A COMBINATION OF CLASSICAL AND FUZZY CLASSIFICATION TECHNIQUES ON A SELF ORGANIZED MEMORIES (SOM)-TYPE NEURAL NETWORK COMPUTATIONAL PLATFORM

Spyros G. Tzafestas[1] *and Sotiris N. Raptis*[2]
Intelligent Robotics And Automation Laboratory (I.R.A.L.)
Department Of Electrical & Computers Engineering
National Technical University Of Athens
Zographou, Athens, GR 157 73, Greece
e-mail: sotrap@softlab.ntua.gr[2], tzafesta@softlab.ece.ntua.gr[1]

ABSTRACT

In this paper a complex classification scheme including a combination of a cluster generator based on a neural network and three different types of classifiers is proposed. The first part of the scheme consists of a self organized memory S.O.M.-type unsupervised neural network which converges quickly to the clusters′ vectors. In the second part, which is the main classification part, three classifiers are being compared. This comparison is made in a such a way so as to study the improvement of the performance and classification when getting from purely classical classification schemes to fuzzy ones and thus emphasize on the usefulness and reliability of the fuzzy set theory. The overall approach aims at showing that neural networks and fuzzy classifiers can be combined in a such a way which exploits the advantages of both approaches in classification problems. Neural networks, on the one hand, operate very quickly after they are trained but they are not capable of recognizing easily information not familiar to them. On the other hand fuzzy systems overcome with success this drawback because they are generalized by nature. However they do not provide us with straight decisions. Instead, they give an estimation of the nature of the problem associating. They thus give us more than one possible solutions. In literature, the most common way of putting together these two concepts of neural networks and fuzzy systems is the adoption of the neuro-fuzzy schemes which in their turn are constructed by fuzzy neurons instead of classical neurons. This paper proposes a two stage scheme, and in addition studies three classical algorithms a classical one such as Nearer Neighbor N.N.R., a generalization of it, the Fuzzy-N.N.R., and a well known fuzzy classifier the Fuzzy C-Means F.C.M. . Results show that fuzzy N.N.R. operates with significantly better performance in terms of classification performance like mean classification error, and operational time expressed by convergence time. However it is shown that the fuzzy schemes are more reliable than the fuzzified classical schemes. The structure of the scheme is simple. There is no need for long convergence delays and complex learning procedures as the scheme is a S.O.M.-type one. It is especially designed for optimal and quick convergence to the vectors of clusters.

I. INTRODUCTION

Neural networks have been widely known and used as classifiers [1], [8], [11]. They are based on their learning capability and their parallel structure which enables them to elaborate massively the information presented to them. A widely known application where neural networks have been mainly used is associating input patterns to other patterns-memories. These new patterns are kept and stored within the network in order to be used later as fast memories which can be quickly associated to a random new input pattern [1], [3], [4]. Hence here we have some kind of classification problem. However, the learning procedure of the neural networks enables them to treat effectively and quickly one kind of information, the one already presented to them, without exhibiting the same performance when dealing with data not familiar to them.

On the other hand fuzzy systems overcome the above disadvantage by subtracting from the data presented to them any detailed information that may, on the price of the accuracy, make them lose their generality and hence their flexibility. When dealing with uncertainty, as fuzzy systems do so, the accuracy in data description is compensated by the capability of elaborating data of a very different nature, which is a virtue of the fuzzy logic and an advantage against all other methods in data processing [11].

Hence, a combination of these two approaches in the field of pattern recognition would be anticipated to exploit and emphasize the advantages of both approaches. In this paper a two modules scheme is described and proposed. Three variations of it are studied. This scheme consists basically of a S.O.M.-type neural network, in its first part, and a classifier in its second part. Furthermore, in the second part three classifiers are compared in terms of their performance within the framework of the general two parts scheme, namely the F.C.M. classifier, the classical N.N.R. classifier and its generalization the Fuzzy-N.N.R. classifier. The basic elements of this scheme have been studied widely in the literature. The proposed combination however differs from the ordinary neuro-fuzzy schemes which are founded on the concept of the fuzzy neuron. The study of this scheme is presented in order to show the performance of putting sequentially fuzzy and neural systems. In the second part the comparison of these three classification schemes aims at a parallel study of the fuzzy clustering methods against classical classification methods. In order to make this comparison more useful the three classification techniques of the second part are chosen in a such a way to show the transition between classical and fuzzy classification techniques. Therefore the second classification technique is a generalization of the classical one using fuzzy logic while the third is a purely fuzzy classification algorithm.

The combination of these two modern signal processing concepts in an effective and useful way is crucial to construct intelligent, robust and reliable algorithms. The framework currently developed in the literature for combining these two concepts is based on the notion of the fuzzy neuron F.N. or fuzzy output cell. Intelligent systems based on the structural element of the fuzzy neuron are called neuro-fuzzy systems N.F. [10]. The F.N. consists of a classical neuron part and a fuzzification part whose output is the membership value. More specifically the output defined as the state-value of the neuron is an input to a fuzzification function whose output is the membership value. Each fuzzy output node is associated to a fuzzy sub-set. The input vector applied to the N.F system is the pattern to be classified. Hence this membership value relates fuzzily the pattern applied to a specific F.N. with the fuzzy cluster associated to the specific F.N. . Mathematically these concepts are represented as follows. Let $\underline{x}_i, i=1,\ldots,N$ be an input pattern, $\underline{w}_j, j=1,\ldots,c$ be a weight vector associated with to an output neuron $y_j, j=1,\ldots,c$, $g(\bullet)$ be the neuron′s j-th activation function. Then the function value over the product $\underline{x}_i^T * \underline{w}_j$ represents the neuron state value ($y_j=g(\underline{x}_i^T * \underline{w}_j)$). This value also represents classical neuron output value. A N.F. system extends these concepts by fuzzifying this value through a fuzzification function $f(\bullet)$ which acts on the classical neuron′s state value y_j, producing the membership value $\mu_{ji}=f(y_j)=f(g(\underline{x}_i^T * \underline{w}_j))$. Network′s learning procedure includes both classical neuron and fuzzy neuron parameters adjustment since the F.N.′s output μ_{ji} is put in some kind of N.F. system convergence function. This mathematical formulation is at the kernel of a whole class of N.F. systems. The vary and differ from each other in terms of the number of layers, of the learning rules, of the parameters to tune, of the bias values, of the types of the activation functions and the types of the fuzzification functions. These computational elements and parameters can be combined in numerous variations. Each such variation is specialized to resolve better a certain family of problems such as those met in the classical neural networks case and more.

However, although these schemes present generally more reliability and effectiveness in many pattern recognition problems like character recognition, medical image applications, shape identification and disease diagnosis, nodules identification within chest radiographs [12] etc., they are computationally complex and demanding and of long learning and convergence time. Of coarse applications such as highly demanding identification tasks require in order to achieve highly ranked decision accuracy and correctness a highly complex computational and intelligent scheme capable of capturing any information and feature that can provide a subtle distinction between patterns that are not easy to discriminate using more simple schemes. An other critical drawback characterizing these complex schemes is their specialization and their task oriented structure and learning process which is mainly due to their classical neural network computational basis.

In this paper we seek to investigate through the study of some simple cases the efficiency of using sequentially and not integrated the one into the other these pattern recognition methods, i.e. the classical neural networks and the fuzzy systems. These two modules operate independently. Hence the learning procedure concerns only the classical neural network parameters. Furthermore, here, an unsupervised scheme is selected so as to make the whole scheme independent and self organizing. We applied our method in a very simple classification task using a chest nodule block and have investigated the flexibility in performance of this scheme by putting it together with different types of classical and fuzzy classifiers.

II. METHODOLOGY

The first module consists of a self organized unsupervised neural network that produces the clusters in the form of its weight vectors. Self organized memories are designed to memorize or encode in their weights their input data series. If the input data have the form of a $N \times p$ matrix where N is the number of input patterns presented to the network and p their dimension, then the network can be designed to associate this series of patterns through a $N \times p \rightarrow N \times p_o$ mapping to a set of N patterns of the same or different dimension p_o. If the memory pattern dimension p_o is chosen to be equal to p then this new set of generated memory patterns can perform as a set of clusters into which original patterns can be attributed. As the S.O.M. scheme is an unsupervised one there is no need for learning sets to be used. The general two stage scheme used is depicted in the diagram contained in Fig. 1.

The type of network used here has the following learning rule :

$$d\underline{w}_i(t) = a(t)(\underline{x}(t) - \underline{w}_i(t)), \; i = 1,\ldots,c$$

where $\underline{x}(t)$ is the image block presented to the network at the time instance t, $\underline{w}_i(t)$ is the weight vector of the i-th output neuron at the time instance t, $\alpha(t)$ is the learning rate at the same time which is taken here as a constant (a=.5), c is the number of output nodes i.e. the number of clusters. The above relation is valid only for the winning node. As such is defined that one whose weight vector is the closer to the current input vector $x(t)$. The algorithm converges quickly and terminates when the weight matrix $W=[\underline{w}_1,\underline{w}_2,\ldots,\underline{w}_c]$ practically remains constant over iterations.

A nearest neighbor classifier module found after the S.O.M. generator is based on the idea that when classifying a single pattern into a set of clusters then that pattern would be classified most likely and accurately into that cluster into which most of its already classified neighboring patterns belong. Thus for each pattern to classify a set of predefined clusters is prescribed. According to crisp-classical set theory the pattern can be classified only into one of those neighboring clusters without having any other relationship or association with the other clusters. Following this logic if for a given pattern $\underline{x}_i$,i=1,…,N a set C_i of neighboring clusters is defined then that pattern has to be classified into one of these uniquely, say $\underline{c}_o$ without dealing with the other clusters of the set C_i.

The next algorithm to consider is the generalized, fuzzified N.N.R., F.N.N.R., algorithm proven to exhibit much better performance than the classical N.N.R. algorithm in terms of representation error and convergence time. To ground F.N.N.R. we extend the concept of neighbors contribution by assuming that each new pattern to classify is attributed a degree of association to all clusters in C_i which is called in fuzzy terminology *membership value*. This value relates $\underline{x}_i$ to each cluster in C_i. The association of the given pattern to the sub-set C_i of clusters can be done through the use of F.C.M. .

The third classifier to consider is the well known F.C.M. . F.C.M. differs from F.N.N.R. in that each pattern to classify is even partially attributed to all clusters produced by the S.O.M. module and not only its neighbors. Thus each pattern is related through a certain degree of membership to all eventual clusters. There are not zero memberships in F.C.M. . Therefore memberships take lower values in F.C.M. than in F.N.N.R. .

A widely known technique for image segmentation and clustering is the fuzzy C-means algorithm (F.C.M.) which is a variation of the classic C-means using fuzzy reasoning. F.C.M. is described in details below.

Given a set of c clusters, F.C.M., classifies input patterns by minimizing a cost function of the type :

$$J(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{N}(\mu_{ji})^m d_{ik}$$

where:

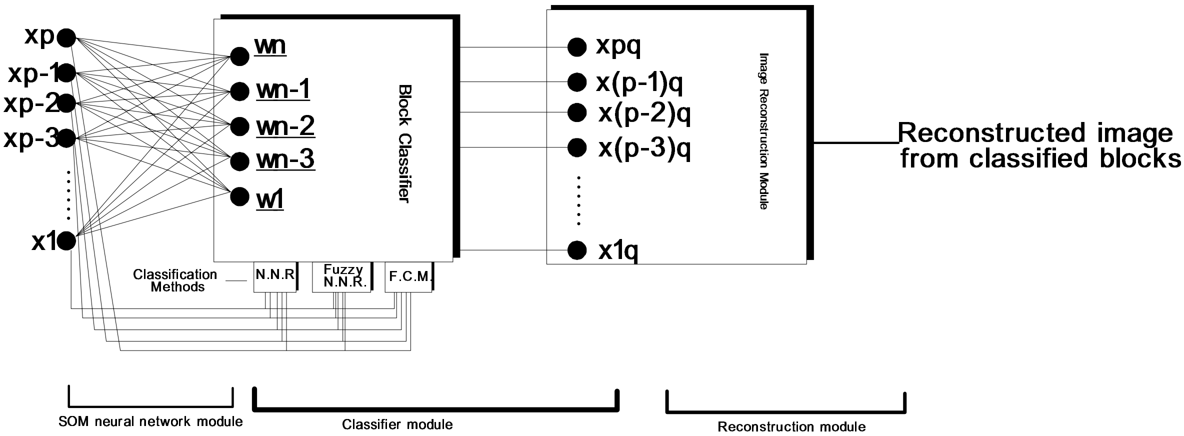
- μ_{ki} : stands for the membership value of pattern i into cluster j
- c : stands for the number of clusters
- U : stands for the fuzzy partition matrix
- N : stands for the number of input patterns
- $\underline{v}_j$,j=1,…,c : represent the vectors-centers of the fuzzy clusters
- d_{ij} : stands for the distance between the patterns $\underline{x}_i$ and $\underline{x}_j$.

The algorithm′s steps are given below:

The algorithm above presents relatively quick convergence and behaves very good mathematically. It is proven that human eye operates about the same way in recognizing colors within an image.

Parameters controlling execution and convergence of the algorithm are:

Step 1	1.1. Define the number of clusters c, $2 \le c \le N$ 1.2. Define the membership factor m, $1 \le m < \infty$ 1.3. Define the metric used in distance measurement $\lvert * \rvert$
Step 2	Initialize fuzzy partition matrix $U^{(0)}$
Step 3	At each step b, b=1,2,… 3.1. Calculate the vectors of the clusters centers $\underline{v}_i$,i=1,…,c using the formula : $\underline{v}_j = \frac{\sum_{j=1}^{c}(\mu_{ji})^m \underline{x}_i}{\sum_{i=1}^{N}(\mu_{ji})^m}$ 3.2. Update the fuzzy partition matrix $U^{(b)}$ and find the next one $U^{(b+1)}$ as follows : 3.2.1. Calculate for each pattern $\underline{x}_i$ the set of clusters that pattern is associated to : $I_i = \left\{k \middle\vert 1 \le k \le c, d_{ik} = \left\Vert \underline{x}_i - \underline{v}_k \right\Vert = 0\right\}$ and the set of clusters T_i=[1,…,c]-I_i the pattern is not associated to. 3.2.2. $\begin{cases} if \quad I_i \neq \emptyset \;\; then \quad \mu_{ji} = \dfrac{1}{\left[\sum_{j=1}^{c}\left(\dfrac{d_{ji}}{d_{ji}}\right)^{2(m-1)}\right]} \\ if \quad I_k = 0 \quad then \quad \mu_{ji} = 0 \;\; \forall k \in T_i \;\; and \\ \sum_{k \in I_i}\mu_{ji} = 1 \end{cases}$
Step 4	Compare $U^{(b)}$ and $U^{(b+1)}$ using some predefined metric and if $\left\Vert U^{(b)} - U^{(b+1)} \right\Vert \le \varepsilon_L$ stop otherwise go to step 3



1. The number of clusters c
2. The convergence and fuzzification factor m
3. The convergence threshold ε_L
4. The initial value of the partition matrix $U^{(0)}$
5. The metric it is used to measure distances $\|*\|$
6. The number of input patterns N

III. RESULTS

All algorithms have been tested using a variable number of initial clusters ranging from 10 to 100 to evaluate performance behavior over changing of clustering conditions. For F.C.M. and F.N.N.R. a convergence tolerance of ($\varepsilon_L \approx$[1e-5 1e-4]) has been selected. For F.N.N.R. and N.N.R. algorithms a neighbors set of 3 or 4 patterns has been selected.

The algorithms have been tested on the same image of a chest nodule which is of reduced size exhibiting a certain degree of complexity so as to easily observe tHE differences in algorithms performance and representation quality. In Fig. 1 is shown the original image ′nodule′.

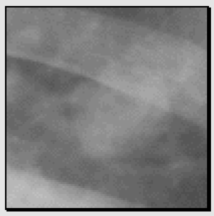
In Table 1 are given some representative and comparative results for all three types of algorithms, namely, N.N.R., F.N.N.R., F.C.M.

Figures Fig.2 to Fig. 31 contain the reconstructed images using the classified image blocks for a variable number of clusters beginning from 10 up to 100. Figs. 2 to 11 contain the classified images using classical N.N.R. algorithm. Figs. 12 to 21 contain the images reconstructed by use of F.N.N.R. algorithm. In the third column are contained these by use of the F.C.M. algorithm.

Table 1
PRESENTATION OF COMPARISON RESULTS

	Nearer Neighbor N.N.R.	Fuzzy N.N.R. F.N.N.R.	Fuzzy C-Means F.C.M.
Number of clusters	Error per block	Error per block	Error per block
10	441	425	9
20	1e3	346	8
30	1.039e3	285	8.7
40	1.093e3	276	8.1
50	1.0597e3	265	8.2
60	1.0575e3	251	8.1
70	1.0431e3	238	0.0014
80	811.4097	230	7.4e-4
90	1.0186e3	221	5.7e-4
100	976.0193	211	4.2e-4

Fig. 1
NODULE ORIGINAL IMAGE



IV. CONCLUSIONS

In this paper an attempt is tried to investigate the capabilities of a novel scheme based on neural networks and fuzzy classifiers. A separated clusters generator procedure has been used in order to achieve better convergence time and to avoid long term learning tasks. In addition in the classification process both classical and fuzzy methods have been compared. Unlike schemes combining neural networks and fuzzy nodes integrated into a compact computational module in this paper the satisfactory performance of putting separately these modules has been evaluated. Of coarse for more demanding applications requiring highly confident decisions in specialized cases the schemes based on the notion of fuzzy neuron exhibit much higher performance. However they still rely behind the current scheme in terms of structural simplicity and fast functioning. An other crucial conclusion that can be drawn is that the easy extension of classical algorithms using fuzzy set theory terms provides better performance at practically now price since membership values can be quickly computed using a standard module like the one used in F.C.M.. . Furthermore one can break down the whole classification procedure in district stages each one of which could be easier and faster implemented by a autonomous module.

REFERENCES

[1] Zheru Chi, Jing Wu and Hong Yan, ′Handwritten Numerical Recognition Using Self-Organizing Maps and Fuzzy Rules′, *Pattern Recognition*, Vol. 28, No. 1, pp. 59-66, 1995

[2] Lei Xu, ′Theories for Unsupervised Learning : P.C.A. and Its Non Linear Extensions′,

[3] Coombes and J. G. Taylor, ′Using generalized principal components analysis to achieve associative memory in a Hopfield net′, *Neural Network 5*, 1994

[4] Francesco Palmieri, ′Hebbian Learning and Self Association in Nonlinear Neural Networks′,

[5] Robert L. Shannon, Jitendra V. Dave, and James C. Bezdek, ′Efficient Implementation of the Fuzzy C-means Clustering Algorithm′, *I.E.E.E. Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 2, March 1986

[6] Richard J. Hathaway, ′Local Convergence of the Fuzz C-means Algorithms′, and James C. Bezdek, *Pattern Recognition*, Vol. 19, No. 6, pp. 477-480, 1986

[7] Nikhil R. Pal and James C. Bezdek, ′On Cluster Validity for the Fuzzy c-Means Model′, *I.E.E.E. Trans. Fuzzy Systems*, Vol. 3, No. 3, August ′95

[8] Hon Keung Kwan, Yaling Cai, ′A Fuzzy Neural Network and its Application to Pattern Recognition′, *I.E.E.E. Trans. Fuzzy Systems*, Vol. 2, No. 3, August ′94

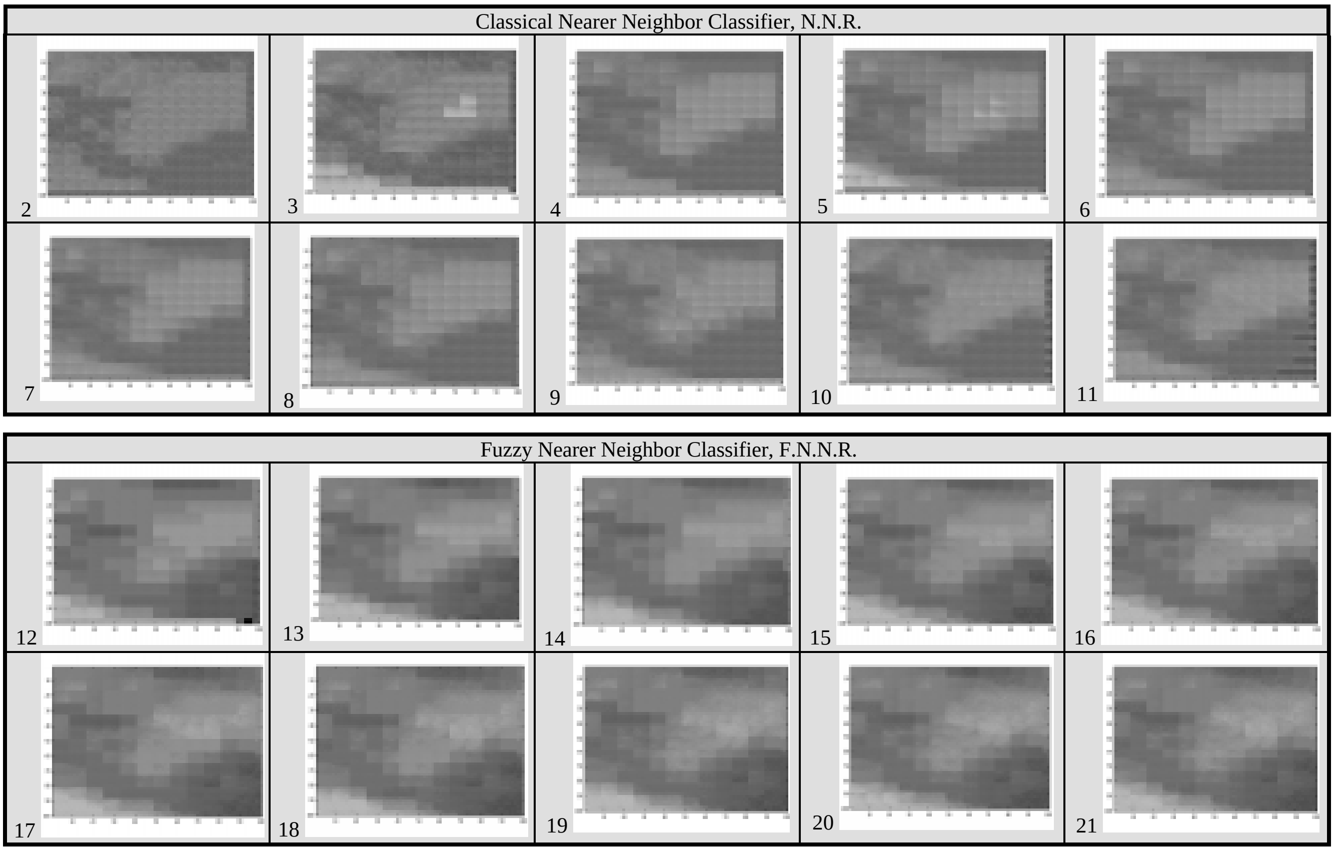
[9] Paul Gaber, James M. Keller, and Juliet Cai, ′A Fuzzy Logic System for the Detection and Recognition of Handwritten Street Numbers′, *I.E.E.E. Trans. Fuzzy Systems*, Vol. 3, No. 1, February ′95, pp. 83-95

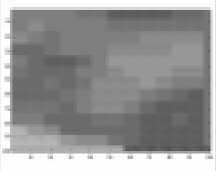
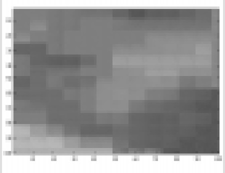
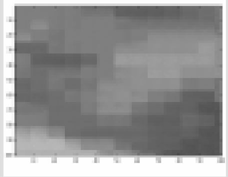
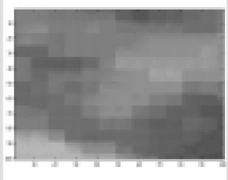
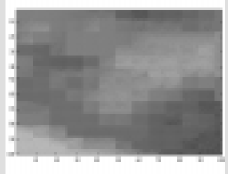
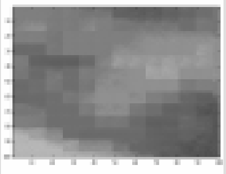
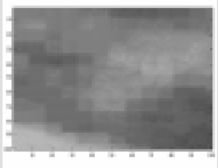
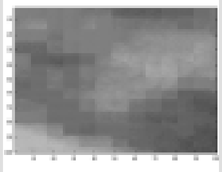
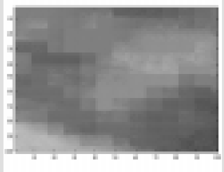
[10] Hon Keung Kwan, Yaling Cai, ′A Fuzzy Neural Network and its Application to Pattern Recognition′, *IEEE Trans. on Fuzzy Systems,* Vol. 2, No. 3, Aug. 1994, pp. 185-193

[11] Timothy J. Ross, ′Fuzzy Logic with Engineering Applications′, eds. McGraw-Hill

[12] Shih-Chung B. Lo, Shyh-Liang A; Lou, Jyh-Shyh Lin, Matthew T. Freedman, Minze V. Chien, and Seong K. Mun, ′Artificial Neural Network Techniques and Applications for Lung Cancer Nodule Detection′, *IEEE Trans. Medical Imaging*, Vol. 14, No. 4, Dec. ′95, pp. 711-718.

Figs 2-31. RECONSTRUCTED CLASSIFIED IMAGES USING N.N.R., F.N.N.R., F.C.M. ALGORITHMS



Fuzzy C-Means, F.C.M.				
22 	23 	24 	25 	26 
27 	28 	29 	30 	31