Classifying Patterns Using Fuzzy Cognitive Maps

G.A. Papakostas and D.E. Koulouriotis

Abstract. This chapter is focused on the use of Fuzzy Cognitive Maps (FCMs) in classifying patterns, as alternative to the traditional classifiers such as neural networks or even as collaborators, in achieving better classification capabilities. By defining the classification procedure as the equilibrium point achieved by applying common inference laws, a FCM can simulate a typical classifier that maps a set of inputs to specific output values.

The classification capabilities of the FCM classifiers are studied in several pattern classification problems, while the ability of the FCM to store knowledge about the problem in hand is investigated in conjunction to the nodes' type of activation function and the inference law used. Appropriate experiments are taken place, in order to analyze the behavior of the FCM-based classifiers, in well known benchmark problems.

1 Introduction

Recently, an increased interest about the theory and application of the Fuzzy Cognitive Maps in engineering science is noted. FCMs are characterized by their ability to model the dynamics of complex systems by incorporating the casual relationships between the main concepts that describe the system. They have been used initially to model complex social (Koulouriotis et al. 2003), strategic (Xirogiannis and Glykas 2007) and financial systems (Koulouriotis et al. 2001b, 2005, Koulouriotis 2004, Xirogiannis and Glykas 2004b, Glykas and Xirogiannis 2004, Chytas et al. 2006), where analytical descriptions do not exist or can not be derived.

Fuzzy Cognitive Maps are fuzzy signed directed graphs with feedback. They were proposed by Kosko (Kosko 1986) as a modeling methodology of complex systems, able to describe the casual relationships between the main factors (concepts) that determine the dynamic behavior of a system. The concepts are interconnected through arcs having weights that denote the cause and effect relationship that a concept has on the others.

G.A. Papakostas and D.E. Koulouriotis
Democritus University of Thrace, Department of Production and Management Engineering, Xanthi, Greece

M. Glykas (Ed.): Fuzzy Cognitive Maps, STUDFUZZ 247, pp. 291–306, 2010.
springerlink.com

At each time, the value of each i^{th} concept A_i is calculated, by summarizing the influences of all the other concepts to this and by squashing the overall impact using a barrier function *f*, according to the following inference rule

$$A_i^{t+1} = f\left(A_i^t + \sum_{i=1, i \neq j} W_{ji} A_j^t \right) \tag{1}$$

where A_i^{t+1} and A_i^t are the values of concept *i* at times *t+1* and *t* respectively, A_j^t the value of concept *j* at time *t*, W_{ji} the weight value of the interconnection with direction from concept *j* to concept *i*, and *f* the barrier function used to restrict the concept value into a specific range, usually in the range [0,1].

In each step, a new state of the concepts is derived according to (1) and after a number of iterations, the FCM may arrive in one of the three states (Kosko 1986) (i) fixed equilibrium point, (ii) limited cycle and (iii) chaotic behavior. When the FCM arrives in a fixed equilibrium point, we can conclude that the FCM has converged and the final state corresponds to the actual system state in which the system concludes, when the initial values of concepts are applied.

Besides the use of FCMs in modeling complex physical systems, their application as part of modern DSS (Decision Support Systems) in decision making, significantly increases through the years (Khan and Quaddus 2004, Xirogiannis et al. 2004a, 2008, Pajares et al. 2006). This motivated the authors (Papakostas et al. 2006, 2008) to investigate the behavior and performance of the FCMs in classifying patterns belonging to pattern recognition problems, under several configurations. The main principles as long as the latest results of this investigation, constitute the subject of the following sections.

2 FCM-Based Classifiers

A crucial part of any modern intelligent system, which learns from its environment and interacts with it, is a pattern recognition process. In general, a pattern recognition process employs four stages: (1) data acquisition (2) data preprocessing (denoising, filtering, etc) (3) feature extraction and finally (4) classification. Among these processing steps, the last one significantly affects the overall performance of the system, comprising the stage where knowledge about the problem in process is stored and takes the decision by classifying the patterns in proper classes. The main functionality of a typical classifier is the mapping of an input set to a specific output one, according to the internal representation of the knowledge stored inside the classifier's structure.

Representative type of commonly used classifiers is the Artificial Neural Networks (ANNs), in which the need of knowledge storing and knowledge based decision have been inspired by the neural networks of a human's brain (Haykin 1999). Due to their knowledge storage capability, neural networks are able to be used for pattern recognition tasks and classification problems, while their ability to repeatedly learn their internal representation makes them very useful to real-time image and signal processing applications. The knowledge in a neural network is stored during the training phase, when its weights take appropriate values, according to the problem being processed. However, since the dynamics of a neural network is still under investigation, one have to make some appropriate trial and error tests in order to decide the optimal architecture of the network to be used, for a specific application.

Due to the fact that the structure of the FCMs resembles this of ANNs, many concepts and procedures from the field of ANNs have been adopted and applied in the case of FCMs. However, while ANNs have been successfully used to classify patterns, the behavior of the FCMs in pattern recognition applications is still unexplored.

Since the major advantage of Fuzzy Cognitive Maps is their ability to describe the behaviour of complex systems, they could be used to model a classifier having appropriate input and output concepts and appropriate interconnections among them. Based on this prospect, the inference procedure of the FCM corresponds to the internal computations the classifier performs in order to map the input data to specific output, to an equilibrium point, according to the stored knowledge.

In order to describe the functionality of the FCM-based classifier in terms of FCMs the following definition has to be declared.

Definition 1. *The Fuzzy Cognitive Map structure used to map a set of input concepts (ICs) to a set of output concepts (OCs), through the execution of the inference formula (1), until an equilibrium point is reached, is called Fuzzy Cognitive Mapper (FCMper).*

The operation of the Fuzzy Cognitive Mapper as a modeling methodology of the dynamics of a classifier can be defined as follows,

Definition 2. *The values of the output concepts (OCs) of the state in which a Fuzzy Cognitive Mapper equilibrates (equilibrium point), when appropriate values of the input concepts (ICs) are applied, correspond to the decision of the classifier being modeled by the FCMper.*

According to the above definitions, a FCM can be used to model the behavior of a classifier, in order to classify patters belonging to specific classes, and its performance depends on several parameters. Some of these parameters are the structure of the FCMper, the nodes' type of activation function, the inference law used to equilibrate the classifier and the learning algorithm used to find the interconnection weights. These factors that affect the performance of a FCMper and determine its classification capabilities are discussed next.

2.1 FCMper – Structure

The first factor that plays important role on the performance of a FCMper, as in the case of ANNs is its structure, which defines the way the input nodes are connected with the output ones. A typical FCMper has the following form (Papakostas et al. 2006, 2008).

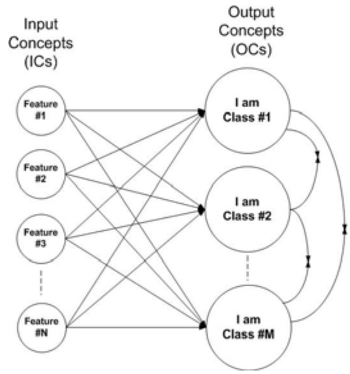

Fig. 1 A typical FCMper – FCMper1

In the above classifier (FCMper1), the input concepts correspond to the features that uniquely describe the patterns and formed by a previous processing step called Feature Extraction Method (FEM), while the output concepts are the classes' labels where the patterns belong. The FCMper1 of Fig.1 is the simplest form of a FCM-based classifier, which present limited performance as compared to the traditional neural classifiers (Papakostas et al. 2008).

More sophisticated hybrid FMCpers structures seem to perform better (Papakostas et al. 2008) in comparison to the previous FCMper1, where one neural classifier and a FCMper1 are operating in a cascade topology, increasing the classification accuracy of the hybrid model. These hybrid structures are illustrated in the following Fig.2 and Fig.3.

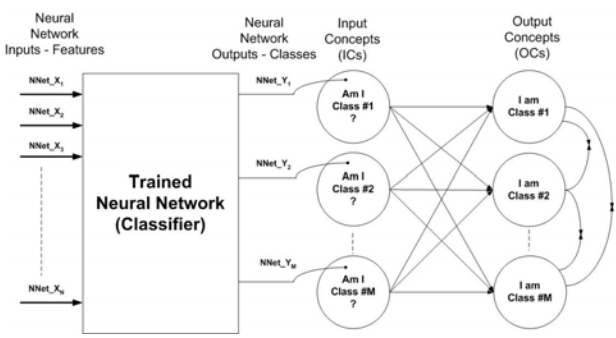

Fig. 2 Hybrid FCMper – FCMper2

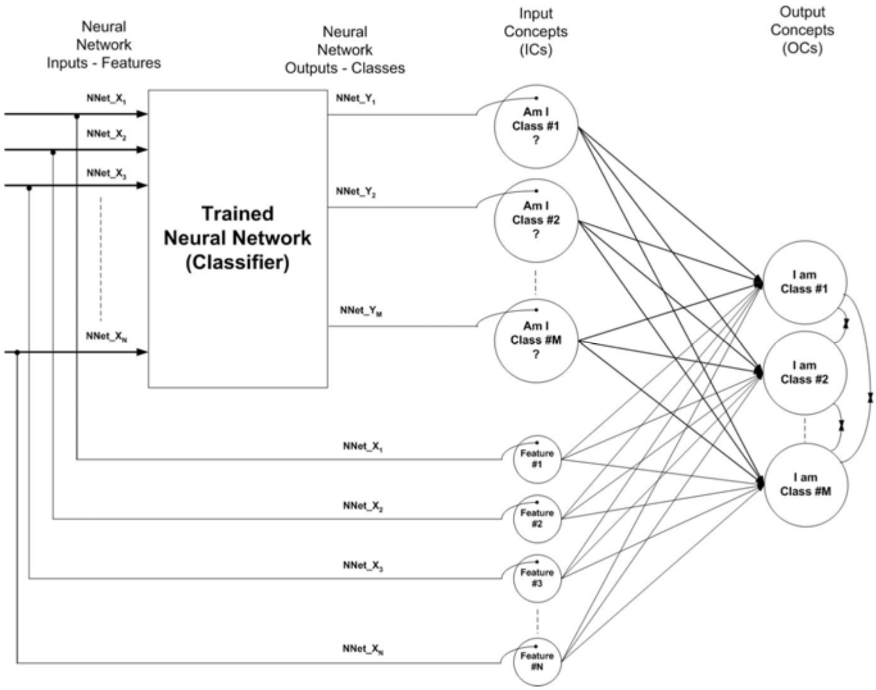

Fig. 3 Hybrid FCMper – FCMper3

The hybrid FCM-based classifiers FCMper2 and FCMper3, take as inputs the problem features (the same as the neural classifier inputs) and the features with the corresponding neural classifier responses respectively and answer to the question which is the appropriate class they belong.

While these hybrid models proved to be superior to the simple FCMper (Papakostas et al. 2008), their complexity make them inappropriate for real-time applications, where a decision about the class belonging of a pattern, has to be obtained in a certain time. This fact makes the usage of simple structures such as the FCMper1 more useful for these applications and therefore new topologies that improve its classification performance need to be developed.

2.2 FCMper – Activation Function

The function used to squash the overall impact of the nodes to a specific node, does not play only the role of a barrier function, but also gives a nonlinearity nature in the reasoning mechanism of a FCMper. This nonlinear behavior of a node is of significant importance in the case of difficult classification problems where the decision boundaries consisting by complex hyperplanes of multiple dimensions.

Some interesting conclusions on the affect of the activation functions on the operation of a FCM has been reported in (Koulouriotis et al. 2001c, 2001d, Tsadiras 2008, Bueno and Salmeron 2009, Boutalis et al. 2009), however this affection on the performance of a FCMper is unknown. For this reason three popular functions are selected as activation functions of the FCMper1 nodes, and their classification capabilities in several pattern recognition benchmark problems are studied in the experimental section. The analytical form of these functions which are the *Step, Trivalent* and *Sigmoid* ones are presented in the following equations.

Step Function (F1)
$$\mathrm{f(x)} = \begin{cases} 0, & \mathrm{x} \le 0 \\ 1, & \mathrm{x}>0 \end{cases} \tag{2}$$

Trivalent Function (F2)
$$\mathrm{f(x)} = \begin{cases} -1, \mathrm{x} \le -0{,}5 \\ 0, -0{,}5 < \mathrm{x} < 0{,}5 \\ 1, \mathrm{x} \ge 0{,}5 \end{cases} \tag{3}$$

Sigmoid Function (F3)
$$\mathrm{f(x)} = \frac{1}{1+\mathrm{e}^{-\mathrm{Ax}}} \tag{4}$$

The parameter *A* in the case of sigmoid function is a parameter that controls the slope of the curve and in the special case where this is equal to 1, the resulted sigmoid functions corresponds to the *logistic sigmoid* function.

2.3 FCMper – Inference Law

The operation of a FCM is based on the repetitive application of a specific inference law, which by taking into account the total impacts of the concepts directly connected to this one, gives the next concept's value, in each iteration. Equation (1) describes an inference law widely used in many FCM applications, but it is not the only one.

Generally, there are four different inference laws regarding the counting or not of the past concept's value or the existence of the self-connection link for each concept node. These inference laws are described in the following formulas and their behavior in classifying patterns by using FCM-based classifiers will analyzed in a next section.

Past & No Self-Connection (Inf1)
$$A_i^{t+1} = f\left(A_i^t + \sum_{i=1, i\neq j}^{N} W_{ji} A_j^t \right) \tag{5}$$

No Self-Connection (Inf2)
$$A_i^{t+1} = f\left(\sum_{i=1, i\neq j}^{N} W_{ji} A_j^t \right) \tag{6}$$

Past & Self-Connection (Inf3)
$$A_i^{t+1} = f\left(A_i^t + \sum_{i=1}^{N} W_{ji} A_j^t \right) \tag{7}$$

$$\textit{Self-Connection(Inf4)} \qquad A_i^{t+1} = f\left(\sum_{i=1}^{N} W_{ji} A_j^t\right) \tag{8}$$

While in the case of using FCMs in modeling complex systems where the concepts correspond to specific property of the system the selection of the appropriate inference law may be an easy task, in the case of the FCMpers this selection is not straightforward. For this reason, an additional procedure of finding the best inference law that yields better classification rates has to be proceed, as part of the overall calibration stage of the system.

2.4 FCMper – Learning Algorithm

In the early years, the interconnection weights of FCMs are decided by a group of experts which knew the relations in force between the concepts, since FCMs were used to model the behavior of a system, as far as its equilibrium from a certain initial state, is concerned. However, the need to equilibrate a system to a specific final state motivated the scientists to enhance the operation of the FCMs by incorporating an additional procedure of finding suitable weight sets called *learning* in conjunction to the similar operation existing in the ANNs training.

Many researchers in the recent years have tried to enhance the functionality of the FCMs by developing learning algorithms, which find the appropriate set of weight interconnections between the concepts. The proposed algorithms are divided into two types, these which are making use of gradient information and are called *gradient-based algorithms* (Papageorgiou et al. 2003, 2004) and those which are based on evolutionary mechanisms (Koulouriotis et al. 2001a, Stach et al. 2005, 2008, Papageorgiou et al. 2005, Ghazanfari et al. 2007). The former type algorithms have the disadvantage of trapping to local optimum by finding a solution far away from the theoretical optimal, while the latter one give more chances to find the best weight set, due to their stochastic nature.

Until now, the algorithms of both kinds search a weight set which can lead the FCM from a specific initial state to a desired equilibrium point. However, in the case of the FCM-based classifiers the learning algorithm should work quite differently. The learning algorithm for the case of FCMpers, have to find a common weight set, which for initial states of the input concepts that belong to the same class, the FCMpers equilibrate to the same points.

This learning procedure defines some new operational notions for the FCM structures, coping with their ability to form internal representations able to lead them to certain equilibrium points. Moreover, this need highlights the significance of storing enough knowledge to distinct the patterns presented as input concepts and also marks the importance to conduct the capacity of the models.

3 Experimental Study

The experiments are organized in two sections where in the first one the influence to the classifier's performance of the type of activation function and the inference law are investigated. In a second phase the classification capabilities of the FCMpers are compared to those of the traditional ANNs in difficult pattern recognition applications.

3.1 Experimental Part 1

For the needs of the first part experiments a set of well known benchmark pattern recognition datasets widely used in the literature are selected from the UCI repository (UCI-Machine Learning Repository) and their properties are summarized in Table 1.

Table 1 Properties of the datasets used in the experiments

Dataset	Number of Classes	Attributes	Instances
Iris Flowers	3	4	150
Pima Indians Diabetes	2	8	768
Wine	3	13	178
Glass Identification	7	9	214
Hepatitis	2	19	155
Echocardiogram	2	11	132
Breast Cancer Wisconsin	2	9	699
Parkinson's	2	22	195
Mammographic Mass	2	5	961

All the experiments are carried out by using a SGA (Simple Genetic Algorithm) as learning algorithm, without having any advanced genetic operator for preserving the diversity in high levels. This is the simplest form of an evolutionary algorithm, where its simplicity allows its application without any specific calibration procedure.

The MSE (Mean Squared Error) between the desired and the actual values of the FCMper1 output concepts is used as objective function, which has to be minimized and its analytical form is as follows.

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(OC_{ij}^{Actual} - OC_{ij}^{Desired} \right)^2 \tag{9}$$

where M is the number of training patterns, N the number of the FCMper output concepts and (OC_{ij}^{Actual} – $OC_{ij}^{Desired}$) the difference between the i^{th} output concept and its corresponding desired value (target), when the j^{th} set of input concepts appears to the FCMper's input.

Each experiment is executed 20 times in order to ensure the statistical accuracy of the outcomes and the corresponding mean results for the case of the first three datasets of Table 1 are summarized in the following Table 2, 3 and 4.

In these tables, the activation functions F4 corresponds to the sigmoid function having slope factor different than 1, which the SGA founds during the learning process the same for each node and the function F5 is the same with F4 but in this case each node has its own slope factor.

The performance measures used to study the classification performance of each combination of activation function and inference law, is the MSE measured during the learning procedure, the PMSE (Performance-MSE) which is the MSE measured by using a small fraction of the dataset called testing set (all the datasets are divided into the training 75% and testing 25% sets) and CR the *Classification Rate* defined as the percentage of the correctly classified patterns over the entire testing set.

Table 2 Classification statistics for the case of Iris Flowers dataset

Activation Function	**Inference Law**	**MSE**	**PMSE**	**CR(%)**
F1	Inf1	0.1730	0.1866	51.78
	Inf2	0.1340	0.1504	57.33
	Inf3	0.1797	0.1926	50.88
	Inf4	0.1314	0.1474	58.88
F2	Inf1	0.1362	0.1563	55.33
	Inf2	0.1384	0.1518	58.66
	Inf3	0.1511	0.1660	54.89
	Inf4	0.1273	0.1444	61.78
F3	Inf1	0.1690	0.1694	66.67
	Inf2	0.1696	0.1701	67.55
	Inf3	0.1712	0.1718	66.89
	Inf4	0.1733	0.1739	66.67
F4	Inf1	0.1078	0.1099	83.33
	Inf2	0.1043	0.1070	82.67
	Inf3	0.1080	0.1104	86.22
	Inf4	0.1097	0.1120	83.78
F5	Inf1	0.1129	0.1140	80.89
	Inf2	**0.1067**	**0.1083**	**89.11**
	Inf3	0.1069	0.1088	88.44
	Inf4	0.1058	0.1081	88.22

Table 3 Classification statistics for the case of Pima Indians Diabetes dataset

Activation Function	**Inference Law**	**MSE**	**PMSE**	**CR(%)**
F1	Inf1	0.3178	0.3236	56.07
	Inf2	0.3242	0.3238	58.60
	Inf3	0.3229	0.3340	56.11
	Inf4	0.3202	0.3312	55.68

Table 3 (*continued*)

F2	Inf1	0.3049	0.3218	52.44
	Inf2	0.3095	0.3207	54.06
	Inf3	0.3044	0.3244	49.65
	Inf4	0.3017	0.3275	48.43
F3	Inf1	0.2202	0.2169	67.07
	Inf2	0.2197	0.2160	66.77
	Inf3	0.2200	0.2165	67.07
	Inf4	0.2204	0.2174	67.16
F4	**Inf1**	**0.2105**	**0.2062**	**70.61**
	Inf2	0.2113	0.2078	69.74
	Inf3	0.2099	0.2051	69.17
	Inf4	0.2100	0.2071	70.35
F5	Inf1	0.2125	0.2080	70.48
	Inf2	0.2142	0.2102	69.48
	Inf3	0.2116	0.2078	69.43
	Inf4	0.2157	0.2125	69.17

Table 4 Classification statistics for the case of Wine dataset

Activation Function	Inference Law	MSE	PMSE	CR(%)
F1	Inf1	0.2403	0.2963	32.96
	Inf2	0.2159	0.2697	33.52
	Inf3	0.2449	0.2901	34.81
	Inf4	0.2645	0.3098	28.33
F2	Inf1	0.3126	0.3808	20.18
	Inf2	0.2325	0.3389	23.52
	Inf3	0.2763	0.3531	33.70
	Inf4	0.3188	0.3605	24.26
F3	**Inf1**	**0.1600**	**0.1600**	**81.67**
	Inf2	0.1541	0.1537	80.74
	Inf3	0.1551	0.1562	80.18
	Inf4	0.1601	0.1624	77.78
F4	Inf1	0.1007	0.1230	76.48
	Inf2	0.0888	0.1064	82.59
	Inf3	0.1145	0.1310	77.22
	Inf4	0.0941	0.1184	78.15
F5	Inf1	0.1225	0.1404	77.22
	Inf2	0.1078	0.1322	76.30
	Inf3	0.1118	0.1292	73.70
	Inf4	0.1225	0.1324	79.81

From the above tables, it is obvious that the Step and Trivalent functions are perform poorly, since they classify wrongly the patters, by giving an average classification rate of 55% for the case of the first two datasets. This fact makes them inappropriate for pattern recognition applications, while their reasoning capabilities are restricted in special cases of the FCM modeling (Tsadiras 2008).

As far as the affection of the inference law is concerned, what is making clear from the above results is that the self-feedback of each node does not play a crucial role in the overall operation of the FCMpe1.

Based on these observations it is useful to study the classification performance of the entire benchmark set only for the cases of F3, F4, F5 with Inf1 and Inf2 as inference laws. The mean classification results are illustrated in the following Table 5.

Table 5 Mean classification statistics for all benchmarks of Table 1

Activation Function	**Inference Law**	**MSE**	**PMSE**	**CR(%)**
F3	Inf1	0.1551	0.1655	72.33
	Inf2	0.1551	0.1643	72.77
F4	**Inf1**	**0.1252**	**0.1521**	**74.68**
	Inf2	**0.1233**	**0.1500**	**74.68**
F5	Inf1	0.1279	0.1558	73.27
	Inf2	0.1259	0.1534	73.91

Table 5, shows that the sigmoid function improves the classification capability of the FCMper1 and in its adaptive form where the slope factor is decided during the learning procedure gives the highest score. Moreover, in the case of using the sigmoid activation function, the type of the used inference law does not play an important role, something which is justified by the fact that the nonlinearity entered into the model, further improves its reasoning operation by making the influence of the inference law less significant.

Consequently, what is concluded by this experimental study is that, the FCM-based classifier the FCMper1, needs a nonlinear activation function in order to construct its internal representations of the patterns consisting the problem. Also, its performance significantly improved when the inference law does not consider self-feedback interconnections, while the past node's value is not important to the overall classification accuracy.

3.2 Experimental Part II

In this section, the most high performance configuration of the FCMper1, along with the two hybrid FCM-based classifiers FCMper2 and FCMper3, are used to classify patterns belonging to complex pattern recognition problems, one artificially generated and one real robotic vision application. Their classification performance is compared with that of conventional neural networks of similar structure.

The selected configuration for the FCM part of these classifiers includes the F4 activation function in conjunction with the Inf1 inference law, since with this settings the FCMper1 present good generalization behavior.

Moreover, the experiments' configuration is the same with the previous ones and the neural networks are trained by using the standard *backpropagation* algorithm (Haykin 1999), for all simulations.

3.2.1 The Two Spiral Problem

The two spiral problem is an extremely hard problem for neural networks to solve. The goal of this problem is to learn to discriminate between two sets of training points, which lie on two distinct spirals in the x-y plane (Langlet et al. 2001). Table 6, summarizes the classification performance of a typical multilayer perceptron neural network with 2 inputs, 10 hidden neurons in one hidden layer and 2 outputs, in comparison to the corresponding one of the FCMper structures, through appropriate indices.

Table 6 Classification Statistics in the case of the Two Spiral problem

Two Spiral				
	Neural Network	***FCM per1***	***FCM per2***	***FCM per3***
Structure	*2-10-2*	*2-2*	*2-2*	*4-2*
MSE	0.1505	0.2485	0.1768	0.1389
PMSE	0.1389	0.2459	0.3158	0.1349
CR	75.27%	56.56%	60.22%	**79.67%**
Stdv(MSE)	0.0922	0.0012	0.0990	0.1093

The experimental pattern set is consisted of 100 training and 50 testing patterns. In the above table *Stdv* is the *Standard Deviation* of the MSE, as an index of the computational stability of the learning algorithms.

The results presented in Table 6, show that the FCMper3 (last column in bold face), outperforms the other FCMpers, even the neural classifier. More precisely, the FCMper3 having smaller structure (only 16 adjustable parameters) than neural network (52 weights and biases), gives better classification rate of almost 80%, with smaller training error.

3.2.2 A Robotic Vision System

In this experiment a typical classification application, where the Zernike moments are used as discriminative features of the objects be classified, is taken place. Initially some test objects (patterns) are selected. Figure 4b shows a wooden pyramidal puzzle, which is used for robot vision tasks in the Control Systems Lab of DUTH (Democritus University of THrace) (Papakostas et al. 2005). The 9 parts of the puzzle, placed in arbitrary positions, are shown in Figure 4a. The (256x256 pixels) images of these parts are the initial nine patterns of our experiments.

(a) (b)

Fig. 4 The nine work pieces that are placed (a) in arbitrary positions on the table and (b) on a 3-D truncated pyramid

The pattern dataset consists of a 594 images from which 324 are used during the training procedure and the rest 270 are used to test the generalization abilities of the structures being compared. More information about the way the dataset is formed can be found in (Papakostas et al. 2005).

For experimental purposes, the first 12 Zernike Moments of each image are computed and used as features to distinguish the 9 parts. From this set the first two moments are omitted because they are constant for all objects and thus they do not constitute discriminative quantities.

After a typical cross-validation procedure (Haykin 1999) it was decided, the neural classifier to be a typical multilayer perceptron containing 1 hidden layer, with 10 hidden neurons and 9 (equal to the number of classes) output neurons.

This problem differs from the previous ones, since the data set has been formed directly from a real environment, the captured images of the objects. The experimental results of this classification problem are presented in Table 7.

Table 7 Classification Statistics in the case of the robotic vision system

Robotic Vision System				
	Neural Network	***FCM per1***	***FCM per2***	***FCM per3***
Structure	*10-10-9*	*10-9*	*9-9*	*19-9*
MSE	0.0849	0.2415	0.1976	0.0739
PMSE	0.0741	0.2390	0.3395	0.0723
CR	87%	56.2%	58.14%	**89.33%**
Stdv(MSE)	0.0863	0.0150	0.1309	0.0894

The superiority of the FCMper3, is still preserved in this real problem, as it can be seen by Table 7.

4 Conclusions – Discussion

A novel application of Fuzzy Cognitive Maps was presented in this chapter. The appropriate principles that define the basic operation of the FCMs as pattern classifiers were demonstrated and the main functional properties of these FCM-based classifiers were discussed. A FCM-based classifier have the advantage to give the capability to adjust its operation by selecting the appropriate set of free parameters, such as the activation function, the inference law, the structure's connectionist in order to take the highest efficiency depending on the specific problem.

The newly introduced classifiers present high classification capabilities even in their simple form of the FCMper1 and they outperform the traditional neural classifiers when they work in more complicated hybrid forms.

While this study constitutes a first attempt to evolve the FCMs in a way they can be used as classifier modules in modern pattern recognition systems, new directions are raised, regarding the increase of the knowledge capacity by adding dummy concepts and finding activation functions or inference laws, that need more research in the name of high classification performance.

References

Boutalis, Y., Kottas, T.L., Christodoulou, M.: Adaptive etimation of fuzzy cognitive maps with proven stability and parameter convergence. IEEE Transactions on Fuzzy Systems 17, 874–889 (2009)

Bueno, S., Salmeron, J.L.: Benchmarking main activation functions in fuzzy cognitive maps. Expert Systems with Applications 36, 5221–5229 (2009)

Chytas, P., Glykas, M., Staikouras, C., Valiris, G.: Performance measurement in a greek financial institute using the balanced scorecard. Journal of Measuring Business Excellence 10, 87–98 (2006)

Ghazanfari, M., Alizadeh, S., Fathian, M., Koulouriotis, D.E.: Comparing simulated annealing and genetic algorithm in learning FCM. Applied Mathematics and Computation 192, 56–68 (2007)

Glykas, M., Xirogiannis, G.: A soft knowledge modeling approach for geographically dispersed financial organizations. Soft Computing 9, 579–593 (2004)

Haykin, S.: Neural Networks – a comprehensive foundation, 2nd edn. Prentice Hall, Englewood Cliffs (1999)

Khan, M.S., Quaddus, M.: Group decision support using fuzzy cognitive maps for causal reasoning. Group Decision and Negotiation 13, 463–480 (2004)

Kosko, B.: Fuzzy Cognitive Maps. International Journal Man-Machine Studies 24, 65–75 (1986)

Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M.: Learning fuzzy cognitive maps using evolution strategies: a novel schema for modeling and simulating high-level behavior. IEEE Congr. Evolutionary Computation, 364–371 (2001a)

Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M.: A fuzzy cognitive map-based stock market model: synthesis, analysis and experimental results. In: IEEE International Conference on Fuzzy Systems (FUZZ 2001), vol. 1, pp. 465–468 (2001b)

Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M.: Realism in fuzzy cognitive maps: incorporating synergies and conditional effects. In: IEEE International Conference on Fuzzy Systems (FUZZ 2001), vol. 3, pp. 1179–1182 (2001c)

Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M.: Anamorphosis of fuzzy cognitive maps for operation in ambiguous and multi-stimulus real world environments. In: IEEE International Conference on Fuzzy Systems (FUZZ 2001), vol. 3, pp. 1156–1159 (2001d)

Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M., Antonidakis, E.N., et al.: Efficiently modeling and controlling complex dynamic systems using evolutionary fuzzy cognitive maps. International Journal of Computational Cognition 1, 41–65 (2003) (invited Paper)

Koulouriotis, D.E.: Investment analysis and decision making in markets using adaptive fuzzy causal relationships. Operational Research: an International Journal 4, 213–233 (2004)

Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M., Zopounidis, C.D.: Development of dynamic cognitive networks as complex systems approximators: validation in financial time series. Applied Soft Computing 5, 157–179 (2005)

Langlet, F., Abdulkader, H., Roviras, D., Mallet, A., et al.: Comparison of neural network adaptive predistortion techniques for satellite down links. In: Proceedings of International Joint Conference on Neural Networks, pp. 709–714 (2001)

Pajares, G., de la Cruz, J.M.: Fuzzy cognitive maps for stereovision matching. Pattern Recognition 39, 2101–2114 (2006)

Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P.: Fuzzy cognitive map learning based on nonlinear Hebbian rule. In: Australian Conference on Artificial Intelligence, pp. 256–268 (2003)

Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P.: Active Hebbian learning algorithm to train fuzzy cognitive maps. International Journal of Approximate Reasoning 37, 219–249 (2004)

Papageorgiou, E.I., Parsopoulos, K.E., Stylios, C.S., Groumpos, P.P., et al.: Fuzzy cognitive maps learning using particle swarm intelligence. Journal of Intelligence Information Systems 25, 95–121 (2005)

Papakostas, G.A., Karras, D.A., Mertzios, B.G., Boutalis, Y.S.: An efficient feature extraction methodology for computer vision applications using wavelet compressed Zernike moments. ICGST International Journal on Graphics, Vision and Image Processing, Special Issue: Wavelets and Their Applications SI1, 5–15 (2005)

Papakostas, G.A., Boutalis, Y.S., Koulouriotis, D.E., Mertzios, B.G.: A first study of pattern recognition by using fuzzy cognitive maps. In: 13th International Workshop on Systems, Signals and Image Processing (IWSSIP 2006), pp. 369–374 (2006)

Papakostas, G.A., Boutalis, Y.S., Koulouriotis, D.E., Mertzios, B.G.: Fuzzy cognitive maps for pattern recognition applications. International Journal of Pattern Recognition and Artificial Intelligence 22, 1461–1468 (2008)

Stach, W., Kurgan, L., Pedrycz, W., Reformar, M.: Genetic learning of fuzzy cognitive Maps. Fuzzy Sets and Systems 153, 371–401 (2005)

Stach, W., Kurgan, L.A., Pedrycz, W.: Numerical and linguistic prediction of time series with the use of Fuzzy Cognitive Maps. IEEE Transactions on Fuzzy Systems 16, 61–72 (2008)

Tsadiras, A.K.: Comparing the inference capabilities of Binary, Trivalent and Sigmoid fuzzy cognitive maps. Information Sciences 178, 3880–3894 (2008)

UCI-Machine Learning Repository, `http://archive.ics.uci.edu/ml/datasets.html`

Xirogiannis, G., Stefanou, J., Glykas, M.: A fuzzy cognitive map approach to support urban design. Expert Systems with Applications 26, 257–268 (2004a)
Xirogiannis, G., Glykas, M.: Fuzzy cognitive maps in business analysis and performance driven change. IEEE Transactions on Engineering Management 51, 334–351 (2004b)
Xirogiannis, G., Glykas, M.: Intelligent modeling of e-business maturity. Expert Systems with Applications 32, 687–702 (2007)
Xirogiannis, G., Chytas, P., Glykas, M., Valiris, G.: Intelligent impact assessment of HRM to the shareholder value. Expert Systems with Applications 35, 2017–2031 (2008)