World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

Neuro-fuzzy Classification System for Wireless-Capsule Endoscopic Images

Vassilis S. Kodogiannis, John N. Lygouras

***Abstract*—**In this research study, an intelligent detection system to support medical diagnosis and detection of abnormal lesions by processing endoscopic images is presented. The images used in this study have been obtained using the M2A Swallowable Imaging Capsule - a patented, video color-imaging disposable capsule. Schemes have been developed to extract texture features from the fuzzy texture spectra in the chromatic and achromatic domains for a selected region of interest from each color component histogram of endoscopic images. The implementation of an advanced fuzzy inference neural network which combines fuzzy systems and artificial neural networks and the concept of fusion of multiple classifiers dedicated to specific feature parameters have been also adopted in this paper. The achieved high detection accuracy of the proposed system has provided thus an indication that such intelligent schemes could be used as a supplementary diagnostic tool in endoscopy.

***Keywords*—**Medical imaging, Computer aided diagnosis, Endoscopy, Neuro-fuzzy networks, Fuzzy integral.

I. INTRODUCTION

IN medical practice, endoscopic diagnosis and other minimally invasive imaging procedures, such as computed tomography, magnetic resonance imaging, are now permitting visualization of previously inaccessible regions of the body. Their objective is to increase the expert's ability in identifying malignant regions and decrease the need for intervention while maintaining the ability for accurate diagnosis.

Conventional diagnosis of endoscopic images employs visual interpretation of an expert physician. Since the beginning of computer technology, it becomes necessary for visual systems to "understand a scene", that is making its own properties to be outstanding, by enclosing them in a general description of an analyzed environment. Computer-assisted image analysis can extract the representative features of the images together with quantitative measurements and thus can ease the task of objective interpretations by a physician expert in endoscopy. A system capable to classify image regions to normal or abnormal will act as a second - more detailed - "eye" by processing the endoscopic video.

Endoscopic images possess rich information [1], which facilitates the abnormality detection by multiple techniques. However, from the literature survey, it has been found that only a few techniques for endoscopic image analysis have been reported and they are still undergoing testing. In addition, most of the techniques were developed on the basis of features in a single domain: chromatic domain or spatial domain. Applying these techniques individually for detecting the disease patterns based on possible incomplete and partial information may lead to inaccurate diagnosis. For example, regions affected with bleeding and inflammation may have different color and texture characteristics. Parameters in the spatial domain related with lumen can be used to suggest the cues for abnormality. For instance, small area of lumen implies the narrowing of the lumen which is often one of the symptoms for lump formation and not the presence of possible bleeding. Therefore, maximizing the use of all available image analysis techniques for diagnosing from multiple feature domains is particularly important to improve the tasks of classification of endoscopic images.

Krishnan, *et al.* [2] has been using endoscopic images to define features of the normal and the abnormal colon. New approaches for the characterization of colon based on a set of quantitative parameters, extracted by the fuzzy processing of colon images, have been used for assisting the colonoscopist in the assessment of the status of patients and were used as inputs to a rule-based decision strategy to find out whether the colon's lumen belongs to either an abnormal or normal category.

Endoscopic images contain rich information of texture. Therefore, the additional texture information can provide better results for the image analysis than approaches using merely intensity information. Such information has been used in CoLD (colorectal lesions detector) an innovative detection system to support colorectal cancer diagnosis and detection of pre-cancerous polyps, by processing endoscopy images or video frame sequences acquired during colonoscopy [3].

Traditional techniques of endoscopic exploration of the esophagus, stomach, and bowel adopt probes which are introduced into the oral or rectal cavities by means of insertion tubes. Such techniques permit detailed and reliable analyses and represent the best nonsurgical tool available today to manage some diseases of the digestive tube. Nevertheless, both gastroscopy and colonoscopy are considerably invasive and frequently ill tolerated by patients. Moreover, they have to

Manuscript received May 20, 2008. This work was supported by the European Commission, 5th Framework Intracorporeal Video-Probe (IVP) research project under Grant No. IST-2001-35169

Dr. Vassilis S. Kodogiannis is in the Centre for Systems Analysis, School of Computer Science, Univ. of Westminster, London HA1 3TP, UK (telephone: +44-777-570976, e-mail: kodogiv@wmin.ac.uk).

Assoc. Prof. J.N. Lygouras is in the Dept. of Electrical & Computer Engineering, Democritus University of Thrace, Greece.

be performed by skilled personnel. As a different technology, the use of the so-called endoscopic capsules for investigations of the digestive apparatus is progressively increasing today [4]. These systems consist of small capsules, which are swollen. They embed several functions, such as illumination of the tract under exploration, image capture, and wireless data transmission towards the exterior of the body [5]. Fig. 1 shows a picture of the most diffused commercial device (M2A/Pillcam, by Given Imaging). Following its swallowing, an endoscopic capsule moves along the digestive tube and a micro-camera mounted onboard takes pictures automatically for several hours. The device is then naturally expelled out of the body. During navigation, pictures are transmitted and recorded into an external unit, so that to be viewed as a video–frame sequence afterwards. Studies have revealed the capsule usefulness for the exploration of the small bowel, which cannot be examined with traditional endoscopic means [6].

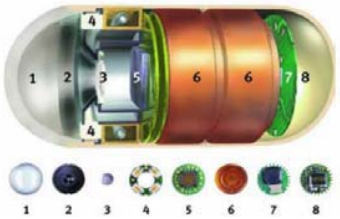

Fig. 1 Given Imaging Capsule- (1) Optical dome, (2) Lens holder, (3) Lens, (4) Illuminating LEDS, (5) CMOS imager, (6) Battery, (7) ASIC transmitter, (8) Antenna.

A computerized system for assisting the analysis of wireless capsule endoscopic (WCE) data using the M2A capsule, and identifying sequences of frames related to small intestine motility has been implemented recently [7]. To encode patterns of intestinal motility, a panel of textural and morphological features of the intestine lumen was extracted. The recognition of contractions was carried out by means of a Support Vector Machines classifier. In order to automatically cluster the different types of intestinal contractions in WCE, a computer vision system which describes video sequences in terms of classical image descriptors has been developed [8]. In an alternative approach, a methodology for detecting bleeding in WCE images has been investigated [9]. The presented methodology used color adaptation using a reference color value of the image. The adaptation was carried out by means of the cone response transform and the reference color was detected using Neural Networks. This method was applied to gastroenterological images in order to detect the presence of blood in a specific frame sequence. In a parallel study, a framework for multiple features analysis of WCE images has been proposed [10]. Several non-color analysis tools are proposed, notably: contrast enhancement, chamfer matching, thresholding, and image similarity. Experimental results suggest that the proposed multiple features analysis can guide the specialist to areas with abnormal patterns and classify the gastrointestinal tract accurately. Further research from statistical analysis of the relevance of various MPEG-7 visual descriptors and subsequent classification results showed that the Scalable Color and Homogenous Texture descriptors are the most adequate for the task of event detection in WCE videos [11].

For the purpose of this research work, which was supported by the "IVP" European research project, endoscopic images have been obtained initially from the M2A microcapsule. Fig. 2 shows a sample of the acquired images. They have spatial resolution of 171x151 pixels, a brightness resolution of 256 levels per color plane (8bits), and consisted of three color planes (red, green and blue) for a total of 24 bits per pixel.

Fig. 2 Selected endoscopic images of normal and abnormal cases

The proposed methodology in this paper is considered in two phases. The first implements the extraction of image features while in the second phase a neuro-fuzzy network is implemented / employed to perform the diagnostic task. Texture analysis is one of the most important features used in image processing and pattern recognition. It can give information about the arrangement and spatial properties of fundamental image elements. Many methods have been proposed to extract texture features, e.g. the co-occurrence matrix [12], and the texture spectrum in the achromatic component of the image [13]. The definition and extraction of quantitative parameters from endoscopic images based on texture information has been proposed. This information was initially represented by a set of descriptive statistical features calculated on the histogram of the original image. Recently, a methodology for extraction texture features from the texture spectra in the chromatic and achromatic domains for a selected region of interest from each color component histogram of wireless-capsule endoscopic images has been presented. The implementation of the diagnostic system was based schemes such as Extended Normalized Radial Basis Function (ENRBF) neural networks [14] and Adaptive Fuzzy Logic Systems [15].

In this research study, an alternative approach of obtaining statistical features/parameters from the texture spectra is proposed both in the chromatic and achromatic domains of the image. The definition of texture spectrum employs the determination of the texture unit (TU) and texture unit number (N_{TU}) values. Texture units characterize the local texture information for a given pixel and its neighborhood, and the statistics of the entire texture unit over the whole image reveal the global texture aspects. However, the crisp definition of TU in many cases is not accurate, hence we propose a Fuzzy TU

which captures the concept of similar enough within all its unit boxes. To provide a more flexible way of assigning values to the Texture Unit Boxes of the TU, the boxes will have three associated membership values, each showing the degree to which the gray level of the corresponding pixel is smaller (0), equal (1) or greater (2) than that of the central pixel.

A clustering algorithm has been applied for the sample data in order to organize feature vectors into clusters such that points within a cluster are closer to each other than vectors belonging to different clusters. The fuzzy rule base then is created, using results obtained from this algorithm. For the diagnostic part, the concept of multiple-classifier scheme has been adopted, where the fusion of the individual outputs was realized using fuzzy integral.

II. Image Features Extraction

A major component in analyzing images involves data reduction which is accomplished by intelligently modifying the image from the lowest level of pixel data into higher level representations. Texture is broadly defined as the rate and direction of change of the chromatic properties of the image, and could be subjectively described as fine, coarse, smooth, random, rippled, and irregular, etc. For this reason, we focused our attention on nine statistical measures (standard deviation, variance, skew, kurtosis, entropy, energy, inverse difference moment, contrast, and covariance) [14]. All texture descriptors are estimated for all planes in both RGB {R (Red), G (Green), B (Blue)} and HSV {H (Hue), S (Saturation), V (Value of Intensity)} spaces, creating a feature vector for each descriptor $D_i=(R_i,G_i,B_i,H_i,S_i,V_i)$. Thus, a total of 54 features (9 statistical measures x 6 image planes) are then estimated. For our experiments, we have used 70 endoscopic images related to abnormal cases and 70 images related to normal ones. Generally, the statistical measures are estimated on histograms of the original image (1st order statistics) [16]. However, the histogram of the original image carries no information regarding relative position of the pixels in the texture. Obviously this can fail to distinguish between textures with similar distributions of grey levels. We therefore have to implement methods which recognize characteristic relative positions of pixels of given intensity levels. An additional scheme is proposed in this study to extract new texture features from the texture spectra in the chromatic and achromatic domains, for a selected region of interest from each color component histogram of the endoscopic images.

A. N_{TU} Fuzzy Transformation

The definition of texture spectrum employs the determination of the texture unit (TU) and texture unit number (N_{TU}) values. Texture unit is may be considered as the smallest complete unit which best characterizes the local texture aspect of a given pixel and its neighborhood in all eight directions of a square raster. In a square raster digital image, each pixel is surrounded by eight neighboring pixels. The local texture information for a pixel can be extracted from a neighborhood of 3x3 pixels, which represents the smallest complete unit (in the sense of having eight directions surrounding the pixel). Texture units thus characterize the local texture information for a given pixel and its neighborhood, and the statistics of all the texture units over the whole image reveal the global texture aspects [13].

Given a neighborhood of $\delta \times \delta$ pixels, which are denoted by a set containing $\delta \times \delta$ elements $P = \{P_0, P_1, ..., P_{(\delta\times\delta)-1}\}$, where P_0 represents the chromatic or achromatic (i.e. intensity) value of the central pixel and P_i $\{i = 1,2,...,(\delta\times\delta)-1\}$ is the chromatic or achromatic value of the neighboring pixel i, the $TU=\{E_0,E_1,....,E_{(\delta\times\delta)-1}\}$,where $E_i\{i=1,2,...,(\delta\times\delta)-1\}$ is determined as follows:

$$E_i = \begin{cases} 0, & if \quad P_i < P_0 \\ 1, & if \quad P_i = P_0 \\ 2, & if \quad P_i > P_0 \end{cases} \tag{1}$$

The element E_i occupies the same position as the i^{th} pixel.

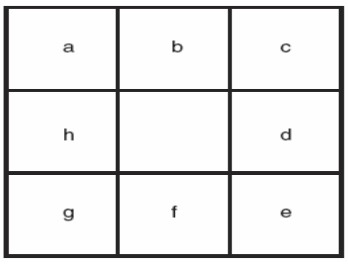

Fig. 3 Eight clockwise, successive ordering ways of the eight elements of the texture unit. The first element E_i may take eight possible positions from a to h

Each element of the TU has one of three possible values; therefore the combination of all the eight elements results in 6561 possible TU's in total. The texture unit number (N_{TU}) is the label of the texture unit and is defined using the following equation:

$$N_{TU} = \sum_{i=1}^{(\delta\times\delta)-1} E_i \times \delta^{i-1} \tag{2}$$

Where, in our case, $\delta = 3$. In addition, the eight elements may be ordered differently. If the eight elements are ordered clockwise as shown in Fig. 3, the first element may take eight possible positions from the top left (a) to the middle left (h), and then the 6561 texture units can be labeled by the Eq. 1, under eight different ordering ways (from a to h). Fig. 4 provides an example of transforming a neighborhood to a texture unit with the texture unit number under the ordering way a.

The previously defined set of 6561 texture units describes the local-texture aspect of a given pixel; that is, the relative grey-level relationships between the central pixel and its

neighbor. However, one of the major inconvenient of this descriptor is the large range of its possible values. In natural images, due to the presence of noise and the different processes of caption and digitation, even if the human eye perceives two neighboring pixels as equal, they rarely have exactly the same intensity value [17].

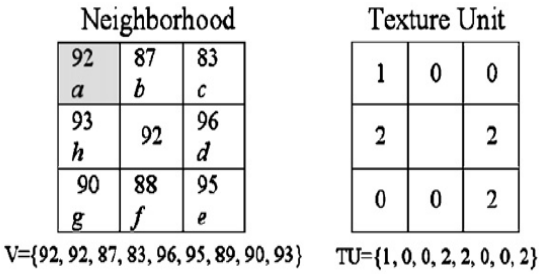

Fig. 4 Transformation of a neighborhood to a Texture Unit and its Texture Unit Number

Thus, when using the Texture spectrum (TS) coding in this type of images, will almost never appear ones as Texture Feature Vectors (TFV) components, but mainly values of 0 and 2; hence the Spectrum obtained will not reflect the human perception of homogeneity.

In texture analysis, we define uniform surface uncertainty, which ranges from 0 to 1, for a point p in the texture as the degree of p belong to uniform physical surface (as defined by the neighborhood average intensity). Therefore, we can transform a grey-scale image into a fuzzy image by suing the uncertainty definition. For a more rough texture, the intensity of pixels in its corresponding fuzzy image will cause a smaller value. The membership distribution of the fuzzy image which transformed from a texture, denoted as Fuzzy Uncertainty Texture Spectrum (FUTS) or fuzzy "N_{TU}" is then used as a distinguishing feature for texture classification.

A grey-scale image f can be transformed into a fuzzy image by a fuzzification function φ. A variety of fuzzification functions can be used to reflect the degree to which pixel intensity represents a uniform physical surface. However, textural properties need neighborhood information about the pixel in order to define adequately membership functions. Here, a simplified triangular membership function used to describe a uniform surface illustrated in Fig. 5 and the *uniform surface uncertainty* is defined as

$$\mu_{ij} = 1 - \left[\frac{\left| f(i,j) - \overline{f}(i,j) \right|}{\max_R f(i,j)} \right] \tag{3}$$

where $\max_R f(i,j)$ is the maximum intensity within the $(\omega \times \omega)$ surface region R centered at point (i,j) and the average intensity is given by

$$\overline{f}(i,j) = \frac{1}{\omega \times \omega - 1} \sum_{m,n \neq i,j}^{R} f(m,n) \tag{4}$$

Note that if $f(i,j)$ is equal to the average neighborhood intensity $\overline{f}(i,j)$ then $f(i,j)$ possesses "full membership" to the surface region R; meaning of the pixel lies on uniform surface, Alternatively, if $f(i,j)$ is significantly different than the average neighborhood intensity $\overline{f}(i,j)$, j), then $\mu_{ij} \to 0$.That means the pixel $f(i,j)$ not lies on uniform surface.

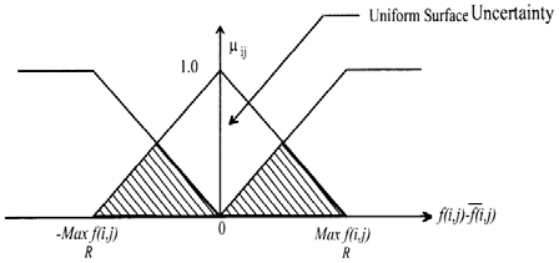

Fig. 5 Fuzzy membership function for a uniform surface

To analyze a texture image, we can transform it into its corresponding fuzzy image by using Eq. 3. As the value in fuzzy image represents the local aspect, the statistics of these values in the fuzzy image should reveal its texture surface information. The texture spectrum histogram $(Hist(i))$ is obtained as the frequency distribution of all the texture units, with the abscissa showing indicating the belief degree or the fuzzy "N_{TU}" and the ordinate representing its occurrence frequency. The texture spectra of various image components {V (Value of Intensity), R (Red), G (Green), B (Blue), H (Hue), S (Saturation)} are obtained from their texture unit numbers. The statistical features are then estimated on the histograms of the "N_{TU}" fuzzy transformations of the chromatic and achromatic planes of the image (R,G,B,H,S,V).

III. Image Features Evaluation

Recently, the concept of combining multiple classifiers has been actively exploited for developing highly reliable "diagnostic" systems [18]. One of the key issues of this approach is how to combine the results of the various systems to give the best estimate of the optimal result. A straightforward approach is to decompose the problem into manageable ones for several different sub-systems and combine them via a gating network. The presumption is that each classifier/sub-system is "an expert" in some local area of the feature space. The sub-systems are local in the sense that the weights in one "expert" are decoupled from the weights in other sub-networks.

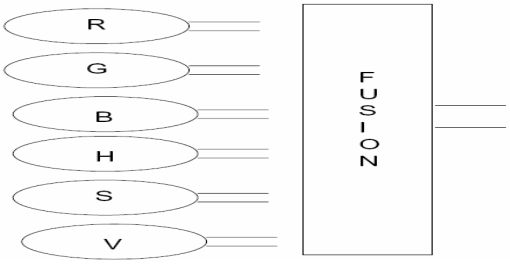

Fig. 6 Proposed fusion scheme

In this study, six subsystems have been developed, and each of them was associated with the six planes specified in the

feature extraction process (i.e. R, G, B, H, S, & V). For each subsystem, 9 statistical features have been associated with, resulting thus a total 54 features space. Each subsystem was modeled with the proposed neurofuzzy learning scheme. This provides a degree of certainty for each classification based on the statistics for each plane. The outputs of each of these networks must then be combined to produce a total output for the system as a whole as can be seen in Fig. 6.

While a usual scheme chooses one best subsystem from amongst the set of candidate subsystems based on a winner-takes-all strategy, the current proposed approach runs all multiple subsystems with an appropriate collective decision strategy. The aim in this study is to incorporate information from each plane/space so that decisions are based on the whole input space. The simplest method is to take the average output from each classifier as the system output. This does not take into account the objective evidence supplied by each of the individual classifiers and confidence which we have in that classifiers results. The fuzzy integral is an alternative method that claims to resolve both of these issues by combining evidence of a classification with the systems expectation of the importance of that evidence.

The fuzzy integral introduced by Sugeno [19] and the fuzzy measures from Yager [20] are very useful in combining information. A fuzzy measure g_λ is a set function such that,

- the fuzzy measure of an empty set is equal to zero - $g(0) = 0$,
- the fuzzy measure of an entire set Q is equal to one - $g(Q) = 1$, and
- the fuzzy measure of set A is less than or equal to that of set B if A is a subset of $B - g(A) \le g(B)$ if $A \subset B$

This function can be interpreted as finding the maximal grade of agreement between networks' outputs and their fuzzy measures for a particular class. If the following additional property is also satisfied, the fuzzy measure is referred to as a $g_\lambda - fuzzy$ measure.

$$\forall A, B \subset Q \text{ and } A \cap B = 0, g(A \cup B) = g(A) + g(B) + \lambda g(A) g(B)$$

$$\lambda \in (-1, \infty) \tag{5}$$

where the λ measure can be given by solving the following non-linear equation [21]

$$\lambda + 1 = \prod_{i=1}^{n} (1 + \lambda g^i) \qquad \lambda > -1 \tag{6}$$

When combining multiple NNs, let g^i denote the fuzzy measure of network i. These measures can be interpreted as quantifying how well a network properly classified the samples/patterns. They must be known and can be determined in different ways i.e., the fuzzy measure of a network could equal the ratio of correctly classified patterns during supervised training over the total number of patterns being classified. In this research, each network's fuzzy measure equaled $1 - K_i$, where K_i was network ith overall testing kappa value [22]. A pattern is being classified to one of m possible output classes, c_j for $j = 1, ..., m$. The outputs of n different networks are being combined, where NN_i denotes the ith network. First, these networks must be renumbered/rearranged such that their a posteriori class probabilities are in descending order of magnitude for each output class j,

$$y_1(c_j) \ge y_2(c_j) \ge .... \ge y_n(c_j)$$

where $y_i(c_j)$ is the ith network's a posteriori class j probability. Next, each network/ set of networks' $g_\lambda - fuzzy$ measure is computed for every output class j and is denoted by $g_j(A_i)$. $A_i = \{NN_1, NN_2, ..., NN_i\}$ is the set of the first i networks ordered correspondingly to class j's associated a posteriori probabilities. These values can then be computed using the following recursive method,

$$g_j(A_1) = g_j(\{NN_1\}) = g^1$$

$$g_j(A_i) = g_j(\{NN_1, ..., NN_i\}) = g^i + g(A_{i-1}) + \lambda g^i g(A_{i-1}) \quad \text{for } 1 < i < n$$

$$g_j(A_n) = g_j(\{NN_1, ..., NN_n\}) = 1$$

Finally, the fuzzy integral for each class j is defined as [23],

$$\max_{i=1}^{n} [\min[y_i(c_j), g_j(A_i)]] \tag{7}$$

The class with the largest fuzzy integral value is then chosen as the output class to which the pattern is classified. Eq. 7 summarizes combining multiple NNs using a fuzzy integral approach.

$$\max_{class} \begin{bmatrix} \max_{network} \begin{bmatrix} \min[y_1(c_1), g_1(A_1)] \\ \min[y_2(c_1), g_1(A_2)] \\ ..... \\ \min[y_n(c_1), g_1(A_n)] \end{bmatrix} \\ \max_{network} \begin{bmatrix} \min[y_1(c_2), g_2(A_1)] \\ \min[y_2(c_2), g_2(A_2)] \\ ..... \\ \min[y_n(c_2), g_2(A_n)] \end{bmatrix} \\ ..... \\ \max_{network} \begin{bmatrix} \min[y_1(c_m), g_m(A_1)] \\ \min[y_2(c_m), g_m(A_2)] \\ ..... \\ \min[y_n(c_m), g_m(A_n)] \end{bmatrix} \end{bmatrix} \tag{8}$$

The classification scheme utilized here is a neuro-fuzzy system that incorporates a two-stages clustering algorithm for finding the initial parameters of rules.

IV. Neuro-fuzzy Classifier

Although Neural Networks (NNs) and Fuzzy Logic (FL) systems have well-established strengths and weaknesses, they are both capable of modeling highly complex non-linear relationships. NNs' highly interconnected structure enables

them to learn the low-level interdependencies of a system. The difficulty is that the resulting model is often as complex as the system itself making thus hard to interpret what has been learnt or to include expert knowledge. FL systems on the other hand are often less able to learn such low level information but the knowledge is represented in the form of IF-THEN type rules. These are readily interpreted by a human expert to validate or expand the learnt information. Fuzzy rules define active areas of the input domain through the product of a number of partitions defined on a single of the domains axis. Neuro-Fuzzy (NF) systems attempt in general to combine the low-level numerical modeling capabilities of NNs with some of the representational transparencies of FLs, but like FL often suffer from the so-called "curse of dimensionality" [24].

The neuro-fuzzy classifier-scheme adopted in this study tries to overcome the increased number of rules by utilizing a two stage process, a clustering sections and a neuro-fuzzy inference network.

A. Clustering Algorithm

The clustering algorithm we apply in this paper consists of two stages. In the first stage the method similar to LVQ algorithm generates crisp c-partitions of the data set. The number of clusters c and the cluster centers v_i, $i = 1,...,c$, obtained from this stage are used by FCM (Fuzzy c-means) algorithm in the second stage. The first stage clustering algorithm determines the number of clusters by dividing the learning data into these crisp clusters and calculates the cluster centers which are the initial values of the fuzzy cluster centers derived the second stage algorithm. Let $Z = [z_1,...,z_n] \in \mathrm{R}^{np}$ be a learning data. The first cluster is created starting with the first data vector from Z and the initial value of the cluster centre is taking as a value of this data vector. Then other data vectors are included into the cluster but only these ones which satisfy the following condition

$$\|z_k - v_i\| < D \tag{9}$$

where $z_k \in Z, k=1,...,n$ and $v_i, i=1,...,c$ are cluster centers, $V=[v_1,...,v_n] \in \mathrm{R}^{cp}$, the constant value D is fixed at the beginning of the algorithm. Cluster centers v_i are modified for each cluster (i.e., $i = 1,...,c$) according to the following equation

$$v_i(t+1) = v_i(t) + a_t(z_k - v_i(t)) \tag{10}$$

where $t = 0,1,2,...$ denotes the number of iterations, $a_t \in [0,1]$ is the learning rate and it is decreasing during performance of the algorithm (depending on the number of elements in the cluster). As a result of performance of this algorithm we get the number of clusters c, we have divided data set into the clusters, and we know values of cluster centers v_i $i = 1,...,c$ which we can use as initial values for the second stage clustering algorithm. In the second stage the fuzzy c-means algorithm has been used. FCM is a constrained optimization procedure which minimizes the weighted within-groups sum of squared errors objective functions J_m with respect to fuzzy membership's u_{ik} cluster centres v_i, given training data z_k, $i = 1,...,c;\ k = 1,..,n$

$$\min_{(U,V)}\{J_m(U,V;Z) = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m \left\|z_k - v_i\right\|^2\} \tag{11}$$

The number of clusters c and the initial values of cluster centers v_i come from the first stage clustering algorithm.

B. Fuzzy Inference Neural Networks

The two-stages clustering algorithm provides the fuzzy c-partition of the sample data. The number of rules in the proposed fuzzy inference neural network (FINN) equals to the number of clusters c obtained from the clustering algorithm. The proposed FINN scheme is a MIMO adaptive fuzzy logic system with centre average as defuzzification concept. The schematic of the FINN scheme which is shown in Fig. 7 consists of four layers. The first two layers LI and L2 correspond to IF part of fuzzy rules whereas layers L3 and L4 contain information about THEN part of these rules, and perform the defuzzification task. There are $c \times q$ elements in layer Ll. They realize the membership functions which are defined by

$$\mu_j^i = \exp\left[-\left(\frac{x_j - v_{ij}}{\sigma_{ij}}\right)^2\right] \tag{12}$$

for $j = 1,...,q$ and $i = 1,...,c$. The values v_{ij} in Eq. (12) denote the centers of the Gaussian membership functions and are equal to the values of the vectors v_i which have been derived from the second stage clustering algorithm. The value σ_{ij} defines the widths of the Gaussian membership functions. These values have been estimated according to

$$\sigma_{ij} = \left(\sum_{k=1}^{n} u_{ik}(z_{kj} - v_{ij})^2 \Big/ \sum_{k=1}^{n} u_{ik}\right)^{1/2} \tag{13}$$

These values are calculated based on the matrix U which elements represent fuzzy memberships of z_k ith cluster and have values obtained from the second stage clustering algorithm.

The second layer L2 has c elements which realize multiplication operation because of using Larsen rule in fuzzy reasoning procedure. Each element in this layer is associated with one fuzzy rule. Outputs of this layer represent the fire strength of the rules, expressed by

$$\tau_i = \prod_{j=1}^{q} \mu_j^i(\overline{x}_j) \tag{14}$$

Layer L3 contains the parameters v_{ip}, for $i = 1,...,c$. The only one element in layer L4 performs division operation. Layers

L3 and L4 serve as the center average defuzzifier.

$$\bar{y} = \frac{\sum_{i=1}^{c} v_{ip}\tau_i}{\sum_{i=1}^{c} \tau_i} \quad (15)$$

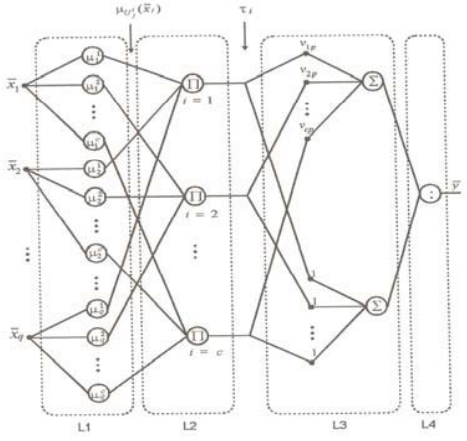

Fig. 7 Schematic of FINN structure

The multi-layer connectionist structures of FINN scheme allow us to apply, learning procedures similar to the back-propagation method which is commonly used as learning algorithm for feed-forward multi-layer artificial neural networks [25]. Based on the idea of the back-propagation the following updates for tuning the parameters of Gaussian membership functions have been derived

$$v_{ip}(t+1) = v_{ip}(t) - \eta(\bar{y} - \hat{y}) \frac{\prod_{j=1}^{q} \exp\left[-\left(\frac{\bar{x}_j - v_{ij}}{\sigma_{ij}}\right)^2\right]}{\sum_{i=1}^{c}\prod_{j=1}^{q} \exp\left[-\left(\frac{\bar{x}_j - v_{ij}}{\sigma_{ij}}\right)^2\right]} \quad (16)$$

$$v_{ij}(t+1) = v_{ij}(t) - \beta \frac{(\bar{x}_j - v_{ij})(v_{ip} - \bar{y})}{(\sigma_{ij})^2} (\bar{y} - \hat{y}) \frac{\prod_{h=1}^{q} \exp\left[-\left(\frac{\bar{x}_h - v_{ih}}{\sigma_{ih}}\right)^2\right]}{\sum_{i=1}^{c}\prod_{h=1}^{q} \exp\left[-\left(\frac{\bar{x}_h - v_{ih}}{\sigma_{ih}}\right)^2\right]} \quad (17)$$

$$\sigma_{ij}(t+1) = \sigma_{ij}(t) - \gamma \frac{(\bar{x}_j - v_{ij})^2 (v_{ip} - \bar{y})}{(\sigma_{ij})^3} (\bar{y} - \hat{y}) \frac{\prod_{h=1}^{q} \exp\left[-\left(\frac{\bar{x}_h - v_{ih}}{\sigma_{ih}}\right)^2\right]}{\sum_{i=1}^{c}\prod_{h=1}^{q} \exp\left[-\left(\frac{\bar{x}_h - v_{ih}}{\sigma_{ih}}\right)^2\right]} \quad (18)$$

V. Results

The proposed approach was evaluated using 140 clinically obtained endoscopic M2A images. For the present analysis, two decision-classes are considered: abnormal and normal. Seventy images (35 abnormal and 35 normal) were used for the training and the remaining ones (35 abnormal and 35 normal) were used for testing. The extraction of quantitative parameters from these endoscopic images is based on texture information. This information is represented by a set of descriptive statistical features. Nine statistical measures for each individual image component are calculated though the related texture spectra after applying the proposed fuzzy N_{TU} transformation. In addition to the proposed FINN scheme, a radial basis function network has been implemented for comparison purposes. Both types of networks (i.e. FINN and RBF) are incorporated into a multiple classifier scheme, where the structure of each individual (for R, G, B, H, S, & V planes) classifier is composed of 9 input nodes (i.e. nine statistical features) and 2 output nodes.

TABLE 1 FUZZY N_{TU}-BASED PERFORMANCES

Modules	RBF Accuracy (70 testing patterns)	FINN Accuracy (70 testing patterns)
R	92.85% (5 mistakes)	94.28% (4 mistakes)
G	95.71% (3 mistakes)	97.14% (2 mistakes)
B	91.42% (6 mistakes)	92.85% (5 mistakes)
H	92.85% (5 mistakes)	94.28% (4 mistakes)
S	90.00% (7 mistakes)	95.71% (3 mistakes)
V	94.28% (4 mistakes)	91.42% (6 mistakes)
Overall	91.42% (6 mistakes)	94.28% (4 mistakes)

Each of the sub-networks was trained in turn using a threshold value of 0.015 to stop training. The values of the learning coefficients in this iteration process have been set as $\eta = 0.3$, $\beta = 0.01$, $\gamma = 0.02$. The value of the weighting exponent in the second stage clustering algorithm has been chosen as m=1.9. The FINN scheme trained on the R feature space and it then achieved an accuracy of 94.28% on the testing data incorrectly classifying 3 of the normal images as abnormal and 1 abnormal as normal ones. The network trained on the G feature space misclassified 2 normal images as abnormal but not the same ones as the R space. The B feature space achieved an accuracy of 92.85% on the testing data with 5 misclassifications, i.e. 3 abnormal as normal ones and the remaining ones as abnormal one. The network trained on the H feature space achieved 94.28% accuracy on the testing data. The network trained on the S feature space achieved an accuracy of only 95.71% on the testing data. Finally, the network for the V feature space misclassified 2 normal case as abnormal and 4 abnormal as normal one, giving it an accuracy of 91.42% on the testing data.

The fuzzy integral (FI) concept has been used here to combine the results from each sub-network and the overall system provided an accuracy of 94.28%. More specifically, 3 normal cases as abnormal and one abnormal as normal one provide us a good indication of a "healthy" diagnostic performance. However the level of confidence/certainty was 0.52 as shown in Fig. 8.

In a similar way, a multi classifier consisting of RBF networks with 9 input nodes and 2 output nodes was trained on each of the six feature spaces. Table 1 illustrates the performances of the network in the individual components.

The RBF network trained on the R feature space and it then achieved an accuracy of 92.85% on the testing data incorrectly classifying 3 of the normal images as abnormal and 2 abnormal as normal ones. The network trained on the G feature space misclassified 2 normal images as abnormal but not the same ones as the R space.

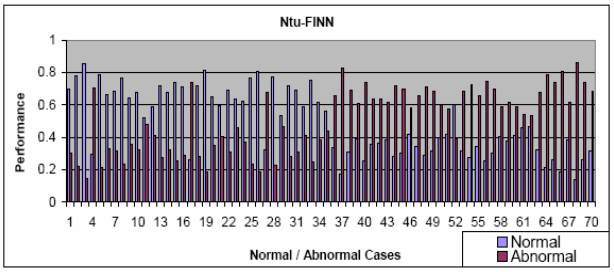

Fig. 8 Performance of FINN structure

The remaining one image was misclassified as normal one. The B feature space achieved an accuracy of 91.43% on the testing data with 6 misclassifications, i.e. 4 abnormal as normal ones and the remaining 2 images as abnormal ones. The network trained on the H feature space achieved 92.85% accuracy on the testing data. The network trained on the S feature space achieved an accuracy of only 90% on the testing data. Finally, the network for the V feature space misclassified 2 normal cases as abnormal and 2 abnormal as normal ones, giving it an accuracy of 94.28% on the testing data. The fuzzy integral (FI) concept has been used here to combine the results from each sub-network and the overall system misclassified 3 normal cases as abnormal and 3 abnormal as normal ones, giving the system an overall accuracy of 91.43% despite the fact that RBF was characterized by a very fast training process. The confidence level for each correct classification was above 0.55.

VI. Conclusion

The major contribution of the proposed system in the process of medical diagnosis is that it can provide additional information to physicians on the characterization of the endoscopic images / tissues, by exploiting its textural characteristics, which are consequently used for the classification of the corresponding image regions as normal or abnormal. An approach on extracting statistical features from endoscopic images using the M2A Given Imaging capsule have been developed by obtaining those quantitative parameters from the texture spectra from the calculation the fuzzy texture unit numbers over the histogram spectrum. In this study, an intelligent decision support system has been developed for endoscopic diagnosis based on a multiple-classifier scheme. This multiple-classifier approach using Fuzzy integral as a fusion method provided encouraging results.

References

[1] K. Nagasako, T. Fujimori, Y. Hoshihara, M. Tabuchi, 1998. *Atlas of Gastroenterologic Endoscopy by High-resolution Video-Endoscopy*, IGAKU-SHOIN Ltd., Tokyo

[2] S. Krishnan, P. Wang, C. Kugean, M. Tjoa, "Classification of endoscopic images based on texture and neural network", *Proc. 23rd Annual IEEE Int. Conf. in Engineering in Medicine and Biology*, (4), pp. 3691-3695, 2001

[3] D.E. Maroulis, D.K. Iakovidis, S.A. Karkanis, D.A. Karras, "CoLD: a versatile detection system for colorectal lesions endoscopy video-frames", *Computer Methods and Programs in Biomedicine,* (70), pp. 151–166, 2003

[4] A.F. Ravens, C.P. Swain, "The wireless capsule: new light in the darkness", in Digestive Diseases, vol. 20, pp. 127-133, 2002.

[5] G. Idden, G. Meran, A. Glukhovsky an€d P. Swain, "Wireless capsule endoscopy", Nature, pp. 405-417, 2000

[6] M. Mylonaki, A. Fritscher-Ravens, C.P. Swain, "Clinical results of wireless capsule endoscopy" Gastrointest. Endosc., vol.55, AB146, 2002

[7] P. Spyridonos, F. Vilariño, J. Vitrià and P. Radeva, "Identification of Intestinal Motility Events of Capsule Endoscopy Video Analysis", Proc. of Advanced Concepts for Intelligent Vision Systems Conf, Antwerp, Belgium, pp. 575-580, 2005

[8] F. Vilariño, P. Spyridonos, J. Vitrià, P. Radeva, "Self Organized Maps for Intestinal Contractions Categorization with Wireless Capsule Video Endoscopy", Proc. of the 3rd European Medical and Biological Engineering Conference, EMBEC'05 Prague, pp. 3443-3447, 2005.

[9] N. Bourbakis, S. Makrogiannis, D. Kavraki, "A Neural Network-based Detection of Bleeding in sequences of WCE images", Proc. of the 5th IEEE Symposium on Bioinformatics and Bioengineering (BIBE'05), pp. 324-327, 2005

[10] P.Y Lau, P.L. Correia, "Detection of bleeding patterns in WCE video using multiple features", Engineering in Medicine and Biology Society, EMBS 2007, 29th Annual International Conference of the IEEE, pp. 5601-5604, 2007

[11] M.T. Coimbra, J.P.S. Cunha, "MPEG-7 Visual Descriptors-Contributions for Automated Feature Extraction in Capsule Endoscopy", IEEE transactions on circuits and systems for video technology, Vol. 16, No. 5, pp. 628-637, 2006.

[12] M. Gletsos, S. Mougiakakou, G. Matsopoulos, K. Nikita, A. Nikita, D. Kelekis, "A computer-aided diagnostic system to characterize CT focal liver lesions: design and optimization of a neural network classifier", IEEE Trans. on Information Technology in Biomedicine, Vol. 7, No. 3, pp. 153-162, 2003.

[13] D-C He, L. Wang, "Texture features based on texture spectrum", Pattern Recognition, Vol. 24, No. 5, pp. 391-399, 1991.

[14] M. Boulougoura, E. Wadge, V.S. Kodogiannis, H.S. Chowdrey, "Intelligent systems for computer-assisted clinical endoscopic image analysis", 2nd IASTED Int. Conf. on Biomedical Engineering, Innsbruck, Austria, pp. 405-408, 2004.

[15] V.S. Kodogiannis, M. Boulougoura, E. Wadge, J.N. Lygouras, "The usage of soft-computing methodologies in interpreting capsule endoscopy", Engineering Applications in Artificial Intelligence, Elsevier 2007, Vol. 20, pp. 539-553.

[16] R. Haralick, "Statistical and structural approaches to texture", IEEE Proc, Vol. 67, pp. 786- 804, 1979

[17] A. Barcelo, E. Montseny, P. Sobrevilla, Fuzzy Texture Unit and Fuzzy Texture Spectrum for texture characterization, Fuzzy Sets and Systems, Vol. 158, pp. 239 – 252, 2007

[18] V.S. Kodogiannis, "Intelligent classification of bacterial clinical isolates in vitro, using an array of gas sensors", J. Intelligent and Fuzzy systems, Vol. 16, No. 1, pp. 1-14, 2005

[19] Sugeno, M. "Fuzzy measures and fuzzy integrals: a survey, in Fuzzy Automata and Decision Processes", (M.M. Gupta, G. N. Saridis, and B.R. Gaines, editors), pp. 89-102, North-Holland, 1977

[20] R. Yager, "A General Approach to Criteria Aggregation using Fuzzy Measures", International Journal of Man-Machine Studies, Vol. 39, No. 2, pp. 187-213, 1993.

[21] M. Grabisch, T. Murofushi, M. Sugeno, M., "Fuzzy measure of fuzzy events defined by fuzzy integrals", Fuzzy Sets and Systems, Vol. 50, pp. 293-313, 1992

[22] S. Mitra, S.K. Pal, P. Mitra, "Data mining in soft computing framework: A survey", IEEE Transactions on Neural Networks, Vol. 13, No. 1, pp. 3-14, 2000
[23] L.I Kuncheva, *Fuzzy Classifier Design*, Physica-Verlag, 2000
[24] C.T. Teng Lin, C.S. George Lee, *Neural Fuzzy Systems*, Prentice Hall, 1996
[25] V. Kodogiannis, "Computer-aided Diagnosis in Clinical Endoscopy using Neuro-Fuzzy System", in Proceedings of IEEE FUZZ 2004, pp. 1425-1429, 2004

Dr. Vassilis S. Kodogiannis was born in Heraklion, Crete, Greece at 1966. He received the Electrical Eng. Degree from the Democritus University of Thrace, Greece in 1990, the MSc in VLSI Systems Engineering from UMIST, UK, in 1991 and the PhD in Electrical Engineering from Liverpool Univ. in 1994. During 1996-1999 he has been with Technical University of Crete and Univ. of Ioannina, Greece as Visiting Assistant Professor. He is currently Principal Lecturer in the School of Computer Science at Univ. of Westminster, London, U.K. His research interests are in the areas of Intelligent Systems, Image Processing, and Control. He is a member of IEEE, INNS, EUSFLAT and the Technical Chamber of Greece. He was the principal investigator for the "Intracorporeal Video Probe" European Research Project.

Dr. John N. Lygouras was born in Kozani, Greece in May 1955. He received the Diploma degree and the Ph.D. in Electrical Engineering from the Democritus University of Thrace, Greece in 1982 and 1990, respectively, both with honors. From 1982 he was a Research Assistant and since 2000 he is an Associate Professor at the Democritus University of Thrace, Department of Electrical and Computer Engineering. In 1997 he spent six months at the University of Liverpool, Department of Electrical Engineering and Electronics as a Honorary Senior Research Fellow. His research interests are in the field of analog and digital electronic systems design and implementation. His interests also include the research on robotic manipulators trajectory planning and execution and position control of underwater remotely operated vehicles.