Applications of the Neurofuzzy and Fuzzy Cognitive Network Models (Advances in Industrial Control)*. Basel, Switzerland: Springer, 2014.

[22] Y. Boutalis, T. Kottas, M. Christodoulou, "On the existence and uniqueness of solutions for the concept values in fuzzy cognitive maps," in *Proc. of 47th IEEE Conf. Decis. and Control*, Cancun, Mexico, 2008, pp. 98–104.

[23] Y. Boutalis, T. Kottas, M. Christodoulou, "Adaptive estimation of fuzzy cognitive maps with proven stability and parameter convergence," *IEEE Transactions on Fuzzy Systems*, Vol. 17, pp. 874–889, 2009.

[24] T.L. Kottas , Y.S. Boutalis , M.A. Christodoulou , "Bi-linear adaptive estimation of fuzzy cognitive networks", *Appl. Soft Comput*., Vol. 12, pp. 3736–3756, 2012.

[25] G. Golub, "Numerical methods for solving linear least squares problems," *Numerische Mathematik*, Vol. 7, pp. 206-216, 1965.

[26] A. S. Householder, "Unitary Triangularization of a Nonsymmetric Matrix," *Journal of the ACM*, Vol 5, pp. 339-342, 1958.

[27] P. Tavner, L. Ran, J. Penman and H. Sedding, "Condition monitoring of rotating electrical machines," *IET Power and Energy Series*, Vol. 56, 2008.

[28] Y. L. Karnavas, I. D. Chasiotis, A. Vrangas, "Fault diagnosis of squirrel-cage induction motor broken bars based on a model identification method with subtractive clustering," in *Proc. of IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED),* Aug. 29 - Sep. 1, 2017, Tinos, Greece, pp. 304-310.

[29] G. Del Vescovo, M. Paschero, A. Rizzi, R. Di Salvo, F.M.F. Mascioli, "Multi-fault diagnosis of rolling-element bearings in electric machines," *XIX Intl. Conference on Electrical Machines*, *(ICEM),* Sep. 6-8, 2010, Rome, Italy, pp. 1-6.

[30] M. Drakaki, Y. L. Karnavas, I. D. Chasiotis and P. Tzionas, "An intelligent Multi-Agent System framework for fault diagnosis of squirrel-cage induction motor broken bars," in *Proc.* of *38th International Conference on Information Systems Architecture and Technology (ISAT)*, Sep. 17-19, 2017, Szklarska Poręba, Poland, pp. 80-89.

[31] F. Immovilli, A. Bellini, R. Rubini, and C. Tassoni, "Diagnosis of bearing faults of induction machines by vibration or current signals: A critical comparison," *IEEE Transactions on Industry Applications*, Vol. 46, pp. 1350 - 1359, 2010.

[32] C. Lessmeier, J. K. Kimotho, D. Zimmer, W. Sextro, "Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification," In *Proc. of the European Conference of the Prognostics and Health Management Society*, 2016.